DISSERTATION

MULTI-COLOR VISUALIZATION AND QUANTIFICATION OF SINGLE RNA TRANSLATION AND HIV-1 PROGRAMMED RIBOSOMAL FRAMESHIFTING IN LIVING CELLS

Submitted by

Kenneth Ray Lyon Jr.

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 2019

Doctoral Committee:

Advisor: Timothy Stasevich

Brian Munsky
Bob Cohen
Jake DeLuca

ABSTRACT

MULTI-COLOR VISUALIZATION AND QUANTIFICATION OF SINGLE RNA TRANSLATION AND HIV-1 PROGRAMMED RIBOSOMAL FRAMESHIFTING IN LIVING CELLS

Sixty years ago, Francis Crick first stated the central dogma of molecular biology: DNA makes RNA, RNA makes protein. At that time the central dogma could only be imagined, but over the past two decades revolutionary advances in fluorescence microscopy have now made it possible to directly image transcription and translation in living cells and organisms. A key breakthrough was the discovery and development of GFP, which can be genetically fused to other proteins to selectively light them up and track their expression *in vivo*. While this powerful technology can illuminate mature protein products in live cells, processes like translation remain in the dark. This is because fluorescent fusion tags take too long to mature and light up. By the time the fluorescence becomes visible, translation of a nascent protein is over and has long since separated from its parental RNA strand. This fundamental challenge has made it difficult to visualize, quantify, and study translational gene regulatory mechanisms in living cells and organisms with fluorescence microscopy.

To achieve this, nascent chain tracking (NCT) was developed, a technique that uses multi-epitope tags and antibody-based fluorescent probes to visualize and quantify protein synthesis dynamics at the single-RNA level. NCT revealed an elongation rate of ~10 amino acids per second, with initiation occurring stochastically every ~30 seconds. Polysomes contain ~1 ribosome every 200 to 900 nucleotides. By employing a multi-color probe strategy, NCT shows that a small fraction (~5%) form multi-RNA sites in which two distinct RNAs are translated simultaneously while spatially overlapping.

NCT was then applied to further understand the dynamics of translating RNAs interactions with ribonucleoprotein (RNP) granules during cellular stress. NCT was used to image real-time single RNAs, their translational output, and RNA-granule interactions during stress. Although translating mRNAs only interact with RNP granules dynamically, non-translating mRNAs can form stable, and sometimes rigid, associations with RNP granules. These stable associations increase with both mRNA length and granule size. Live and fixed-cell imaging demonstrated that mRNAs can extend beyond the protein surface of a stress granule, which may facilitate interactions between RNP granules. Thus, the recruitment of mRNPs to RNP granules involves dynamic, stable and extended interactions affected by translation status, mRNA length and granule size that collectively regulate RNP granule dynamics.

Finally, NCT was utilized to quantify the dynamics of multiple open reading frames at the individual RNA level in a living system. An NCT multi-color imaging modality was used to investigate ribosomal frameshifting during translation. Frameshifts occur through a ribosomal +1 (-2) or -1 (+2) nucleotide(s) "slip" into another frame and are implicated in both human disease and viral infections. While previous work has uncovered many mechanistic details about single-RNA frameshifting kinetics *in vitro*, very little is known about how single RNAs frameshift in living systems. Applying multi-frame NCT technology to RNA encoding the -1 programmed ribosomal frameshift (-1PRF) sequence of HIV-1 revealed that a small subset (~8%) of translating RNAs frameshift in living cells. This multi-color NCT method also revealed that frameshifting RNA are preferentially in multi-RNA sites, are translated at about the same rate as non-frameshifting RNA, and can continuously frameshift for more than four rounds of translation. Interestingly, fits to a bursty model of frameshifting constrain frameshifting kinetic rates and demonstrate how ribosomal traffic jams contribute to the persistence of the frameshifting state. These data provide fresh insight into retroviral frameshifting and could lead to alternative strategies to perturb the process in living cells.

ACKNOWLEDGEMENTS

I would like to thank the Stasevich lab for the refinement of my scientific techniques, the opportunity to learn coding languages for analysis, and helping me develop a sharp philosophical evaluation of my work and others.

I would also like to thank my student advisor committee members Jake DeLuca, Bob Cohen, and Brian Munsky for their time, wisdom, and perspectives.

Finally, I want to acknowledge Timothy Stasevich, Tatsuya Morisaki, Luis U. Aguilera, Brian Munsky, Roy Parker, Anthony Khong, Keith DeLuca, Jake DeLuca, and Stephanie Moon for their insight and contribution to the following work(s).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Visualizing translation

The central dogma of molecular biology was first hypothesized sixty years ago: DNA makes RNA, RNA makes protein [1]. The synthesis of protein encoded by messenger RNA (mRNA), or translation, is a crucial step of gene expression across all domains of life [2]. Translation is an energy-intensive process [3] that can be divided into four key steps [4, 5]: (i) ribosome search and recruitment; (ii) initiation; (iii) nascent peptide elongation; and (iv) termination. Each of these steps is subject to tight regulation by a variety of mechanisms [6, 7]. In eukaryotes, the small (43S) subunit of the ribosome binds to and scans along RNA during the search and recruitment phase [5]. Once a start codon is located, it initiates and assembles with other factors into the large (80S) elongation complex capable of decoding mRNA [5]. The 80S ribosomal complex is composed of three sites: an A site which allows for codons to pair with an anti-codon of the correct aminoacylated tRNA, a P site where peptidyl transfer occurs from the tRNA at this site to the A site, and finally an E site where tRNA exit from the complex [8]. Elongation occurs as the ribosome decodes the mRNA one codon (three nucleotides) at a time to assemble a nascent peptide one amino acid at a time [4]. As synthesis of the chain is completed, the ribosome terminates at a stop codon and the nascent protein is released [9].

In vitro and population-based studies have established most of what is known about translation. Using cell-free, live-cell, single molecule fluorescence resonance energy transfer (smFRET), and crystallography, these studies have uncovered: the minimal components necessary for translation [10, 11], translation kinetics [12–17], and different ribosomal rotational states [18]. In addition, single-ribosome resolution studies have revealed that nascent peptide co-translational folding can generate forces to release arrested ribosomes from allosteric exit tunnel interactions [19] and that ribosomes have two distinct active helicase mechanisms to overcome RNA structure during

translation [20]. Deep-sequencing techniques, such as ribosome profiling, are currently used to achieve sub-codon resolution of initiating and elongating ribosomes [21]. This technique relies on deep-sequencing of cDNA generated from ribosomal protected RNA fragments. Ribosome profiling has provided insight into genome-wide translational output, sub-codon positioning of ribosomes on mRNA, genome-wide translation kinetics [22], and the effects of mRNA isoforms on translational output [23].

These studies have provided valuable population-based measurements, *in vitro* single-molecule kinetics, and sub-codon snapshots to infer translation dynamics. However, these methods cannot be performed on intact cells, causing spatial information to be lost and the heterogeneity of potentially key regulatory mechanisms to be obscured in ensemble measurements [24, 25]. Without the ability to directly visualize translation in a living system, dissecting complex translational regulatory dynamics has been challenging [24, 26]. Not only has it been difficult to distinguish different co-translational regulatory mechanisms from each other, but, more importantly, it has been difficult to distinguish translational regulatory heterogeneity in general. For example, if the level of a particular protein product in a cell was measured to be low, it was not clear whether this was due to a subset of RNA that was translationally regulated differently or was due to RNA being translationally regulated similarly. Understanding this distinction in a live-cell could help uncover potentially key regulatory steps or mechanisms that were previously not fully understood using population-based measurements.

Although labeling of mRNA using viral coat proteins, e.g., MS2 coat proteins and PP7 coat proteins [27, 28], has been successfully used to study transcription at the single molecule level [29, 30], labeling of nascent peptide chains at translation sites cannot be easily done with fluorescent fusion tags because they take too long to mature and light up (Figure 1.1). By the time the fluorescence becomes visible, translation is over and the protein has long since separated from its parental RNA strand [31]. This fundamental challenge has made it difficult to visualize, quantify, and study translational gene regulatory mechanisms in living cells and organisms with fluorescence microscopy.

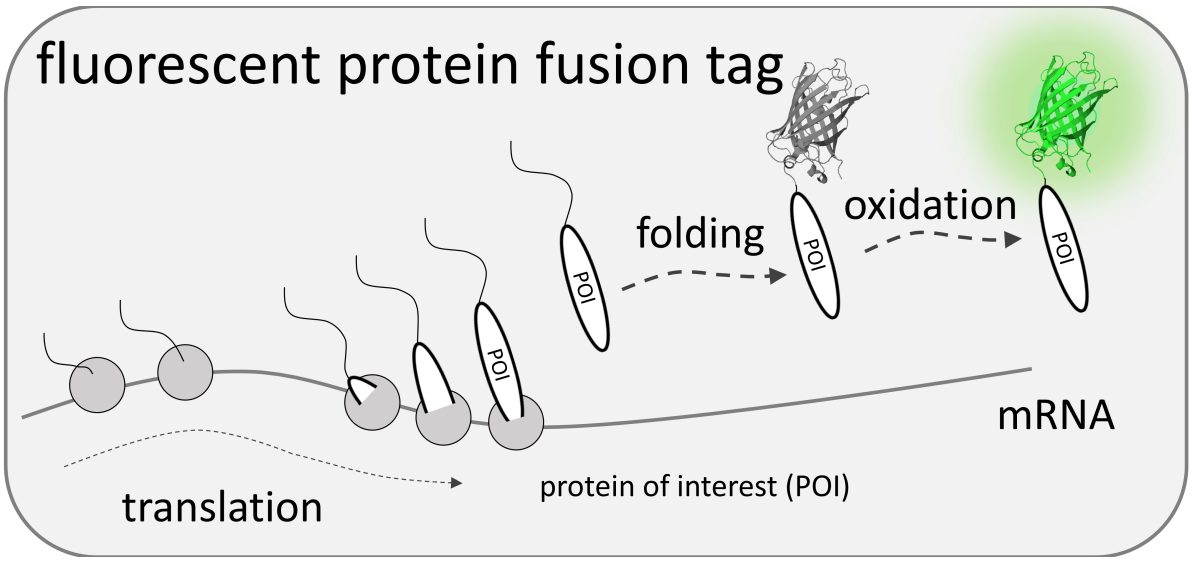


Figure 1.1: Fluorescent fusion tags incapable of visualizing translation. Fluorescent fusion tags such as GFP [shown as a glowing beta-barrel structure; Protein Data Base (PDB) ID: 4KW4] take time to fluoresce and are not sufficient to observe translation dynamics of a protein of interest (POI). Adapted from [31]. K. Lyon created and contributed to the design of the cartoon schematic.

Despite these challenges, researchers have attempted to develop methods to visualize translation using fluorescence microscopy. For example, the FlAsH and ReAsH dyes have been used to rapidly bind and light up tagged nascent proteins and to detect the sites of translation in living cells [32]. However, photobleaching of dyes and non-specific binding limited the signal to noise ratio and prevented real-time imaging. Some of the signal to noise issues were addressed by recent work using a 3x repeat tag for signal amplification [33], but solubility issues persist therefore FlAsH and ReAsH have not been widely adopted. More recently, the Translating RNA Imaging by Coat protein Knock-off (TRICK) assay was used to image and distinguish translated RNA from untranslated RNA [26]. In this assay, RNA fluorescence is lost upon the first round of translation. While useful for imaging the pioneering round of translation, it is not possible to track multiple rounds of translation.

To address these issues, multiple research groups have developed a methodology to visualize and quantify single RNA through multiple rounds of translation. The method employed within this dissertation is nascent chain tracking (NCT) and is based off of work in which I was a co-author

(Appendix A and D). Briefly, NCT relies on repeated non-maturing epitopes encoded in the N-terminus of a nascent peptide chain. We target these epitopes with fluorescently labeled probes to mark the nascent chains of elongating ribosomes and combine this with the MS2 system to mark RNA [27]. This method has enabled other studies of translation heterogeneity and regulation in living cells [34–36].

In this dissertation, I will highlight a few applications of NCT. In Appendix A, I discuss the establishment and validation of NCT, of which I was a co-author (see Appendix D for permission of use). In Chapter 2, I discuss probing translation dynamics using fragmented antibodies (Fabs) or single-chain variable fragments (scFv), of which is based on a review where I was first author. In Chapter 3, I discuss the dynamics of mRNA interactions with stress granules, of which I was a co-author (see Appendix D for permission of use). Finally, in Chapter 4, I discuss the dynamics of single RNA HIV-1 programmed frameshifting, of which I was first author. A graphical summary of the main chapters of this dissertation can be seen in Figure 1.2.

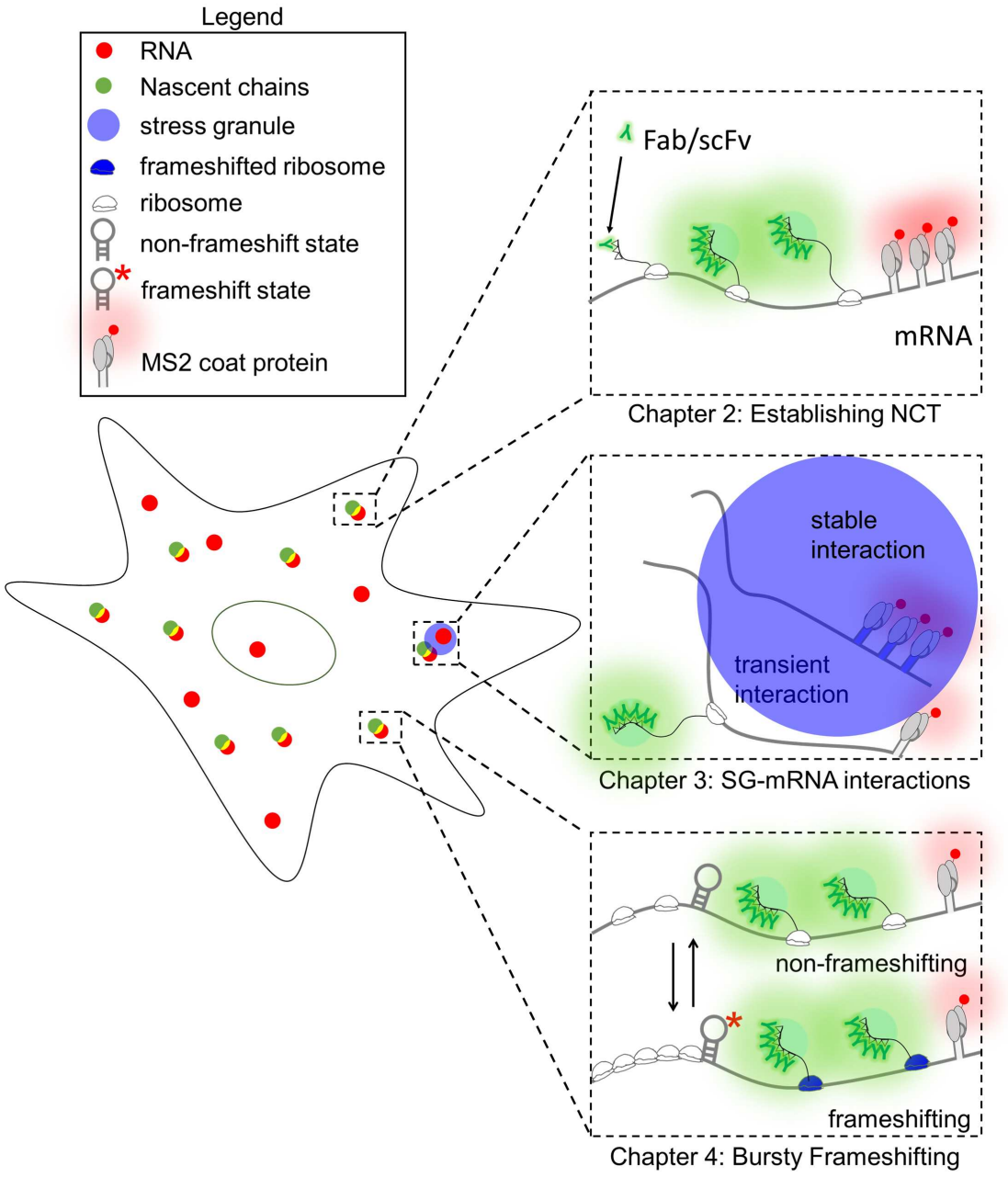


Figure 1.2: Graphical abstract summary of the main chapters of this dissertation. Fab/scFv can label nascent chains of ribosomes during elongation on individual RNAs. Using this method we established Nascent Chain Tracking (NCT) which enabled us to study translational heterogeneity (Chapter 2 and Appendix A), the dynamics of mRNA interactions with stress granules (Chapter 3), and HIV-1 programmed ribosomal frameshifting (Chapter 4). K. Lyon created and contributed to the design of the cartoon schematic.

Chapter 2

Establishment of Nascent Chain Tracking

For the validation of the Nascent Chain Tracking (NCT) employed in this dissertation, of which I was a co-author, see Appendix A. The following chapter is based on a review article where I was first author. [1]

2.1 Probing translational dynamics in living cells with Fab and scFv

Recently, five research groups have independently established a methodology to visualize and quantify single RNA through multiple rounds of translation [37–40]. Remarkably, all solved the problem in essentially the same way: by using antibody-based probes that bind to repeat epitope tags. This burst of papers attests to the robustness of the approach, which allowed each group to bring preformed fluorescence to the growing nascent peptide chain without the need for the chain to fold, mature, or fluoresce on its own (Figure 2.1).

In general, each group amplified the signal from translation sites by fusing tags encoding tandem repeats of short, linear epitopes to the N termini of proteins of interest (POIs), as shown in Figure 2.1. This made it possible to recruit multiple fluorescent antibody-based probes to a single elongating nascent chain. Since more than one ribosome can translate an RNA at a time (in polysomes), the amplified signal from a single chain is multiplied by the number of elongating ribosomes, yielding amplification of amplification. To visualize this, four of the groups used the SunTag system [38–41]. Here, the probe is an scFv derived from a GCN4 antibody and evolved via ribosome display to tightly bind a 33-amino acid (aa) epitope [42]. This was later adapted for intracellular stability within yeast [43]. To prevent aggregation in mammalian cells, a protein G (GB1) domain was further fused to the C terminus of the scFv and the 33-aa epitope was shortened to just 19 aa (in

[1] This chapter comes from a review in which I was first author [31].

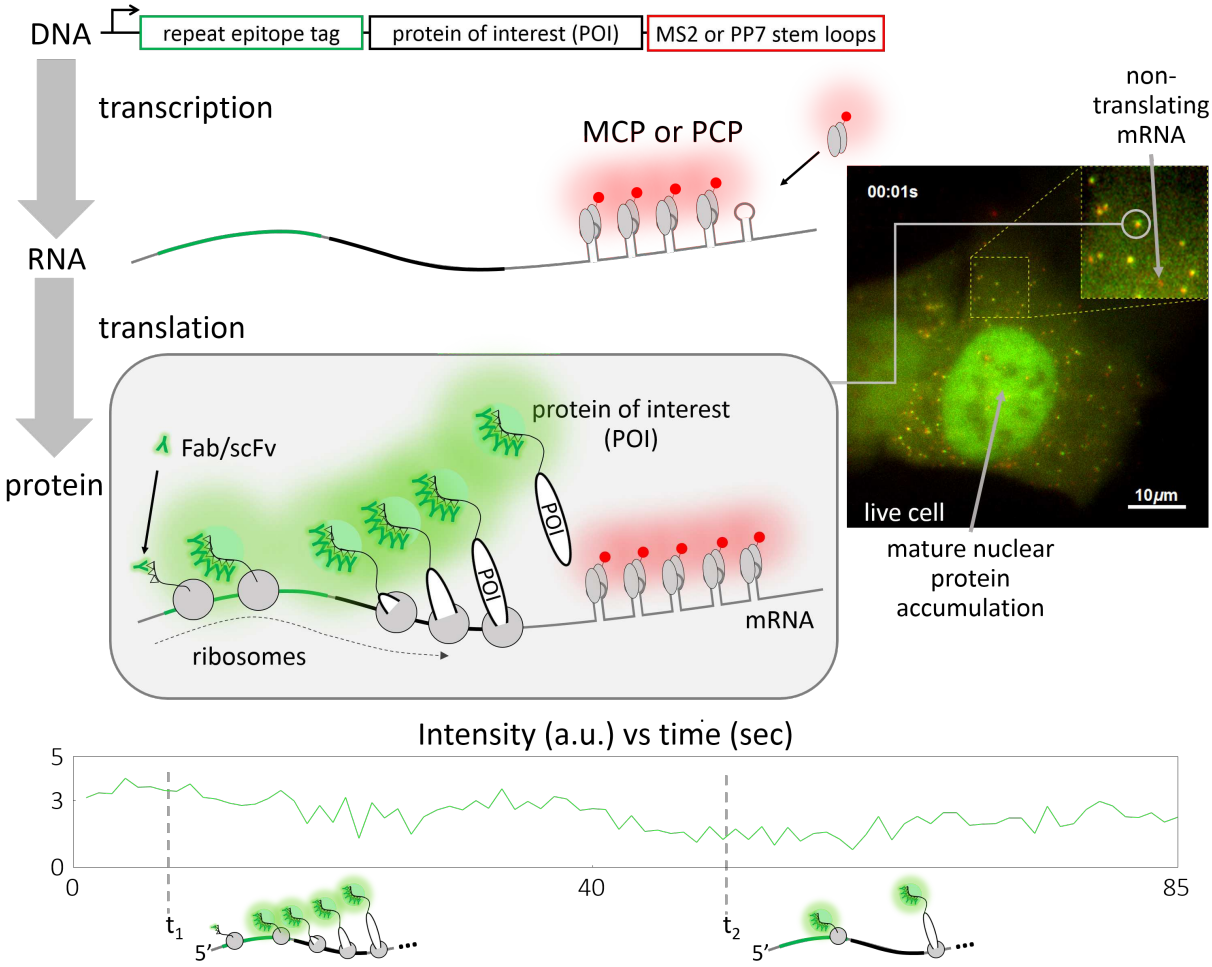


Figure 2.1: Imaging gene regulatory dynamics in live cells using antibody-based probes. To successfully image live-cell translation dynamics, 30 untranslated region (UTR) stem loops and a repeat epitope tag need to be encoded into the gene of interest. During transcription, the stem loops will form (gray loops) and quickly be bound by a coat protein (MS2 or PP7; gray ovals) fused to a fluorescent protein tag (red circle). This enables the visualization of RNA molecules in living cells. When these marked RNA are translated, the repeat epitopes (triangles) will emerge from the ribosome and be quickly bound by a fluorescent probe [antibody fragments (Fabs) or single-chain variable fragments (scFvs); glowing 'Y']. This will quickly bring light to the growing nascent chain of the protein of interest (POI) while it is being translated. The co-migration of the red RNA signal and the green nascent chain signal will mark sites of translation. A sample live-cell image is shown, where the green signal corresponds to Fabs labeling spaghetti monster-tagged lysine demethylase (KDM5B) and the red signal corresponds to MS2-labeled RNA. These signals colocalize to form yellow spots that mark the sites of translation. The green fluorescence time series from the circled yellow spot in the inset is shown below, along with cartoons depicting how ribosomes might be distributed along the RNA at the two specified time points. Scale bar 10 μm. Adapted from [31]. K. Lyon contributed to the design of the cartoon schematics and measured signals from the sample fluctuating translation site that T. Morisaki acquired.

version 4) [44]. These epitopes are repeated in the SunTag and separated from one another by short linkers (either 24 repeats [38, 40, 45] or 56 repeats [41], totaling 571 aa or 1419 aa, respectively). These linkers allow every epitope in the SunTag to be bound by an scFv, resulting in very bright fluorescence signals. In a similar but distinct manner, we used the spaghetti monster system to visualize translation [37]. Here, the antibody-based probes were Fabs that were generated from classic antibodies, either FLAG or HA [46, 47]. The spaghetti monster tag (325 aa) comprised ten FLAG (8 aa) or HA (9 aa) epitopes distributed throughout a 3D protein scaffold to prevent steric hindrance between bound Fab or full antibody [48]. In addition to the repeated peptide epitope tag, a repeated RNA stem loop tag (MS2 [27] or PP7 [28]) was also used by all groups to label and track the POI-encoding RNA. This was necessary to distinguish nascent peptide chains still being translated from mature protein, since only the former would colocalize with RNA.

To quantify the kinetics of translation, each group used a variety of experimental approaches. Despite these differences, there was consistency between the studies, as shown in Table 2.1. An important parameter was the translation elongation rate. This rate can be measured in several different ways, exactly analogous to how transcription elongation rates are measured in living cells using the MS2 RNA labeling system [49–51]. Whether measured by run-off inhibitor treatments [39–41], photobleaching [37, 38, 41], fluorescence correlation analysis [37, 38], or the tracking of RNA being translated by single ribosomes [39], all studies concluded that the translation elongation rate was on the order of 10 aa/sec. Importantly, this rate is close to what has been measured *in vitro* and also via ribosome profiling [21, 22]. Another key measurable was the number of nascent chains (i.e., ribosomes) per RNA. Each group measured this by calibrating the intensity of individual translation sites. Although this number varied more from group to group than did the elongation rate, the variability diminishes when considering the density of ribosomes along RNA. In this case, all groups measured about one ribosome every 200-900 nucleotides. By combining this density estimate with the measured elongation rate, it is possible to calculate the translation initiation time. Again, there was excellent agreement between studies, being somewhere around 30 sec.

Table 2.1: A comparison of live-cell measurements of single RNA translation kinetics. Live-Cell Measurements of Single RNA Translation Kinetics. Adapted from [31]. K. Lyon, T. Stasevich.

Feature	Ref, [37]	[38]	[39]	[40]	[41]
Cell Line	U-2OS	U-2OS and primary hippocampal neurons	U-2OS and HEK293	HeLa and primary hippocampal neurons	HeLa
Tags	Spaghetti monster 10x FLAG or 10x HA	24x SunTag	5x, 10x, and 24x SunTag	24x SunTag	32x and 56x SunTag
Probes	Bead-loaded FLAG-Cy3 or HA-A488 Fab	Stably expressed GCN4 scFv-GFP	Stably expressed GCN4 scFv-GFP	Stably expressed GCN4 scFv-GFP	Stably expressed GCN4 scFv-GFP
Elongation Rate	~10 aa/sec	~5 aa/sec	~3-6 aa/sec	~4 aa/sec	~13-18 aa/sec
Initiation Rate (sec/ribosome)	~30 sec	~50 sec (U-2 OS) or ~30 sec (PHN)	~30 sec	~20 sec	~15 sec
Nucleotides per ribosome	~200-900	~700	~300	~250	~600-800

Beyond the basic translation initiation and elongation kinetics that were quantified, different groups also observed a variety of other unique translational dynamics. One surprising observation was that RNA being translated can be highly mobile and move about the cell just as rapidly as RNA that are not being translated [37, 38, 41]. Since polysomes are so massive, this would suggest that the mobility is active rather than passive. Indeed, in neurons, many motoring RNAs were observed to be undergoing translation [38], in opposition to the earlier belief that transcripts are repressed while being motored to their destination [52]. Another interesting observation that emerged from different groups was that translation can be organized in higher-order structures comprising two or more translation sites clustered into small groups [37] or multi-RNA translation sites [41]. Other forms of spatial organization were also observed, including ER-bound transcripts, which displayed greatly restricted mobility [40], and neuronal transcripts, which localized to the far edge of distal dendrites [38].

All groups also perturbed translation using a variety of different treatments, including: (i) puromycin [53], which caused premature release of the fluorescently labeled nascent peptide chains; (ii) cycloheximide [37, 38], which stalled translation (although the number of ribosomes per transcript did not notably increase, as anticipated by earlier work in fixed cells [54]); (iii) harringtonin [40, 41], which prevented new ribosome initiation (permitting run-off assays); and (iv) DTT and arsenic stress, which dynamically altered start codon selection in a burst [40]. The origin of these diverse observations is not yet clear. Some possibilities include differences in 3' and 5' untranslated regions (UTRs), codon usage, or the local cellular environment.

Although the SunTag scFv and Fab-based translation imaging systems are similar, they have distinct advantages and disadvantages. The SunTag scFv can be stably expressed in cells and even put into an endogenous locus using CRISPR [41]. Thus, imaging could in principle be done throughout whole transgenic animals or within sensitive cell types (such as neurons [38, 40]) that are difficult to load with Fab. By contrast, Fabs currently offer more flexibility for multi-color imaging applications because Fabs can be generated from a range of commercially available monoclonal antibodies [37]. Therefore, it is easy to make orthogonal probes (such as FLAG and HA), whereas an orthogonal scFv probe that can be used in combination with the SunTag system has been developed but is much more challenging to design and validate in a living system [55].

Chapter 3

Quantification of Single RNA Interactions with mRNP Granules During Stress

The following chapter is based off of work where I was a co-author. [2]

3.1 Translation ends prior to interaction with stress granules and resumes following disassembly

During cellular stress translation is halted for many RNA. These RNAs go on to accumulate in ribonucleoproteins (RNPs) granules, e.g., stress granules (SGs) and p-bodies (PBs), which helps ensure translational priority of stress responsive genes [56]. Experiments employing fixed-cell immunostaining and fluorescence microscopy of fluorescently labeled SG proteins have established mRNA within stress granules are translationally repressed and that RNA-binding proteins within SGs have biphasic mobility being either in a diffuse state (mobile, liquid-like) or a bound state (immobile, nanocores) [57–62]. Although these past studies have shown that mRNAs accumulate within SGs and are translationally repressed, the dynamics of translating and non-translating mRNA within and near the granule surface remain poorly characterized. Establishing these general interaction principles are important as they play critical roles in stress response, maternal mRNA storage in early development, synaptic plasticity, tumor progression, and neurodegeneration [63–71].

To simultaneously visualize and quantify the exit of single mRNAs from translation and their dynamic interactions with ribonucleoprotein (RNP) granules in living cells, we adapted the newly

[2]Chapter 3 is based on the following work [35], of which I was a co-author, with the following author contributions: S.L.M. and R.P. conceptualized the study; S.L.M., T.M., T.J.S., A.K. and K.L. developed and designed methodology; T.M. and T.J.S. developed and implemented software; S.L.M. validated findings; T.M., T.J.S., S.L.M. and A.K. performed formal analyses; S.L.M., A.K., K.L., T.M. and T.J.S. performed experiments and/or collected data; R.P., T.J.S., T.M., A.K., K.L. and S.L.M. provided resources; T.M., T.J.S. and S.L.M. provided data curation; S.L.M. and R.P. wrote the original draft; R.P., S.L.M., T.J.S., T.M. and A.K. reviewed and edited the drafts; T.M., T.J.S., S.L.M. and A.K. visualized data; R.P., T.J.S., T.M. and S.L.M. supervised the project; S.L.M. managed and coordinated the project. See Appendix D for permission of use.

developed nascent chain tracking technique [37]. The translation of individual mRNAs labelled with fluorescent MS2 coat protein (MCP-HaloTag with Janelia Fluor 646 (JF646) dye [72]) was monitored via the binding of fluorescent antibody fragments (Fab conjugated with the Cy3 dye) to epitopes at the N terminus of the nascent peptide (Figure 3.1A). Thus, translating mRNAs were labelled with both MS2 coat protein (MCP) and Fab, whereas non-translating mRNAs were only labelled with MCP. By imaging these constructs in U-2 OS cells stably expressing the stress granule (SG) marker green fluorescent protein (GFP)-G3BP11 [64], we could examine the translation status of single mRNAs during arsenite stress in relation to their interactions with individual SGs.

We first imaged cells at a lower temporal resolution (one cell volume every 3 min for 1 h) to determine the timing of translation repression during stress. To avoid background from single mature proteins and limit photobleaching, we lowered the laser powers so that only the mRNAs translating in polysomes were detectable. As the majority of translation output is thought to occur in polysomes, we considered mRNAs without a detectable associated nascent peptide to be in a translationally repressed state. The translation of single KDM5B reporter mRNAs declined following stress, with SGs forming after 10 min of stress (Figure 3.1B). The interaction of KDM5B mRNAs with SGs lagged slightly behind SG assembly, with about 30-40% of the mRNAs associating with SGs 40-60 min after the addition of arsenite, which is similar to what is predicted for endogenous KDM5B mRNAs from the SG transcriptome (~39%) [73]. Ninety-eight per cent of the KDM5B SG-associated mRNAs were translationally repressed, which implies that the repression of translation is a general prerequisite for the recruitment of mRNA to SGs. Surprisingly, 1-2% of SG-associated mRNAs retained nascent chains, suggesting that mRNAs can interact with SGs while still associated with polysomes. To verify these observations, we imaged samples at a higher temporal resolution (one cell volume every 2 s for 10 min) between 10 and 30 min post-stress, a period when both translating mRNAs and SGs are observed (Figure 3.1B). This confirmed that most of the mRNAs entered SGs in a translationally repressed state (Figure 3.1C). Between one and two per cent of the mRNAs interacting with SGs were associated with nascent chains, but these mRNAs only interacted

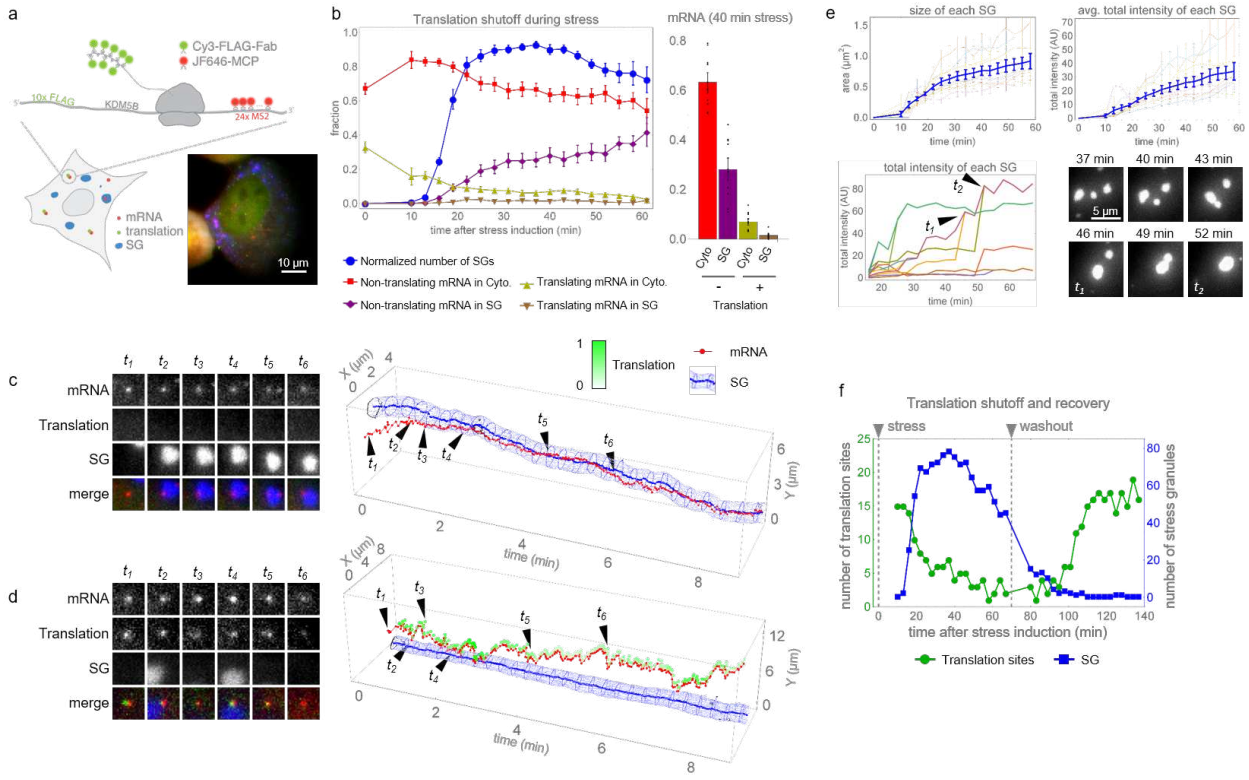


Figure 3.1: mRNAs are translationally repressed before entering SGs and resume translation following SG disassembly. a, Single mRNAs with 24x MS2 stem loops in the 3' UTR were visualized with JF646-MCP (red) and translation was observed by anti-FLAG Cy3-Fab (green) binding to 10x FLAG tags in the N-terminal KDM5B open reading frame (ORF) in U-2 OS cells that express the SG marker GFP-G3BP1 (blue). Inset, a representative cell (from ten we analysed in detail in b) is shown. Scale bar, 10 μm. b, Simultaneous detection of mRNA localization, translation activity and SG formation in arsenitestressed cells. Left, normalized number of SGs per cell (that is, fraction of the maximum number of SGs observed in each cell throughout the stress), fraction of cytoplasmic (Cyto.) non-translating and translating mRNAs, and fraction of SG-localized translating and non-translating mRNAs. Right, fraction of cytoplasmic or SG-localized translating or non-translating mRNAs at 40 min post-stress are shown. The mean $\pm$ s.e.m. is shown from n = 10 cells collected from three independent experiments. c,d, Representative images of single mRNAs (red), their translation activity (green) and SG (blue) interactions; scale bars, 1 μm (left); and graphical representations of the interactions with the intensity of translation foci represented as white-green (right). c, A representative long-term mRNA-SG interaction (> 3 min) from 82 non-translating mRNAs tracked from n = 9 cells collected from three independent experiments. d, An example of transient translating mRNA-SG interactions from 334 translating mRNAs tracked from n = 9 cells collected from three independent experiments. e, Growth of SGs over time, as average SG size (upper left) and average SG intensity (upper right). Growth of individual SGs (lower left) with fusion events shown graphically (t_1 and t_2) and as a representative time series (lower right; scale bar, 5 μm). Mean $\pm$ s.e.m. is shown from n = 10 cells collected from three independent experiments. f, Representative data from a single cell (from n = 4 cells collected from two independent experiments) showing the resumption of translation after arsenite washout. The number of translation foci (green) and SGs (blue) during stress (0-70 min) and following washout (80-140 min) are shown. Taken from [35]. K. Lyon prepared reagents, designed experiments, and performed experiments contributing to Figure 3.1a-e

with SGs transiently for a few seconds (Figure 3.1D). Out of a total of 336 interactions detected, only two transcripts associated with nascent chains interacted with SGs for longer than two minutes.

Translation repression positively correlates with the growth of SGs, which increases throughout the stress response [57, 74, 75]. Stress granules grew by both incremental accumulation of material and SG fusion (Figure 3.1E). Incremental growth dominated during the early phases of stress, up to ~40 min post-stress, when the number and size of SGs continued to increase, whereas fusion events dominated at later time points, after ~40 min post-stress, when the number of SGs stopped increasing but SG size continued to increase. The growth of SGs also correlated with increased KDM5B mRNAs in SGs, which is consistent with mRNA recruitment contributing to SG growth.

During recovery from arsenite stress, SG disassembly (quantified as the number of individual SGs) occurred a few minutes before translation resumed in single cells (Figure 3.1F). We verified that these translating spots in the recovery phase were bona fide translating mRNAs by observing their disappearance following puromycin treatment (data not shown). KDM5B mRNAs did not resume translation until the SGs had dissolved completely, suggesting that SGs must disassemble before translation resumes during recovery from stress.

3.2 RNAs interact transiently and stably with stress granules

To determine how mRNAs interact with RNP granules, we measured the duration of individual SG interactions with mRNAs that encoded KDM5B, H2B or p300 (Figure 3.2A,B). In all three cases, the probability that a SG-mRNA interaction lasted a set duration of time (that is, the interaction time survival probability) could be fit by a two-component model with fast interactions on the order of seconds and slow interactions on the order of many minutes (Figure 3.2C,Figure 3.3A). These results suggest that mRNAs frequently "sample" SGs and occasionally enter a stable association.

Generally, three parameters influenced the nature of the SG-mRNA interactions. First, we compared the dynamic interactions of reporter mRNAs containing the KDM5B ORF (4,632 nt), the shorter H2B ORF (375 nt) or the longer p300 ORF (8,265 nt) with SGs by determining the dwell time of mRNAs in SGs and found that mRNA ORF length correlated with the duration of

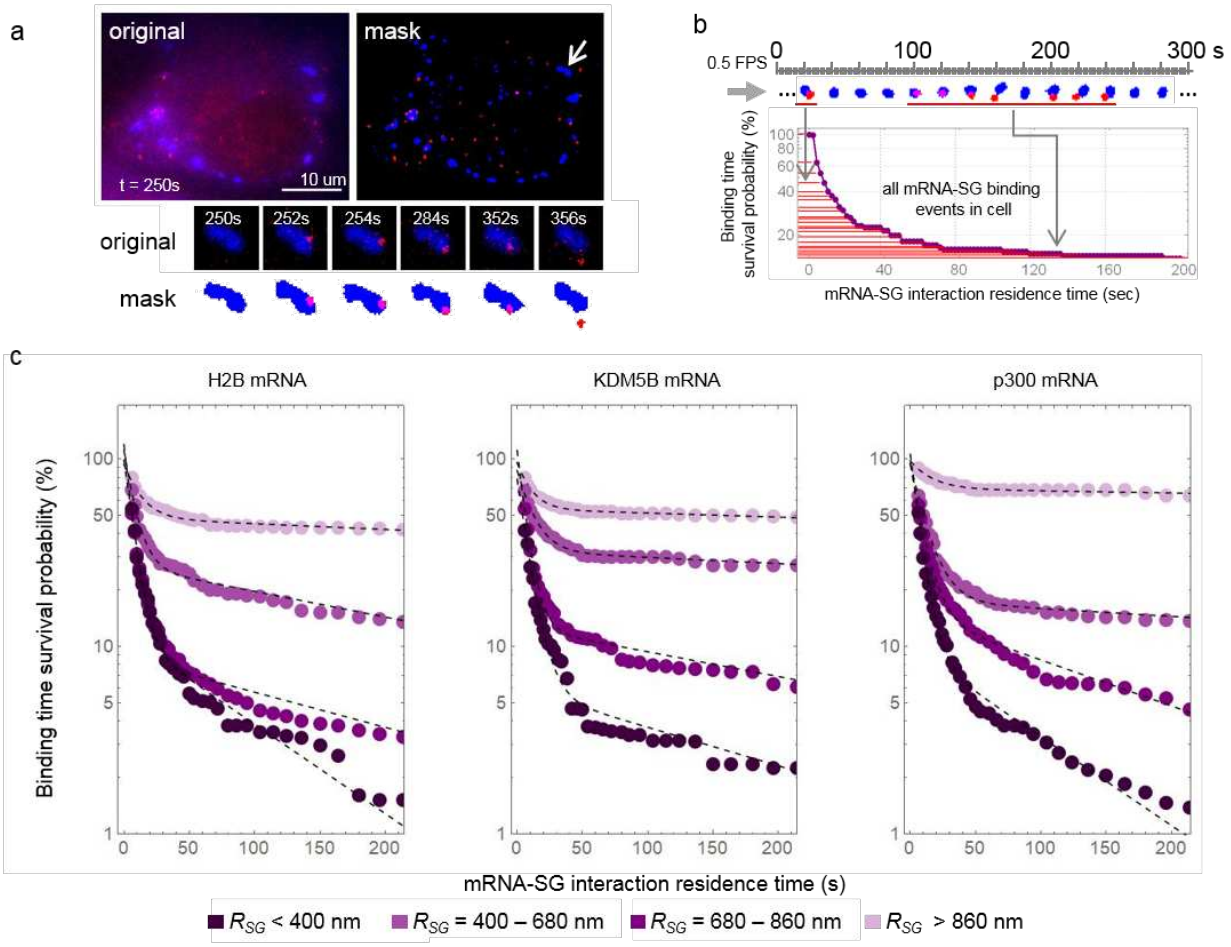


Figure 3.2: mRNAs interact transiently and stably with SGs. a, Determination of mRNA-SG interaction times. Representative image (from n = 9 cells expressing KDM5B mRNA that we analysed in detail in c) showing mRNAs (red) and SGs (blue) (upper left; scale bar, 10 μm) and the masked image (upper right) with a representative mRNA-SG interaction (arrow) shown below (scale bar, 1 μm). b, Survival probability distribution of mRNA-SG interaction times (red) from one representative cell from n = 9 cells expressing KDM5B mRNA that we analysed in detail in c. The diagram above the graph depicts representative masked mRNA (red) and SG (blue) interactions derived from imaging cells at 0.5 frames per second (FPS), with grey arrows indicating how these interactions are represented graphically below. c, The binding-time survival probability of H2B (left; data were calculated from 492 tracked SGs from n = 11 cells collected from three independent experiments), KDM5B (middle; data were collected from 409 tracked SGs from n = 9 cells collected from four independent experiments) and p300 (right; data were calculated from 824 tracked SGs from n = 16 cells from four independent experiments) mRNA-SG interactions are shown partitioned by SG size (legend at the bottom). Adapted from [35]. K. Lyon prepared reagents and performed experiments contributing to Figure 3.2a-c (for H2B mRNA and KDM5B mRNA).

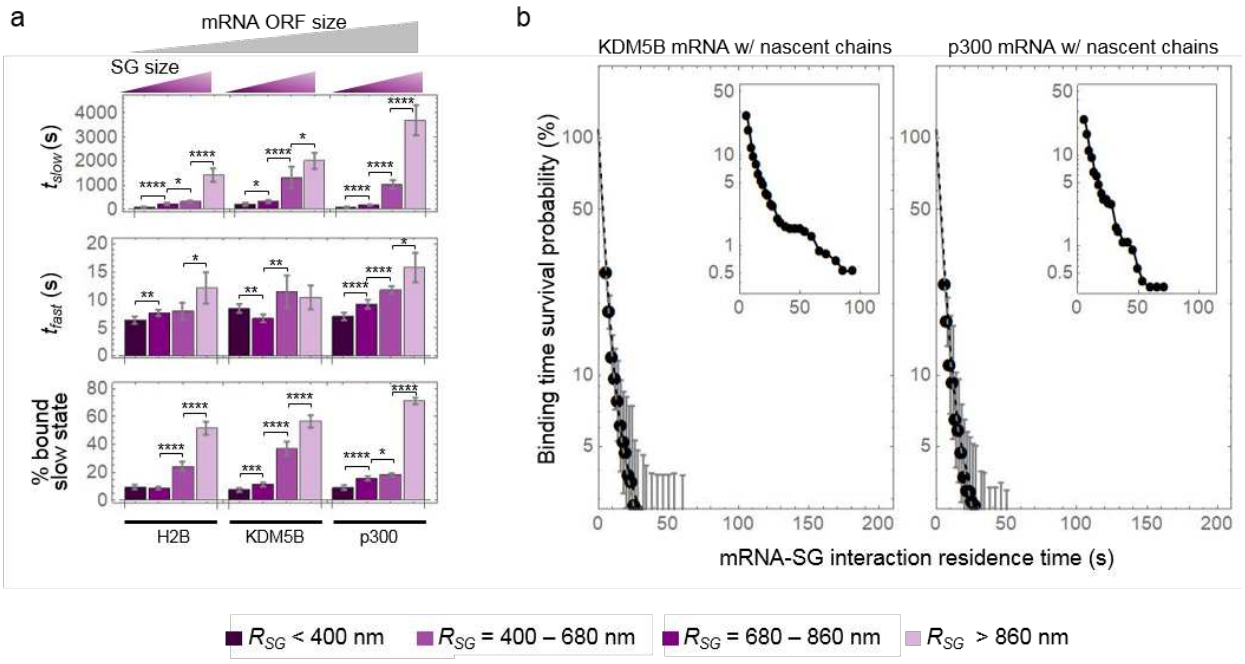


Figure 3.3: Longer mRNAs interact longer with SGs and can engage with polysomes. a, The extracted mRNA-SG interaction data in c were fit to a two-component model, resulting in average (± 90% confidence interval (CI)) fitted slow and fast mRNA-SG interaction times, t_{slow} (top) and t_{fast} (middle), respectively, and percentage mRNA bound to SGs in the slow interaction mode ("slow state"). The fitted results are shown for H2B, KDM5B and p300 mRNAs for a given effective SG radius (RSG) depicted as dotted lines. *P < 0.05, **P < 0.01, ***P < 0.005, ****P < 0.001 and d.f. = 44; two-sided t-tests. Fitting was performed once to the collective data set shown in Figure 3.2c. b, The average binding-time survival probability (± 90% CI) of KDM5B (left; data were calculated from 326 tracked SGs from n = 9 cells collected from three independent experiments) and p300 (right; data were calculated from 336 tracked SGs from n = 10 cells collected from four independent experiments) mRNAs associated with nascent chains is shown with graphical insets showing the full distribution (adjusted x and y axes). Adapted from [35]. K. Lyon prepared reagents and performed experiments contributing to Figure 3.3a and b (H2B and KDM5B mRNAs only).

interactions with large SGs (Figure 3.2C,Figure 3.3A; t_{slow}(H2B) = 1,400 $\pm$ 300 s, t_{slow}(KDM5B) = 2,000 $\pm$ 300 s and t_{slow}(p300) = 3,700 $\pm$ 600 s). We also observed that the duration of the slow interactions and the fraction of mRNAs engaged in slow interactions with SGs for the H2B, KDM5B and p300 mRNAs varied depending on the size of the SG. Larger SGs had more and longer stable interactions with mRNAs than smaller SGs (Figure 3.2C,Figure 3.3A). The observation that longer mRNAs interact more stably with SGs explains why longer mRNAs were more likely to be highly enriched in the steady-state SG transcriptome [73, 76]. Finally, we observed that KDM5B and p300 mRNAs still engaged with polysomes, as assessed by the detection of associated nascent peptides, and could interact with SGs transiently but did not engage in the stable interaction mode (Figure 3.3E). We did not analyse translating H2B transcripts because they were more difficult to track due to their dim signals and low numbers per cell, even before stress [37]. Translating mRNAs in the cytoplasm moved slower than translationally repressed mRNAs (Figure 3.4A), ruling out the possibility that the transient association of translating mRNAs with SGs was due to faster overall mobility. Together, these observations suggest that mRNA length, SG size and mRNA translation status are strongly correlated with mRNA partitioning into SGs, although there are probably other mRNP factors that influence the dynamics of mRNA-SG interactions.

3.3 mRNAs interact transiently and stably with PBs and traffic bidirectionally between PBs and SGs

To determine whether other RNP granules also exhibit bimodal interactions with mRNAs, we examined the dynamics of mRNA interactions with P-bodies (PBs), marked with mRFP-DCP1a in stressed U-2 OS cells that also expressed the SG marker GFP-G3BP1. We observed that H2B, KDM5B and p300 mRNAs all showed rapid and stable interactions with PBs (Figure 3.5A), mRNA ORF length correlated with the duration of mRNA interaction with PBs, as p300 interacted most stably with PBs and H2B interacted most unstably with PBs (Figure 3.5A,B), and the degree of stable interactions for all mRNAs analysed increased with the size of the PB (Figure 3.5A,B). These observations indicate that mRNAs can dock and undock to SGs and PBs through transient

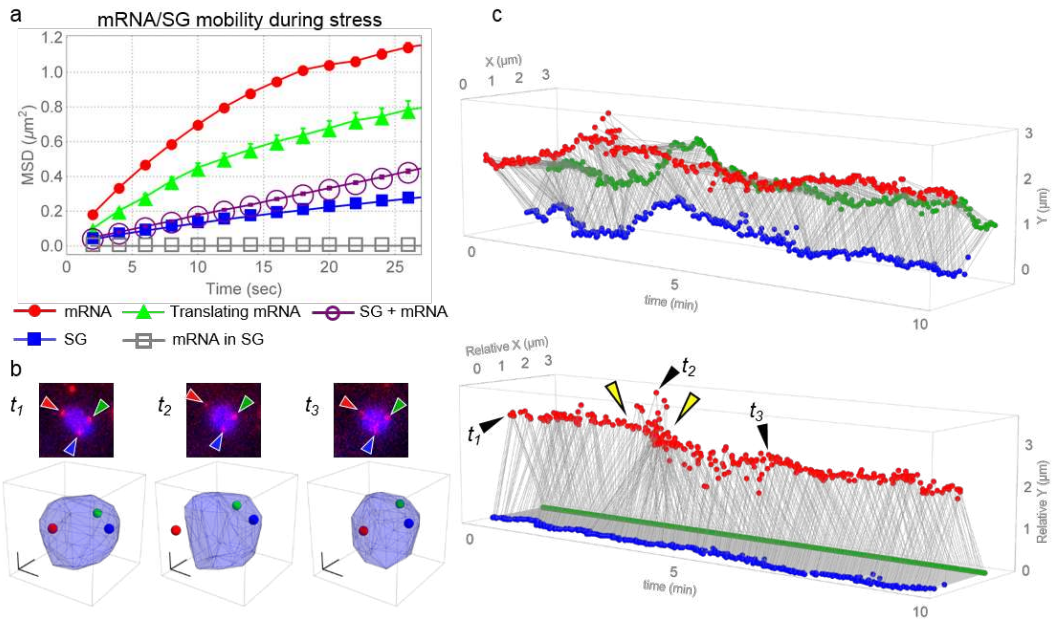


Figure 3.4: mRNAs can be rigidly positioned within SGs and/or tethered to them. a, Average mean squared displacement (MSD; ± 90% CI) of cytoplasmic KDM5B mRNAs (fitted diffusion constant (D) = 0.016 ± 0.004 $\mu m^2 s^{-1}$), mRNAs with nascent chains (translating RNA; fitted D = 0.011 ± 0.002 $\mu m^2 s^{-1}$), SGs (fitted D = 0.003 ± 0.0006 $mum^2 s^{-1}$), SGs containing mRNAs (fitted D = 0.004 ± 0.0004 $\mu m^2 s^{-1}$) and mRNAs in SGs (fitted D = 10 ± 2 $nm^2 s^{-1}$) from n = 7 cells from two independent experiments; 1,243 mRNAs, 108 mRNAs with nascent chains, 1,049 SGs, and 92 SGs containing mRNAs were analysed, and 3 mRNAs were tracked in three dimensions (3D) within one SG. b, Representative time series (from n = 2 cells collected from two independent experiments) showing three KDM5B mRNAs (top; foci marked by red, blue and green arrows; scale bar, 1 μm) and their 3D positions (bottom; red, green and blue dots) in a SG. c, Projected two-dimensional positions (top) and relative 3D positions (bottom) of the three mRNAs in b plotted over time (green mRNA position fixed and blue mRNA oriented relative to green mRNA). The frames in b t_1, t_2 and t_3 are indicated (black arrows). The yellow arrows indicate exit and entrance of the red mRNA from the SG (from n = 2 cells collected from two independent experiments). Adapted from [35]. K. Lyon prepared reagents and performed experiments contributing to Figure 3.4a-c.

interactions and in some cases form a stable set of interactions that "locks" the mRNA into the granule for prolonged periods. A stable association could be more prevalent on longer mRNAs and in larger SGs and PBs due to increased sites for additional interactions. This model implies that the association of an mRNA with a granule can be influenced by *cis* or *trans* inputs that modify either the docking/undocking step or the rate of entry into a stable, locked interaction state, which is supported by the observation that knockdown of the RNA binding protein LARP1 can reduce the association of some mRNAs with PBs and SGs [77].

Non-translating mRNAs can accumulate in both SGs and PBs. A general model is that following translation repression, mRNAs first associate with SGs and then can be sorted for targeting to PBs [78]. In contrast, it has been proposed that mRNAs move from PBs to SGs following glucose deprivation in yeast [79]. To determine whether there is a preferred movement of mRNAs between SGs and PBs, we simultaneously imaged SGs, PBs and KDM5B mRNAs during stress. We observed mRNA exchange between SGs and PBs in a bidirectional manner (Figure 3.5C,D). These interactions included cases of mRNAs transitioning from a SG to a PB and back to a different SG within as little as ~30 s (Figure 3.5C) and instances where the mRNA was localized between a PB and a SG for a few minutes (Figure 3.5D). Transcripts were considered to be associated with a SG or PB when they moved with these granules over several frames (or at least 4 s). These rapid bidirectional movements between SGs and PBs indicate that there is no obligate path for mRNA movement between these granules, which also demonstrates that an individual mRNA with its associated proteins is capable of interacting with either SGs or PBs, or exchanges proteins that enables SG or PB interaction within seconds.

3.4 mRNAs can be rigidly positioned within SGs

It has been proposed that SGs, and other RNP granules, are liquid-like compartments with components showing rapid diffusion within the compartment [59, 60] and/or contain liquid-like areas surrounding densely compacted "cores" [57, 58]. Those mRNAs that were stably associated with RNP granules for many minutes adopted the motility of SGs (Figure 3.4A) and could be

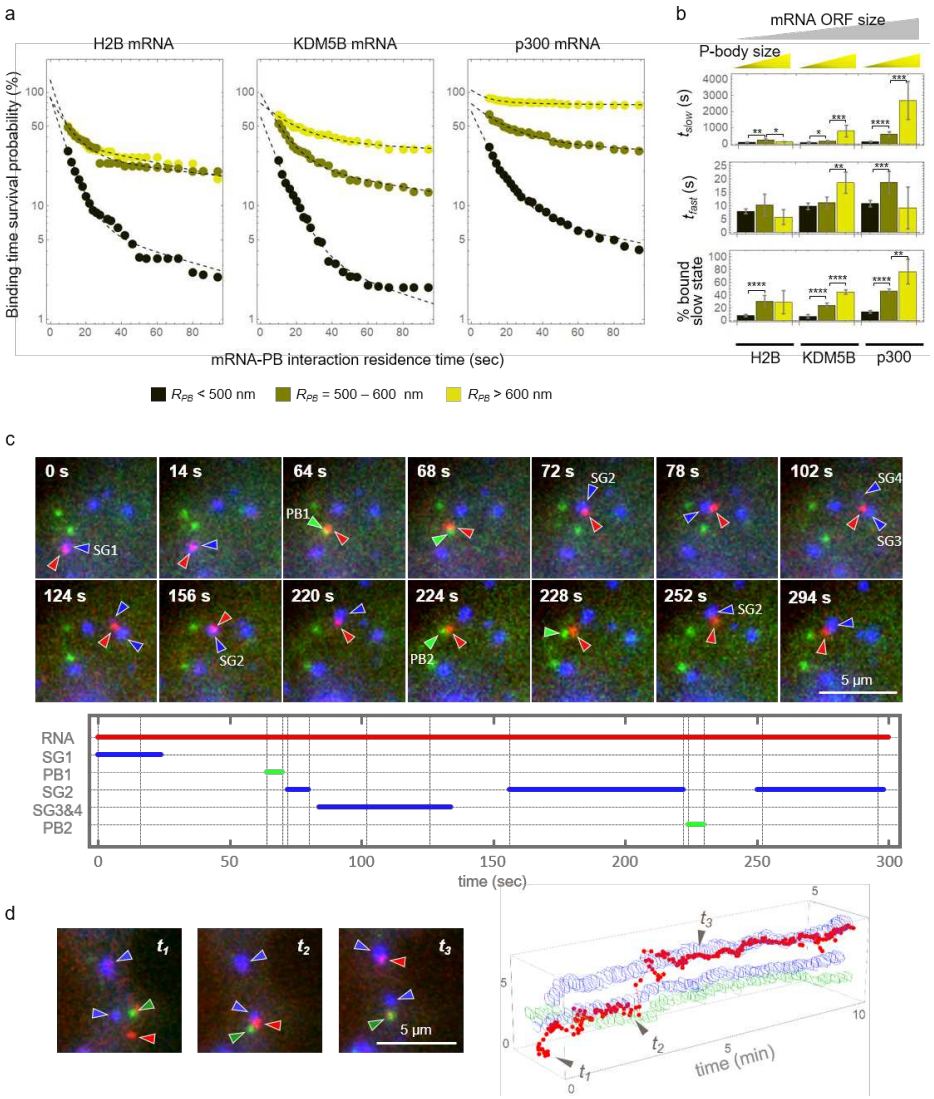


Figure 3.5: mRNAs interact transiently and stably with PBs in stressed cells and traffic bidirectionally between PBs and SGs. a, Average binding-time survival probability distributions of H2B mRNA-PB (left; data were calculated from 106 tracked PBs of n = 4 cells from one experiment), KDM5B mRNA-PB (middle; data were calculated from 137 tracked PBs of n = 7 cells collected from three independent experiments) and p300 mRNA-PB (right; data were calculated from 244 tracked PBs of n = 16 cells collected from four independent experiments) for a given effective PB radius (RPB). b, The data in a were fit to a two-component model, resulting in average fitted fast (middle) and slow (top) interaction times; the percentage of mRNA bound to PBs in the slow interaction mode are shown. The fitted results are shown ($\pm$ 90% CI) for H2B, KDM5B and p300 mRNAs for each effective PB radius (legend in a). *P < 0.05, **P < 0.01, ***P < 0.005, ****P < 0.001, d.f. = 31; two-sided t-tests. Fitting was performed once to the collective data set shown in a. c, Representative images (top; from n = 6 cells collected from four independent experiments; scale bar, 1 μm) of a single KDM5B mRNA that interacted with four SGs and two PBs. The duration of each mRNA-RNP granule interaction was plotted (bottom). d, A representative mRNA trajectory (from n = 3 cells collected from two independent experiments) between two SGs and a PB visualized by plotting the position of the mRNA, SG and PB over time (right). Example cropped images corresponding to time points t_1 to t_3 are shown (left; scale bar, 1 μm). SGs, blue; mRNA, red and PB, green; the same colour codes apply to the arrows in c. Taken from [35]. K. Lyon prepared reagents and performed experiments contributing to Figure 3.5a-d (H2B and KDM5B mRNAs only).

observed at the surface or embedded within SGs (Figure 3.4B). To examine the liquid-like nature of SGs, we examined the movement of mRNAs relative to one another within SGs. We observed two cases where the relative localization of mRNAs, as assessed by quantifying the distance and angle between three individual KDM5B mRNAs in one SG, remained relatively constant over time (Figure 3.4B,C), consistent with limited movement of the mRNA within the larger assembly. Quantification revealed the intra-SG diffusion of mRNAs to be 280 times slower than the SG as a whole and over 1,600 times slower than free mRNAs within the cytoplasm (Figure 3.4A). These observations demonstrate that SGs contain solid or gel-like components and mRNAs within large SGs can be rigidly positioned, with relatively limited movements of mRNAs.

Unexpectedly, we observed two cases where the MCP signal, which represents the 3' end of a KDM5B mRNA, exited the SG and then returned to approximately the same position within the SG (yellow arrows, Figure 3.6A). This observation suggested that the 3' end of this mRNA might exit the GFP-G3BP1-marked portion of the SG while the remainder of the mRNA is still within the SG. Consistent with the body of an mRNA extending beyond the G3BP1-marked SG, we observed that endogenous RNAs that localize to SGs (AHNAK and NORAD) are often detected near, but not in, SGs in U-2 OS cells by single-molecule fluorescence in situ hybridization (smFISH; Figure 3.6A) with the shorter NORAD RNA (~5 kb) closer to the periphery of a SG than the longer AHNAK RNA (18 kb). Furthermore, using smFISH probes to the 5' or 3' ends of the AHNAK mRNA, we observed that ~15% of ~1,000 AHNAK mRNAs were at the SG periphery with one end of the mRNA outside the SG while the other end co-localized with GFP-G3BP1, with an approximately two-fold bias for the 3' end to be within the SG (Figure 3.6B). These observations suggest that the surfaces of these membraneless organelles are not uniform and can have RNP extensions into the cytosol beyond boundaries as defined by protein components. Such surface extensions could thus provide a set of interactions to promote the fusion of smaller SGs and docking of SGs and PBs, which is consistent with our observations of individual mRNAs localizing between a PB and a SG for several minutes (Figure 3.5D).

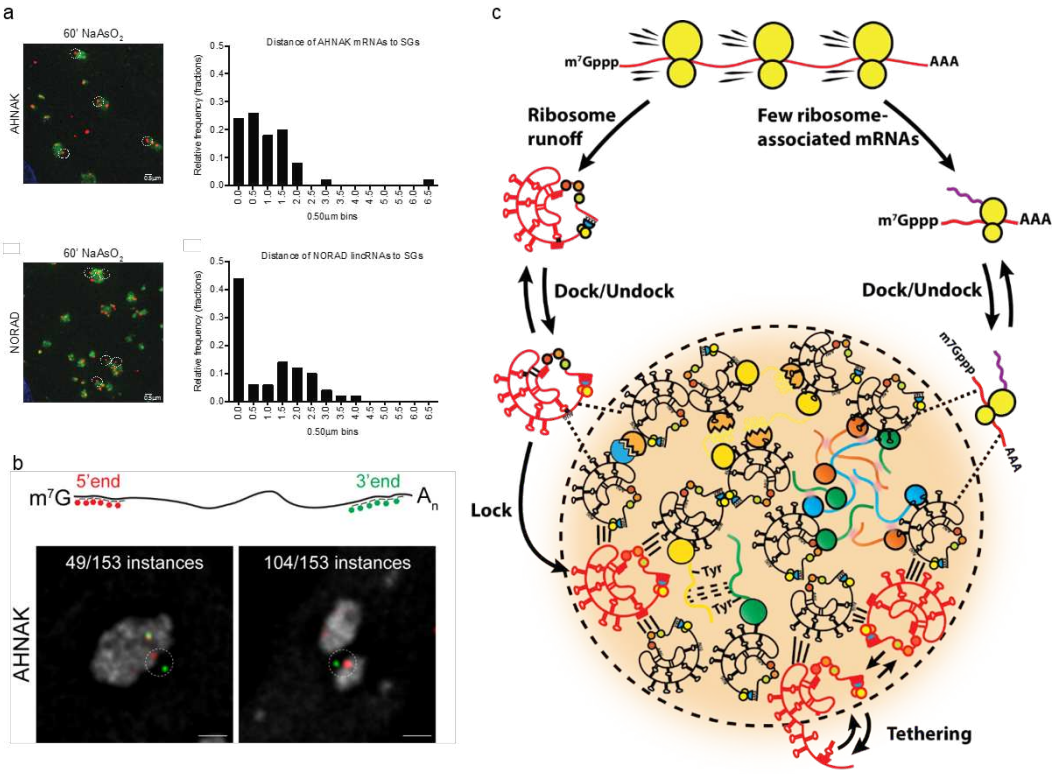


Figure 3.6: The 3' end of mRNA docks into SGs; overall mRNA-SG interaction model. a, Representative immunofluorescence smFISH images of AHNAK (n = 4 cells) and NORAD (n = 4 cells) RNAs (red) and SGs (G3BP1, green) at 60 min post-arsenite stress (left); white circles indicate RNAs clustered near SGs. The images shown represent data from n = 4 cells collected from one experiment. Scale bar, 0.5 μm. The relative frequency of distances of 50 AHNAK (top right) and NORAD (bottom right) RNAs to the nearest SG are shown. b, Schematic of AHNAK smFISH probe positions (top). Representative smFISH images where one end of the AHNAK mRNA was outside the SG while the other end was inside the SG are shown below. The 5' probes are indicated in red, 3' probes in green and SGs in grey; n = 14 cells collected from one experiment. Scale bars, 1 μm. c, Model depicting dynamic mRNA-SG interactions. Translation repression causes ribosomes to run off transcripts that can then interact transiently with SGs via docking and undocking. Some transcripts then enter a stable association (lock) with SGs and may engage in multivalent interactions with other RNAs and proteins. The transcripts may be tethered to the SG surface, possibly facilitating the growth of SGs or docking of PBs. The transcripts that remain in translation complexes can only transiently interact with SGs. Adapted from [35]. A. Khong designed experiments, performed experiments, and analyses for Figure 3.6a-b and R. Parker and S. Moon designed and created the cartoon schematic

Chapter 4

Live-Cell Single RNA Imaging Reveals Bursts of Translational Frameshifting

The following chapter is based off of work where I was first author. [3]

4.1 Visualization of multiple frames, a useful tool for studying regulatory translation dynamics

Frameshifting is a fundamental biological process in which a ribosome translating an RNA "slips" by either +1 (-2) or -1 (+2) nucleotide(s), resulting in the translation of an entirely different peptide sequence from that point forward. While frameshifting is generally detrimental to protein fidelity [80, 81], the process effectively creates two distinct proteins from a single RNA [82, 83]. Viruses exploit this aspect of frameshifting to minimize their genomes and to successfully replicate in host cells [84–89]. Endogenous mammalian frameshifting has also been found to occur at genes such as the Paternally Expressed 10 (PEG10) and, the polyamine-stimulated frameshifting gene, Ornithine Decarboxylase Antizyme 1 (OAZ1) [82, 90].

A prototypical example of programmed ribosomal frameshifting, a type of frameshifting that uses unique RNA structures, is the virus HIV-1. This virus has infected over ~70 million people in the past several years averaging ~5,000 infections per day [91]. HIV-1 utilizes frameshifting to translate the Gag-Pol proteins from a single viral RNA [92]. This programmed frameshift of HIV-1 has been extensively studied and characterized using *in vitro* and population-based methods. One thoroughly studied example is the -1 programmed ribosomal frameshifting (-1PRF)

[3] Chapter 4 is based on the following work [36], of which I was first author, with the following author contributions: Conceptualization, TJS, KL, and BM; Methodology: KL, LA, TM, BM, and TJS; Software, KL, LA, TM, BM, and TJS; Formal Analysis, KL and LA; Experimentation, KL, TM, and TJS; Computational modeling, LA and BM; Resources, BM and TJS; Writing, Review, and Editing, KL, LA, TM, BM, and TJS; Supervision, BM and TJS; Funding Acquisition, BM and TJS.

of HIV-1. Through rigorous mutational analyses the -1PRF of HIV-1 was found to be comprised of two specialized RNA elements: a "slippery" sequence followed by a downstream stem-loop structure [93, 94]. Together, these structures are known as a frameshift stimulatory sequence (FSS). Further characterization of the FSS concluded that it is capable of pausing elongating ribosomes upstream of its stem-loop structure. Although this pause is necessary for frameshifting, it alone is insufficient to cause frameshifting [95, 96]. The frameshift translational efficiency induced by the FSS of HIV-1 has been measured to be around 5-10%. This efficiency reflects the percentage of total -1 mature protein product accumulation and is commonly measured using a luciferase bioluminescent reporter [85, 89, 97, 98]. These experiments are done by inserting the HIV-1 FSS upstream of a luciferase encoding region and measuring the bioluminescent ratio of -1 frame to 0 frame. This assay has been useful in uncovering trans-acting factors that can influence the frameshifting translational efficiency of HIV-1 [99] and uncovering that the -1PRF of HIV-1 is functionally distinct from other -1PRF such as the mammalian gene PEG10 [88]. Recent studies, employing FRET and ensemble quench-flow, reveal that ribosomes can adopt non-canonical intersubunit rotated states upon encountering these -1PRF RNA structures allowing a frameshift to be more favored [100, 101]. These non-canonical ribosomal rotational states, which can be amplified by specialized -1PRF RNA structures, have also been implicated to influence overall translation fidelity [102].

Although the FSS of HIV-1 has been extensively characterized, many basic questions about frameshifting remain unresolved. In particular, it is not clear how heterogenous frameshifting is from one RNA to another, nor is it clear if single RNA continuously frameshift in a constitutive fashion or if instead they frameshift in prolonged bursts, as has been observed for transcription [103] and translation [38]. Finally, the localization of frameshifting has never been investigated, so it is not clear if frameshifting occurs all throughout the cell or is instead preferentially localized to specific sub-cellular regions. Furthermore, in the case of HIV-1 Gag-Pol previous assays have shown that 5-10% of translated protein product is frameshifted [85, 89, 97, 98]. Whether or not this occurs constitutively and with equal probability on all HIV-1 RNA or if instead it occurs on a specialized subset that are in the right place, at the right time, and with the right factors remains

to be determined. Therefore, observing HIV-1 frameshifting mechanisms at the single-RNA level could provide fresh insight into this process and yield alternative strategies to perturb this process in living cells.

To visualize and quantify single RNA frameshifting dynamics in living cells we designed a multi-frame tag strategy. This strategy uses multi-frame repeat epitopes, complementary high-affinity fluorescent probes that selectively bind the epitopes, and multicolor single-molecule microscopy. This allows us to simultaneously monitor the translation of single RNA into two unique nascent polypeptide chains encoded in shifted open reading frames. Inserting the HIV-1 FSS into this construct uncovers unexpected heterogeneity in the production of frameshifted product and implicates a bursty frameshifting mechanism. Besides frameshifting, our technology can now be used to examine other translational regulatory dynamics, including upstream open reading frame selection, non-canonical initiation, and ribosomal shunting. In particular, the high spatiotemporal resolution of our technology makes it possible to detect and quantify even the smallest sub-populations of translating RNA. Whereas ribosome profiling [21] and single-cell assays [104] can provide detailed snapshots of the average translational state of a cell, they lack the spatiotemporal resolution required to capture translation dynamics at the single-RNA level. We therefore anticipate multi-frame nascent chain tracking will be a powerful tool to dissect complex translational regulatory dynamics in living cells and organisms.

4.2 Developing a method to monitor single RNA translation in two reading frames simultaneously

We created a multi-frame (MF) tag to monitor, in living cells, the translation of single RNAs with overlapping open reading frames (ORFs). The tag builds off earlier technology to visualize translation using repeat FLAG or SunTag epitopes labeled by fluorescent Fab or scFv, respectively [25, 31]. In the MF tag, FLAG epitopes in the 0 frame are separated from one another by SunTag epitopes in the -1 frame. With this arrangement, as shown in Figure A.1A, single RNAs with ribosomes translating the 0 frame will produce FLAG epitopes labeled by Fab (green), while those

with ribosomes translating the -1 frame will produce SunTag epitopes labeled by scFv (blue). Thus, depending on the chosen frame(s), polysomes will appear all green (all ribosomes translating the 0 frame), all blue (all ribosomes translating the -1 frame), or some combination of the two.

To ensure both frames of the MF tag encode functional proteins and have similar space for ribosomes, we inserted the first exon of the human GNAS locus downstream of the FLAG and SunTag epitopes. The GNAS locus contains two overlapping ORFs of roughly equivalent lengths that encode peripheral membrane proteins in the 0 and -1 frames: XXLb1 and AlexX [105, 106] (Figure 4.1A). In combination with the epitopes, this arrangement has several advantages: 1) FLAG epitopes are interspersed between SunTag epitopes. Thus, when FLAG epitopes are translated, the out-of-frame SunTag epitopes act as linkers between consecutive FLAG epitopes (and vice versa). This is an optimal arrangement in that additional linkers between epitopes would further space them out and lower their density within the tag; 2) signals are digital, so frameshifted and non-frameshifted species are marked by two distinct probes/colors; and 3) epitopes are placed in nearly equivalent positions, so signals appear at roughly the same time and with similar amplification when translated with similar kinetics.

As a first application of the MF tag, we focused on -1 programmed ribosomal frameshifting caused by the HIV-1 frameshift sequence (FSS). We inserted the FSS upstream of our MF tag and transiently transfected the resulting construct into U-2 OS cells. The FSS contains a slippery poly-U stretch nine nucleotides upstream of a stem loop. Two to ten hours post transfection, we observed cells with tens or hundreds of individual RNA diffusing throughout the nucleus and cytoplasm (Figure 4.1B). Nascent Chain Tracking (NCT) [37] of the RNA revealed a high degree of RNA-to-RNA heterogeneity, with a subset of RNA labeled by Fab only – indicative of FLAG epitopes from translation in the 0 frame – and a smaller subset labeled by both Fab and scFv – indicative of both FLAG and SunTag epitopes from canonical and frameshifted translation in the 0 and -1 frames, respectively (Figure 4.1B,C).

To confirm these RNA were active translation sites, or polysomes, we performed two experiments. First, we re-imaged cells 12-24 hours after transfection. At these later time points, Fab

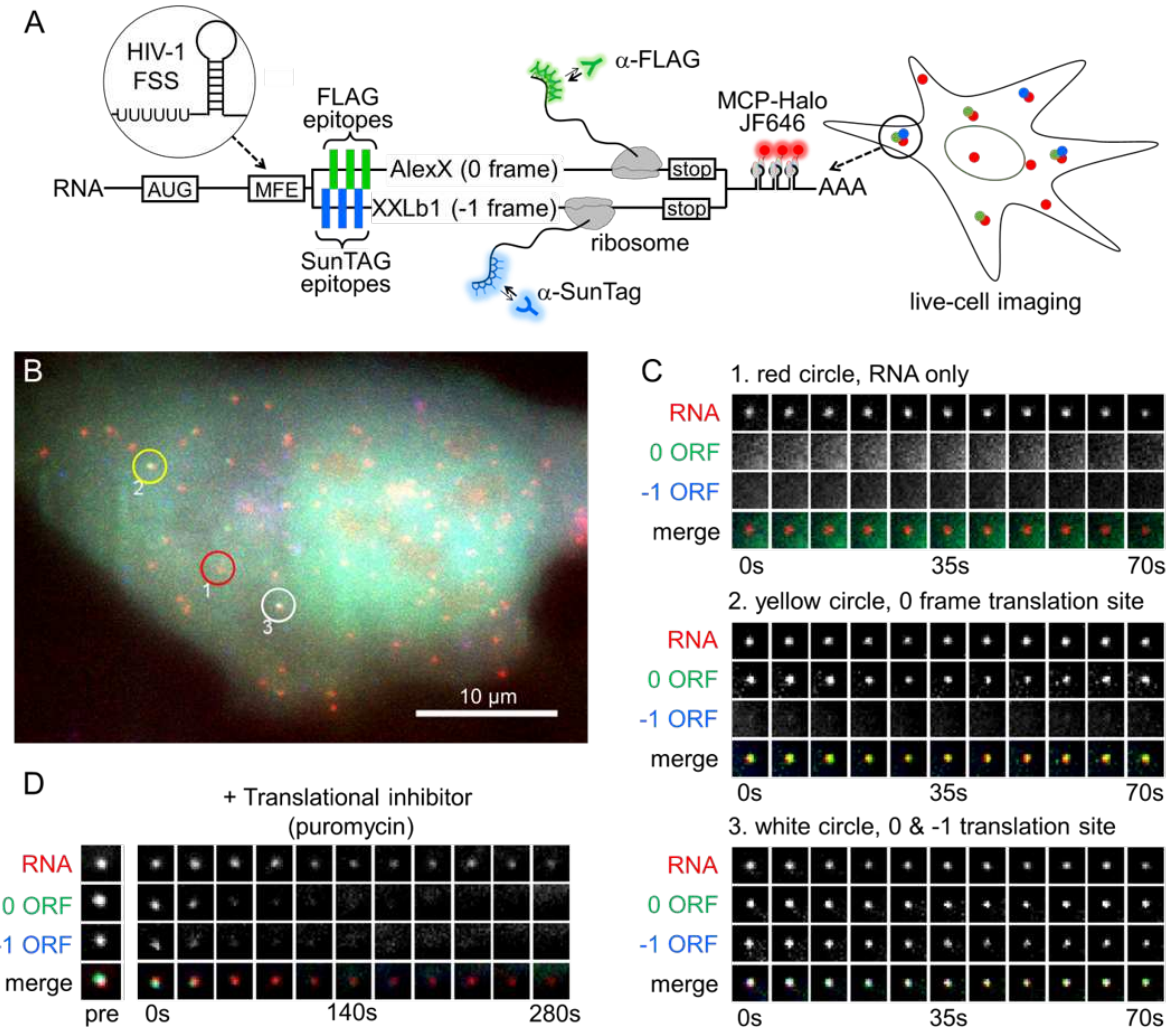


Figure 4.1: A multi-frame tag to image single RNA frameshifting dynamics in living cells. (A) The multi-frame (MF) tag contains 12 repeated FLAG epitopes in the 0 frame interspaced between 12 repeated SunTag v4 epitopes in the -1 frame. Depending on which frame is translated, nascent epitopes are labeled by fluorescent α-FLAG antibody fragments (Cy3-Fab, green 'Y') or α-SunTag single chain variable fragments (scFv-GFP, blue 'Y'). Following the repeat epitopes is exon 1 of the GNAS locus, in which the peripheral membrane proteins AlexX (689 aa) and XXLb1 (690 aa) were placed in the 0 and -1 frames, respectively. Preceding the multi-frame tag is a multi-frame element (MFE), where the HIV-1 frameshift stimulatory sequence (FSS) is placed. In this study, the HIV-1 frameshift sequence was used as the MFE. To facilitate single-RNA tracking, a 24x MS2 stem-loop tag was also placed in the 3' UTR. This tag is labeled by MCP-HaloTag (with JF646-HaloTag ligand, red). (B) A representative cell ~10 hours after transient transfection with the multi-frame tag depicted in A. The red circle (labeled '1') highlights a non-translating RNA, the yellow circle (labeled '2') highlights a 0-frame translation site (TS), and the white circle (labeled '3') highlights a 0 & -1 TS. Scale bar, 10 μm. (C) Montages showing the temporal evolution of the RNA spots circled in B. (D) A representative montage showing the loss of signal from the 0 and -1 open reading frames upon addition of the translational inhibitor puromycin (100 μg/mL). Taken from [36]. K. Lyon created cartoon schematics, prepared reagents, cloned constructs, performed experiments, and analyses for Figure 4.1.

and scFv began to accumulate in the cell membrane (Figure 4.2, left panels), as would be expected if they labeled mature and functional XXLb1 and AlexX proteins [106]. In cells transfected with the -FSS control tag, little or no -1 frame product accumulation was observed (Figure 4.2, middle panels), despite this frame encoding a functional protein, as demonstrated by shifting the sequence by one nucleotide into the 0 frame (Figure 4.2, right panels). Second, we treated cells with the translational inhibitor puromycin. Just minutes after treatment, we observed a dramatic decrease in the number of Fab- and/or scFv-labeled RNA, consistent with the premature release of nascent chains (Figure 4.1D). Together, these data provide strong evidence that we are able to detect single RNA frameshifting dynamics with the MF tag.

4.3 Using the multi-frame tag to quantify HIV-1 frameshifting efficiency

The HIV-1 FSS structure has been previously shown to produce frameshifted protein with an efficiency of 5-10% based on the dual luciferase assay and similar bulk assays [85, 89, 97, 98]. However, it remains unclear how this percentage is established. One possibility is that all RNA behave more or less the same and their ribosomes frameshift with 5-10% probability. On the opposite end of the spectrum, it is possible that RNA display a high degree of heterogeneity, so that just 5-10% of RNA have ribosomes that frameshift with nearly 100% probability. A third possibility is that frameshifting is common on all RNA, but frameshifted proteins are less stable and degraded faster than non-frameshifted proteins.

As a first step to quantify frameshifting dynamics at the single-molecule level, we tracked thousands of individual RNA approximately two to ten hours post transfection with and without the FSS sequence present ($\pm$FSS; (Figure 4.3A-C). We observed all possible combinations of translation sites (Figure 4.5). In addition, we observed many non-translating RNA as well as mature protein punctae. These single-color spots served as convenient internal controls that demonstrated no fluorescence bleed-through at our imaging conditions (Figure 4.5A,B). In the ($\pm$FSS cells, we found 92 ($\pm$ 1.3% of translation sites were translating the canonical 0 frame alone, while 6.2

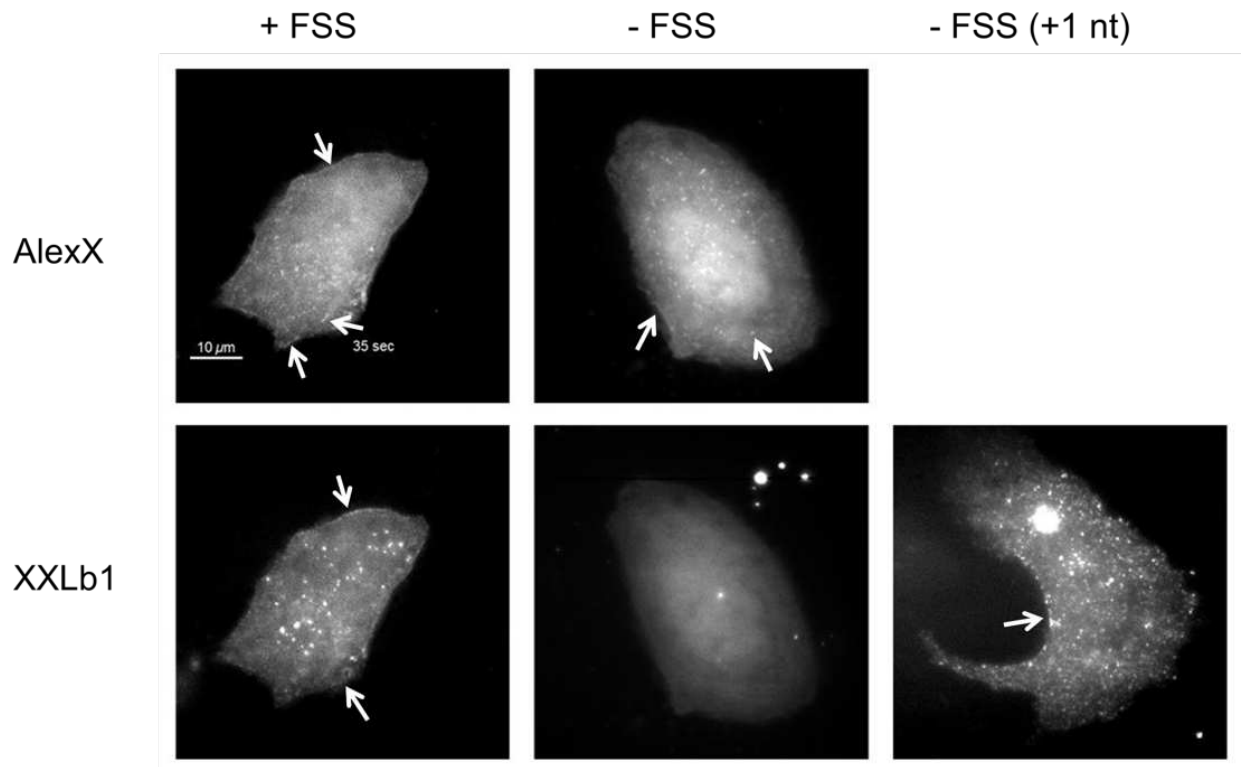


Figure 4.2: Accumulation of mature AlexX and XXLb1 proteins from multi-frame tags. Mature AlexX (upper) and XXLb1 (lower) membrane proteins accumulate in distinct punctae that coalesce to form elongated structures that dynamically move within the cellular membrane. These structures are enriched at the top and bottom of cells and around the edge of the cytoplasm (arrows), consistent with membrane localization. In the +FSS multi-frame reporter, mature AlexX comes from 0 frame canonical translation, while mature XXLb1 comes from -1 frameshifting translation. In the -FSS control reporter, only mature AlexX accumulates, indicating little to no frameshifting. The -FSS (+1 nt) column shows accumulation of XXLb1 when an extra nucleotide is inserted in the -FSS control reporter following the start codon. This extra nucleotide pushes XXLb1 into the 0 frame, demonstrating it can be expressed. Adapted from [36]. K. Lyon prepared reagents, cloned constructs, and performed experiments for Figure 4.2.

($\pm$ 1.1% were translating both the 0 and -1 frames. Only rarely did we observe translation sites translating just the -1 frame (1.6 ($\pm$ 0.5%, Figure 4.5A and C). To ensure these results were not influenced by the multi-frame tag, we reversed the FLAG and SunTag epitopes in the tag. Repeating the experiments with the reversed tag yielded the same fraction of translating and frameshifting RNA (p=0.18 and 0.46, respectively, Figure 4.6), confirming the tag order and/or epitope positioning did not bias measurements. The consistency also suggests we are able to detect all or nearly all translation sites with either tag. We then repeated experiments in cells transfected with the control -FSS construct. In this case, we observed virtually no frameshifting sites (0.9 ($\pm$ 0.7%) (Figure 4.3B, C, and Figure 4.5D). Taken together, these data suggest the FSS alone causes ~8% of translation sites to frameshift. Finally, we wanted to know if the HIV-1 stem-loop could shift ribosomes into the +1 (or -2) frame. Although there is a naturally occurring stop codon in the +1 frame of the HIV-1 stem-loop, we designed a new construct without this stop codon in the +1 frame. This construct revealed a small degree of +1 frameshifting (3.8 $\pm$ 1.3% total +1 frameshifting sites) which we interpret as a slip into the -2 frame (Figure 4.7).

To further characterize the efficiency of frameshifting, we quantified the precise number of frameshifted versus non-frameshifted nascent chains (or ribosomes) per translation site. To do this, we imaged a calibration reporter harboring a single FLAG or single SunTag epitope, but otherwise identical in length and sequence to the -FSS control tag (Figure 4.8). With the calibration construct, each nascent chain in a translation site is labeled by a single fluorophore. Since the fluorescence of a single fluorophore can be unambiguously measured via the observation of single-step photobleaching, the ratio of translation site fluorescence to single-fluorophore fluorescence provides an estimate for the number of nascent chains (in units of mature protein). After calibration, we found frameshifting sites with translation in both the 0 and -1 frames had 11.4 ($\pm$ 1.6 nascent chains total, 2.8 ($\pm$ 0.5 of which were frameshifted, while 0-frame only translation sites had 6.1 ($\pm$ 0.2 non-frameshifted nascent chains, and -1-frame only translation sites had 5 ($\pm$ 2 frameshifted nascent chains (Figure 4.3D). Thus, although just ~8% of translation sites contained frameshifted nascent chains, within these subset of sites, anywhere between 25% (in 0- and -1-frame translation

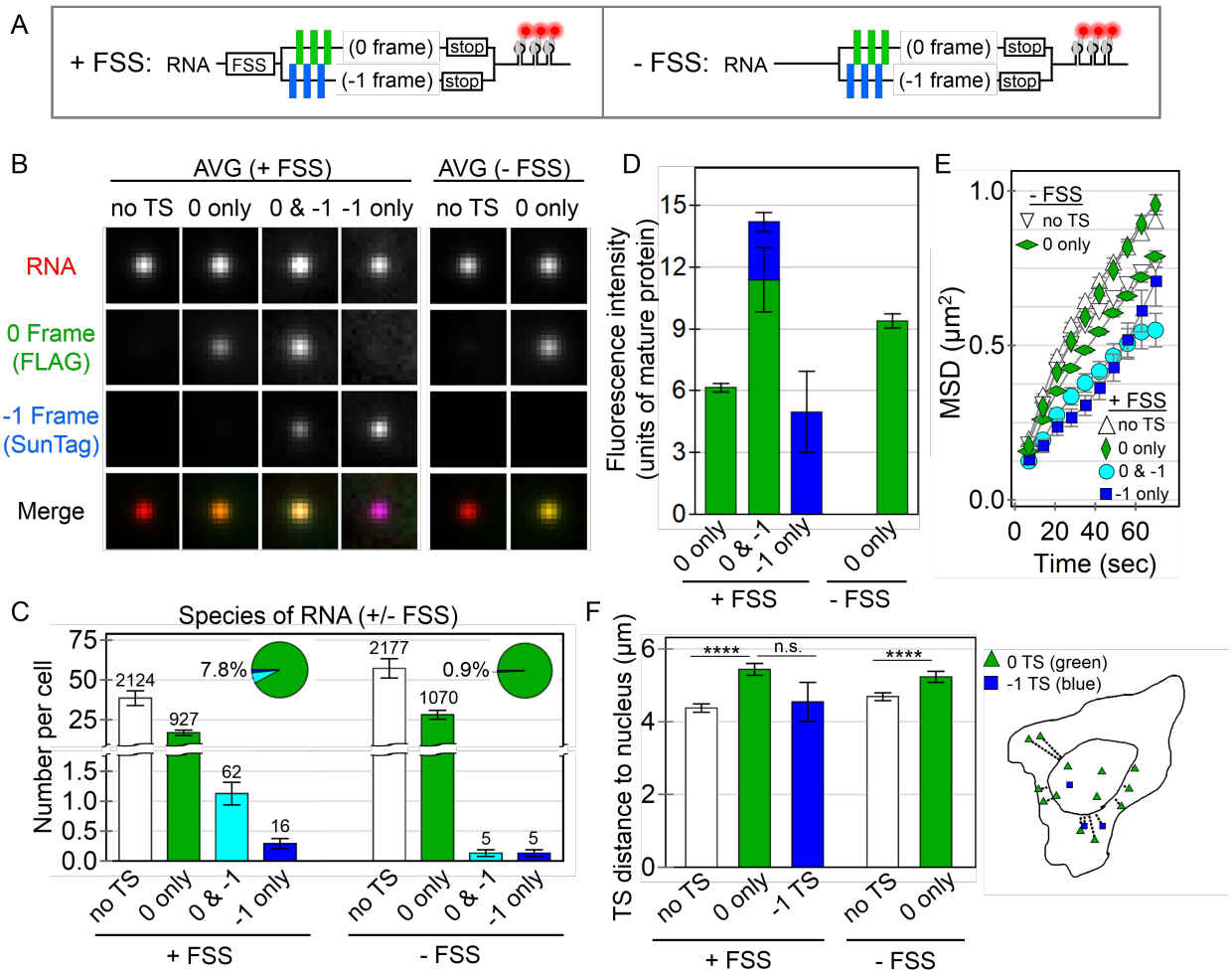


Figure 4.3: Quantification of HIV-1 stimulated frameshifting. (A) A schematic of the MF (+FSS) and control (-FSS) tags. (B) Average image trims of all non-translating RNA sites (no TS), 0-frame only translation sites (0 only), 0- & -1-frame translation sites (0 & -1), and -1-frame translation sites (-1 only), with their respective merges. (C) The number of RNA detected per cell transfected with either the +FSS MF tag (60 cells, 3129 total RNA) or the -FSS control tag (49 cells, 3257 total RNA). The pie charts highlight the percentage of all frameshifting species per transfected cell. The number of detected RNA for each species is shown above each bar. Error bars represent S.E.M. among cells. (D) The fluorescence intensity (in units of mature protein) of nascent chains per translation site for the +FSS MF and -FSS control tags. Error bars represent S.E.M. among RNA. (E) The mean squared displacement (MSD) of tracked RNA species as a function of time. Error bar represent S.E.M. among RNA. The diffusion coefficient (D_{RNA}) was estimated from a linear fit to the first four timepoints (95% CI). For the +FSS MF tag: no TS (D_{RNA} = (1.8 $\pm$ 0.09) x 10^{-2} $\mu m^2 s^{-1}$); 0 only (D_{RNA} = (1.6 $\pm$ 0.06) x 10^{-2} $\mu m^2 s^{-1}$); 0 & -1 (D_{RNA} = (1.0 $\pm$ 0.04) x 10^{-2} $\mu m^2 s^{-1}$); and -1 only (D_{RNA} = (0.7 $\pm$ 0.06) x 10^{-2} $\mu m^2 s^{-1}$). For the -FSS control tag: no TS (D_{RNA} = (1.6 $\pm$ 0.1) x 10^{-2} $\mu m^2 s^{-1}$; and 0 only (D_{RNA} = (1.3 $\pm$ 0.07) x 10^{-2} $\mu m^2 s^{-1}$). (F) The average distance (μm) of detected translation sites from the nucleus (Error bars represent S.E.M. among RNA). An outline of a representative cell on the right shows all detected translating RNA within the cell and their measured distance from the nuclear border (inner curve). P-values are based on the Mann Whitney U test; * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001. Taken from [36]. K. Lyon prepared reagents, cloned constructs, performed experiments, and analyses for Figure 4.3.

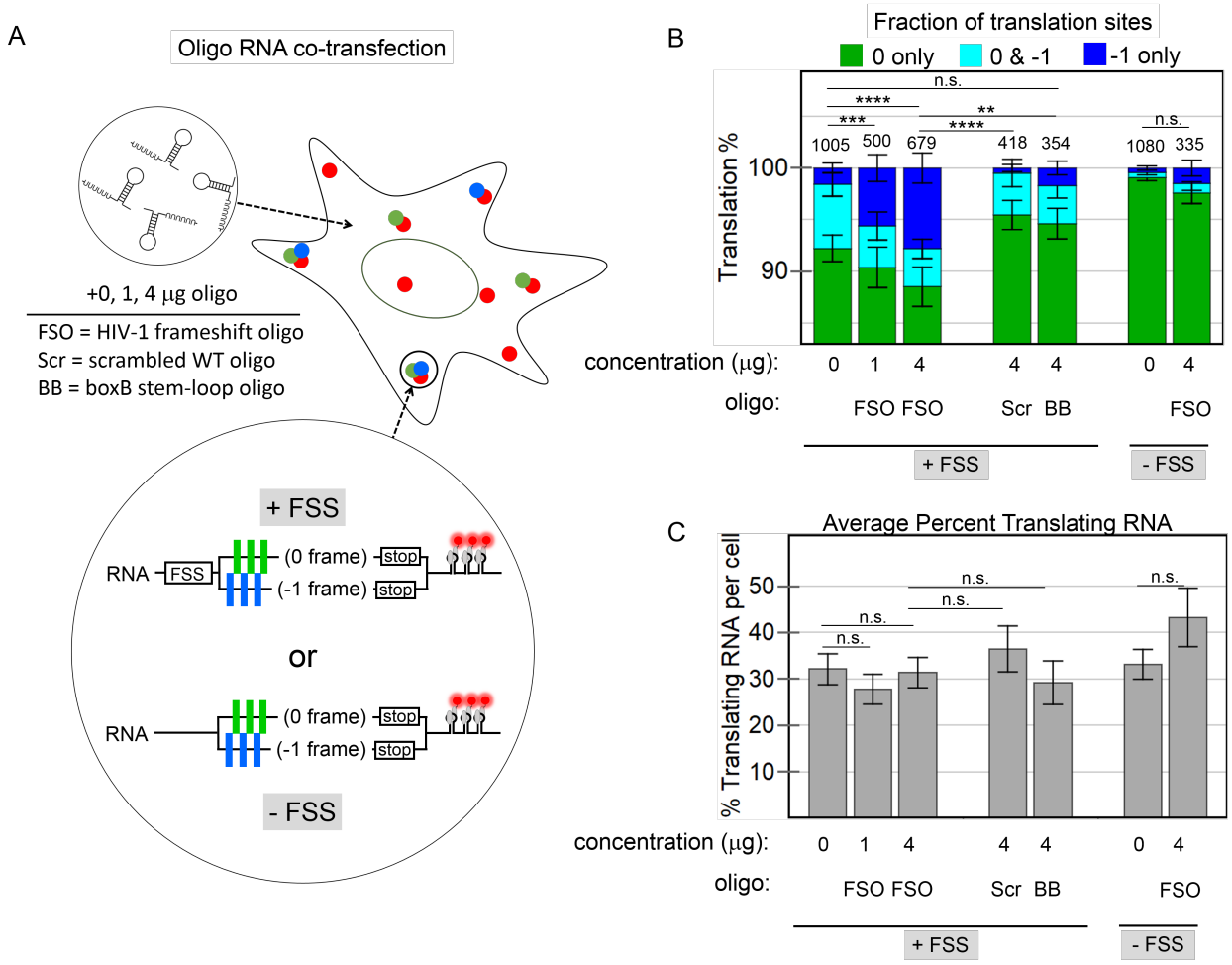


Figure 4.4: Frameshifting can be stimulated by an oligo encoding the FSS. (A) Cells were co-transfected with different concentrations of short oligo RNAs encoding just the frameshift sequence (FSO), a scrambled FSO sequence (Scr), or the boxB stem-loop sequence (BB), together with either the +FSS MF tag or -FSS control tag. (B) The percentage of translation sites translating just the 0 frame only (green, 0 only), the 0 and -1 frames (cyan, 0 & -1), or just the -1 only frame (dark blue, -1 only). Below each plot the concentration of loaded oligo (0, 1, or 4 μg), oligo type (FSO, Scr, or BB), and construct (+FSS or -FSS) are indicated. (C) Average percentage of translating RNA per cell for the experiments in B. The error represents S.E.M. between cells. The p-values are based on the Mann Whitney U test of -1 frame percentages between experiments; * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001. Taken from [36]. K. Lyon created cartoon schematics, prepared reagents, designed experiments, performed experiments, and analyses for Figure 4.4.

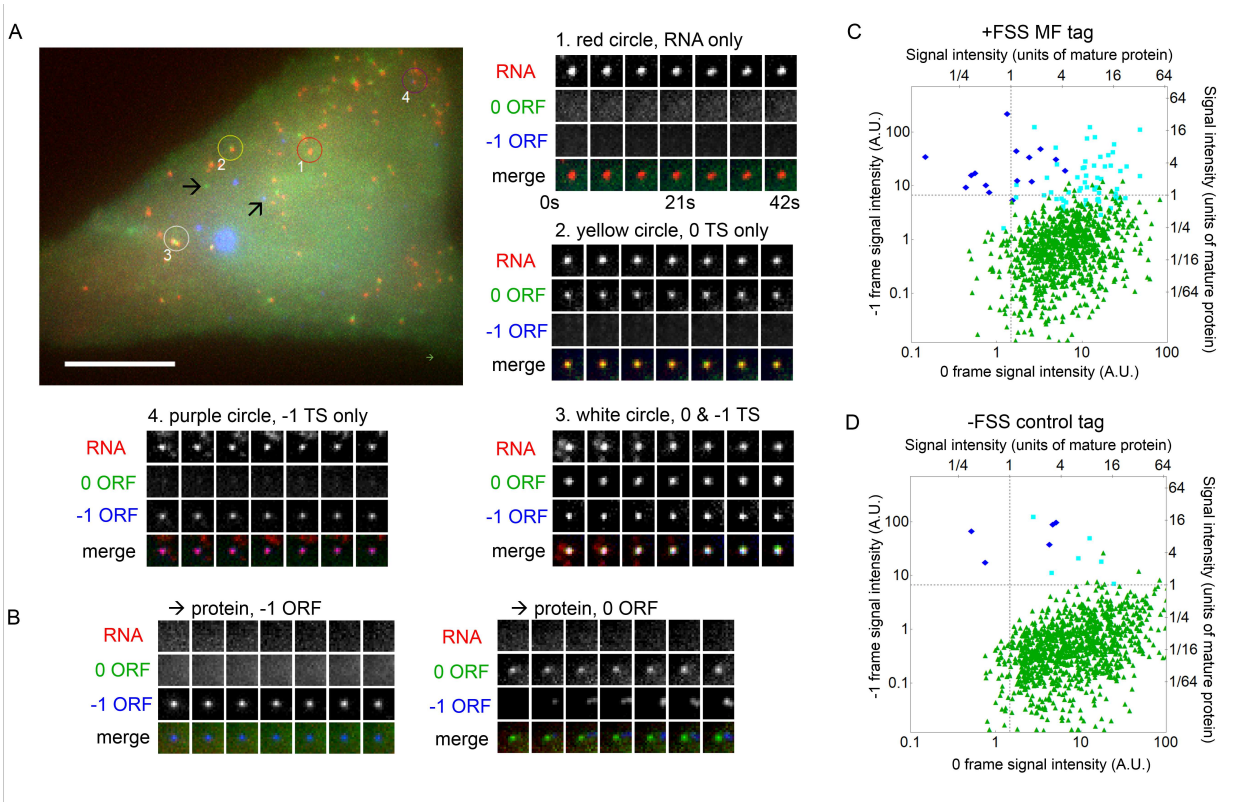


Figure 4.5: A cell with all possible types of translation sites. (A) A sample cell exhibiting all four possible translation sites (TS): (1) untranslated TS (RNA only); (2) 0-frame only TS (0 TS only); (3) 0- and -1-frame TS (0 & -1); and (4) -1 frame only TS (-1 TS only). Scale bar = 10 μm. These images also demonstrate no bleed-through of signals. (B) Additional images of mature protein spots, indicated by black arrows, demonstrating no bleed-through of 0 ORF and -1 ORF signals into the other channels. (C) Scatter plot of 0- and -1-frame signal intensities (in arbitrary units and units of mature proteins) for all TS detected in cells transfected with the +FSS MF tag. All signals were semi-manually hand-checked to remove false-positives. Colors/shapes correspond to the hand-checked identification of the spots: green triangles correspond to 0 TS only signals, cyan squares to 0 & -1 TS signals, and dark blue diamonds to -1 TS only signals. (D) Same as C but in cells transfected with the -FSS control tag. Adapted from [36]. K. Lyon prepared reagents, performed experiments, and analyses for Figure 4.5.

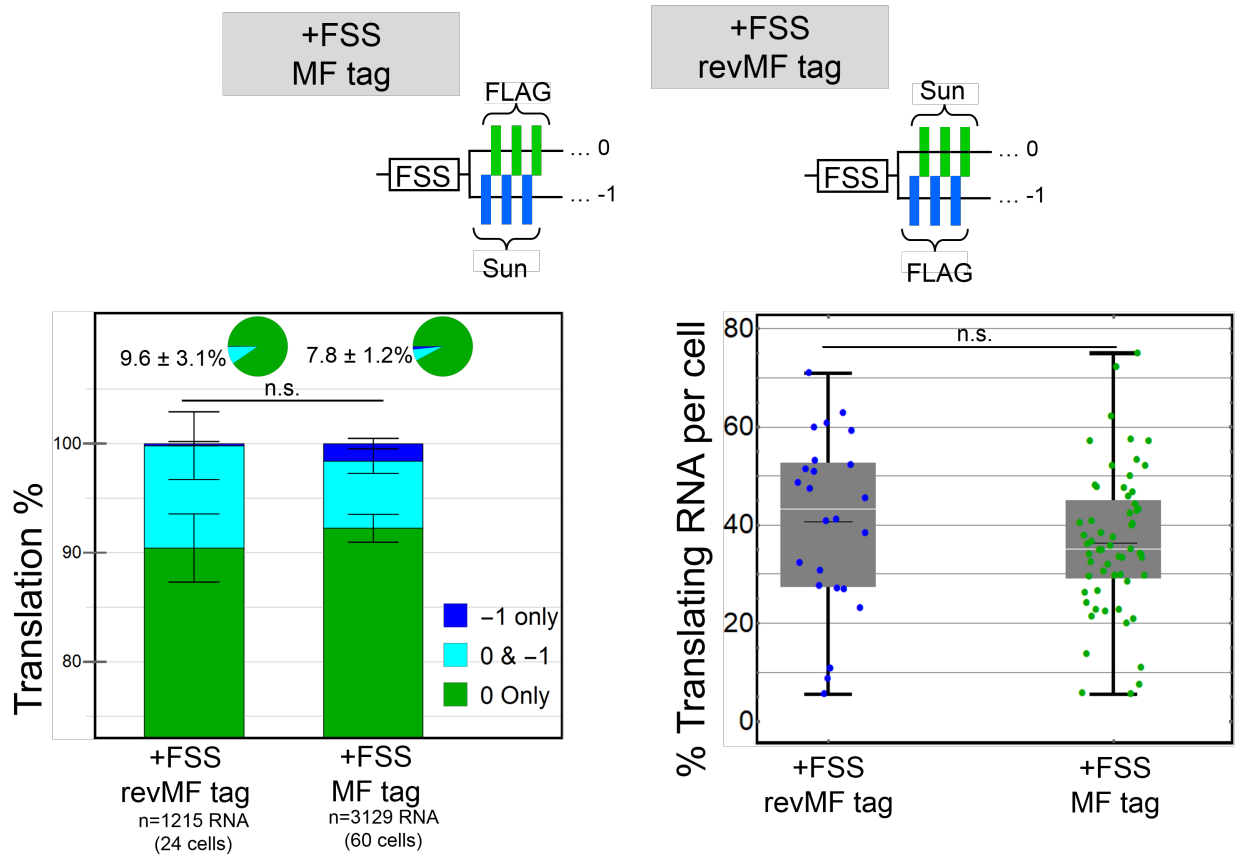


Figure 4.6: Epitope order in tags has no significant impact on percentage of translating RNA. Left, the percentage of RNA translating the -1 frame only (dark blue, -1 only), the 0 and -1 frames (cyan, 0 & -1), and the 0 frame only (green, 0 only) for the +FSS MF tag and the +FSS revMF tag (in which FLAG and SunTag epitopes are reversed). Pie charts show the percentage of translating species per cell. Error bars represent S.E.M. among cells. Right, box whisker plot of the percent translating species per cell for the +FSS revMF tag (dark blue points) and +FSS MF tag (light green points). Each point represents one cell. According to the Mann Whitney U test, there is no statistical difference between the MF and revMF tag frameshifting fractions (left, p=0.46, n.s.) and % translating species per cell (right, p=0.18, n.s.). Taken from [36]. K. Lyon prepared reagents, cloned constructs, performed experiments, and analyses for Figure 4.6.

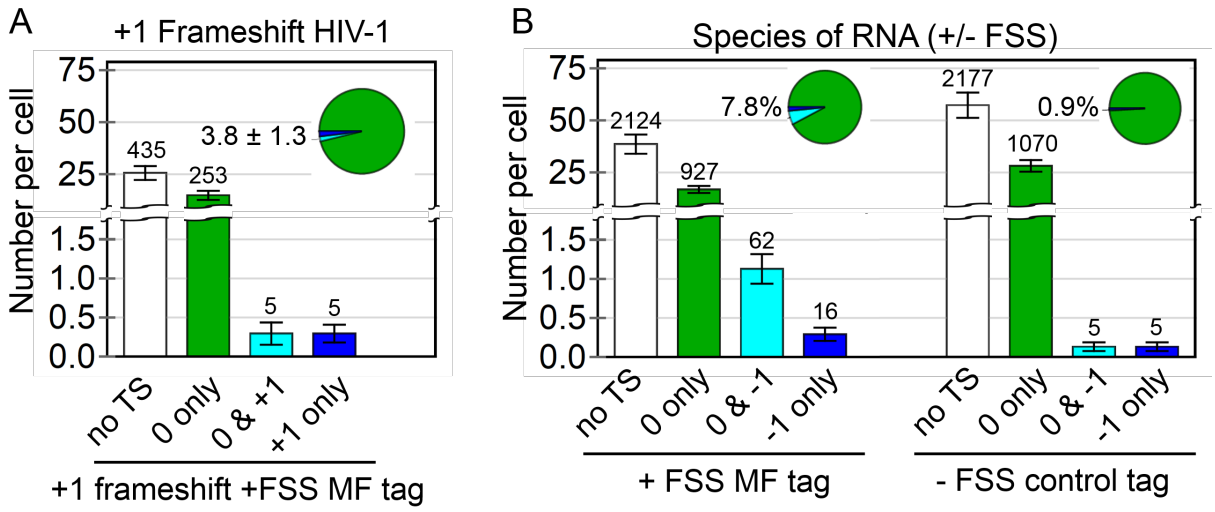


Figure 4.7: Quantification of HIV-1 +1 (or -2) stimulated frameshifting. (A) Quantification of +1 frameshifting using the +1 +FSS MF tag. The number of RNA detected per cell transfected (17 cells, 698 total RNA). The pie charts highlight the percentage of all frameshifting species per transfected cell. The number of detected RNA for each species is shown above each bar. Error bars represent S.E.M. among cells. (B) Quantification of -1 frameshifting using +FSS MF tag and the control tag from Figure 4.3C. K. Lyon prepared reagents, cloned constructs, performed experiments, and analyses for Figure 4.7.

sites) to 100% (in -1-frame only translation sites). These data support a heterogenous RNA model in which frameshifting occurs on a small subset of RNA with high probability.

4.4 Characterizing the mobility, location, and local environment of frameshifting

Given the heterogeneity of observed frameshifting, we hypothesized that the frameshifting state could be stimulated by a specific sub-cellular environment. To test this hypothesis, we performed a statistical analysis of all tracks to see if any biophysical parameters correlated with frameshifting. This revealed frameshifting sites diffuse more slowly than other translation sites or non-translating RNAs ($p<0.05$, Figure 4.3E). This could be due to a preference for frameshifting in the peri-nuclear endoplasmic reticulum, where RNA have been shown to be less mobile and more efficiently translated [40, 107]. In contrast to this, we did not find a preference for frameshifting RNA to be closer to the nucleus compared to non-frameshifting RNA ($p=0.21$, Figure 4.3F).

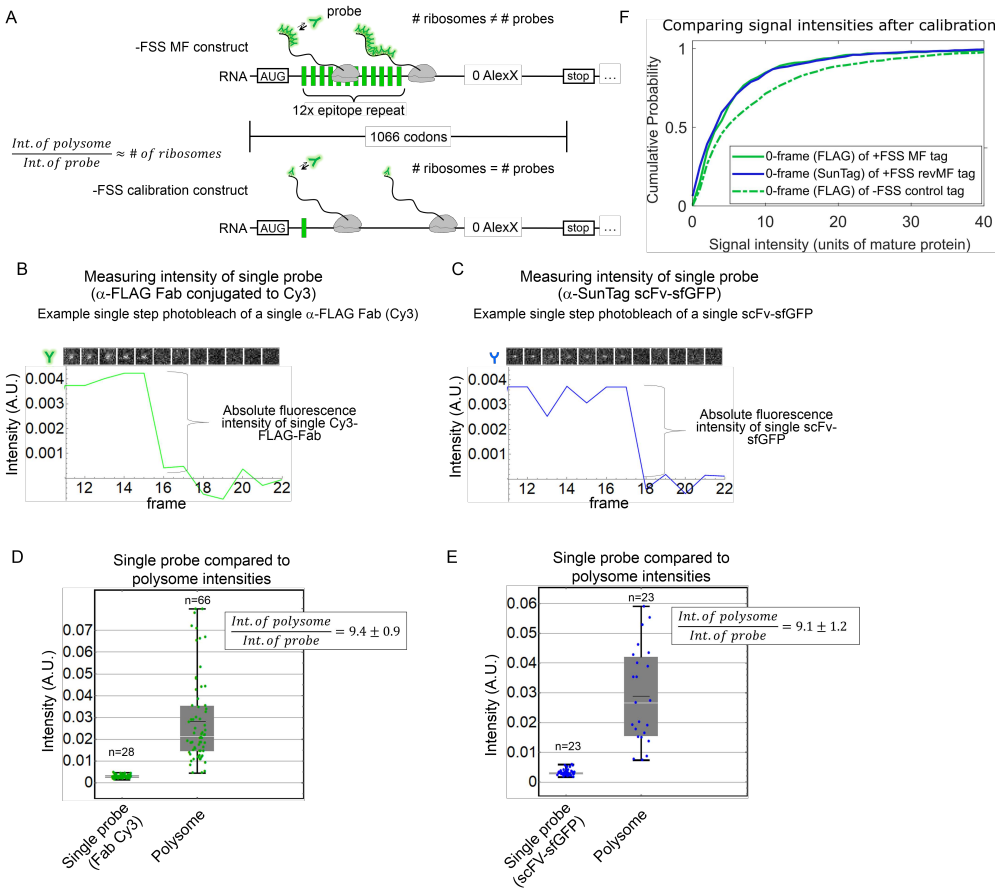


Figure 4.8: Calibrating translation site intensity for Figure 4.3D. (A) Schematic of calibration constructs designed to be of equivalent length to the -FSS control tag. Each calibration construct has either a single FLAG or SunTag epitope in the 0 frame translated with AlexX. Since a single epitope is labeled by a single probe, the number of nascent chains (or ribosomes) within a polysome is approximately the fluorescence intensity of nascent chain signals in polysomes divided by the fluorescence intensity of a single probe. This calibration converts fluorescence intensity to units of mature protein. (B) Example trim and intensity trace showing single-step photobleaching of a single α-FLAG Fab probe conjugated with Cy3. The intensity drop gives the intensity of a single probe for calibrating nascent chain signals (C) Same as B but for a single α-SunTag scFv-sfGFP probe. (D) Box and whisker plot showing single α-FLAG Fab Cy3 probe intensities compared to polysomes (labeled by the Fab). Each green point represents a single detected spot. (E) Box and whisker plot showing single α-SunTag scFv-sfGFP probe intensities compared to polysomes (labeled by the scFv). Each blue point represents a single detected spot. For the box and whisker plots, median is shown by a white line, mean is shown by a black line, the box indicates 25-75% range, and whiskers indicate 5-95% range. There is no statistical difference between the ratio of polysome to mean single fluorophore intensity in D and E (p=0.825, Mann-Whitney U test). (F) A plot comparing signal intensities after calibration. The cumulative probability distributions for the 0-frame signal intensity from the +FSS MF tag (FLAG, light green line), the 0-frame signal intensity from the +FSS revMF tag (SunTag, dark blue line), and the 0-frame signal intensity from the +FSS control tag (FLAG, light green dashed line). The cumulative distributions from the +FSS MF tags are not statistically different (p=0.42, 2-sample KS test), but both are statistically different from that of the FSS- control tag (p<1e-6, 2-sample KS test). Taken from [36]. K. Lyon prepared reagents, cloned constructs, performed experiments, and analyses for Figure 4.8.

Of the parameters we quantified, one of the strongest correlates of frameshifting was RNA signal intensity. Specifically, sites translating the -1 frame had an average RNA signal intensity that was nearly 30% brighter than RNA only spots (Figure 4.3B and Figure 4.9A). This was not due to a few outliers, as >80% of all such translation sites had above average RNA signal intensities (Figure 4.9B). Since these sites have nearly twice as many nascent chains as other sites (Figure 4.3D), it is likely they contain multiple translating RNA. To corroborate this, we chose the 50% dimmest RNA (which are presumably single RNA). In this gated RNA sub-population, the intensity distribution of the 0-frame nascent chains got significantly dimmer for 0- and -1-frameshifting sites ($p=0.037$; Figure 4.9B). In contrast, the intensity distributions of the 0 frame only and -1 frame only spots did not change significantly upon gating (Figure 4.9B, $p=0.89$ and 0.71, respectively). These data imply frameshifting sites are ~80% (Figure 4.9B) of the time composed of two (or more) translating RNAs, at least one of which is frameshifting.

We wondered if this might be an artifact due to the aggregation of probes. Self-aggregation of Fab was unlikely as this would cause 0-frame translation sites to also be brighter, which was not the case (Figure 4.3A, $p<0.0001$). Similarly, it was unlikely to be due to self-aggregation of scFv, as this would cause 0-frame translation sites to be brighter in the reverse MF tag, which was also not the case (Figure 4.9D, $p<0.0001$). This left the possibility that Fab aggregate with scFv. To rule this out, we re-imaged the MF tag without Fab. As frameshifting translation sites remained brighter than other sites (Figure 4.9E, $p<0.0001$), we conclude the brighter RNA signal is not a tagging artifact, but instead represents a propensity for frameshifting sites to associate with other RNA to form higher-order multimer sites. Indeed, although difficult to capture, we have observed such multimers dissociate upon puromycin treatment (Figure 4.10).

4.5 Frameshifting can be stimulated by an oligo encoding the FSS

Unlike experiments with the +FSS MF tag, we did not observe a significant number of translation sites with increased RNA signal intensity in experiments with the -FSS control tag (compare

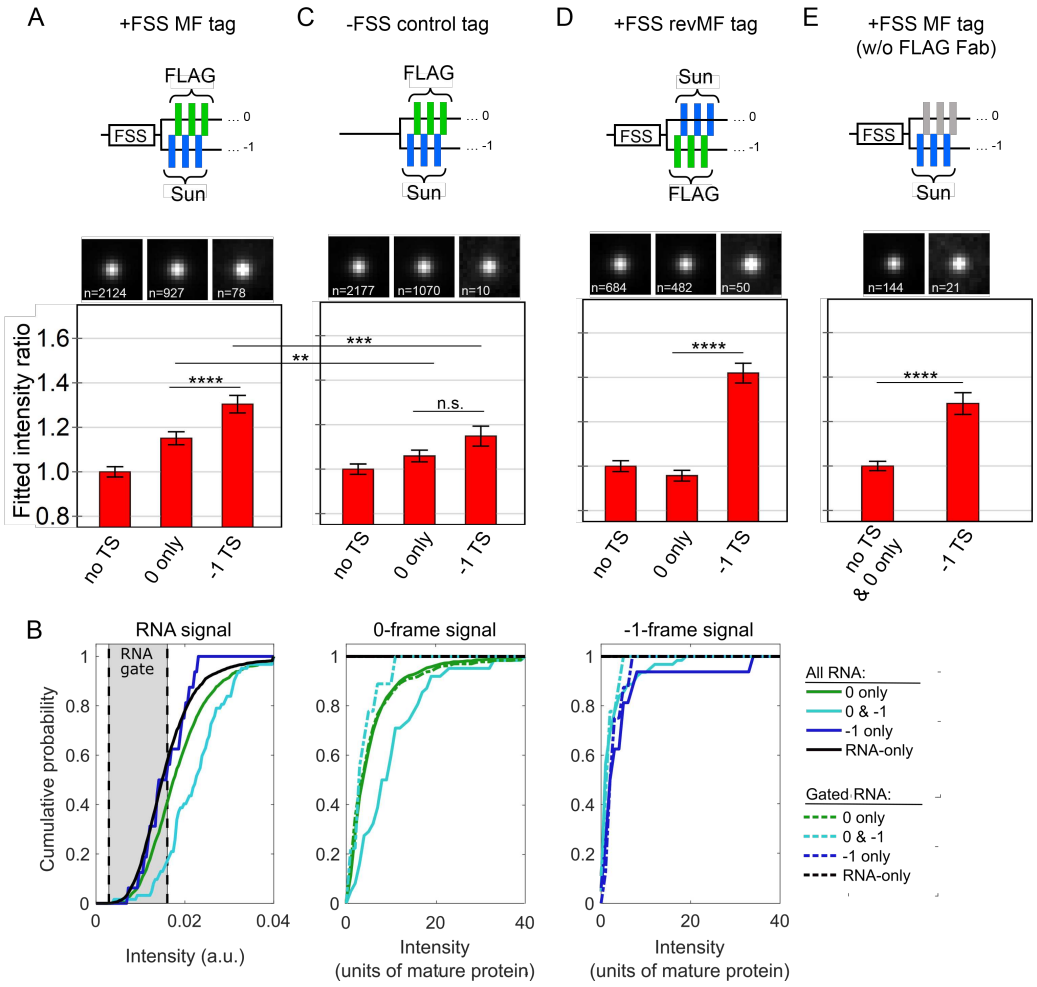


Figure 4.9: Frameshifting sites have brighter RNA signal intensities. (A) Average +FSS MF tag RNA signals for non-translating RNA (no TS), 0-frame only translation sites (0 only), and -1-frame translation sites (-1 TS). The number of RNA used to generate each average image is shown. The bar graphs below show the Gaussian fit intensites (normalized to non-translating RNA, i.e. no TS sites, error bars are fitted 95% CI). The p-values are based on confidence intervals from the fits; * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001. (B) Left, cumulative distribution of +FSS MF tag RNA signal intensity (arbitrary units, a.u.) for 0-frame only translation sites (green line; 0 only), 0- and -1-frame translation sites (cyan line; 0 & -1), -1-frame only translation sites (dark-blue line; -1 only), and RNA-only sites (black line; RNA-only). Significance (2-sample KS test) of the differences between the RNA-only sites and the other translation sites are: 0 & -1 (N=62, p<1e-10), 0 only (N=927, p<1e-16), and -1 only (N=16, p=0.93). An RNA gate (gray region) was applied that selects the 50% dimmest RNA-only sites. Middle, cumulative distribution from all RNA (solid lines) or gated RNA (dashed lines) of the 0-frame signal for the four types of sites. Significance before vs. after gating (2-sample KS test) are: 0 & 1 (Ngated = 9, Ntotal = 62, p=0.037), 0 only (Ngated = 367, Ntotal = 927, p=0.89). Right, same as middle but for the -1-frame signal. Significance before vs. after gating are: -1 only (Ngated = 8, Ntotal = 16, p=0.71), and 0 & -1 (Ngated = 9, Ntotal = 62, p=0.78). (C-E) Same as A but using different constructs: (C) The -FSS control tag, (D) the +FSS revMF tag, and (E) the +FSS MF tag imaged without α-FLAG Fab (error bars are fitted 95% CI. Aside from the -FSS control tag, frameshifting sites (-1 TS) have significantly brighter RNA signals compared to canonical translation sites (0 only). The 0 only and -1 TS RNA signal intensities from C are significantly dimmer than those in A (p<0.01 and p<0.001, respectively). Taken from [36]. K. Lyon prepared reagents, cloned constructs, performed experiments, and analyses for Figure 4.9.

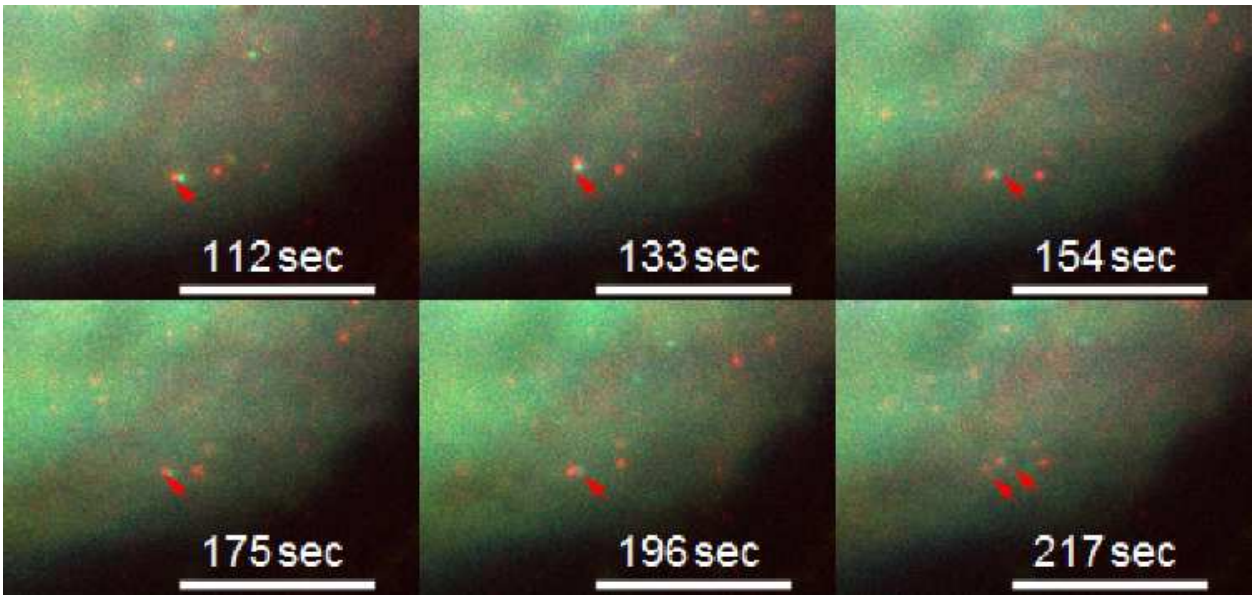


Figure 4.10: Frameshifting site splitting upon puromycin treatment. Max projection of a 13 z-stack movie showing an example U-2 OS cell expressing the + FSS MF tag and bead-loaded with probes: MCP-HaloJF646 (Ch1 red, RNA), FLAG Fab Cy3 (Ch2 green, 0 ORF), and scFv-sfGFP (Ch3 blue, -1 ORF). Images were acquired every 7 seconds for a total of 217 (of 280) seconds total. Puromycin addition was done at 105 seconds. This montage depicts a zoom-in from a live cell to accentuate a frameshifting site splitting (red arrows). Scale bar, 10 μm. Adapted from [36]. K. Lyon prepared reagents, performed experiments, and analyses for Figure 4.10.

Figure 4.9A to Figure 4.9C, 0 only p<0.01; -1 TS p<0.001). We therefore hypothesized the FSS sequence could be involved in multimerization. To test this, we co-transfected cells expressing the +FSS MF tag with a short oligo RNA encoding just the FSS sequence (FSO; frameshift oligo, Figure 4.4A). Remarkably, this led to a significant increase in the fraction of frameshifting sites translating just the -1 frame, from 1.6 to 5.6% when 1 μg FSO was added (Figure 4.4B, p<0.001), and up to 9.5% when 4 μg FSO was added (Figure 4.4B, p<0.0001). However, in contrast to our expectation, the FSO did not significantly impact the RNA signals within frameshifting sites. Instead, irrespective of FSO concentration, the distributions of RNA signals within frameshifting (and non-frameshifting) sites remained statistically unchanged (Figure 4.11). For controls, we repeated experiments, first with non-specific oligos and second with the FSO in cells expressing the -FSS control tag. In both cases, we did not see a significant increase in frameshifting (Figure 4.4B). Furthermore, in all experiments the fraction of translating RNA remained statistically constant (Figure 4.4C), indicating cellular stress was not a factor. We therefore conclude the FSS can somehow interact with other translation sites to facilitate frameshifting. While it remains unclear if the interaction is direct or indirect, by itself the interaction does not appear to alter the multimerization of frameshifting sites.

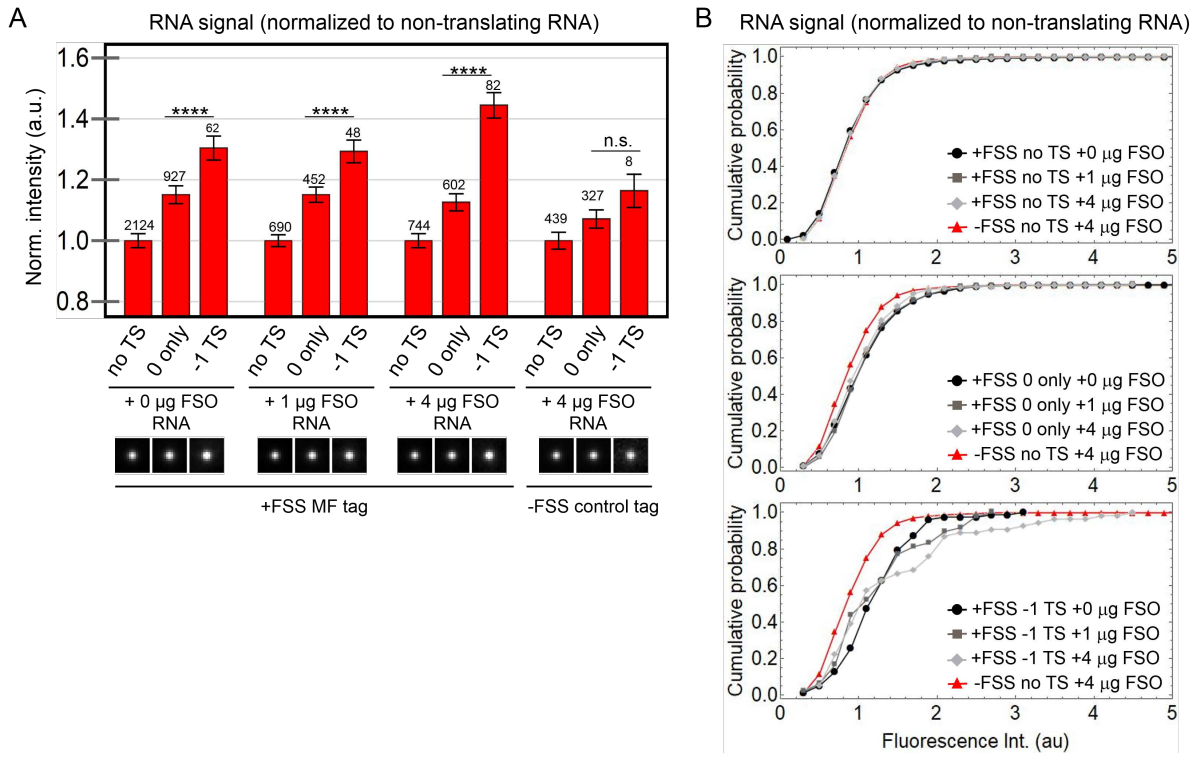


Figure 4.11: RNA signal intensities in FSO co-transfection experiments. (A) Average RNA signals from non-translating RNA (no TS), 0-frame only translation sites (0 only), and all frameshifting translation sites (-1 TS) when different concentrations (0, 1, and 4 μg) of frameshift oligo (FSO) were co-transfected into cells with the +FSS MF tag or the -FSS control tag. The bar graph above shows the Gaussian fit intensity (normalized to non-translating RNA, i.e. no TS sites, error bars are fitted 95% CI). The p-values are based on confidence intervals from the fits; * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001. (B) Cumulative probability distributions from the RNA signal intensities in A (upper plot, no TS RNA; middle plot, 0 only RNA; lower plot, -1 TS RNA). A reference curve (red) corresponds to the no TS RNA signal intensities from the -FSS control tag. In each plot the grayscale curves (black circles, dark gray squares, and light gray diamonds) are not statistically different from one another according to the 2-sample KS test. Taken from [36]. K. Lyon prepared reagents, designed experiments, performed experiments, and analyses for Figure 4.11.

4.6 Translational output of frameshifted ribosomes

The ~8% of frameshifted translation sites we observed is consistent with previous measurements of 5-10% frameshifted protein product [85, 89, 97, 98]. All else equal, this implies that frameshifting alone can explain the steady-state levels of frameshifted protein, without the need for other regulatory mechanisms, such as protein degradation. To test this hypothesis, we performed a ribosomal run-off experiment with a doubled +FSS MF tag construct (+FSS 2xMF tag) to increase the signal

amplification so we could lower the laser powers and thereby eliminate observable photobleaching. This allowed us to roughly estimate the elongation rates of frameshifted and non-frameshifted ribosomes. Fits to the post-tag portion of run-off curves yielded similar run-off times (Figure 4.12A and Figure 4.13). Fluorescence recovery after photobleaching experiments further confirmed these estimates without the use of drugs (Figure 4.14). Based on these similar post-tag elongation rates for both non-frameshifting and frameshifting sites, a single round of translation would take ~9 minutes. Accounting for the number of ribosomes per translation site and their relative fractions, we calculate a cell with 100 RNA would produce ~170 frameshifted protein per hour compared to ~4200 canonical proteins. In other words, frameshifted proteins would account for ~4 $\pm$ 1% of the total, in agreement with earlier measurements. Thus, the presence of the FSS sequence alone can be sufficient to account for the steady-state levels of frameshifted protein in living cells, without the need for additional regulatory mechanisms, such as protein degradation.

4.7 Evidence for ribosomal traffic jams at HIV-1 frameshifting translation sites

Despite the similar elongation rates, we noticed a slight but significant delay in the run-off response at frameshifting translation sites (compare 'Sun sites' to 'FLAG only sites' Figure 4.12A, $p<0.001$). Given earlier work showing the potential of the FSS to pause ribosomes [98], we envisioned this delay could be due to a queue or traffic jam of ribosomes upstream of the FSS within frameshifting sites. As the backed-up ribosomes clear the traffic jam, they replenish the loss of ribosomes running-off. Only after the traffic jam is fully cleared does the number of ribosomes beyond the FSS (with labeled nascent chains) begin to decay.

To test this possibility, we added a 10x HA epitope repeat upstream of the FSS in the +FSS MF tag. We call this new tag the HA MF tag (Figure 4.12B). This served two purposes: first, it allowed us to monitor both ribosomes upstream of the FSS (translating HA epitopes) and ribosomes downstream (translating either 0-frame FLAG or -1-frame SunTag epitopes); second, the arrangement more closely mimicked the natural placement of the FSS between the Gag-Pol polyproteins. In particular,

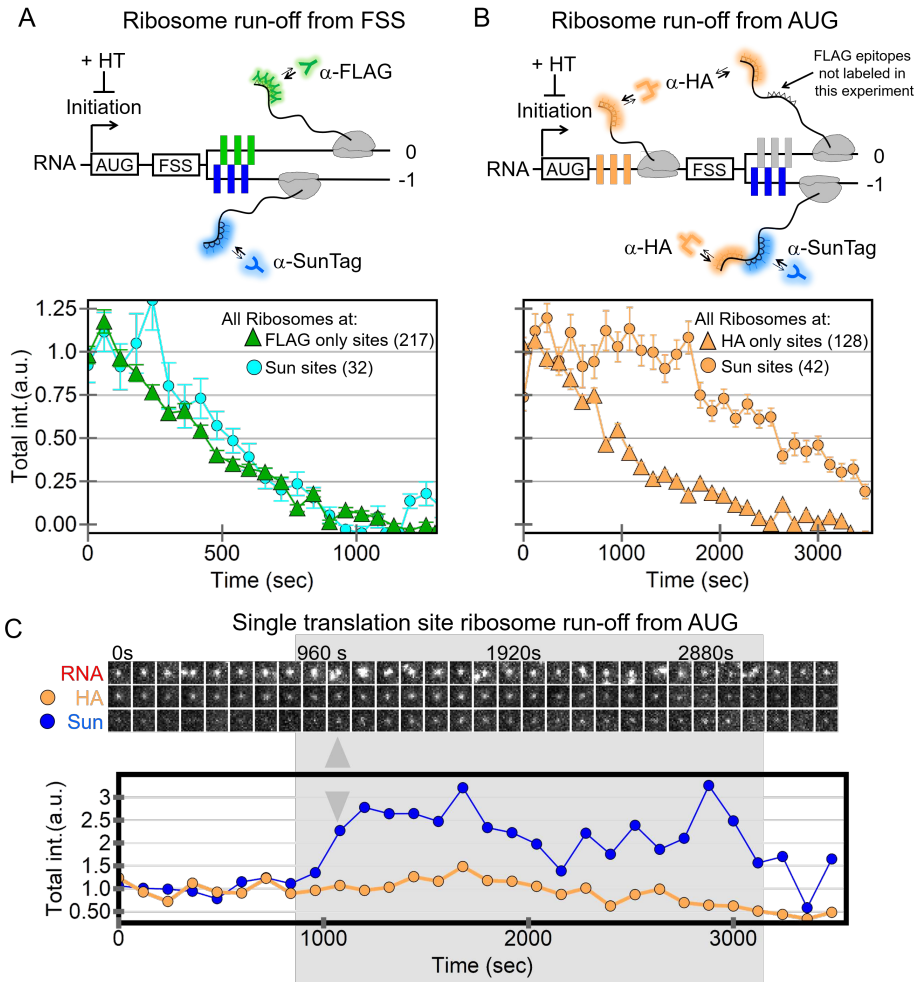


Figure 4.12: Ribosomal run-off at frameshifting and non-frameshifting translation sites. (A) A schematic showing harringtonine-induced ribosomal run-off from the +FSS 2xMF tag with FLAG (green) and SunTag epitopes (blue) in the 0 and -1 frames, respectively. The normalized total intensity (a.u.) of nascent chain signals within non-frameshifting translation sites (green triangles, 217 FLAG only sites initially) and frameshifting sites (cyan circles, 32 Sun sites initially). Frameshifting translation sites are distinguished by the presence of α-SunTag scFv. As these sites contain both FLAG- and SunTag nascent chains, the intensity is the sum of the α-FLAG Fab and α-SunTag scFv fluorescence. There is a small but significant difference between the run-off of non-frameshifting vs. frameshifting ribosomes ($p<0.001$ for all timepoints up to ~600 sec, after which the two curves begin to converge to zero intensity, Mann Whitney U test, 19 cells). (B) Similar to A, but with a modified MF tag with the addition of 10 HA epitopes (orange, HA MF tag) upstream of the frameshift sequence. The α-HA Fab signals in non-frameshifting translation sites (orange triangles, 128 HA only sites initially) and frameshifting translation sites (orange circles, 42 Sun sites initially). Frameshifting translation sites are distinguished by the presence of α-SunTag scFv. The non-frameshifting and frameshifting HA run-offs were significantly different ($p<0.0001$ for all timepoints, Mann Whitney U test, 27 cells). (C) A sample single translation site encoding the modified HA MF tag (shown in B) after addition of harringtonine. A montage of image trims shows the detected RNA-, HA-, and Sun-signals through time. Below, the normalized total intensity of the α-HA Fab signal (marking all ribosomes) and the α-SunTag scFv signal (marking frameshifting ribosomes) is plotted through time. Gray arrows and gray box signify a burst of frameshifting. Taken from [36]. K. Lyon created cartoon schematics, prepared reagents, cloned constructs, designed experiments, performed experiments, and analyses for Figure 4.12.

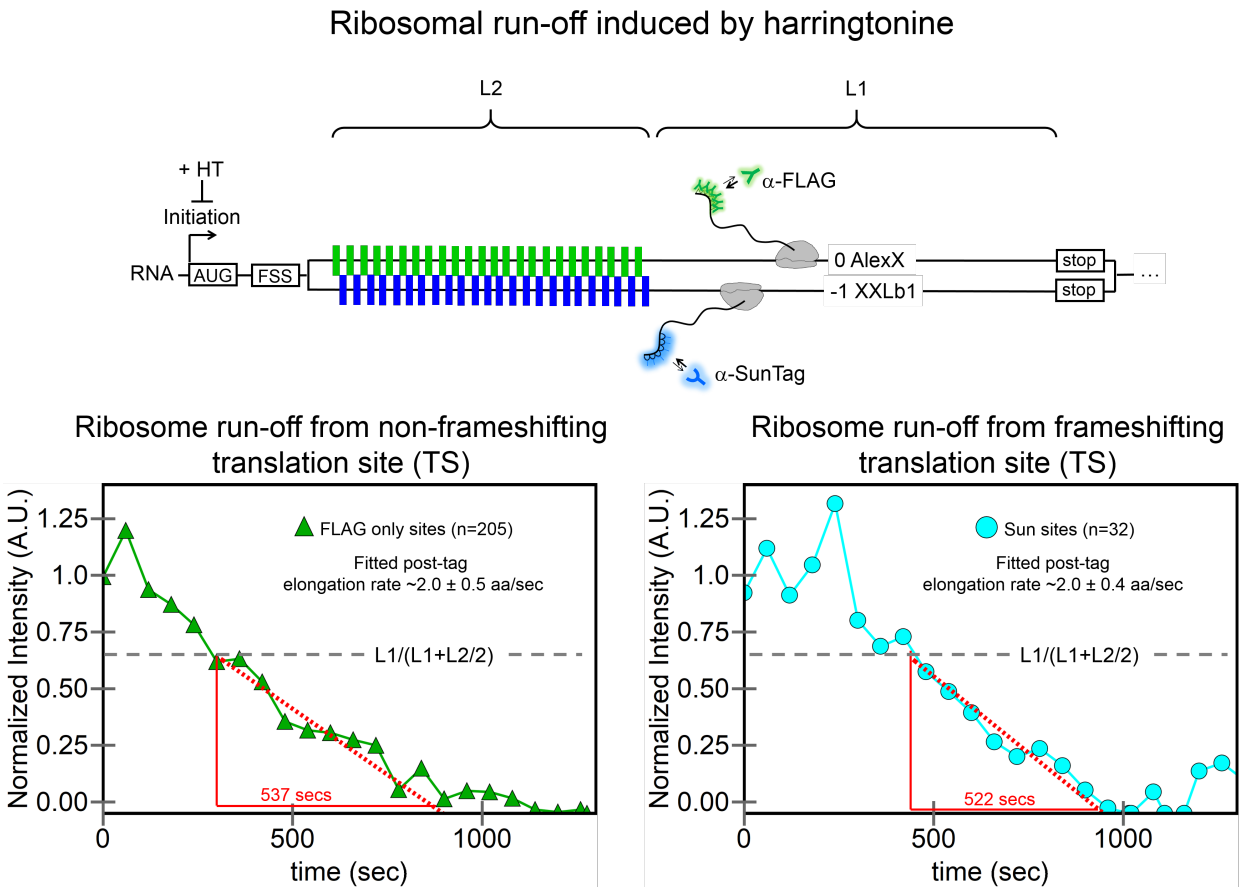


Figure 4.13: Fits to the post-tag linear portion of ribosomal run-offs. The total intensity of detected non-frameshifting and frameshifting translation sites decays with time after addition (at t = 0 sec) of the translational initiation inhibitor harringtonine. Data is taken from Figure 4.12A. If L2 is the length of the tagged portion of the open reading frame and L1 is the length of the non-tagged portion, then the linear (post-tag) portion of the decay begins from approximately L1/(L1+L2/2). This portion of the curve provides an estimate of the elongation rate during the part of the run-off where no new epitopes are being translated. Therefore, the run-off fit is independent of the frameshifting kinetics that occur upstream of the epitopes. Curves are normalized to their initial values. Fitted values are reported with 95% CI. Taken from [36]. K. Lyon prepared reagents, cloned constructs, designed experiments, performed experiments, and analyses for Figure 4.13.

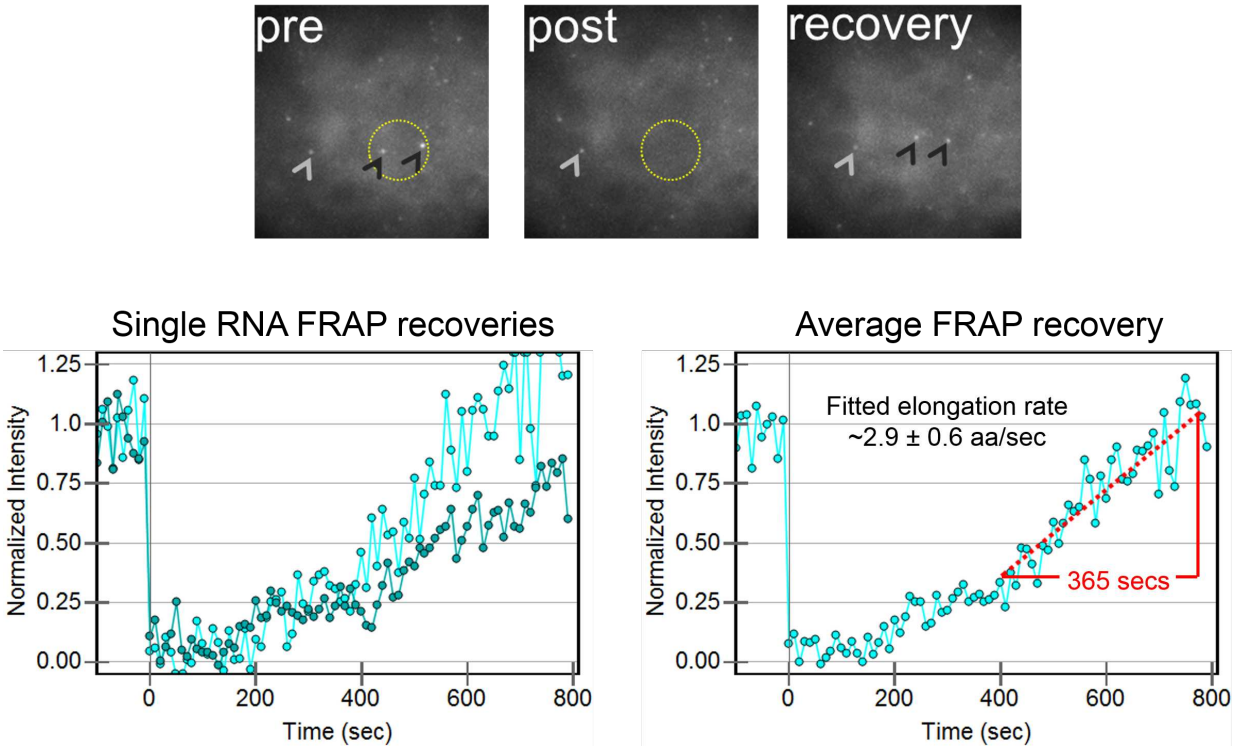


Figure 4.14: Fluroescence recovery after photobleach of frameshifting sites. (A) Fluorescence recovery after photobleaching (FRAP) experiments were performed at frameshifting translation sites (yellow circle marks the photobleach spot). The fluorescence recovery of the 0-frame signal within these sites was quantified as a function of time, with sample pre, post, and recovery frames shown above. (B) The average FRAP recovery time can be fit to estimate the elongation rate. This rate is similar to what was measured with harringtonine in Figure 4.13. Fitted values are reported with 95% CI. Taken from [36]. K. Lyon prepared reagents, designed experiments, performed experiments, and analyses for Figure 4.14.

the HA MF tag includes 368 codons upstream of the FSS compared to just 23 codons with the original +FSS MF tag. This extra space could theoretically accommodate longer ribosomal traffic jams (up to ~40 ribosomes), should they occur. We hypothesized that longer ribosomal traffic jams would lead to longer run-off delays. Consistent with this, ribosomes within frameshifting sites took much longer to run-off compared to ribosomes in non-frameshifting sites (compare "Sun sites" to "HA only sites" in Figure 4.12B, $p<0.0001$), with frameshifted ribosome levels remaining high for upwards of 3000 s, despite an overall ribosome loss (Figure 4.15 and Figure 4.16). We observed this trend even at the single-molecule level, where the fluorescence signal intensity of HA (marking all ribosomes) decreased through time, but the fluorescence signal intensity of SunTag scFv-sfGFP (marking only the frameshifted ribosomes) remained high compared to the initial signal intensity throughout the entire imaging session (Figure 4.12C and Figure 4.17). These fluctuations appear to arise from the stochastic release of stalled ribosomes within the traffic jam and their subsequent translation of frameshifted epitopes. Such a release can be seen in the single-molecule track at the ~1000 s time point, when the frameshift signal gets significantly brighter. Although difficult to capture, we observed this type of dynamic in another single frameshifting RNA track as well (Figure 4.17).

4.8 Computational modeling of HIV-1 frameshifting bursts at the single-RNA level

To quantify the kinetics of frameshifting, we developed two candidate models and attempted to fit each to our four main observations: (i) the percentages of single RNA engaged in 0 frame, -1 frame, and both frame translation, (ii) nascent chain intensities in 0 and/or -1 frames on these RNA, (iii) the average total ratio of -1 frame to 0 frame protein production; and (iv) run-off kinetics for original and extended HA constructs. To disentangle effects of single-RNA translation and aggregation of translating RNA to form translation site multimers, we down-selected to the 50% dimmest RNA (Figure 4.9B) before fitting items (i) and (ii). We reiterate that gating on the dimmest RNA led to no significant difference for the intensities of 0 only or -1 only translation sites, but

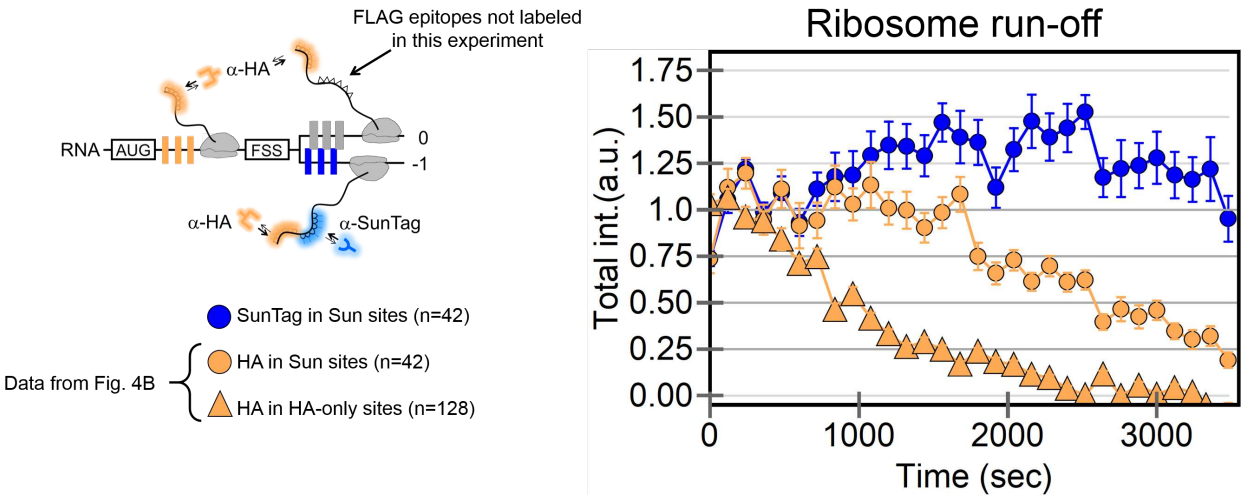


Figure 4.15: Ribosome run-off from the HA MF tag. Ribosome run-off curve showing Sun signal in HA&Sun sites from the experiment performed in Figure 4.12B. The Sun signal comes from frameshifted ribosomes that have run past the frameshift sequence (FSS). Error bars represent S.E.M. Taken from [36]. K. Lyon created cartoon schematics, prepared reagents, cloned constructs, designed experiments, performed experiments, and analyses for Figure 4.15.

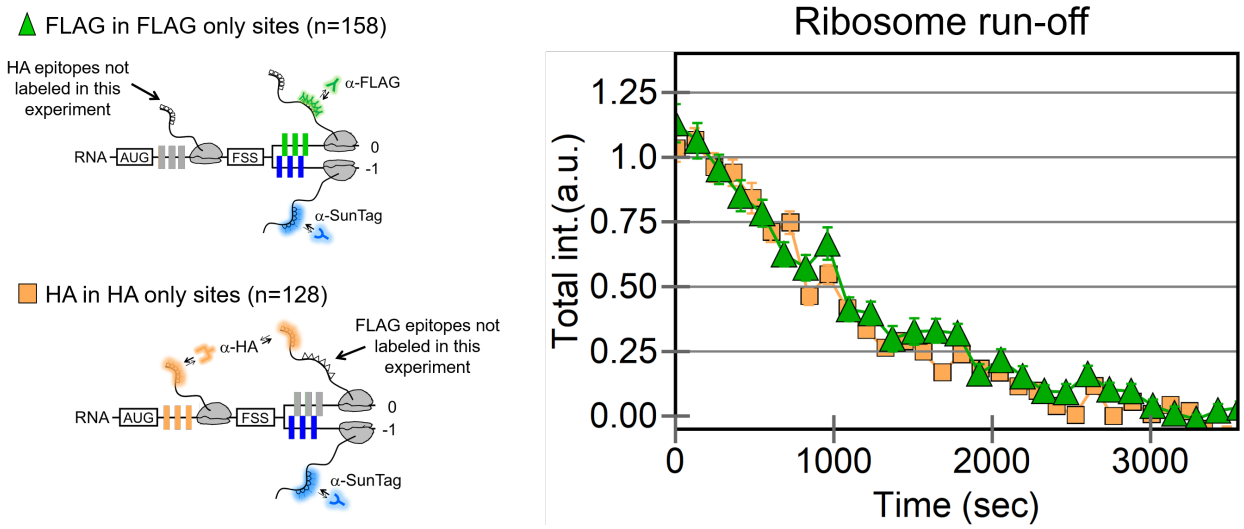


Figure 4.16: Ribosome run-off from FLAG epitopes in HA MF tag. For completeness, an additional experiment was performed to generate the ribosome run-off curve from FLAG epitopes in the HA MF tag. Because we can only image two epitopes at the same time (since RNA is imaged in a third color), we examined FLAG and Sun epitopes (Sun epitopes were required to distinguish frameshifting and non-frameshifting sites). The FLAG run-off from this experiment is on the right (green triangles). This curve is compared to the run-off curve of HA in HA only sites (orange squares) from Figure 4.12B. Both experiments represent non-frameshifting sites (no Sun signal detected). Error bars represent S.E.M. Taken from [36]. K. Lyon created cartoon schematics, prepared reagents, cloned constructs, designed experiments, performed experiments, and analyses for Figure 4.16.

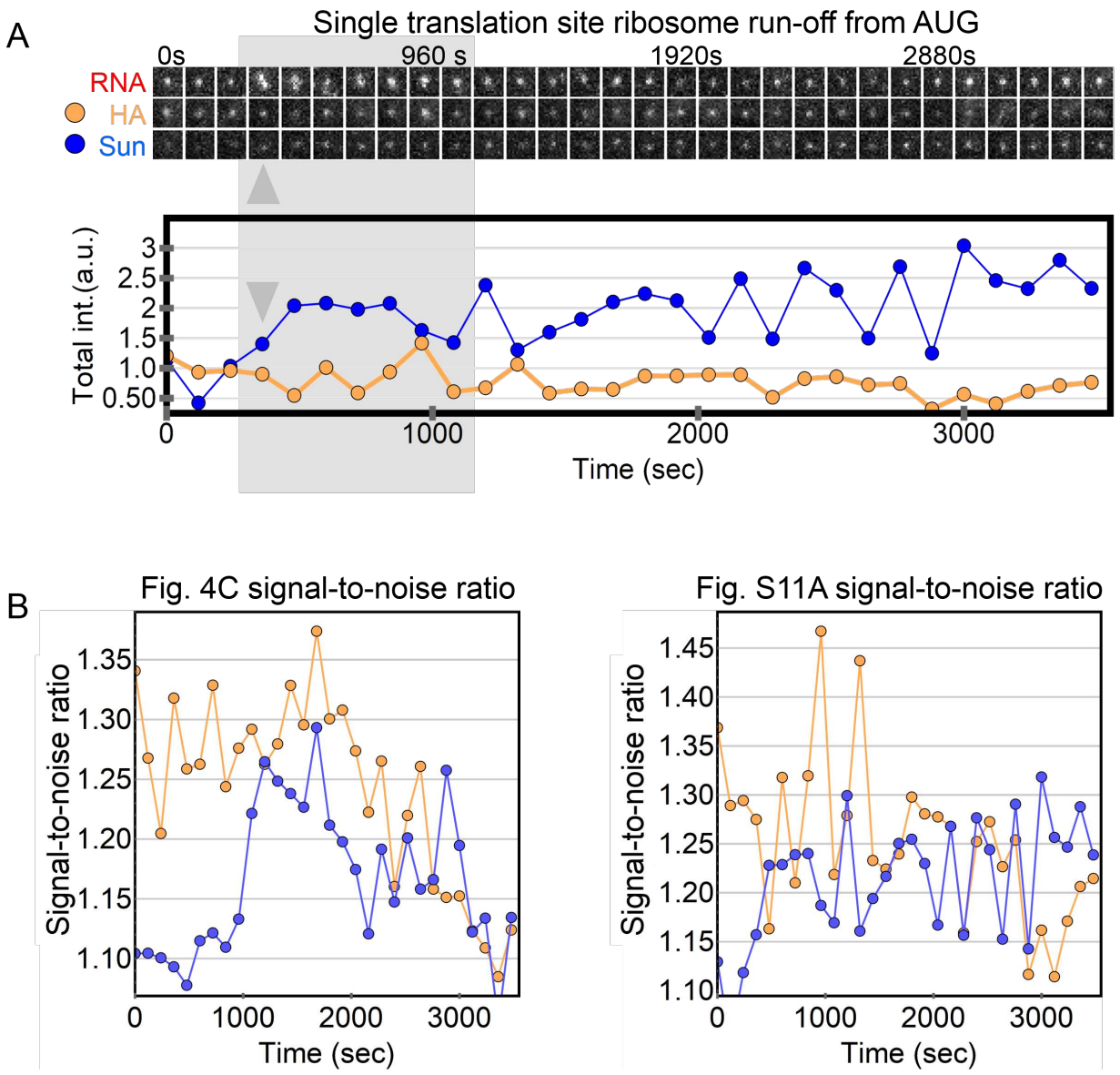


Figure 4.17: Track of a frameshifting burst at a single translation site. (A) A single frameshifting translation site encoding the HA MF tag was tracked after harringtonine addition. On the top, a montage of image trims shows the detected RNA, HA, and Sun signals through time. Below, the normalized total intensity of the HA Fab signal (marking all ribosomes, orange circles) and the SunTag scFv signal (marking frameshifting ribosomes, dark blue circles) is plotted through time. At about the fourth timepoint (gray arrows), another non-translating RNA interacts with the frameshifting RNA, coincident with a burst of frameshifting signal (Sun). (B) Plots showing the signal to noise ratio of the HA and SunTag signals from Figure 4.12C (left) and data in A (right). The signal is the average intensity of an inner circle of diameter four pixels centered on the translation site, while the background is the average intensity of an outer ring surrounding the inner circle of width twelve pixels. Taken from [36]. K. Lyon prepared reagents, designed experiments, performed experiments, and analyses for Figure 4.17.

removed 80% of the 0- & -1-frame translation sites, providing more evidence that these spots are multiple RNA. Items (iii) and (iv) refer to the total 0 and -1 frame translation and were fit without gating.

Both models include initiation of ribosomes, codon-dependent elongation of proteins along the RNA template and ribosomal exclusion to block ribosomes from passing or occupying the same place on the RNA. The only difference in the two models is the treatment for how ribosomes shift from the 0 to the -1 frame. The first model assumes constitutive frameshifting, in which each ribosome can frameshift at the FSS with a fixed and equal probability. This model could capture either observation (i) or (ii), but not both simultaneously; frameshifting either led to excessively large fractions of frameshifting sites or excessively small ribosomal loading, in disagreement with our observations that a relatively small fraction of RNA frameshift with relatively high ribosomal occupancies. Even with addition of distinct pauses in elongation at the FSS in both frames, the constitutive model was unable to fit our data (Figure 4.18 S12).

The second model is inspired by two-state gene models that are commonly used to describe heterogeneous transcription [108]. In this "bursty" model, RNA stochastically switch between non-frameshifting and frameshifting states in which either 0% or 100% of ribosomes produce frameshifted proteins (Figure 4.19A). In the bursty model, the RNA frameshift state is assumed to switch ON and OFF at rates $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$, respectively, and the steady state fraction of RNA in the ON state is given by $\mathrm{f} = \mathrm{k_{on}}/(\mathrm{k_{on}} + \mathrm{k_{off}})$.

To estimate the timescale of switching $(\mathrm{k_{off}})$, we tracked translation sites for longer periods of time. To achieve this tracking, we used the brighter +FSS 2xMF tag and changed our imaging strategy to sample the RNA signal at all time points and the 0 and -1 translation signals once every fifth time point. This arrangement substantially reduced photobleaching and allowed us to continuously track and monitor the translational status of single translation sites in 3D for nearly an hour. Figure 4.19B shows the frameshifting state survival times for the seven translation sites we tracked in this manner, including one site that frameshifted for longer than 40 minutes (Figure 4.20), representing at least four rounds of translation at our estimated elongation rate. Remarkably, this

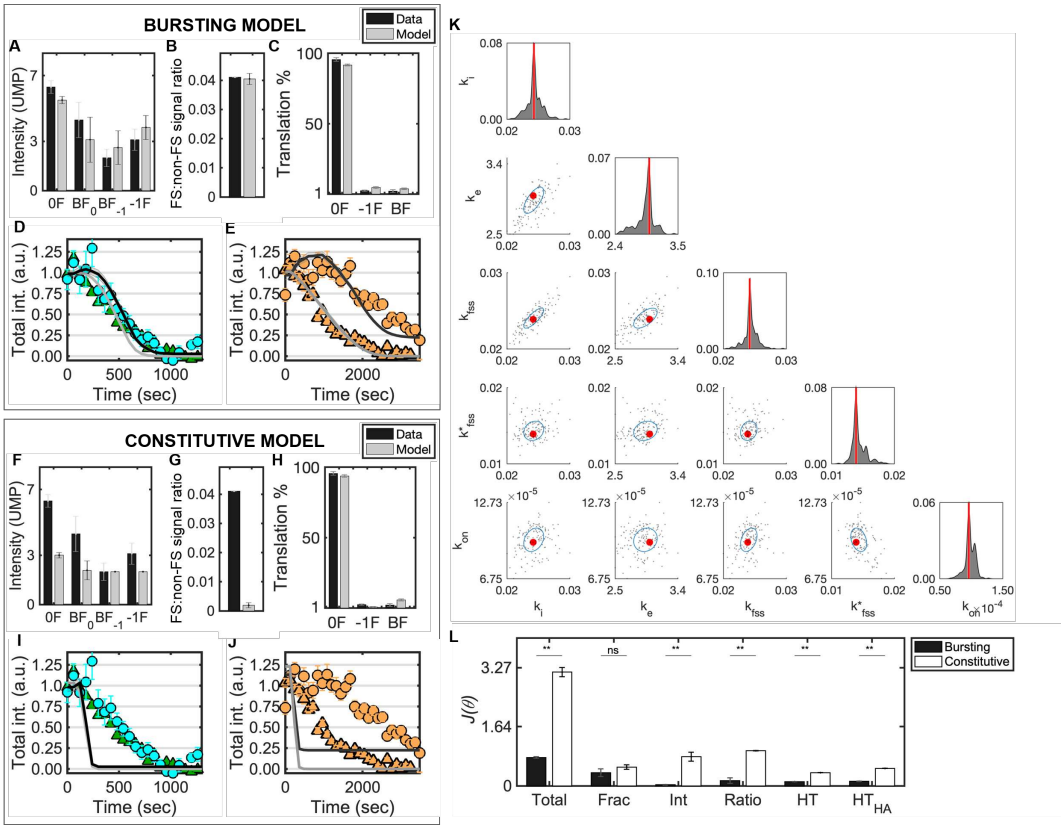


Figure 4.18: Comparison between the bursting and constitutive models. Model dynamics for the bursting (A-E) and constitutive models (F-J). Simulations were performed using the best parameter values obtained from the optimization process. (A,F) A bar graph comparing the measured (black) and best-fit model predicted (gray) intensity (units of mature protein, u.m.p.) within 0-frame (0F), -1-frame (-1F) and 0- and -1-frame (both frames, BF) single RNA translation sites. Error bars represent S.E.M. among RNA. (B,G) A bar graph comparing the measured (black) and best-fit model predicted (gray) ratio between the total frameshifted and non-frameshifted signal intensity (FS to non-FS signal ratio). (C,H) A bar graph comparing the measured (black) and best-fit model predicted (gray) percentage of non-frameshifting (0F), frameshifting (-1F), and both frames (BF) translation sites. Error bars represent S.E.M. among cells. (D,I) Simulated run-off (solid lines) from the frameshift sequence (FSS) of all ribosomes in non-frameshifting (gray) and frameshifting sites (black), plotted with data in Figure 4.12A (triangles and circles). (E,J) Simulated run-off (solid lines) from the start site (AUG) of all ribosomes in non-frameshifting (gray) and frameshifting sites (black), plotted with data in Figure 4.12B (triangles and circles). Error bars represent S.E.M. among RNA. (K) Parameter uncertainty analyses for the bursty model. Optimized parameters: k_{ini}, k_{el}, k_{on}, k_{FSS}, k^*_{FSS}. Parameter k_{off} was estimated from the FS-spot survival times and was fixed during the optimization process. Red lines represent the parameter values reported in the main text. Ellipses represent 90% confidence intervals. (L) Results of fitting the constitutive and bursting models with the Pattern Search Algorithm using the objective function (J). Where lower J(θ) values are better fits to the observational data. Mean values from 8 independent optimization processes, error bars represent S.E.M. Significance was calculated using the Mann Whitney U test. * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001. Taken from [36]. K. Lyon generated data that was fit in Figure 4.18.

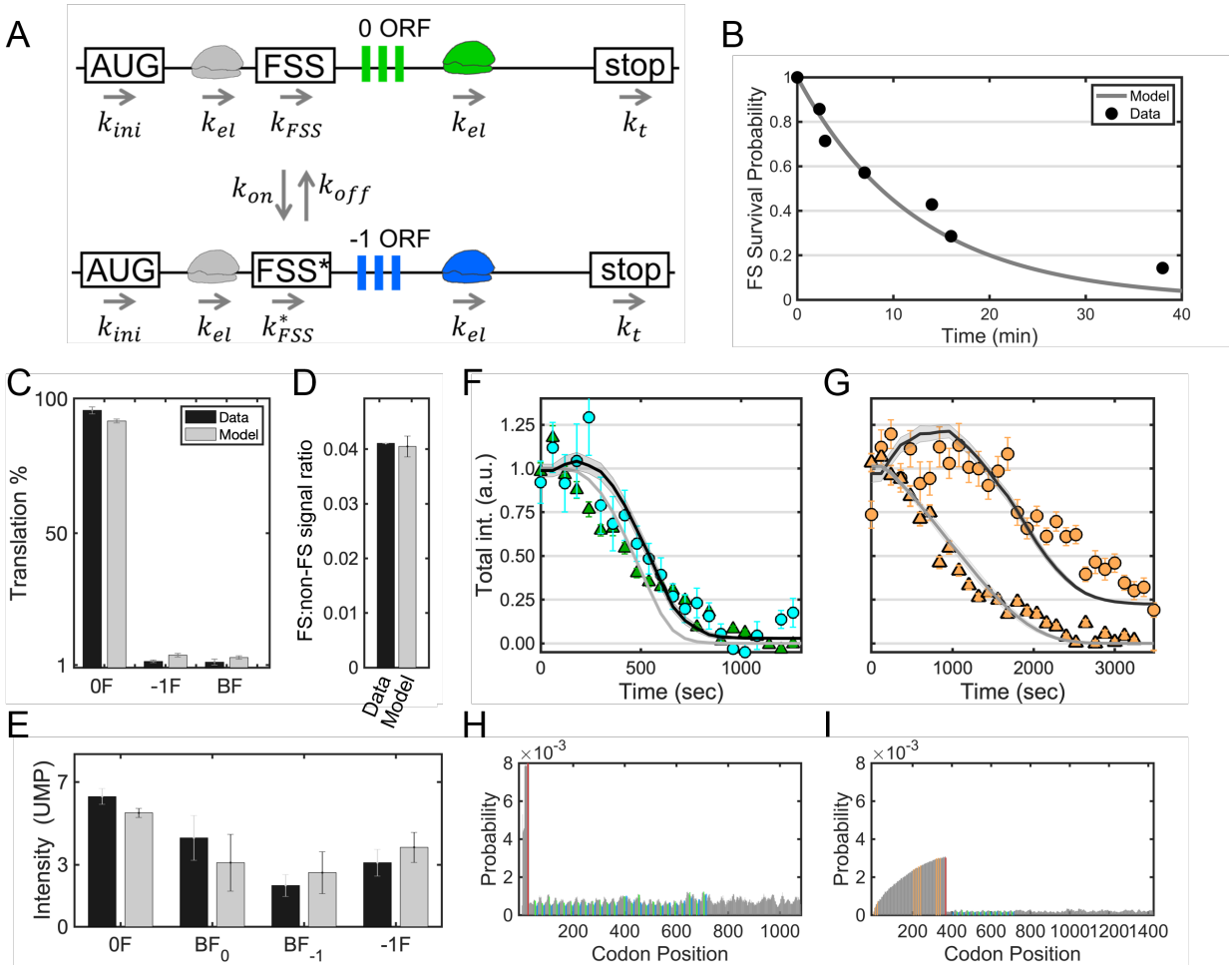


Figure 4.19: A model for bursty frameshifting. (A) A schematic of the model: k_{ini} is the translation initiation rate, k_{el} is the translation elongation rate, k_{on} is the rate at which RNA switch to the frameshifting state, k_{off} is the rate at which RNA switch to the non-frameshifting state, k_{FSS} is the pause rate at the FSS in the non-frameshifting state, k^*_{FSS} is the pause rate at the FSS in the frameshifting state, and k_t is the termination rate (assumed equal to k_{el}). (B) The survival probability of frameshifting sites through time (gray dots) is fit with a single exponential decay (black line). (C-G) Simultaneous fit of all data. (C) A bar graph comparing the measured (black) and best-fit model predicted (gray) percentage of non-frameshifting (0F), frameshifting (-1F), and both frames (BF) translation sites. Error bars represent S.E.M. among cells. (D) A bar graph comparing the measured (black) and best-fit model predicted (gray) ratio between the total frameshifted and non-frameshifted signal intensity (FS to non-FS signal ratio). (E) A bar graph comparing the measured (black) and best-fit model predicted (gray) intensity in units of mature protein (u.m.p.) within non-frameshifting (0F), frameshifting (-1F), and both frames (BF) translation sites. Error bars represent S.E.M. among RNA. (F) Best-fit model (solid lines) of the data from Figure 4.12A. Error bars represent S.E.M. among RNA. (G) Best-fit model prediction of the data from Figure 4.12B. Error bars represent S.E.M. among RNA. (H) The predicted ribosomal occupancy along the MF tag is shown. The positions of the FSS (red), FLAG (green), and SunTag (blue) epitopes are shown in color. (I) The predicted ribosomal occupancy along the HA MF tag is shown. The positions of the FSS (red), HA (orange), FLAG (green), and SunTag (blue) epitopes are shown in color. Taken from [36]. K. Lyon created cartoon schematic and generated data that was fit in Figure 4.19.

frameshifting translation site associated with another for a large part of the 40-minute imaging window. This supports the notion that the brighter RNA signal intensity at frameshifting sites comes from more than one RNA. From the frameshift state survival times, we fit the rate of $(\mathrm{k_{off}})$ to be ~0.0013 (sec^{-1}) (Figure 4.19B), corresponding to an average frameshift persistence time of 1/$(\mathrm{k_{off}})$ ~12.8 min.

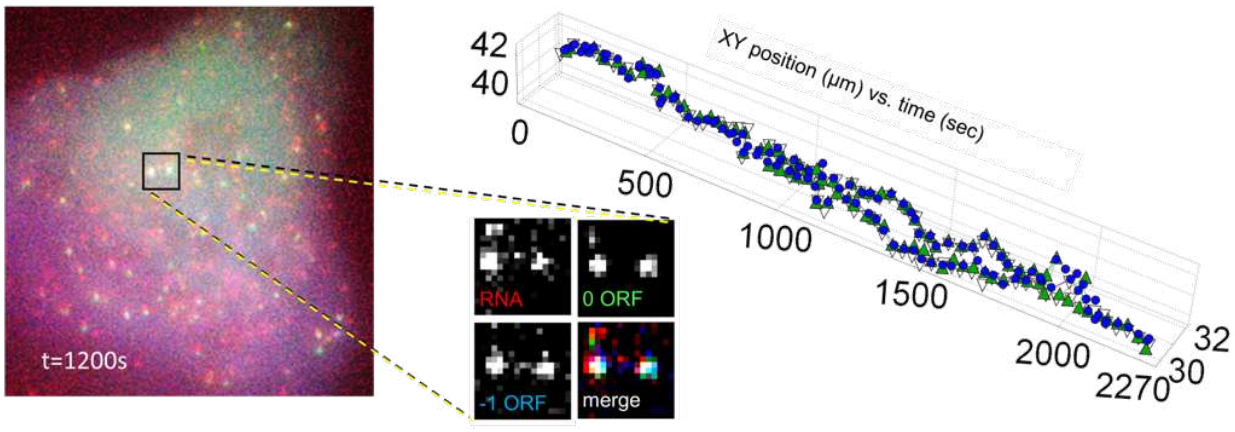


Figure 4.20: Single frameshifting RNA track persisting for >35 minutes. A sample frameshifting translation site encoding the +FSS 2xMF tag was tracked for >35 minutes. Left shows the cell at the 1200 second time point. The zoom shows the channels separated to highlight two frameshifted RNA that split from a single translation site. On the right, the position of the spot is plotted through time. Taken from [36]. K. Lyon designed experiments, performed experiments, and analyses for Figure 4.20.

Using this constrained value for (k_{off}), we then fit to find the remaining parameters (k_{on}), (k_{ini}), (k_{el}), (k_{FSS}), and (k^*_{FSS}) (Table 4.1), with which the bursty frameshifting model could simultaneously reproduce all of our observations (Figure 4.19C-G), in contrast to the constitutive model (Figure 4.18). From these parameters, we calculate $1/(k_{on})$ ~170 min, meaning that RNA encoding the HIV-1 frameshift sequence switch to a frameshifting state rarely, on the timescale of a few hours (Table 4.1, $1/(k_{on})$). Once an RNA is in the frameshifting state, it remains there for ~13 mins on average (Table 4.1, $1/(k_{off})$), occasionally lasting up to an hour or more. We also predicted an elongation rate of ~3 aa/sec for both non-frameshifting and frameshifting sites. This is slower than previous results using autocorrelation (Figure A.4F). The discrepancy between the elongation rates may be due to cell-to-cell heterogeneity in stress levels. Once the translation initiation inhibitor, harringtonine, is added it stresses cells further. This could lead to slower responses whereby the elongation rate of ribosomes is slowed or stalled completely in cells that are too stressed. To confirm

the slower rates were due to harringtonine treatment, we repeated the harringtonine run-off with FLAG SM-tagged KDM5B and measured a slower elongation rate of ~3 aa/sec (data not shown).

To account for the different run-off delays seen at frameshifting and non-frameshifting sites (Figure 4.19F,H), the model required elongation pauses of $1/(\mathrm{k_{FSS}})$ ~43 s at non-frameshifting sites and $1/(\mathrm{k^*_{FSS}})$ ~72 s at frameshifting sites. We explored if codon usage could also explain the differences in run-off times. However, according to the codon adaptation index [109, 110], which is related to the speed at which each codon is translated in the simulation, there is no notable difference between the 0 and -1 frames (Figure 4.21). Moreover, the distinct pauses predicted by the model are comparable to those previously measured using *in vitro* and *in vivo* bulk assays [95, 101].

Table 4.1: Bursty frameshifting model parameters from simultaneous fit. Best fit bursty frameshift model parameters. Adapted from [36]. L. Aguilera, B. Munsky, T. Stasevich, K. Lyon.

Bursty frameshifting model parameters from simultaneous fit	
Ribosome elongation rate ($\mathrm{k_{el}}$)	3 ± 0.15 aa/sec
Ribosome initiation rate ($\mathrm{k_{ini}}$)	0.0244 ± 0.0015 $\mathrm{sec^{-1}}$
Switching on rate to frameshifting state ($\mathrm{k_{on}}$)	9.6x10-5 $\pm$ 7.8x10-5 $\mathrm{sec^{-1}}$
Switching off rate from frameshifting state ($\mathrm{k_{off}}$)	1.3x10-3 $\pm$ 3.5x10-4 $\mathrm{sec^{-1}}$
Pause rate at FSS in non-frameshifting state ($\mathrm{k_{FSS}}$)	0.0234 ± 0.0012 aa/sec
Pause rate at FSS* in frameshifting state ($\mathrm{k^*_{FSS}}$)	0.0139 ± 0.0011 aa/sec
Calculated quantities for bursting model with HA MF tag	
Fraction of frameshifted RNA	$7 \pm 5\%$
Average time in the non-shifting state $\langle \mathrm{b} \rangle$	170 ± 140 min
Average burst time in the shifting state $\langle \mathrm{b^*} \rangle$	12.8 ± 3.5 min
Number of ribosomes to initiate in a shifting state ($\mathrm{r_i^*}$)	19 ± 5
Number of ribosomes to initiate in a non-shifting state ($\mathrm{r_i}$)	250 ± 200
Time for a ribosome in shifting state to clear the FSS (τ^*_{FSS})	75 ± 6 sec
Time for a ribosome in non-shifting state to clear the FSS (τ_{FSS})	46 ± 2 sec
Note: Formulas given in the methods.	

Because the estimated average pause time, $1/\mathrm{k^*_{FSS}}$, is greater than the average initiation time, $1/\mathrm{k_{ini}}$, ribosomes could initiate faster than they clear the FSS with an excess rate of $\mathrm{k_{ini}} - \mathrm{k^*_{FSS}}$ = 0.011 per second and create upstream traffic jams in frameshifting sites. These traffic jams would continue to build for as long as the RNA remains in the frameshifted state, or approximately $1/\mathrm{k_{off}}$ = 12.8 min on average. Occasionally, traffic jams can extend all the way back to the start codon, as

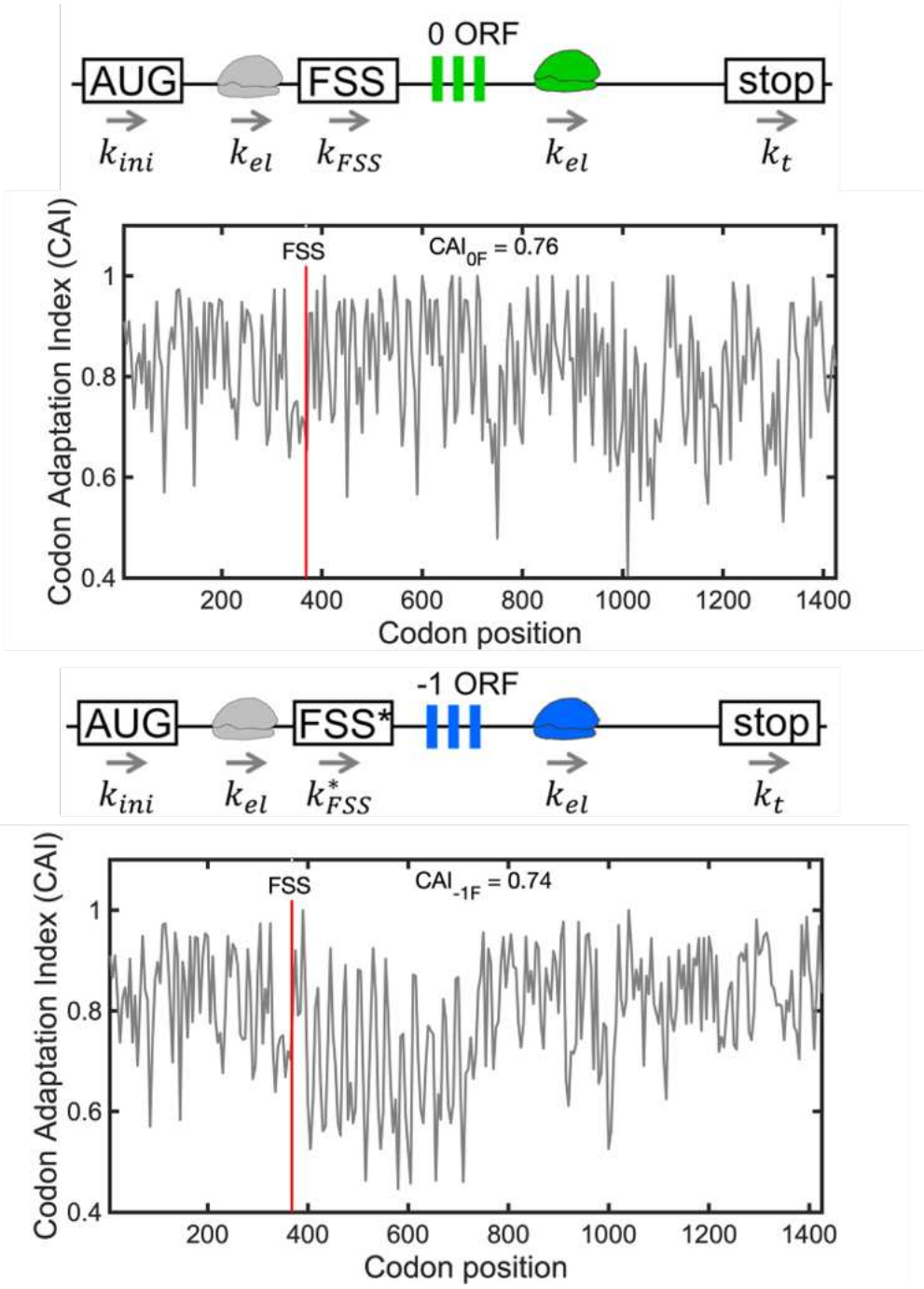


Figure 4.21: Codon usage plays a minor role in traffic jam formation. The codon adaptation index (CAI) for the 0 frame (top) and -1 frame (bottom) of the +FSS multi-frame tag. The y-axis shows the CAI calculated using the codon frequency in the human genome. The x-axis shows the length of the genes, in codons. In the plots, rare codons have low CAIs and common codons have high CAIs. The vertical red line represents the location of the frameshift sequence (FSS). Similar CAIs are obtained for the sequences in the 0 frame ($\mathrm{CAI}_{0\mathrm{F}} = 0.76$) and the -1 frame ($\mathrm{CAI}_{-1\mathrm{F}} = 0.74$). Taken from [36]. K. Lyon created cartoon schematics and was involved with conception of Figure 4.21.

seen in a sample simulations (Figure 4.22). Moreover, due to the long time it can take to clear traffic jams, frameshifting can persist for hours following the global shut down of translation initiation (with harringtonine, for example), as can be seen in simulations of the best-fit model Figure 4.22 and consistent with what we observed in Figure 4.12C and Figure 4.17. Thus, the final bursty model suggests a mechanism by which frameshifting can persist for long periods in the absence of translation initiation.

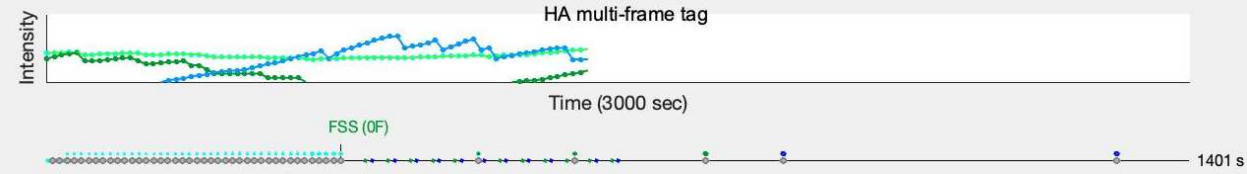


Figure 4.22: Sample simulation of best-fit bursty model with large ribosomal traffic jam. A best-fit bursty model simulation of a translation site encoding the HA MF tag. On the top, the simulated intensity of the site is tracked through time, with green showing signal from FLAG epitopes downstream of the FSS, blue showing signal from SunTag epitopes downstream of the FSS, and cyan showing signal from HA epitopes upstream of the FSS. On the bottom, individual ribosomes (gray dots) are shown moving along the reporter (gray line), with green, blue, and cyan squares marking the positions of FLAG, SunTag, and HA epitopes, respectively. As the ribosomes move along the repeat epitopes, the translated nascent chains appear as green (FLAG-tagged nascent chains), blue (Sun-tagged nascent chains), and cyan (HA-tagged nascent chains) dots that increase in size as more epitopes are translated. When the RNA is in the non-frameshifting state, the FSS is labeled 'FSS(OF)' in green; when the RNA is in the frameshifting state, the FSS is labeled 'FSS(-1F)' in blue. The simulated time is shown to the right of the reporter. A large ribosomal traffic jam is created when the RNA switches into the frameshifting state. Adapted from [36]. L. U. Aguilera and B. Munsky performed simulations and modeling to create Figure 4.22.

Chapter 5

Discussion

In this dissertation, I have discussed the establishment and validation of the Nascent Chain Tracking (NCT) method. This method was utilized in this dissertation to visualize and quantify individual RNA translation interactions with stress granules and individual RNA HIV-1 programmed frameshifting in living cells. In the following chapter, I will further explore the results presented in this dissertation. I will discuss the considerations of using antibody-based probes to study single-RNA translation, the observed dynamic interactions between RNA and stress granules, and finally I will discuss the dynamics of single RNA HIV-1 programmed ribosomal frameshifting.

5.1 Considerations of antibody-based probes for the study of active translation

Translation is a key process within the central dogma and is crucial for cellular viability, function, and development. Thus, it is critical that translation is well regulated and maintained for every cell type. Misregulation of translation can lead to the progression of disease states [80]. To truly understand biologically relevant regulation dynamics of translation, *in vivo* based single-RNA investigations are needed. Regulation dynamics of individual translating mRNA are of particular interest because these regulatory dynamics could drastically affect a cells fate, similar to how single-mRNA transcriptional dynamics can dramatically affect cell fate [111, 112].

To successfully image translational gene regulatory dynamics in living cells, a probe must be able to distinguish unique peptides irrespective of whether they are fully folded or mature. Fabs and scFvs both fit this criterion and have been successfully used for these purposes [37–41]. Both Fabs and scFvs can bind short peptide epitopes with high specificity, similar to the full antibodies from which they are derived. In addition, Fabs and scFvs have two key advantages over full antibodies for live-cell imaging purposes: (i) their small size; and (ii) their monovalency [113, 114]. First,

their small size allows them to quickly and efficiently access target epitopes in the complex and crowded cellular environment. For example, Fabs in living cells can pass through the nuclear pore and immediately bind target proteins within the nucleus [114]. By contrast, full antibodies cannot pass through the nuclear pore and, therefore, require cell division and nuclear envelope breakdown to access the nucleus [114]. Second, their monovalency prevents aggregation and interference. This is because Fabs and scFvs have a single binding domain that transiently binds only one target epitope at a time. By contrast, full antibodies are multivalent and, therefore, can bind multiple target epitopes at a time with high avidity. This means that target epitopes can not only be blocked for an extended period, but may also form aggregate chains of "-target-antibody-target-antibody-" [114].

While Fabs and scFvs are both derived from full antibodies, there are key differences in their production and form (Figure 5.1). Briefly, Fabs are generated directly from an antibody [115], whereas scFvs are genetically encoded Fab "mimics" [116]. To generate a Fab, the constant Fc base of a Y-shaped immunoglobulin (Ig) is digested away. This leaves two separate and identical "sticky" epitope-binding arms, Fabs, each approximately 50 kDa (or approximately one-third the weight of the full Ig). There are commercial kits for generating Fabs from Ig [117]; thus, it is relatively straightforward to make Fabs, but can be expensive. After the digestion is complete, each Fab is a heterodimer comprising one constant and one variable portion of the heavy and light chains from the parental Ig. These chains are held together by disulfide bonds and residual electrostatic interactions [118]. For live-cell imaging, Fabs are typically conjugated with small synthetic dyes [119].

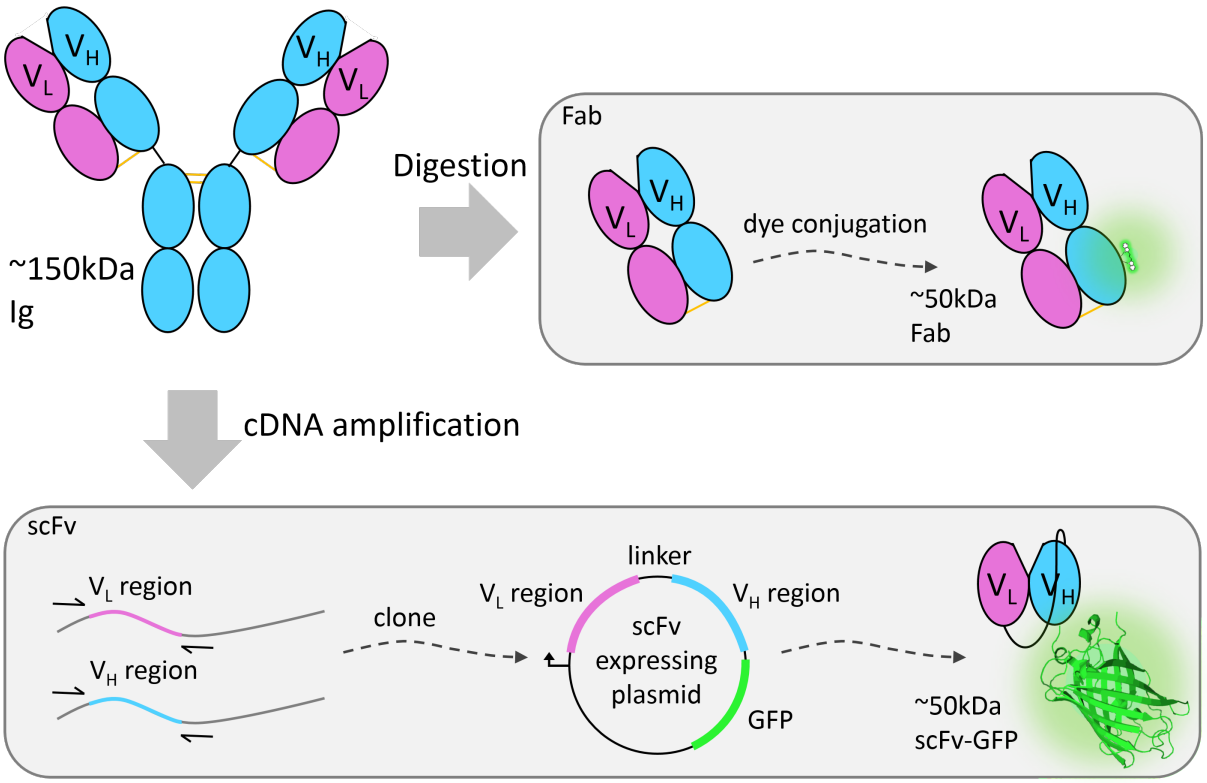


Figure 5.1: Antibody fragments (Fabs) and single-chain variable fragments (scFvs) IgG origins. Immunoglobins (Igs) comprise heavy (blue ovals) and light (pink ovals) chains. IgG binding to a target is facilitated by the heavy (VH) and light chain variable regions (VL). Fab fragments can be generated by digesting any commercially available IgG. Fabs retain their ability to bind the targets of IgG and can be conjugated with a small synthetic dye. Fab with a conjugated dye is approximately 50 kDa in size. By starting with the cDNA of the heavy and light variable regions of an IgG, they can be cloned into a plasmid vector to create an scFv. This allows variable regions to be connected by a polypeptide linker and even fused to a fluorescent protein, such as GFP [glowing beta-barrel structure; Protein Data Base (PDB) ID: 4KW4]. Once the scFv is fused to a GFP, their collective size is also approximately 50 kDa. Taken from [31]. K. Lyon created and contributed to the design of the cartoon schematics for Figure 5.1.

For an scFv, the variable regions of the heavy and light chains of a Fab are genetically linked via the N and C termini in a single chain [120]. The flexible linker acts in place of both the constant regions and the disulfide bond in Fabs. For live-cell imaging, scFvs are typically fused to fluorescent proteins [121], resulting in a probe that is comparable in size to a fluorescently conjugated Fab. One point to keep in mind is that scFvs are often unstable and insoluble in the cytoplasm of cells, in part because antibodies are naturally folded in a different environment, the endoplasmic reticulum (ER). In combination with the unnatural linker and fluorescent fusion tag, this makes it challenging

to engineer a functional scFv for intracellular expression [42, 43, 122]. For this reason, there are only a limited number of usable scFvs for live-cell imaging purposes [44, 121, 122]. However, a strategy implemented by the Stasevich lab has overcome many of the challenges associated with probe development by using a diverse set of scFv scaffolds. These scaffolds have been proven to fold properly and function within the reduced cytoplasm of living cells. Onto these scaffolds, they loop-grafted all six complementarity determining regions (CDRs) from an epitope specific antibody (HA) [123]. Nevertheless, once a functional intracellular scFv is generated, it is a more convenient, inexpensive, and reliable (no lot-to-lot variability [124]) probe than a Fab because the parental antibodies are no longer required.

Before imaging, Fabs and scFvs have to first be introduced into living cells. Since scFvs are genetically encoded, they can be transiently or stably expressed inside living cells or even whole organisms. By contrast, Fabs are not genetically encoded and, therefore, need to be physically loaded. Bead-loading is one of the simplest and most efficient way to load Fab into a living cell [114, 125–127]. This technique can load thousands of cells in a few seconds with less than a microgram of Fab. Other methods to get Fabs into cells include microinjection [37], electroporation [128], and, most recently, treatment with the pore-forming bacterial enzyme streptolysin O (SLO) [129]. All of these techniques are variations on the same theme: permeabilizing the membrane of cells so that purified Fabs can get in. Therefore, care should be taken to ensure cells remain healthy after loading.

While Fabs and scFvs offer unique advantages for imaging translational and post-translational gene regulatory dynamics in living cells compared with fluorescent fusion tags, they also suffer from several unique limitations that can make quantification difficult. Most of these limitations arise from the background signal from freely diffusing probes that are not bound to targets. This is necessary in antibody-based imaging because unbound probes are always present to quickly label newly made targets which are made during translation.

To mitigate artifacts that might arise from background signals, it is important to maintain an appropriate probe to target epitope ratio. The ratio of probe to target is specific, and dependent

on the desired assays application. For translational imaging, it is generally necessary to keep the concentration of probes higher than the concentration of targets. An appropriate concentration can be estimated from the number of target epitopes per cell using quantitative western blots [119]. In this method, a high background signal is less problematic because it is easily overcome by adding more repeat epitopes in the tag to amplify the signal above the noise. Therefore, it is more important to ensure there is an excess of free probe to quickly label the nascent peptides being translated. To illustrate, when single-RNA translation was imaged with the spaghetti monster repeat epitope tag, imaging was best between 2 to 10 hours after transient transfection of the construct. After 24 hours of expression, Fab was soaked up by all the mature tagged protein, making the new nascent peptide chains being translated more difficult to detect [37]. One way to address this issue is to fuse a degron to the tagged protein so that it is quickly degraded [38]. In this way, the probe to target ratio can be maintained at high levels indefinitely.

A related problem with live-cell imaging using antibody-based probes is interference from the epitope tag, the probe, or both. Since the epitopes are exogenous and not bound by other endogenous proteins, competitive interference is not an issue while imaging translation. However, the large number of probes recruited to the repeated epitope tags can create a large macromolecular complex on the order of megadaltons. It is easy to imagine that this large, unnatural complex would somehow interfere with normal translation. So far, this does not appear to be the case, because there is agreement between translation rates measured with repeat epitope tags of different sizes, including a 5x, 10x, 24x, 32x, and 56x SunTag [38–41] and a 10x FLAG or HA spaghetti monster tag [37]. In fact, the largest (56x SunTag) and one of the smallest (10x spaghetti monster) tags both gave relatively fast elongation rates (15-18 aa/sec and 8-12 aa/sec, respectively), whereas medium-sized tags had lower rates (3-6 aa/sec), suggesting little correlation between tag size and measured translation kinetics. Nevertheless, tags and probes may still interfere with normal transcript and/or protein behavior. As an example, when a 56x SunTag was inserted into the 5' end of one of the endogenous alleles encoding RNA polymerase II, even though transcripts appeared and were translated, the tagged polymerases were unable to enter the nucleus [41]. Thus, depending

on the specific application, it is important to check multiple tag sizes and their positioning within genes to properly assess whether tags interfere with the dynamics being measured.

Regardless of the challenges, we have successfully established nascent chain tracking (NCT), a multi-color method to visualize translation of multiple different reading frames on single RNA within living cells and in real-time. By imaging translation, a core process within the central dogma, with spatial and temporal precision, we expect NCT will positively contribute to the study of translational gene regulatory dynamics. Using this, we were able to measure the following on individual RNA: i) the average number of ribosomes within polysomes (given the length of a gene, 1 ribosome every 200 to 900 nucleotides), ii) the average elongation rate (3-10 aa/sec, depending on the gene), iii) the average initiation rate (initiate stochastically every 30-60 seconds), and iv) the formation of mutli-RNA translation sites (~5% form such sites).

5.2 A general mechanism for RNA-granule interactions

Applying the NCT methodology to the study of stress granules, we demonstrated that mRNAs interact with SGs and PBs in both stable and transient ways. In model assemblies, the exchange rates of components are related to both the strength of individual interactions and the valency, with more stable interactions occurring at higher valencies [130]. This suggests that the transient interactions with SGs or PBs represent a low-valency interaction mode but provide a docked state whereby mRNPs can form additional interactions with SG components to perhaps enter a stable locked state with higher valency (Figure 3.6C). Higher valency would be expected on longer mRNAs and larger SGs or PBs, providing an explanation for why longer mRNAs and larger RNP granules have more stable RNA-granule interactions. Our experiments provide clear evidence that mRNAs stably associated with SGs are translationally repressed, although mRNAs that are associated with polysomes can transiently interact with SGs. One hypothesis is that the 3' untranslated region (UTR) and/or partially exposed coding regions may be sufficient to form transient protein-protein or RNA-RNA interactions, but a fully exposed coding region is required to make sufficient interactions for stable SG association. This work identifies the docking and locking steps as two distinct steps in

RNP association with any RNP granule that can be modulated to affect RNA-granule association. We propose this bimodal nature of RNA-granule interaction as a general principle of any self organized RNP granule where RNP recruitment requires entry into a multivalent state.

5.3 Aspects of single RNA -1PRF of HIV-1 dynamics revealed by NCT

We employed a modified NCT tagging technology to study HIV-1 programmed ribosomal frameshifting, an aberrant form of translation. Frameshifting is a common tactic used by viruses to minimize their genomes for faster, more efficient replication in host cells, effectively getting two viral proteins for the price of one viral RNA. While the general architecture of frameshift sequences is well characterized and the dynamics of frameshifting have been measured with single-nucleotide precision *in vitro*, frameshifting has not been directly observed in a living system until now. To achieve this, we created a multi-frame repeat epitope tag that can light up single RNA translation sites in different colors depending on which open reading frame is being translated. Together with sensitive single-molecule microscopy and computational modeling, we have demonstrated six aspects of HIV-1 frameshifting: (i) frameshifted proteins originate from a small subset of RNA that frameshift with high efficiency; (ii) frameshifting occurs in bursts on single RNA that can last for several rounds of translation; (iii) ribosomes that frameshift are paused for longer at the frameshifting sequence than ribosomes that do not frameshift; (iv) pauses at the frameshift sequence induce ribosomal traffic jams that can maintain the production of frameshifted proteins despite global inhibition of translation; (v) frameshifting RNAs have reduced mobility and are often found in multi-RNA translation sites; and (vi) frameshifting can be stimulated by an oligo encoding the FSS. Figure 5.2 summarizes our findings.

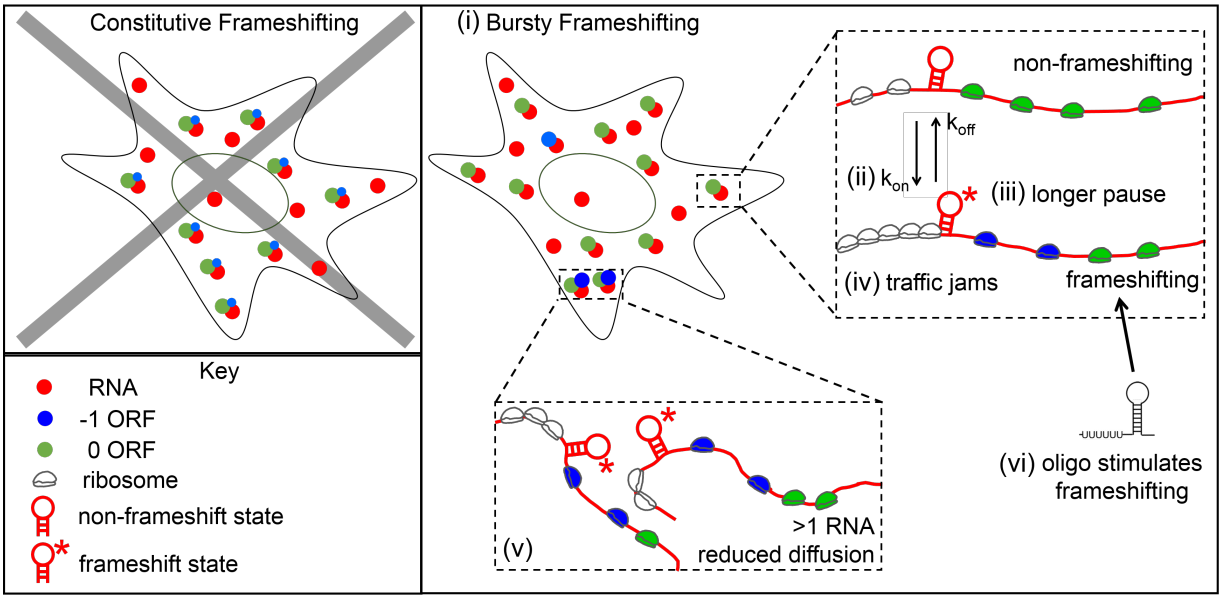


Figure 5.2: Summary of observed bursty frameshifting dynamics. (i) In contrast to constitutive frameshifting, in which all RNA behave the same, bursty frameshifting leads to heterogeneity between RNA. (ii) A state of persistent frameshifting is initiated with rate k_{on} and leads to relatively large bursts of frameshifted ribosomes translating the RNA. (iii) While in the frameshifting state, ribosomes experience a longer pause at the frameshift site. (iv) Increased pausing in the frameshifting state leads to traffic jams upstream of the frameshift site. (v) Frameshifting RNA have reduced mobility and are often found in multi-RNA translation sites. (vi) Frameshifting can be stimulated by an oligo encoding just the frameshift sequence. Taken from [36]. K. Lyon created and contributed to the design of the cartoon schematics for Figure 5.1.

In contrast to constitutive frameshifting on any RNA, our data indicate that frameshifting occurs in bursts on a subset of RNA. Bursty expression has been demonstrated by others, both at single transcription sites as well as at translation sites in bacteria [103] and eukaryotes [38]. The origin of frameshifting bursts is difficult to pinpoint. It is tempting to speculate that there is a unique structure the RNA takes that enhances frameshifting, particularly given the distinct pause times in our best-fit model. How this precisely occurs remains unclear, but our observation that the FSO (frameshift oligo) alone can stimulate frameshifting may offer some clues. In particular, this observation is consistent with the recently discovered repressor of frameshifting, shiftless [131]. Shiftless is thought to directly bind the FSS sequence and cause frameshifted ribosomes to be prematurely released. The FSO would presumably sequester shiftless and thereby indirectly lead to an increase

in frameshifting globally. More generally, the destabilization or unbinding of shiftless from the FSS could therefore be responsible for the bursts of frameshifting we observed. These possibilities will be interesting to explore in the future.

According to our best-fit model, pauses always occur at the frameshift sequence, with longer pauses associated with frameshifting RNA compared to non-frameshifting RNA. Pausing is therefore only a weak predictor of frameshifting, as others have shown [96, 132]. Nevertheless, the longer pauses associated with frameshifting suggest the frameshift sequence may adopt more than one state or structure, similar to what has been shown *in vitro* with sequences encoding pseudoknots [133]. Longer pauses result in longer ribosomal traffic jams. Our model predicts these jams can extend all the way back to the initiation site, involving up to ~40 ribosomes, as shown in Figure 4.19I. Our model predicts an elevated ribosomal occupancy as far back as ~400 codons from the FSS, a length that coincides almost perfectly with the length of the FSS-upstream sequence in HIV-1 Gag-Pol. This is not unprecedented, as ribosome profiling experiments have also found evidence for relatively high ribosomal densities as far back as the start codon of the gag protein [134]. These data suggest the strength of the HIV-1 pause may have evolved to, on occasion, produce the longest possible traffic jams. As clearance of these traffic jams takes time, frameshifting can persist on RNA for hours despite a global shut-down of translation initiation. In effect, the traffic jam acts like a battery that continually fuels the production of downstream frameshifted protein. This unique mechanism would allow viral proteins to continue to increase in numbers during cellular stress. An open question is how these long traffic jams manage to evade protein quality control [135, 136], which recently was shown to target the interface of di-ribosomes collisions [137].

Chapter 6

Future Perspectives

Nascent Chain Tracking (NCT) is a powerful tool that is capable of visualizing and quantifying unique translational regulatory dynamics at the single-RNA level, such as frameshifting, highlighted in this dissertation. NCT is now well established and characterized. It can therefore be immediately employed in a variety of other contexts. In this chapter I will discuss general implications of single RNA HIV programmed ribosomal frameshifting dynamics, +1 programmed ribosomal frameshifting, endogenous programmed ribosomal frameshifting, and finally, other unique translational regulatory dynamics. Within these sections I will explore experiments that can be done with NCT, their potential results, and relevance.

6.1 General implications of single RNA HIV programmed ribosomal frameshifting dynamics

Although there are structurally different RNA which induce -1 programmed frameshifting (-1PRF), they all share the same underlying mechanisms of a ribosomal pause and then a shift in the open reading frame during translation. All -1PRFs share a X XXY YYZ heptameric sequence or "slippery" sequence, where X is any nucleotide; Y is A or U; and Z is A, U, or C. In the 0 frame the codons are read as XXY YYZ and in the -1 frame as XXX YYY. The slippery sequence is where the frameshift occurs and is adjunct to another structured element downstream, either being a stem-loop or a pseudoknot [94].

How to target -1PRFs specifically has been a central question in the field of -1PRFs [80]. This inspired researchers to find small molecule modulators that would affect the viral -1PRF structure of HIV-1, but would not affect the structurally different endogenous -1PRF of PEG10 [88]. Despite these distinct responses, the generic -1PRF mechanisms are the same for the viral -1PRF of HIV-1 and the endogenous -1PRF of PEG10. It was discovered recently that an interferon-stimulated

repressor, shiftless, binds to a broad-spectrum of -1PRF inducing structures [131]. This would imply -1PRFs share a unique feature that allows them to be broadly targeted by shiftless. Together, this would argue that the model of bursty frameshifting presented in this dissertation may be broadly applicable. This could be validated by visualizing and quantifying another -1PRF, such as the -1PRF for PEG10, using NCT.

Viruses that employ the -1PRF genetic recoding mechanism, such as HIV-1, rely on a specific ratio of canonically translated product (Gag) to frameshifted translated product (Gag-Pol). If this ratio is disrupted, it can have profound defects on viral replication. The observations and model of -1PRF of HIV-1 presented in this dissertation will help guide future therapies to better target the subset of frameshifting RNAs. The bursty behavior of frameshfiting we observed suggests frameshifting RNA undergo a unique change in either their structure or bound factors. According to our best fit model, these changes are transient. By targeting these changes, new therapeutics can be developed to fight viral -1PRF frameshifting. Here, imaging the MF tag will be a powerful approach to rigorously test new therapeutics.

Because RNA encoding the -1PRF of HIV are mostly (>90%) non-frameshifting RNA, it could be most RNA adopt a structure that does not efficiently pause ribosomes. In contrast, frameshifting RNA may take on a structure that efficiently pauses. This would lead to di-ribosome collisions. This is in agreement with our measurements of distinct pauses at non-frameshifting RNA versus frameshifting RNA (Figure 4.1) and our observations that frameshifting RNA are more likely to induce ribosomal traffic jams (Figure 4.22). We also found frameshifting sites are often composed of more than one RNA. As shown in Figure 4.9B, the -1 frame signal within multi-RNA frameshifting is the same as within single RNA sites, whereas the 0 frame signal is reduced by nearly a factor of two. Thus, the other RNA in multi-RNA frameshifting sites is most often a non-frameshifting RNA.

Although the origin of multi-RNA frameshifting sites is not yet clear, one possible explanation is the additional non-frameshifting RNA somehow shields the frameshifting site from quality control surveillance. This would argue that the frameshift stimulatory sequence (FSS) of the -1PRF HIV can affect global ribosomal fidelity when collided ribosomes are involved and may be a general

mechanism shared by all viral -1PRFs. Although co-transfecting the frameshift sequence oligo (FSO) did not affect multi-RNA associations at frameshifting sites, at least according to RNA signal intensity, it did slightly increase frameshifting in cells expressing the control tag (Figure 4.4). This slight increase is not significant when compared to the addition of no FSO in cells expressing the control tag. However, this effect may be amplified and become more significant if more than 4 μg of FSO is co-transfected into cells with the control tag. If co-transfecting more FSO does increase overall frameshifting in cells expressing the control tag, this would suggest that the FSO may have an uncharacterized trans-acting effect that may interfere with quality control surveillance. -1 PRFs would then arise from both from traditional cis-acting effects as well as trans-acting effects. In support of this, HIV-1 RNAs are known to dimerize through interaction via their long terminal repeats (LTRs) located at either end of the RNAs [138]. Thus, these dimerized RNA may translate together and this trans-acting effect may allow both frameshifting RNA to successfully evade translational quality control within host cells.

Understanding the origin of the interactions between RNA in frameshifting sites is now important because it could reveal a common structure -1PRF RNAs take to bind di-ribosome collision interfaces and hide these interfaces from host cell quality control factors, such as ZNF598 and RACK1. This type of feature would likely be general and utilized by a broad-spectrum of -1PRF RNA. This is implied because, despite any structural differences, pausing is necessary for a -1PRF to occur. This pausing would ultimately lead to ribosomal traffic jams upstream that would then result in di-ribosome collisions. These collisions would need to be hidden from translational quality control factors for successful translation of the -1 frame. To test if FSSs of -1PRFs are capable of interfacing with di-ribosome collisions, native gels and structural studies could be carried out using either Electrophoretic Mobility Shift Assays (EMSAs) or cryoEM with purified components. An EMSA could be done by adding radiolabeled viral FSSs in the presence of purified di-ribosome collisions. If indeed the viral FSSs interact with the structure, then a shift will be seen as more di-ribosome is titrated into the assay. Furthermore, EMSAs could access whether quality control factors compete directly for the same interface as the FSSs. This could be done by adding in purified

shiftless, ZNF598, RACK1 or some combination. If competitive, then a shift back down in size would be seen. If no such shift occurs when the quality control factors are present, then the FSSs could be binding to the di-ribosomes allosterically. Structural studies could further uncover the exact type of interactions between the FSSs and the di-ribosome collisions. This could be revealed with cryoEM using purified mRNA harboring di-ribosome collisons together with viral FSSs.

A second, alternative explanation for the presence of additional non-frameshifting RNA within frameshifting sites is the notion of a translational quality control "hub." From our results, it is not clear if -1PRF causes RNA multimerization or if instead the RNA multimerization is a consequence of another tangential process induced by frameshifting, such as ribosome collisions. To test this, the -1PRF FSS could be removed within the MF tag and replaced with a structured stem-loop sequence capable of pausing ribosomes. If these multi-RNA associations are induced only by ribosomal traffic jams, then the RNA signal intensity of translating sites would be as bright as non-translating sites. Another way to test this would be to use the MF tag with the HIV-1 FSS in combination with sub-inhibitory concentrations of the elongation inhibitor emetine. Presumably, these sub-inhibitory concentrations would cause the elongation rate of ribosomes to slow enough to avoid the formation of traffic jams upstream of the pause induced by the -1PRF FSS. This type of strategy was used in recent work to understand di-ribosome collisions in the context of quality control surveillance [137]. Quantifying the RNA signal intensities using either of these strategies would provide insight into whether or not the multi-RNA association is the direct result of -1PRF or if, instead, it is only a consequence of ribosomal traffic jams.

It has been speculated that ribosomes can sense the status of nascent chains and thereby act as regulatory hubs for the recruitment of quality control factors [139]. Ribosomes that recruit similar quality control machinery may assemble into larger hubs, similar to ATP-dependent SG fusion events, as shown in this dissertation (Figure 3.1e). Quality control factors, such as ZNF598 or RACK1, have been shown to target di-ribosomal collisions, which would likely occur within the large traffic jams induced by the -1PRF pause. The newly discovered frameshift repressor shiftless would also likely accumulate in these hubs [131]. One quick way to test if shiftless is directly

involved would be to induce more shiftless expression by adding interferon to cells expressing the MF tag. Depending on the timescale for induction of shiftless, we may see frameshifting sites assemble together in real-time. To further distinguish the effects of multimerization, NCT could be used to track frameshifting sites and measure the RNA signal intensities in cell lines that have either ZNF598, RACK1, shiftless or some combination of the three knocked out. If any of these proteins are involved with the multimerization of frameshifting sites, then the RNA signal within frameshifting sites should become similar to non-frameshifting sites. Investigating this further could provide some hints of how quality control distinguishes, targets, and assembles RNA with large ribosomal traffic jams together.

6.2 +1 ribosomal programmed frameshifting

Unlike -1PRF, which involves a generic slippery site and pause site, +1 programmed ribosomal frameshifting (+1PRF) signals and mechanisms appear to be case-specific [140]. Despite this, +1PRFs all share a slippery sequence with a rare codon that is important to shift the open reading frame by +1 nucleotide, toward the 3' direction of the mRNA. Studying single viral RNA could help clarify case-by-case mechanisms and help identify new strategies to perturb these processes.

Recently, the endemic influenza A virus (IAV), a virus that infects ~49 million people per year, was shown to harbor a +1PRF that is important for the expression of an endonuclease which represses host cellular gene expression [141–143]. The +1PRF of IAV was found to have unique sequence motifs that influence the +1 frameshift efficiency. This is an interesting development because canonical +1 frameshifting in bacteria or yeast is thought only to be induced by a rare codon whose tRNA is limiting within the slippery sequence. In addition, these unique sequence motifs were found to be upstream of the +1 frameshift slippery sequence, whereas such sequence motifs for -1 frameshifts are downstream of the slippery sequence. This suggests that the upstream structures in +1 frameshifts may interact with the back of ribosomes to provide a barrier that makes a +1 nucleotide frameshift more favorable, similar to the downstream -1 frameshift structures that interact with the front of ribosomes. Because the barrier for +1PRFs is not a structure the ribosome

needs to unwind and translate through, the pause induced by +1 frameshifting may be much shorter than that of -1 frameshift pauses.

Using the MF tag to visualize, quantify, and ultimately model the IAV +1PRF could reveal that the pause for +1 frameshifting is not as extreme as -1PRFs and therefore does not induce ribosomal traffic jams. This could also mean that +1 frameshifting sites do not multimerize, which would support distinct mechanisms for +1 frameshifts versus -1 frameshifts, as other have concluded with *in vitro* assays [140].

Another interesting +1 frameshifting mechanism that could be investigated immediately with the MF tag is the polyamine-stimulated +1 ribosomal frameshift found at the endogenous eukaryotic gene OAZ1. When polyamine concentrations are elevated, a +1 nucleotide frameshift directs ribosomes to translate an inhibitor of polyamine synthesis [90]. Structures important for this +1 frameshift are upstream and downstream of the slippery sequence, with the upstream sequences being more important for the overall +1 frameshift and the downstream sequences being more relevant for the inducibility of the +1 frameshift efficiency. Consistent with the idea that +1 frameshifts are mechanistically different from -1PRFs, it was found that +1 frameshifts did not produce pauses detectable with *in vitro* toe-printing assays [144]. Therefore, visualizing +1PRFs with the MF tag may reveal there to be no observable multi-RNA associations at +1 frameshifting sites.

By visualzing +1PRFs at the individual RNA level, multimerization of frameshifting sites can be directly tested and distinct mechanisms can be worked out on a case-by-case basis. Any common features observed would furthermore help constrain general models for +1PRFs and possibly lead to new strategies to perturb the +1PRF of IAV. In addition, understanding this +1 inducible system at the single RNA level could help with pharmaceutical and biotechnology applications, wherein two proteins or enzymes could be encoded and induced on one mRNA.

6.3 Endogenous -1 ribosomal programmed frameshifting

Nearly ~10% of eukaryotic genes are predicted to harbor a slippery sequence as well as downstream structures to induce a -1PRF. While viral -1PRFs direct ribosomes to translate a -1 open reading frame to create a fusion polyprotein, nearly all endogenous -1PRFs appear to direct ribosomes to a premature termination codon [140]. This suggests that most cellular -1PRFs are present to regulate gene expression on specific mRNA through the nonsense-mediated decay (NMD) pathway. In this sense, targeting the general mechanism of -1PRF as a viral therapeutic could be detrimental to not only the virus, but also the host cell. However, because of the diversity of structures and sequences used to pause and "slip" ribosomes at these endogenous mRNA, it is likely that the -1PRF signals could be regulated specifically. In support of this, it has been shown that viral -1PRFs can be specifically targeted and their -1 frameshift efficiency effected differently than that of endogenous -1PRFs with the use of small molecule modulators, presumably because of different structures used to induce the ribosomal pausing [88]. By characterizing endogenous -1PRF using the multi-frame (MF) NCT technology presented in this dissertation, distinct features could be revealed and these could lead to more specific targeting of -1PRFs overall.

An interesting observation on the specific regulation of cellular -1PRFs comes from the mammalian gene CCR5. This gene encodes a cytokine receptor that has a strong predicted -1PRF signal within the coding region directing ribosomes to a premature termination codon. The -1PRF efficiency of CCR5 can be enhanced by small miRNAs that can interact specifically with the -1PRF structure of CCR5 [80]. Upregulation of the -1PRF of CCR5 destabilizes the CCR5 mRNA via the NMD pathway, leading to an overall reduction in the cytokine receptor within the cellular membrane. This effect could be exploited by viruses during infection. Visualizing the -1PRF of CCR5 using the MF NCT tag would help further characterize this modulation and lead to a better understanding of miRNA targeting of -1PRFs. Since miRNAs target the -1PRF signal of the CCR5 sequence directly and enhance its frameshifting, then performing NCT using the MF tag would help characterize the dynamics of miRNA action. For example, miRNA may enhance frameshifting

either by extending the length of frameshifting bursts or altering their frequency. These possibilities will be easy to distinguish via tracking mRNA encoding MF-tagged CCR5.

Similar to CCR5, the MF NCT tag can immediately be used to study another interesting endogenous -1PRF. Unlike most endogenous -1PRFs, which cause a premature termination codon, the -1PRF of PEG10 encodes a polyprotein homologous to the Gag and Pol of HIV-1 [82]. PEG10 is a mammalian gene that is a potent cell growth regulator that promotes cell-cycle progression. Upregulation of PEG10 has been shown to promote the progression of lethal cancers types [145,146]. For that reason, understanding unique characteristics of this endogenous -1PRF could aid in the development of drugs targeting it or other alternative strategies to specifically perturb its frameshifting. Like CCR5, NCT on PEG10 would be useful to better understand PEG10 regulation. Overall, more single-molecule observations of different types of -1PRF are now needed, not just to better understand -1PRF dynamics, but also to guide the design of more specific targets of viral -1PRFs that will not impact endogenous -1PRFs. [80].

6.4 Unique translational regulatory dynamics that can be studied with multi-color NCT

Beyond frameshifting, any other non-canonical translation process involving more than one open reading frame can be investigated with NCT using the MF tag. Possible applications include: start-codon selection, leaky scanning, ribosomal shunting, and general translation fidelity.

In fact, another paper using a similar MF tag was used to examine upstream and downstream open reading frame selection [147]. Like us, the authors also saw a high degree of heterogeneity between translating RNAs with bursts of translation initiation in multiple open reading frames, similar to the bursts of frameshifting we observed at the HIV-1 frameshift sequence. Translational heterogeneity may therefore be far more common than originally thought, particularly when it comes to non-canonical translation.

To extend these observations, the MF tag could be used to help better predict non-AUG translation initiation and verify non-AUG translation initiation events. One current central question in this

field is whether it is possible to predict non-AUG translation initiation at any given transcript. Although non-AUG translation initiation is thought to represent errors involving initiation machinery, there is growing evidence that non-AUG translation initiation is important for a number of cellular and viral genes and that misregulation can lead to cancer and neurodegeneration [148]. Traditional assays do not have the sensitivity required to give further insight into the underlying mechanisms involving non-AUG initiation. Using the MF NCT tag, we could provide such insight even if these events are rare. By placing a canonical start codon near a downstream non-AUG start codon, we can begin to observe in what specific contexts it is preferentially initiated. This context is important for diseases like Huntingtin, where large CAG expansions induce Repeat-Associated Non-AUG (RAN) translation initiation, which leads to a progression of the disease. By imaging and quantifying these events in a live cell, in combination with computational modeling, we could better predict non-AUG translation initiation events on any transcript.

Ribosomal shunting is a recently observed translational phenomenon whereby ribosomes translate a small upstream open reading frame (uORF) and terminate near a large structured region. This large structured region is bypassed by ribosomes that terminate at the uORF and subsequently "shunt" onto a canonical start site downstream of the large structured region [149]. Shunting has been shown to be used by mammalian genes and viruses, such as human papillomavirus. Due to the complexity of the large structured region, ribosomal shunts are thought to be regulated by many trans-acting factors [150]. A question still unknown is what trans-acting factors help stimulate this process in the viral and cellular contexts. A multi-color NCT approach to understand this mechanism further could be achieved by encoding small epitope tags in both the uORF and the downstream coding sequence. This would not only allow the visualization of ribosomal shunting, but provide a way to perform immunoprecipitations of the upstream and downstream ribosomes and thereby reveal the trans-acting factors involved in either the cellular or viral processes. This would also help guide biotechnological and pharmaceutical applications or shunting. Dissecting the translational mechanisms involved in shunting would help the general understanding for stress

induced genes and help reveal alternative treatments for viruses that employ this mechanism, such as human papillomavirus.

Finally, one highly debated topic that NCT could be applied to is the study of nuclear translation. Ribosomes are large ribonucleoproteins, made up of RNA and proteins, that assemble in the nucleolus, a dense region in the nucleus where rRNA is transcribed. The active form of the ribosome is thought to be present in the cytoplasm, where the bulk of translational co-factors are found and where the maturation steps of the ribosome exclusively occur [151]. However, because the ribosome is made up of RNA and proteins, it is easy to imagine that some ribosomal proteins may need to be co-translationally folded with rRNA during ribosome biogenesis. Co-translational folding and assembly has been shown for many other large proteins [152] and could be potentially the case for the ribosome. If true, some ribosomal proteins may harbour zipcodes in the untranslated region (UTR) which direct them to the nucleolus. Zipcodes are unique structured RNA sequences that are bound by proteins that shuttle the RNA to a specific site in the cell, a well studied example being the beta-actin zipcode [153]. To test this, different 3' UTRs of ribosomal proteins could be inserted into the 3' UTR of a NCT reporter mRNA. Nuclear translation could then be confirmed by using a multi-color approach of marking a clear nuclear boundary, an mRNA of interest, and the nascent peptides of elongating ribosomes. This could reveal a unique zipcode structure or sequence needed to bring mRNA to the nucleolus for translation. The dynamics of this process would likely differ from their cytosolic counterparts in mobility and kinetics due to the the density of the nucleolus and limited translation factors, respectively.

Other theories for nuclear translation involve non-sense mediated decay (NMD) and immuno-surveillance. NMD is used for detection of mRNAs which have a premature stop codon. Briefly, splicing machinery present at sites of splicing on a nascent mRNA are "knocked-off" by the first successful round of translation. If splicing machinery is not removed upon the first round of translation, then the mRNA is shuttled to the NMD pathway, where it is ultimately degraded [154]. There are studies that suggest the initial round of translation, which removes splicing machinery on nascent mRNA, occurs in the nucleus [155–157]. Likewise, another plausible argument for nuclear

translation is in the context of immunosurveillance, which involves specialized ribosomes referred to as "immunoribosomes," which are localized in the nucleus. Immunoribosomes translate short defective ribosomal products (DRiPs), from either exonic or intronic RNA, and then transfer the DRiPs to MHC class I molecules that are involved in displaying the DRiPs on the cell surface to enable T cell immunosurveillance [158]. To test this, a multi-color construct encoded with one repeated epitope within an exon and another distinct repeated epitope within an intron would need to be used. With such a construct, if an immunoribosome does exist, then it would result in intronic translation events occurring exclusively within the nucleus, while the exonic translation events would be found more often in the cytosol. This result would be interesting as it would support the idea that there may be sub-classes of ribosomes that are not as processive as cytosolic counterparts.

Another strategy to test if nuclear translation can even occur would be to tether canonically capped mRNA in the nucleus. Tethering mRNA in the nucleus could be done using the LambdaN-BoxB system. This is done by encoding BoxB stem-loops into the 3' UTR, along with MS2 stem-loops for visualization. In combination with a histone-LambdaN fusion, newly translated mRNAs with the BoxB stem-loops would bind histone-LambdaN in the nucleus and thereby be captured before nuclear export. NCT on this construct would allow one to directly visualize whether or not mRNA tethered within the nucleus are still translating. If so, this could shed insight into the kinetics of immunoribosomes.

In summary, as illustrated here, NCT has many potential applications that will shed light on canonical and non-canonical translation. It should therefore become a powerful tool for dissecting complex RNA regulatory dynamics in a variety of important contexts.

Bibliography

[1] Francis Crick. On protein synthesis. In *Symp Soc Exp Biol*, volume 12, page 8, 1958.

[2] Anthony P Schuller and Rachel Green. Roadblocks and resolutions in eukaryotic translation. *Nature Reviews Molecular Cell Biology*, page 1, 2018.

[3] DF Rolfe and Guy C Brown. Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiological reviews*, 77(3):731–758, 1997.

[4] Kivie Moldave. Eukaryotic protein synthesis. *Annual review of biochemistry*, 54(1):1109–1149, 1985.

[5] Alan G Hinnebusch. The scanning mechanism of eukaryotic translation initiation. *Annual review of biochemistry*, 83:779–812, 2014.

[6] Gloria A Brar. Beyond the triplet code: context cues transform translation. *Cell*, 167(7):1681–1692, 2016.

[7] Craig R Stumpf, Melissa V Moreno, Adam B Olshen, Barry S Taylor, and Davide Ruggero. The translational landscape of the mammalian cell cycle. *Molecular cell*, 52(4):574–582, 2013.

[8] Venki Ramakrishnan. Ribosome structure and the mechanism of translation. *Cell*, 108(4):557–572, 2002.

[9] G Zhouravleva, L Frolova, X Le Goff, R Le Guellec, S Inge-Vechtomov, L Kisselev, and M Philippe. Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *The EMBO journal*, 14(16):4065–4072, 1995.

[10] Yoshihiro Shimizu, Akio Inoue, Yukihide Tomari, Tsutomu Suzuki, Takashi Yokogawa, Kazuya Nishikawa, and Takuya Ueda. Cell-free translation reconstituted with purified components. *Nature biotechnology*, 19(8):751, 2001.

[11] Elena Z Alkalaeva, Andrey V Pisarev, Lyudmila Y Frolova, Lev L Kisselev, and Tatyana V Pestova. In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. *Cell*, 125(6):1125–1136, 2006.

[12] Mark C Capece, Guy L Kornberg, Alexey Petrov, and Joseph D Puglisi. A simple real-time assay for in vitro translation. *RNA*, 21(2):296–305, 2015.

[13] Albert Tsai, Alexey Petrov, R Andrew Marshall, Jonas Korlach, Sotaro Uemura, and Joseph D Puglisi. Heterogeneous pathways and timing of factor departure during translation initiation. *Nature*, 487(7407):390, 2012.

[14] Chunlai Chen, Haibo Zhang, Steven L Broitman, Michael Reiche, Ian Farrell, Barry S Cooperman, and Yale E Goldman. Dynamics of translation by single ribosomes through mRNA secondary structures. *Nature structural & molecular biology*, 20(5):582, 2013.

[15] Ruth M Barrett, Hui-wen Liu, Haihong Jin, Richard H Goodman, and Michael S Cohen. Cell-specific profiling of nascent proteomes using orthogonal enzyme-mediated puromycin incorporation. *ACS chemical biology*, 11(6):1532–1536, 2016.

[16] Robert AJ Signer, Jeffrey A Magee, Adrian Salic, and Sean J Morrison. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature*, 509(7498):49, 2014.

[17] Jing Liu, Yangqing Xu, Dan Stoleru, and Adrian Salic. Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proceedings of the National Academy of Sciences*, 109(2):413–418, 2012.

[18] Wen Zhang, Jack A Dunkle, and Jamie HD Cate. Structures of the ribosome in intermediate states of ratcheting. *Science*, 325(5943):1014–1017, 2009.

[19] Daniel H Goldman, Christian M Kaiser, Anthony Milin, Maurizio Righini, Ignacio Tinoco, and Carlos Bustamante. Mechanical force releases nascent chain–mediated ribosome arrest in vitro and in vivo. *Science*, 348(6233):457–460, 2015.

[20] Xiaohui Qu, Jin-Der Wen, Laura Lancaster, Harry F Noller, Carlos Bustamante, and Ignacio Tinoco. The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature*, 475(7354):118, 2011.

[21] Nicholas T Ingolia, Sina Ghaemmaghami, John RS Newman, and Jonathan S Weissman. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, 324(5924):218–223, 2009.

[22] Nicholas T Ingolia, Liana F Lareau, and Jonathan S Weissman. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, 147(4):789–802, 2011.

[23] Stephen N Floor and Jennifer A Doudna. Tunable protein synthesis by transcript isoforms in human cells. *Elife*, 5:e10921, 2016.

[24] Jeffrey A Chao, Young J Yoon, and Robert H Singer. Imaging translation in single cells using fluorescent microscopy. *Cold Spring Harbor perspectives in biology*, 4(11):a012310, 2012.

[25] Tatsuya Morisaki and Timothy J Stasevich. Quantifying single mRNA translation kinetics in living cells. *Cold Spring Harbor perspectives in biology*, 10(11):a032078, 2018.

[26] James M Halstead, Timothée Lionnet, Johannes H Wilbertz, Frank Wippich, Anne Ephrussi, Robert H Singer, and Jeffrey A Chao. An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science*, 347(6228):1367–1671, 2015.

[27] Edouard Bertrand, Pascal Chartrand, Matthias Schaefer, Shailesh M Shenoy, Robert H Singer, and Roy M Long. Localization of ASH1 mRNA particles in living yeast. *Molecular cell*, 2(4):437–445, 1998.

[28] Jeffrey A Chao, Yury Patskovsky, Steven C Almo, and Robert H Singer. Structural basis for the coevolution of a viral RNA–protein complex. *Nature structural & molecular biology*, 15(1):103, 2008.

[29] Antoine Coulon, Carson C Chow, Robert H Singer, and Daniel R Larson. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nature reviews genetics*, 14(8):572, 2013.

[30] Antoine Coulon, Matthew L Ferguson, Valeria de Turris, Murali Palangat, Carson C Chow, and Daniel R Larson. Kinetic competition during the transcription cycle results in stochastic rna processing. *Elife*, 3:e03939, 2014.

[31] Kenneth Lyon and Timothy J Stasevich. Imaging translational and post-translational gene regulatory dynamics in living cells with antibody-based probes. *Trends in Genetics*, 33(5):322–335, 2017.

[32] Alexis J Rodriguez, Shailesh M Shenoy, Robert H Singer, and John Condeelis. Visualization of mRNA translation in living cells. *The Journal of cell biology*, 175(1):67–76, 2006.

[33] Schuyler B Van Engelenburg, Theresa Nahreini, and Amy E Palmer. FACS-Based Selection of Tandem Tetracysteine Peptides with Improved ReAsH Brightness in Live Cells. *Chembiochem*, 11(4):489–493, 2010.

[34] Shintaro Iwasaki and Nicholas T Ingolia. Seeing translation. *Science*, 352(6292):1391–1392, 2016.

[35] Stephanie L Moon, Tatsuya Morisaki, Anthony Khong, Kenneth Lyon, Roy Parker, and Timothy J Stasevich. Multicolour single-molecule tracking of mRNA interactions with RNP granules. *Nature cell biology*, 21:162–168, 2019.

[36] Kenneth R Lyon, Luis U Aguilera, Tatsuya Morisaki, Brian Munsky, and Timothy J Stasevich. Live-cell single RNA imaging reveals bursts of translational frameshifting. *Molecular Cell*, 2019.

[37] Tatsuya Morisaki, Kenneth Lyon, Keith F DeLuca, Jennifer G DeLuca, Brian P English, Zhengjian Zhang, Luke D Lavis, Jonathan B Grimm, Sarada Viswanathan, Loren L Looger,

et al. Real-time quantification of single RNA translation dynamics in living cells. *Science*, 352(6292):1425–1429, 2016.

[38] Bin Wu, Carolina Eliscovich, Young J Yoon, and Robert H Singer. Translation dynamics of single mRNAs in live cells and neurons. *Science*, 352(6292):1430–1435, 2016.

[39] Xiaowei Yan, Tim A Hoek, Ronald D Vale, and Marvin E Tanenbaum. Dynamics of translation of single mRNA molecules in vivo. *Cell*, 165(4):976–989, 2016.

[40] Chong Wang, Boran Han, Ruobo Zhou, and Xiaowei Zhuang. Real-time imaging of translation on single mRNA transcripts in live cells. *Cell*, 165(4):990–1001, 2016.

[41] Xavier Pichon, Amandine Bastide, Adham Safieddine, Racha Chouaib, Aubin Samacoits, Eugenia Basyuk, Marion Peter, Florian Mueller, and Edouard Bertrand. Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *J Cell Biol*, 214(6):769–781, 2016.

[42] Jozef Hanes, Lutz Jermutus, Susanne Weber-Bornhauser, Hans Rudolf Bosshard, and Andreas Plückthun. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proceedings of the National Academy of Sciences*, 95(24):14130–14135, 1998.

[43] Arne Wörn, Adrian Auf der Maur, Dominik Escher, Annemarie Honegger, Alcide Barberis, and Andreas Plückthun. Correlation between in vitro stability and in vivo performance of anti-GCN4 intrabodies as cytoplasmic inhibitors. *Journal of Biological Chemistry*, 275(4):2795–2803, 2000.

[44] Marvin E Tanenbaum, Luke A Gilbert, Lei S Qi, Jonathan S Weissman, and Ronald D Vale. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, 159(3):635–646, 2014.

[45] Jian-Rong Yang, Xiaoshu Chen, and Jianzhi Zhang. Codon-by-codon modulation of translational speed and accuracy via mRNA folding. *PLoS biology*, 12(7):e1001910, 2014.

[46] Nicola Green, Hannah Alexander, Arthur Olson, Stephen Alexander, Thomas M Shinnick, J Gregor Sutcliffe, and Richard A Lerner. Immunogenic structure of the influenza virus hemagglutinin. *Cell*, 28(3):477–487, 1982.

[47] Thomas P Hopp, Kathryn S Prickett, Virginia L Price, Randell T Libby, Carl J March, Douglas Pat Cerretti, David L Urdal, and Paul J Conlon. A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/technology*, 6(10):1204, 1988.

[48] Sarada Viswanathan, Megan E Williams, Erik B Bloss, Timothy J Stasevich, Colenso M Speer, Aljoscha Nern, Barret D Pfeiffer, Bryan M Hooks, Wei-Ping Li, Brian P English, et al. High-performance probes for light and electron microscopy. *Nature methods*, 12(6):568, 2015.

[49] Xavier Darzacq, Jie Yao, Daniel R Larson, Sébastien Z Causse, Lana Bosanac, Valeria De Turris, Vera M Ruda, Timothee Lionnet, Daniel Zenklusen, Benjamin Guglielmi, et al. Imaging transcription in living cells. *Annual review of biophysics*, 38:173–196, 2009.

[50] Robert A Coleman, Zhe Liu, Xavier Darzacq, Robert Tjian, Robert H Singer, and Timothée Lionnet. Imaging transcription: past, present, and future. In *Cold Spring Harbor symposia on quantitative biology*, volume 80, pages 1–8. Cold Spring Harbor Laboratory Press, 2015.

[51] Antoine Coulon and David R Larson. Fluctuation analysis: dissecting transcriptional kinetics with signal theory. In *Methods in enzymology*, volume 572, pages 159–191. Elsevier, 2016.

[52] Michael Doyle and Michael A Kiebler. Mechanisms of dendritic mRNA transport and its role in synaptic tagging. *The EMBO journal*, 30(17):3540–3552, 2011.

[53] Daniel Nathans. Puromycin inhibition of protein synthesis: incorporation of puromycin into peptide chains. *Proceedings of the National Academy of Sciences of the United States of America*, 51(4):585, 1964.

[54] Tilman Schneider-Poetsch, Jianhua Ju, Daniel E Eyler, Yongjun Dang, Shridhar Bhat, William C Merrick, Rachel Green, Ben Shen, and Jun O Liu. Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nature chemical biology*, 6(3):209, 2010.

[55] Sanne Boersma, Deepak Khuperkar, Bram MP Verhagen, Stijn Sonneveld, Jonathan B Grimm, Luke D Lavis, and Marvin E Tanenbaum. Multi-color single molecule imaging uncovers extensive heterogeneity in mRNA decoding. *bioRxiv*, page 477661, 2018.

[56] David SW Protter and Roy Parker. Principles and properties of stress granules. *Trends in cell biology*, 26(9):668–679, 2016.

[57] Joshua R Wheeler, Tyler Matheny, Saumya Jain, Robert Abrisch, and Roy Parker. Distinct stages in stress granule assembly and disassembly. *Elife*, 5:e18413, 2016.

[58] Benedikt Niewidok, Maxim Igaev, Abel Pereira da Graca, Andre Strassner, Christine Lenzen, Christian P Richter, Jacob Piehler, Rainer Kurre, and Roland Brandt. Single-molecule imaging reveals dynamic biphasic partition of RNA-binding proteins in stress granules. *J Cell Biol*, 217(4):1303–1318, 2018.

[59] Ming-Tzo Wei, Shana Elbaum-Garfinkle, Alex S Holehouse, Carlos Chih-Hsiung Chen, Marina Feric, Craig B Arnold, Rodney D Priestley, Rohit V Pappu, and Clifford P Brangwynne. Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nature chemistry*, 9(11):1118, 2017.

[60] Marina Feric, Nilesh Vaidya, Tyler S Harmon, Diana M Mitrea, Lian Zhu, Tiffany M Richardson, Richard W Kriwacki, Rohit V Pappu, and Clifford P Brangwynne. Coexisting liquid phases underlie nucleolar subcompartments. *Cell*, 165(7):1686–1697, 2016.

[61] Amandine Molliex, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P Kanagaraj, Hong Joo Kim, Tanja Mittag, and J Paul Taylor. Phase separation by low complexity domains

promotes stress granule assembly and drives pathological fibrillization. *Cell*, 163(1):123–133, 2015.

[62] Aristeidis P Sfakianos, Alan J Whitmarsh, and Mark P Ashe. Ribonucleoprotein bodies are phased in. *Biochemical Society Transactions*, 44(5):1411–1416, 2016.

[63] J Ross Buchan and Roy Parker. Eukaryotic stress granules: the ins and outs of translation. *Molecular cell*, 36(6):932–941, 2009.

[64] Nancy Kedersha, Sarah Tisdale, Tyler Hickman, and Paul Anderson. Real-time and quantitative imaging of mammalian stress granules and processing bodies. *Methods in enzymology*, 448:521–552, 2008.

[65] Jennifer A Schisa. Effects of stress and aging on ribonucleoprotein assembly and function in the germ line. *Wiley Interdisciplinary Reviews: RNA*, 5(2):231–246, 2014.

[66] Indulekha P Sudhakaran, Jens Hillebrand, Adrian Dervan, Sudeshna Das, Eimear E Holohan, Jörn Hülsmeier, Mihail Sarov, Roy Parker, K VijayRaghavan, and Mani Ramaswami. FMRP and Ataxin-2 function together in long-term olfactory habituation and neuronal translational control. *Proceedings of the National Academy of Sciences*, 111(1):E99–E108, 2014.

[67] Amal M El-Naggar and Poul H Sorensen. Translational control of aberrant stress responses as a hallmark of cancer. *The Journal of pathology*, 244(5):650–666, 2018.

[68] Elda Grabocka and Dafna Bar-Sagi. Mutant KRAS enhances tumor cell fitness by upregulating stress granules. *Cell*, 167(7):1803–1813, 2016.

[69] Ke Zhang, J Gavin Daigle, Kathleen M Cunningham, Alyssa N Coyne, Kai Ruan, Jonathan C Grima, Kelly E Bowen, Harsh Wadhwa, Peiguo Yang, Frank Rigo, et al. Stress granule assembly disrupts nucleocytoplasmic transport. *Cell*, 173(4):958–971, 2018.

[70] Manuela Neumann, Deepak M Sampathu, Linda K Kwong, Adam C Truax, Matthew C Micsenyi, Thomas T Chou, Jennifer Bruce, Theresa Schuck, Murray Grossman, Christopher M

Clark, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*, 314(5796):130–133, 2006.

[71] Yun R Li, Oliver D King, James Shorter, and Aaron D Gitler. Stress granules as crucibles of ALS pathogenesis. *J Cell Biol*, 201(3):361–372, 2013.

[72] Jonathan B Grimm, Brian P English, Jiji Chen, Joel P Slaughter, Zhengjian Zhang, Andrey Revyakin, Ronak Patel, John J Macklin, Davide Normanno, Robert H Singer, et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nature methods*, 12(3):244, 2015.

[73] Anthony Khong, Tyler Matheny, Saumya Jain, Sarah F Mitchell, Joshua R Wheeler, and Roy Parker. The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Molecular cell*, 68(4):808–820, 2017.

[74] Daisuke Ohshima, Kyoko Arimoto-Matsuzaki, Taichiro Tomida, Mutsuhiro Takekawa, and Kazuhisa Ichikawa. Spatio-temporal dynamics and mechanisms of stress granule assembly. *PLoS computational biology*, 11(6):e1004326, 2015.

[75] Nancy Kedersha, Michael R Cho, Wei Li, Patrick W Yacono, Samantha Chen, Natalie Gilks, David E Golan, and Paul Anderson. Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *The Journal of cell biology*, 151(6):1257–1268, 2000.

[76] Sim Namkoong, Allison Ho, Yu Mi Woo, Hojoong Kwak, and Jun Hee Lee. Systematic characterization of stress-induced RNA granulation. *Molecular cell*, 70(1):175–187, 2018.

[77] Johannes H Wilbertz, Franka Voigt, Ivana Horvathova, Gregory Roth, Yinxiu Zhan, and Jeffrey A Chao. Single-molecule imaging of mRNA localization and regulation during the integrated stress response. *Molecular cell*, 73(5):946–958, 2019.

[78] Nancy Kedersha, Georg Stoecklin, Maranatha Ayodele, Patrick Yacono, Jens Lykke-Andersen, Marvin J Fritzler, Donalyn Scheuner, Randal J Kaufman, David E Golan, and

Paul Anderson. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol*, 169(6):871–884, 2005.

[79] Clare E Simpson, Jennifer Lui, Christopher J Kershaw, Paul FG Sims, and Mark P Ashe. mRNA localization to P-bodies in yeast is biphasic with many mRNAs captured in a late Bfr1p dependent wave. *J Cell Sci*, 127(6):1254–1262, 2014.

[80] Ashton Trey Belew, Arturas Meskauskas, Sharmishtha Musalgaonkar, Vivek M Advani, Sergey O Sulima, Wojciech K Kasprzak, Bruce A Shapiro, and Jonathan D Dinman. Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature*, 512(7514):265, 2014.

[81] Mi Ryoung Choi, Nam Jin Yoo, Chang Hyeok An, and Sug Hyung Lee. Frameshift mutations in mammalian target of rapamycin pathway genes and their regional heterogeneity in sporadic colorectal cancers. *Human pathology*, 46(5):753–760, 2015.

[82] Michael B Clark, Martina Jänicke, Undine Gottesbühren, Torsten Kleffmann, Michael Legge, Elizabeth S Poole, and Warren P Tate. Mammalian gene PEG10 expresses two reading frames by high efficiency–1 frameshifting in embryonic-associated tissues. *Journal of Biological Chemistry*, 282(52):37359–37369, 2007.

[83] Sezen Meydan, Dorota Klepacki, Subbulakshmi Karthikeyan, Tõnu Margus, Paul Thomas, John E Jones, Yousuf Khan, Joseph Briggs, Jonathan D Dinman, Nora Vázquez-Laslop, et al. Programmed ribosomal frameshifting generates a copper transporter and a copper chaperone from the same gene. *Molecular cell*, 65(2):207–219, 2017.

[84] John F Atkins, Gary Loughran, Pramod R Bhatt, Andrew E Firth, and Pavel V Baranov. Ribosomal frameshifting and transcriptional slippage: from genetic steganography and cryptography to adventitious use. *Nucleic acids research*, 44(15):7007–7078, 2016.

[85] Ian Brierley and Francisco J Dos Ramos. Programmed ribosomal frameshifting in HIV-1 and the SARS–CoV. *Virus research*, 119(1):29–42, 2006.

[86] Ian Brierley, Paul Digard, and Stephen C Inglis. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell*, 57(4):537–547, 1989.

[87] Neva Caliskan, Frank Peske, and Marina V Rodnina. Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. *Trends in biochemical sciences*, 40(5):265–274, 2015.

[88] Tony S Cardno, Yosuke Shimaki, Brad E Sleebs, Kurt Lackovic, John P Parisot, Rebecca M Moss, Caillan Crowe-McAuliffe, Suneeth F Mathew, Christina D Edgar, Torsten Kleffmann, et al. HIV-1 and human PEG10 frameshift elements are functionally distinct and distinguished by novel small molecule modulators. *PloS one*, 10(10):e0139036, 2015.

[89] Kathryn D Mouzakis, Andrew L Lang, Kirk A Vander Meulen, Preston D Easterday, and Samuel E Butcher. HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic acids research*, 41(3):1901–1913, 2012.

[90] Claudia Rato, Svetlana R Amirova, Declan G Bates, Ian Stansfield, and Heather M Wallace. Translational recoding as a feedback controller: systems approaches reveal polyamine-specific effects on the antizyme ribosomal frameshift. *Nucleic acids research*, 39(11):4587–4597, 2011.

[91] K Srinath Reddy. Global Burden of Disease Study 2015 provides GPS for global health 2030. *The Lancet*, 388(10053):1448–1449, 2016.

[92] Santiago Guerrero, Julien Batisse, Camille Libre, Serena Bernacchi, Roland Marquet, and Jean-Christophe Paillart. HIV-1 replication and the cellular eukaryotic translation apparatus. *Viruses*, 7(1):199–218, 2015.

[93] Tyler Jacks, Michael D Power, Frank R Masiarz, Paul A Luciw, Philip J Barr, and Harold E Varmus. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature*, 331(6153):280, 1988.

[94] Ian Brierley, Nicola J Rolley, Alison J Jenner, and Stephen C Inglis. Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *Journal of molecular biology*, 220(4):889–902, 1991.

[95] John D Lopinski, Jonathan D Dinman, and Jeremy A Bruenn. Kinetics of ribosomal pausing during programmed- 1 translational frameshifting. *Molecular and cellular biology*, 20(4):1095–1103, 2000.

[96] Harry Kontos, Sawsan Napthine, and Ian Brierley. Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Molecular and cellular biology*, 21(24):8657–8670, 2001.

[97] Guido Grentzmann, Jennifer A Ingram, Paul J Kelly, Raymond F Gestaland, and John F Atkins. A dual-luciferase reporter system for studying recoding signals. *RNA*, 4(4):479–486, 1998.

[98] Dominic Dulude, Martin Baril, and Léa Brakier-Gingras. Characterization of the frameshift stimulatory signal controlling a programmed -1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic acids research*, 30(23):5094–5102, 2002.

[99] Johanie Charbonneau, Karine Gendron, Gerardo Ferbeyre, and Léa Brakier-Gingras. The 5' UTR of HIV-1 full-length mRNA and the Tat viral protein modulate the programmed -1 ribosomal frameshift that generates HIV-1 enzymes. *RNA*, 18(3):519–529, 2012.

[100] Neva Caliskan, Vladimir I Katunin, Riccardo Belardinelli, Frank Peske, and Marina V Rodnina. Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell*, 157(7):1619–1631, 2014.

[101] Jin Chen, Alexey Petrov, Magnus Johansson, Albert Tsai, Seán E O'Leary, and Joseph D Puglisi. Dynamic pathways of -1 translational frameshifting. *Nature*, 512(7514):328, 2014.

[102] Shannon Yan, Jin-Der Wen, Carlos Bustamante, and Ignacio Tinoco Jr. Ribosome excursions during mrna translocation mediate broad branching of frameshift pathways. *Cell*, 160(5):870–881, 2015.

[103] Timothée Lionnet and Robert H Singer. Transcription goes digital. *EMBO reports*, 13(4):313–321, 2012.

[104] Kyuho Han, Ariel Jaimovich, Gautam Dey, Davide Ruggero, Oded Meyuhas, Nahum Sonenberg, and Tobias Meyer. Parallel measurement of dynamic changes in translation rates in single cells. *Nature methods*, 11(1):86, 2014.

[105] Joel Abramowitz, Dagoberto Grenet, Mariel Birnbaumer, Hector N Torres, and Lutz Birnbaumer. XLαs, the extra-long form of the α-subunit of the Gs G protein, is significantly longer than suspected, and so is its companion Alex. *Proceedings of the National Academy of Sciences*, 101(22):8366–8371, 2004.

[106] Cumhur Aydin, Nurgul Aytan, Mathew J Mahon, Hesham AW Tawfeek, Neil W Kowall, Alpaslan Dedeoglu, and Murat Bastepe. Extralarge XLαs (XXLαs), a variant of stimulatory G protein α-subunit (Gsα), is a distinct, membrane-anchored GNAS product that can mimic Gsα. *Endocrinology*, 150(8):3567–3575, 2009.

[107] Franka Voigt, Hui Zhang, Xianying A Cui, Désirée Triebold, Ai Xin Liu, Jan Eglinger, Eliza S Lee, Jeffrey A Chao, and Alexander F Palazzo. Single-molecule quantification of translation-dependent association of mRNAs with the endoplasmic reticulum. *Cell reports*, 21(13):3740–3753, 2017.

[108] Brian Munsky, Gregor Neuert, and Alexander Van Oudenaarden. Using gene expression noise to understand gene regulation. *Science*, 336(6078):183–187, 2012.

[109] Barbara Gorgoni, Luca Ciandrini, Matthew R McFarland, M Carmen Romano, and Ian Stansfield. Identification of the mRNA targets of tRNA-specific regulation using genome-wide simulation of translation. *Nucleic acids research*, 44(19):9231–9244, 2016.

[110] Paul M Sharp and Wen-Hsiung Li. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic acids research*, 15(3):1281–1295, 1987.

[111] Gene-Wei Li and X Sunney Xie. Central dogma at the single-molecule level in living cells. *Nature*, 475(7356):308, 2011.

[112] Marvin E Tanenbaum, Noam Stern-Ginossar, Jonathan S Weissman, and Ronald D Vale. Regulation of mRNA translation during mitosis. *Elife*, 4:e07957, 2015.

[113] Hiroshi Kimura, Yoko Hayashi-Takanaka, Timothy J Stasevich, and Yuko Sato. Visualizing posttranslational and epigenetic modifications of endogenous proteins in vivo. *Histochemistry and cell biology*, 144(2):101–109, 2015.

[114] Yoko Hayashi-Takanaka, Kazuo Yamagata, Teruhiko Wakayama, Timothy J Stasevich, Takashi Kainuma, Toshiki Tsurimoto, Makoto Tachibana, Yoichi Shinkai, Hitoshi Kurumizaka, Naohito Nozaki, et al. Tracking epigenetic histone modifications in single cells using fab-based live endogenous modification labeling. *Nucleic acids research*, 39(15):6475–6488, 2011.

[115] Alan Coulter and Rodney Harris. Simplified preparation of rabbit Fab fragments. *Journal of immunological methods*, 59(2):199–203, 1983.

[116] James S Huston, Douglas Levinson, Meredith Mudgett-Hunter, Mei-Sheng Tai, Jiří Novotnỳ, Michael N Margolies, Richard J Ridge, Robert E Bruccoleri, Edgar Haber, and Roberto Crea. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proceedings of the National Academy of Sciences*, 85(16):5879–5883, 1988.

[117] M Mariani, M Camagna, L Tarditi, and E Seccamani. A new enzymatic method to obtain high-yield F (ab) 2 suitable for clinical use from mouse IgGl. *Molecular immunology*, 28(1-2):69–77, 1991.

[118] Hongcheng Liu and Kimberly May. Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function. In *MAbs*, volume 4, pages 17–23. Taylor & Francis, 2012.

[119] Timothy J Stasevich, Yuko Sato, Naohito Nozaki, and Hiroshi Kimura. Quantifying histone and RNA polymerase II post-translational modification dynamics in mother and daughter cells. *Methods*, 70(2-3):77–88, 2014.

[120] Gorawit Yusakul, Seiichi Sakamoto, Benyakan Pongkitwitoon, Hiroyuki Tanaka, and Satoshi Morimoto. Effect of linker length between variable domains of single chain variable fragment antibody against daidzin on its reactivity. *Bioscience, biotechnology, and biochemistry*, 80(7):1306–1312, 2016.

[121] Yuko Sato, Masanori Mukai, Jun Ueda, Michiko Muraki, Timothy J Stasevich, Naoki Horikoshi, Tomoya Kujirai, Hiroaki Kita, Taisuke Kimura, Seiji Hira, et al. Genetically encoded system to track histone modification in vivo. *Scientific reports*, 3:2436, 2013.

[122] Yuko Sato, Tomoya Kujirai, Ritsuko Arai, Haruhiko Asakawa, Chizuru Ohtsuki, Naoki Horikoshi, Kazuo Yamagata, Jun Ueda, Takahiro Nagase, Tokuko Haraguchi, et al. A genetically encoded probe for live-cell imaging of H4K20 monomethylation. *Journal of molecular biology*, 428(20):3885–3902, 2016.

[123] Ning Zhao, Kouta Kamijo, Philip Fox, Haruka Oda, Tatsuya Morisaki, Yuko Sato, Hiroshi Kimura, and Timothy J Stasevich. A genetically encoded probe for imaging HA-tagged protein translation, localization, and dynamics in living cells and animals. *bioRxiv*, page 474668, 2018.

[124] Manish J Gandhi, Danielle M Carrick, Sarah Jenkins, Steven De Goey, Nancy A Ploeger, Gregory A Wilson, Jar How Lee, Jeffrey L Winters, James R Stubbs, Pearl Toy, et al. Lot-to-lot variability in HLA antibody screening using a multiplexed bead-based assay. *Transfusion*, 53(9):1940–1947, 2013.

[125] Yoko Hayashi-Takanaka, Kazuo Yamagata, Naohito Nozaki, and Hiroshi Kimura. Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. *The Journal of cell biology*, 187(6):781–790, 2009.

[126] Erik MM Manders, Hiroshi Kimura, and Peter R Cook. Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. *The Journal of cell biology*, 144(5):813–822, 1999.

[127] Paul L McNeil and Elizabeth Warder. Glass beads load macromolecules into living cells. *Journal of cell science*, 88(5):669–678, 1987.

[128] Pouyan E Boukany, Andrew Morss, Wei-ching Liao, Brian Henslee, HyunChul Jung, Xulang Zhang, Bo Yu, Xinmei Wang, Yun Wu, Lei Li, et al. Nanochannel electroporation delivers precise amounts of biomolecules into living cells. *Nature nanotechnology*, 6(11):747, 2011.

[129] Kai Wen Teng, Yuji Ishitsuka, Pin Ren, Yeoan Youn, Xiang Deng, Pinghua Ge, Sang Hak Lee, Andrew S Belmont, and Paul R Selvin. Labeling proteins inside living cells using external fluorophores for microscopy. *Elife*, 5:e20378, 2016.

[130] Salman F Banani, Allyson M Rice, William B Peeples, Yuan Lin, Saumya Jain, Roy Parker, and Michael K Rosen. Compositional control of phase-separated cellular bodies. *Cell*, 166(3):651–663, 2016.

[131] Xinlu Wang, Yifang Xuan, Yuling Han, Xiang Ding, Kai Ye, Fuquan Yang, Pu Gao, Stephen P Goff, and Guangxia Gao. Regulation of HIV-1 Gag-Pol expression by shiftless, an inhibitor of programmed -1 ribosomal frameshifting. *Cell*, 176(3):625–635, 2019.

[132] Chialing Tu, Tzy-Hwa Tzeng, and Jeremy A Bruenn. Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proceedings of the National Academy of Sciences*, 89(18):8636–8640, 1992.

[133] Brian Houck-Loomis, Michael A Durney, Carolina Salguero, Neelaabh Shankar, Julia M Nagle, Stephen P Goff, and Victoria M D'souza. An equilibrium-dependent retroviral mRNA switch regulates translational recoding. *Nature*, 480(7378):561, 2011.

[134] Sawsan Napthine, Roger Ling, Leanne K Finch, Joshua D Jones, Susanne Bell, Ian Brierley, and Andrew E Firth. Protein-directed ribosomal frameshifting temporally regulates gene expression. *Nature communications*, 8:15582, 2017.

[135] Claudio AP Joazeiro. Ribosomal stalling during translation: providing substrates for ribosome-associated protein quality control. *Annual review of cell and developmental biology*, 33:343–368, 2017.

[136] Szymon Juszkiewicz and Ramanujan S Hegde. Initiation of quality control during poly (A) translation requires site-specific ribosome ubiquitination. *Molecular cell*, 65(4):743–750, 2017.

[137] Szymon Juszkiewicz, Viswanathan Chandrasekaran, Zhewang Lin, Sebastian Kraatz, V Ramakrishnan, and Ramanujan S Hegde. ZNF598 is a quality control sensor of collided ribosomes. *Molecular cell*, 72(3):469–481, 2018.

[138] Michael D Moore and Wei-Shau Hu. HIV-1 RNA dimerization: It takes two to tango. *AIDS reviews*, 11(2):91, 2009.

[139] Sebastian Pechmann, Felix Willmund, and Judith Frydman. The ribosome as a hub for protein quality control. *Molecular cell*, 49(3):411–421, 2013.

[140] Jonathan D Dinman. Mechanisms and implications of programmed translational frameshifting. *Wiley Interdisciplinary Reviews: RNA*, 3(5):661–673, 2012.

[141] AE Firth, BW Jagger, HM Wise, CC Nelson, K Parsawar, NM Wills, S Napthine, JK Taubenberger, P Digard, and JF Atkins. Ribosomal frameshifting used in influenza a virus expression occurs within the sequence ucc_uuu_cgu and is in the+ 1 direction. *Open biology*, 2(10):120109, 2012.

[142] BW Jagger, HM Wise, JC Kash, K-A Walters, NM Wills, Y-L Xiao, RL Dunfee, LM Schwartzman, A Ozinsky, GL Bell, et al. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science*, 337(6091):199–204, 2012.

[143] Lenee Blanton, Vivien G Dugan, Anwar Isa Abd Elal, Noreen Alabi, John Barnes, Lynnette Brammer, Alicia P Budd, Erin Burns, Charisse N Cummings, Shikha Garg, et al. Update: InfluenzaActivityâĂŤUnited States, September 30, 2018–February 2, 2019. *Morbidity and Mortality Weekly Report*, 68(6):125, 2019.

[144] Martina M Yordanova, Cheng Wu, Dmitry E Andreev, Matthew S Sachs, and John F Atkins. A nascent peptide signal responsive to endogenous levels of polyamines acts to stimulate regulatory frameshifting on antizyme mRNA. *Journal of Biological Chemistry*, 290(29):17863–17878, 2015.

[145] Wai-Ki Ip, Paul B-S Lai, Navy L-Y Wong, Shirley M-H Sy, Ben Beheshti, Jeremy A Squire, and Nathalie Wong. Identification of PEG10 as a progression related biomarker for hepatocellular carcinoma. *Cancer letters*, 250(2):284–291, 2007.

[146] Shusuke Akamatsu, Alexander W Wyatt, Dong Lin, Summer Lysakowski, Fan Zhang, Soojin Kim, Charan Tse, Kendric Wang, Fan Mo, Anne Haegert, et al. The placental gene PEG10 promotes progression of neuroendocrine prostate cancer. *Cell reports*, 12(6):922–936, 2015.

[147] Sanne Boersma, Deepak Khuperkar, Bram MP Verhagen, Stijn Sonneveld, Jonathan B Grimm, Luke D Lavis, and Marvin E Tanenbaum. Multi-color single molecule imaging uncovers extensive heterogeneity in mRNA decoding. *Cell*, 2019.

[148] Michael G Kearse and Jeremy E Wilusz. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes & development*, 31(17):1717–1731, 2017.

[149] Mikhail M Pooggin and Lyubov A Ryabova. Ribosome shunting, polycistronic translation, and evasion of antiviral defenses in plant pararetroviruses and beyond. *Frontiers in microbiology*, 9:644, 2018.

[150] Kyle W Sherrill and Richard E Lloyd. Translation of cIAP2 mRNA is mediated exclusively by a stress-modulated ribosome shunt. *Molecular and cellular biology*, 28(6):2011–2022, 2008.

[151] Katherine E Sloan, Ahmed S Warda, Sunny Sharma, Karl-Dieter Entian, Denis LJ Lafontaine, and Markus T Bohnsack. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA biology*, 14(9):1138–1152, 2017.

[152] Ivanka Kamenova, Pooja Mukherjee, Sascha Conic, Florian Mueller, Farrah El-Saafin, Paul Bardot, Jean-Marie Garnier, Doulaye Dembele, Simona Capponi, HT Marc Timmers, et al. Co-translational assembly of mammalian nuclear multisubunit complexes. *Nature communications*, 10(1):1740, 2019.

[153] Anthony F Ross, Yuri Oleynikov, Edward H Kislauskis, Krishan L Taneja, and Robert H Singer. Characterization of a beta-actin mRNA zipcode-binding protein. *Molecular and cellular biology*, 17(4):2158–2165, 1997.

[154] Søren Lykke-Andersen and Torben Heick Jensen. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nature reviews Molecular cell biology*, 16(11):665, 2015.

[155] Francisco J Iborra, Alexandre E Escargueil, Kon Y Kwek, Alexandre Akoulitchev, and Peter R Cook. Molecular cross-talk between the transcription, translation, and nonsense-mediated decay machineries. *Journal of cell science*, 117(6):899–906, 2004.

[156] Yao-Fu Chang, J Saadi Imam, and Miles F Wilkinson. The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.*, 76:51–74, 2007.

[157] Maximilian Wei-Lin Popp and Lynne E Maquat. Organizing principles of mammalian nonsense-mediated mRNA decay. *Annual review of genetics*, 47:139–165, 2013.

[158] Luis C Antón and Jonathan W Yewdell. Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors. *Journal of leukocyte biology*, 95(4):551–562, 2014.

[159] Günter Blobel and David Sabatini. Dissociation of mammalian polyribosomes into subunits by puromycin. *Proceedings of the National Academy of Sciences*, 68(2):390–394, 1971.

[160] Jonathan R Warner, Alexander Rich, and Cecil E Hall. Electron microscope studies of ribosomal clusters synthesizing hemoglobin. *Science*, 138(3548):1399–1403, 1962.

[161] A Kent Christensen and Carol M Bourne. Shape of large bound polysomes in cultured fibroblasts and thyroid epithelial cells. *The Anatomical Record: An Official Publication of the American Association of Anatomists*, 255(2):116–129, 1999.

[162] Gabriella Viero, Lorenzo Lunelli, Andrea Passerini, Paolo Bianchini, Robert J Gilbert, Paola Bernabò, Toma Tebaldi, Alberto Diaspro, Cecilia Pederzolli, and Alessandro Quattrone. Three distinct ribosome assemblies modulated by translation are the building blocks of polysomes. *J Cell Biol*, 208(5):581–596, 2015.

[163] Daniel R Larson, Daniel Zenklusen, Bin Wu, Jeffrey A Chao, and Robert H Singer. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*, 332(6028):475–478, 2011.

[164] Hiroshi Kimura and Peter R Cook. Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *The Journal of cell biology*, 153(7):1341–1354, 2001.

[165] Florian Mueller, Tatsuya Morisaki, Davide Mazza, and James G McNally. Minimizing the impact of photoswitching of fluorescent proteins on FRAP analysis. *Biophysical Journal*, 102(7):1656–1665, 2012.

[166] Daniel R Larson, Robert H Singer, and Daniel Zenklusen. A single molecule view of gene expression. *Trends in cell biology*, 19(11):630–637, 2009.

[167] Ji Yu, Jie Xiao, Xiaojia Ren, Kaiqin Lao, and X Sunney Xie. Probing gene expression in live cells, one protein molecule at a time. *Science*, 311(5767):1600–1603, 2006.

[168] Makio Tokunaga, Naoko Imamoto, and Kumiko Sakata-Sogawa. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature methods*, 5(2):159, 2008.

[169] Brian P English and Robert H Singer. A three-camera imaging microscope for high-speed single-molecule tracking and super-resolution imaging in living cells. In *Biosensing and Nanomedicine VIII*, volume 9550, page 955008. International Society for Optics and Photonics, 2015.

[170] Arthur D Edelstein, Mark A Tsuchida, Nenad Amodaj, Henry Pinkard, Ronald D Vale, and Nico Stuurman. Advanced methods of microscope control using μmanager software. *Journal of biological methods*, 1(2), 2014.

[171] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676, 2012.

[172] Philippe Thevenaz, Urs E Ruttimann, and Michael Unser. A pyramid approach to subpixel registration based on intensity. *IEEE transactions on image processing*, 7(1):27–41, 1998.

[173] Arthur Edelstein, Nenad Amodaj, Karl Hoover, Ron Vale, and Nico Stuurman. Computer control of microscopes using μmanager. *Current protocols in molecular biology*, 92(1):14–20, 2010.

[174] Biosearch Technologies. Stellaris RNA FISH: Protocol for Sequential IF + FISH in Adherent Cells, 2015. https://biosearchassets.blob.core.windows.net/assets/bti_custom_stellaris_immunofluorescence_seq_protocol.pdf.

[175] Yasukazu Nakamura, Takashi Gojobori, and Toshimichi Ikemura. Codon usage tabulated from the international DNA sequence databases; its status 1999. *Nucleic acids research*, 27(1):292–292, 1999.

[176] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[177] Robert Hooke and Terry A Jeeves. "Direct Search" Solution of Numerical and Statistical Problems. *Journal of the ACM (JACM)*, 8(2):212–229, 1961.

Appendix A

Real-time Quantification of Single RNA Translation Dynamics in Living Cells

The following appendix chapter is based off of work where I was a co-author. [4]

A.1 Developing a method to visualize translation

To allow for the study of translation at the single-RNA level in real-time with fluorescence microscopy, we developed a system based on bright photostable small-molecule dyes, antibody-based probes, and multi-epitope protein tags. First, a plasmid encoding the large nuclear protein KDM5B, N-terminally tagged with a 10x FLAG tag (refer to as the spaghetti monster, SM) [48] and containing a 24x MS2 tag [27] in the 3' untranslated region (UTR) (Figure A.1A), was created to validate this system. The idea was that nascently translated FLAG SM tag permits high avidity for binding of fluorescently labeled fragments of antibody to FLAG (anti-FLAG Fab), whereas the MS2 stem-loop repeats allowed visualization of RNA with labeled MS2 coat protein (MCP) [48,49] (Figure A.1B). Transient transfection of this plasmid was done into U-2 OS cells that were subsequently bead-loaded with Cy3-labeled anti-FLAG Fab and Halo-tagged MCP [labeled with the far-red JF646 fluorophore [72]. Twenty-four hours after transfection, MCP marked RNA in the cytoplasm and Fab marked KDM5B in the nucleus, suggesting that neither the FLAG SM tag nor the presence of Fab interfered with RNA and KDM5B production and localization (Figure A.1C). To determine if Fab could mark nascently translated FLAG SM-tagged KDM5B, we imaged ~6 hours after transfection. At this time, Fab was diffuse suggesting that very little KDM5B had been

[4]The following appendix chapter is based on the following work [37], of which I was a co-author, with the following author contributions: T.M. and T.J.S. planned, analyzed, and carried out all experiments except for microinjections. T.J.S., T.M., and K.L. wrote the manuscript. K.L. helped with plasmid construction and transfections, cell culture, analysis, Western blotting, and figures. K.F.D. and J.G.D. performed microinjections. B.P.E. provided guidance for the multicolor single-molecule microscopy. Z.Z. purified Halo-MCP. L.L.L. and S.V. provided HA and FLAG SM constructs. L.D.L. and J.B.G. provided fluorophores. T.L. provided help with experimental design. See Appendix D for permission of use.

synthesized (Figure A.1D). Fab also colocalized and co-moved with many MCP-labeled RNAs in the cytoplasm (Figure A.1D and E). These co-moving spots displayed RNA-like diffusivity, were stably persisting for >2 hours and were present in the cytoplasm of cells that were both transfected with FLAG SM-tagged KDM5B and bead-loaded with Fab. To confirm if these co-moving protein-RNA spots were bona fide translation sites we treated cells with 50 μg/mL of puromycin, an inhibitor of translation that releases nascent chains from ribosomes [159]. Within minutes of drug addition, the number of co-moving spots decreased rapidly (Figure A.1F and G). By tracking the co-moving spots, we could see the disappearance of Fab-labeled protein, despite the persistence of the RNA (Figure A.1F and G).

KDM5B has a large coding region (1544 codons or amino acids) allowing more ribosomes to occupy the transcript and amplify the nascent chain signal, therefore translation should be relatively easy to detect. To test whether we could detect translation of smaller transcripts encoding smaller proteins, we constructed two plasmids (Figure A.2A) encoding either β-actin (374 amino acids) or the core histone H2B (125 amino acids). As with KDM5B, neither the FLAG SM tag nor Fab disturbed the localization of these proteins (Figure A.2B and C). Furthermore, we could again observe translation sites ~6 hours after transient transfection (Figure A.2D), indicating that this system is useful for visualizing translation of protein coding genes of varying sizes with fluorescence microscopy.

A.2 Biophysical characteristics of translation

To better understand the dynamics and biophysical characteristics of these single translation sites we tracked them using a technique we refer to as nascent chain tracking (NCT). NCT is based on a custom written *Mathematica* code, more details can be found in B.3.6. With NCT, we followed individual H2B, β-actin, and KDM5B translation sites for thousands of seconds in full cell volumes (300 timepoints x 2 colors x 13 z planes = 7800 images per movie). We lowered laser powers to focus exclusively on translation sites rather than on fully translated single protein products (which can interfere with tracking). As shown in the inset of Figure A.2D, this allowed us

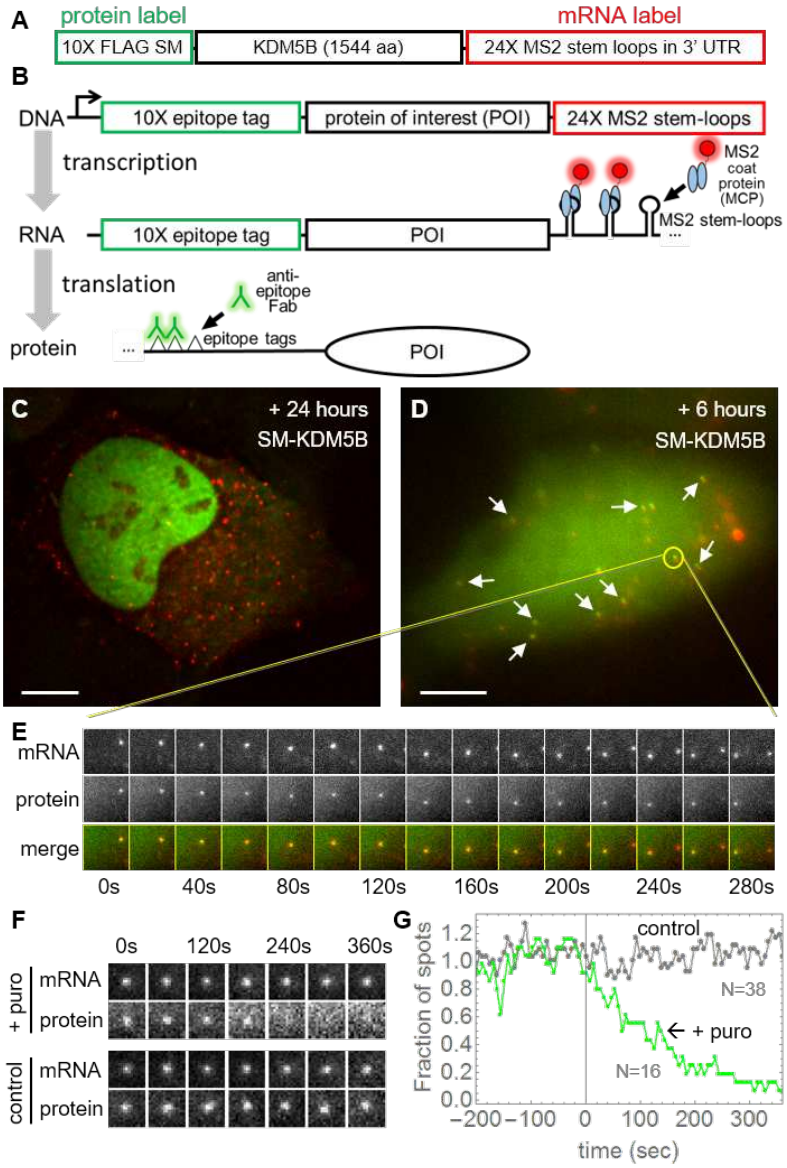


Figure A.1: A system for imaging single RNA translation kinetics. (A) Plasmid encodes KDM5B with an N-terminal 10x FLAG SM tag followed by a 24x MS2 tag in the 3' UTR (aa, amino acids). (B) Schematic of the system. RNA (red) is marked by MCP (labeled with JF646) that binds to repeated MS2 stem loops in the 3' UTR; protein (green) is labeled by Fab (conjugated to Cy3) that binds to peptide epitopes in the N terminus. (C) A deconvolved image showing that the FLAG SM-tagged KDM5B protein ('SM-KDM5B') localizes to the nucleus, whereas its mRNA localizes to the cytoplasm 24 hours after transfection and bead-loading with Fab and MCP. (D) Six hours after transfection, FLAG SM-tagged KDM5B colocalizes with mRNA in punctae (arrows). (E) Example co-movement of FLAG SM-tagged KDM5B and mRNA punctae circled in yellow in (D). (F) Addition of puromycin (50 μg/ml) leads to a loss of the FLAG SM-tagged KDM5B signal in punctae; this does not occur in control cells loaded with vehicle. (G) Quantification of the loss of the FLAG SM-tagged KDM5B signal from punctae (lower green curve) in a single cell as a function of time after addition of puromycin (which takes effect at 0 s), compared with the signal from a control cell (upper gray curve). Curves are normalized to pre-puromycin levels. Scale bars, 10 μm. Taken from [37]. K. Lyon contributed to the quantification of Figure A.1G

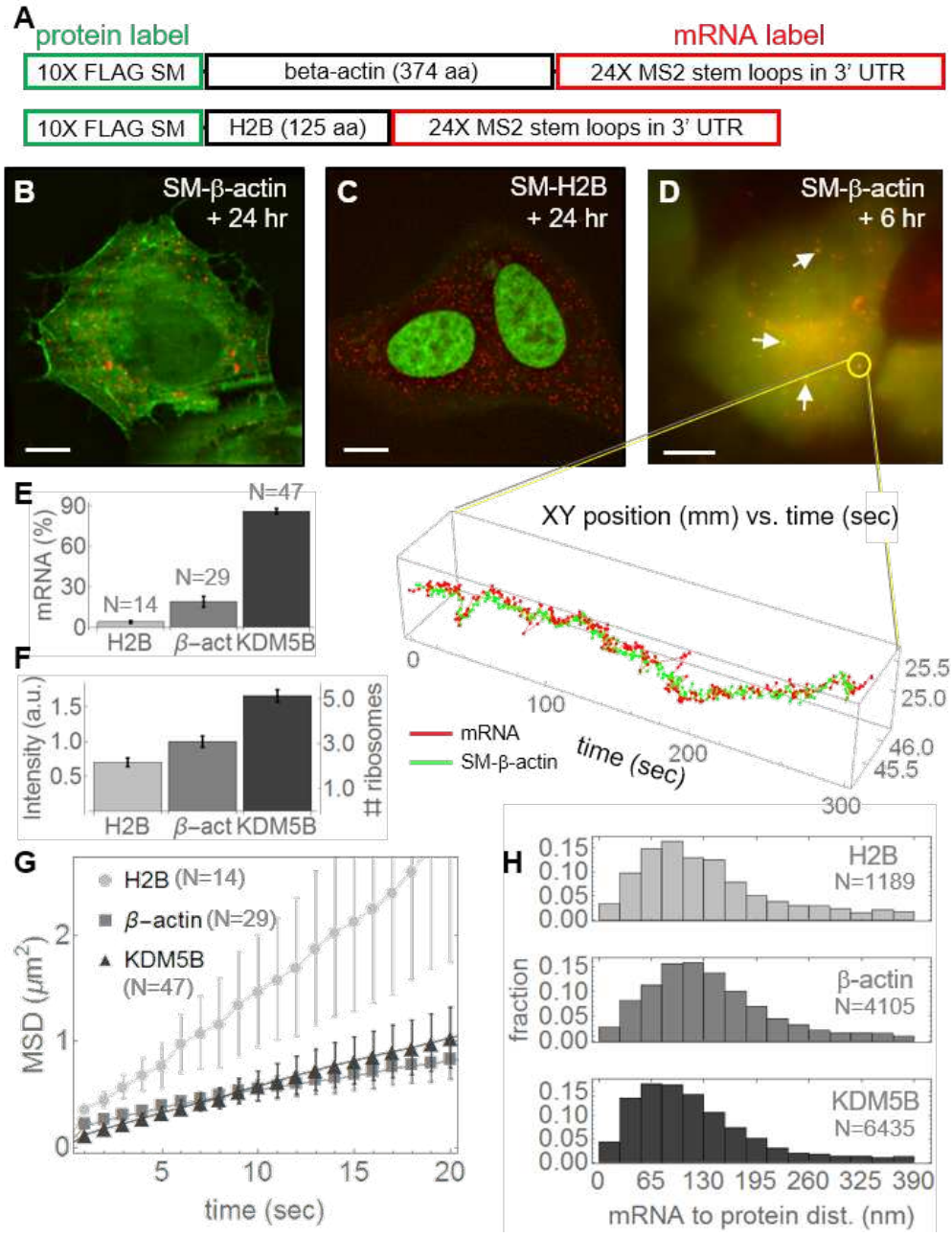


Figure A.2: Biophysical characteristics of canonical translation sites. (A) Plasmids encoding β-actin (SM-β-actin) and H2B (SM-H2B), analogous to the FLAG SM-tagged KDM5B construct in Figure A.1A. (B and C) Deconvolved images showing that SM-β-actin [(B), green] and SM-H2B [(C), green] localize to the cytoplasm and nucleus, respectively, whereas their RNAs (red) both localize to the cytoplasm 24 hours after transfection and bead-loading with Fab and MCP. (D) Six hours after transfection, SM-β-actin translation sites can be seen (arrows) and tracked (yellow circle and inset). (E and F) Quantification of the frequency (E) and intensity (F) of translation sites. Normalization to the number of ribosomes is shown on the right axis in (F) (a.u., arbitrary units). (G) Measured mean squared displacements (MSD) of tracked polysomes as a function of time. (H) Histograms of the Gaussian fit distances between RNA and protein in tracked polysomes. Error bars show S.E.M. Scale bars, 10 μm. Taken from [37]. K. Lyon prepared reagents, cloned constructs, and performed experiments contributing to Figure A.2C-H.

to accurately compare (i) the appearance frequency and brightness, (ii) the mobility, and (iii) the size of polysomes at translation sites. Of these parameters, frequency and brightness varied the most, tending to increase with construct length. We detected the translation of 86 $\pm$ 2% of KDM5B RNA but just 19 $\pm$ 4% of β-actin RNA and 4 $\pm$ 1% of H2B RNA (error, S.E.M.) (Figure A.2E). Furthermore, KDM5B translation sites, as marked by Fab, were over 1.5 times as bright as β-actin sites, which themselves were nearly 1.5 times as bright as H2B sites (Figure A.2F).

This difference of brightness at translation sites was most likely due to a difference in the number of nascent peptide chains per RNA (interpreted as number of ribosomes per RNA), as would be the case for larger polysomes on larger transcripts [160]. To determine the number of nascent chains that exist per site, we calibrated fluorescence by imaging a new β-actin plasmid containing a single 1x FLAG tag rather than the 10x FLAG SM tag (Figure A.3A). With this plasmid, only one Fab can be bound per peptide chain, allowing a direct comparison between translation site fluorescence and single Fab fluorescence. By imaging cells transfected with this plasmid at high laser powers, both single Fabs and translation sites could be detected and tracked (Figure A.3B and C), revealing translation sites to be on average 3.1 $\pm$ 0.5 times as bright as single Fabs (Figure A.3D). Therefore, we estimated 3.1 $\pm$ 0.5 nascent peptide chains per β-actin translation site, 2.1 $\pm$ 0.4 per H2B site, and 5.1 $\pm$ 0.9 per KDM5B site (Figure A.2F, right axis). Combining these data and assuming one ribosome per nascent chain, we conclude that detected canonical translation sites can contain as few as one ribosome every 900 RNA bases or as many as one ribosome every 200 RNA bases.

In addition to differences in their brightness, NCT also exposed differences in the mobility of translation sites. We quantified this by measuring the mean squared displacement (MSD) of tracked translation sites as a function of time. For the nuclear proteins H2B and KDM5B, mobility was modeled well by diffusion. Not only did MSD increase linearly with time for at least 20 s, the H2B translation sites with less ribosomes and smaller transcript size moved significantly faster than KDM5B translation sites that had more ribosomes and a larger transcript length (Figure A.2G). In contrast, β-actin translation sites displayed constrained diffusion, with jump sizes that were initially between those of KDM5B and H2B (consistent with ribosomal number and transcript length) but,

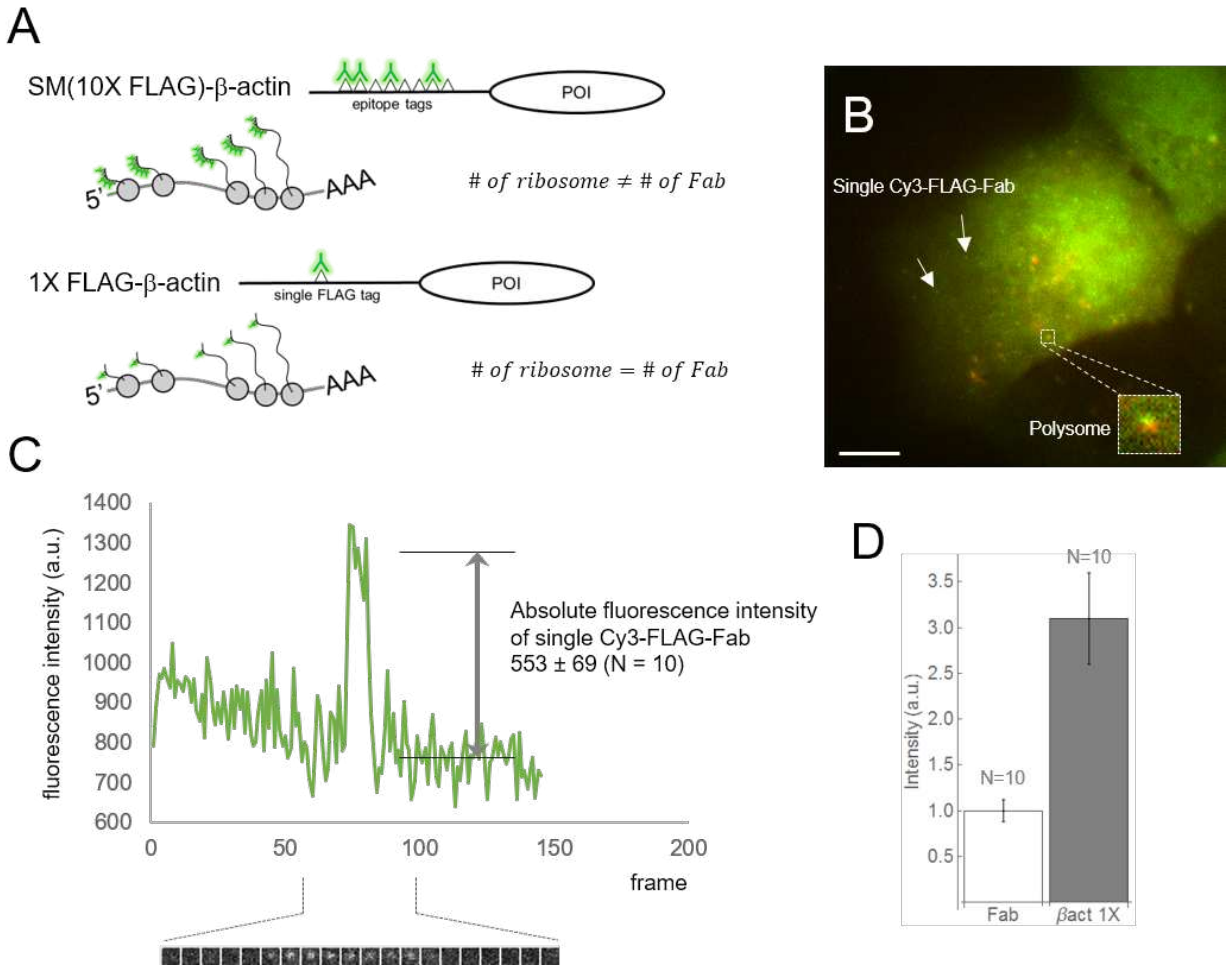


Figure A.3: Calibrating fluorescence intensity for experiments in Figure A.2F. (A) Schematic showing a 1x FLAG tag construct (bottom) designed to estimate the number of ribosomes (gray circles) per polysome since only one Fab (green 'Y' shapes) can bind per ribosome. This provides a lower-bound estimate for the number of ribosomes because the single FLAG sites may not be saturated by Fab. (B) When the 1x FLAG plasmid is transfected into cells bead-loaded with MCP (red, for labeling mRNA) and Fab (green, for labeling the 1x FLAG tag) and imaged at high laser powers, single Fab could be seen alongside polysomes (marked by arrows). (C) A representative time series showing the intensity of a single detected Fab through time. This time series shows the location of the spot prior to the single Fab appearing (frame ~70) and after the Fab photobleached in one step (frame ~78), indicating a single fluorophore. The height of the signal above background was used to calculate the average single Fab intensity. Bottom are images of the single Fab fluorescence shown in the time series. (D) The average single Fab intensity (white bar) compared to polysome intensity (gray bar) reveals approximately 3.1 $\pm$ 0.5 Fab per β-actin polysome, indicating 3.1 $\pm$ 0.5 ribosomes per polysome. Data are normalized to signal from single Fab. Error bars show S.E.M. Scale bar, 10 μm. Taken from [37]. K. Lyon contributed to the quantification of Figure A.3B-D.

ultimately lagged behind both at longer times. This constrained movement of β-actin could be due to interactions with cytoplasmic binding partners (other β-actin proteins or chaperones). Despite these trends, there was substantial variability in mobility between RNA, sometimes rapidly moving KDM5B translation sites (up to 6 $\mu m^2/s$) and nearly immobile H2B translation sites (~0.01 $\mu m^2/s$. This made it nearly impossible to identify a translated RNA based on mobility alone, implying that the translation machinery only weakly alters RNA movement in our system. Unlike their brightness and mobility, the size of polysomes at translation sites was less variable between constructs. To quantify sizes, we measured the distance between the 3' UTR of polysomal RNA (labeled with MCP) and the nascent peptide chains (labeled with Fab). The fluorescence from polysomes was within diffraction-limited spots, so we determined their mean positions with super resolution by using Gaussian fitting. According to hairpin models of polysome organization [161], this distance should grow as the length of the RNA unwinds, presumably while being translated. Instead, we found that the distance was shortest in KDM5B polysomes, typically around 65 nm, compared with 85 nm for H2B and 105 nm for β-actin (Figure A.2H). This suggests that the polysomes that we imaged at translation sites are organized in a globular shape rather than an elongated shape, consistent with recent atomic force microscopy images [162].

A.3 Quantifying translation site kinetics

Having measured the basic biophysical properties of translation sites, we next focused on translation kinetics. In particular, we wondered how the number of ribosomes at translation sites is controlled. This number reflects a balance between incoming and out going ribosomes and therefore depends on the ribosome elongation rate. One way to non-invasively estimate this rate is to examine the correlation of fluctuations in NCT data by means of fluorescence correlation spectroscopy (FCS), similar to how transcription elongation rates have been measured by using MS2 fluorescence fluctuations [163]. However, if Fabs bind nascent chains on the time scale of the elongation times their dynamics would contribute to the fluctuations and would distort the FCS measurements. To ensure this was not the case, we measured how quickly Fabs bind nascent chains by microinjecting

them into cells transfected 6 hours prior with our KDM5B construct and preloaded with MCP. Just 3 s after microinjection (as soon as we could image), many translation sites were labeled by Fab, implying that the binding time is less than 3 s (Figure A.4A and B). To measure the lifetime of Fab binding, we performed fluorescence recovery after photobleaching (FRAP) experiments in cells transfected with the H2B construct and bead-loaded with Fab and MCP 24 hours prior. We chose H2B, because it is known to remain bound for hours [164, 165], so any fluorescence recovery on the time scale of minutes would be exclusively due to Fab turnover. As Figure A.4C shows, there was little FRAP recovery in 4 min, implying that most Fabs are bound much longer. These binding kinetics (Figure A.4D) make Fabs ideal tools for measuring translation elongation times on time scales ranging from ~10 s to ~5 min.

Knowing the limits of Fab, we analyzed the fluorescence fluctuations of translation sites by extracting their intensity time series from NCT tracking data. We began with KDM5B translation sites, because these were the brightest and most numerous. The intensity of translation sites fluctuated with time (Figure A.4E), reflecting changes in the number of elongating ribosomes. From each intensity time series, the correlation curve was computed, details found in B.1.10, and averaged these together. There was significant heterogeneity among RNAs, but the average correlation curve revealed a clear linear drop to zero at 149 ± 20 s (Figure A.4F). In direct analogy to transcription correlation analyses [163], the time at which the correlation is zero marks the total elongation dwell time. To confirm this, we treated cells with 100 μg/ml of cycloheximide to stall translation. As expected, the correlation disappeared (Figure A.5). To corroborate these measurements, we performed FRAP on KDM5B translation sites. We intentionally photobleached a large section of the cytoplasm, where many translation sites were present (Figure A.6). By tuning the powers of the photobleaching laser, we could selectively photobleach just the Fab, leaving the RNA signals. This allowed us to monitor the fluorescence recovery of the relatively slower-moving translation sites. On average, translation sites recovered 80 to 90% of their fluorescence in 125 to 180 s, although there was again significant heterogeneity among RNAs, just as with FCS. Nevertheless, the average recovery time was on the same time scale as the FCS measurements. Given this consistency, we

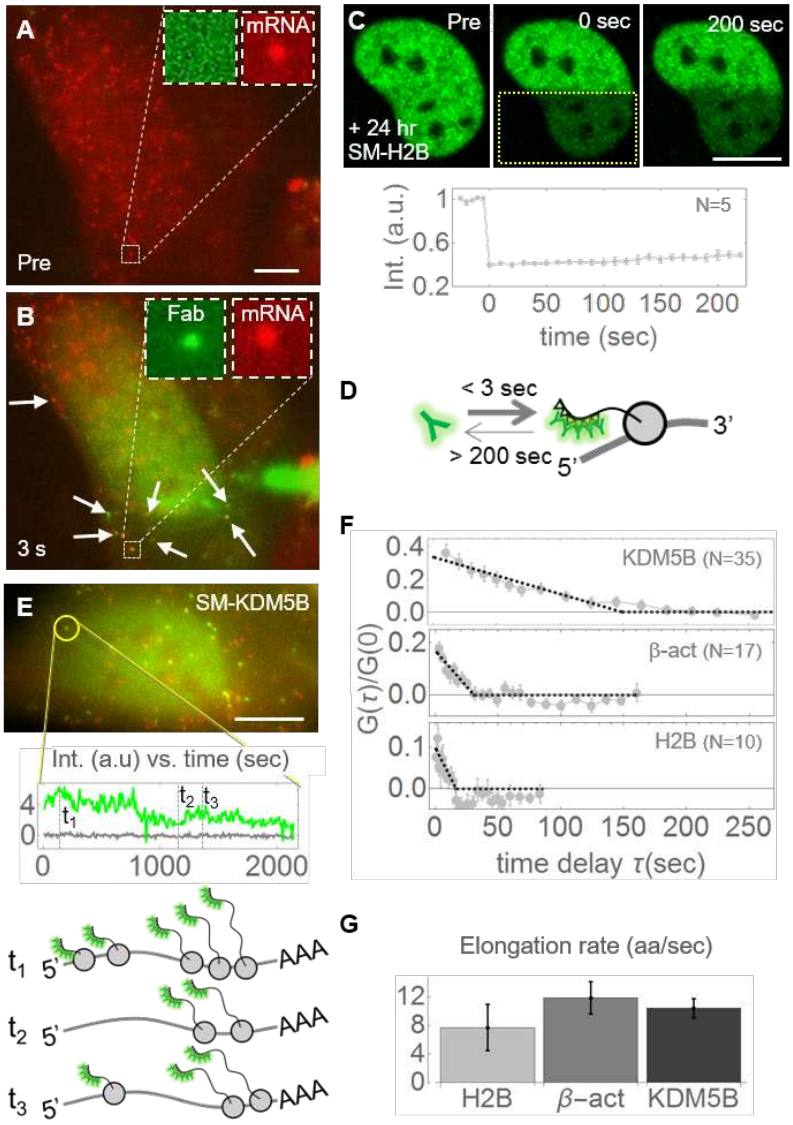


Figure A.4: Quantifying the kinetics of tracked translation sites. (A) Sample cell transfected with FLAG SM-tagged KDM5B and loaded with MCP before Fab microinjection. Many RNAs (red) can be seen. Their fluorescence does not bleed into the green channel (inset). (B) Three seconds after microinjection, Fabs (green) co-localize with mRNAs (arrows and inset). The site of microinjection can be seen on the right (bright green smear). (C) Sample FRAP experiment on a cell transfected with SM-H2B and bead-loaded with MCP and Fab 24 hours earlier. There is little recovery in 200 s (lower curve; int., intensity). Error bars show S.E.M. (D) A cartoon of results from (A) to (C) showing fast on rates and slow off rates for Fab (green 'Y' shapes) binding to SM epitopes (triangles) as they emerge from a ribosome (circle) translating RNA (thick line). (E) The intensity of a tracked FLAG SM-tagged KDM5B translation site (yellow circle and inset) can be measured as a function of time. The cartoon below shows how movement of ribosomes along RNA and the emergence of elongating peptide chains can produce intensity fluctuations at the indicated times t1, t2,and t3 (AAA, poly-A tail). (F) The average correlation curves calculated from intensity fluctuation data for each construct (error bars show S.E.M.; G, autocorrelation function). The time at which the correlation hits zero can be obtained from fits (dashed lines) to estimate the elongation dwell time. (G) Calculated elongation rates (amino acids per second) from fits in (F) (error bars show 95% confidence intervals). Scale bars, 10 μm. Taken from [37]. K. Lyon prepared reagents, performed experiments, and contributed to the quantification of Figure A.4A-D.

next performed FCS on the shorter β-actin and H2B constructs. Again, the correlation curves were linear and went to zero at distinct elongation dwell times. As expected, the dwell times decreased with transcript length, being 32 ± 9 s for β-actin and just 16 ± 7 s for H2B (Figure A.4F, lower panels). Importantly, for all constructs, the correlation remained zero at times greater than the dwell time. This implies that initiation is random, so there is no memory between initiation events, similar to transcription initiation [163] and in contrast to bursting [166, 167]. To calculate the elongation rate, we divided the length of the encoded protein by the elongation dwell time of each construct. The calculated rates were all within error (Figure A.4G), yielding a single consistent elongation rate of 10 ± 2.3 amino acids/s, which is fairly close to what has been measured for genome-wide ribosomal profiling (5.6 amino acids/s) [21, 22]. The difference could be due to differences in RNA codon usage of these specific transcripts [45].

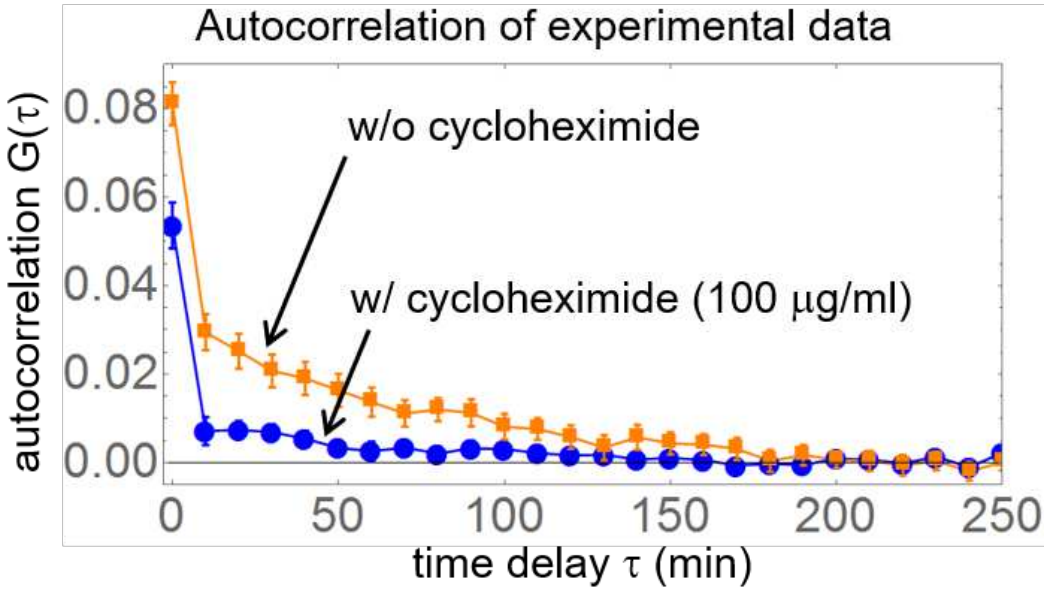


Figure A.5: Loss of correlation after treatment with cycloheximide. Cells expressing SM-KDM5B were treated with 100 μg/ml cycloheximide to stall elongation. This resulted in a loss in the autocorrelation seen in Figure A.4F. Taken from [37]. T. Morisaki designed experiments, performed experiments, and analyses for Figure A.5

With a consistent elongation rate, we can unify these observations. First, assuming KDM5B elongation occurs at 10 ± 2.3 amino acids/s, a new ribosome would have to initiate on average every 30 ± 9 s to maintain the measured 5.1 ± 0.9 ribosomes per translation site.

A.4 Translation sites can form multi-RNA sites

One advantage of using Fab to mark translation sites is the large number of high-affinity antibodies for multi-color applications. To demonstrate this, we generated new Fabs from hemagglutinin (HA) antibodies and labeled these with Alexa488 (A488) dye. In parallel, we engineered a new KDM5B construct with a 10x HA SM tag (HA-KDM5B) [48] to complement FLAG SM-tagged KDM5B (hereafter, FLAG-KDM5B), as shown in Figure A.7A. As a first application of this technology, we wanted to test whether translation sites can interact with each other to form higher-order structures that can translate two distinct RNAs at the same time. For this, we cotransfected cells with HA- and FLAG-KDM5B and bead-loaded them with MCP and anti-HA and anti-FLAG Fab.

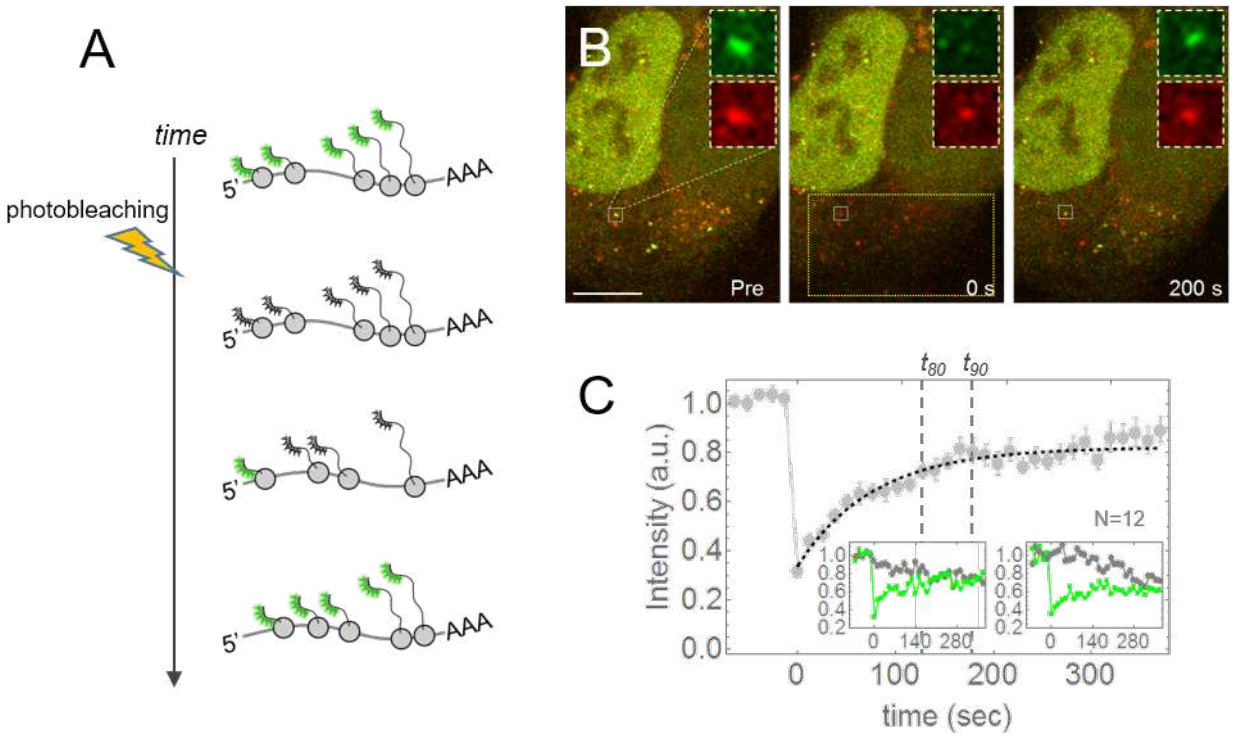


Figure A.6: Fluorescence recovery after photobleach of SM-KDM5B translation sites. (A) Cartoon depicting a FRAP experiment on SM-KDM5B. Initially Fab (green 'Y' shapes) fully label the nascent chains emerging from the ribosomes (gray circles) that make up the translation site (RNA shown as thicker line). Upon photobleaching, translation site fluorescence only recovers to its original value when photobleached ribosomes are replaced by newly initiated ribosomes. (B) A sample FRAP experiment in cells transiently transfected with SM-KDM5B and bead-loaded with MCP (red) and Fab (green) 6 hours earlier. The FRAP laser power was adjusted to only bleach the Fab. The image on the left shows the cell before FRAP. The middle image shows the cell immediately after photobleaching the area enclosed by the yellow box. The right frame shows the cell 200 seconds after FRAP. Insets are zoomed images of the polysome within the indicated white box. Although the Fab fluorescence photobleached, it recovered within 200 seconds. (C) Quantification of FRAP experiments like the one shown in B. The fitted 80% and 90% recovery times are shown (t_80 and t_90 dashed lines). The polysome fluorescence recovery was normalized to the intensity of an unbleached translation site in the same cell (dark gray curves in the examples in the inset). Depending on the polysome, recoveries could be somewhat variable, as demonstrated by the faster (left) and slower (right) recoveries in the inset. Error bars represent S.E.M. Scale bar, 10 μm. Taken from [37]. T. Morisaki designed experiments, performed experiments, and analyses for Figure A.6

As anticipated, cotransfected cells contained two types of translation sites in equal fractions (Figure A.7B-E), one type labeled by anti-HA Fab (HA-KDM5B), and the other labeled by anti-FLAG Fab (FLAG-KDM5B) (Figure A.7B). For the most part, there was little interaction between the two, providing direct evidence that the vast majority of KDM5B translation sites act independently of one another. However, a small fraction (~5%) of KDM5B translation sites formed complexes that co-moved for hundreds of seconds (Figure A.7C and D) and that produced both HA- and FLAG-tagged nascent peptide chains. By measuring the distance between the nascent HA and FLAG chains, we found the complexes to be roughly twice the size of a single translation site polysome (Figure A.7F), suggesting that the component translation sites remain compartmentalized. These complexes could reflect a general strategy to assemble higher-order complexes co-translationally [41] or co-regulate the expression of two genes.

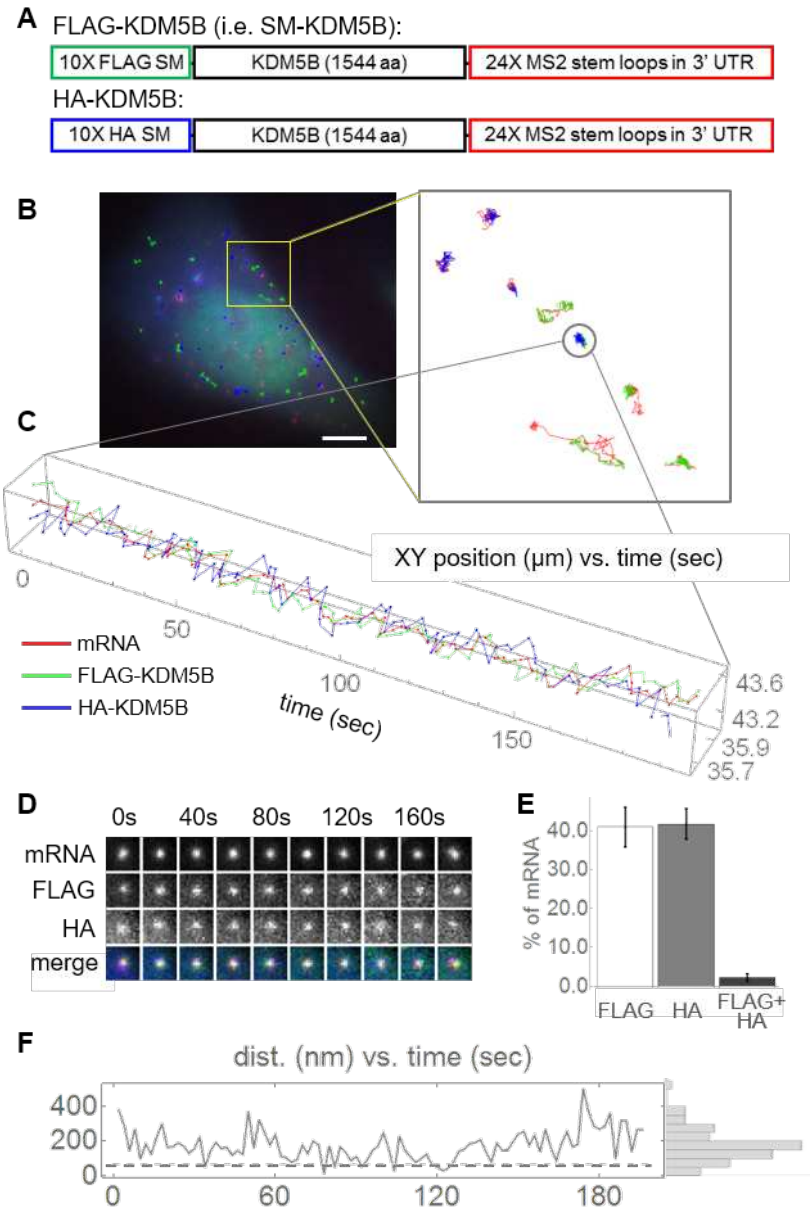


Figure A.7: Canonical translation sites can associate. (A) Complementary plasmids for imaging the translation of FLAG-KDM5B (green) and HA-KDM5B (blue). (B) Sample tracks from a cell expressing FLAG- and HA-KDM5B that was bead-loaded 3 hours earlier with MCP and anti-FLAG and anti-HA Fab. A translation site (circled) harboring both FLAG- and HA-KDM5B translation sites is tracked in (C). (D) The fluorescence from the spot in (C) (cropped and centered on mRNA) shows strong spatial overlap of FLAG- and HA-KDM5B. (E) The percentage of FLAG- and HA- KDM5B RNA single translation sites and the percentage of RNA in multi-RNA site. Error bars show S.E.M. (F) The Gaussian fit distance between the colocalized FLAG- and HA-KDM5B in the circled spot in (B) as a function of time. The distribution of these distances is shown on the right. It peaks at ~130 nm, twice the distance reported in Figure A.2H between KDM5B nascent chains and RNA in a single translation site (~65 nm, dashed line). Taken from [37]. K. Lyon prepared HA and FLAG Fab reagents contributing to Figure A.7B-F.

Appendix B

Experimental Material and Methods

B.1 Methods For Appendix A

B.1.1 Plasmid construction

The coding region of the spaghetti monster 10X FLAG-tag (SM) [48] was obtained by polymerase chain reaction (PCR) of a pCAG_mRuby2_smFP_FLAG (#59760, Addgene; mRuby2-based) template, using the following primers: 5'-GAG GAG GAG GCG GCC GCC ACC ATG GAC TAC AAG GAC GAC GAC GAC AAA GG-3'; 5'-CTC CTC CTC CTG CAG TGA ACC TCC TCC ACC TGA TCC ACC GCC TCC CTT ATC ATC ATC ATC CTT GTA ATC C-3'. The PCR product was flanked by NotI and PstI, and fused to the N terminal of beta-actin followed by the beta-actin zipcode and 24 MS2 stem loops in the 3' UTR [26] to obtain SM-beta-actin. The beta-actin coding region of SM-beta-actin was then digested out with PstI and NheI to construct SM-KDM5B and SM-H2B. For SM-KDM5B construction, the following oligo DNAs were annealed and inserted between PstI and NheI of digested SM-beta-actin to introduce AsiSI and PmeI sites: 5'-GGC GAT CGC CAT GGC CGT TTA AAC G-3'; 5'-CTA GCG TTT AAA CGG CCA TGG CGA TCG CCT GCA-3'. This product was then digested with AsiSI and PmeI, and replaced by the coding sequence of KDM5B acquired by digesting pFN21AAE9635 (Kazusa DNA Res. Inst.) with AsiSI and PmeI. For SM-H2B construction, the H2B coding region was PCR amplified from GFP-H2B (#11680, Addgene) using the following primers: 5'-GGG GCG GCC GCC ACC ATG CTG CAG CCA GAG CCA GCG AAG TCT GCT CCC G-3'; 5'-GGG GCT AGC CTA CAT ATG CTT AGC GCT GGT GTA CTT GGT GAT GGC CT-3'. The PCR product was flanked by PstI and NheI, and inserted between PstI and NheI of digested SM-beta-actin. HA-KDM5B was constructed via HA-beta-actin because of the availability of unique restriction enzyme sites. To obtain HA-beta-actin, the coding region of spaghetti monster 10X HA-tag was PCR amplified from pCAG_smFP_HA (#59759, Addgene, GFP-based) using the following primers: 5'-GGT TCG GCT TCT GGC GTG TGA

CC-3'; 5'-CTC CTC CTC CTG CAG TGA ACC TCC TCC ACC TGA TCC ACC GCC TCC AGC GTA GTC CGG GAC ATC GTA CGG GTA ACC G-3' and replaced into the coding region of 10X FLAG-tag of SM-beta-actin by using NotI and PstI. The beta-actin coding region was then replaced with the KDM5B coding region to construct HA-KDM5B following the same manner described above.

B.1.2 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 (Wako, 012-22384 Anti DYKDDDDK mouse IgG2b monoclonal) and HA (Sigma-Aldrich, H3663 HA-7 mouse IgG1 monoclonal; required clean-up with NAb Protein A column (Thermo Scientific)) antibodies 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 either Alexa Fluor 488 (Alexa488) or Cy3. Alexa488 tetrafluorophenyl ester (Invitrogen) or Cy3 N-hydroxysuccinimide ester (Invitrogen) were dissolved in DMSO and stored at -20 °C. 100 μg of Fab were diluted into 100 μl of 100 mM NaHCO3 (pH 8.5). 1.33 μl of Cy3 or 5.33 μl Alexa488 was added to this solution and incubated with end-over-end rotation for 1-2 hours at room 2 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 to dye ratio was calculated using the absorbance at 280 and 495 or 550 nm, and using the extinction coefficient of Fab with the dye correction factor at 280 nm provided by the manufacturers (0.11 or 0.08 for Alexa488 and Cy3, respectively). The degree of labeling was calculated using the following formula:

$$\mathrm{DOL} = \left(\frac{\epsilon_{\mathrm{protein}}}{\epsilon_{\mathrm{dye}}}\right)\left(\frac{1}{(\mathrm{A}_{\mathrm{ratio,measured}} - \mathrm{CF})^{-1}}\right) \tag{B.1}$$

Only Fab with DOL of ~1 were used in experiments.

B.1.3 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 $MgCl_2$, 100 mM KCl, 0.1 mM EDTA, 0.01 % NP-40 detergent, and 1 mM DTT) and stored at -80 °C after snapfreezing by liquid nitrogen.

B.1.4 Cell culture, transfection, and bead-loading

U-2 OS cells were grown in DMEM (Thermo Scientific) supplemented with 10% (v/v) FBS, 1 mM L-glutamine and 1% (v/v) Penicillin-streptomycin. Prior to experiments, cells were plated into a 35 mm MatTek chamber (MatTek) and transiently transfected using Lipofectamine 3000 (Thermo Scientific) according to the manufacturer's instructions. 2-3 hours after transfection, cells were bead-loaded with fluorescently labeled Fab and purified MCP-HaloTag protein as previously described [119, 127]. Briefly, 100 μg/ml of fluorescently labeled Fab and 33 μg/ml of purified MCP-HaloTag protein were prepared in 4 μl of PBS. After removing DMEM, this 4 μl solution was pipetted on top of the cells and ~106 μm glass beads (Sigma Aldrich) were then evenly distributed over the cells. The chamber was then tapped firmly 6 times, and DMEM was added back to the cells. 2 hours after bead-loading, the cells were washed three times with phenol-red-free complete DMEM to remove glass beads, and 200 nM of JF646-HaloTag ligand (a cell permeable fluorogenic ligand [72]) was added to label MCP-HaloTag protein. After 30 min of incubation, the cells were washed three times with phenol-red-free complete DMEM to remove the unliganded fluorophores. Cells were immediately imaged for translation experiments, whereas cells were incubated an additional 18 hours before imaging for testing localization of constructs and SM-H2B FRAP experiments.

B.1.5 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 [168] based on a previously described design [169]. 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 (iXon Ultra 888, Andor) cameras by focusing with 300 mm tube lenses (this lens combination produces 100X images with 130 nm/pixel). Live cells were placed into a custom-built incubation chamber at 37 °C and 5% CO2 on a piezoelectric stage (PZU-2150, Applied Scientific Instrumentation). The focus was maintained with the CRISP Autofocus System (CRISP-890, Applied Scientific Instrumentation). The lasers, the cameras, and the piezoelectric stage were synchronized by an Arduino UNO board (Arduino). Image acquisition was performed using open source Micro-Manager software [170]. For two-color single molecule tracking, the far-red signal from mRNA visualized with MCP-Halo-JF646 and the red signal from elongating protein visualized with Cy3-FLAG-Fab were simultaneously imaged by the two cameras without any time delay. Note that all colors we describe in the text and we show in the figures are based on the color of excitation lasers, namely mRNA in red (JF646) and protein in green (Cy3). Imaging size was set to 512 x 512 pixels^2 (66.6 x 66.6 μm^2), and exposure time was selected as 53.64 msec. The readout time of the cameras from the combination of our imaging size and the vertical shift speed we selected was 23.36 msec, resulting in our imaging rate of 13 Hz (70 msec per image). The excitation laser lines were digitally synched such that they illuminate the cells only when the camera is exposing in order to avoid any excess observational photobleaching. To capture the whole thickness of the cytoplasm of U2OS cells, 13 z-stack of step size 500 nm (6 μm in total) were imaged using the piezoelectric stage, resulting in our total cellular imaging rate of 1 Hz (1 sec per volume). Laser power was set to 160 μW and 340 μW at the back focal plane of the objective for 561 nm and 637 nm, respectively, throughout experiments. With this relatively weaker laser power setting, only signal from more than one SM could be detected which allowed us

to (1) avoid detecting already translated single protein products, (2) image only polysomes, and (3) reduce excess observational photobleaching. For single-FLAG experiments were performed with 1.5 mW of 561 nm to visualize single fluorophore molecules. KDM5B fluorescence correlation spectroscopy (FCS) experiments were carried out with 50 μW and 170 μW of 561 nm and 637 nm, respectively (this was possible because KDM5B polysomes are relatively bright).

For three-color experiments, the far-red signal of mRNA was imaged on one camera, and the red and the green signal of proteins visualized by Cy3-FLAG-Fab and Alexa488-HA-Fab, respectively, were imaged on the other camera. Note again that all colors described in the text and figures are based on the color of excitation lasers as mentioned above, namely mRNA in red (JF646), FLAG-protein in green (Cy3), and HA-protein in blue (Alexa488). Image acquisitions were performed with the same conditions described above, except that an additional 70 μW of 488 nm laser was used to excite Alexa488, and that Cy3 and Alexa488 signals were imaged alternatively. For this, the piezoelectric stage was moved to the next position every 2 images (Cy3 and Alexa488), resulting in our imaging rate of 0.5 Hz (2 sec per volume in 3 color). With these settings, there was only a 77 msec time delay between the alternating Cy3 and Alexa488 image acquisitions.

B.1.6 Particle tracking

Collected images were first pre-processed with Fiji [171]. Briefly, the 3D images were projected to 2D images by a maximum intensity projection, background subtracted, and corrected for observational photobleaching (using the built-in Bleach Correction tool). Pre-processed images were then analyzed by a custom-written Mathematica (Wolfram Research) routine to detect and track particles. Specifically, 3-frame rolling average movies were created and the averaged images were processed with a band-pass filter to accentuate particles so their positions could be detected using an appropriate threshold intensity value. Detected particles were linked through time by allowing a maximum displacement between consecutive frames of 3 pixels corresponding to 390 nm in length. Tracks lasting at least 25 seconds were selected and the precise coordinates and intensity of each particle were determined by fitting (using the built-in Mathematica routine NonlinearModelFit) the

original, pre-processed images to 2D Gaussians of the following form:

$$\mathrm{I}(\mathrm{x},\mathrm{y}) = \mathrm{I}_{\mathrm{BG}} + \mathrm{I}_{\mathrm{e}}^{-\frac{(\mathrm{x}-\mathrm{x}_0)^2}{2\sigma_{\mathrm{x}}^2} - \frac{(\mathrm{y}-\mathrm{y}_0)^2}{2\sigma_{\mathrm{y}}^2}} \tag{B.2}$$

where I_{BG} is the background fluorescence, I the particle intensity, and $(\mathrm{x}_0,\ \mathrm{y}_0)$ the particle location. For fitting, σ_{x} and σ_{y} were fixed at $\sigma_{\mathrm{x}} = \sigma_{\mathrm{y}} = 1.5$ pixels (195 nm), values determined by fitting diffraction limited spots acquired from images of 200 nm diameter beads (Tetraspeck, Molecular Probes/Thermo-Fisher) taken with the same imaging conditions. Fitted parameters were saved for each track for further analysis, along with 95% confidence intervals of the fitted parameters. The offset between the two cameras was registered using the built-in Mathematica routine FindGeometricTransform to find the transform function that best aligned the fitted positions of 200 nm diameter Tetraspeck beads evenly spread out across the image field-of-view. Note that we did not register the images, but only the fitted positions in order to avoid introducing any distortion into images. This is why a slight offset can be observed between the red and the green particles even though they are supposed to be in the same diffraction limited spots from our calculations. The only exceptions being the mRNA-centered cropped images showing particles in Figure A.1F and Figure A.7D, where the protein image has been corrected by the transformation function.

B.1.7 Puromycin treatment

Cells transiently-transfected with SM-KDM5B, bead-loaded with Cy3-conjugated α-FLAG Fab and MCP-HaloTag protein (labeled with JF646-HaloTag ligand) were imaged as described above, except that a 6.5 second interval was introduced every 13 frames (every one volume). After acquiring 10 time points of pre-treatment images, cells were treated with a final concentration of 50 μg/ml puromycin. After treatment, the cells were imaged under the same settings used for the pre-treatment imaging. For the control, the same experiments were performed except that the cells were treated with the vehicle of puromycin (H_2O). The number of detected polysomes (co-moving

SM-KDM5B and mRNA labeled by Fab and MCP, respectively) was then normalized by the average number of polysomes detected during pre-treatment and plotted as a function of time.

B.1.8 Micro-injection

Cy3 conjugated α-FLAG Fab were micro-injected into cells transiently-transfected with SM-KDM5B, bead-loaded with MCP-HaloTag protein alone (without Fab), and labeled with JF646-HaloTag ligand. Fab was diluted into Injection Buffer (IB: 10 mM Na2HPO4, pH 7.4, 100 mM KCl, and 1 mM MgCl2) and loaded into a final needle concentration of 0.1 mg/ml (Eppendorf, Femtotips II). Transiently-transfected and beadloaded cells were identified by their characteristic MCP punctae in the far-red channel. These were then micro-injected with a FemtoJet II Injector (Eppendorf) on an inverted Nikon TE-2000 microscope equipped with a 100X, NA 1.42 oil immersion objective (Nikon) and a Photometrics HQ2 camera (Roper Scientific) after acquiring a preinjecting image. Immediately after the injection, two-color (543 nm and 646 nm) timelapse imaging (at 1 Hz) was manually started. The earliest time point achievable was ~3 seconds after the injection.

B.1.9 Fluorescence recovery after photobleaching (FRAP)

For examination of koff of Cy3-FLAG-Fab, FRAP experiments were performed on cells transiently transfected with SM-H2B and bead-loaded with Cy3 conjugated α-FLAG Fab using a Zeiss LSM 880 confocal microscope equipped with a 63X, NA 1.40 oil immersion objective (Carl Zeiss). After acquiring 5 pre-bleach images (265 x 256 pixels^2; pixel size = 263 nm) every 5 seconds, half of the nucleus was photobleached using 100% 561 nm laser illumination. The fluorescence recovery in the photobleached region was then monitored for 50 frames under the same settings used for pre-bleach imaging. Cell movement was corrected using the StackReg ImageJ plugin [172]. After subtracting background and correcting for observational photobleaching, the average fluorescence intensity in the photobleached region was plotted as a function of time.

For examination of translation elongation, FRAP on SM-KDM5B polysomes was performed on the LSM880. Cells transiently-transfected with SM-KDM5B and beadloaded with Cy3-conjugated α-FLAG Fab and MCP-HaloTag protein (labeled with JF646-HaloTag ligand) were imaged every

12.8 sec in z-stacks containing 10 planes with a step size of 500 nm (265 x 256 pixels^2; voxel size = 166 x 166 x 500 nm). After acquiring 5 pre-bleach images, a fraction of the cytoplasm containing several translation sites was photobleached using 561 nm laser illumination, such that only the signal from Fab was photobleached, not mRNA. Subsequently cells were imaged for 30 frames (6.4 min) with the same settings used for pre-bleach imaging. Collected 3D images were then projected to 2D images by a maximum intensity projection. Translation sites were tracked manually following the mRNA signals and the fluorescence intensity of Cy3 was measured. The control translation sites that were not photobleached were also tracked, and their Cy3 intensities were used to correct observational photobleaching. This introduced some additional noise into our measurement, making it difficult to precisely determine the full recovery time. To approximate this time, we averaged the photobleach corrected curves and fit with a single exponential to estimate the 80% and 90% recovery times.

B.1.10 Fluorescence correlation spectroscopy (FCS)

SM-H2B, SM-beta-actin, and SM-KDM5B were transfected into cells that were subsequently bead loaded 3 hours later with MCP-HaloTag (labeled with JF646-HaloTag ligand) and Cy3-conjugated α-FLAG Fab. 3-6 hours after bead-loading, single polysomes were imaged in two colors (mRNA and protein) and tracked for 300 time points total at 1, 1/3, or 1/10 Hz, respectively, as described above. The mRNA intensity signal was stable and relatively easy to track, so its position was used as a starting guess for fitting the protein signal with a Gaussian. This allowed us to measure protein intensity in polysomes even if the intensity dropped below background levels. This resulted in an intensity time series that was stored for each tracked polysome. These were processed individually as follows: First, any detectable decay in polysome intensity due to unintentional photobleaching while imaging was corrected for by dividing out fits to an exponential decay. This yielded a new intensity time series $\mathrm{I(t)}$ without decay that was re-normalized to the average intensity of the original time series and stored. Second, the intensity fluctuations about the average intensity $\delta\mathrm{I(t)}$ are also stored. Third, the auto-correlation curves $\mathrm{G}(\tau)$ were generated from these time series

using the following formula:

$$\mathrm{G}(\tau) = \frac{\langle \delta \mathrm{I(t)} \delta \mathrm{I(t+\tau)} \rangle}{\langle \mathrm{I(t)} \rangle^2} \tag{B.3}$$

where the brackets represent an average of the time series data over the discrete time points t in the time series and τ is a discrete time shift in units of the discrete t (using the built-in Mathematica function ListCorrelate). Fourth, for fitting, data was binned so as not to overweight the end of the autocorrelation compared to the beginning (where most of the change occurs). Specifically, the first 8 data points in $\mathrm{G}(\tau)$ were not binned, points 9-24 were in bins of length two (i.e. (9,10), (11,12), (13,14), ..., (23,24)), time points 25-56 were in bins of length four (i.e. ((25,26,27,28), (29,30,31,32), ..., (53,54,55,66)), time points 57-120 were in bins of length 8 (i.e. (57, 58, 59, 60, 61, 62, 63, 64), (65, 66, 67, 68, 69, 70, 71, 72), ... (113, 114, 115, 116, 117, 118, 119, 120)), and so on. The average value of τ and $\mathrm{G}(\tau)$ in each bin was calculated and stored for subsequent fitting. To fit the elongation dwell time T (the average time it takes a ribosome to translate the protein) from the measured auto-correlation, the following formula was applied [163]:

$$\mathrm{G}(\tau) = \frac{\mathrm{T}-\tau}{\mathrm{cT}^2}\mathrm{H}(\mathrm{T}-\tau) \tag{B.4}$$

where c is the translation initiation rate and $\mathrm{H(t)}$ is the Heaviside step function equal to 1 for $\mathrm{t} >$ 0 and zero otherwise. This is an approximation of the exact correlation function that was derived in [163] to describe MCP intensity fluctuations due to transcription at a gene with repeated MS2 stem loops in the 5' UTR. The approximation is valid when the gene is long compared to the MS2 stem loop tag. Due to the analogies between our Fab-based peptide labeling system and the MS2-based RNA labeling system and between transcription and translation, the model is directly applicable to our experimental data. For shorter proteins, like H2B, the approximation is not as good because translation of the epitopes in the tag contribute more to $\mathrm{G}(\tau)$. Nevertheless, fits still provide

a reasonable upper bound on the elongation dwell time. Fitting was performed with Mathematica using the built-in function NonlinearModelFit. Error bars show the 95% confidence intervals.

B.2 Methods For Chapter 3

B.2.1 Cell culture

U-2 OS cells stably expressing the SG marker GFP-G3BP1 and the PB marker mRFP-DCP1a (a gift from the Kedersha lab [64]) were maintained in DMEM supplemented with 10% fetal bovine serum and 1% streptomycin/ penicillin. Cells harbouring the GFP-G3BP1 SG marker alone were isolated by cell sorting of the GFP-G3BP1 and mRFP-DCP1 mixed cell pool (BioFrontiers Institute Flow Cytometry Core). U-2 OS cells were validated by STR (short tandem repeat) profiling and morphological assessments.

B.2.2 Fab generation and dye conjugation

As described above, (B.1.2).

B.2.3 MCP purification

As described above, (B.1.3).

B.2.4 Nascent chain tracking

Nascent chain tracking was done essentially as described by Morisaki and colleagues [37]. Briefly, a plasmid expressing a reporter mRNA encoding the FLAG spaghetti monster tag N-terminal to H2B (Addgene plasmid no. 81082), KDM5B (Addgene plasmid no. 81084) or p300 ORFs, followed by the 3' UTR of beta-actin with 24x MS2 stem loops, was either transfected into cells 2 h before bead-loading or into cells together with purified Fab and/or MCP. Fluorescently labelled Fab (100 μg ml^{-1}) and/or purified MCP-HaloTag protein (33 μg ml^{-1}) in 4 μl PBS were bead loaded. Cells were incubated for 2-3 h in DMEM lacking phenol red with 10% fetal bovine serum, 1% streptomycin/penicillin and 1% glutamine before imaging. Cells were incubated for 30

min in the presence of JF646-HaloTag ligand to label the MCP-HaloTag protein and washed three times immediately before imaging.

B.2.5 Imaging Conditions

To track single mRNAs and their translation status with RNP granules, we used a custom-built widefield fluorescence microscope with a highly inclined illumination scheme [37, 168]. 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 (x60 numerical aperture 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 (iXon Ultra 888, Andor) cameras by focusing with 300 mm tube lenses (this lens combination produces 100x images with 130 nm pixel^{-1}).

Live cells were placed into an incubation chamber at 37 °C with 5% CO2 (Okolab) on a piezoelectric stage (PZU-2150, Applied Scientific Instrumentation). The focus was maintained with the CRISP Autofocus System (CRISP-890, Applied Scientific Instrumentation). The lasers, cameras and piezoelectric stage were synchronized by an Arduino Mega board. Image acquisition was performed using the open source Micro-Manager software (1.4.22) [170, 173]. The imaging size was set to 512 x 512 pixels^2 (66.6 x 66.6 μm^2) and the exposure time was selected as 53.64 ms. The readout time of the cameras from the combination of our imaging size and the vertical shift speed we selected was 23.36 ms, which resulted in an imaging rate of 13 Hz (70 ms per image). The excitation laser lines were digitally synced so that they illuminated the cells only when the camera was exposing, to avoid any excess observational photobleaching. The far-red signal of mRNA visualized by JF646-MCP was imaged on one camera and the red signal of translation (or PBs) visualized by Cy3-FLAG-Fab (or mRFP-DCP1) and the green signal of SGs visualized by GFP-G3BP1 were imaged on the other camera. Cy3/mRFP and GFP signals were imaged alternately. To minimize the bleed through, the JF646 signal was simultaneously imaged with the GFP signal. We also employed a filter wheel (HS-625 HSFW TTL, Finger Lakes Instrumentation) equipped with a filter for Cy3 and mRFP (593/46 nm BrightLine, Semrock), and for GFP (510/42

nm BrightLine, Semrock). The filter position was changed during the camera read-out time (23.36 ms) by the Arduino Mega board.

To capture the entire thickness of the cytoplasm of U-2 OS cells, 13 z stacks of step size 500 nm (6 μm in total) were imaged using the piezoelectric stage so that the z position changed every two images (Cy3 and GFP). This resulted in our total cellular imaging rate of 0.5 Hz (2 s per volume). For Figure 3.1b,e, the cells were first imaged at a single time point with 13 z stacks before arsenite addition (Time point 0). Arsenite was then added (0.5 mM) and after 10 min cells were imaged every 3 min with 13 z stacks per time point for up to 60 min. For Figure 3.1f, image acquisitions were performed with the same conditions described above, except that the cells were washed with fresh media ten times at 70 min post-stress. Ten minutes after washout, cells were imaged again every 3 min with 13 z stacks per time point for another 60 min. For all other figures, the cells were imaged continuously at 0.5 Hz with 13 z stacks per time point for 10 min beginning between 10 and 30 min post-stress (when SGs were visible). Note that all of the colours described in the text and figures are based on the colour of excitation lasers as mentioned above, namely mRNA in red (JF646), translation or PB in green (Cy3 or mRFP) and SG in blue (GFP). Following this colour scheme, purple corresponds to mRNA (red) in SGs (blue), whereas yellow corresponds to mRNA in PBs and brown corresponds to translating mRNAs in SGs.

B.2.6 Analysis of SG formation, mRNA localization and translation

Single particle detection was performed on XY maximum projections of the stacks to detect mRNA, translation and SGs, using custom Mathematica code (version 11.2.0.0). Briefly, for each image channel, a bandpass filter was used to highlight particles within a given size range and the resulting image was binarized using a reasonable threshold to create a mask of the cell in which the vast majority of mRNA, translation or SGs were detectable. The Mathematica command ComponentMeasurements was then used to measure the intensity-centroid position, area, total intensity and convex vertices of each masked object, either mRNA, translation or SGs. The precise coordinates (super-resolved locations) of mRNAs and nascent peptides were determined by fitting

(using the built-in Mathematica routine NonlinearModelFit) the original, pre-processed sub-images of the detected particles to two-dimensional Gaussians of the following form:

$$\mathrm{I}(\mathrm{x},\mathrm{y}) = \mathrm{I}_{\mathrm{BG}} + \mathrm{I}_{\mathrm{e}}^{-\frac{(\mathrm{x}-\mathrm{x}_0)^2}{2\sigma_{\mathrm{x}}^2}-\frac{(\mathrm{y}-\mathrm{y}_0)^2}{2\sigma_{\mathrm{y}}^2}} \tag{B.5}$$

where I_{BG} is the background fluorescence, I is the particle intensity, σ_{x}, σ_{y} are the spreads of the particle and x_0, y_0 represent the particle location. The offset between the two cameras was registered using the built-in Mathematica routine FindGeometricTransform to find the transform function that best aligned the fitted positions of 200 nm diameter Tetraspeck beads evenly spread out across the image field-of-view. Translating mRNAs were found by selecting mRNAs that colocalize with a translation signal location within 390 nm of the mRNA. Messenger RNAs within the SGs were found by selecting mRNAs within the detected SG-masked regions. The number of translating mRNAs in the cytosol or SGs, and the number of non-translating mRNAs in the cytosol or SGs were counted and divided by the total number of mRNAs at each time point to determine the fraction of each population throughout the time course during stress. The number of SGs was renormalized to the fraction of the maximum number of SGs detected in the cell at a single time point. As the timing of the beginning of SG formation varies from cell to cell, all of the the curves were aligned first before averaging the data from different cells. This alignment was done as follows. First, the SG formation curve was fit with the following equation:

$$\mathrm{N}_{\mathrm{SG}} = \mathrm{A} \times \tanh(\mathrm{B} \times \mathrm{t} - \mathrm{C}) \tag{B.6}$$

where NSG is the number of SGs, A, B and C are normalization factors, and t is the frame number. Then, from the fitted results, all of the single-cell curves were shifted so that their values at 10% of their maxima were aligned. These aligned curves were then averaged.

B.2.7 Analysis of mRNA interactions with SGs and PBs

The location and the area of mRNA, SGs and PBs were detected and tracked as described above. The effective radius was then calculated from the area as the radius of an equivalent-area circle. Although the effective radius is an acceptable approximation of the size of the object, it can be influenced by the brightness of the object as it is calculated from a binarized mask. Thus, small objects that are extremely bright (particularly PBs) may have effective radii that are a bit larger than they are in reality. The option 'CornerNeighbors âĘŠ False' was used to prevent the detection of any SGs or mRNAs that may have been clipped near the edge of the image frame. From these tracks and the binarized masks, the presence or absence of mRNA inside tracked SGs or PBs could be easily determined by the overlap of mRNA masks with the tracked SG or PB mask. The length of consecutive frames for which a tracked SG or PB mask had an overlapping mRNA mask was then set as equal to the length of the interaction time. Interactions that lasted for less than two frames (1 s) were ignored, as they could be influenced by imaging noise. If more than one mRNA mask overlapped with a SG or PB mask, it was counted as a single interaction (as most of the time these mRNA masks themselves overlapped, thus making it difficult to know how many mRNAs were actually present). This tends to slightly underestimate the number of long interactions. Finally, to validate our image analysis pipeline, we manually assessed the interaction times of individual KDM5B mRNAs in SGs and PBs using Fiji [171]. In all cases, manual measurements agreed well with the image analysis results in Mathematica.

To calculate the probability of an interaction lasting a certain amount of time (that is the survival probability of the interaction), we counted the fraction of tracks that lasted longer than N seconds, with N ranging from 6 to 300. A curve of the fraction of tracks versus time was then fit to a double exponential curve of the form $\mathrm{Ae}^{-\mathrm{t}/\mathrm{t}_1} + \mathrm{Be}^{-\mathrm{t}/\mathrm{t}_2}$. The data was then renormalized according to the fit so tt that at t = 0 s the fraction was 100% (that is A + B = 100). The renormalized data represents the survival probability versus time and is displayed in the plots in the main text along with the fitted fast interaction time t_1, slow interaction time t_2 and the slow fraction B/(A + B). The 90%

CI was calculated for the fitted values, from which the s.d. was estimated. From the s.d. and fitted mean values from each fit, effect size and P values between fits were calculated.

B.2.8 Analysis of MSD and 3D movement of mRNAs within SGs

The MSD was calculated using the tracked mRNA super-resolved locations for mRNA in the cytoplasm or in SGs. The MSD of mRNA inside SGs was calculated by dividing the MSD of the position of one mRNA with respect to the other by a factor of two. The MSD of SGs was calculated using the intensity-centroid of the SG. The 3D coordinates of mRNAs were determined by fitting (using the built-in Mathematica routine NonlinearModelFit) the original, pre-processed sub-image stacks of the detected particle to 3D Gaussians of the following form:

$$\mathrm{I(x,y)} = \mathrm{I_{BG}} + \mathrm{I_e}^{-\frac{(\mathrm{x}-\mathrm{x_0})^2}{2\sigma_\mathrm{x}^2} - \frac{(\mathrm{y}-\mathrm{y_0})^2}{2\sigma_\mathrm{y}^2} - \frac{(\mathrm{z}-\mathrm{z_0})^2}{2\sigma_\mathrm{z}^2}} \tag{B.7}$$

The convex vertices of SGs and the locations of mRNAs were plotted. For these analyses, mRNA spots were linked to the nearest neighbour mRNA in the consecutive frame. With the sparse mRNA density in the analysed SGs and the high temporal imaging frequency, it is unlikely that different mRNAs will be linked frame to frame. Additionally, the quality of automated tracking was validated by eye to ensure proper tracking.

B.2.9 Sequential immunofluorescence and smFISH

The protocol was performed as described previously [73]. Briefly, U-2 OS cells were seeded on sterilized coverslips in six-well tissue-culture plates. At ~80% confluency, the media was exchanged with fresh media 60 min before stress. Cells were then stressed with the addition of sodium arsenite (0.5 mM) for 60 min. Media was then removed and cells were washed with pre-warmed PBS. Cells were fixed for 10 min at room temperature with 500 μl 4% paraformaldehyde. After fixation, cells were washed twice with PBS, permeabilized in 0.1% Triton X-100 in PBS for 5 min and washed once with PBS. Coverslips were transferred to a humidifying chamber and cells were

incubated in 5 μg ml$_{-1}$ mouse α-G3BP1 antibody (ab56574, Abcam) in PBS for 60 min at room temperature. The coverslips were next transferred to a 6-well plate and washed three times with PBS. The coverslips were then transferred back to the humidifying chamber and incubated in goat α-mouse FITC-conjugated antibody in PBS (1:1,000 dilution; ab6785, Abcam) for 60 min at room temperature. The coverslips were transferred to 6-well plate and washed three times with PBS. Antibodies that were bound to cells were fixed on cells by incubating the coverslips with 500 μl 4% paraformaldehyde for 10 min at room temperature.

After immunofluorescence, smFISH was performed as described previously [73] which was adapted from the protocol provided by Biosearch Technologies website [174]. Biosearch Technologies Stellaris Buffers were also used (SMF-HB1-10, SMF-WA1-60 and SMF-WB1-20). Specific smFISH probes were created by Biosearch Technologies. The probes to the 5' end of the AHNAK mRNA and for NORAD were described in [57]. Imaging was performed using a widefield DeltaVision Elite microscope with a ÃŮ 100 objective and a PCO Edge sCMOS camera and softWoRx. The images shown in Figure 3.6a were deconvolved and maximum intensity projections made from at least 30 z stacks to span the entire cell using Fiji [171]. Figure 3.6b shows images from a single z plane.

B.2.10 Image analysis with Bitplane Imaris analysis software

Image analysis of SGs and smFISH spots was conducted using Bitplane Imaris analysis software (8.4.1) as described previously [73]. To measure the distances between smFISH spots to the nearest SG, smFISH diffraction limited spots and SGs were first identified by Imaris imaging software using the spot and cell component, respectively. Recognition of the smFISH diffraction limited spot by the spot component was determined using the following parameters: diameter, 0.2 μm and thresholding, manually determined for each image. Recognition of SGs (G3BP1 staining) by the cell component was determined using the following parameters: minimum size, $\geq$ 1 voxel) and thresholding, manually determined for each image. A measuring tool (measuring points) was then

manually applied to measure the distance between the centre of the spot and the surface of the SG. This was applied to 50 individual NORAD or AHNAK RNAs as described in Figure 3.6a.

B.2.11 Statistics and reproducibility

For live cell imaging, all experiments were performed with at least ten cells collected from three or four independent experiments to account for cell-to-cell variability. Images were discarded if they were of insufficient quality to accurately analyse. The number of images that were deemed to be of sufficient quality for inclusion in each analysis are described below. For the mRNA-SG interaction experiments: H2B, data were calculated from 492 tracked SGs from 11 cells collected from 3 independent experiments; KDM5B, data were calculated from 409 tracked SGs from 9 cells collected from 4 independent experiments; p300, data were calculated from 824 tracked SGs from 16 cells from 4 independent experiments. For mRNA-PB interaction experiments: H2B, data were calculated from 106 tracked PBs of 4 cells from 1 experiment; KDM5B, data were calculated from 137 tracked PBs of 7 cells collected from 3 independent experiments; p300, data were calculated from 244 tracked PBs of 16 cells collected from 4 independent experiments. For fixed cell experiments, each cell was considered as an independent biological replicate. The smFISH experiments were performed once to detect AHNAK and once to detect NORAD and the distance of 50 AHNAK or 50 NORAD RNAs from the nearest SG was determined in four cells each. The analysis of AHNAK 5' and 3' ends within SGs and outside SGs was performed using data from one experiment, with 153 mRNAs analysed in 14 cells.

For live-cell imaging experiments, the s.e.m. representing variation between individual cells was calculated, except for Figure 3.3a and Figure 3.5b. Figure 3.3a and Figure 3.5b were created using the fitted results from the plots in Figure 3.2c and Figure 3.5a, respectively. Each bar represents a fitted value and each error bar represents a 90% CI of the fit, calculated using Mathematica's built-in function NonlinearModelFit. The d.f. of the fits was 44 and 31 for Figure 3.3a and Figure 3.5b, respectively. The s.e.m. was calculated from the 90% confidence intervals (that is the 90% CIs were divided by 1.645), from which the two-sided t-test was performed and the P values were calculated.

B.3 Methods For Chapter 4

B.3.1 Plasmid construction

The HIV-1 frameshift sequence (FSS) followed by either the multi-frame (MF) tag or the reverse multi-frame (revMF) tag were synthesized by GeneArt gene synthesis service (Thermo Fisher Scientific). The gene fragments were flanked by NotI and NheI and fused upstream of the beta-actin zipcode and 24x MS2 stem loops in the 3' UTR of plasmid pUB_smFLAG_ActB_MS2 (Plasmid #81083, addgene) to obtain the FSS-MF and FSS-revMF, respectively. To double the MF tag, the MF tag region was digested out from FSS-MF with XbaI and AgeI, and then ligated into FSS-MF flanked with NheI and AgeI. The open reading frame encoding human XXLb1/AlexX [105, 106] was amplified from U-2 OS cells cDNA with the primers: 5'- GTT GTC ATA TGG GCG TGC GCA ACT -3'; 5'- GAT GTA GCT AGC CTA GAA GCA GCA GGC GGT G -3'. The amplified XXLb1/AlexX was flanked with NdeI and NheI, and then inserted into the C-terminal region of the FSS-MF, FSS-2xMF and FSS-revMF to obtain FSS-MF-AlexX (i.e. the +FSS MF tag), FSS-2xMF-AlexX (i.e. the +FSS 2xMF tag), and FSS-revMF-AlexX (i.e. the +FSS revMF tag), respectively. To produce smHA-FSS-MF-AlexX (i.e. the HA MF tag), the spaghetti monster HA (smHA) [48] was flanked with NotI and PstI, and then inserted into the N-terminal region of FSS-MF-AlexX. For the control constructs, FSS was removed using KpnI and XbaI. To keep the same frame for MF and AlexX, the following sequence was ligated between KpnI and XbaI to obtain MF-AlexX (i.e. the -FSS control tag): 5'- GGT ACC GGG AAT TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC GCA CCG TTT CTA GA -3'. To shift the -1 frame into 0 frame for MF and AlexX, the following sequence was ligated between KpnI and XbaI to obtain MF-AlexX (the -FSS(+1nt) tag): 5'- CGG GAA TTT TCT TCA GAG CAG ACC AGA GCC AAC AGC CGC ACC GTT CT -3'.

scFv-sfGFP was amplified from pHR-scFv-GCN4-sfGFP-GB1-dWPRE (Plasmid #60907, addgene) using primers: 5'- GCG CGC ATA TGA TGG GCC CCG ACA TC -3'; 5'- GCC GGA ATT CGC CGC CTT CGG TTA CCG TGA AGG T -3'. The amplified scFv-sfGFP was flanked with NdeI and EcoRI, and then inserted into a pET21 vector backbone for expression and purification from E.coli.

For HA MF tag experiments, the HA epitope encoded in the scFv-sfGFP plasmid (Plasmid #60907) from Addgene was removed by site-directed mutagenesis with QuikChange Lightning (Agilent Technologies) per the manufacturer's instruction using primers: 5'- CCT CCG CCT CCA CCA GCG TAA TCT GAA CTA GCG GTT CTG CCG CTG CTC ACG GTC ACC AGG GTG CCC -3'; 5'- GGG CAC CCT GGT GAC CGT GAG CAG CGG CAG AAC CGC TAG TTC AGA TTA CGC TGG TGG AGG CGG AGG -3'.

B.3.2 Fab generation and dye conjugation

As described above, (B.1.2).

B.3.3 MCP and scFv-sfGFP purification

His-tagged MCP or scFv-sfGFP was purified over a Ni-NTA-agarose (Qiagen) per the manufacturer's instructions with minor modifications. Briefly, bacteria were lysed in a PBS-based buffer with a complete set of protease inhibitors (Roche). Binding to the Ni-NTA resin was done in the presence of 10 mM imidazole. The resin was washed with 20 and 50 mM imidazole in PBS. The protein was then eluted in 300 mM imidazole in PBS. The eluted his-tagged MCP was dialyzed in a HEPES-based buffer (10% glycerol, 25 mM HEPES pH 7.9, 12.5 mM $MgCl_2$, 100 mM KCl, 0.1 mM EDTA, 0.01 % NP-40 detergent, and 1 mM DTT), snap-frozen in liquid nitrogen, and stored at -80C.

B.3.4 Cell culture, transfection, and bead-loading

U-2 OS cells were grown using DMEM (Thermo Scientific) supplemented with: 10% (v/v) FBS, 1 mM L-glutamine and 1% (v/v) Penicillin-streptomycin. Before experiments, cells were plated on a 35 mm MatTek chamber (MatTek) and DNA was either transiently transfected with Lipofectamine LTX (Thermo Scientific) per the manufacturer's instructions or transiently transfected via bead-loading. As described previously [25, 114], bead-loading involved the following six steps: First, 100 μg/ml of fluorescently labeled Fab, 250 μg/ml of purified (GCN4) scFv-sfGFP and 33 μg/ml of purified MCP-HaloTag protein were mixed with PBS to a final volume of 4 μl. Second, in a cell

culture hood, DMEM was aspirated from the MatTek chamber and the 4 μl mix was pipetted on top of cells. Third, ~106 μm glass beads (Sigma Aldrich) were evenly sprinkled over the cells. Fourth, the chamber was tapped carefully ~10 times on the cell culture hood bench top. Fifth, DMEM was immediately added back to the cells. Sixth, cells were returned to the incubator for at least an hour to recover from the loading procedure. In most experiments, we also bead-loaded DNA, which we added to the initial 4 μl mix (so that DNA had a final concentration close to 1 mg/ml). On occasion, DNA was transiently transfected ~2 hours before bead-loading. Around one hour before experiments began, bead-loaded cells were washed with phenol-red-free complete DMEM to remove glass beads, and 200 nM of JF646-HaloTag ligand (a cell permeable fluorogenic ligand [72]) was added to label MCP-HaloTag protein. After 30 mins of incubation, cells were washed three times using phenol-red-free complete DMEM to remove any unconjugated fluorophores. Cells were then immediately imaged for experiments.

B.3.5 Single molecule tracking microscopy

To track single molecule translation sites, a custom-built widefield fluorescence microscope based on a highly inclined illumination scheme [168] was used [37]. Briefly, the excitation beams, 488, 561 and 637 nm solid-state lasers (Vortran), were coupled and focused at the back focal plane of the objective lens (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 using two EM-CCD (iXon Ultra 888, Andor) cameras via focusing with 300 mm tube lenses (producing 100X images with 130 nm/pixel). With this setting, one camera detected far-red signals and the other detected either red or green signals. Far red signals were detected with the 637 nm laser and the 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. Live cells

were placed into a stage top incubator set to a temperature of 37°C and supplemented with 5% CO2 (Okolab) on a piezoelectric stage (PZU-2150, Applied Scientific Instrumentation). The focus was maintained using the CRISP Autofocus System (CRISP-890, Applied Scientific Instrumentation). The lasers, the cameras, the filter wheel, and the piezoelectric stage were synchronized via an Arduino Mega board (Arduino). Image acquisition was done with open source Micro-Manager software [170]. Imaging size was set to 512 x 512 pixels^2 (66.6 x 66.6 μm^2), and exposure time was set to 53.64 msec. Readout time for the cameras from the combination of our imaging size, readout mode, and the vertical shift speed was 23.36 msec, resulting in an imaging rate of 13 Hz (70 msec per image). The excitation laser lines were digitally synched to ensure they only illuminated cells when the camera was exposing in order to avoid excessive photobleaching. To capture the entire volume of the cytoplasm of U-2 OS cells, 13 z-stacks with step size of 500 nm (6 μm in total) were imaged using the piezoelectric stage such that the z-position changed every 2 images (one image for Cy3 and one for sfGFP/A488 + JF646). 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 sec 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/A488).

B.3.6 Particle tracking

Images were first pre-processed using either Fiji [171] or a custom-written batch processing Mathematica code (Wolfram Research) to make 2D maximum intensity projections from 3D images. Pre-processed images were then analyzed with a custom-written Mathematica code to detect and track particles. Specifically, particles were emphasized with a band-pass filter so the positions could be detected using the built-in Mathmatica 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 5-10 frames were selected and their precise coordinates were determined by fitting (using the built-in Mathematica routine NonlinearModelFit) the original 2D maximum intenstiy projected images to a 2D Gaussians of the following form:

$$\mathrm{I}(\mathrm{x},\mathrm{y}) = \mathrm{I}_{\mathrm{BG}} + \mathrm{I}_{\mathrm{e}}^{-\frac{(\mathrm{x}-\mathrm{x}_0)^2}{2\sigma_{\mathrm{x}}^2} - \frac{(\mathrm{y}-\mathrm{y}_0)^2}{2\sigma_{\mathrm{y}}^2}} \tag{B.8}$$

where I_{BG} is the background intensity, I the particle peak intensity, $(\sigma_{\mathrm{x}}, \sigma_{\mathrm{y}})$ the spread of the particle, and $(\mathrm{x}_0, \mathrm{y}_0)$ the particle location. The offset between the two cameras was registered using the built-in Mathematica routine FindGeometricTransform to find the transform function that best aligned the fitted positions of 100 nm diameter Tetraspeck beads evenly spread out across the image field-of-view. We did not register the images, but only the fitted positions in order to avoid introducing any distortion into images. This is why a slight offset can be observed between the red and the green/blue particles even though they are within a diffraction limited spot, according to our registration.

For visualization and some quantification, average intensity image trims were created by averaging images of all detected particles of a given species (each centered by their intensity centroid). To compute the average RNA signal intensity in a translation site (see Figure 4.9 and Figure 4.11), the average image trims of RNA were fitted to the equation above. The average intensities were calculated by integrating the fitted Gaussian. Fits were susceptible to noise, so we also used an alternative strategy to determine the average intensity of translation sites and RNA that was robust to noise (in all Figs aside from Figure 4.9 and Figure 4.11). Specifically, the intensity of centered images of RNA or translation sites was calculated to be the average intensity within a centered radial spot of four pixels in diameter minus the average background intensity from a centered ring with an outer diameter of twelve pixels and an inner diameter of eight pixels.

B.3.7 Calibrating translation site intensity

To calibrate the nascent chain intensity signals within translation sites to units of mature protein (i.e. to units that are roughly equivalent to the number of nascent chains or ribosomes within the translation site), two new calibration constructs were cloned. The calibration constructs are nearly identical in length and sequence to the original -FSS control construct (which lacks the frameshift

sequence), but each harbors just a single epitope in the 0 frame: either a single FLAG epitope or a single SunTag epitope. In this way, each nascent chain in a translation site is labeled by just one fluorophore (from a single anti-FLAG Fab or a single anti-SunTag scFv). Thus, the intensity of a translation site will be equal to the number of nascent chains multiplied by the intensity of a single fluorophore. Because the intensity of a single fluorophore can be unambiguously measured via the observation of single step photobleaching, it is possible to estimate the number of nascent chains (or ribosomes) per translation site, i.e. convert the intensity signal into units of mature protein, as shown in Figure 4.8.

To clone the calibration constructs, the FSS was removed from both the MF tag and the revMF tag using KpnI and XbaI digestion. Second, small oligos encoding a single SunTag epitope and single FLAG epitope were ligated between KpnI and XbaI of the MF tag and the revMF tag constructs, resulting in the 1XSunTag-MF and 1XFLAG-revMF constructs, respectively. The amplified XXLb1/AlexX was flanked with NdeI and NheI and then inserted into the C-terminal region of the 1XSunTag-MF and the 1XFLAG-revMF to obtain the 1XSunTag-MF-AlexX and 1XFLAG-revMF-AlexX calibration constructs. By design, these two constructs can be imaged in the same imaging session in separate dishes to directly compare 1X epitope (FLAG or SunTag) fluorescence to 12X epitope fluorescence in the MF or revMF tags (Figure 4.8A).

To calibrate translation site fluorescence intensity, cells transfected with the calibration constructs and bead-loaded with either anti-FLAG Fab (Cy3) or anti-SunTag scFv (sfGFP) were imaged in a single plane at high laser powers (50 mW for 561nm and 100mW for 488nm laser). A short movie was acquired, after which cells were continually imaged (without acquiring a movie) 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) from the 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 probe just prior to single-step photobleaching. Examples of this procedure are shown in Figure 4.8B,C. We performed calibration in this way to ensure at the beginning of the

movie the epitopes would be close to saturation. Had we started from a lower concentration of probe (although this would enable single probe tracking without the need for photobleaching), the epitopes in polysomes would be less saturated, which would lead to an underestimate of the polysome fluorescence. From these measurements, the average number of nascent chains (or ribosomes) in a translation sites can be estimated from the intensity ratio of polysomes to single probes. Assuming the two calibration tags harbor the same number of ribosomes per translation site, their intensity ratios provide the conversion of the intensity of the MF tag or revMF tag to units of mature protein (see Figure 4.8A).

B.3.8 Translation site species identification

After RNA particles were identified and tracked using the custom Mathematica code described above, an average centered image of the first five frames from each track was created for RNA (JF646), 0 ORF (FLAG-Cy3 in the +FSS MF tag), and -1 ORF (scFv-sfGFP in the +FSS MF tag). The trims were then hand checked to remove any trims with artifacts, e.g., smears or non-diffraction limited spots. Next, a custom Mathematica code was used to detect particles in the 0 ORF or -1 ORF trim channels, sorting the spots into RNA only, 0 frame translation sites (0 TS only), 0 and -1 TS, and -1 only TS. For all cases, RNA always had to be present. Finally, frameshifting translation sites (the 0 and -1 TS or the -1 only TS) were validated by eye, to further remove artifacts. For example, RNA that briefly colocalized with a mature protein punctae were removed at this stage. After all sites were validated, the total count of each type of species was used to determine the percentage of non-translating RNA (no TS), 0 frame translation sites (0 TS), 0 and -1 TS, and -1 only TS.

B.3.9 Puromycin treatment

To confirm active translation elongation, puromycin (Sigma Aldrich) was used to release nascent chains from elongating ribosomes, leading to a rapid loss of nascent chain signal at translation sites. Bead-loaded cells with visible translation sites were imaged at a rate of one volume every seven seconds. After acquiring 16 pre-images, cells were treated with a final concentration of 100 μg/ml puromycin and continuously imaged for an additional 100 time points. As a control, the

same imaging conditions were performed except that the cells were treated with vehicle (H2O). In this case, nascent chain signals did not disappear (data not shown). Both frameshifting and non-frameshifting translation sites were monitored through time using the tracking code described above.

B.3.10 Oligo FSS RNA co-transfection

Uncapped RNA oligos containing the FSS from HIV-1 (FSO), a scrambled form of the FSO sequence (Scr), or the boxB stem-loop sequence (BB) were synthesized from IDTDNA with the following RNA sequences, respectively: 5'- UUU UUU AGG GAA GAU CUG GCC UUC CCA CAA GGG AAG GCC AG -3', 5'- GAC GAA CUC AGG AUC GCC UUA GCG GAG UCU UAU UGA AUG GC -3', or 5'- AUU CCU GGG CCC UGA AGA AGG GCC CCU CGA CUA AGU CCA AC -3'. Co-transfection of FSS-MF-AlexX and the RNA oligo was carried out via bead-loading, as described above. Briefly, 1 μg plasmid FSS-MF-AlexX construct DNA, 100 μg/ml of fluorescently labeled Fab, 250 μg/ml of purified (GCN4) scFv-sfGFP or 1 μg of (GCN4) scFv-sfGFP plasmid, 1 or 4 μg of RNA oligo, and 33 μg/ml of purified MCP-HaloTag protein were mixed with PBS to a final volume of ~4 μl of PBS.

B.3.11 Ribosome run-off experiments and fits

To measure the average elongation rate, harringtonine (Cayman Chemical) was used to block translation initiation and iduce the run-off of all actively elongating ribosomes, leading to a gradual loss of nascent chain signal at translation sites. Bead-loaded cells with visible translation sites were imaged as described above except that cell volumes were acquired every 60 or 120 seconds. Laser powers were lowered (the 488 nm laser from 500 μW to 150 μW (3.3 times lower), the 561 nm laser from 900 μW to 500 μW (1.8 times lower), and the 647 nm laser from 1.5 mW to 1 mW (1.5 times lower)) to eliminate observable photobleaching of the green (Cy3 conjugated Fab) and blue (SunTag-scFv-sfGFP) channels. After acquiring 5 pre-images, cells were treated with a final concentration of 3 μg/ml Harringtonine and continuously imaged for 30 more time points. As a photobleaching control, cells with translation sites were imaged with the same image settings and

number of frames, revealing no loss of nascent chain signal (data not shown). After experiments, the intensities of translation sites were measured as described above. The intensity of all translation sites in each frame (and in all cells) were then totaled to produce the run-off curve. Run-off curves were normalized to the mean of the total intensity of the first four time points after treatment of harringtonine began. These curves were then fit to a linear regression to roughly estimate run-off times (see Figure 4.13). The linear portion of the run-off decay begins when the normalized run-off intensity reaches a fraction $\mathrm{f_0}$:

$$\mathrm{f_0} = \frac{\mathrm{L1}}{\mathrm{L1+L2/2}} \tag{B.9}$$

where L1 is the length of untagged portion of the open reading frame and L2 is the length of the tagged portion (i.e. the length of the repeated epitopes), as described [25]. The linear portion of the decay was then interpolated to background levels to estimate the run-off time $\mathrm{R_T}$. The elongation rate $\mathrm{k_{el}}$ was calculated as follows:

$$\mathrm{k_{el}} = \frac{\mathrm{L1+L2/2}}{\mathrm{R_T}} \tag{B.10}$$

B.3.12 Fluorescence recovery after photobleaching (FRAP)

To confirm elongation rates of frameshifting translation sites, FRAP experiments were performed. Bead-loaded cells with visible translation sites were imaged once every 10 seconds, with an intentional photobleach at frame 10. The photobleach was performed with a 405 nm laser focused to a spot roughly 5 μm in diameter and operating at a minimal power such that the nascent chain signal did not completely vanish. This allowed us to track the translation site continously throughout the experiment. Following the intentional photobleach, the fluorescence recovery of translation sites within the bleach zone were monitored for an additional 80 time points. These translation sites were tracked and their intensities quantified, as described above. To correct for unintentional photobleaching, the loss of signal from the control translation site was fit to a single exponential

decay and this decay was divided out from the FRAP recovery curves. The FRAP recovery curve can be thought of as the inverse of the harringtonine run-off curve [25]. In this way, the FRAP recovery curve was fit to determine the average translation elongation rate (Figure 4.14).

B.3.13 Statistical Analyses

For comparing cumulative distributions, we use the 2-sample KS test. For comparing mean values, we use the Mann Whitney U test. For fitted parameters, we use the fitted confidence intervals from Mathematica's built in NonlinearModelFit routine. In all figures, P-values are displayed as: * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$.

Appendix C

Frameshift Modeling

C.1 General aspects of modeling frameshift kinetics

The frameshift models describe the stochastic dynamics of nascent translation with single-codon resolution. Both the constitutive and bursting models are formulated to track an arbitrary number of individual ribosomes that can perform three possible reactions: (i) A new ribosome can initiate translation at the start codon at rate $\mathrm{k_{ini}}$. (ii) An existing ribosome can elongate at rate $\mathrm{k_{el}}$ to incorporate one amino acid into the nascent peptide chain. The rate $\mathrm{k_{el}}$ is sequence-specific with each codon's rate scaled by the genomic copy of the corresponding codon [175]. When the ribosome completes the final amino acid, translation is terminated and that ribosome is eliminated. (iii) A ribosome at the frameshift sequence can shift from the 0 to the -1 frame. In addition to these three reactions, ribosomes at the frameshift sequence can pause for an average time of $1/\mathrm{k^*_{FSS}}$ or $1/\mathrm{k_{FSS}}$ for the shifted or non-shifted states. For the bursting model, each RNA is assumed to switch back and forth between non-frameshifting and frameshifting states with rates $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$. With these mechanisms and parameters, the models can be analyzed using either simplified approximations or detailed simulations.

Several approximate features can be derived directly from the bursting model parameters (see Table 4.1). The average burst time or time spent in a frameshifting state is:

$$\langle \mathrm{b^*} \rangle = 1/\mathrm{k_{off}} \tag{C.1}$$

The average time spent in a non-frameshifting state is:

$$\langle \mathrm{b} \rangle = 1/\mathrm{k_{on}} \tag{C.2}$$

The fraction of RNA in a frameshifting state is:

$$\mathrm{f_s^*} = \mathrm{k_{on}}/(\mathrm{k_{on}} + \mathrm{k_{off}}) \tag{C.3}$$

The number of ribosomes that initiate during the frameshifting state is:

$$\mathrm{r_i^*} = \mathrm{k_{ini}}/\mathrm{k_{off}} \tag{C.4}$$

The number of ribosomes that initiate during the non-frameshifting state is:

$$\mathrm{r_i} = \mathrm{k_{ini}}/\mathrm{k_{on}} \tag{C.5}$$

The time for a ribosome to clear the FSS in the frameshifting state is:

$$\tau_{\mathrm{FSS}}^* = 1/\mathrm{k_{FSS}^*} + \delta \times \mathrm{r_e}/\mathrm{k_{el}} \tag{C.6}$$

where the second term accounts for ribosome pileup in the 9-codon ribosomal exclusion region [21], $\mathrm{r_e}$, upstream from the FSS. Similarly, the time for a ribosome to clear the FSS in the non-frameshifting state is:

$$\tau_{\mathrm{FSS}} = 1/\mathrm{k_{FSS}} + \delta \times \mathrm{r_e}/\mathrm{k_{el}} \tag{C.7}$$

One can approximate δ = 1 or 0 if initiation is faster $\mathrm{k_{ini}} > \mathrm{k_{FSS}}$ or slower $\mathrm{k_{FSS}} > \mathrm{k_{ini}}$ than the FSS pause.

The average elongation time for a ribosome in the frameshifting state is:

$$\langle \tau_{\mathrm{e}}^{*} \rangle = \mathrm{gl}/\mathrm{k_{el}} + \tau_{\mathrm{FSS}}^{*} \tag{C.8}$$

and the number to clear in a non-frameshifting state is:

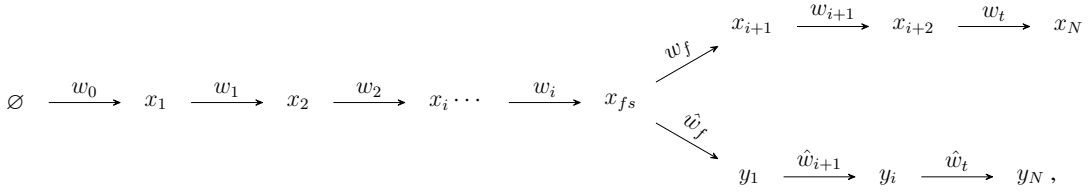
$$\langle \tau_{\mathrm{e}} \rangle = \mathrm{gl}/\mathrm{k_{el}} + \tau_{\mathrm{FSS}} \tag{C.9}$$

where gl is the gene length in codons.

Simulations were started at t = -10,000 seconds to approximate steady state at t = 0 using the Gillespie algorithm [176]. Ribosome densities were found by collecting position statistics for multiple simulations. Simulated ribosome numbers and positions and multi-frame tag probe locations were combined to estimate translation site intensities. Harringtonine assays were simulated by preventing the initiation reaction at the time of treatment. Parameter estimation was performed using genetic algorithms and a multiple-objective cost function that considers the frameshifting efficiency, the number of ribosomes per RNA and the Harringtonine assays. A detailed description of the computational methods and codes is given in the âĂŸComputational Details' section below.

C.2 Computational Details

Both the constitutive model and the bursting model consist of three general ribosomal reaction types: initiation $\mathrm{w_0}$, elongation $\mathrm{w_i}$, $\hat{\mathrm{w}}_{\mathrm{i}}$ and frameshifting $\mathrm{w_{FSS}}$, $\hat{\mathrm{w}}_{\mathrm{FSS}}$ as depicted in Eq. C.10,



(C.10)

C.2.1 Ribosome initiation

Ribosomes are assumed to bind to the RNA with rate $\mathrm{k_{ini}}$. To account for the fact that ribosomes are large biomolecules that occupy around 20 to 30 nuclear bases on the RNA [21], initiation is blocked by any downstream ribosome within $\mathrm{n_f}$= 9 codons (in either frame). Therefore, the initiation rate is set at:

$$w_0 = \begin{cases} k_{ini} & \text{if codons } \{1, \ldots, n_f\} \text{ unoccupied,} \\ 0 & \text{otherwise.} \end{cases} \tag{C.11}$$

C.2.2 Ribosome elongation

Ribosome Elongation: Each ribosome moves along the RNA codon by codon in the 5' to 3' direction. The elongation rate for each i^{th} codon, $\hat{k}_{el}$l is assumed to be:

$$w_i = \begin{cases} k_{el} \cdot (u_i/\bar{u}) & \text{if codons } \{i+1, \ldots, i+n_f\} \text{ unoccupied,} \\ 0 & \text{otherwise,} \end{cases} \tag{C.12}$$

where u_i denotes the codon usage frequency in the human genome obtained from [175], and $\bar{u}$ represents the average codon usage frequency in the human genome. The fit parameter k_{el} specifies the average elongation rate. Ribosomal termination is assumed to be equivalent to elongation of the final codon.

C.2.3 Frameshifting and pausing

When the ribosome reaches the frameshift site, n_{FSS}, it pauses, and may shift from the 0 frame to the -1 frame. For the constitutive model, the ribosome can continue in the 0 frame with rate

$$w_f = \begin{cases} k_{FSS} & \text{if codons } \{n_{FSS}+1, \ldots, n_{FSS}+n_f\} \text{ unoccupied,} \\ 0 & \text{otherwise,} \end{cases} \tag{C.13}$$

or the ribosome can continue in the -1 frame at rate:

$$\hat{w}_f = \begin{cases} k^*_{FSS} & \text{if codons } \{n_{FSS}+1, \ldots, n_{FSS}+n_f\} \text{ unoccupied,} \\ 0 & \text{otherwise.} \end{cases} \tag{C.14}$$

For the bursting model, the decision to continue in the 0 or -1 frame depends upon the frameshifting state with the rate given by:

$$\hat{w}_f = \begin{cases} k_{FSS} & \text{if OFF and codons } \{n_{FSS}+1, \ldots, n_{FSS}+n_f\} \text{ unoccupied,} \\ k^*_{FSS} & \text{if ON and codons } \{n_{FSS}+1, \ldots, n_{FSS}+n_f\} \text{ unoccupied,} \\ 0 & \text{otherwise.} \end{cases} \tag{C.15}$$

C.2.4 Relating model dynamics to experimental fluorescence intensity

To relate ribosome occupancy to experimental translation spots, we proscribe an assumed fluorescence to each ribosome based upon its position. This intensity is proportional to the number of peptide probes upstream from the ribosome location and in the appropriate frame or frames. Ribosomes in the 0 frame include all upstream probes in the 0 frame. Ribosomes in the -1 frame include probes in the 0 frame between the start codon and the FSS (if any) and probes in the -1 frame between the FSS and the current ribosome position. The total intensity vector for the j^{th} color is given by

$$I_j(t) = \sum_{i=1}^{N} c_j(R_i(t)), \tag{C.16}$$

where $R_i(t)$ is the position (i.e., frame and codon location) of the i^{th} ribosome, and $c_j(R_i)$ is the corresponding intensity of the j^{th} color.

C.2.5 Comparison of Data and Models

Fraction of translation spots

Experimental data was measured between 2 to 6 hours post bead-loading for a period of 120 seconds. To reproduce these experimental data, the model was solved using 2000 trajectories of 120 seconds starting at steady state. Spots were classified and their prevalence was reported as percentages for 0 frame only ($\mathrm{P_{0F}}$), -1 frame only ($\mathrm{P_{-1F}}$) and both frames ($\mathrm{P_{BF}}$).

Intensities of translation sites per RNA

Nascent chain fluorescence in single translation sites was measured in units of mature protein (u.m.p.), as described in Figure 4.8. Intensities were reported for translation spots containing -1 frame only ($\mathrm{I_{-1F}}$), 0 frame only ($\mathrm{I_{0F}}$), both frames in the green channel ($\mathrm{I_{BF/0}}$) and both frames in the blue channel ($\mathrm{I_{BF/-1}}$) (in Figure 4.3). In the bursting model, spots expressing both frames are those that have recently switched between 0 and -1 frame expression, but for which the ribosomes in the previous frame have not yet completed translation.

Ratio of intensities per RNA

To measure the average relative expression of frameshifted proteins, we calculated the ratio between the total intensity in the frameshifting and non-frameshifting translation spots, $\mathrm{R_{F:nFS}}$, as follows:

$$R_{F:nFS} = \frac{\sum_{i=1}^{N} I_{-1F} + \sum_{i=1}^{N} I_{BF/-1}}{\sum_{i=1}^{N} I_{0F} + \sum_{i=1}^{N} I_{BF/0}} \tag{C.17}$$

where $\mathrm{I_{\cdot}}$ denotes the steady state intensity in each frame of the $\mathrm{i^{th}}$ translation spot. Under an assumption that the 0 frame and -1 frame proteins have equal degradation rates the production ratio

$\mathrm{R_{F:nFS}}$ would be expected to match the average steady state ratio of mature protein in cells, which has previously been reported as 20:1 for the HIV Gag/Pol proteins [85, 89, 97, 98].

Harringtonine assays

Harringtonine inhibits translation by binding to the ribosomal 60S sub-unit, which blocks new initiation events. Experimental data showed that harringtonine causes the intensity in translating spots to drop to a basal intensity value after a run-off time (in Figure 4.12). To mimic the effects of harringtonine in our model, we modified the initiation rate as:

$$w_0 = \begin{cases} k_{ini} & \text{if } t < t_p \text{ and codons } \{1, \ldots, n_f\} \text{ unoccupied,} \\ 0 & \text{otherwise,} \end{cases} \tag{C.18}$$

where t_p is the time of application of Harringtonine. After Harringtonine application, spots simulated from the original construct were classified as 0 frame only ('0F') or both 0 frame and -1 frames ('BF'). After classification, average spot intensities were quantified as:

$$H_{0F}(t) = \langle I_{0F}(t) \rangle + b_{0F}, \tag{C.19}$$

$$H_{BF}(t) = \langle I_{BF}(t) \rangle + b_{BF}, \tag{C.20}$$

where $b_{(.)}$ is experimental background expression obtained at the end of the experimental time of the run-off assays. Similarly, spots simulated for the extended construct with upstream HA tags and downstream -1 tags were classified as HA only ('HA') or both HA and -1 ('HA/-1').

$$H_{HA}(t) = \langle I_{HA}(t) \rangle + b_{HA}, \tag{C.21}$$

$$H_{HA/-1}(t) = \langle I_{HA/-1}(t) \rangle + b_{HA/-1}. \tag{C.22}$$

Parameter estimation

For each model, we sought to find a single parameter set that reproduces all experimental data. Given the diversity of sources and types of experimental data, we estimated the parameter

values using a multi-objective optimization strategy to simultaneously compare the fractions of shifted, non-shifted, and transiently shifting spots, the spot intensities for both frames, ratio of spot intensities and the Harringtonine run-off data, as follows:

$$J(\theta) = \underbrace{w_P \sum_{f=1}^{F} |\log_{10}(\hat{P}_f(\theta)/\bar{P}_f)|}_{\text{fractions per frame}} + \underbrace{w_I \sum_{f=1}^{F} |\bar{I}_f - \hat{I}_f(\theta)|}_{\text{Intensities}} + \underbrace{w_R \sum_{f=1}^{F} |\bar{R}_{F:nFS} - \hat{R}_{F:nFS}(\theta)|}_{\text{Intensity Ratios}} + \underbrace{w_H \sum_{c=1}^{C} \sum_{f=1}^{F} \sqrt{\frac{\sum_{t=1}^{T} (\bar{H}_{t,f,c} - \hat{H}_{t,f,c}(\theta))^2}{T}}}_{\text{Harringtonine assays}}, \quad \text{(C.23)}$$

where experimental data are denoted as $\bar{x}$ and simulations results as $\hat{x}$. Weights (w_P, w_I, w_R, and w_H) were used to balance constraints by the different experiments and time points. T represents the number of experimental points in the Harringtonine assays, F is the number of studied frames, C is the number of different gene constructs. The first term in the objective function constrains the model to fit approximately to the 0 only, -1 only, and both frame fractions of translating spots and was expressed in terms of the sum of absolute log10 differences (i.e., fold changes rather than absolute differences) between the model and data, with weight $w_P = 1$. The weight used on the Harringtonine data set was defined as $w_H = 1/(C \times F)$. The weight used to compare the ratio of intensities was defined as $w_R = 1/\bar{R}_f$. The weight used to compare intensities was defined as: $w_I = \sum_{f=1}^{F} \bar{I}_f$.

Parameter searches

Parameter searches consisted of optimization routines based on the pattern search algorithm. Pattern search optimization is an iterative approach that directs the search of parameters by evaluating the effects of varying one parameter at a time in the objective function. The size of the variation and direction of the search are directed by the changes in the objective function [177].

The constitutive model has a total of four fitting parameters (k_{el}, k_{ini}, k_{FSS} and k^*_{FSS}). Although the bursting model has two additional parameters (k_{on}, k_{off}), the parameter k_{off} was directly determined from data in Figure 4.19B, and parameter k_{on} was included in the optimization routine, leading to five parameters total for that model. The parameter set that best reproduces the data was selected as:

$$\theta_{fit} = \underset{\theta}{argmin} J(\theta) \tag{C.24}$$

Optimized parameter values are given in full detail in Table 4.1 for the bursting model.

Simulation details

To simulate the model's stochastic dynamics, we used the direct method from Gillespie's algorithm [176] coded in MATLAB. The pattern search algorithm was used for parameter optimization [177]. Parameter uncertainty analyses for the bursty model were calculated by building parameter distributions that reproduce results within a 10% error, calculated from 10,000 independent simulations using randomly selected parameter values. Simulations were performed on the W. M. Keck High Performance Computing Cluster at Colorado State University.

Appendix D

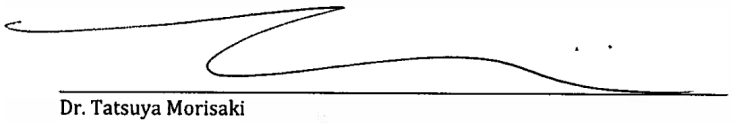
Permission of Use

D.1 Morisaki et al., 2016

2/08/19

Kenneth Lyon has my permission to include material published, of which I was a first author and he was a co-author, in his doctoral dissertation.

Morisaki, Tatsuya, et al. "Real-time quantification of single RNA translation dynamics in living cells." *Science* 352.6292 (2016): 1425-1429.



Dr. Tatsuya Morisaki

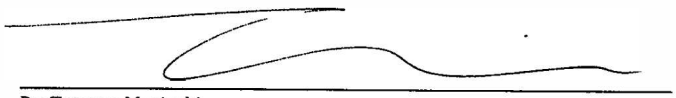
D.2 Moon and Morisaki et al., 2018

2/08/19

Kenneth Lyon has my permission to include material published, of which we were first authors and he was a co-author, in his doctoral dissertation.

Moon, Stephanie L., et al. "Multicolour single-molecule tracking of mRNA interactions with RNP granules." *Nature cell biology* (2019).

Dr. Stephanie Moon



Dr. Tatsuya Morisaki