

Visualizing mRNA quality control

Mechanistic insights from translational dynamics

COLOFON

The work described in this thesis was performed at the Hubrecht Institute for Developmental Biology and Stem Cell Research (Royal Netherlands Academy of Arts and Sciences, KNAW) within the framework of the research school Cancer, Stem Cells & Developmental Biology (CS&D), which is part of the Utrecht Graduate School of Life Sciences (Utrecht University). Tim Hoek was financially supported by a PhD fellowship from the Boehringer Ingelheim Fonds.

Print: Ridderprint | www.ridderprint.nl

Layout: Lotte Bruens

Cover design: Tim Hoek

Cover: a U2OS cell in which translation is visualised using the SunTag system. Green spots are translating mRNAs while red spots visualize mRNAs independent of their translation status.

ISBN: 978-94-6416-419-0

Copyright © 2021 by Tim Hoek. All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without prior permission of the author.

Visualizing mRNA quality control

Mechanistic insights from translational dynamics

mRNA kwaliteitscontrole visualiseren: Mechanistische inzichten verkregen uit de dynamiek van translatie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 9 maart 2021 des middags te 2.30 uur

door

Tim Andreas Hoek

geboren op 11 februari 1992 te Utrecht

PROMOTOR:

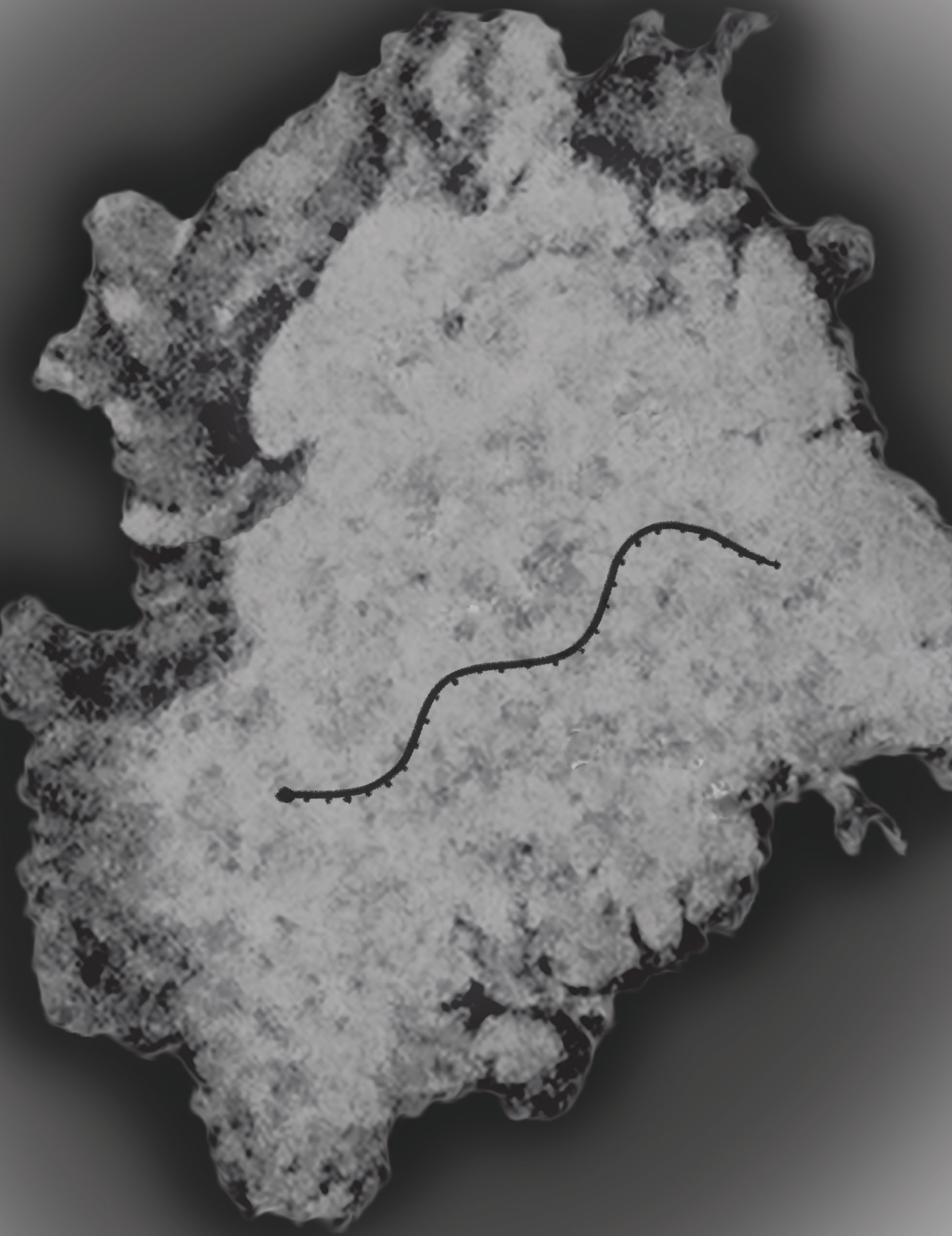
Prof. dr. W. L. de Laat

COPROMOTOR:

Dr. M. E. Tanenbaum

TABLE OF CONTENTS

Chapter 1	General introduction	7
Chapter 2	Dynamics of Translation of Single mRNA Molecules In Vivo	29
Chapter 3	Imaging Translation Dynamics of Single mRNA Molecules in Live Cells	61
Chapter 4	Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay	85
Chapter 5	Single-Molecule Imaging of Degradation of Non-Stop mRNAs by the Exosome	131
Chapter 6	Quantification of mRNA Translation in Live Cells using Single-Molecule Imaging	159
Chapter 7	An Optimized Procedure for High-Throughput RT-qPCR-Based COVID-19 Diagnostics	207
Chapter 8	Summarizing discussion	225
Addendum	Nederlandse samenvatting Curriculum vitae List of publications Acknowledgements	237



Chapter 1

General Introduction

THE CENTRAL DOGMA OF BIOLOGY

The central dogma of biology describes how the genetic information that is stored in the DNA is used in cells by copying the DNA into mRNA and then translating the mRNA into a protein. This flow of information is tightly controlled at multiple levels (Schwanhäusser et al., 2011), as it is of key importance that each gene is always expressed at the correct level. In addition, it is vital that all aspects of gene expression happen accurately to prevent detrimental effects of errors in gene expression. Gene expression consists of several steps that together both enable precise tuning of the expression level of each gene and ensure that errors in the DNA or RNA are not propagated into protein (Figure 1). First, the information of a DNA sequence is copied into a pre-mRNA by RNA polymerase II during transcription. Then, during processing, several modifications are made to the pre-mRNA to form a mature mRNA (Moore and Proudfoot, 2009): 1) A 7-methylguanylate cap is added to the 5' end of the mRNA, which protects the mRNA from degradation, 2) the non-coding regions of the pre-mRNA, or introns, are removed through splicing, and 3) a poly(A)-tail is added to the 3' end of the mRNA. The poly(A) tail is a sequence of ~150-250 adenine nucleotides that protects the mRNA from degradation and also stimulates translation of the mRNA (Eisen et al., 2020; Imataka, 1998). The processed mRNA is then exported from the nucleus and ready to be translated. During translation, the information present in an mRNA is read and translated into protein by ribosomes, and the proteins are then able to perform their function in the cell. Finally, both the mRNA and protein are degraded at some moment, which is necessary to decrease protein levels when expression of the gene needs to be reduced.

Over the years, a wealth of information has become available on how different steps of gene expression are regulated to ensure correct expression levels. For example, this has revealed how expression of genes can be activated through binding of transcription factors to the DNA (Tjian, 1978), how the genetic code is translated from a 4-letter nucleic acid code into a 20-letter amino acid code (Nirenberg and Matthaei, 1961), and how gene expression can be activated only in specific regions of a cell through localized translation (Holt and Schuman, 2013). Although it has been extensively studied how gene expression regulation ensures that genes are expressed at the right level, it is becoming increasingly apparent that there is also extensive regulation to ensure that genes are expressed without any errors. DNA quality control has been appreciated for a long time, and studies on DNA quality control have for example revealed how the accuracy of DNA replication is ensured by proofreading activity of DNA polymerase (Brutlag and Kornberg, 1972) and that double-strand breaks in the DNA are repaired to prevent DNA damage from propagating (Kanaar et al., 1998). Quality control also occurs at the level of mRNA, and various errors can be present in an mRNA that prevent protein production or impair protein function. Examples of defective mRNAs are mRNAs that

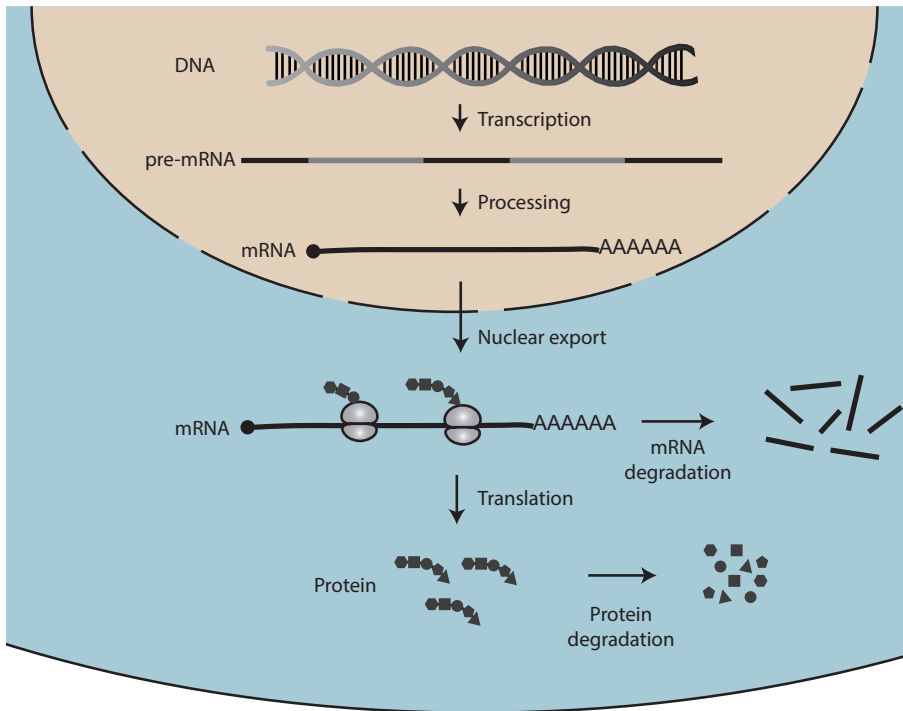


Figure 1. The central dogma of biology. Schematic overview of different steps of gene expression. The information in the DNA is first copied into a pre-mRNA molecule during transcription. The pre-mRNA is then processed into a mature mRNA and exported from the nucleus. Proteins are made by ribosomes that translate the mRNA. Finally, mRNAs and proteins are degraded.

do not have a stop codon and thus produce extended proteins (Frischmeyer et al., 2002; Van Hoof et al., 2002), mRNAs that have a premature stop codon and thus produce truncated proteins (He and Jacobson, 2015), or mRNAs with oxidized nucleotides that impair translation by ribosomes (Simms et al., 2014). These errors often result in activation of mRNA quality control mechanisms. Importantly, errors in the mRNA are not directly detected by quality control mechanisms. Instead, errors in mRNA are detected through the effect they have on ribosomes translating the mRNA, and mRNA quality control is therefore intrinsically coupled to translation. Understanding how defective mRNAs are detected thus requires understanding of how translation occurs normally, how translation is affected by errors in the mRNA, and how these changes in translation are detected by mRNA quality control mechanisms in order to degrade the defective mRNAs. In this introduction, I will describe the different steps of translation (initiation, elongation, and termination), describe how they are connected to mRNA quality control mechanisms, and explain why new methods are required to study the interplay between translation and mRNA quality control.

TRANSLATION INITIATION

Translation initiation refers to the assembly of a ribosome on an mRNA to begin translation. Translation initiation requires binding of the cytoplasmic cap-binding complex eIF4F, consisting of eIF4E, eIF4G, and eIF4A, to the 5' cap of an mRNA (Figure 2) (Sonneveld et al., 2020). eIF4F recruits a 43S pre-initiation complex (PIC) to the mRNA, which consists of a 40S small ribosomal subunit, a methionyl initiator tRNA, and eIF2-GTP. The PIC scans the 5'UTR of the mRNA until it encounters a start codon, which is recognized by the anticodon of the met-tRNA_i. Initiation does not always happen on the first start codon; instead, the likelihood of initiating on a start codon depends on how strongly the nucleotide sequence surrounding the start codon resemble the Kozak consensus sequence (gccrccAUGg), with a high resemblance resulting in efficient initiation (Kozak, 1990). When a start codon is recognized, GTP is hydrolysed into GDP, and additional factors bind the PIC to form the 48S-PIC. eIF2-GDP is then released, and a 60S large ribosomal subunit binds the 48S-PIC to form a 80S ribosome initiation complex that is ready to continue with translation elongation. Translation initiation is generally considered to be the rate-limiting step in translation (Shah et al., 2013) and is therefore the key step in determining the amount of protein that is produced from an mRNA. The rate of translation depends mainly on aspects of the 5' UTR, such as the presence of upstream open reading frames (uORFs) (Johnstone et al., 2016), secondary structures (Babendure et al., 2006), and the Kozak sequence (Kozak, 1990). Besides mRNA intrinsic regulation, translation initiation is also regulated at a cellular level. Examples of this are the global downregulation of translation that happens as a response to cellular stress (Liu and Qian, 2014), or a global upregulation of translation that has been observed in certain cancers and is associated with increased metastatic potential of these tumours (Ebright et al., 2020).

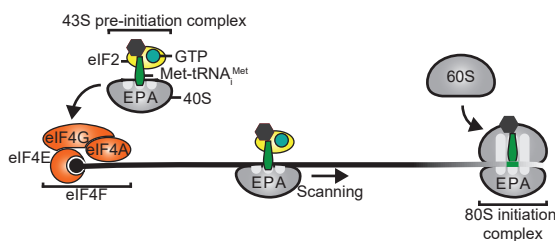


Figure 2. Translation initiation. Schematic overview of translation initiation. A 43S-preinitiation complex is recruited to the eIF4F complex at the 5' cap. The 43S complex scans the mRNA until it encounters a start codon (green horizontal bar). At the start codon, GTP is hydrolyzed into GDP, eIF2 releases from the ribosome, and a 60S ribosomal subunit binds to form a 80S initiation complex that is ready to begin translation. Figure adapted from (Sonneveld et al., 2020).

Although there are no mRNA quality mechanisms that detect errors during translation initiation, the efficiency of translation initiation does affect mRNA quality control processes indirectly as defective mRNAs are more efficiently degraded when they are actively translated.

TRANSLATION ELONGATION

Once translation initiation is completed, the ribosome is ready to start translation elongation, during which the nucleotide sequence of the mRNA is translated into a polypeptide sequence. Translation elongation starts when a ternary complex consisting of eEF1A, GTP and an aminoacyl-tRNA (AA-tRNA) binds in the empty A-site of the ribosome (Figure 3A) (Dever et al., 2018). When the 3-nucleotide anticodon of the tRNA is complementary to the codon in the A-site of the ribosome, GTP is hydrolysed by eEF1A, which results in release of eEF1A-GDP and accommodation of the AA-tRNA into the A-site of the ribosome. Next, a peptide bond is formed between the peptidyl-tRNA and the AA-tRNA, resulting in a peptide that has been extended by a single amino acid. The two tRNAs then move from the P and A site to the E and P sites, respectively, during the translocation step. Finally, the cycle is completed with the binding of a new aminoacyl-tRNA and simultaneous release of the E-site tRNA. This cycle then keeps repeating until a stop codon is encountered that indicates that the polypeptide is finished and that translation should be terminated.

Although translation elongation is a robust process, translation elongation does not always happen with the same speed. Instead, the efficiency by which amino acids are incorporated can vary between different mRNAs (Yan et al., 2016) or codons (Ingolia et al., 2009). In general, codons are translated faster than the average codon when their corresponding tRNA is abundant in the cell, and slower when their tRNA is scarce (Presnyak et al., 2015; Wu et al., 2019). Since multiple codons can often encode the same amino acid, the speed by which an amino acid is incorporated depends on which codon is used to encode it. mRNAs with a high ‘codon optimality’ are enriched for codons that are translated efficiently, and these mRNAs are therefore translated more rapidly than mRNAs with low codon optimality. In addition, the speed of translation elongation can also be affected by the polypeptide that is produced. For example, translation can be slowed down when long stretches of lysines are present in the ribosomal exit tunnel (Arthur et al., 2015; Koutmou et al., 2015), or through binding of the signal recognition particle (SRP) to the signal peptide that is present in many membrane proteins (Lakkaraju et al., 2008). In most cases, the speed of translation elongation is tuned to support the function of the protein that is being produced. For example, mRNAs with high codon optimality can support most efficient protein production (Presnyak et al., 2015), stretches of nucleotides with low codon optimality can decrease the speed of translation to allow more time for accurate protein folding (Spencer et al., 2012; Thanaraj and Argos, 1996), and binding of the SRP to the signal peptide stalls translation until an mRNA is translocated to the ER (Lakkaraju et al., 2008). It is important to note that when the speed of translation is reduced in these examples, translation is only slowed down, but in the end the polypeptide will be synthesized. However, in other cases the translating ribosomes can be stalled so strongly on an mRNA that no functional protein

can be produced. This can for example happen when strong secondary structures in the mRNA block the ribosome (Doma and Parker, 2006) or when ribosomes are stuck at the 3' end of the mRNA (Frischmeyer et al., 2002; Van Hoof et al., 2002). Prolonged stalling of ribosomes is a toxic situation for the cell, as mRNAs and ribosomes are sequestered away and the unfinished nascent peptides can potentially be toxic to the cell if they are released from the ribosome. Two different quality control mechanisms exist that target stalled ribosomes. No-go decay is activated by ribosomes that are stalled internally on the mRNA (i.e. not at the 3' end) (Doma and Parker, 2006), while non-stop decay targets mRNAs that lack a stop codon and have a ribosome stalled at their 3' end (Frischmeyer et al., 2002; Van Hoof et al., 2002).

Activation of No-go decay by stalled ribosomes

Ribosomes that stall for long periods of time can activate a quality control response called no-go decay (Figure 3B), which results in removal of the ribosome from the mRNA and can also result in mRNA degradation (Doma and Parker, 2006). Various sequences can be present in the mRNA that can induce ribosome stalling. For example, stalling can happen when strong secondary structures are present in the mRNA that cannot be removed by the helicase activity of the ribosome, such as large stem-loops or pseudoknots (Tholstrup et al., 2012; Wen et al., 2008). Stalling can also occur when long stretches of consecutive adenines are present in the mRNA, as this leads to a helical confirmation of the nucleotides of the mRNA in the ribosome that inhibits translocation (Arthur et al., 2015; Chandrasekaran et al., 2019; Koutmou et al., 2015). Besides mRNA-intrinsic sequences that block the ribosome, chemical modifications of the mRNA, such as oxidation of guanosine into 8-oxoguanosine, which happens as a consequence of oxidative stress, can also induce ribosome stalling (Simms et al., 2014; Yan et al., 2016). Removal of stalled ribosomes and defective mRNAs by no-go decay is important to prevent accumulation of defective mRNAs and to prevent sequestering of ribosomes on these mRNAs (Simms et al., 2014).

An important step in no-go decay is distinguishing ribosomes that are permanently stalled from ribosomes that are transiently stalled. Stalled ribosomes do not have a clear molecular structure that is used to distinguish a stalled ribosome from a non-stalled ribosome. Instead, stalling is detected when additional non-stalled ribosomes that translate the mRNA collide with the stalled ribosome (Juszkiewicz and Hegde, 2017; Juszkiewicz et al., 2018; Simms et al., 2017). Collision of two ribosomes creates a unique interface between the two 40S subunits of the ribosomes that can be bound by ZNF598, which ubiquitinates both the stalled ribosome and the colliding ribosome (Juszkiewicz and Hegde, 2017; Juszkiewicz et al., 2018). The ubiquitination is required for removal of the stalled ribosome and degradation of the mRNA, although it is not known by which factor the ubiquitination is recognized. Removal of the stalled

ribosome is accomplished through activation of the ribosomal quality control (RQC) complex (Brandman et al., 2012). In yeast, no-go decay degrades the mRNA through both endonucleolytic cleavage by Cue2 and exonucleolytic degradation by XRN1 (D’Orazio et al., 2019), resulting in efficient degradation of defective mRNAs. In humans, it is less clear whether stalled ribosomes only induce removal of the ribosome, or whether they also induce mRNA degradation. A homologue of Cue2 (N4BP2) is present in mammals (D’Orazio et al., 2019), but it remains to be determined whether this protein plays a role in mRNA degradation through no-go decay.

Ribosomes at the 3’ end of the mRNA induce non-stop decay

The second mRNA quality control mechanism that acts during translation elongation is non-stop decay (Figure 3C). Non-stop decay is activated by ribosomes that are stalled at the 3’ end of an mRNA, which can happen when an mRNA lacks an in-frame stop codon (Frischmeyer et al., 2002; Van Hoof et al., 2002). A major source of these ‘non-stop mRNAs’ are truncated mRNAs that arise from endonucleolytic cleavage, which is the cleavage of an mRNA into two cleavage fragments. When this cleavage occurs upstream of the stop codon, the resulting 5’ cleavage fragment will not have a stop

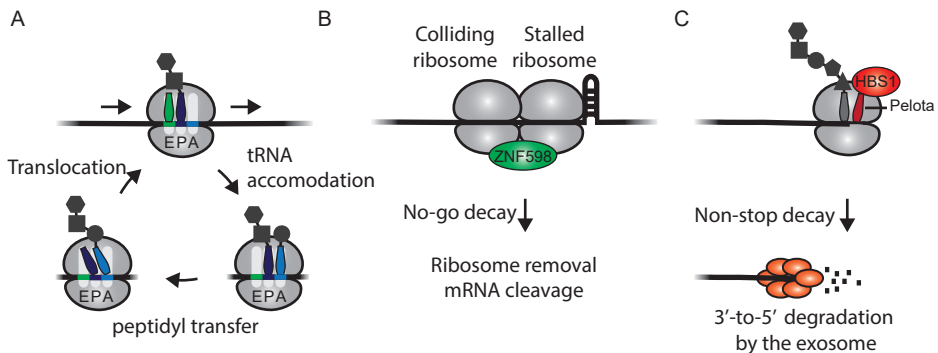


Figure 3. Translation elongation, no-go decay and non-stop decay. Schematic overview of translation elongation (A), no-go decay (B), and non-stop decay (C). A) Different stages of the translation elongation cycle. An aminoacyl-tRNA (blue shape) binds and is accommodated in the A-site of the ribosome. During peptidyl transfer, the aminoacyl-tRNA is added to the nascent peptide, extending the peptide by a single amino acid. The two tRNAs then move to the E- and P-site during the translocation step. The translation elongation cycle is completed by dissociation of the tRNA in the E-site. B) Strong stalling of ribosomes during translation is detected when a second ribosome collides with the stalled ribosome. Ribosome collision creates a new interface that can be bound by ZNF598, which ubiquitinates the ribosomes and activates no-go decay. C) Ribosomes that reach the 3’ end of the mRNA induce non-stop decay. Non-stop decay involves removal of the ribosome by Pelota and Hbs1 and subsequent degradation of the mRNA by the exosome. Figure 3A is adapted from (Sonneveld et al., 2020).

codon and is therefore a non-stop mRNA. Endonucleolytic cleavage is important in various mRNA degradation mechanisms, such as siRNA mediated cleavage, nonsense-mediated mRNA decay, and no-go decay (Arribere and Fire, 2018; Hashimoto et al., 2017; Tsuboi et al., 2012), and all these cleavages can create substrates for non-stop decay (Arribere and Fire, 2018). In addition to non-stop mRNAs created by endonucleolytic cleavage, non-stop mRNAs can also arise when the canonical stop codon of the mRNA is not present. This can for example happen as a consequence of a mutation of the stop codon in the DNA or through premature poly-adenylation. In contrast to non-stop mRNAs created by endonucleolytic cleavage, these non-stop mRNAs do have a poly(A) tail. It is currently unclear whether ribosomes reach the 3' end of polyadenylated non-stop mRNAs or whether ribosomes stall permanently upon translation of the polyadenine sequence of the poly(A) tail (Arthur et al., 2015; Chandrasekaran et al., 2019; Koutmou et al., 2015). This is important, as ribosomes stalled on the poly(A) tail are likely to induce no-go decay and these mRNAs might thus be degraded through a different mechanism than truncated non-stop mRNAs (that lack a poly(A) tail).

Ribosomes that are stalled at the 3' end of an mRNA can be distinguished from other ribosomes, as 3' stalled ribosomes are the only species of translating 80S ribosomes that do not have mRNA in the A-site. The empty A-site can be recognized by Pelota and Hbs1. These two proteins are structurally similar to the eukaryotic release factors eRF1 and eRF3, which are required for translation termination at a stop codon, but Pelota and Hbs1 preferentially bind when no mRNA is present in the A-site of the ribosome (Becker et al., 2011; Chen et al., 2010). Pelota and Hbs1 stimulate release of the nascent chain from the stalled ribosome, which is followed by ribosome recycling via ABCE1 (Pisareva et al., 2011; Shoemaker et al., 2010) and degradation of the mRNA through exonucleolytic decay by the exosome (Van Hoof et al., 2002). In addition, non-stop mRNAs can also be degraded through endonucleolytic cleavage when upstream ribosomes collide with the stalled ribosome and trigger a no-go decay response. Nevertheless, endonucleolytic cleavage of non-stop mRNAs has mainly been observed upon depletion of Pelota and/or Hbs1 (Arribere and Fire, 2018; Guydosh and Green, 2017; Tsuboi et al., 2012), suggesting that exosomal degradation is the main mechanism by which non-stop mRNAs are degraded.

Exosomal degradation

Structural studies of the exosome have given great insight into the mechanism of degradation by the exosome and revealed that the exosome consists of a barrel-like core complex that channels the mRNA to the catalytic subunit of the complex, Dis3L, which processively degrades the mRNA (Liu et al., 2006). Exosome activity is dependent on binding of the Ski2-complex to the exosome, which stimulates degradation by channelling the mRNA into the exosome, resulting in efficient degradation (Halbach

et al., 2013; Zinoviev et al., 2020). Interestingly, the Ski2-complex is normally in a closed conformation in which mRNA cannot pass through it. However, binding of the 40S ribosomal subunit induces a conformational change that allows mRNA to be channelled through the complex (Schmidt et al., 2016). Since non-stop mRNAs have ribosomes at the 3' end of the mRNA, it is possible that non-stop decay is activated by the opening of Ski2-complex by ribosomes at the 3' end of the mRNA, as this would activate the exosome while being in close proximity of a 3' end of an mRNA. However, no nucleotides extend from 3' end-stalled ribosomes, so there are no nucleotides available for channelling into the exosome when the ribosome is bound to the mRNA. This situation creates an apparent contradiction, in which a ribosome needs to be present to activate the exosome, but the exosome cannot degrade the mRNA when a ribosome is present. It therefore remains to be determined how the exact interplay between the ribosome and the exosome ultimately leads to efficient degradation of non-stop mRNAs

TRANSLATION TERMINATION

The third and final step of translation is translation termination, during which the newly synthesized peptide is released from the ribosome and the ribosome is recycled (Figure 4A). Translation termination occurs when a ribosome encounters a stop codon (also called termination codon, UAA, UGA, or UAG). Stop codons are recognized by eRF1, which is structurally similar to tRNA and binds in the form of a ternary complex consisting of eRF1, eRF3 and GTP (Brown et al., 2015; Taylor et al., 2012). After binding of the ternary complex to the ribosome, eRF3 hydrolyses GTP, which causes dissociation of eRF3 and accommodation of eRF1 in the A-site of the ribosome (Frolova et al., 1996). eRF1 then promotes hydrolysis of the ester bond between the nascent peptide and the tRNA, thereby releasing the nascent peptide. Finally, ABCE1 binds to eRF1 and promotes dissociation of the large ribosomal subunit (Pisarev et al., 2010). Release of the nascent chain and the ribosomal subunits completes the translation of an mRNA, allowing the newly made protein to exert its function in a cell and allowing the ribosomal subunits to be recycled for new rounds of translation. Translation termination is the final step during which the quality of the mRNA is monitored. In most cases, termination occurs with very high accuracy as ribosomes terminate with close to 100% efficiency at the first termination codon they encounter, but there are situations in which termination does not happen in the canonical way. In some cases, stop codons can be decoded as sense codons, resulting in the addition of extra amino acids to the C-terminus of the polypeptide. The frequency by which 'stop codon readthrough' happens depends on the nucleotide sequence of both the stop codon and the surrounding nucleotides, and in certain specific nucleotide contexts readthrough can

happen by 30% of all ribosomes (Loughran et al., 2014). In other nucleotide contexts, UGA stop codons are recoded to allow the incorporation of selenocysteine amino acids into the polypeptide (Berry et al., 1993). Another example in which translation termination occurs differently is during re-initiation. During re-initiation, the 40S ribosome continues to scan the mRNA until it encounters another start codon and re-initiates translation, resulting in the production of two distinct polypeptides from a single mRNA (Hinnebusch et al., 2016). These examples demonstrate how translation termination can be regulated to produce multiple proteins from a single mRNA. Besides regulated changes in termination, translation termination can also be affected when stop codons are erroneously introduced into the mRNA, for example because of mutations in the DNA. In these cases, termination at the newly introduced stop codon leads to production of defective protein. However, production of defective proteins is reduced by an mRNA quality control mechanism called nonsense-mediated decay (NMD) that monitors the fidelity of translation termination and degrades mRNAs that have acquired a premature stop codon.

Nonsense mediated mRNA decay (NMD)

mRNAs that have acquired a premature stop codon (or premature termination codon, PTC), which is a stop codon upstream of their normal stop codon, are degraded through NMD (Figure 4B). PTC-containing mRNAs produce C-terminally truncated proteins, which can have a negative impact on the cell in two different ways. First, truncated proteins are often missing one or multiple protein domains, which can make the protein non-functional. Second, truncated proteins are prone to aggregation or dominant negative effects. A major source of PTCs are mutations in the DNA, such as nonsense mutations, missense mutations, or mutations that induce alternative splicing. Strikingly, about ~20% of all disease-causing monogenetic mutations are nonsense or frameshift mutations that introduce a stop codon (Mort et al., 2008). PTCs can also be stochastically introduced in mRNAs that do not have a mutation in their DNA. For example, alternative splicing can lead to the inclusion of introns that have a stop codon, or errors in transcription can lead to a frameshift in the mRNA that introduces a PTC (Mendell et al., 2004). For mutations that are introduced during mRNA production, loss-of-function effects are not detrimental, as the errors occur at low frequency for many genes, but the accumulation of many different truncated proteins could still have negative effects. The detrimental effects of truncated proteins are limited by degradation of PTC-containing mRNAs through NMD. The importance of NMD is highlighted by the observation that mutations in the NMD pathway are associated with neurodegenerative diseases and cancer (Liu et al., 2014; Tarpey et al., 2007). The key step of NMD is distinguishing PTCs from normal termination codons. PTCs and normal stop codons have the same nucleotide sequence (UAA, UGA, or UAG), so the

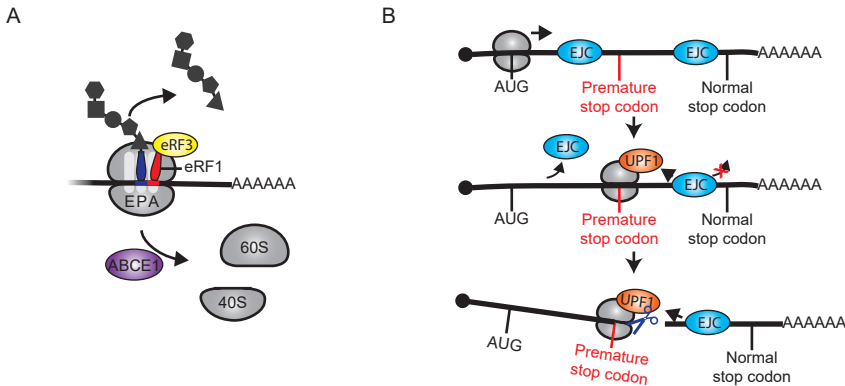


Figure 4. Translation termination and nonsense-mediated mRNA decay. Schematic overview of translation termination (A) and nonsense-mediated mRNA decay (B). A) During translation termination, a stop codon (red horizontal bar) is recognized by the release factors eRF1 and eRF3. Binding of eRF1 and eRF3 results in hydrolysis of the bond between the peptide and the tRNA, resulting in dissociation of the newly synthesized peptide from the ribosome. After peptide release, the ribosome subunits are split by ABCE1. B) During splicing, protein complexes called exon-junction complexes (EJCs, cyan shapes) are deposited on exon-exon junctions. EJCs are removed by translating ribosomes in the first round of translation. Most normal stop codons are located in the last exon of the gene, so all EJCs are removed by ribosomes before translation termination happens. Premature stop codons are often located upstream of the last exons, and EJCs downstream of a premature stop codon are not removed by translating ribosomes. Translation termination in the presence of mRNA-bound EJCs can activate NMD via the key NMD factor UPF1, which results in endonucleolytic cleavage of the mRNA by SMG6 (blue scissors). Figure 4A is adapted from (Sonneveld et al., 2020).

PTCs cannot be recognized simply by mRNA sequence alone. Instead, the distinction between PTCs and normal stop codons depends on protein complexes that are deposited on the mRNA during mRNA processing in the nucleus.

In the canonical NMD pathway, the distinction between PTCs and normal termination codons relies on the presence of exon-junction complexes (EJCs) that are bound to the mRNA. EJCs are multi-subunit protein complexes that are deposited approximately 20-24 nucleotides upstream of each exon-exon junction of an mRNA during splicing (Le Hir et al., 2000; Saulière et al., 2012; Singh et al., 2012). EJCs that are upstream of the stop codon are removed from the mRNA by translating ribosomes during the first round of translation, while EJCs downstream of the stop codon remain associated with the mRNA. In the vast majority of genes, the normal termination codon is located in the last exon. Therefore, all EJCs are upstream of the stop codon and are thus removed from the mRNA during the first round of translation. However, PTCs are often introduced in one of the internal exons (i.e. not the first or last exon). In these situations, EJCs will be deposited on the mRNA downstream of the PTC and will therefore not be removed by translating ribosomes. Translation termination on PTC-

containing mRNAs will thus happen in the presence of an mRNA-bound EJC, which is a trigger for activation of the NMD pathway (Ferraiuolo et al., 2004; Palacios et al., 2004). Activation of NMD involves the key NMD factor UPF1, which binds to eRF1 and eRF3, and together with SMG1 forms the SURF (SMG1-UPF1-release factor) complex (Kashima et al., 2006). When termination happens in the presence of an mRNA-bound EJC, the EJC complex can bind to the SURF complex via UPF3b and UPF2 (Kashima et al., 2006). This interaction enables phosphorylation of UPF1 by SMG1, which then recruits the endonuclease SMG6 that cleaves the mRNA into two cleavage fragments that are further degraded by the exosome and XRN1 (Eberle et al., 2009; Gatfield and Izaurralde, 2004; Huntzinger et al., 2008). Alternatively, the mRNA can also be degraded through exonucleolytic decay via recruitment of SMG5 and SMG7 to phosphorylated UPF1 (Loh et al., 2013; Unterholzner and Izaurralde, 2004). In addition to EJC-dependent NMD, mRNAs can also be degraded in an EJC-independent way. EJC-independent NMD is the main mechanism of NMD in yeast, as most yeast genes do not have introns, but also happens in humans (Buhler et al., 2006). Although EJC-independent NMD does not seem to degrade transcripts with disease-causing mutations in humans (Lindeboom et al., 2016), it does play a role in controlling expression levels of various human genes (Buhler et al., 2006; Yepiskoposyan et al., 2011).

Even though it is clear that NMD leads to degradation of PTC-containing mRNAs, not all NMD substrates are degraded with the same efficiency. In most studies, residual levels of NMD-substrates can be detected in steady state measurements, and the amount of residual mRNA varies between different NMD substrates (Cheng and Maquat, 1993; Cheng et al., 1994; Lindeboom et al., 2016; Thermann et al., 1998; Trcek et al., 2013). NMD efficiency depends on the gene and the position of the PTC. For example, mRNAs containing a PTC located less than 50 nucleotides upstream of the last exon-exon junction are not efficiently targeted NMD, as ribosomes translating these mRNAs are still able to displace the EJC bound to the last exon-exon junction (Lindeboom et al., 2016; Thermann et al., 1998). However, even efficient NMD substrates show some residual levels of mRNA. Understanding what causes (lack of) NMD efficiency is important as it provides insight into the molecular mechanisms of NMD. Residual NMD substrates could be a result of heterogeneity between mRNA molecules, in which some mRNAs are susceptible to NMD and others are not. However, residual substrates could also remain when NMD is not always induced when a ribosome terminates on a PTC, which could happen when NMD induction is stochastic and each ribosome has a certain probability of inducing NMD. Importantly, these scenarios would result in different degradation kinetics, and precise observation of the kinetics of NMD could therefore be useful to gain insight in the molecular mechanisms of NMD. Understanding of the molecular mechanisms of NMD efficiency could have a major impact on diseases that are caused by PTCs. This is especially true for PTC-introducing mutations that truncate

the protein without affecting its function. For example, mutations in the penultimate exon of dystrophin do not always make the protein non-functional, but since the PTC-containing mRNA is degraded by NMD, no protein is produced, effectively resulting in a loss-of-function effect (Kerr et al., 2001). In these cases, reducing the efficiency of NMD on the mutated mRNAs could have strong therapeutic benefit, and it is therefore important to understand what determines NMD efficiency.

METHODS TO STUDY mRNA QUALITY CONTROL

Since mRNA quality control involves an interplay between translation and mRNA degradation, valuable insight on mechanisms of mRNA quality control can be gained by methods that can simultaneously assess both translation and mRNA degradation. Classically, northern blotting has been used to study how mRNA stability was affected by various modifications to the mRNA or to changes in of the translation. These methods have given great insight into mRNA quality control and have been used to first identify nonsense-mediated mRNA decay, non-stop decay, and no-go decay (Doma and Parker, 2006; Frischmeyer et al., 2002; Van Hoof et al., 2002). However, Northern blotting only measures the average abundance of an mRNA in all cells at a single time point, precluding insights in kinetics of degradation or heterogeneity between mRNAs. Ribosome profiling gives a genome-wide view of the translational status of all mRNAs (Ingolia et al., 2009), and has been recently adopted to detect targets of endonucleolytic cleavage by detecting footprints of ribosomes at the 3' end of the mRNA (Guydosh and Green, 2017). Ribosome profiling can reveal which mRNAs or mRNA sequences induce endonucleolytic cleavage, but provides limited information on the mechanism by which mRNA decay is induced. In-vitro reconstitutions of translation can provide very precise insight into a process as each component of the reconstitution can be modulated separately. In vitro reconstitutions have for example been used to show how Pelota and Hbs1 preferentially remove ribosomes stalled at the 3' end of the mRNA rather than ribosomes stalled internally on an mRNA (Pisareva et al., 2011). However, in vitro reconstitutions are hard to establish and not always mimic in vivo situations as not all factors that are present in a cell are present in the reconstitution. Live-cell imaging approaches have been used to gain both spatial and temporal insight in mRNA quality control. By introducing binding sites in an mRNA that recruit fluorescent RNA-binding proteins (such as MS2 or PP7), the location of mRNA molecules in cells can be followed in a real-time, and these approaches have been used to assess the location of mRNA degradation as well as their timing relative to nuclear export (Trcek et al., 2013). More recently, new methods have been developed that combine two fluorescent RNA-binding proteins to visualize the first round of translation or

endonucleolytic cleavage of reporter mRNAs (Halstead et al., 2015; Horvathova et al., 2017). Live cell imaging approaches have greatly aided in understanding of kinetics of mRNA degradation. However, current live-cell imaging methods do not simultaneously visualize translation and mRNA degradation, and could therefore be improved to study mRNA quality control. Because of the lack of methods for simultaneous real-time observation of translation and decay, many questions on mRNA quality control have been challenging to address. For example, it has been difficult to address if mRNA quality control is limited to a certain time window, as has been proposed for NMD, or if it can always happen, if mRNA degradation is affected by the spatial localization of the mRNA, and if all mRNA molecules are targeted with equal efficiency or some mRNAs are preferentially degraded, for example because of differences in mRNA processing or RNA modifications. Therefore, new techniques are required that enable real time observation of both translation and mRNA degradation of individual mRNA molecules.

THESIS OVERVIEW

In **chapter 2**, we describe the development of a new method that uses the SunTag for real-time observation of the translation of single mRNA molecules. We use this method to determine the rates of translation initiation and elongation for individual mRNAs, and observe how certain mRNA sequence can induce stalling of ribosomes. In addition, we find that there can be strong heterogeneity between translation efficiency of mRNA molecules derived from the same DNA. In **chapter 3**, we go into more details of this method of SunTag translation imaging. In **chapter 4**, we further develop our translation imaging method in order to simultaneously visualize translation of single mRNA molecules and their degradation by NMD. Using this new method, we discover that NMD happens with equal efficiency during each round of translation, but that a fraction of mRNA molecules is insensitive to NMD. We also gain insight into the dynamics of the first round of translation and its relation to NMD, and find out how the efficiency of NMD is affected by the number and position of introns in a gene. Finally, we gained insight in how the 3' cleavage fragment is degraded by XRN1. In **chapter 5**, we apply our method in a different way to gain insight in degradation of mRNAs by the exosome through non-stop decay. Although non-stop decay is stimulated by ribosomes at the 3' end of the mRNA, we observe that exosomal degradation is not only stimulated by ribosomes at the 3' end of the mRNA, but also by ribosomes that terminate translation in close proximity of the 3' end of the mRNA. In addition, we find that that exosome degradation is delayed by ribosomes that are stalled by translation inhibitors, indicating that ribosomes can block the exosome. In **chapter**

6, we dive deeper in the protocols behind the new applications of SunTag imaging method. Finally, in **chapter 7**, we try to help in fighting the COVID-19 pandemic by establishing a new procedure for diagnostic testing. This procedure enables testing at a much higher throughput than current procedures and thereby aids in screening large populations for COVID-19.

REFERENCES

- Arribere, J.A., and Fire, A.Z. (2018). Nonsense mRNA suppression via nonstop decay. *Elife* 7, 1–23.
- Arthur, L.L., Pavlovic-Djuranovic, S., Koutmou, K.S., Green, R., Szczesny, P., and Djuranovic, S. (2015). Translational control by lysine-encoding A-rich sequences. *Sci. Adv.* 1.
- Babendure, J.R., Babendure, J.L., Ding, J.H., and Tsien, R.Y. (2006). Control of mammalian translation by mRNA structure near caps. *RNA* 12, 851–861.
- Becker, T., Armache, J.-P., Jarasch, A., Anger, A.M., Villa, E., Sieber, H., Motaal, B.A., Mielke, T., Berninghausen, O., and Beckmann, R. (2011). Structure of the no-go mRNA decay complex Dom34–Hbs1 bound to a stalled 80S ribosome. *Nat. Struct. Mol. Biol.* 18, 715–720.
- Berry, M.J., Banu, L., Harney, J.W., and Larsen, P.R. (1993). Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J.* 12, 3315–3322.
- Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.W., Zhou, S., King, D., Shen, P.S., Weibezahn, J., et al. (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* 151, 1042–1054.
- Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. *Nature* 524, 493–496.
- Brutlag, D., and Kornberg, A. (1972). Synthesis of Deoxyribonucleic. *J. Biol. Chem.* 247, 241–248.
- Buhler, M., Steiner, S., Mohn, F., Paillusson, A., and Muhlemann, O. (2006). EJC-independent degradation of nonsense immunoglobulin- μ mRNA depends on 3' UTR length. *Nat. Struct. Mol. Biol.* 13, 462–464.
- Chandrasekaran, V., Juskiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. *Nat. Struct. Mol. Biol.* 26, 1132–1140.
- Chen, L., Muhrad, D., Hauryliuk, V., Cheng, Z., Lim, M.K., Shyp, V., Parker, R., and Song, H. (2010). Structure of the Dom34-Hbs1 complex and implications for no-go decay. *Nat. Struct. Mol. Biol.* 17, 1233–1240.
- Cheng, J., and Maquat, L.E. (1993). Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. *Mol. Cell. Biol.* 13, 1892–1902.
- Cheng, J., Belgrader, P., Zhou, X., and Maquat, L.E. (1994). Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance. *Mol. Cell. Biol.* 14, 6317–6325.
- D'Orazio, K.N., Wu, C.C.-C., Sinha, N., Loll-Krippelber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. *Elife* 8, 1–27.
- Dever, T.E., Dinman, J.D., and Green, R. (2018). Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* 10, a032649.
- Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* 440, 561–564.
- Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* 16, 49–55.
- Ebright, R.Y., Lee, S., Wittner, B.S., Niederhoffer, K.L., Nicholson, B.T., Bardia, A., Truesdell, S., Wiley, D.F., Wesley, B., Li, S., et al. (2020). Deregulation of ribosomal protein expression and translation promotes

breast cancer metastasis. *Science* 367, 1468–1473.

Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., and Bartel, D.P. (2020). The Dynamics of Cytoplasmic mRNA Metabolism. *Mol. Cell* 77, 786–799.e10.

Ferraiuolo, M.A., Lee, C.-S., Ler, L.W., Hsu, J.L., Costa-Mattioli, M., Luo, M.-J., Reed, R., and Sonenberg, N. (2004). A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. *Proc. Natl. Acad. Sci.* 101, 4118–4123.

Frishmeyer, P.A., Van Hoof, A., O'Donnell, K., Guerrero, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* 295, 2258–2261.

Frolova, L., Goff, X.L.E., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA* 2, 334–341.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* 429, 575–578.

Guydosh, N.R., and Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. *RNA* 23, 749–761.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The yeast ski complex: Crystal structure and rna channeling to the exosome complex. *Cell* 154, 814–826.

Halstead, J.M., Lionnet, T., Wilbertz, J.H., Wippich, F., Ephrussi, A., Singer, R.H., and Chao, J.A. (2015). An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science* 347, 1367–1671.

Hashimoto, Y., Takahashi, M., Sakota, E., and Nakamura, Y. (2017). Nonstop-mRNA decay machinery is involved in the clearance of mRNA 5'-fragments produced by RNAi and NMD in *Drosophila melanogaster* cells. *Biochem. Biophys. Res. Commun.* 484, 1–7.

He, F., and Jacobson, A. (2015). Nonsense-mediated mRNA decay: degradation of defective transcripts is only part of the story. *Annu. Rev. Genet.* 49, 339–366.

Hinnebusch, A.G., Ivanov, I.P., and Sonenberg, N. (2016). Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* 352, 1413–1416.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *EMBO J.* 19, 6860–6869.

Holt, C.E., and Schuman, E.M. (2013). The Central Dogma Decentralized: New Perspectives on RNA Function and Local Translation in Neurons. *Neuron* 80, 648–657.

Van Hoof, A., Frishmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* 295, 2262–2264.

Horvathova, I., Voigt, F., Kotrys, A. V., Zhan, Y., Artus-Revel, C.G., Eglinger, J., Stadler, M.B., Giorgetti, L., and Chao, J.A. (2017). The dynamics of mRNA turnover revealed by single-molecule imaging in single cells. *Mol. Cell* 68, 615–625. e9.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* 14, 2609–2617.

Imataka, H. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* 17, 7480–7489.

- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* *324*, 218–223.
- Johnstone, T.G., Bazzini, A.A., and Giraldez, A.J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J.* *35*, 706–723.
- Juszkiewicz, S., and Hegde, R.S. (2017). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol. Cell* *65*, 743–750.e4.
- Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., and Hegde, R.S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol. Cell* *72*, 469–481.e7.
- Kanaar, R., Hoeijmakers, J.H.J., and van Gent, D.C. (1998). Molecular mechanisms of DNA double-strand break repair. *Trends Cell Biol.* *8*, 483–489.
- Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* *20*, 355–367.
- Kerr, T.P., Sewry, C.A., Robb, S.A., and Roberts, R.G. (2001). Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay? *Hum. Genet.* *109*, 402–407.
- Koutmou, K.S., Schuller, A.P., Brunelle, J.L., Radhakrishnan, A., Djuranovic, S., and Green, R. (2015). Ribosomes slide on lysine-encoding homopolymeric A stretches. *Elife* *2015*, 1–18.
- Kozak, M. (1990). Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci.* *87*, 8301–8305.
- Lakkaraju, A.K.K., Mary, C., Scherrer, A., Johnson, A.E., and Strub, K. (2008). SRP Keeps Polypeptides Translocation-Competent by Slowing Translation to Match Limiting ER-Targeting Sites. *Cell* *133*, 440–451.
- Lindeboom, R.G.H., Supek, F., and Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat. Genet.* *48*, 1112–1118.
- Liu, B., and Qian, S.-B. (2014). Translational reprogramming in stress response. *Wiley Interdiscip. Rev. RNA* *5*, 301–305.
- Liu, C., Karam, R., Zhou, Y., Su, F., Ji, Y., Li, G., Xu, G., Lu, L., Wang, C., Song, M., et al. (2014). The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenocarcinoma. *Nat. Med.* *20*, 596–598.
- Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. *Cell* *127*, 1223–1237.
- Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5–SMG7 heterodimer directly recruits the CCR4–NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.* *27*, 2125–2138.
- Loughran, G., Chou, M.-Y., Ivanov, I.P., Jungreis, I., Kellis, M., Kiran, A.M., Baranov, P. V., and Atkins, J.F. (2014). Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res.* *42*, 8928–8938.
- Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martínez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat. Genet.* *36*, 1073–1078.
- Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA Processing Reaches Back to Transcription and Ahead to Translation. *Cell* *136*, 688–700.

- Mort, M., Ivanov, D., Cooper, D.N., and Chuzhanova, N.A. (2008). A meta-analysis of nonsense mutations causing human genetic disease. *Hum. Mutat.* *29*, 1037–1047.
- Nirenberg, M.W., and Matthaei, J.H. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci.* *47*, 1588–1602.
- Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* *427*, 753–757.
- Pisarev, A. V., Skabkin, M.A., Pisareva, V.P., Skabkina, O. V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U.T., and Pestova, T. V. (2010). The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling. *Mol. Cell* *37*, 196–210.
- Pisareva, V.P., Skabkin, M.A., Hellen, C.U.T., Pestova, T. V., and Pisarev, A. V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J.* *30*, 1804–1817.
- Presnyak, V., Alhusaini, N., Chen, Y.-H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., et al. (2015). Codon Optimality Is a Major Determinant of mRNA Stability. *Cell* *160*, 1111–1124.
- Saulière, J., Murigneux, V., Wang, Z., Marquet, E., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Crollius, H.R., and Le Hir, H. (2012). CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct. Mol. Biol.* *19*, 1124–1131.
- Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., et al. (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. *Science* *354*, 1431–1433.
- Schwahnüsser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337–342.
- Shah, P., Ding, Y., Niemczyk, M., Kudla, G., and Plotkin, J.B. (2013). Rate-Limiting Steps in Yeast Protein Translation. *Cell* *153*, 1589–1601.
- Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* *330*, 369–372.
- Simms, C.L., Hudson, B.H., Mosior, J.W., Rangwala, A.S., and Zaher, H.S. (2014). An Active Role for the Ribosome in Determining the Fate of Oxidized mRNA. *Cell Rep.* *9*, 1256–1264.
- Simms, C.L., Yan, L.L., and Zaher, H.S. (2017). Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol. Cell* *68*, 361–373.e5.
- Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* *151*, 750–764.
- Sonneveld, S., Verhagen, B.M.P., and Tanenbaum, M.E. (2020). Heterogeneity in mRNA Translation. *Trends Cell Biol.* *30*, 606–618.
- Spencer, P.S., Siller, E., Anderson, J.F., and Barral, J.M. (2012). Silent Substitutions Predictably Alter Translation Elongation Rates and Protein Folding Efficiencies. *J. Mol. Biol.* *422*, 328–335.
- Tarpey, P.S., Lucy Raymond, F., Nguyen, L.S., Rodriguez, J., Hackett, A., Vandeleur, L., Smith, R., Shoubridge, C., Edkins, S., Stevens, C., et al. (2007). Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat. Genet.* *39*, 1127–1133.

- Taylor, D., Unbehaun, A., Li, W., Das, S., Lei, J., Liao, H.Y., Grassucci, R.A., Pestova, T. V., and Frank, J. (2012). Cryo-EM structure of the mammalian eukaryotic release factor eRF1-eRF3-associated termination complex. *Proc. Natl. Acad. Sci.* *109*, 18413–18418.
- Thanaraj, T.A., and Argos, P. (1996). Protein secondary structural types are differentially coded on messenger RNA. *Protein Sci.* *5*, 1973–1983.
- Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeyer, C., Hentze, M.W., and Kulozik, A.E. (1998). Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.* *17*, 3484–3494.
- Tholstrup, J., Oddershede, L.B., and Sørensen, M.A. (2012). mRNA pseudoknot structures can act as ribosomal roadblocks. *Nucleic Acids Res.* *40*, 303–313.
- Tjian, R. (1978). The binding site on SV40 DNA for a T antigen-related protein. *Cell* *13*, 165–179.
- Trcek, T., Sato, H., Singer, R.H., and Maquat, L.E. (2013). Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes Dev.* *27*, 541–551.
- Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3' End of Aberrant mRNA. *Mol. Cell* *46*, 518–529.
- Unterholzner, L., and Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell* *16*, 587–596.
- Wen, J. Der, Lancaster, L., Hodges, C., Zeri, A.C., Yoshimura, S.H., Noller, H.F., Bustamante, C., and Tinoco, I. (2008). Following translation by single ribosomes one codon at a time. *Nature* *452*, 598–603.
- Wu, Q., Medina, S.G., Kushawah, G., DeVore, M.L., Castellano, L.A., Hand, J.M., Wright, M., and Bazzini, A.A. (2019). Translation affects mRNA stability in a codon-dependent manner in human cells. *Elife* *8*, 1–22.
- Yan, X., Hoek, T.A., Vale, R.D., and Tanenbaum, M.E. (2016). Dynamics of Translation of Single mRNA Molecules in Vivo. *Cell* *165*, 976–989.
- Yepiskoposyan, H., Aeschmann, F., Nilsson, D., Okoniewski, M., and Muhlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* *17*, 2108–2118.
- Zinoviev, A., Ayupov, R.K., Abaeva, I.S., Hellen, C.U.T., and Pestova, T. V. (2020). Extraction of mRNA from Stalled Ribosomes by the Ski Complex. *Mol. Cell* *77*, 1340-1349.e6.



Chapter 2

Dynamics of Translation of Single mRNA Molecules In Vivo

Xiaowei Yan, Tim A. Hoek, Ronald D. Vale and Marvin E. Tanenbaum

ABSTRACT

Regulation of mRNA translation, the process by which ribosomes decode mRNAs into polypeptides, is used to tune cellular protein levels. Currently, methods for observing the complete process of translation from single mRNAs *in vivo* are unavailable. Here, we report the long-term (>1 hr) imaging of single mRNAs undergoing hundreds of rounds of translation in live cells, enabling quantitative measurements of ribosome initiation, elongation and stalling. This approach reveals a surprising heterogeneity in the translation of individual mRNAs within the same cell, including rapid and reversible transitions between a translating and non-translating state. Applying this method to the cell cycle gene *Emi1*, we find strong overall repression of translation initiation by specific 5'UTR sequences, but individual mRNA molecules in the same cell can exhibit dramatically different translational efficiencies. The ability to observe translation of single mRNA molecules in live cells provides a powerful tool to study translation regulation.

INTRODUCTION

Precise tuning of the expression of each gene in the genome is critical for many aspects of cell function. The level of gene expression is regulated at multiple distinct steps, including transcription, mRNA degradation and translation (Schwanhausser et al., 2011). Regulation of all of these steps in gene expression is important, though the relative contribution of each control mechanism varies for different biological processes (Brar et al., 2012; Jovanovic et al., 2015; Peshkin et al., 2015; Tanenbaum et al., 2015; Vardy and Orr-Weaver, 2007).

Measuring the translation rate from individual mRNAs over time provides valuable information on the mechanisms of translation and translational regulation. *In vitro* experiments, mainly using bacterial ribosomes, have revealed exquisite information on ribosome translocation dynamics at the single molecule level (Blanchard, 2009; Chen et al., 2012; Cornish et al., 2008; Fei et al., 2008; Wen et al., 2008; Zaher and Green, 2009), but such methods have not yet been applied *in vivo*. In contrast, a genome-wide snapshot of the translational efficiency of endogenous mRNAs *in vivo* can be obtained through the method of ribosomal profiling (Ingolia et al., 2009; Ingolia et al., 2011). However, this method requires averaging of many cells and provides limited temporal information because of the requirement to lyse cells to make these measurements. Single cell imaging studies have succeeded in measuring average protein synthesis rates (Aakalu et al., 2001; Brittis et al., 2002; Han et al., 2014; Leung et al., 2006; Tanenbaum et al., 2015; Yu et al., 2006), observing the first translation event of an mRNA (Halstead et al., 2015), localizing sub-cellular sites of translation by co-localizing mRNAs and ribosomes (Katz et al., 2016; Wu et al., 2015), and staining nascent polypeptides with small molecule dyes (Rodriguez et al., 2006).

While ribosomal profiling and other recently developed methods have provided many important new insights into the regulation of translation, many questions cannot be addressed using current technologies. For example, it is unclear to what extent different mRNA molecules produced in a single cell from the same gene behave similarly. Many methods to study translation *in vivo* require averaging of many mRNAs, masking potential differences between individual mRNA molecules. Such differences could arise from differential post-transcriptional regulation, such as nucleotide modifications (Choi et al., 2016; Wang et al., 2015), differential transcript lengths through use of alternative transcriptional start sites (Rojas-Duran and Gilbert, 2012) or polyadenylation site selection (Elkon et al., 2013; Gupta et al., 2014), differences in ribonucleic protein (RNP) composition (Wu et al., 2015), distinct intracellular localization (Huttelmaier et al., 2005), or different states of RNA secondary structure (Babendure et al., 2006; Kertesz et al., 2010). Heterogeneity among mRNA molecules could have a profound impact on the total amount of polypeptide produced, as well as the localization of protein synthesis, but remains poorly studied. Furthermore, the extent to which translation

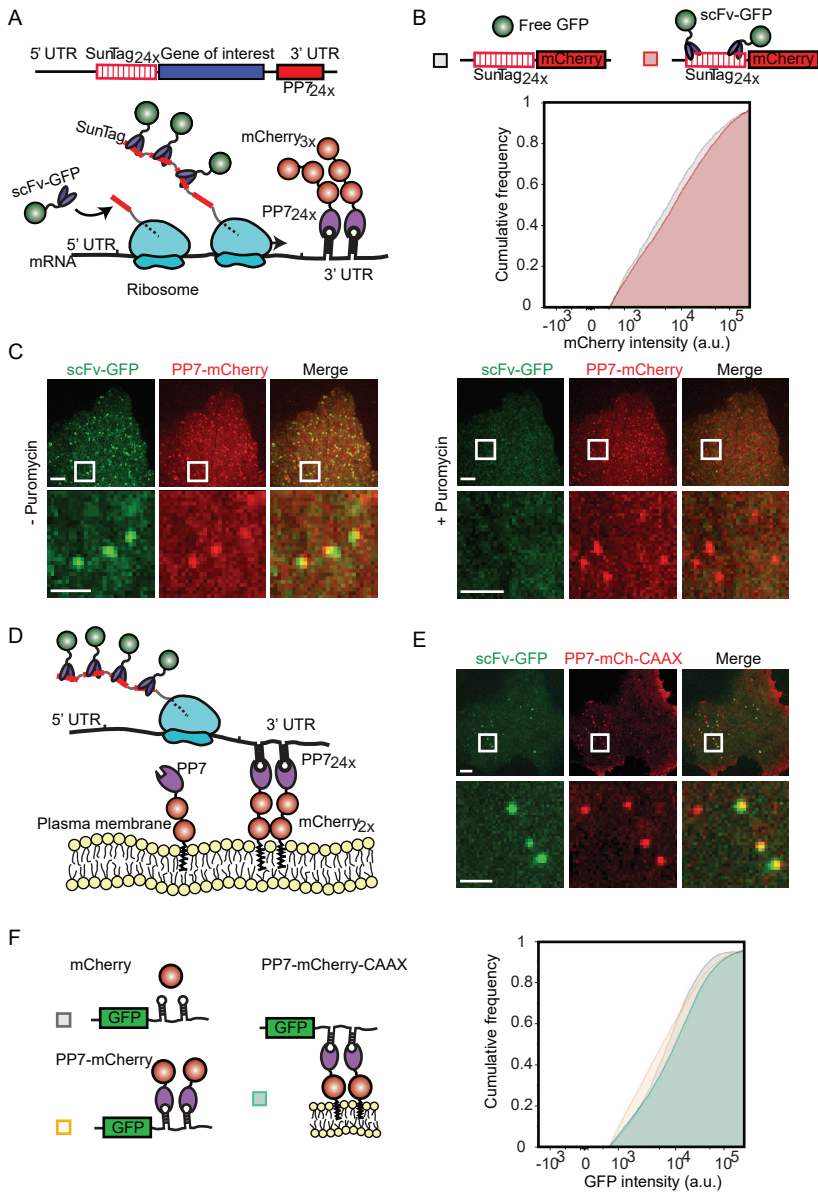


Figure 1. Fluorescence labeling of nascent chains to visualize translation of single mRNA molecules. A) Schematic of nascent polypeptides labeling using the SunTag system and mRNA labeling (A) and membrane tethering (D) using the PP7 system. B) An mCherry-SunTag_{24x} reporter gene was co-transfected with either GFP or scFv-GFP, and the expression of the SunTag_{24x}-mCherry reporter was determined by FACS (See Methods). Binding of the scFv-GFP to the SunTag nascent chain did not detectably alter protein expression. C) A representative U2OS cell is shown expressing scFv-GFP, PP7-3xmCherry and the translation reporter (SunTag_{24x}-Kif18b-PP7_{24x}). Cytosolic translation sites (scFv-GFP) co-localize with mRNAs (PP7-3xmCherry). Ribosomes were dissociated from mRNA by addition of puromycin (right panel).

on single mRNA molecules varies over time also is largely unknown. For example, translation may occur in bursts, rather than continuously (Tatavarty et al., 2012; Yu et al., 2006), and regulation of protein synthesis may occur by modulating burst size and/or frequency, which could occur either globally or on each mRNA molecule individually. In addition, the ability of an mRNA molecule to initiate translation may vary with time or spatial location, for example as cells progress through the cell cycle (Stumpf et al., 2013; Tanenbaum et al., 2015) or undergo active microtubule-based transport to particular cellular destinations (Holt and Schuman, 2013). Such regulation could involve changes in the rates of translation initiation and/or the ribosome elongation. To address these questions, new methods are required for visualizing translation on single mRNA molecules in live cells over time.

Here, we present a method, based on the SunTag fluorescence tagging system that we recently developed (Tanenbaum et al., 2014), for measuring the translation of single mRNA molecules over long periods of time. Using this system, we have measured initiation, elongation and stalling on individual mRNA molecules and have uncovered unexpected heterogeneity among different mRNAs molecules encoded by the same gene within a single cell. Our system will be widely applicable to the study of mRNA translation in live cells.

RESULTS

An assay for long term observation of translation of individual mRNAs

Observing the synthesis of a genetically-encoded fluorescent protein, such as GFP, *in vivo* is difficult because of the relatively long maturation time required to achieve a fluorescent state. Thus, a GFP-fusion protein typically will not fluoresce until after its translation is completed. To overcome this temporal challenge and to create a sufficiently bright signal to observe protein synthesis from single mRNAs *in vivo*, we used our recently developed SunTag system (Tanenbaum et al., 2014). In this assay, cells are co-transfected with a reporter transcript containing an array of 24 SunTag

Figure 1 continued... Note that translation sites and mRNA do not perfectly overlap, because of the brief time difference in acquiring GFP and mCherry images. D) Schematic of nascent polypeptides labeling and membrane tethering of the mRNA using the PP7 system. E) U2OS cells expressing scFv-GFP (green), PP7-2xmCherry-CAAX (red) and the translation reporter (SunTag_{24x}-Kif18b-PP7_{24x}). A single time-point of the cell (top panel), and a zoomed-in view from the white boxed area containing a few mRNAs (lower panel) are shown. F) U2OS cells were transfected with mCherry, PP7-mCherry or PP7-mCherry-CAAX together with a GFP reporter transcript with 24 PP7 binding sites in the 3'UTR, and GFP expression was analyzed by FACS (see Methods). Cumulative distribution of GFP expression levels from GFP-mCherry double positive cells are shown in (B) and (F) ($n = 3$ independent experiments). Scale bars, 5 μm (upper panels) and 2 μm (lower panels). See also Figure S1 and Movies S1-3.

peptides followed by a gene of interest, along with a second construct expressing a GFP-tagged single-chain intracellular antibody (scFv-GFP) that binds to the SunTag peptide with high affinity. As the SunTag peptides are translated and emerge from the ribosome exit tunnel, they are rapidly bound by the soluble and already fluorescent scFv-GFP (Figure 1A). Importantly, labeling of nascent chains using the SunTag antibody did not detectably alter protein synthesis rates of a reporter mRNA in human U2OS cells, as determined by FACS analysis (Figure 1B). At the same time, the mRNA was fluorescently labeled by introducing 24 copies of a short hairpin sequence into the 3'UTR, and co-expressing the PP7 bacteriophage coat protein (Chao et al., 2008), which binds with high affinity to the hairpin sequence, fused to three copies of mCherry (PP7-mCherry) (Figure 1A).

When observed by spinning disk confocal microscopy, the co-expression of a reporter construct (SunTag_{24x}-Kif18b-PP724x, with Kif18b being a kinesin motor with a 2.5 kb coding sequence (Tanenbaum et al., 2011)), scFv-GFP and PP7-mCherry resulted in the appearance of a small number (10-50) of very bright green and red fluorescent spots per cell that co-migrated in time-lapse movies (Figure 1C and Movie S1). Spot tracking revealed that these spots diffused with a diffusion coefficient of $0.047 \mu\text{m}^2/\text{s}$, which is slightly slower than previous measurements of mRNA diffusion ($0.1 - 0.4 \mu\text{m}^2/\text{s}$) (Katz et al., 2016), consistent with the fact that our reporter mRNA contains a larger open reading frame (4.4 kb versus 1.1 kb) and thus more associated ribosomes. In addition, we observed many dim GFP spots that did not co-migrate with a mCherry signal in time-lapse movies. The bright spots rapidly disappeared upon terminating translation by addition of a protein synthesis inhibitor puromycin, which dissociates nascent polypeptides and ribosomes from mRNA (Figure 1C and Movie S2), indicating that they are sites of active translation where multiple ribosomes are engaged on a single mRNA molecule. The dim spots were unaffected by puromycin treatment, suggesting that they represent individual, fully synthesized SunTag24x-Kif18b proteins that had already been released from the ribosome. Thus, this translation imaging assay allows visualization of ongoing translation of single mRNA molecules.

Rapid 3-D diffusion of mRNAs makes it difficult to track single mRNAs for >1 min, as mRNAs continuously diffuse in and out of the z-plane of observation, and mRNAs regularly cross paths, complicating identification and tracking of individual mRNA molecules over time. To track mRNAs unambiguously for long periods of time, we added a CAAX sequence, a prenylation sequence that gets inserted into the inner leaflet of the plasma membrane, to the PP7-mCherry protein which served to tether mRNAs to the 2-D plane of the plasma membrane (Figure 1D, E). As a result of many PP7-mCherry molecules clustering through their interaction with the multiple recognition sites on a single mRNA, bright red dots appeared on the plasma membrane at the bottom on cell, representing a tethered mRNA molecule (Figure 1E). Tethered mRNA molecules co-migrated with scFv-GFP foci, indicating that they are sites of active translation (Figure

1E and Movie S3). The membrane tethering of the mRNA had minimal effects on the protein expression of a GFP reporter construct as analyzed by FACS (Figure 1F). While membrane tethering greatly improves the ability to visualize translation on single mRNA molecules over long periods of time and does not appear to grossly perturb mRNA translation, it is important to note that some aspects of translation, especially localized translation, may be altered due to tethering (See Discussion).

We first analyzed the PP7-mCherry spots observed on the plasma membrane to confirm that they contained only a single mRNA molecule. The fluorescence intensities of PP7-mCherry foci were very homogeneous (Figure S1A); their absolute intensity was ~ 1.4 -fold brighter, on average, than single, membrane-tethered SunTag_{24x}-CAAX proteins bound with scFv-mCherry, which is expected to contain 24 mCherry molecules (Figure S1B). PP7 binds as a dimer to the RNA hairpin and each PP7 was tagged with two tandem copies of mCherry. Thus mRNAs spots could be expected to be four times as bright as single scFv-mCherry-SunTag_{24x}-CAAX spots, but previous studies suggested that only about half of PP7 binding sites may be occupied (Wu et al., 2015); thus, mRNA spots would be about 2-fold brighter than single mCherry-SunTag_{24x} spots if they contain a single mRNA molecule, but 4- or more fold brighter if they contained 2 or more mRNAs. These results are therefore most consistent with the mCherry-PP7 foci being single mRNA molecules rather than multiple copies of mRNAs. Further supporting this idea, we tracked 63 single mRNA foci for 30-45 min and did not find a single case in which one spot split into two, which would have been indicative of more than one mRNA molecule being present in a single spot.

Because single mRNAs were tethered to the plasma membrane through multiple PP7 molecules, and thus through many CAAX membrane insertion domains, the 2-D diffusion of mRNAs was extremely slow ($1.06 \times 10^{-3} \mu\text{m}^2/\text{s}$, $n = 211$ mRNAs). This slow diffusion made it possible to track individual mRNAs and their associated translation sites for extended periods of time (mean tracking time >1 hr, Figure S1C). Furthermore, the very slow diffusion rate of tethered mRNAs allowed us to image tethered translation sites using long exposure times (500-1000 ms); during this time interval, rapidly diffusing, non-tethered fully synthesized polypeptides only produced a blurred, diffuse image on the camera sensor, which enabled sites of translation to be easily distinguished from fully synthesized molecules (Figure S1D). Finally, to confirm that the scFv-GFP was binding to nascent SunTag peptides, we replaced the SunTag epitope peptides in our reporter mRNA with an unrelated nucleotide sequence (encoding BFP) and found no GFP foci formation near mRNAs (Figure S1E).

In conclusion, we have developed assays that enable both single mRNAs and their associated nascent translating polypeptides to be imaged over time. This general SunTag-based method can be performed with either freely diffusing mRNAs or mRNAs tethered to the plasma membrane, each of which has unique advantages depending on the specific biological question (See Discussion). For further experiments in this study,

we used the membrane-tethered system to follow translation for long periods of time.

Measurement of ribosome number, initiation rate and elongation rate on single mRNAs

To estimate the number of ribosomes translating each mRNA, we compared the scFv-GFP fluorescence intensity of translation sites with that of the single, fully synthesized SunTag_{24x}-Kif18b molecules present in the same cell (Figure S2A, B). Several considerations need to be taken into account to calculate ribosome number from the fluorescence intensities of translation sites and fully synthesized single SunTag proteins (See also Extended Experimental Procedures). First, ribosomes present at the 5' end of the reporter transcript have translated only a subset of the 24 SunTag peptides, so the nascent polypeptide associated with these ribosomes will have a lower fluorescence intensity due to fewer bound scFv-GFPs. We generated a mathematical model to correct for the difference in fluorescence intensity for ribosomes at different positions along the transcript (See Extended Experimental Procedures). Second, if scFv-GFP-peptide has a slow on-rate for the epitope *in vivo*, a lag time could exist between the synthesis of a SunTag peptide and binding of a scFv-GFP, which could result in the underestimation of the number of ribosomes per mRNA. To test this, cells were treated with the translation inhibitor cycloheximide (CHX), which blocks ribosome elongation by locking ribosomes on the mRNA and prevents the synthesis of new SunTag peptides, while allowing binding of scFv-GFP to existing peptides to reach equilibrium. The translation site scFv-GFP signal did not substantially increase after CHX treatment (Figure S2C), indicating that under our experimental conditions, the lag time between peptide synthesis and scFv-GFP binding does not detectably affect translation site intensity. Based on the above controls and our mathematical model, we could estimate the ribosome number per mRNA from the fluorescence intensity of the translation site. Approximately 30% of the mRNAs did not have a corresponding GFP signal, suggesting that they were not actively translating. For the remaining 70% of the mRNAs that were translating, the majority (76%) had between 10-25 ribosomes (Figure 2A; see Extended Experimental Procedures), corresponding to an average inter-ribosome distance of ~200-400 nucleotides (nt). We also compared translation site intensity of two additional reporter mRNAs with either 5x or 10x SunTag peptides with the 24x peptide reporter. This analysis revealed that ribosome density was very similar on the 5x and 10x reporter (1.26-fold and 1.00-fold, respectively) (Figure S2D), indicating that the long 24x SunTag array does not grossly perturb ribosome loading on the reporter mRNA.

Next, we measured the translocation speed of ribosomes on single mRNAs by treating cells with harringtonine, a small molecule inhibitor of translation that stalls new ribosomes at the start of the mRNA coding sequence without affecting ribosomes further downstream (Ingolia et al., 2011). As mRNA-bound ribosomes complete translation one-by-one after harringtonine treatment, the GFP signal on mRNAs

decreases (Figure 2B-D and Movie S4). Using a simple mathematical model to fit the decay in fluorescence of a cumulative curve from many mRNAs (see Extended Experimental Procedures and Figure S7), we estimate a ribosome translocation rate of 3.5 ± 1.1 codons/s. In a parallel approach, we also measured the total time required for run-off of all ribosomes from individual mRNAs (Figure S2E), from which we calculated a similar translation elongation rate (3.1 ± 0.14 codons/s) as the one obtained through our model (See Extended Experimental Procedures). A reporter with only 5 instead of 24 SunTag peptides showed similar elongation kinetics (3.1 ± 0.4 codons/s; Figure S2F), indicating that translocation rates are likely not affected by SunTag labeling of the nascent chain. Finally, we measured elongation rates of a shorter and codon-optimized reporter gene, which revealed a somewhat faster elongation rate of 4.9 codons/s (Figure S2G), indicating that elongation rates may differ on different transcripts. Using the elongation rate and ribosome density described above, we were able to estimate the translation initiation rate to be between $1.4 - 3.6 \text{ min}^{-1}$ on the Kif18b reporter (see Extended Experimental Procedures).

Together, these results provide the first *in vivo* measurements of the rates of ribosome initiation and translocation on single mRNA molecules in live cells.

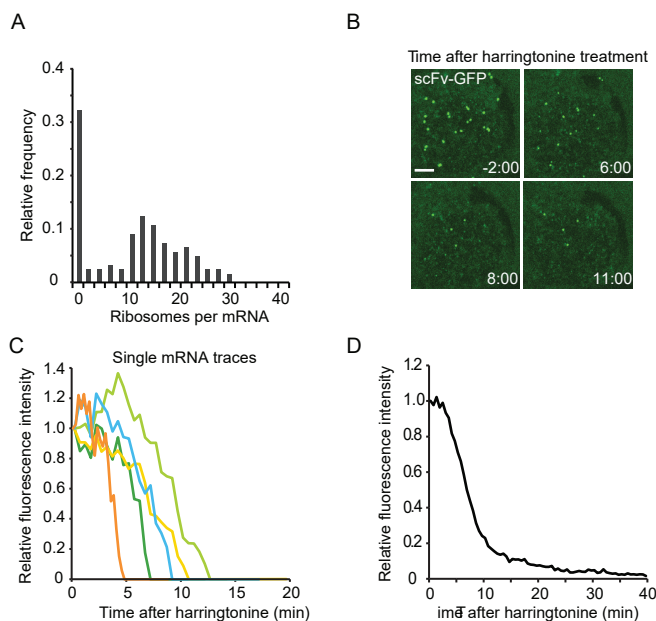


Figure 2: Measurements of ribosome initiation and elongation rates on single mRNA molecules.

U2OS cells expressing scFv-GFP, PP7-2xm-Cherry-CAAX and the translation reporter (SunTag_{24x}-Kif18b-PP7_{24x}) A) Distribution of the number of ribosomes bound to single mRNAs of the translation reporter (SunTag_{24x}-Kif18b-PP7_{24x}) (n = 2 independent experiments, 16 cells and 124 mRNAs), see Extended Experimental Procedures. B-D) *U2OS* cells expressing the translation reporter (SunTag_{24x}-Kif18b-PP7_{24x}) were treated with harringtonine at t = 0. B) Representative images from a time lapse movie. C) 5

representative traces of fluorescence decay on single mRNAs (of >100 analyzed). D) Normalized quantification of the decrease in fluorescence over time from many translation sites (n = 4 independent experiments, 37 cells, 536 mRNAs). Scale bars, 5 μm . See also Figure S2 and Movie S4.

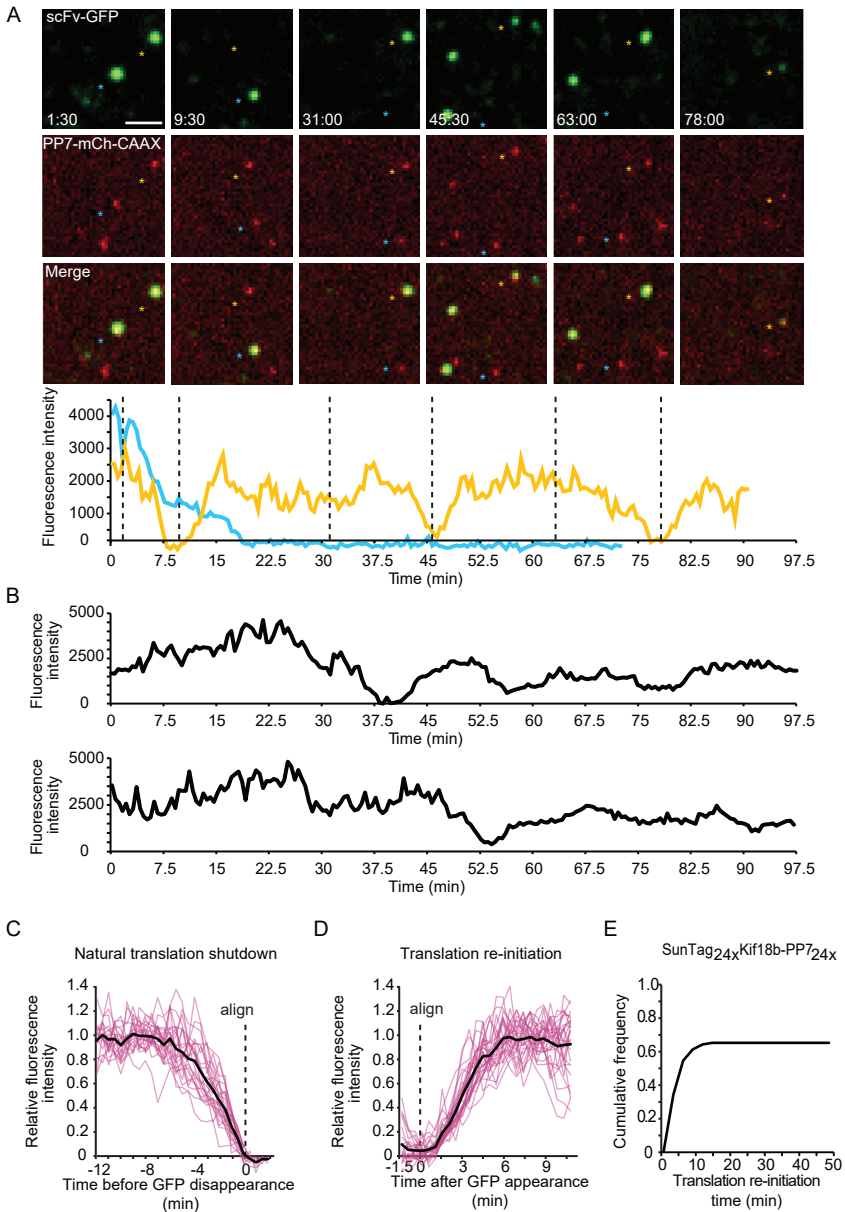


Figure 3. Long-term dynamics of translation of single mRNA molecules. U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the translation reporter (SunTag_{24x}-Kif18b-PP7_{24x}) A) U2OS cell expressing the SunTag_{24x}-Kif18b-PP7_{24x} reporter was imaged by time-lapse microscopy. Blue and yellow asterisks mark two different mRNAs undergoing changes in translation over time (upper panel). Intensity of scFv-GFP was measured over time for the two mRNAs (bottom panel). Colors of lines correspond to scFv-GFP intensity of translation sites marked by asterisk with the same color. B) ScFv-GFP intensity traces of two additional mRNA molecules. C) mRNAs undergoing permanent translation shutdown. Fluorescence

Temporal changes in translation of single mRNA molecules

To study translation over time, we imaged cells for 2 hrs and quantified the scFv-GFP signal from single mRNA molecules that could be tracked for >1 hr (Figure 3A, B and S3A). The results show considerable fluctuations in the translational state of individual mRNAs over time (Figure 3A, B and S3A). Such large fluctuations were not observed when cells were treated with the translation inhibitor CHX (Figure S3B), indicating they were due to changes in translation initiation and/or elongation rather than measurement noise. We also observed heterogeneity of behavior between different mRNAs. Some remained in a high translating state for >1 hr (e.g. Figure S3A, panels 12 and 13). Others shut down translation initiation and lost their scFv-GFP signal (e.g. Figure 3A, B and S3A, panels 1, 3-11, 14), which may account for the population of non-translating mRNAs observed in steady state measurements (Figure 2A). From the progressive decline in scFv-GFP fluorescence (Figure 3C and Movie S5), we could estimate a ribosome run-off rate of 3.0 codons/s (Figure 3C), which is similar to that measured after addition of harringtonine (3.5 ± 1.1 codons/s; Figure 2). Interestingly, a subset (67 of 104 mRNAs, 3 independent experiments, 19 cells) of these mRNAs later reinitiated translation and largely recovered their original scFv-GFP fluorescence (Figure 3A, B, D and S3A, panels 1, 3, 5, 8-10). Individual mRNAs even showed repeated cycling between non-translating and translating states (Figure 3A, yellow line, and S3A, panels 3, 5 and 8). Such cycles of complete translational shutdown and re-initiation occurred 0.29 ± 0.10 times per mRNA per hour ($n = 4$ independent experiments, 27 cells, 106 mRNAs), suggesting that most mRNAs will undergo one or more translational shutdown and re-initiation events in their lifetime. Thus, single mRNA imaging reveals reversible switching between translational shutdown and polysome formation.

After synchronized expression of the reporter construct using an inducible promoter, we often observed the initial binding events of newly transcribed mRNAs to the PP7-mCherry at the membrane (Figure 4A, B). Of these initial binding events, 44% of the mRNAs was associated with scFv-GFP fluorescence, indicating that they had already begun translation. However, the majority, 56% of mRNAs, initially appeared at the membrane in a non-translating state and subsequently converted to a translating state, usually within 1-5 min (Figure 4C and Movie S6). These mRNAs are likely newly transcribed mRNAs that are translating for the first time, rather than mRNAs that have already undergone translation but transitioned temporarily to a non-translating

Figure 3 continued... *intensity quantification is shown ($n = 24$ mRNAs). Average (black line), and single traces (pink lines) are shown. D) mRNAs undergoing translation re-activation after shutdown. Average (black line), and single traces (pink lines) are shown ($n = 30$ mRNAs). E) Time to reappearance of the first scFv-GFP fluorescence from translation sites that underwent complete translational shutdown. ~60% of the mRNAs re-initiated translation after complete shutdown and did so within 10 min ($n = 104$ translational sites analyzed). Scale bar, 2 μ m. See also Figure S3 and Movie S5.*

state. In support of this argument, long term (>1 hr) imaging of single mRNAs reveals that mRNAs spend on average only 2.5% of their life-time in such a temporary non-translating state ($n = 4$ independent experiments, 27 cells, 106 mRNAs), which is not sufficient to explain the 56% non-translating mRNAs that appeared at the membrane after synchronized transcription of the reporter. Rapid initiation of translation on newly transcribed mRNAs was described recently (Halstead et al., 2015), but our assay additionally allows an analysis of polysome build-up on new mRNAs (Figure 4B). Our analysis of the increase in scFv-GFP fluorescence indicates that once the first ribosome begins chain elongation, additional ribosomes initiate translation with a rate indistinguishable from that on polysomes at steady state (see Extended Experimental Procedures). We also examined the rate of fluorescence recovery (corresponding to polysome buildup) after complete shutdown of translation and subsequent re-initiation (Figure 4D). The polysome build-up on new transcripts was comparable to that observed for mRNAs that were cycling between translating and non-translating states (Figure 4D).

Ribosome Stalling

Several studies reported that ribosomes can pause or stall at a defined nucleic acid sequence with a regulatory function (Walter and Blobel, 1981; Yanagitani et al., 2011),

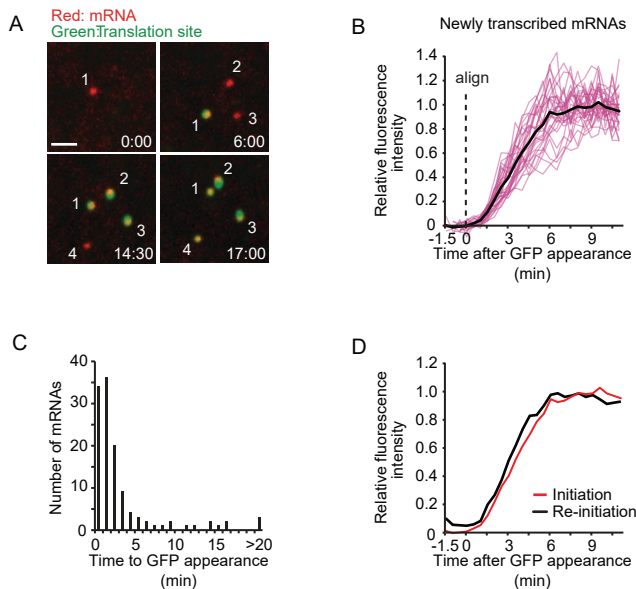


Figure 4. Analysis of polysome build-up on newly transcribed mRNAs. U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the translation reporter (*Sun-Tag_{24x}-Kif18b-PP7_{24x}*). **A**) Images from a time-lapse movie of newly transcribed mRNAs undergoing the first rounds of translation. **B**) Quantification of the fluorescence intensity increase, aligned at the first time-point at which scFv-GFP signal was detected ($n = 30$ individual mRNAs (pink lines) and average (black line) are shown). **C**) Quantification of the time between mRNA appearance and the first detection of translation by scFv-GFP fluorescence. **D**) Comparison of scFv-GFP fluorescence buildup on either new transcripts (red line) or on re-initiating mRNAs (black line). Data is re-plotted from figures 3D and 4B. Scale bar, 2 μ m. See also Movie S6.

Scale bar, 2 μ m. See also Movie S6.

at chemically modified or damaged nucleotides (Simms et al., 2014), or at regions in the RNA with a strong secondary structure (Tholstrup et al., 2012; Wen et al., 2008). We found that a subset (~5-10%) of mRNAs retained a bright scFv-GFP signal 15 min after harringtonine treatment (Figure 2B, D), a time at which ribosomes translocating at ~3 codons/s should have finished translating the reporter. A similar percentage of stalled ribosomes was observed on two additional reporter transcripts, both of which were designed using optimal codon usage (Figures S2G and S4A). Ribosome stalling also was observed using hippuristanol (Figure S4B), a translation initiation inhibitor with a different mechanism of inhibition (Bordeleau et al., 2006), indicating that the stalling was not caused by harringtonine. We also observed stalls when examining ribosome run-off from non-tethered cytosolic mRNAs lacking PP7 binding sites (Figure S4C). Importantly, stalls were not observed after puromycin treatment (Figure S4D, E) and the prolonged (>15 min) scFv-GFP signal on mRNAs from harringtonine-treated cells rapidly disappeared upon the addition of puromycin, confirming that the observed signal indeed represents stalled ribosomes (Figure S4F). The majority of mRNAs with stalled ribosomes (33 of 43) could be tracked for > 40 min, the typical duration of our harringtonine run-off experiments, indicating that they were not readily targeted by the no-go mRNA decay machinery within this time-frame.

Ribosome stalls could be due to defective ribosomes causing roadblocks on the mRNA or by defects in the mRNA. These models can potentially be distinguished by examining how such stalls are resolved; a single defective ribosome will inhibit ribosome run-off until the stalled ribosome is removed, after which the remaining ribosomes will run-off at a normal rate. In contrast, if the stalls are caused by defects to the mRNA, such as chemical damage, then each ribosome passing over the damaged nucleotide will be delayed, resulting in an overall slower scFv-GFP decay rate (Figure 5A). Long-term tracking of stalled ribosomes on single mRNAs was consistent with the latter model, indicating that ribosome stalling is likely caused by defective mRNA (Figure 5B). Consistent with the hypothesis that chemical damage to mRNA causes ribosome stalling, treatment of cells with 4-nitroquilone-1-oxide (4NQO), a potent nucleic acid damaging agent which causes 8-oxoguanine modifications and stalls ribosomes *in vitro* (Simms et al., 2014), resulted in a slow run-off on the majority of mRNAs, indicating widespread ribosome stalling (Figure 5C). Thus, chemical damage to mRNAs stalls ribosome elongation *in vivo*.

Regulated ribosome pausing occurs both *in vitro* and *in vivo* at asparagine 256 in the stress related transcription factor Xbp1 (Ingolia et al., 2011; Yanagitani et al., 2011), and this ribosome pausing is important for membrane targeting of the mRNA (Yanagitani et al., 2011). To test whether our translation imaging system could recapitulate such translation pausing, we introduced a strong ribosome pausing sequence (a point mutant of the wildtype Xbp1 pausing sequence that shows enhanced ribosome pausing (Yanagitani et al., 2011)) into the 3' region of the coding sequence of our reporter

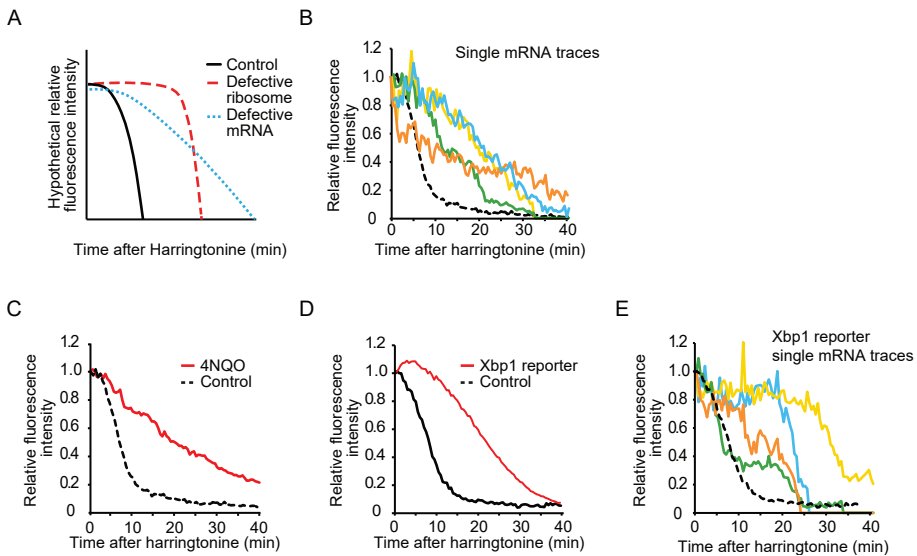


Figure 5. Dynamics of ribosome stalling. U2OS cells expressing *scFv-GFP*, *PP7-2xmCherry-CAAX* and the *SunTag_{24x}-Kif18b-PP7_{24x}* translation reporter (A-C) or the *Xbp1* translation reporter (D-E). A-B) Ribosome stalling likely results from mRNA defects, model (A) and experiment (B). B) Fluorescence intensity over time is shown for 4 representative stalled translation sites (colors) (of 20 analyzed). Since intensity values of single mRNAs were derived from the experiments presented in figure 2D, the average fluorescence decay presented in figure 2D is re-plotted here for comparison (dashed black line). C) Nucleic acid damage through 4NQO treatment (red line) induces ribosome stalling ($n = 3$ independent experiments, 40 cells, 455 mRNAs). For comparison, the harringtonine run-off from control cells with the *SunTag_{24x}-Kif18b-PP7_{24x}* reporter from Figure 2D is re-plotted, as these experiments were performed in parallel. D-E) Harringtonine run-off for the *Xbp1* pause site (red line, $n = 3$ independent experiments, 31 cells, 990 mRNAs) and control reporter (black dashed lines, $n = 3$ independent experiments, 27 cells, 437 mRNAs). See also Figure S4.

(hereafter referred to as *Xbp1* reporter). Harringtonine ribosome run-off experiments on the *Xbp1* reporter revealed a delay in ribosome run-off (Figure 5D), confirming that our reporter faithfully reproduced the ribosome pausing phenotype. To study the behavior of individual ribosomes on the *Xbp1* ribosome pausing sequence, we tracked single mRNAs during ribosome run-off. Surprisingly, the fluorescence decay was not linear, as would be expected if each ribosome paused a similar amount of time on the pause site. Rather, fluorescence decay occurred in bursts interspaced with periods in which no decay was detectable (Figure 5E, 4 representative traces shown out of 25 analyzed). These results indicate that most ribosomes are only briefly delayed at the *Xbp1* pause site, but a small subset of ribosomes remain stalled for an extended (> 10 min) period of time, explaining the strong ribosome stalling phenotype observed in ensemble experiments.

Translational regulation of the cell cycle regulator Emi1

We also applied our assay to study the transcript-specific translational regulation of Emi1, a key cell cycle regulatory protein. Our recent work reported strong translational repression of Emi1 during mitosis and found that the 3'UTR of Emi1 is involved in this regulation (Tanenbaum et al., 2015), but a role of its 5'UTR in translational regulation was not established. Interestingly, Emi1 has at least two splicing isoforms that differ in their 5'UTR sequence; NM_001142522.1 (hereafter referred to as 5'UTR_long) and NM_012177.3 (hereafter referred to as 5'UTR_short) (Figure 6A). We found that a GFP protein fused downstream of the 5'UTR_long was expressed at 40-fold lower levels than a GFP fused to the 5'UTR_short (Figure 6B). Such difference in protein expression could be due to a difference in transcription rate, mRNA stability, or reduced translation initiation or elongation rates. To distinguish between these possibilities, we prepared translation reporter constructs bearing either the short or long 5'UTR of Emi1. Robust translation was observed on ~50% of mRNAs encoding the short 5'UTR (Figure 6C). In contrast, the majority (~80%) of transcripts encoding the Emi1 5'UTR_long showed no detectable translation (not shown), and of the translating mRNAs, only very weak scFv-GFP fluorescence was usually detected (Figure 6C). Surprisingly, however, a very small fraction of mRNAs containing the 5'UTR_long (~2%) was associated with a bright scFv-GFP signal (Figure 6C, >92 bin), indicating they are bound to many ribosomes. This was not due to ribosome stalling and subsequent (slow) accumulation of ribosomes on a subset of mRNAs, as this bright scFv-GFP signal rapidly dissipated upon harringtonine treatment (Figure S5), indicating that these mRNAs were translated at high levels. Calculation of the total number of ribosomes associated with the mRNAs based upon scFv-GFP fluorescence intensity, revealed that 52% of all ribosomes translating the Emi1 5'UTR_long reporter were associated with the minor (2%) fraction associated with the highest scFv-GFP intensity. These results indicate that the great majority of 5'UTR_long transcripts are strongly translationally repressed but that a small subset of these mRNAs escape repression and undergo robust translation. Thus, substantial heterogeneity in translational efficiency can exist among different mRNA molecules within the same cell.

Observation of translation by single ribosomes

Interestingly, with the Emi1 5'UTR_long reporter, we often observed the abrupt appearance of a weak scFv-GFP signal on a transcript that was previously translationally silent. The GFP signal initially increased over time, plateaued, and then was abruptly lost after 6-8 min (Figure 7A-C and Movie S7). This type of signal is best explained by a single ribosome sequentially decoding the 24 SunTag peptides on the mRNA, followed by the release of the newly synthesized polypeptide upon completion of translation. Consistent with this hypothesis, the absolute fluorescence intensity of such translation

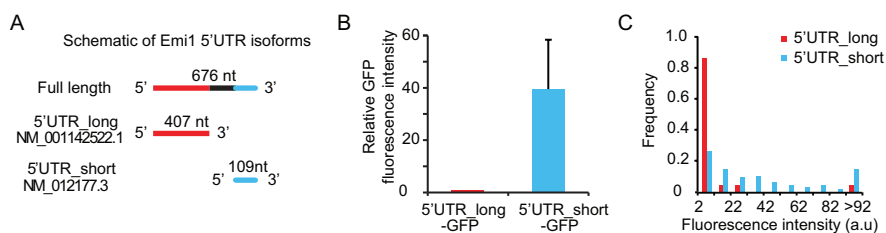


Figure 6. Differential control of translation initiation by two Emi1 splicing isoforms. A) Schematic of the 5'UTR of two Emi1 splicing isoforms. B) Fluorescence intensity of a GFP reporter under control of the two Emi1 isoforms (5'UTR_long and 5'UTR_short) expressed in HEK293 cells was measured by microscopy for single cells. Mean intensities were determined, which was corrected for background fluorescence in untransfected cells. At least 20 cells were measured per experiment per condition. Error bars, standard deviation between experiments. C) Fluorescence intensity distributions of single translation sites of indicated reporters ($n = 283$ mRNAs, 14 cells (5'UTR_long) and $n = 433$ mRNAs, 16 cells (5'UTR_short) from 3 independent experiments). Background from adjacent regions was subtracted. Only mRNAs are plotted that had translation signal above background (with an intensity value of more than 2; 16 and 53% of mRNAs for 5'UTR_long and 5'UTR_short, respectively). See also Figure S5.

events at the plateau phase (when all 24 SunTag peptides have been synthesized) was very similar to the intensity of a single fully synthesized SunTag24x-Kif18b protein (Figure S6A, B). The duration of the scFv-GFP signal per translation event could be converted to a translocation speed of single ribosomes (See Extended Experimental Procedures), which revealed an average elongation rate of 3.0 codons/s (Figure 7D). This value is similar to that determined from our bulk measurements of harringtonine-induced ribosome run-off or natural translational initiation shut-down and runoff (3-3.5 codons/s), indicating that ribosome elongation was not affected by the Emi1 5'UTR_long. Comparison of translocation rates obtained from single ribosome translation events also revealed heterogeneity in the decoding speed of individual ribosomes *in vivo* (Figure 7D).

DISCUSSION

Using the SunTag system, we have developed an imaging method that measures the translation of individual mRNAs in living cells. Immobilization of mRNAs on the plasma membrane allows the long-term (>1 hr) observation of translation of single mRNA molecules, which enables analyses of translational initiation, elongation and stalling in live cells for the first time. Under conditions of infrequent translational initiation, we can even observe a single ribosome decoding an entire mRNA molecule. Our observations reveal considerable and unexpected heterogeneity in the translation

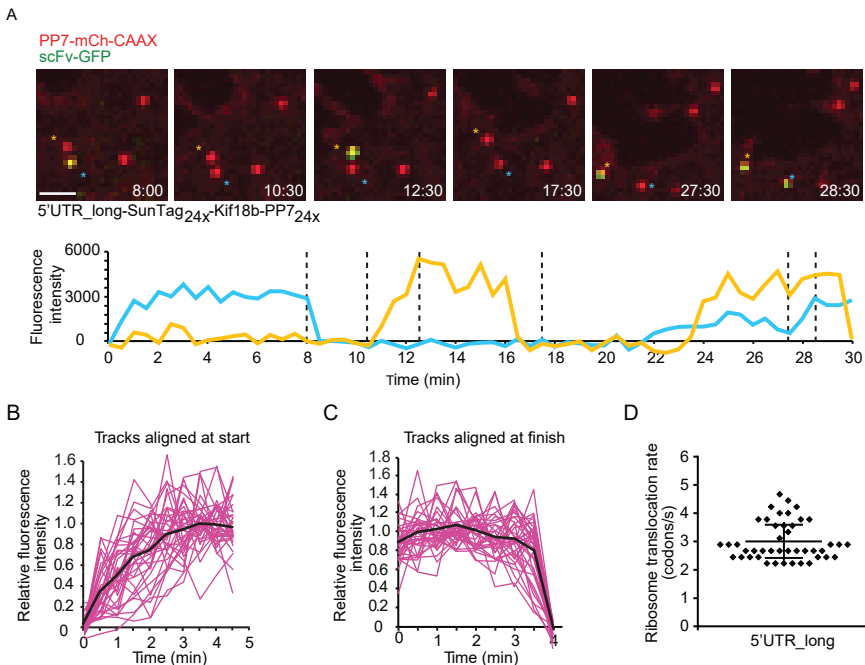


Figure 7. Visualizing single ribosomes decoding an mRNA molecule. A-D), Analysis of single ribosomes on the *Emi1* 5'UTR_long reporter mRNA. A) Representative images of multiple single ribosome translation events of individual mRNAs (upper panel). ScFv-GFP intensity was quantified over time for the two mRNAs marked by asterisks with the same color (bottom panel). B) Increase in scFv-GFP fluorescence from single ribosome translation events aligned at the first detectable scFv-GFP signal ($n = 35$ individual mRNAs in pink and average in black). C) Steady state and then abrupt decrease in scFv-GFP fluorescence from single translating ribosomes ($n = 35$ individual mRNAs in pink and average in black). D) Single ribosome elongation rates ($n = 44$) (see Extended Experimental Procedures). Mean \pm SD is shown in (D). Scale bar, $2 \mu\text{m}$. See also Figure S6 and Movie S7.

properties of different mRNA molecules derived from the same gene in a single cell, with some not translating, others actively translating with many ribosomes, and others bound to stalled ribosomes. The SunTag translation imaging assay should be applicable to many different cell types, including neurons and embryos where the localization and control of protein translation is thought to play an important role in cell function.

Comparison of methods to study translation in vivo

Ribosome profiling, a method in which fragments of mRNAs that are protected by the ribosome are analyzed by deep sequencing (Ingolia et al., 2009), has found widespread use in measuring translation. The strength of ribosomal profiling lies in its ability to measure translation on a genome-wide scale of endogenous mRNAs. However, a

limitation of ribosome profiling is the need to pool mRNAs from many thousands of cells for a single measurement. Thus, ribosome profiling in its present form cannot be used to study translation heterogeneity between different cells in a population or among different mRNA molecules in the same cell. Furthermore, since ribosome profiling requires cell lysis, only a single measurement can be made for each sample, limiting studies of temporal changes.

A number of single cell translation reporters have been developed based on fluorescent proteins (Aakalu et al., 2001; Brittis et al., 2002; Han et al., 2014; Raab-Graham et al., 2006; Tanenbaum et al., 2015; Tataavarty et al., 2012; Yu et al., 2006). Such reporters generally rely on the accumulation of new fluorescence after the assay is initiated. Advantages of these systems are that they are generally easy to use and have single cell sensitivity. However, they do not provide single mRNA resolution, often do not allow continuous measurement of translation, and do not report on ribosome initiation and elongation rates.

Finally, two methods were developed recently to image translation on single mRNAs *in vivo*. In one approach, the first round of translation is visualized (Halstead et al., 2015)). This method, however, does not allow continuous measurements of translation. The second approach involves measurements of the number of ribosomes bound to an mRNA using fluorescence fluctuation spectroscopy (Wu et al., 2015). The advantage of this method is that it can detect binding of a single fluorescent protein to an mRNA and different subcellular sites can be probed to study spatial differences in translation. The limitation of this method though is the inability to follow translation of single mRNAs over time, as these mRNAs cannot be tracked in the cell.

SunTag-based translation imaging assays are unique thus far in their ability to follow translation of individual mRNAs over time. This translation assay can be employed with either freely diffusing or tethered mRNAs, the choice of which will depend on the biological question to be addressed. In the study by Wang et al. (co-submitted), translation is observed in distinct spatial compartments in neurons using a similar SunTag-based translation imaging method with non-tethered mRNAs. In contrast, for studying ribosome translocation dynamics, the tethering assay provides the ability to track a single mRNA throughout the duration of the ribosome elongation cycle. Using this assay, we could measure polysome build-up rates over time, observe mRNAs cycling between translating and non-translating states, uncover heterogeneity in translation initiation rates (*e.g.* with the Emi1 5'UTR) and even observe a single ribosomes translating an entire transcript. These measurements were aided by the vastly improved signal-to-noise of the tethered assay and the ability to easily track slowly diffusing tethered mRNAs for an hour or more. These long-term observations allowed us to discover that mRNAs can reversibly switch between a translating and non-translating state and high variability in pause duration at the Xbp1 site. Thus, the untethered and tethered SunTag assays provide means to study translation of single mRNA molecules, which

will be applicable to a wide variety of biological questions and will be complementary to existing methods of studying translation.

A drawback of our assay is the need to insert an array of SunTag peptide repeats into the mRNA of interest to fluorescently label the nascent polypeptide, and to insert an array of PP7 binding sites in the 3'UTR to label the mRNA. As is true of any tagging strategy, these modifications could interfere with translation and/or mRNA stability under certain conditions. We have performed a number of control experiments to ensure that binding the scFv-GFP to the nascent chain and tethering of the transcript to the membrane do not grossly perturb translation (*e.g.* Figure 1B, F). We have also shown that ribosome translocation rates and ribosome density are similar when using a reporter with a very short (5X) or long (24X) SunTag peptide array and comparing tethered and non-tethered mRNAs (Figure 2D, S2F and S4C), indicating that many aspects of translation are not perturbed in our assay. Nevertheless, tethering of certain mRNAs to the plasma membrane may influence translation, especially for those mRNAs that undergo local translation in a specific compartment of the cell. Thus, our assay has unique advantages for certain types of measurements of translation, but appropriate controls should be performed for each experimental system or objective.

Heterogeneity in translation of single mRNAs: possible molecular mechanisms

Using our system, we measured the ribosome translocation speed on single mRNA molecules. Ribosome translocation rates have been measured in bulk previously in mouse embryonic stem cells (Ingolia et al., 2011), which yielded a translocation rate of 5.6 codons/s. Our values of 3-5 codons/s are in general agreement with those published values, and very similar to those measured by Wang et al (4 codons/s; co-submitted manuscript). Our experiments, and those of Wang et al. are the first to measure ribosome translocation rates for a single mRNA species, in single cells and on single mRNAs, which provides new opportunities to study regulation of translation elongation.

We also found that translation initiation can shut down temporarily on individual mRNAs and rapidly restart (Figure 3). Such shutdown of translation initiation could be due to transient loss of eIF4E binding to the mRNA cap, mRNA decapping followed by recapping (Mukherjee et al., 2012), or transient binding of regulatory proteins. Using our mRNA tethering assay, binding and unbinding of single proteins to translating mRNA could potentially be observed using Total Internal Reflection Fluorescence (TIRF), which could open up a many additional possibilities for studying translational regulation at the single molecule level.

The pioneer round of translation, the first ribosome to initiate translation on a newly transcribed mRNA, may be an especially important, as it is thought to detect defects in the mRNA, including premature stop codons (Ishigaki et al., 2001). A recently developed translation biosensor can detect the location of this pioneer round of

translation (Halstead et al., 2015). However, what happens after the first ribosome initiates translation is unknown. We found that the translation initiation rate on our reporter mRNA was similar on newly transcribed, recently shut down and re-initiating mRNAs and polysomal mRNAs (Figure 4; See Extended Experimental Procedures), indicating that the initiation rate is independent of the number of ribosomes bound to the mRNA. The presence of introns in a gene may also affect translation initiation on newly transcribed mRNAs (Le Hir et al., 2016), which could be tested in future studies. A subset of ribosomes stall on mRNAs in a sequence-independent fashion (Figure 2D, S2G and S4A). One possible explanation for this is that ribosome stalling is caused by naturally occurring mRNA “damage” (*i.e.* chemical modifications of the nucleotides). Previous studies have found that the 8-oxoguanine modification occurs on mRNA *in vivo*, and such modifications cause ribosome stalling *in vitro* (Simms et al., 2014) and *in vivo* in cells treated with 4NQO (Figure 5C). Alternatively, while we have performed numerous control experiments (Figure 5 and S4), we cannot completely exclude that the observed stalling on a small subset of mRNAs is an artifact of our construct or assay. We also observe ribosome pausing in a sequence-dependent fashion on the pause site of the Xbp1 transcription factor. Such pausing had been observed previously in bulk measurements (Ingolia et al., 2011; Yanagitani et al., 2011), but our quantitative analysis of single mRNAs revealing a high degree of variability in ribosome pausing at this site.

Finally, we show that the 5'UTR sequence of one Emi1 transcript isoform severely inhibits translation initiation. A likely explanation for this effect is the presence of several upstream open reading frames (uORFs) in this sequence. Surprisingly, a small number of mRNA molecules encoding this 5'UTR do undergo high levels of translation. It is possible that highly translating mRNAs are generated through alternative downstream transcription start site selection, which generates an mRNA that lacks the repressive sequence (for example, the uORFs). Alternatively, translation could occur if the 5'UTR repressive sequence is cleaved off followed by recapping after transcription, or if a repressive protein factor dissociates or an inhibitory RNA secondary structure unfolds. Further studies will be required to distinguish between these possibilities.

In summary, here we have developed an imaging method that enables the measurement of ribosome initiation and translocation rates on single mRNA molecules in live cells. Future developments of this technology could include simultaneous observation of single translation factors or other regulatory molecules together with mRNAs and nascent polypeptides, which would provide a very powerful system to dissect the molecular mechanisms of translational control.

MATERIAL & METHODS

Cell culture and drug treatment

U2OS and HEK293 cells were grown in DMEM/5% with Pen/Strep. Plasmid transfections were performed with Fugene 6 (Roche) and stable transformants were selected with zeocin (Life Technologies). Unless noted otherwise, reporter transcripts were expressed from a doxycycline inducible promoter, and expression of the reporter was induced with 1 $\mu\text{g}/\text{ml}$ doxycycline (Sigma) for 1 hr before imaging. Harringtonine (Cayman Chemical) was used at 3 $\mu\text{g}/\text{ml}$. 5 μM 4NQO (Sigma) was added to cells for 1 hr before imaging. Puromycin (Life Technologies) was used at 100 $\mu\text{g}/\text{ml}$. Hippuristanol (a kind gift of dr. J. Tanaka) was used at 5 μM . Cycloheximide (Sigma) was used at 200 $\mu\text{g}/\text{ml}$.

Plasmid sequences

Sequences of constructs used in this study are provided in the Extended Experimental Procedures.

Microscopy

Cells were grown in 96-well glass bottom dishes (Matriplate, Brooks). Images were acquired using a yokogawa CSU-X1 spinning disk confocal attached to an inverted Nikon TI microscope with Nikon Perfect Focus system, 100x NA 1.49 objective, an Andor iXon Ultra 897 EM-CCD camera and Micro-Manager software (Edelstein et al., 2010). Single z-plane images were acquired every 30 s unless noted otherwise. During image acquisition, cells were maintained at a constant temperature of 36-37° C. Camera exposure times were generally set to 500 ms, unless noted otherwise. We note that stable expression of PP7-mCherry, either with or without the CAAX domain, also resulted in an accumulation of mCherry signal in lysosomes, but lysosomes could be readily distinguished from mRNA foci based on signal intensity and mobility.

FACS

GFP and scFv-GFP (Figure 1B) or mCherry, PP7-mCherry or PP7-2xmCherry-CAAX (Figure 1F) were expressed from a constitutive promoter, while the two reporters, SunTag_{24x}-mCherry and GFP-PP7_{24x} (Figure 1B and 1F, respectively) were expressed from an inducible promoter in U2OS cells expressing the Tet repressor protein and their expression was induced 24 hrs after transfection using doxycycline (1 $\mu\text{g}/\text{ml}$). This ensured that the reporters were translated in the presence of high levels of the scFv-GFP and PP7-2xmCherry-CAAX proteins. Cells were collected one day after doxycycline induction and analyzed by FACS. Cells were gated for GFP and mCherry double positivity, and the mCherry and GFP levels (Figure 1B and 1F, respectively)

were analyzed using Flowjo v10.1.

Image analysis and quantification

For detailed description of Image analysis and quantification, see Extended Experimental Procedures.

Supplemental information

Supplemental information includes Supplemental Experimental Procedures, seven figures, and seven movies and can be found online at <http://dx.doi.org/10.1016/j.cell.2016.04.034>

Author contributions

M.E.T. conceived of the project with input from R.D.V.; X.Y., T.A.H. and M.E.T. performed the experiments and analyzed the data. All authors interpreted the results. X.Y. developed the mathematical model. X.Y., M.E.T. and R.D.V. wrote the manuscript with input from T.A.H..

Acknowledgments

We would like to thank Nico Stuurman for developing the spot_counter plugin for imageJ and for general help with microscopy, dr. J. Tanaka for providing hippuristanol, and S. Ruijtenberg and L. Krenning for helpful discussions. This work was supported by a European Research Council (ERC) Starting grant (ERC-STG 677936-RNAREG). MET was supported by a fellowship of the Dutch Cancer Society (KWF), RDV was supported by the Howard Hughes Medical Institute.

REFERENCES

- Aakalu, G., Smith, W.B., Nguyen, N., Jiang, C., and Schuman, E.M. (2001). Dynamic visualization of local protein synthesis in hippocampal neurons. *Neuron* *30*, 489-502.
- Babendure, J.R., Babendure, J.L., Ding, J.H., and Tsien, R.Y. (2006). Control of mammalian translation by mRNA structure near caps. *Rna* *12*, 851-861.
- Blanchard, S.C. (2009). Single-molecule observations of ribosome function. *Current opinion in structural biology* *19*, 103-109.
- Bordeleau, M.E., Mori, A., Oberer, M., Lindqvist, L., Chard, L.S., Higa, T., Belsham, G.J., Wagner, G., Tanaka, J., and Pelletier, J. (2006). Functional characterization of IRESes by an inhibitor of the RNA helicase eIF4A. *Nature chemical biology* *2*, 213-220.
- Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T., and Weissman, J.S. (2012). High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* *335*, 552-557.
- Brittis, P.A., Lu, Q., and Flanagan, J.G. (2002). Axonal protein synthesis provides a mechanism for localized regulation at an intermediate target. *Cell* *110*, 223-235.
- Chao, J.A., Patskovsky, Y., Almo, S.C., and Singer, R.H. (2008). Structural basis for the coevolution of a viral RNA-protein complex. *Nature structural & molecular biology* *15*, 103-105.
- Chen, J., Tsai, A., O'Leary, S.E., Petrov, A., and Puglisi, J.D. (2012). Unraveling the dynamics of ribosome translocation. *Current opinion in structural biology* *22*, 804-814.
- Choi, J., Jeong, K.W., Demirci, H., Chen, J., Petrov, A., Prabhakar, A., O'Leary, S.E., Dominissini, D., Rechavi, G., Soltis, S.M., *et al.* (2016). N(6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nature structural & molecular biology* *23*, 110-115.
- Cornish, P.V., Ermolenko, D.N., Noller, H.F., and Ha, T. (2008). Spontaneous intersubunit rotation in single ribosomes. *Molecular cell* *30*, 578-588.
- Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010). Computer control of microscopes using microManager. *Current protocols in molecular biology* / edited by Frederick M Ausubel [et al] *Chapter 14*, Unit14 20.
- Elkon, R., Ugalde, A.P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. *Nature reviews Genetics* *14*, 496-506.
- Fei, J., Kosuri, P., MacDougall, D.D., and Gonzalez, R.L., Jr. (2008). Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation. *Molecular cell* *30*, 348-359.
- Gupta, I., Clauder-Munster, S., Klaus, B., Jarvelin, A.I., Aiyar, R.S., Benes, V., Wilkening, S., Huber, W., Pelechano, V., and Steinmetz, L.M. (2014). Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. *Molecular systems biology* *10*, 719.
- Halstead, J.M., Lionnet, T., Wilbertz, J.H., Wippich, F., Ephrussi, A., Singer, R.H., and Chao, J.A. (2015). Translation. An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science* *347*, 1367-1671.
- Han, K., Jaimovich, A., Dey, G., Ruggiero, D., Meyuhas, O., Sonenberg, N., and Meyer, T. (2014). Parallel measurement of dynamic changes in translation rates in single cells. *Nature methods* *11*, 86-93.
- Holt, C.E., and Schuman, E.M. (2013). The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* *80*, 648-657.

- Huttelmaier, S., Zenklusen, D., Lederer, M., Dichtenberg, J., Lorenz, M., Meng, X., Bassell, G.J., Condeelis, J., and Singer, R.H. (2005). Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* *438*, 512-515.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* *324*, 218-223.
- Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* *147*, 789-802.
- Ishigaki, Y., Li, X., Serin, G., and Maquat, L.E. (2001). Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* *106*, 607-617.
- Jovanovic, M., Rooney, M.S., Mertins, P., Przybylski, D., Chevrier, N., Satija, R., Rodriguez, E.H., Fields, A.P., Schwartz, S., Raychowdhury, R., *et al.* (2015). Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. *Science* *347*, 1259038.
- Katz, Z.B., English, B.P., Lionnet, T., Yoon, Y.J., Monnier, N., Ovrzyn, B., Bathe, M., and Singer, R.H. (2016). Mapping translation 'hot-spots' in live cells by tracking single molecules of mRNA and ribosomes. *eLife* *5*.
- Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y., and Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. *Nature* *467*, 103-107.
- Le Hir, H., Sauliere, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. *Nature reviews Molecular cell biology* *17*, 41-54.
- Leung, K.M., van Horck, F.P., Lin, A.C., Allison, R., Standart, N., and Holt, C.E. (2006). Asymmetrical beta-actin mRNA translation in growth cones mediates attractive turning to netrin-1. *Nature neuroscience* *9*, 1247-1256.
- Mukherjee, C., Patil, D.P., Kennedy, B.A., Bakthavachalu, B., Bundschuh, R., and Schoenberg, D.R. (2012). Identification of cytoplasmic capping targets reveals a role for cap homeostasis in translation and mRNA stability. *Cell reports* *2*, 674-684.
- Peshkin, L., Wuhr, M., Pearl, E., Haas, W., Freeman, R.M., Jr., Gerhart, J.C., Klein, A.M., Horb, M., Gygi, S.P., and Kirschner, M.W. (2015). On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development. *Developmental cell* *35*, 383-394.
- Raab-Graham, K.F., Haddick, P.C., Jan, Y.N., and Jan, L.Y. (2006). Activity- and mTOR-dependent suppression of Kv1.1 channel mRNA translation in dendrites. *Science* *314*, 144-148.
- Rodriguez, A.J., Shenoy, S.M., Singer, R.H., and Condeelis, J. (2006). Visualization of mRNA translation in living cells. *The Journal of cell biology* *175*, 67-76.
- Rojas-Duran, M.F., and Gilbert, W.V. (2012). Alternative transcription start site selection leads to large differences in translation activity in yeast. *Rna* *18*, 2299-2305.
- Schwanhauser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337-342.
- Simms, C.L., Hudson, B.H., Mosior, J.W., Rangwala, A.S., and Zaher, H.S. (2014). An active role for the ribosome in determining the fate of oxidized mRNA. *Cell reports* *9*, 1256-1264.
- Stumpf, C.R., Moreno, M.V., Olshen, A.B., Taylor, B.S., and Ruggero, D. (2013). The translational landscape of the mammalian cell cycle. *Molecular cell* *52*, 574-582.
- Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system

for signal amplification in gene expression and fluorescence imaging. *Cell* *159*, 635-646.

Tanenbaum, M.E., Macurek, L., van der Vaart, B., Galli, M., Akhmanova, A., and Medema, R.H. (2011). A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by Aurora kinases. *Current biology* : *CB* *21*, 1356-1365.

Tanenbaum, M.E., Stern-Ginossar, N., Weissman, J.S., and Vale, R.D. (2015). Regulation of mRNA translation during mitosis. *eLife* *4*.

Tatavarty, V., Ifrim, M.F., Levin, M., Korza, G., Barbarese, E., Yu, J., and Carson, J.H. (2012). Single-molecule imaging of translational output from individual RNA granules in neurons. *Molecular biology of the cell* *23*, 918-929.

Tholstrup, J., Oddershede, L.B., and Sorensen, M.A. (2012). mRNA pseudoknot structures can act as ribosomal roadblocks. *Nucleic acids research* *40*, 303-313.

Vardy, L., and Orr-Weaver, T.L. (2007). Regulating translation of maternal messages: multiple repression mechanisms. *Trends in cell biology* *17*, 547-554.

Walter, P., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *The Journal of cell biology* *91*, 557-561.

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* *161*, 1388-1399.

Wen, J.D., Lancaster, L., Hodges, C., Zeri, A.C., Yoshimura, S.H., Noller, H.F., Bustamante, C., and Tinoco, I. (2008). Following translation by single ribosomes one codon at a time. *Nature* *452*, 598-603.

Wu, B., Buxbaum, A.R., Katz, Z.B., Yoon, Y.J., and Singer, R.H. (2015). Quantifying Protein-mRNA Interactions in Single Live Cells. *Cell* *162*, 211-220.

Yanagitani, K., Kimata, Y., Kadokura, H., and Kohno, K. (2011). Translational pausing ensures membrane targeting and cytoplasmic splicing of XBP1u mRNA. *Science* *331*, 586-589.

Yu, J., Xiao, J., Ren, X., Lao, K., and Xie, X.S. (2006). Probing gene expression in live cells, one protein molecule at a time. *Science* *311*, 1600-1603.

Zaher, H.S., and Green, R. (2009). Quality control by the ribosome following peptide bond formation. *Nature* *457*, 161-166.

SUPPLEMENTAL FIGURES

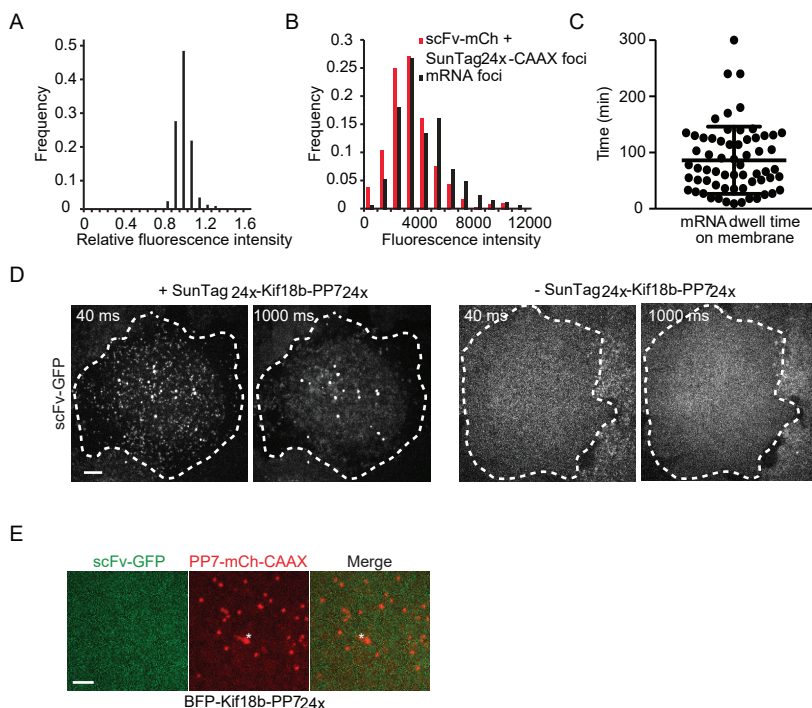


Figure S1. Related to figure 1. Validation of single molecule translation visualization assay. A-C) U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the SunTag_{24x}-Kif18b-PP7_{24x} translation reporter. A) Fluorescence intensity of mRNA foci was measured and was corrected for background fluorescence. The average corrected mRNA fluorescence intensity was set to 1 for each separate cell ($n = 3$ experiments, 14 cells, 278 mRNAs). B) Intensity of single mRNA foci was measured and corrected for background, but intensity was not normalized as in (A) to allow comparison of absolute intensities (Black bars, $n = 3$ independent experiments, 22 cells, 377 mRNAs). In parallel, U2OS cells co-expressing SunTag_{24x}-CAAX and scFv-mCherry were imaged and the intensities of single membrane bound scFv-mCherry-SunTag_{24x} foci was measured (Red bars, $n = 4$ independent experiments, 24 cells, 162 mRNAs). C) Dwell time of tethered mRNAs on the membrane. The time between mRNA appearance at the focal plane of the membrane and its disappearance was scored. mRNA disappearance was due to mRNA detachment or degradation, not photobleaching. Mean and SD are indicated. D) Cells expressing scFv-GFP with (left two images) or without (right two images) the SunTag_{24x}-Kif18b-PP7_{24x} reporter were imaged with indicated exposure time. Dotted line shows outline of the cell. E) U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the BFP-Kif18b-PP7_{24x} translation reporter. Representative image is shown. Asterisk indicates lysosome. Scale bars are 5 μm (D) and 2 μm (E).

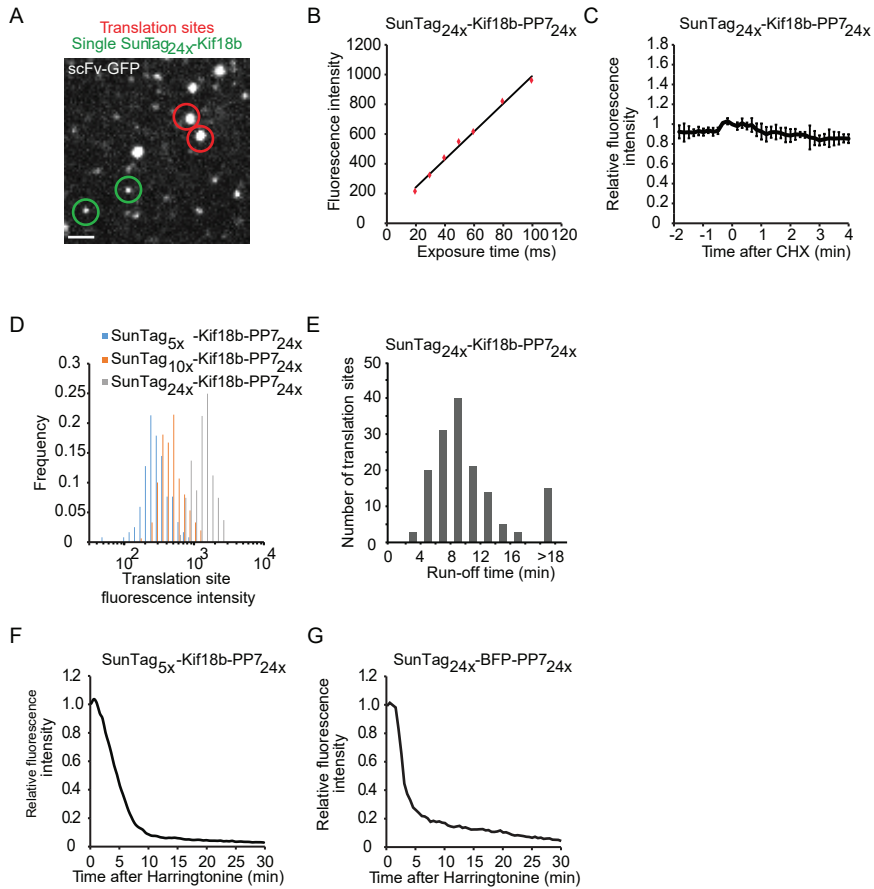


Figure S2. Related to figure 2. Quantification of ribosome number and elongation speed on single mRNAs. U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and indicated translation reporters. A) Images were acquired using short exposure times (40 ms), limiting motion blurring of fast moving particles, so both translation sites (red circle) and single, fully synthesized, freely diffusing SunTag proteins (green circles) could be observed as distinct foci. Fluorescence intensity of single SunTag_{24x}-Kif18b foci and single translation site was quantified in the same cell using a ROI with fixed size. (n = 45 translation sites, 15 cells, 3 experiments). B) To determine whether the exposure time of 40 ms used in (A) was sufficiently short to prevent a reduction in fluorescence intensity of foci due to motion blurring, we measured the intensity of single fully synthesized SunTag_{24x}-Kif18b foci at different exposure times. Fluorescence intensities of the ~25 brightest foci per image were measured. Results show a linear relationship between exposure time and fluorescence intensity at short exposure times, indicating that exposure times were short enough to prevent reduction in fluorescence intensity of foci due to motion blurring (n = 3 independent experiments, 18 cells and 400-500 spots). C) Cells were treated with 200 µg/ml CHX at t = 0 and fluorescence intensities of translation sites were measured over time. Note that fluorescence does not increase upon CHX treatment (n = 3 independent experiments, 31 cells, 209 mRNAs). Error bars indicate standard deviation. D) scFv-GFP fluorescence intensity of translation sites using reporters with varying numbers of SunTag peptides (5x, 10x and 24x)... Continued on next page.

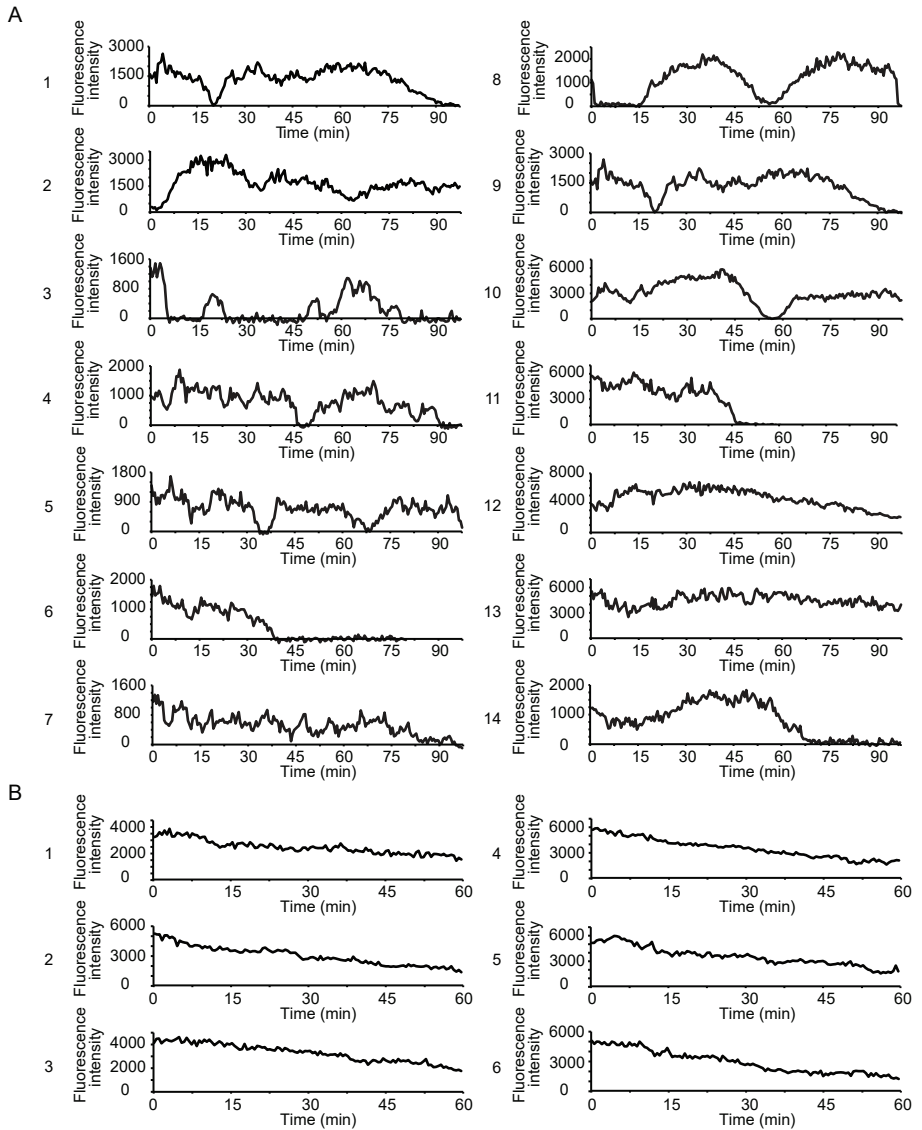


Figure legend on next page →

Figure S2 continued... ($n = 117, 149, 80$ translation sites for the 5x, 10x and 24x reporters, respectively) E-G) Cells were treated with harringtonine at $t = 0$ and translation site intensity was quantified over time. E) Histogram of the total run-off time, measured from the time of harringtonine treatment to the final disappearance of the scFv-GFP signal. 60 s was subtracted from all times to correct for the time required for harringtonine to enter the cell. F, G) ScFv-GFP fluorescence intensity was measured over time after harringtonine addition (F, $n = 3$ independent experiments, 39 cells, 1883 mRNAs) (G, $n = 3$ independent experiments, 30 cells, 378 mRNAs). Scale bar, $2 \mu\text{m}$.

Figure S3. Related to figure 3. Translation dynamics on single mRNA molecules. U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the translation reporter (SunTag_{24x} -Kif18b-PP7 $_{24x}$) were imaged by time-lapse microscopy for 2 hrs (A) or 1 hr (B) and the fluorescence intensity of single translation sites was tracked over time. 14 traces of untreated cells (A) or 6 traces of CHX treated cells (B) are shown. Note that the intensity of translation sites in CHX-treated cells slowly decreases over time, which is likely due to a decrease in the ribosome number per mRNA after prolonged CHX treatment.

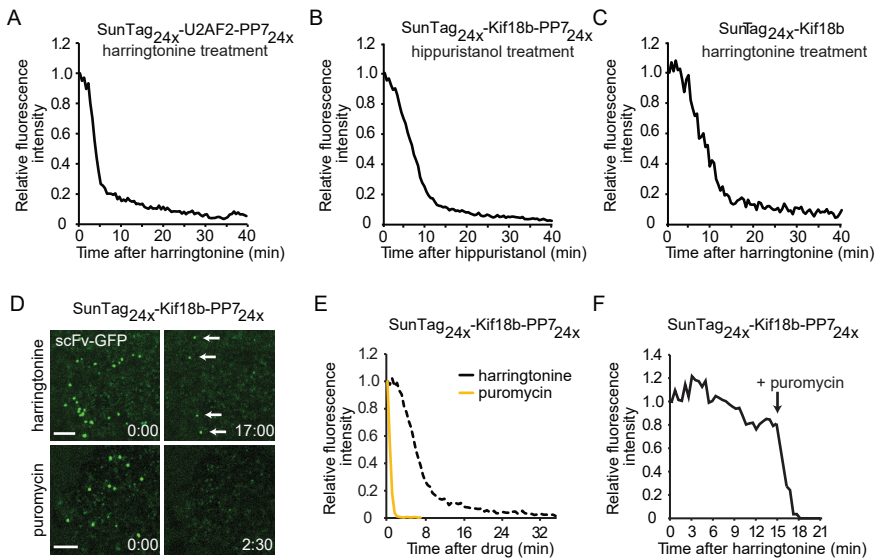


Figure S4. Related to figure 5. Validation of the ribosome stalling phenotype. A, C-F) U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and indicated translation reporters were treated with harringtonine at $t = 0$ and translation site intensity was quantified over time. A) Reporter containing a codon optimized version of the U2AF2 coding sequence ($n = 3$ independent experiments, 29 cells, 512 mRNAs). B) U2OS cells expressing the translation reporter (SunTag_{24x} -Kif18b-PP7 $_{24x}$) were treated with another translation initiation inhibitor (hippuristanol) and translation site intensity was quantified over time ($n = 2$ independent experiments, 14 cells, 515 mRNAs). C) Harringtonine run-off experiments were also performed on a translation reporter lacking PP7 binding sites (SunTag_{24x} -Kif18b) ($n = 2$ independent experiments, 19 cells, 248 mRNAs). D) Representative images in which stalled ribosome can be observed after harringtonine treatment (arrows). No ribosome stalling is observed after puromycin treatment (lower panel). E) At $t = 0$, either harringtonine (re-plotted from figure 2D) or puromycin ($n = 3$ independent experiments, 22 cells, 403 mRNAs) was added and translation site intensity was quantified over time. F) Sequential addition of harringtonine and then puromycin ($n = 7$ mRNAs). Scale bar, 5 μm .

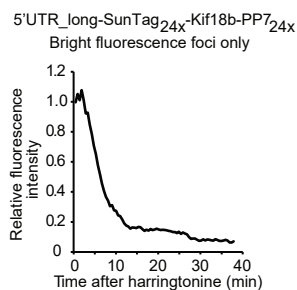


Figure S5. Related to figure 6. Ribosome elongation rates on Emi1 5'UTR_long containing mRNAs. U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the Emi1 5'UTR_long translation reporter were treated with harringtonine at $t = 0$. The fluorescence intensity of very bright translation sites was quantified over time ($n = 3$ independent experiments, 29 cells, 39 mRNAs).

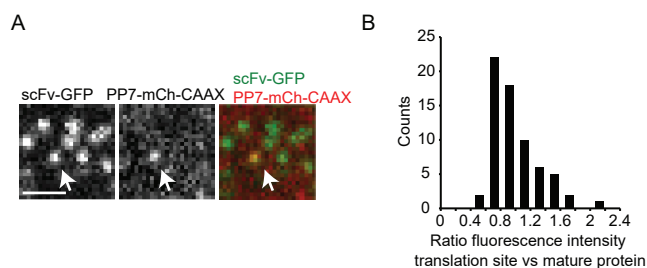


Figure S6. Related to figure 7. Single ribosomes translate the Emi1 5'UTR_long mRNA. U2OS cells expressing scFv-GFP, PP7-2xmCherry-CAAX and the Emi1 5'UTR_long translation reporter were imaged using a very short (30 ms) exposure time, so fully synthesized, freely diffusing mature SunTag-Kif18b molecules can be observed together with translation sites. Translation sites could be distinguished from fully synthesized SunTag molecules, as they co-migrated with mRNAs for multiple (> 5) consecutive time points. A) Representative image of a single translation site (arrow) surrounded by multiple mature SunTag molecules. B) Quantification of fluorescence intensities of translation sites and mature protein. The fluorescence intensity of a single translation sites was compared to the average fluorescence intensity of 5 nearby mature SunTag molecules.

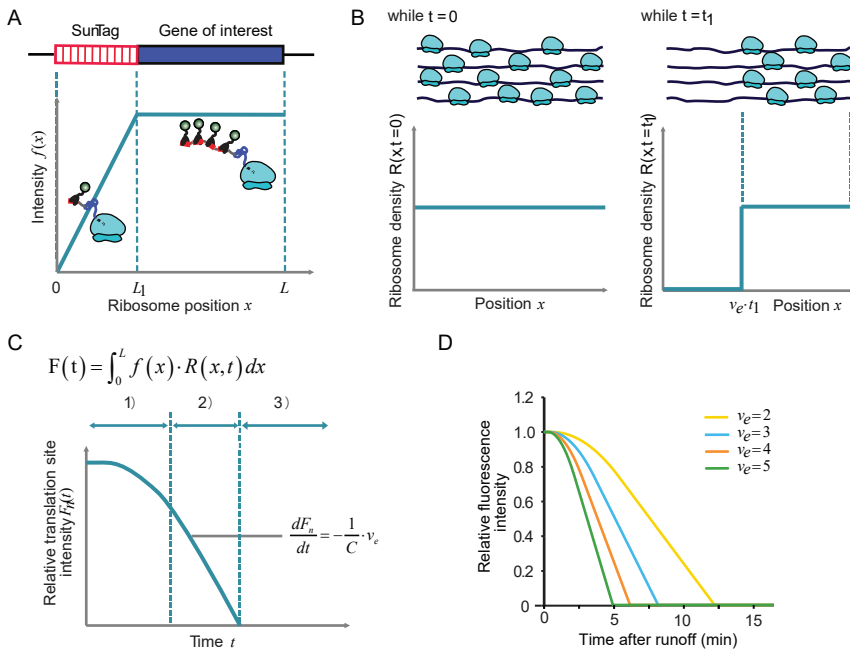
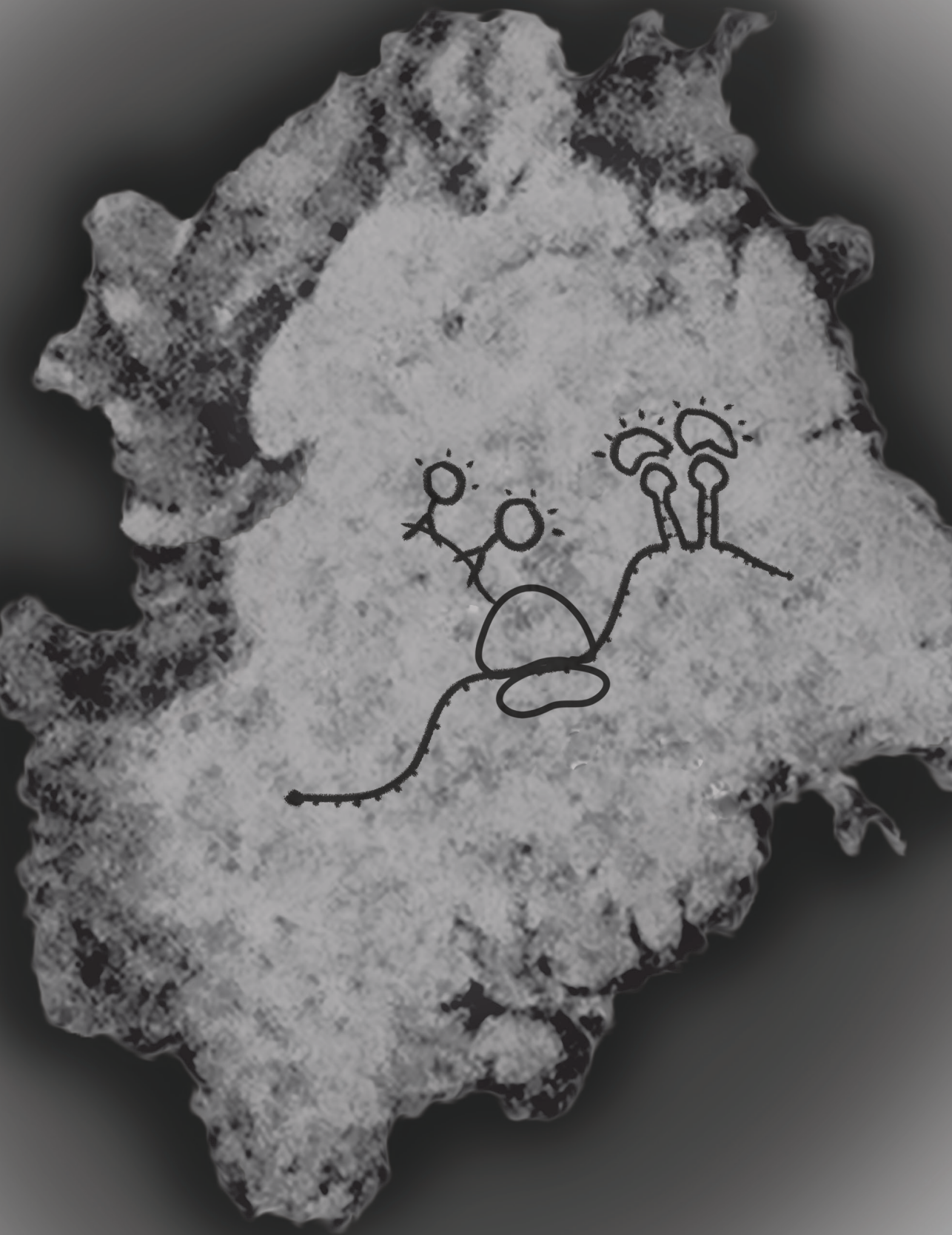


Figure S7. Related to Methods. Modeling of translation site intensity. A) Intensity from a single ribosome mainly depends on ribosome location on the mRNA. Due to the synthesis of SunTag peptides, ribosome intensity will increase initially as the ribosome moves towards the 3' end until SunTag peptides are fully synthesized and exposed. A typical curve for intensity function $f(x)$ is shown. For simplicity, a linear function was used to simulate the intensity increase. B) Ribosome density changes during ribosome run-off. When there are no new initiation events, already bound ribosomes will run off the mRNA from 5' to 3' end. Examples of the ribosome density function at $t=0$ and $t=t_1$ are shown. C) Translation site intensity is dependent on both intensity from single ribosomes as well as ribosome density throughout the mRNA. A formula describing translation site intensity is shown on top. A typical curve of intensity change during ribosome run-off process is shown at the bottom with three clear stages labeled using numbers. Intensity decreases linearly during the second stage, whose first order derivative could be used to derive elongation rate. D) Example results of simulations of harringtonine run-off from the Kif18b reporter (SunTag_{24x}-Kif18b), which were run using different elongation rates (2, 3, 4, 5 codons/s). Run-off starts at $t=0$.



Chapter 3

Imaging Translation Dynamics of Single mRNA Molecules in Live Cells

Suzan Ruijtenberg, Tim A. Hoek, Xiaowei Yan, Marvin E. Tanenbaum

ABSTRACT

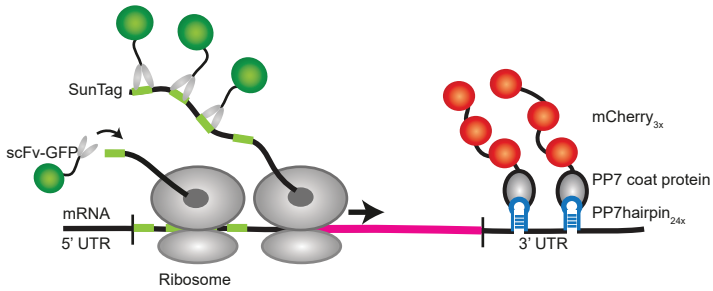
mRNA translation is a key step in decoding the genetic information stored in DNA. Regulation of translation efficiency contributes to gene expression control and is therefore important for cell fate and function. Here, we describe a recently developed microscopy-based method that allows visualization of translation of single mRNAs in live cells. The ability to measure translation dynamics of single mRNAs will enable a better understanding of spatio-temporal control of translation, and will provide unique insights into translational heterogeneity of different mRNA molecules in single cells.

1. INTRODUCTION

Translation of mRNAs into proteins is a key step in gene expression and of critical importance for fine-tuning cellular protein levels. In recent years, different methods have provided many new and important insights into the regulation of translation, yet many questions remain. For example, it is still unclear whether all mRNAs transcribed from the same gene are translated with similar efficiencies, or whether translational heterogeneity exists among such mRNAs. Similarly, it is largely unknown how translation efficiencies are controlled in space and time. An important reason for our limited understanding of translational control is that many current methods to assess translation efficiency rely on population-based measurements and frequently require fixation or lysis of the cells to obtain a measurement of translation efficiency. As a consequence, mainly snapshots of average translation efficiencies of thousands of mRNA molecules have been obtained. A major advance in measuring translation efficiency of single mRNAs in live cells was recently achieved by our lab, as well as several other labs (1, 2, 3, 4, 5).

Here, we describe a microscopy-based method, which allows quantitative measurements of ribosome initiation and elongation on individual mRNA molecules in live cells. This approach provides a powerful and widely applicable tool to study dynamics and regulation of translation. In this method, a reporter mRNA is designed which encodes a protein of interest (POI) fused at its N-terminus to an array of antibody peptide-epitopes, derived from the SunTag system that we previously developed (6). The SunTag peptide epitopes are recognized by a single chain antibody fragment fused to superfolderGFP (sfGFP) (6), which is co-expressed in cells with the reporter mRNA. When the reporter mRNA is translated, the peptide-epitopes emerge from the ribosome while the fused POI is undergoing synthesis (Figure 1A (upper panel)). Binding of the GFP-fused antibodies to the nascent peptide-epitopes results in a bright green labeling of the nascent polypeptide, which can be observed under the microscope as a bright fluorescent dot at the site of translation (Figure 1B), providing a real-time readout of the translation of the reporter mRNA. In addition to fluorescent labeling of the nascent polypeptide, the mRNA molecule is fluorescently labeled in a second color through the MS2- or PP7-based labeling system (7, 8) (Figure 1A (upper panel), B). In order to improve the long term tracking of mRNA molecules, we have devised an mRNA tethering system, which reduces mRNA mobility and allows tracking of individual mRNA molecules for extended periods of time (>1 hr) (Figure 1A (lower panel), B). In this chapter we provide details on how to design, carry out, and interpret experiments to image translation dynamics of an mRNA of interest and in a cell type of choice.

A Untethered Assay



Membrane Tethered Assay

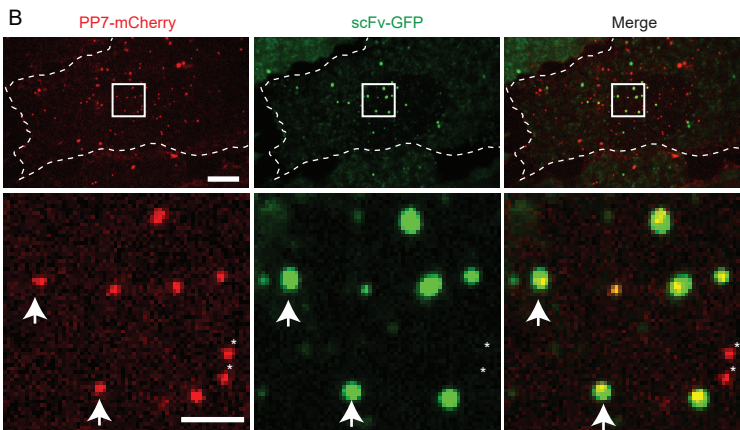
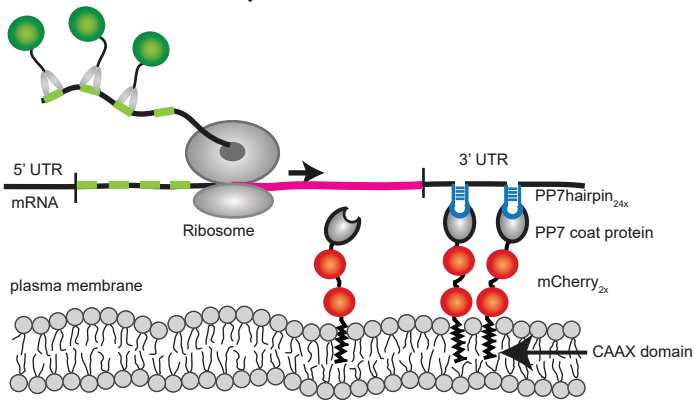


Figure 1. Imaging translation of single mRNA molecules. (A) Schematic representation of the fluorescence labeling of nascent polypeptides and the mRNA using the SunTag system and PP7 system, respectively. A gene of interest (pink) is located downstream of an array of SunTag peptides (green stripes). When translated, the peptide-epitopes will emerge from the ribosome first, while the gene of interest is still undergoing translation. The scFv-GFP antibody will bind to these newly synthesized SunTag peptides

2. MATERIALS

2.1 Plasmids

1. Translational reporter, containing the 24 SunTag peptides and 24 PP7 binding sites (For example, Addgene #74928)
2. scFv-GFP (For example, Addgene #60907)
3. PCP-mCherry (For example, Addgene #74926 or #74925)

2.2 Cell culture

1. Glass bottom cell culture dishes suitable for live-cell microscopy. Most high magnification microscope objectives are designed for glass with a thickness of 0.17 mm. We routinely use 96-wells glass bottom dishes (Matriplate, Brooks)
2. Cell type specific cell culture medium
3. Live-cell imaging cell culture medium; we use Leibovitz's-L15 imaging medium (Gibco, life technologies)
4. Transfection reagent: Fugene (Promega)

2.3 Small molecules useful for translation imaging

1. Doxycycline (stock solution of 10 mg/ml in DMSO, used at a final concentration of 1 $\mu\text{g}/\text{ml}$)
2. Puromycin (stock solution of 10 mg/ml in DMSO, used at final concentration of 100 $\mu\text{g}/\text{ml}$)

Figure. 1 continued... as soon as they emerge from the ribosome, resulting in a bright green fluorescence signal at the site of translation, allowing live observation of protein synthesis. To visualize the mRNA independently of translation, 24 PP7 binding sites (blue) are inserted in the 3' UTR of the mRNA. These sites are recognized by the PP7 bacteriophage coat protein, which is fused to 3 copies of mCherry and expressed in the same cell. As a result, the mRNAs can be observed as mCherry positive foci. By fusing a CAAX-motif to PP7-mCherry (lower panel), mRNAs can be tethered to the plasma membrane, which allows tracking of individual mRNAs for longer time periods. (B) A representative U2OS cell is shown expressing scFv-GFP, PP7-2xmCherry-CAAX and the translational reporter (SunTag_{24x}-kif18b-PP7_{24x}). mRNAs are visible in red, sites of translation in green. The dotted line indicates the outline of a cell. A zoomed-in view of the white-boxed area is shown below. Arrows indicate examples of mRNAs undergoing translation and asterisks indicate mRNAs which are not translated. Scale bars, 5 μm (upper panels) or 2 μm (lower panels).

3. Cycloheximide (stock solution of 50 mg/ml in DMSO, used at a final concentration of 200 µg/ml)
4. Harringtonine (stock solution of 3 mg/ml in DMSO, used at a final concentration of 3 µg/ml).

2.4 Microscopy

1. Either a wide-field, confocal, spinning disk confocal or Total Internal Reflection Fluorescence (TIRF) microscope containing a 40x, 60x or 100x objective.

3. METHODS

3.1 Plasmids and Plasmid design

In order to visualize single mRNAs and their translation by the method described in this protocol, three different plasmids are required (all plasmids are illustrated in Figure 2 and available on Addgene):

3.1.1 The translation reporter

(Available from Addgene: pcDNA4TO-24xGCN4_v4-kif18b-24xPP7 (#74928))

In general, a translation reporter consists of several elements; the SunTag peptide array, a sequence encoding a POI (see Note 1 for further discussion on choosing the POI), and an array of 24 PP7 binding motifs.

1. Clone the POI in the translational reporter available on Addgene (#74928). Using the enzymes AgeI and EcoRV the original POI can be removed (a fragment of 2587 bp) and replaced by any POI to create a new translational reporter (see Note 1 for further discussion on choosing the protein of interest).
2. Alternatively, start with a vector of choice and a POI of choice and clone the individual elements of the translational reporter (i.e. SunTag and PP7 binding sites) into this vector.
 - a. Since the SunTag peptide array forms a repetitive sequence, it is difficult to amplify the array by PCR. We therefore recommend cloning strategies based on enzymatic digestion and ligation (see Note 2 for more information about PCR-based cloning strategies of the SunTag array). To insert the SunTag array (consisting of 5-24 SunTag peptides (see Note 3 about the use of different numbers of SunTag pep-

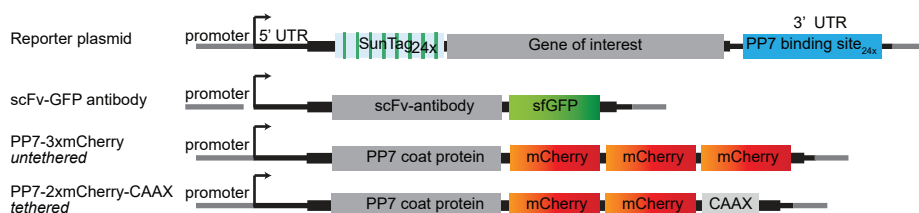


Figure 2: Schematic overview of the plasmids used for the translation imaging method. The SunTag peptides are shown in light blue with green stripes representing individual peptides (8 shown), the PP7 binding sites are shown in dark blue, sfGFP in green and mCherry in red. All plasmids are available on Addgene (see text for catalog number).

tides)) at the N-terminus of the POI, use HindIII and AgeI restriction enzymes to digest the translational reporter plasmid available on Addgene (#74928). This results in two fragments (9141 bp and 1819 bp). The smaller fragment contains the 24x SunTag peptides, which can be cloned into the desired vector.

- b. Clone the 24 PP7 binding motifs into the 3' UTR of the mRNA to label the mRNA independently of translation. Use BamH1 and EcoR1 to digest the translational reporter plasmid available on Addgene (#74928). This results in two fragments (9492 bp and 1468 bp), the smaller of which contains the PP7 binding motives. Since this array of short hairpin sequences is highly repetitive, we recommend cloning methods based on digestion and ligation rather than through PCR-based cloning.
- c. Clone a promoter of choice upstream of the reporter coding sequence. Expression of the reporter mRNA is typically driven by a doxycycline-inducible promoter to allow temporal control of reporter mRNA expression (see Note 4 about the use of an inducible promoter). This promoter can be obtained from plasmid (#74928), by using the enzymes MluI and HindIII or by PCR.

3.1.2 sfGFP-tagged antibody that binds the SunTag peptide (scFv-GFP)

(Available from Addgene: pHR-scFv-GCN4-sfGFP-GB1-dWPRES, #60907)

The scFv-GFP binds with high affinity to the SunTag peptides, which results in the fluorescent labeling of nascent polypeptides as soon as they emerge from the ribosome exit tunnel (2-5).

1. PCR amplify the scFv-GFP coding sequence and clone it into a plasmid containing the promoter appropriate to your cell type of choice.

3.1.3 *mCherry-fused PCP (PP7-mCherry)*

(Available from Addgene: pHR-tdPP7-3xmCherry, #74926, or pHR-PP7-2xmCherry-CAAX, #74925)

Dependent on whether or not the aim is to tether the mRNAs to the plasma membrane, plasmid #74926 (untethered) or plasmid #74925 (tethered) can be used as a template. Using PCR-based methods, either PP7 alone, PP7 fused to mCherry or PP7 fused to mCherry and a CAAX domain can be amplified and placed into a vector of choice (see Note 5 on tethering the mRNA to the membrane). Note that plasmids containing multiple copies of mCherry cannot be amplified by PCR and need to be cloned by digestion-ligation methods.

3.1.4 *Exchanging fluorescent proteins*

The three plasmids described above enable visualization of both the reporter mRNA (PP7-mCherry, red) and its translation (scFv-GFP, green) in live cells (Figure 1B). In principal, the color of the fluorescent proteins (e.g. GFP and mCherry) can be changed, but the functionality of the newly designed constructs needs to be carefully tested, as, for example, addition of sfGFP to the antibody has been shown to be important for preventing scFv aggregation in mammalian cells (6) (see Note 6 for further comments on fluorescent proteins fused to the scFv antibody).

1. Exchange fluorescent protein using PCR-based cloning methods.
2. Test expression level of newly created fusion protein by transient transfection in cell type of choice.
3. Examine aggregation state in cells of fusion protein after transient transfection using widefield or confocal microscopy (bright fluorescent foci in transfected cells indicate protein aggregation).

3.2 **Creating a cell line for imaging translation.**

3.2.1 *Choosing a suitable cell type.*

We have performed the majority of our experiments in U2OS cells. However, similar SunTag-based translation imaging has been successfully performed in neurons and HeLa cells (2-4), suggesting that the translation imaging approach

described here can be performed in most cell types (see Note 7 on how to choose the best cell type for your experiment).

3.2.2 Delivering the plasmids

To create a cell line in which translation can be imaged, the plasmids described above can be delivered into cells with standard methods such as transfection or viral transduction. Because of the repetitive sequences present in the reporter mRNA plasmid (e.g. the PP7 binding sites and the sequence encoding the Sun-Tag peptides), virus titers may be low, and transfection may be preferred over viral transduction. To transfect the plasmids into U2OS cells in a 6 cm cell culture dish (transient or stable transfection; see Note 8 about transient versus stable transfections) the following Fugene (Promega) transfection protocol can be used. Please note that other cell types may require other transfection protocols.

1. Warm up DMEM medium without serum and without antibiotics (DMEM -/-) to 37°C.
2. Make a mastermix (number of reactions +1) of Fugene (Promega) containing 100 μ l DMEM -/- and 2 μ l Fugene per 6 cm dish.
3. Mix by tapping vigorously.
4. Spin down 3s to collect medium in bottom of the tube.
5. Incubate the master mix for ~5 min at room temperature.
6. Make the DNA mix containing 1 μ g total DNA per transfection.
7. Add 100 μ l of DMEM/Fugene mastermix to the DNA and mix by pipetting.
8. Incubate for 5 to 15 min at room temperature.
9. Add 3 ml of fresh cell culture medium to the cells.
10. Add the transfection mix to the cells.
11. After 24 hours wash the cells (note: this is not essential).

3.3 Preparing cells for imaging

1. Approximately 12-24 hours before imaging plate the cells containing the reporter plasmid, PP7-mCherry (-CAAX) and the scFv-GFP, in a glass bottom dish (we routinely use 96-wells glass bottom dishes) at the intended densities (~50% confluency). Depending on the experimental design and cell type, cells can also grow for longer time periods on the glass surface.

2. Immediately before transferring the glass-bottom dish containing the transfected cells to the microscope, replace the culture medium with pre-warmed imaging medium per 96-well (see Note 9 about the use of imaging medium).
3. Set the temperature at the microscope to 37°C for mammalian cells. Changes in temperature may result in cellular stress, which could influence the process of translation. We found temperatures between 36°C and 37.5°C to be acceptable for most human cell lines
4. When imaging for longer time periods, it is important to prevent evaporation of the cell culture medium, as this may result in changes in medium composition. We recommend keeping a lid on the imaging plate whenever possible to prevent evaporation.
5. When using a doxycycline-inducible promoter to express the translation reporter, add doxycycline (used at a final concentration of 1 µg/ml) approximately 10 min before the start of imaging. Addition of doxycycline to the cells will generally induce expression of the reporter mRNAs within 15-30 min (see Note 10 about how to add doxycycline to the cells). Adding doxycycline immediately before the start of imaging, allows imaging the first rounds of translation of newly transcribed mRNAs as well. In addition, adding doxycycline just before the start of imaging prevents cytoplasmic depletion of scFv-GFP, which occurs when high levels of SunTag protein are present in the cell (see below and Note 11 about the levels of scFv-GFP).
6. Select cells for imaging that have the correct levels of GFP- and mCherry-tagged proteins, and in which the translation reporter is expressed. (See Note 12 for more details about the expression levels of mCherry, See Note 11

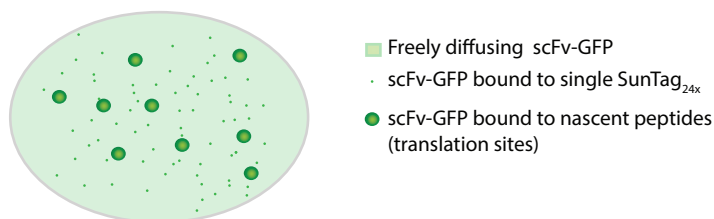


Figure 3: Different pools of scFv-GFP present in cells. Expression of scFv-GFP in the presence of a reporter plasmid results in the appearance of three different pools of GFP in the cell: 1) a pool of unbound, freely diffusing scFv-GFP, 2) a pool of scFv-GFP bound to mature SunTag proteins, which have 24 SunTag peptides and are thus ~24 times brighter than the single scFv-GFP molecules and 3) a pool of scFv-GFP bound to the nascent SunTag-peptides, which represents the sites of translation. Translation sites are much brighter than single SunTag proteins, as multiple ribosomes (and thus multiple nascent SunTag polypeptides) are present at a single mRNA molecule.

and Figure 3 for more details about scFv-GFP levels, see Note 13 for more information on how to select cells with the correct levels of translational reporter). Note that the number of cells that can be imaged at the same time is limited when a high time resolution is required.

3.4 Imaging and image acquisition

Different optical imaging techniques, including widefield, point scanning confocal, spinning disc confocal, TIRF and light sheet microscopy can be used for imaging translation (see Note 14 about the advantages and disadvantages of different imaging techniques).

1. Select the appropriate microscope objective. In general, we image with a 100x NA1.49 oil-immersion objective to obtain a high resolution and good sensitivity. Objectives with lower NA or magnification may also be used to image translation sites with multiple ribosomes, but might fail to reliably detect the mCherry-labeled mRNA or single ribosomes translating an mRNA.
2. Set the appropriate laser power and exposure time. Laser power and exposure time settings depend on the objective, microscope, camera and specifics of the experimental design. In the case of the tethered mRNA assay, low laser power in combination with a long exposure time provides the highest image quality and signal-to-noise ratio. Long exposure times (in the range of 500 ms) will cause motion blurring of the highly motile, GFP-labeled mature SunTag proteins and will therefore result in a more homogenous background signal. Since tethered mRNAs diffuse more slowly and are therefore not motion blurred at 500 ms exposure times, long exposure times will help to distinguish translation sites from background signal.
3. Find the correct focal plane to imaging mRNAs and translation sites. When the mRNA is tethered to the plasma membrane, it is important to focus the objective slightly above the plasma membrane of the cells during the imaging, as this is where the fluorescence associated with both the mRNA and the ribosomes translating the tethered mRNAs is located (see Note 15 on how to focus on both the mRNA and translation sites).
4. Set the time interval for time-lapse imaging. In order to image translation dynamics and allow measurements of translation initiation and elongation, we usually acquire an image every 30 s. In general, shorter time intervals make it easier to detect short lived events or to track individual mRNAs over time, while imaging with longer time intervals results in reduced photobleaching and phototoxicity, which allows imaging for longer time periods. In experi-

ments in which mRNAs are tethered, we found a 30 second time interval to be a good compromise between high temporal resolution and photobleaching for most of our experiments. However, when mRNAs are not tethered, a higher temporal resolution may be required to achieve accurate mRNA tracking.

5. Start image acquisition.
6. Add drugs which interfere with translation, as required by experimental setup. In most experiments, we recommend to add drugs after 10-30 min of imaging, when 10-50 translations sites are present per cell. Adding the drugs during imaging allows one to observe their immediate effects on translation. After addition of the drugs (See section 3.5), cells should be imaged for another 5-30 min to observe the effect on translation.

3.5 Drugs that interfere with translation as tools to study translation dynamics

Several drugs that are known to interfere with translation can be added to the cells to measure specific aspects of translation dynamics. Note that when adding drugs to the medium, it is advisable to pre-dilute the drug in large volume (~20% of final volume) (see Note 10 on adding drugs to the cells).

1. Puromycin (used at 100 $\mu\text{g}/\text{ml}$). Puromycin binds the elongating nascent polypeptide chain, thereby releasing the nascent polypeptide from the ribosome and dissociating the ribosomal subunits from the mRNA. Addition of puromycin to the cells results in the disappearance of bright GFP spots (translation sites) within 1 min after addition. Puromycin can therefore be used as a tool to verify whether the observed GFP spots are active translation sites. Make a 10 mg/ml stock concentration of puromycin in DMSO, which can be diluted to a concentration of 700 $\mu\text{g}/\text{ml}$ (7x the final concentration) in imaging medium. Of this dilution, add 50 μl to 300 μl imaging medium present in a well of a 96 well plate, to create a final concentration of 100 $\mu\text{g}/\text{ml}$.
2. Cycloheximide (used at 200 $\mu\text{g}/\text{ml}$). Cycloheximide (CHX) binds to the E-site of the ribosome, preventing release of the ribosome-bound tRNA and ribosomal translocation along the mRNA. Thus, CHX treatment results in stalling of ribosomes on the mRNA and should lead to the stabilization of GFP signal at translation sites. CHX may therefore be used to address whether changes in GFP intensity at sites of translation are caused by altered ribosomal occupancy. Make a 50 mg/ml stock concentration of cycloheximide in DMSO, which can be diluted to a concentration of 1400 $\mu\text{g}/\text{ml}$ (7x the final concentration) in imaging medium. Of this dilution, add 50 μl to 300 μl imaging medium

present in a well of a 96 well plate, to create a final concentration of 200 $\mu\text{g}/\text{ml}$.

3. Harringtonine (used at 3 $\mu\text{g}/\text{ml}$). Harringtonine is a small molecule translation inhibitor that specifically blocks translocation of ribosomes at the initiation codon, without affecting downstream ribosomes. As a consequence, upon harringtonine treatment ribosomes downstream of the start codon will complete translation normally and dissociate from the mRNA one-by-one after translation termination, resulting in a gradual decrease of the translation site GFP signal. Measuring the decay rate of GFP-fluorescence from single mRNAs, and fitting the data to a simple mathematical model (5), allows estimation of ribosome translocation rates on a given mRNA transcript. The duration of ribosome run-off is dependent on the length of the POI. In case of the translational reporter Addgene #74928, run-off can be observed within 5-15 min after harringtonine addition. Starting from a 3 mg/ml stock concentration of harringtonine in DMSO, make a dilution of 21 $\mu\text{g}/\text{ml}$ (7x the final concentration) in imaging medium. Of this dilution, add 50 μl to 300 μl imaging medium present in a well of a 96 well plate, to create a final concentration of 3 $\mu\text{g}/\text{ml}$.

3.6 Image analysis

Using the translation imaging method described here, the GFP-intensity of translation sites can be used to determine quantitative features of translation, including ribosome density on mRNAs, translation initiation rates, ribosome translocation rates and ribosome stalling. Determining these characteristics of translation requires 1) precise quantification of GFP intensities and 2) careful interpretation of the GFP fluorescence intensity.

3.6.1 Measuring GFP intensities and determining the number of ribosomes on an mRNA

The GFP intensity reports on the number of ribosomes on an mRNA. In some cases, it is sufficient to determine the relative number of ribosomes on each mRNA, in which case comparing GFP intensities between translation sites is possible. However, for other experiments it is useful to determine the absolute number of ribosomes on an mRNA. Below, we describe a step-wise protocol on how to measure GFP intensities and how these fluorescence intensity measurements can be used to calculate the number of ribosomes present on an mRNA.

In order to determine the number of ribosomes on an mRNA based on the GFP intensity of the translation site, it is important to compare the observed GFP

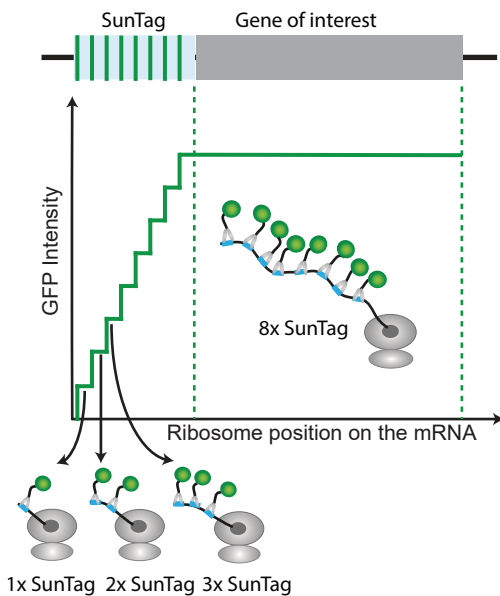


Figure 4: Interpreting GFP fluorescence intensities of translation sites. The scFv-GFP intensity associated with a single translating ribosome depends on the location of the ribosome on the reporter mRNA. The GFP intensity will initially increase as the ribosome synthesizes successive SunTag peptides (illustrated in the figure by the binding of 1, 2 or 3 scFv-GFP antibodies). Ribosome-associated fluorescence reaches a maximum once all SunTag peptides have been synthesized, and will remain constant while the ribosomes translates the remaining sequence of the mRNA. As a consequence, ribosomes at the 3' end of the mRNA are labeled more brightly, than those at the 5' end. When the number of ribosomes on an mRNA is calculated based on GFP intensity, these position-dependent effects of GFP intensity need to be taken into account. A simple mathematical model can be used, as described in the text.

intensity of a translation site with the GFP intensity of a single “mature” SunTag protein. Visualizing single mature SunTag proteins requires imaging with a short exposure time (10-30 ms) and a sensitive camera as a single SunTag protein contains at most 24 GFP molecules (in contrast to translation sites, which can contain >100 GFPs). Use a laser power that is sufficiently high to allow detection of single SunTag proteins, but without saturating camera pixel intensities at the much brighter translation sites. Single SunTag molecules are detectable on either EMCCD or sCMOS cameras.

1. Measure the GFP intensity of a single mature protein.
 - a. Draw a region of interest (ROI) around each fluorescent spot that represents a freely diffusing mature SunTag protein (See Note 16 on how to select foci representing mature SunTag proteins). Use an ROI that is as small as possible, but large enough to also accommodate translation sites.
 - b. Measure the mean fluorescence intensity in the ROI.
 - c. Measure the mean background GFP intensity in the cell, by using a large ROI which does not contain any GFP foci.
 - d. Subtract the mean background intensity from the mean spot intensity to obtain the intensity of a single SunTag protein.

2. Measure GFP intensity of a translation site.
 - a. Draw a ROI (of the same size as used for measuring single mature proteins) around each translation site.
 - b. Measure the average fluorescence intensity of each translation site.
 - c. Subtract the mean background intensity (measured in 1C) from the mean translation site intensity to calculate the mean intensity of a translation site.
3. In experiments where substantial photobleaching is observed, correction for photobleaching of the fluorescence intensities is critical (see Note 17 for further discussion about photobleaching)
4. Divide the mean GFP intensity of the translation sites by the mean intensity of the single mature SunTag proteins.
5. The value calculated in (3) provides an estimate of the number of SunTag arrays present at a translation site. However, as ribosomes at the 5' end of the mRNA (which haven't synthesized all the SunTag peptides yet) are dimmer than ribosomes at the 3' end of the mRNA, a correction needs to be applied to estimate the number of ribosomes present on the mRNA. To calculate the number of ribosomes per mRNA we have generated a mathematical model, which is described in detail in Yan et al 2016 (See Note 18 and Figure 4 for a further description on how to interpret GFP intensity).

3.6.2 Image analysis software to measure GFP intensities

In order to analyze the images obtained by microscopy, different image analysis software packages can be used, including Matlab, Python and ImageJ. The choice for a specific software package mainly depends on the experimenter's previous experience and personal preference. For unique or complex questions, custom analysis software may be required, making Matlab and Python good options. However, for many simple types of analysis, existing ImageJ plugins can be used. Currently, several plugins are available which allow, for example, counting of the number of translation spots per cell, measuring the intensities of individual translation spots or tracking translation spots over time. One simple ImageJ plugin that is useful for the analysis of translation dynamics is the `spot_counter` ImageJ plugin (<http://fiji.sc/SpotCounter>) developed by Nico Stuurman. This plugin counts the number of translation spots in a cell over time, and determines the fluorescence intensity of individual spots.

4. NOTES

1. **Choosing the gene of interest in the translation reporter.**

In principal, any gene can be introduced in the translation reporter, and the choice will mainly depend on the goal of the experiment. However, it is important to take into account that other regulatory sequences outside the coding region, such as the 3' and 5' UTR, may influence translation efficiency as well and can also be inserted into the reporter construct, either on their own or together with the coding sequence. If the goal of the experiment is not related to the regulation of a specific gene, but rather to study global translational control mechanisms, the specific mRNA sequence inserted downstream of the SunTag sequence may not be critical and different sequences can be inserted. The *length* of the reporter sequence is, however, an important parameter to take into account. The longer the reporter gene, the more ribosomes can be present on the mRNA simultaneously, affecting the brightness of the translation sites. The use of longer reporter genes (1-2kb) is therefore advisable and will facilitate further analysis.

2. **Cloning the SunTag peptide array sequence.**

The SunTag peptide array contains a somewhat repetitive sequence and is therefore difficult to amplify by PCR. The SunTag sequence is codon scrambled (i.e. different codons are used to encode the same amino acid sequence in each peptide) to minimize the degree of repeated nucleotide sequences. Codon scrambling allows PCR amplification to some extent, but some clones after PCR-based cloning will have small deletions. Therefore, cloning strategies that circumvent PCR-based amplification of this sequence are preferable.

3. **The number of SunTag peptides in the reporter gene.**

The reporter construct that is used by us and others (*β-5*) generally contains 24 SunTag peptides. However, shorter arrays (5 or 10 copies of the SunTag peptide (*5*)) or longer arrays (56 copies (*2*)) can also be used to image translation. Comparison between reporter mRNAs containing either 5, 10 or 24 SunTag copies revealed that ribosome density on these mRNAs was very similar, indicating that increased number of SunTag peptides does not detectably alter translation initiation or elongation rates of the reporter mRNA. Lowering the number of SunTag peptides will make the imaging of translation sites, especially at low ribosome occupancy, more challenging as it decreases the translation-associated GFP signal. Therefore, having a high copy number of SunTag peptides will be favorable in most situations. However, some specific experimental set-ups, such as the integration of the SunTag sequence into an endogenous gene locus by CRISPR/Cas9, may benefit from the use of shorter and less repetitive peptide arrays.

4. Using an inducible promoter to express a reporter mRNA.

The advantage of using an inducible promoter is that it allows temporal control of mRNA synthesis. Expressing the reporter mRNA only during the imaging experiment has two main advantages; first, it prevents accumulation of high level of relatively bright mature SunTag proteins, which hinders the imaging of translation sites. (See also Note 13 about the expression level of the translational reporter). Second, limiting the levels of mature SunTag protein will also prevent depletion of the freely diffusible pool of scFv-GFP antibody from the cytoplasm. If mature SunTag protein levels become too high, the majority of scFv-GFP is bound to mature protein and is therefore not available to bind nascent SunTag peptides at translation sites. As a consequence, newly made SunTag peptides emerging from the ribosome are labeled incompletely, limiting the fluorescence of translation sites.

5. Tethering the mRNA to the plasma membrane.

PP7-mCherry binds with high affinity to the hairpin sequence present in the 3' UTR of the reporter plasmid, resulting in fluorescent labeling of mRNA molecules. However, due to the rapid diffusion of single mRNAs, it is challenging to track single mRNAs over longer periods of time. In order to improve long-term tracking of individual mRNAs, a CAAX prenylation motif can be fused to the PCP. The CAAX motif anchors the PCP to the plasma membrane, which results in tethering of the reporter mRNAs containing the PCP binding sites to the plasma membrane. Membrane tethering of mRNAs reduces their mobility, facilitating long-term tracking of individual mRNA molecules (Figure 1A (lower panel), B; (5)). Importantly, membrane tethering of mRNAs also allows a specialized form of microscopy, called TIRF microscopy, which significantly improves the signal-to-noise ratio of the image. So far, we have not observed any differences in translation dynamics between tethered and untethered mRNAs ((5) and unpublished results). However, under certain conditions tethering of the mRNA to the membrane could affect translation, for example when translation of the reporter mRNA is spatially regulated. Appropriate controls should therefore be performed in experiments involving mRNA tethering.

6. The importance of fusing sfGFP to the SunTag antibody to create scFv-GFP.

The SunTag antibody has a tendency to aggregate at high expression levels in the cytoplasm of mammalian cells. To optimize intracellular expression of the antibody, a variety of N- and C-terminal fusion proteins known to enhance protein solubility were tested. Fusion of one variant of GFP, called sfGFP (9), to the C-terminus of the antibody resulted in soluble expression of the antibody even at high expression levels (5). Therefore, when changing the fluorophore fused to the antibody, it is important to

test whether the newly created antibody-fluorescent protein fusion does not aggregate in cells.

7. Choosing a suitable cell type.

In principal, most cell types can be used to study translation using the method described here. However, there are some features which may be worthwhile to take into account when choosing a cell type. The most important aspect of a cell type is whether the plasmids described above can be efficiently delivered into the cells. In addition, it is important to consider whether the cell type can be imaged using high resolution microscopy. For example, cells grown in suspension may be more difficult to image than flat, adherent cells.

8. Transient versus stable transfection.

In general, the plasmids required for translation imaging (See 2.1) can be either transiently transfected or used to generate cell lines stably expressing the genes of interest. While the use of a stable cell line will generally make results slightly more reproducible, a transient transfection will save time, as it allows for imaging of translation of a specific reporter one day after transfection. Since PP7-mCherry and scFv-GFP are used in every experiment, we recommend making a cell line in which both PP7-mCherry and scFv-GFP are stably expressed. This cell line can then be used to introduce different reporter mRNAs to study their translation. Note that in some cases we observed that stable expression of PP7-mCherry led to lysosomal accumulation of mCherry signal. Lysosomal accumulation was not detected when PP7 was fused to other fluorophores, such as GFP. mCherry-positive lysosomes are readily distinguishable from reporter mRNAs based on size, shape, intensity and motile properties, and therefore do not hinder the imaging of mRNAs, unless their localization overlaps with an mRNA molecule.

9. Using CO₂-independent, phenol red free imaging medium.

Replacing the normal cell growth medium (generally CO₂-dependent media containing phenol red) with CO₂-independent, phenol red-free imaging medium ensures a correct pH over the course of the experiment and, in addition prevents that phenol red from interfering with fluorescence imaging. We routinely use Leibovitz-L-15 imaging medium which is CO₂-independent and free of phenol red, and therefore ideal for live-cell microscopy. Note that appropriate levels of serum and antibiotics should still be added to the imaging medium. Alternatively, an optimal pH during the experiment can be achieved by having a CO₂ supply in the microscope imaging chamber.

10. Diluting drugs before addition to the cells.

We recommend pre-diluting drugs that need to be added to the cells during the imaging experiment in a large volume (~20% of final volume) before adding it to the cells. This ensures quick diffusion through the imaging medium so that the drug rapidly reaches the cells without repeated pipetting (which can cause cellular stress). We usually make a 7x dilution, of which we add 50 μ l to the 300 μ l imaging medium present per 96 well.

11. Expression levels of scFv-GFP.

Expression of scFv-GFP in the presence of a reporter mRNA results in the appearance of three different populations of GFP particles in the cell: 1) a pool of freely diffusing scFv-GFP, 2) a pool of scFv-GFP bound to SunTag proteins which have completed translation and have been released from the ribosome (and thus contain a single SunTag peptide array), referred to as “mature proteins”, and 3) a pool of scFv-GFP bound to the nascent SunTag peptides which represent sites of translation (Figure 3). scFv-GFP expression levels need to be sufficiently high to bind all SunTag peptides present in both the mature SunTag proteins and the newly synthesized nascent SunTag polypeptides as they emerge from the ribosomes. If scFv-GFP levels are too low, all antibody will be bound to mature SunTag protein, and the newly synthesized, nascent SunTag peptides will not be (completely) fluorescently labeled. On the other hand, if scFv-GFP levels are too high, this will give a strong background fluorescence of unbound scFv-GFP that can mask the signal of translation sites.

12. Expression levels of PP7-mCherry.

The expression levels of PP7-mCherry should be high enough to saturate binding to the PP7 binding sites, but low enough to prevent high background fluorescence of PP7-mCherry not bound to an mRNA. In addition, high expression of PP7-mCherry can cause accumulation in lysosomes, resulting in bright red dots in the cell (See also Note 8 about lysosomal accumulation of PP7-mCherry).

13. Expression level of the translational reporter.

For most experiments it is useful to select cells or a cell line for imaging with a high number of transcripts, as this enables the imaging of many translation events in one experiment. However, very high expression of the reporter also results in rapid antibody depletion (i.e. a situation in which the majority of antibody is bound to mature protein, resulting in weak labeling of translation sites, See note 11) and might impair long-term tracking of mRNAs (as moving mRNAs are more likely to cross paths). Optimal expression levels of the reporter mRNA therefore depends on the specific

experimental conditions. Inducible expression of the reporter reduces some of the problems of high expression levels of the mRNA reporter (i.e. antibody depletion) and is therefore generally beneficial. An additional approach to prevent high levels of labeled, mature SunTag protein in the cells involves fusion of the SunTag protein to a degron to ensure its rapid degradation after its synthesis is completed. This approach has been successfully used to increase the signal-to-noise ratio during imaging experiments (3, 4). When selecting cells with the correct level of the translational reporter, it is important to note that in case of very high levels of mature SunTag protein expression, the GFP signal in the cell may appear homogenous throughout the cell without clear SunTag punctae, because each SunTag protein is labeled with so few scFv-GFP molecules that individual SunTag-labeled proteins can no longer be distinguished from unbound scFv-GFP molecules based on their fluorescence intensity.

14. Advantages and disadvantages of different imaging systems.

Widefield microscopy: Widefield microscopy allows imaging of relatively thick Z-sections, facilitating mRNA tracking in 3D with limited number of Z-slices acquired. The reduced number of Z-slices required for tracking mRNAs in 3D may cause less photobleaching, and increase the time resolution that can be achieved. A drawback of widefield imaging, however, is that the signal-to-noise ratio is lower due to increased out-of-focus light, and therefore only moderately to strongly translating mRNAs (i.e. mRNAs translated by multiple ribosomes) will be detectable.

Point scanning confocal microscopy: With point scanning confocal microscopy, a higher signal-to-noise ratio can be achieved (compared to widefield microscopy). However, point scanning confocal microscopy is slow and causes relatively high levels of photobleaching and phototoxicity, and is therefore less suitable for long-term live-cell imaging.

Spinning disc confocal microscopy: Similar to point scanning confocal microscopy, the use of spinning disc confocal microscopy allows imaging with a high signal-to-noise ratio. However, spinning disc confocal imaging is faster than point scanning confocal microscopy and may cause lower levels of phototoxicity to the sample. It is therefore suitable for imaging translation of single mRNA molecules with high sensitivity over longer time periods, and is our system of choice for the majority of experiments involving the translation imaging system.

Total Internal Reflection Fluorescence (TIRF) microscopy: TIRF microscopy greatly reduces background signal and therefore increases the signal-to-noise ratio. However, only fluorescent molecules that are located close to the glass surface (e.g. near the plasma membrane at the bottom of the cell) can be observed, so tethering of the mRNAs to the

membrane is recommended when imaging translation sites using TIRF. An additional disadvantage of TIRF microscopy is that the illumination of cells is generally uneven, which hinders quantitative measurements of fluorescence intensities.

Light sheet microscopy: Light sheet microscopy can potentially be used to image translation in thick samples, such as tissues and embryos of various model organisms.

15. Finding the correct focal plane for imaging mRNAs and translation sites using the plasma membrane tethering approach.

Although red mRNAs and green translation spots largely co-localize, we noted that the mRNA fluorescence signal localizes slightly below (i.e. closer to the plasma membrane) than the translation signal. This is expected as mCherry is fused directly to the CAAX domain, and thus very close to the plasma membrane, whereas the GFP is connected to the membrane through mCherry, the mRNA, and the nascent chain, and will thus localize slightly further towards the cell interior (Figure 1A (lower panel)). Therefore, care should be taken to ensure both the mRNA and translation signal are in focus.

16. Selecting foci representing mature SunTag proteins to determine their fluorescence intensity.

Since mature proteins are not tethered to the plasma membrane and are relatively small compared to translation sites, they diffuse rapidly in 3D throughout the cell. Imaging mature SunTag proteins and measuring their fluorescence intensity is complicated by their fast diffusion, as the rapid movement of single SunTag protein causes motion blurring of the fluorescent foci in the image. In addition, rapid movement of mature SunTag proteins in the Z-axis causes many spots to be slightly out of focus when images are acquired. Both issues involving focus and motion blurring affect the fluorescence intensity of mature SunTag proteins. To minimize abovementioned issues, imaging with a short (10-30 ms) exposure time is recommended, which will reduce motion blurring. In addition, manually selecting foci for quantification that appear in focus will help alleviate abovementioned issues.

17. Correcting for photobleaching.

As a consequence of exposing fluorophores to excitation light, photobleaching occurs over time, reducing the intensity of GFP measured at translation sites. The rate at which photobleaching occurs can be determined by measuring the GFP signal of a large area in the cell (potentially the whole cell or field of imaging). Choosing an area of the cell lacking translation spots to measure bleaching rates ensures that such measurements

are not affected by appearance and disappearance of mRNAs. In experiments where substantial photobleaching is observed, correction for photobleaching of the fluorescence intensities of translation sites is critical.

18. Interpreting the scFv-GFP fluorescence intensity to measure translation dynamics.

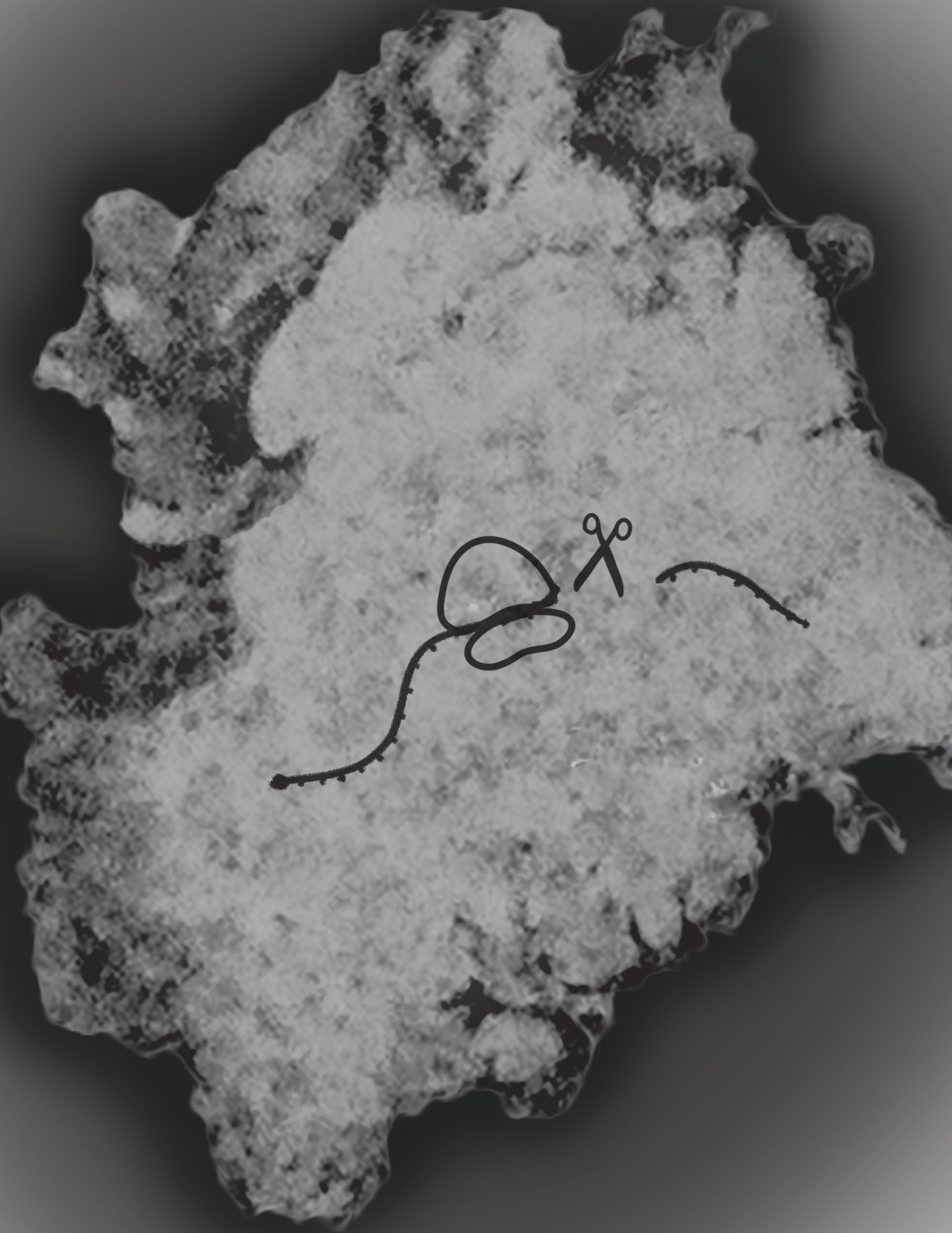
The GFP signal observed at translation sites is a result of the nascent SunTag peptides bound by scFv-GFP antibodies. Importantly, ribosomes on the 5' end of the mRNA have not yet translated all the SunTag peptides, and thus have fewer antibodies and fewer GFPs associated with them as compared to ribosomes at the 3' end of the mRNA (Figure 4). As a result, the GFP intensity associated with a ribosome at the 5' end of the mRNA is lower than with a ribosome at the 3' end of the mRNA (which has synthesized the entire SunTag peptide array). As a consequence, the measured GFP intensity at a translation site is not directly related to the number of ribosomes on the mRNA. To correctly translate GFP intensity to ribosome number, both ribosome density and ribosome location along the mRNA need to be taken into account. A simple mathematical model can be used to calculate the number of ribosomes from the measured GFP intensities (5).

ACKNOWLEDGEMENT

This work was supported by a fellowship of the Dutch Cancer Society (KWF) to M.E.T and an NWO Veni grant (0.16. Veni. 171.050) to S.R.

REFERENCES

1. Morisaki T, Lyon K, Deluca KF et al. (2016) Real-time quantification of single RNA translation dynamics in living cells. *Science (New York, N.Y.)* 352:1425-1429
2. Pichon X, Bastide A, Safieddine A et al. (2016) Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *The Journal of cell biology* 214:769-781
3. Wang C, Han B, Zhou R et al. (2016) Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell* 165:990-1001
4. Wu B, Eliscovich C, Yoon YJ et al. (2016) Translation dynamics of single mRNAs in live cells and neurons. *Science (New York, N.Y.)* 352:1430-1435
5. Yan X, Hoek TA, Vale RD et al. (2016) Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* 165:976-989
6. Tanenbaum ME, Gilbert LA, Qi LS et al. (2014) A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159:635-646
7. Bertrand E, Chartrand P, Schaefer M et al. (1998) Localization of ASH1 mRNA particles in living yeast. *Molecular cell* 2:437-445
8. Chao JA, Patskovsky Y, Almo SC et al. (2008) Structural basis for the coevolution of a viral RNA-protein complex. *Nature structural & molecular biology* 15:103-105
9. Pedelacq JD, Cabantous S, Tran T et al. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology* 24:79-88



Chapter 4

Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay

Tim A. Hoek*, Deepak Khuperkar*, Rik G. H. Lindeboom, Stijn Sonneveld, Bram M. P. Verhagen, Sanne Boersma, Michiel Vermeulen and Marvin E. Tanenbaum

** These authors contributed equally to this work*

Adapted from Mol Cell (2019)

ABSTRACT

Nonsense-mediated decay (NMD) is a surveillance system that degrades mRNAs containing a premature termination codon (PTC) and plays important roles in protein homeostasis and disease. The efficiency of NMD is variable, impacting the clinical outcome of genetic mutations. However, limited resolution of bulk analyses has hampered the study of NMD efficiency. Here, we develop an assay to visualize NMD of individual mRNA molecules in real time. We find that NMD occurs with equal probability during each round of translation of an mRNA molecule. However, this probability is variable and depends on the exon sequence downstream of the PTC, the PTC-to-intron distance, and the number of introns both upstream and downstream of the PTC. Additionally, a subpopulation of mRNAs can escape NMD, further contributing to variation in NMD efficiency. Our study uncovers real-time dynamics of NMD, reveals key mechanisms that influence NMD efficiency, and provides a powerful method to study NMD.

INTRODUCTION

Nonsense mutations (i.e. point mutations that create a premature termination codon (PTC)) are responsible for ~20% of all disease-associated single-base pair substitutions (Mort et al., 2008). In addition to genetic mutations, a PTC can also be introduced into an mRNA molecule stochastically, through errors in transcription or splicing. Understanding the fate of mRNAs containing a nonsense mutation is critical to understand the phenotypic outcome of such mutations. Transcripts harboring a PTC are rapidly degraded by a process called nonsense-mediated mRNA decay (NMD), which prevents the synthesis of truncated, and potentially toxic, proteins (Kurosaki and Maquat, 2016, Karousis et al., 2016, Lykke-Andersen and Jensen, 2015, He and Jacobson, 2015).

Pre-mRNA splicing is a critical regulator of NMD. Most mammalian genes contain multiple introns, which are spliced out before nuclear export of the mRNA (Sakharkar et al., 2004). Concomitant with intron splicing, the Exon Junction Complex (EJC) is loaded onto the mRNA 20-24 nt upstream of the exon-exon junction (Le Hir et al., 2000, Singh et al., 2012, Saulière et al., 2012). Since stop codons are generally located in the last exon of a gene, all EJCs will usually be deposited in the coding sequence of the mRNA. During translation, these EJCs are removed by the first ribosome translating the mRNA (Dostie and Dreyfuss, 2002, Sato and Maquat, 2009, Lejeune et al., 2002), so translation termination occurs in the absence of EJCs bound to the mRNA. In contrast, PTCs are frequently located upstream of one or more introns, and translation termination on PTC-containing transcripts can thus occur while one or more EJCs are still bound to the mRNA. These mRNA-bound EJCs are thought to communicate with the ribosome during translation termination through the NMD factor UPF1 and the translation termination factors eRF1/3 (Kashima et al., 2006). EJC-ribosome communication triggers degradation of the PTC-containing mRNA, which occurs either through endonucleolytic cleavage of the mRNA by SMG6 followed by exonucleolytic decay of the cleavage fragments (Huntzinger et al., 2008, Eberle et al., 2009, Gatfield and Izaurralde, 2004), or through deadenylation, decapping and exonucleolytic decay stimulated by the SMG5/7 complex (Unterholzner and Izaurralde, 2004, Loh et al., 2013).

For many NMD substrates, residual levels of PTC-containing transcripts can be detected in steady-state measurements (Cheng and Maquat, 1993, Cheng et al., 1994, Belgrader et al., 1994, Trcek et al., 2013, Thermann et al., 1998, Boehm et al., 2014, Lindeboom et al., 2016), suggesting that mRNA molecules with the same sequence display heterogeneity in timing of decay and/or susceptibility to NMD. Interestingly, the amount of residual mRNA for an NMD substrate appears to vary depending on the gene and the position of the PTC (Lindeboom et al., 2016, Thermann et al., 1998, Cheng et al., 1994). Uncovering the underlying cause of this variability, as well as the

factors that influence it, would provide important insights into the mechanisms that control NMD efficiency, and would provide a better understanding of the clinical outcome of disease-associated genetic mutations. However, current technologies such as Northern blot analysis or qRT-PCR only report the average mRNA levels of an NMD target in a population of cells, and thus preclude insight into differences within mRNA populations. Differences in levels of NMD targets could arise not only from differences in the rate of NMD-dependent mRNA decay, but also from differences in the fraction of mRNA molecules that is susceptible to NMD or even differences in NMD efficiency between subpopulations of cells. Furthermore, as NMD requires translation of the target mRNA in the cytoplasm, steady-state mRNA levels are likely also affected by differences in nuclear export rates and differences in the onset and efficiency of translation of different mRNAs. Finally, NMD has been proposed to occur preferentially during a ‘pioneer round’ of translation on mRNAs bound to the nuclear cap-binding complex (CBC) (Ishigaki et al., 2001, Maquat et al., 2010), which is replaced by the cytoplasmic cap-binding protein eIF4E after nuclear export, although this model is debated (Durand and Lykke-Andersen, 2013, Rufener and Mühlemann, 2013). The inability to precisely determine the timing of both NMD and CBC-replacement by eIF4E has hampered the understanding of the effect of CBC-replacement on NMD. To overcome the technical challenges associated with bulk and ‘snapshot’ analyses, we have developed an imaging method that allows real-time visualization of both mRNA translation and NMD of single mRNA molecules in living cells. Using this system, we precisely define the timing of NMD, describe a subpopulation of mRNA molecules of variable size that is resistant to NMD, uncover the key parameters that control NMD efficiency and provide the first real-time kinetic measurements of degradation of the 3’ mRNA cleavage fragment generated by NMD. Taken together, this single molecule imaging approach reveals key determinants of NMD variability and efficiency and provides a powerful assay to study NMD.

RESULTS

An assay to visualize NMD of single mRNA molecules in real-time

To analyze NMD of single mRNA molecules in real-time, we modified the SunTag fluorescence labeling approach that we (Tanenbaum et al., 2014, Yan et al., 2016) and others (Wang et al., 2016, Morisaki et al., 2016, Wu et al., 2016, Pichon et al., 2016) have recently developed for studying translation of single mRNA molecules. Briefly, this system uses a reporter mRNA that encodes (i) an array of 24 SunTag peptides near the 5’ end of the coding sequence to monitor translation, and (ii) 24 binding sites for the PP7 bacteriophage coat protein (PCP) within the 3’ UTR to monitor the mRNA itself (Chao et al., 2008) (Fig. 1A). Upon translation, the SunTag peptides recruit GFP-tagged

antibodies (scFv-GFP, stably expressed in the cell) that fluorescently label the nascent protein (Fig. 1A-C). The reporter mRNA is fluorescently labeled and tethered to the plasma membrane via PCP-mCherry-CAAX (Fig. 1A-C); we previously showed that tethering improves visualization without affecting translation (Yan et al., 2016).

To generate NMD reporter mRNAs, we introduced the sequence encoding the model NMD target Triose Phosphate Isomerase (TPI) in the reporter (Boehm et al., 2014, Belgrader et al., 1994), either with the native stop codon (TPI^{WT}), or with a PTC at amino acid 160 (TPI^{PTC160}) or at amino acid 1 (TPI^{PTC1}) (Fig. 1B). Furthermore, we included a sequence encoding BFP upstream of TPI to confirm expression of the reporter and placed the reporter mRNA under control of a doxycycline-inducible promoter (Fig. 1B). Human U2OS cells expressing TPI^{PTC160} or TPI^{PTC1} displayed a 4- or 12-fold lower level of mRNA expression, respectively, than cells expressing TPI^{WT}, as determined by qRT-PCR (Fig. S1A). TPI^{PTC160} and TPI^{PTC1} abundance increased 3-6 fold upon depletion of the key NMD factor UPF1, consistent with NMD of TPI^{PTC160} and TPI^{PTC1} (Fig. S1B, C). Degradation of NMD reporter mRNAs was not affected by expression of scFv-GFP and PCP-mCherry-CAAX (Fig. S1A), demonstrating that fluorescence labeling of mRNA and nascent polypeptides or mRNA tethering do not interfere with NMD.

To observe NMD of single mRNA molecules in real-time, human U2OS expressing TPI^{WT} or TPI^{PTC160} reporter mRNAs were followed using time-lapse spinning disk confocal microscopy with a 30 s time interval. Cells expressing the TPI^{WT} reporter displayed red mRNA foci in the cytoplasm within 15-30 minutes of transcription induction by doxycycline addition. The majority of mRNAs ($86 \pm 1.1\%$, mean \pm SEM) initially appeared without associated green translation signal, but rapidly initiated translation (after 2.3 ± 0.2 min, mean \pm SEM). Translation was generally maintained for the remainder of the experiment (\sim 30-60 min) (Fig. 1C, top panel, S1D-F and Movie S1) (see STAR methods). Quantitative analysis revealed that the initial GFP appearance on mRNAs represented the first round of translation for almost all mRNAs ($>99\%$) (see STAR methods).

Kinetics of transcription and of translation initiation of TPI^{PTC160} mRNAs were similar to those of TPI^{WT} mRNAs (Fig. S1D-F). However, the green and red foci associated with individual translating TPI^{PTC160} mRNAs often separated rapidly after translation had initiated ($79 \pm 3\%$ in 20 min, mean \pm SEM) (Fig. 1C, bottom panel and Movie S2). In contrast, only $13 \pm 3\%$ of TPI^{WT} mRNAs showed foci separation during this time-period (Fig. 1D). We also observed rapid foci separation for TPI^{PTC1} reporter mRNAs ($90 \pm 3\%$, Movie S3), or when TPI was replaced by another NMD model substrate, β -globin^{PTC39} ($84 \pm 3\%$, compared to $8 \pm 1\%$ for the β -globin^{WT} reporter) (Fig. 1E). Separation of red and green foci likely represents endonucleolytic cleavage of the mRNA, rather than translation shutdown, because multiple ribosomes (5-30) associate with a single mRNA molecule (Fig. S1G), and the entire GFP spot (i.e. all 5-30

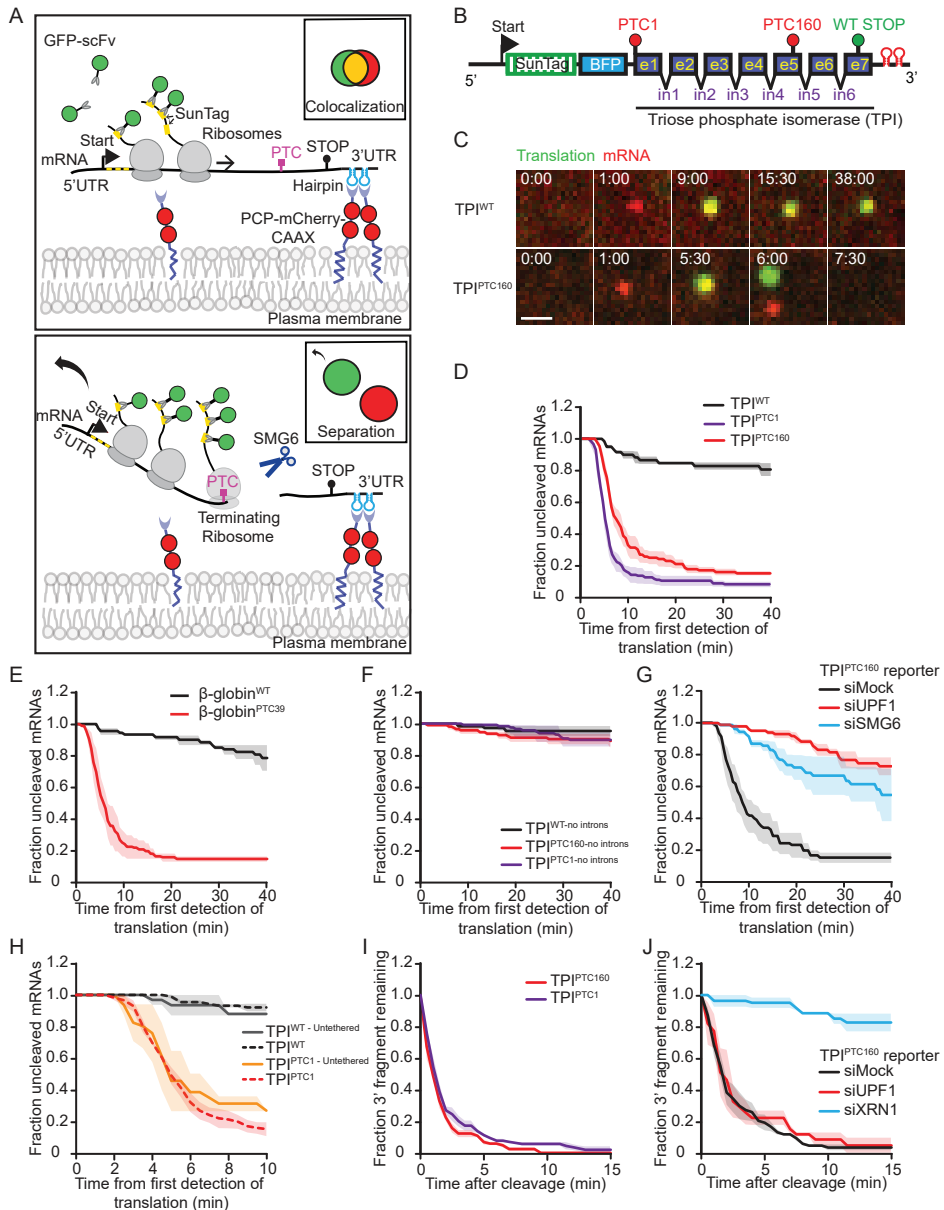


Figure 1. An assay for real-time visualization of NMD of single mRNA molecules. A) Schematic of NMD single molecule imaging assay before (top) or after (bottom) NMD induction. Green and red spots (insets) show nascent proteins and reporter mRNA, respectively, as observed through the microscope. B) Schematic of the NMD reporter constructs. e = exon; in = intron. PTC160 and PTC1 - indicate PTC at amino acids 160 and 1, respectively. (C-J) U2OS cells expressing scFv-sfGFP and either PCP-mCherry-CAAX (C-G, I, J) or PCP-HaloTag (H) were transfected with indicated reporter constructs (C-J) and siRNAs (G, J), and were analyzed by time-lapse microscopy. C) Representative images of mRNA molecules of

ribosomes) separates from the mRNA signal in a single step (Fig. S1H, S1I and STAR methods). The following observations support that the observed mRNA cleavage is induced by NMD; first, cleavage only occurred after the first ribosome reached the stop codon, consistent with an essential role of translation termination at the PTC in NMD induction (Fig. 1D, S1J and see STAR methods). Second, very little cleavage was observed in the absence of a PTC (Fig. 1D, E). Third, mRNA cleavage required the presence of introns in the mRNA (Fig. 1F). Finally, cleavage was strongly reduced after depletion of the important NMD factor UPF1 and the endonuclease SMG6 (9% and 28% cleavage in 20 min, respectively, compared to 78% in mock-treated cells, Fig. 1G). Surprisingly, the small amount of cleavage observed with the TPI^{WT} reporter was also reduced upon UPF1 depletion (Fig. S1K), suggesting that a small fraction of mRNA molecules are targeted for NMD even in the absence of a PTC, possibly due to stochastic errors in transcription or splicing of TPI^{WT} mRNAs. Together, these results show that red and green foci separation is an accurate readout of NMD induction of single mRNA molecules.

Interestingly, for all NMD reporters, a small subset of mRNAs molecules appeared resistant to cleavage (approximately 20%, 10% and 10% for TPI^{PTC160}, TPI^{PTC1} and β -globin^{PTC39} reporters, respectively, Fig. 1D, E), consistent with heterogeneous NMD kinetics for different subpopulations of mRNAs (Trcek et al., 2013). This cleavage-resistant population could not be explained by heterogeneity in NMD efficiency among different cells (Fig. S2A-D, see STAR methods). The cleavage resistant population was also not a result of differences in translation efficiency between NMD-sensitive and resistant mRNAs (Fig. S2E). Therefore, these data suggests that different mRNA molecules are heterogeneous with respect to NMD.

Long 3' UTRs can stimulate NMD under certain conditions (Singh et al., 2008, Boehm et al., 2014, Buhler et al., 2006), so we tested whether the relatively long 3' UTR in our reporter mRNA (1986 nt, including the 24x PP7 binding sites) might affect NMD induction. However, reducing the length of the TPI^{PTC160} 3' UTR length to 765 nt (containing 5x PP7 binding sites) did not alter NMD kinetics (Fig. S2F, G, and see STAR methods). Tethering the reporter mRNAs to the plasma membrane also did not detectably alter the kinetics of NMD (Fig. 1H, Movie S4). Furthermore, we observed similar cleavage kinetics upon transient transfection or stable integration of the reporter gene (Fig. S2H, I). Therefore, to facilitate the experimental setup and analyses, we used

Figure 1 continued... indicated reporters are shown. Scale bar, 1 μ m. Time is shown in min:sec. D-J) The time from first detection of translation until separation of red and green foci (i.e. mRNA cleavage) (D-H) or the time from mRNA cleavage until disappearance of the 3' cleavage fragment (red spot) (I-J) was quantified. Solid lines and corresponding shaded regions in (D-J) represent mean \pm SEM. Dotted lines in (H) indicates that the data is replotted from an earlier figure panel for comparison. Number of measurements for each experiment are listed in Table S1. See also Figure S1, S2 and Movies S1-S4.

transiently transfected reporters, which contained 24x PP7 binding sites and were tethered to the plasma membrane, unless indicated otherwise.

Finally, we examined the fate of the two mRNA fragments that are produced by cleavage of the reporter mRNA. The 5' fragment rapidly diffused out of the plane of observation after endonucleolytic cleavage, precluding analysis of its degradation. The 3' fragment, which remains tethered to the membrane, disappeared rapidly after mRNA cleavage (Fig. 1C, 1I). SiRNA-mediated depletion of the 5'-to-3' exonuclease XRN1, but not UPF1, reduced the rate by which red foci disappeared by over 10-fold (Fig. 1J and S2J), indicating that red foci disappearance represents XRN1-mediated decay of the 3' cleavage fragment. In conclusion, our imaging approach allows us to monitor NMD of individual mRNA molecules from start-to-finish in real-time.

Each ribosome that terminates translation at the PTC has an equal probability of inducing NMD

Earlier studies have suggested that NMD occurs preferentially during the pioneer round of translation, which is generally defined as the first, or first few ribosomes that translate an mRNA, while the mRNA is bound by CBC (Maquat et al., 2010), while other studies provided evidence that NMD can occur during any round of translation (Durand and Lykke-Andersen, 2013, Rufener and Mühlemann, 2013). Since our method allows precise measurements of the timing of the first round of translation and NMD, we wished to determine which ribosome induced NMD.

When examining the cleavage kinetics of TPI^{PTC1} in more detail, we noticed three distinct phases in the distribution of cleavage times (Fig. S3A); in the first phase (0-3 min after GFP appearance), which represents the time it takes for the first ribosome to translate the coding sequence and reach the PTC, very little cleavage occurred (Fig. S3A). In the second phase (from 3 to 10 minutes for TPI^{PTC1}), the curve showed a very steep downward slope, indicating that most mRNA molecules were cleaved during this phase. Finally, during the third phase (>10 min) only few cleavage events occurred, indicating that the mRNAs that had not yet been cleaved (~5-10% of molecules for TPI^{PTC1}) were largely resistant to NMD. Similar phases were observed for other reporters as well (Fig. 1D, E). The first ribosome was predicted to arrive at the PTC of TPI^{PTC1} after ~3 minutes, yet cleavage was observed over a period of 3-10 minutes. Two possible models could explain the observed variability in timing of cleavage; first, it is possible that the first ribosome translating an mRNA always induces NMD, and that the variability is caused by variation in arrival time of the first ribosome at the PTC. Alternatively, the first ribosome may arrive at the PTC around 3 minutes on all mRNAs, and the variability in cleavage time is caused by variation in which ribosome induces NMD (i.e. the first ribosomes or one of the following ribosomes). To distinguish between these two models, we precisely determined the average time as well as the variation in time it takes ribosomes to reach the PTC, which we found to be 2.6 ± 0.8 min (mean

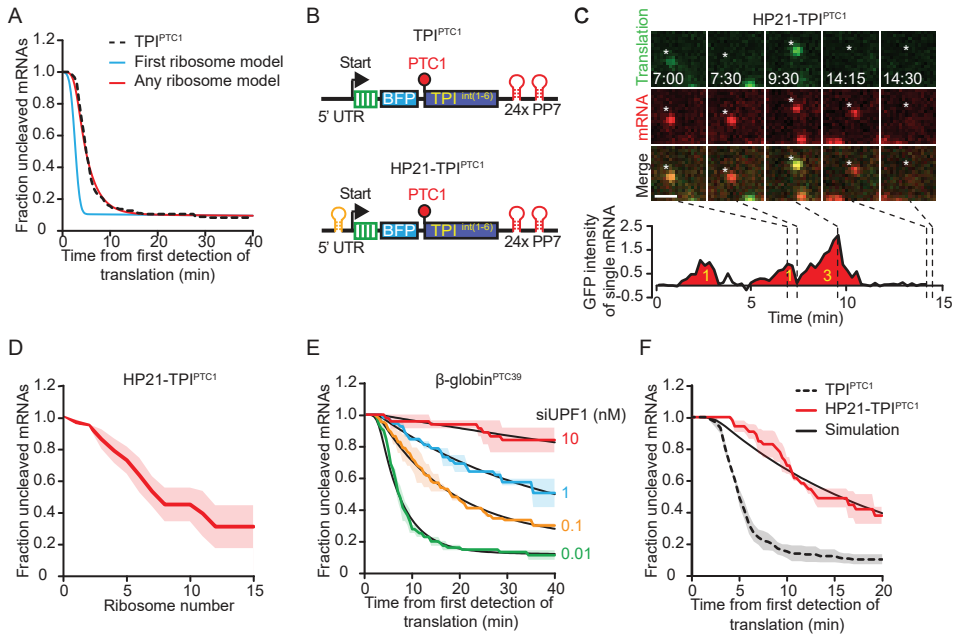


Figure 2. NMD occurs with equal probability during each round of translation. A) Experimentally-determined cleavage time distribution of TPI^{PTC1} is shown (black dotted line). Predicted cleavage time distributions based on stochastic simulations are shown for indicated models. B) Schematic of indicated reporters. C-F) U2OS cells expressing *scFv-sfGFP* and *PCP-mCherry-CAAX* were transfected with indicated reporter plasmids (C-F) and siRNAs (E) and were analyzed by time-lapse microscopy. C) Representative images of a single $HP21-TPI^{PTC1}$ mRNA molecule. Scale bar, 1 μm . Time is shown in min:sec. Graph shows GFP fluorescence intensity of the mRNA over time. Red filled areas represent peaks that were called as translation events. Yellow numbers indicate the number of ribosomes that contributed to the peak. D) Quantification of the number of ribosomes that translated $HP21-TPI^{PTC1}$ mRNAs before cleavage occurred. E-F) The time from first detection of translation until mRNA cleavage was quantified. Black lines indicate the best fit from simulations. Dotted lines (A, F) indicate that the data is replotted from an earlier figure panel for comparison. Solid lines and corresponding shaded regions in (D-F) represent mean \pm SEM. Number of measurements for each experiment are listed in Table S1. See also Figure S3 and Movie S5.

\pm SD) (Fig. S3B, see STAR methods). We then performed stochastic simulations to determine if the experimentally determined variation in rate of ribosome translocation could explain the observed cleavage time distribution. In brief, we developed a model to describe the cleavage time distribution using two parameters: 1) the time of arrival of the first ribosome, for which we used our experimentally determined values (Fig. S3B), and 2) the fraction of NMD-resistant mRNAs (see STAR methods). We found that the observed cleavage kinetics of the TPI^{PTC1} reporter were poorly described by a model in which NMD is induced by the first ribosome (Fig. 2A, compare blue and black lines). We therefore tested an alternative model, in which the cleavage rate reflects both the

time of arrival of the first ribosome and the probability that a ribosome will induce NMD upon translation termination. To this end, we added a third parameter to the model; a probability for each terminating ribosome of inducing NMD. This parameter requires knowledge of the frequency of ribosome termination events, which we calculated to be 3.2 ribosomes per minute based on the translation elongation rate and ribosome occupancy of TPI^{PTC1} (Fig. S1G, S3B, see STAR methods). The second model resulted in a very good fit with the data (Fig. 2A, compare red and black lines, Akaike information criterion (AIC) = -639 compared to AIC = -308 for the ‘first ribosome’ model described above, see STAR methods) and revealed that $90 \pm 3\%$ (mean \pm S.E.M) of TPI^{PTC1} mRNAs were degraded in the rapidly degrading population, and that each terminating ribosome induced NMD with a probability of 0.11 ± 0.01 (i.e. a 11% chance of inducing NMD per ribosome) (see STAR methods). Together, these results strongly suggest that each terminating ribosome has an equal probability of inducing NMD.

To unambiguously determine which ribosome induced NMD of each mRNA, we engineered a TPI^{PTC1} reporter with a strongly reduced initiation rate by introducing a 21-nucleotide hairpin in the 5' UTR (which reduced translation by ~ 30 -fold (Fig. S3C)) that allowed us to count individual ribosomes translating the reporter mRNA (HP21-TPI^{PTC1}, Fig. 2B). When tracking GFP intensity over time on HP21-TPI^{PTC1} mRNAs, we observed clear peaks of GFP signal that lasted several minutes, interspersed by periods lacking detectable GFP signal, and we could assign the precise number of ribosomes that made up each peak (Fig. 2C, Fig. S3E, Movie S5 and see STAR methods).

When mRNAs are translated by a single ribosome, nascent chain release during translation termination and mRNA cleavage both result in complete separation of the GFP and mCherry foci and are thus indistinguishable. Therefore, to monitor NMD of the HP21-TPI^{PTC1} mRNAs we defined mRNA cleavage as red and green foci separation rapidly followed by the disappearance of the red foci (i.e. decay of the 3' cleavage fragment), which was a reliable readout for NMD of HP21-TPI^{PTC1} mRNAs (Fig. S3F, G). When counting individual ribosomes translating HP21-TPI^{PTC1} reporter mRNAs, we found that 8 ribosomes (median) typically translated the HP21-TPI^{PTC1} reporter mRNA before NMD was triggered (Fig. 2D). This corresponds to a probability of 0.10 ± 0.03 (mean \pm SEM) of inducing NMD for each terminating ribosome (see STAR methods), which is in good agreement with the probability of 0.11 per ribosome determined through our stochastic simulation approach for the TPI^{PTC1} reporter lacking the hairpin sequence (Fig. 2A).

NMD does not occur preferentially on CBC-bound mRNAs

If NMD occurs preferentially on CBC-bound mRNAs, a progressively slower NMD decay rate over time should be observed due to gradual replacement of CBC by eIF4E within the population of mRNAs. However, the cleavage rate of TPI^{PTC1}, TPI^{PTC160} and

β -globin^{PTC39} was constant over time (i.e. fit an exponential decay distribution; excluding the subset of mRNAs that are resistant to NMD) (Fig. 2A). Decay of these reporters was, however, very rapid (<10 min), so it is possible that all of these mRNAs are degraded while the mRNA is still bound to CBC because the CBC-eIF4E exchange occurs at a time-scale of >10 min. To determine if the NMD decay rate remains constant over longer time periods, we reduced the NMD efficiency to extend the time window during which the decay rate can be analyzed. First, we performed partial depletion of UPF1 in cells expressing either TPI^{PTC160} or β -globin^{PTC39}, which revealed that the decay rate remained largely constant over the entire 40 min experiment, as evident from a good fit of the data with an exponential decay distribution (Fig. 2E and S3H). Second, we examined NMD efficiency of an mRNA with a lower translation initiation rate to reduce the speed of NMD. For this, we re-plotted the cleavage kinetics of the HP21-TPI^{PTC1} reporter, and found that it also showed a similar rate of decay over the entire experiment (Fig. 2F). Together, these results show that NMD occurs at a constant rate over at least 40 min (~100 rounds of translation).

To definitively determine the efficiency of NMD of both CBC- and eIF4E-bound mRNAs, we wished to determine the exact moment of CBC-to-eIF4E replacement. For this, we made use of a specific inhibitor of eIF4E-dependent translation; a hyperactive, non-phosphorylatable, non-degradable variant of the protein 4E-BP1 (ha4E-BP1) (Yanagiya et al., 2012, Durand and Lykke-Andersen, 2013). As CBC does not bind to 4E-BP1, we reasoned that translation initiation rates of single mRNAs would be unaffected by overexpression of ha4E-BP1 as long as mRNAs were bound to CBC (Durand and Lykke-Andersen, 2013). However, at some point in time, translation rates of single mRNAs would decrease in ha4E-BP1 overexpressing cells compared to control cells, and this time-point would represent the moment of replacement of CBC by eIF4E. In the absence of ha4E-BP1 expression, the majority of newly transcribed mRNAs rapidly initiated translation, as evident by the appearance of a green fluorescence signal, and continued translating for the remainder of the movie (20-45 min) (Fig. 3A, B, S4A). In cells overexpressing ha4E-BP1, initial appearance of green fluorescence occurred with similar kinetics as in control cells (Fig. 3C, S4B), suggesting that initial translation is likely driven by CBC on most mRNAs. However, in ha4E-BP1 expressing cells, translation was rapidly shut down on the majority of mRNAs within minutes of initial translation initiation (Fig. 3A-B, S4A), indicative of the CBC-to-eIF4E switch.

To determine the precise moment of translation initiation of each ribosome, we made use of a fluorescence fitting algorithm, RiboFitter, which we recently developed (Boersma et al., 2018). In control cells, we found an average translation initiation rate of 2-3 ribosomes per minute, which remained mostly constant over time (Fig. 3D, F). In contrast, in ha4EBP1-expressing cells most mRNAs showed a brief burst of translation initiation during which the initiation rate was similar as in control cells, followed by a period without initiation (Fig. 3E, F). Note that after the period with

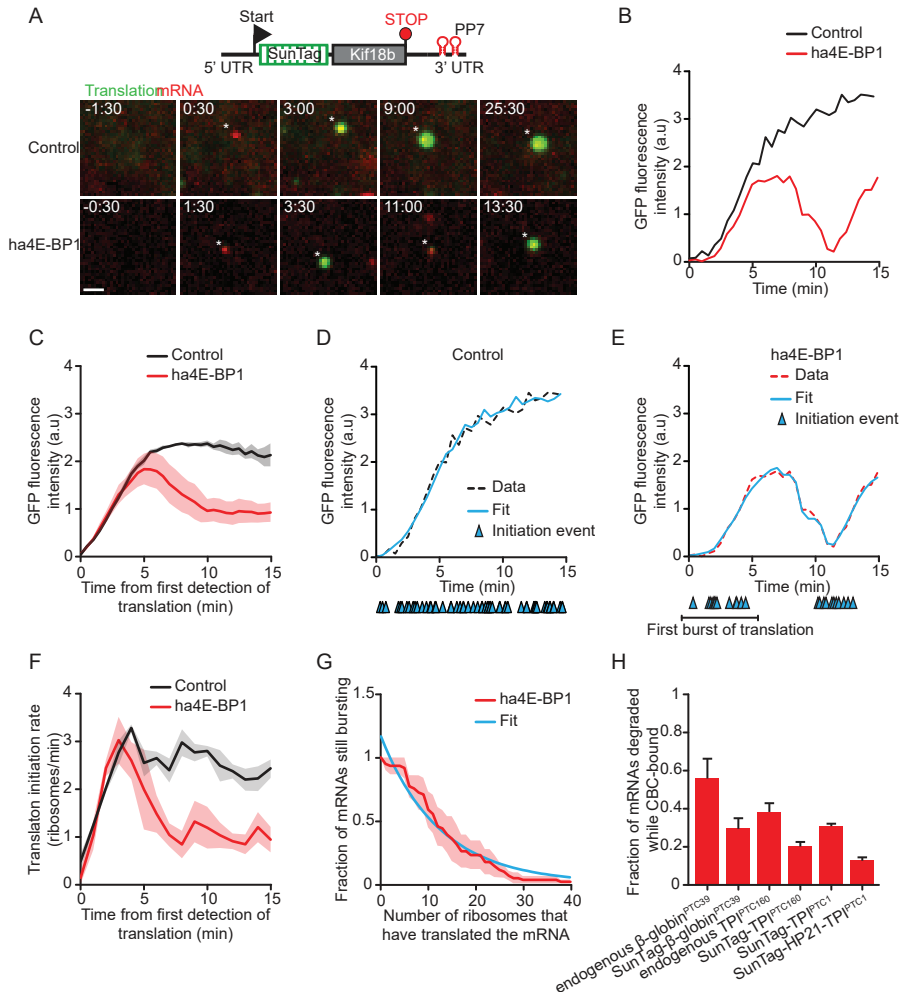


Figure 3. NMD does not occur preferentially on CBC-bound mRNAs. A-G) U2OS cells expressing scFv-sfGFP, PCP-mCherry-CAAX and the translation reporter shown in (A) were transfected with ha4E-BP1 or mock transfected and analyzed by time-lapse microscopy. A) (top) Schematic of standardized translation reporter. Representative images of a single mRNA molecule of either mock (upper image panel) or ha4E-BP1 (lower image panel) transfected cells are shown. Scale bar, 1 μ m. Time is shown in min:sec. B-E) GFP fluorescence intensity over time of a representative mRNA (B, D, E) or of the average of all mRNAs (C). Blue lines (D-E) indicate the best fit from simulations. Blue triangles indicate translation initiation events. F) Quantification of the mean translation initiation rate determined by the fitting approach illustrated in (D, E). G) Quantification of the number of ribosomes that initiated in the first burst of translation. Data was fit with a single exponential decay distribution (blue line). H) The calculated fraction of mRNAs that is targeted for NMD while CBC is bound to the mRNA cap (mean \pm SEM). Dotted lines in (D, E) indicate that the data is replotted from an earlier figure panel for comparison. Solid lines and corresponding shaded regions in (C, F-G) represent mean \pm SEM. Number of measurements for each experiment are listed in Table S1. See also Figure S4.

no translation initiation events, additional bursts of initiation were often observed, which might represent removal of ha4E-BP1 from the cap or binding of a new eIF4E molecule that was not bound to ha4E-BP1 to the mRNA cap (Fig. 3E, S4C). The first burst of translation initiation in ha4E-BP1 expressing cells likely reflects CBC-bound translation, and the end of the translation initiation burst thus reflects the transition of CBC-to-eIF4E and binding of ha4E-BP1 to eIF4E. Analysis of the size of the first translation initiation burst in ha4E-BP1 expressing cells revealed that the median number of ribosomes that initiated on a CBC-bound mRNA before CBC was replaced by eIF4E was 8.7 (average of 3 experiments, Fig. 3G). As the NMD cleavage rate remained constant for >40 minutes (~100 rounds of translation, much longer than the duration of CBC-dependent translation), it follows that NMD efficiency is similar on CBC- and eIF4E-bound mRNAs.

Using the quantitative data on the kinetics of the CBC-to-eIF4E switch, we developed a model to calculate the fraction of mRNAs on which NMD was induced while CBC was bound to the mRNA cap. This fraction not only depends on the efficiency of NMD induction, but also on the length of the ORF; ribosomes require more time to reach the stop codon of long ORFs, which increases the time during which CBC-to-eIF4E replacement could occur. Our modelling approach suggests that for highly efficient NMD substrates with a short ORF, such as the endogenous β -globin mRNA with a PTC at position 39 (ORF of 117 nt), 56% of mRNAs would be targeted for NMD while CBC is bound to the cap (Fig. 3H). However, on mRNAs that are less rapidly targeted for NMD, such as weaker NMD substrates (e.g. TPI^{PTC160}), mRNAs with a lower translation initiation rate (e.g. HP21-TPI^{PTC1}) or substrates with a longer ORF (e.g. endogenous mRNAs with a long ORF, or reporters mRNAs containing the SunTag sequence), NMD will be induced more frequently when eIF4E is bound to the cap. For example, our modelling suggests that only 13% of cleavage events of the HP21-TPI^{PTC1} reporter mRNA (ORF length of 2517 nt) occur while the mRNA is bound to CBC (Fig. 3H).

To confirm that a subset of TPI^{PTC160} or TPI^{PTC1} mRNAs are bound by eIF4E at the moment of NMD induction, we also analyzed their cleavage rate in ha4E-BP1 expressing cells. As expected, NMD was delayed in these cells compared to control cells, and the magnitude of this delay was close to the magnitude predicted by our modeling (Fig. S4D, E). Taken together, these data demonstrate that NMD occurs with equal probability on CBC- and eIF4E-bound mRNAs and that the fraction of mRNAs undergoing NMD while bound to CBC depends on the NMD efficiency, translation initiation rate and ORF length..

Exon sequence downstream of the PTC influences NMD efficiency

Our real-time imaging approach revealed that both the cleavage rate and the fraction of NMD-resistant mRNAs can vary between different reporter mRNAs (e.g. fig. 1D). This

assay therefore provides a unique opportunity to determine the mechanisms regulating these parameters, and thus NMD efficiency.

First, we focused on the role of gene-specific nucleotide sequences on NMD efficiency. We found that the sequence of the PTC sequence itself did not affect NMD efficiency (Fig. S5A). Next, we examined the effect of nucleotide sequences downstream of the PTC. To directly compare different sequences, we generated a standardized reporter that contains a constant coding (Kif18b) and PTC (TAA) sequence, but variable sequences downstream of the PTC. We then inserted 14 randomly selected single introns with their native flanking exons (“exon-intron-exon” referred to as ‘intron cassettes’) in the standardized reporter downstream of the PTC (Fig. 4A). First, we assessed splicing efficiency of the introns in each of these reporters, either using a two-color fluorescence splicing reporter (Fig. S5B and see STAR methods), or by qPCR (Fig. S5C, D). As expected, reporters that were efficiently spliced generally showed NMD, while unspliced reporters did not (Fig. S5E) and hence were excluded from further analysis. Two reporters, ERAL1 and MITD1, showed moderate to strong splicing in the FACS-based reporter, but no splicing by qPCR, yet both reporters showed cleavage in the NMD assay. The observed cleavage was dependent on splicing of the intron, as cleavage was eliminated upon removal of the intron or mutation of the splice sites (Fig. S5G, S5H), indicating that they are undergoing splicing-dependent NMD as well.

Among well-spliced reporters, we observed striking differences in the cleavage rates, with the probability of inducing NMD for each terminating ribosome ranging from 0.008 to 0.14 (for MITD1 and TPI, respectively, Fig. 4B, 4C, S5F). For three reporters with varying cleavage rates, we confirmed that cleavage was dependent on UPF1 and on splicing (Fig. 4D, S5G, S5H), confirming that the observed cleavage is caused by NMD. The differences in the observed cleavage efficiencies were not caused by differences in intron sequences, as swapping introns did not affect NMD efficiency (Fig. 4E, F), suggesting that the exon sequences downstream of the PTC determine the NMD cleavage rate. For reporters that showed a relatively slow rate of cleavage (ERAL1, MITD1, DNAL4 and NUBP2) we examined whether the mRNA molecules that were not cleaved, were instead degraded through exonucleolytic decay, but found no evidence for exonucleolytic decay within the time-frame of our experiments (Fig. 4G). These results show that NMD cleavage rates can vary substantially depending on the mRNA sequence downstream of the PTC.

The PTC-to-intron distance affects both the cleavage rate and fraction of NMD-resistant mRNAs

Genome-wide studies revealed that in long exons, a large PTC-to-intron distance can result in a reduced NMD efficiency (Lindeboom et al., 2016). However, if a PTC is close (<50-55nt) to the last intron, NMD efficiency is also reduced, likely because EJC are displaced from the mRNA by the translocating ribosome (Dostie and Dreyfuss,

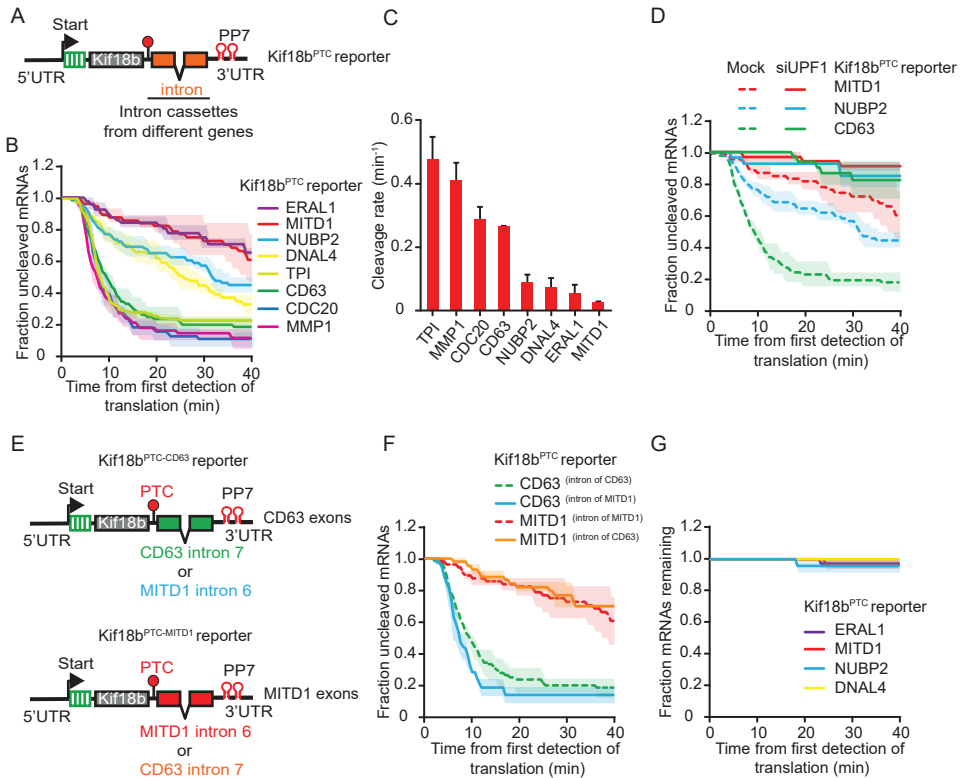


Figure 4. The NMD cleavage rate is variable and depends on exon sequences downstream of the PTC. A, E) Schematic of indicated reporters. B-G) U2OS cells expressing scFv-sfGFP and PCP-mCherry-CAAX were transfected with indicated reporter plasmids (B-G) and UPF1 siRNA (D) and analyzed by time-lapse microscopy. The time from first detection of translation until mRNA cleavage was determined (B-F). B) Reporters were designed as in (A) and encoded an intron cassette of indicated genes. C) Cleavage rates for each reporter shown in (B). G) Quantification of the duration that individual mRNAs that were not cleaved could be observed over time. Solid lines and corresponding shaded regions in (B, D, F, G) represent mean \pm SEM. Dotted lines in (D, F) indicate that the data is replotted from an earlier figure panel for comparison. Error bars in (C) represent SEM. Number of measurements for each experiment are listed in Table S1. See also Figure S5.

2002, Lejeune et al., 2002). Using our assay, we can distinguish whether PTC-to-intron distance affects the cleavage rate or fraction of NMD-resistant mRNAs, which could shed new light on the mechanisms underlying these observations. We introduced linker sequences of different lengths (100, 175, 250 and 1000 nt) between the PTC and downstream intron of the Kif18b^{PTC}-TPI-intron 6 reporter, in which the PTC is located 91 nt upstream of intron (Fig. 5A). Increasing the linker length decreased the cleavage rate up to 7-fold (Fig. 5B-D), suggesting that the probability that a terminating ribosome

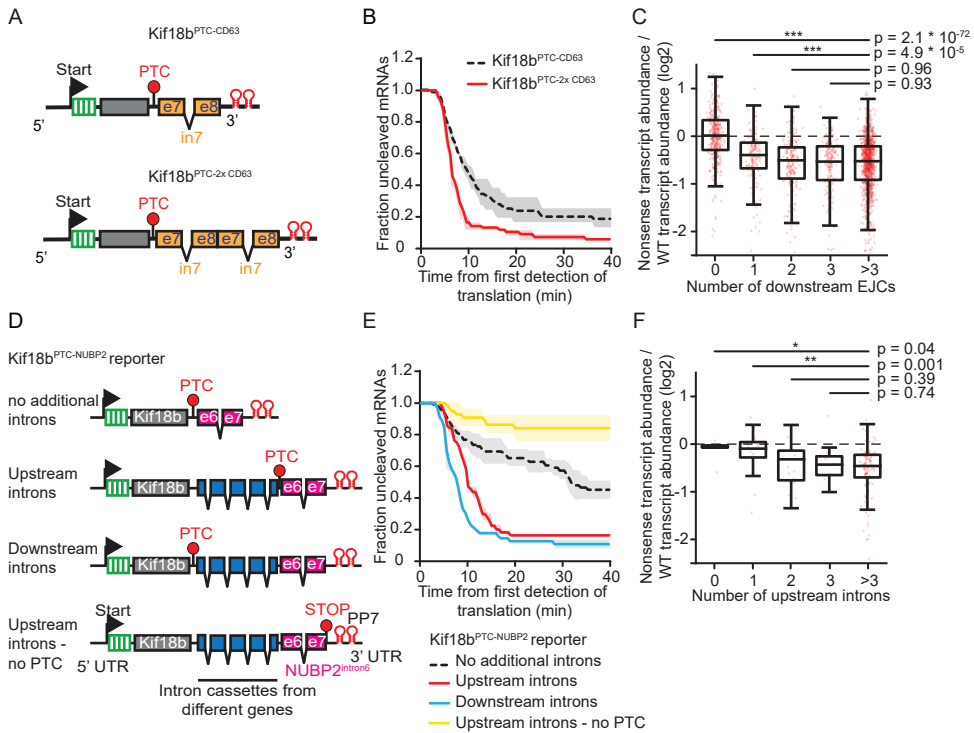


Figure 6. The number and position of introns relative to the PTC affects the NMD decay rate and fraction NMD-resistant mRNA molecules. A, D) Schematic of indicated reporters. B, E) U2OS cells expressing scFv-sfGFP and PCP-mCherry-CAAX were transfected with indicated reporter plasmids and analyzed by time-lapse microscopy. The time from first detection of translation until mRNA cleavage was determined. C, F) Genome-wide analysis of the effect of the number of downstream (C) or upstream (F) EJCs on NMD efficiency in a large cohort of cancer samples. Only PTCs with a single downstream EJC were included in (F). Solid lines and corresponding shaded regions in (B, E) represent mean \pm SEM. Dotted lines in (B, E) indicate that the data is replotted from an earlier figure panel for comparison. In boxplots of (C, F), the boxes represent the interquartile range with the central line depicting the median, and the whiskers extend to the extreme values after removing outliers. P values are indicated as * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) by two-tailed Mann-Whitney U tests. Number of measurements for each experiment are listed in Table S1. See also Figure S7.

will induce NMD depends on its proximity to the downstream EJC. The dependence of the cleavage rate on the PTC-to-intron distance was also observed with other linker sequences and reporters (Fig. S6A-E). Interestingly, although larger PTC-to-intron distances resulted in reduced cleavage rates, the fraction of NMD-resistant mRNAs was not substantially unaffected (Fig. 5C, D).

To explore the effects of very short PTC-to-intron distances on NMD, we introduced new PTCs in the TPI gene (Fig. 5E). Decreasing the distance between the PTC and the

downstream intron to 73, 52 or 40 nt (TPI^{PTC185}, TPI^{PTC192} and TPI^{PTC196}, respectively), led to a strong reduction in NMD (Fig. 5F), as expected. Interestingly, the reduction in NMD efficiency was mostly caused by a substantial increase in the fraction of NMD-resistant mRNAs (Fig. 5F), consistent with a model in which the EJC can be displaced from the mRNA by the translating ribosome if it is too close to the PTC, thereby preventing NMD during future rounds of translation. Surprisingly, we found that a fraction of mRNAs was still susceptible to NMD in all 3 reporters (TPI^{PTC185/192/196}) (Fig. 5F), suggesting that the EJC was not displaced by translating ribosomes from all mRNAs, even when it is positioned very closely to the PTC.

The number of introns both upstream and downstream of the PTC affects the NMD cleavage rate and the fraction of NMD-resistant mRNAs.

Next, we examined the effect of altering the number of introns in an mRNA on the NMD decay rate and the fraction of NMD-resistant mRNAs. We added an extra copy of a single intron cassette to our standardized reporter (Kif18b^{PTC-CD63}) to create the Kif18b^{PTC-2xCD63} reporter (Fig. 6A). These reporters have an identical PTC-to-intron distance, and the same nucleotide sequence immediately downstream of the PTC, but a different number of introns. Introduction of a second intron downstream of the PTC resulted in a faster decay rate and also a reduced fraction of NMD-resistant mRNA molecules (Fig. 6B), indicating that multiple downstream introns enhance NMD through two parallel mechanisms. Similar results were obtained with a second set of reporters (Fig. S7A-B). To support these findings, we performed analysis of the effects of nonsense mutations on mRNA levels in a large cohort of previously sequenced cancer samples (Lindeboom et al., 2016). Genome-wide analysis revealed that PTCs with only a single EJC downstream of the PTC had a significantly lower NMD efficiency than PTCs with multiple (>3) downstream EJCs ($p = 4.9 * 10^{-5}$, Mann-Whitney U test, Fig. 6C).

We considered two possible models to explain the enhanced NMD cleavage rate of mRNAs containing multiple introns. First, deposition of multiple EJCs downstream of the PTC could increase the probability that a terminating ribosome interacts with an EJC and induces NMD. Alternatively, multiple introns could enhance NMD through altered mRNA processing, for example by enhancing the loading of a protein factor onto the mRNA that stimulates NMD. In the first model, NMD efficiency is only enhanced when introns are placed downstream of a PTC, while in the second model, NMD efficiency could also be affected when introns are inserted upstream of a PTC. To distinguish between these models, we placed additional introns either upstream or downstream of the PTC in the Kif18b^{PTC-NUBP2} reporter (Fig. 6D) (which had a moderate cleavage rate and thus allows detection of both increases and decreases in the cleavage rate). Interestingly, inserting 4 additional intron cassettes upstream of the PTC enhanced NMD efficiency (Fig. 6E). The observed cleavage was still mostly dependent

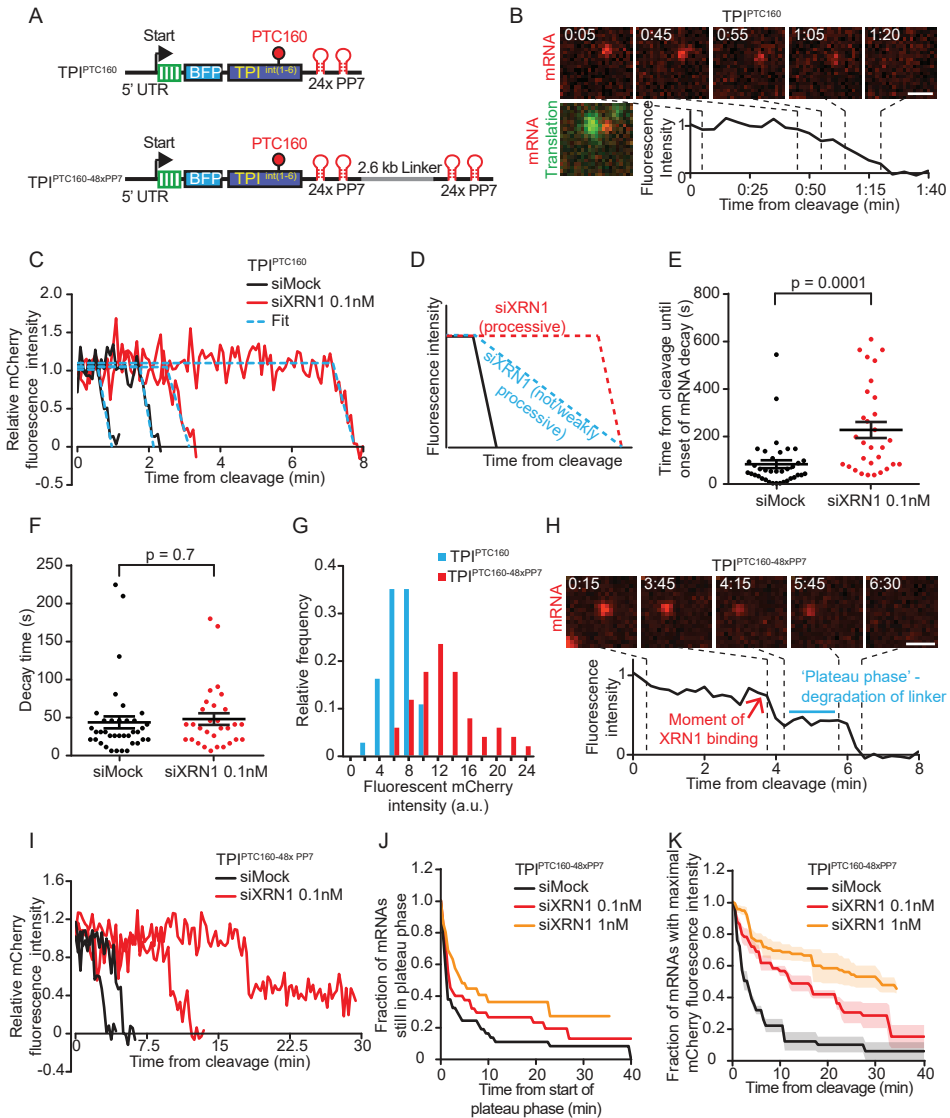


Figure 7. mRNA decay kinetics of XRN1. A) Schematics of indicated reporters. B, C, E-K) U2OS cells expressing scFv-sfGFP and PCP-Cherry-CAAX were transfected with either TPI^{PTC160} (B, C, E, F) or with $TPI^{PTC160-48xPP7}$ (G-K) reporter plasmids and with indicated siRNAs (C, E, F, I-K). Cells were analyzed by time-lapse microscopy at 5 s (D-F) or 15 s (H-K) time interval. B, H) Representative images of single mRNA molecules are shown. Scale bar, 1 μ m. Time is shown in min:sec. B, C, H, I) mCherry fluorescence intensity over time of representative example mRNAs. Dashed blue lines in (C) indicate best fit from simulations (see STAR methods). D) Schematic illustrating the expected fluorescence intensities over time upon XRN1 depletion. E-F, J, K) Quantification of the time between cleavage and onset of 3' fragment degradation (E, J), the time between onset of degradation and complete disappearance of the 3' cleavage fragment (F), and of the duration of the plateau phase (K). 3 experiments in (J) were grouped because of the low number of

on the presence of the PTC, suggesting that upstream introns did not trigger NMD through generation of new PTCs due to splicing errors (Fig. 6E). Inserting 4 intron cassettes downstream of the PTC stimulated the NMD decay rate even further (Fig. 6E), suggesting that the presence of multiple EJCs downstream of the PTC also enhances NMD induction during translation termination. Similar results were obtained when using multiple different sets of reporters (Fig. S7C-F). Enhancement of NMD by upstream introns was also observed in the genome-wide cancer dataset; when the PTC was located in the penultimate exon, we observed a significantly lower NMD efficiency when 0 or 1 upstream introns were present, than when 4 or more upstream introns were present, ($p = 0.04$ and $p = 0.001$ respectively, Mann-Whitney U test, Fig. 6F). Upstream introns did not further increase NMD efficiency when two or more introns were present downstream of the PTC (Fig. S7G). Together, these results suggest that the presence of multiple introns in an mRNA enhances NMD through two distinct mechanisms.

Decay kinetics of the 3' fragment after mRNA cleavage

After NMD-dependent mRNA cleavage, the 3' cleavage fragment is degraded by XRN1 ((Gatfield and Izaurralde, 2004) and Fig. 1J). While *in vitro* studies have shown that XRN1 displays high processivity (Jinek et al., 2011, Chang et al., 2011, Stevens, 1980), very little is known about the speed and processivity by which XRN1 degrades mRNAs *in vivo*. To examine the kinetics of XRN1-mediated mRNA degradation *in vivo*, we imaged TPI^{PTC160} mRNAs with high temporal resolution (5s interval) (Fig. 7A, B) and quantified the decrease in red fluorescence intensity of individual mRNAs over time (Fig. 7B, C). This analysis revealed an exonucleolytic decay speed of 38 nt/s (median, see STAR methods and Fig. 7C). To examine processivity of XRN1, we determined how partial depletion of XRN1 affects the onset and rate of decay; if XRN1 is highly processive, reduced levels of XRN1 would be expected to slow the onset, but not the rate of decay. In contrast, if XRN1 is non- or weakly processive, decay should initiate at approximately the same time, but show a decreased rate (Fig. 7D). We found that the onset of decay was delayed by XRN1 depletion (160 s vs 50 s in XRN1-depleted vs control cells, Fig. 7D, E), but that the decay rate was similar (median rate 38 vs 31 nucleotides/s, Fig. 7F, see STAR methods), suggesting that most TPI^{PTC160} 3' cleavage fragments are degraded by a single, processive XRN1 molecule.

To analyze XRN1 processivity more precisely, we generated a new reporter with a

Figure 7 continued... mRNAs that could be analyzed. G) mCherry fluorescence intensities of individual mRNA foci of indicated reporters were measured at the time point immediately before the mRNA was cleaved. Dots in (E, F) represent single mRNA molecules, lines show mean \pm SEM. Solid lines and corresponding shaded regions in (K) represent mean \pm SEM. Number of measurements for each experiment are listed in Table S1. See also Movie S6.

second 24x PP7 array followed by a 2.6 kb linker sequence upstream of the original 24x PP7 array (TPI^{PTC160-48xPP7}), which showed an approximately 2-fold increased mCherry fluorescence, as expected (Fig. 7A, G). The TPI^{PTC160-48xPP7} reporter has a much longer 3' UTR than TPI^{PTC160} (5kb vs 1.5 kb), increasing the likelihood of XRN1 dissociation from the 3' cleavage fragment, thus providing a more sensitive readout for XRN1 processivity. Decay of the first 24x PP7 array results in a ~50% decrease in fluorescence, after which the fluorescence intensity remains constant during the decay of the 2.6 kb linker sequence (referred to as the plateau phase), followed by another decrease in fluorescence when the second 24x PP7 array is degraded (Fig. 7H, I and Movie S6). Interestingly, when analyzing many mRNA molecules, the duration of the plateau phase showed a clear bimodal distribution; ~60% of molecules displayed rapid degradation of the linker (median time 0.8 min, Fig. 7J, black line), which is in good agreement to the predicted decay time of ~1.1 min for the linker sequence based on our measured XRN1 decay speed. The remaining molecules displayed much slower decay of the linker (median time 8.5 min). The rapidly degrading mRNA population likely represents processive degradation of the linker via a single XRN1 binding event, whereas the slowly degrading mRNAs could represent degradation that requires two or more XRN1 binding events. Consistent with this hypothesis, depletion of XRN1 did not substantially affect the fraction of rapidly degraded mRNA molecules (~50%), but dramatically increased the decay time of the slowly degrading mRNAs (Fig. 7J). Furthermore, a strong delay in the initial recruitment of XRN1 was observed by XRN1 depletion, and this delay was similar to the delay in the degradation time of the linker (Fig. 7J, K, see STAR methods). This further suggests that the slowly decaying population of mRNAs represent mRNAs in which XRN1 fell off the mRNA before completing degradation, and that the increased time required for degradation of the linker in XRN1-depleted cells is caused by slower recruitment of a new XRN1 molecule when cellular XRN1 levels are low. Together, these results show that XRN1 is a fast and highly processive enzyme *in vivo*, but many mRNAs nonetheless require two or more XRN1 binding events for complete mRNA degradation.

DISCUSSION

SunTag translation imaging – a method to study NMD

A major strength of our single molecule imaging NMD assay is that it accounts for many variable factors that can influence the steady-state mRNA levels of an NMD target. For example, delayed nuclear export or delayed translation initiation would extend the lifetime of mRNAs, making them appear less sensitive to NMD in bulk mRNA decay measurements, while our assay distinguishes between these possibilities. In addition, single molecule measurements can uncover different mRNA subpopulations, for

instance with distinct sensitivities to NMD. Although exogenous mRNA sequence elements are required for the assay (i.e. SunTag and PP7 binding sites), we find that our NMD reporter faithfully recapitulates key aspects of NMD, including the dependence on a PTC and an EJC downstream of the PTC, a requirement for the key NMD factors UPF1 and SMG6, and exonucleolytic decay of the 3' cleavage fragment by XRN1. We note that a small percentage of mRNAs (~5%) is cleaved even in the absence of a PTC or EJCs, so it is possible that these mRNAs are cleaved through EJC-independent NMD or through a mechanism other than NMD. Long 3'UTRs can stimulate NMD by increasing the distance between a PTC and the poly(A) tail (Buhler et al., 2006, Singh et al., 2008, Eberle et al., 2008). However, for the reporters tested here, the length of the 3'UTR did not play a major role in inducing NMD (Fig. 1F, S2F, G, S5G, H). Thus, our single molecule imaging method faithfully recapitulates most, if not all, aspects of NMD, and therefore adds a unique tool to study NMD timing, kinetics and heterogeneity. The mRNA cleavage and exonucleolytic decay assays developed here may also be adapted to study other forms of mRNA quality control, and more generally other aspects of RNA biology involving mRNA translation and decay.

Probability that a ribosome induces NMD during termination

We provide multiple lines of evidence that NMD occurs with equal probability during each round of translation (Fig. 2) and that NMD efficiency is not preferentially induced on CBC-bound mRNAs (Fig. 3). Nonetheless, we find that on efficient NMD-substrates up to 56% of mRNA molecules are targeted for NMD while the mRNA is still bound to the CBC. Together, these findings could reconcile the apparently contradictory observations that NMD is preferentially induced during the 'pioneer' round of translation (Ishigaki et al., 2001) and observations that NMD can be efficiently induced on eIF4E-bound mRNAs (Durand and Lykke-Andersen, 2013, Rufener and Mühlemann, 2013), thus potentially providing a unifying model for NMD induction. The observation that only a relatively small subset of termination events results in NMD may also explain why translation of upstream open reading frames (uORFs) does not result in substantial NMD (Calvo et al., 2009). If a ribosome translates a uORF before the main ORF has been translated, it would terminate on the uORF stop codon while EJCs are still associated with the main ORF, potentially triggering NMD. Since only a few termination events would occur on the uORF stop codon before the main ORF is translated and EJCs are removed, a low probability of inducing NMD for each translation termination event would largely prevent NMD on uORF-containing mRNAs, while allowing rapid NMD of mRNAs containing a bona fide PTC.

An NMD-resistant subpopulation of mRNA molecules

Most mRNA molecules are efficiently degraded by NMD, but a subpopulation of

mRNA molecules, generally ranging from 5-30%, is resistant to NMD-dependent mRNA cleavage, consistent with a previous report (Trcek et al., 2013). One possibility is that NMD-resistant mRNA molecules arise from a (stochastic) failure in splicing or EJC deposition on a subset of mRNA molecules. Consistent with this, our results show that insertion of additional introns downstream of a PTC reduces the fraction of NMD-resistant mRNAs (Fig. 6B). The presence of multiple introns would increase the fraction of mRNAs that contains at least one EJC complex downstream of the PTC. In this study, all mRNAs were expressed from a single promoter, so an interesting open question is to what extent the promoter and chromatin context affects the fidelity of splicing and EJC loading, and thus potentially the fraction of NMD-resistant mRNAs. While variability in splicing and/or EJC deposition may explain part of the NMD-resistant mRNAs, a small fraction of mRNAs escapes NMD even in reporter mRNAs that contain multiple introns located downstream of the PTC, suggesting that an additional mechanism may also contribute to NMD escape.

XRN1 speed and processivity

We found that XRN1 degrades 3' cleavage fragments with high speed and processivity, but occasionally dissociates from the mRNA. The observed XRN1 degradation speed (38-55 nt/s) is substantially higher than the translocation speed of ribosomes (9-15 nt/s), which explains why XRN1 trails ribosomes during co-translational mRNA decay (Pelechano et al., 2015). We also found that human XRN1 efficiently degrades structured RNAs with tightly bound RBPs (i.e. PP7 binding site with bound PCP), consistent with a previous study, but contrasting results obtained with yeast XRN1 (Garcia and Parker, 2015, Horvathova et al., 2017, Kim et al., 2019). Finally, a recent study found that degradation of PTC-containing mRNAs in zebrafish can lead to transcriptional adaptation, a process in which genes with sequence similarity to the degraded mRNA are upregulated, and the study showed that transcriptional adaptation was dependent on XRN1-mediated decay of NMD substrates (El-Brolosy et al., 2019). Upregulation of these genes was sequence specific, and the authors speculated that mRNA decay intermediates may play a role. Since we found that XRN1 occasionally dissociates from an mRNA during degradation, which results in production of decay intermediates, XRN1 dissociation from mRNAs during decay could be important for this process of transcriptional adaptation.

MATERIALS & METHODS*Key research table*

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	Cat# 31966021
Leibovitz's L15 medium	Gibco	Cat# 21083-027
Penicillin-Streptomycin	Gibco	Cat# 15140-122
Fetal Bovine Serum (FBS)	Sigma-Aldrich	Cat# F7524
Doxycycline	Sigma-Aldrich	Cat# D9891-1G
Opti-MEM	Sigma-Aldrich	Cat# 11058-021
FuGENE 6 Transfection Reagent	Promega	Cat# E231A
Lipofectamine RNAi-MAX	Invitrogen	Cat# 13778-075
Polyethylenimine	Polysciences Inc	Cat# 23966
Zeoцин	Invitrogen	Cat# R25001
Trimethoprim	Sigma-Aldrich	Cat# T7883-5G
TRIsure	Bioline	Cat# Bio-38033
Bioscript Reverse Transcriptase	Bioline	Cat# Bio-27036
Halo-TMR ligand	Promega	Cat# G8252
RNase Inhibitor	New England Biolabs (NEB)	Cat# M0307L
Polybrene	Santa Cruz Biotechnology, Inc	Cat# sc-134220
Critical Commercial Assays		
iQ SYBR Green SuperMix	Bio-Rad	Cat# 1708885
Experimental Models: Cell Lines		
Human U2OS cells	Tanenbaum lab	Cat# HTB-96
HEK293T cells	Tanenbaum lab	Cat# CRL-3216
Oligonucleotides		
siRNA targeting sequence-UPF1: GCAGUCCGCUCCAUUUUGAU	Dharmacon	N/A

Key resource table continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
siRNA targeting sequence-XRN1: AGAUGAACUUACCGUAGAAUU	Dharmacon	N/A
siRNA targeting sequence-SMG6: GGGUCACAGUGCUGAAGUAAU	Dharmacon	N/A
See Table S1 for all primers used for RT-qPCRs.	This study	N/A
<hr/> Recombinant DNA <hr/>		
See Table S1 for all plasmids used in the paper	This study	N/A
<hr/> Software and Algorithms <hr/>		
ImageJ	NIH	https://imagej.nih.gov/ij/
Graphpad Prism 7	GraphPad Software Inc	http://www.graphpad.com/scientific-software/prism/
Matlab R2012b	The Mathworks, Inc.	https://nl.mathworks.com/products/matlab.html
Micromanager for microscope control	Micro-Manager 1.4.22	http://micro-manager.org
NIS elements 5.11.01	Nikon	https://www.microscope.healthcare.nikon.com/en_EU/products/software
TransTrack (MatLab)	Boersma et al., 2018	https://github.com/TanenbaumLab
RiboFitter (R)	Boersma et al., 2018	https://github.com/TanenbaumLab
<hr/> Other <hr/>		
96-well glass bottom imaging plates-(Matriplates)	Brooks Life Science Systems	Cat# MGB096-1-2-LG-L
<hr/> Deposited Data <hr/>		
Raw data of imaging experiments	Mendeley data	http://dx.doi.org/10.17632/bw255hcw7h.1

U2OS and HEK293T cell culture

Human U2OS cells and HEK293T cells (ATCC) were grown in DMEM (4.5g/L glucose, Gibco) containing 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco). Cells were grown at 37°C and with 5% CO₂.

Plasmids

The complete list and sequence of all plasmids are available in the STAR methods section online at <https://doi.org/10.1016/j.molcel.2019.05.008>.

Plasmid and siRNA transfections

For imaging experiments, plasmid transfection of U2OS cells was performed in 96-well glass-bottom imaging plates 24 hr before imaging, using 0.5 µl FuGENE 6 (Promega) and 100-200 ng DNA per well. In experiments in which ha4E-BP1 was overexpressed, cells were transfected with ha4E-BP1 plasmid 16h before the start of imaging to reduce toxicity associated with overexpression of this protein. The transfection mix was prepared in OptiMEM (Sigma-Aldrich) and added to the cells in a total volume 150-200 µl of medium. Transfections in 24-well plates were performed using 1 µl FuGENE and 200-400 ng DNA per well in a total volume of 300 µl.

For experiments in which siRNA transfections and plasmid transfections were combined, U2OS cells were first reverse transfected with siRNAs at a final concentration of 10 nM (unless stated otherwise) using Lipofectamine RNAiMAX (Invitrogen) and seeded in plastic 24-well plates. After 24hr, the cells were trypsinized, transfected with a second dose of 10 nM siRNA and re-plated in 96-well glass-bottom imaging plates. 48 hr after the first siRNA transfection, cells were transfected with plasmid DNA, as described above. 24 hr after DNA transfection, cells were analyzed by time-lapse microscopy. The sequences of the siRNAs used in this study are listed in the Key Resource table.

For generation of cells stably expressing reporter mRNAs, U2OS cells were transfected with indicated reporter plasmids. 24 hr after transfection, selection for stable integration was performed using 0.4mg/ml Zeocin (Invitrogen) for 10 days.

Lentivirus production and infection

For lentivirus production, HEK293T cells were transfected with the lentiviral vector along with lentiviral packaging plasmids pMD2.g and pspax2 using Polyethylenimine (PEI) (Polysciences Inc). The medium was replaced the day after transfection with fresh culture medium, and 72 hr after transfection, viral supernatant was collected. For lentiviral infections, cells were seeded in a 6-well plate at about 70% confluency. Viral supernatant was added to the cells along with Polybrene (10µg/ml) (Santa Cruz Biotechnology Inc) and the cells were spun at 2000 rpm for 90 min at 22°C (Spin-infection). After the spin-infection, the culture medium was replaced with fresh medium, and cells were incubated for at least 48 hr before further analysis.

Microscopy

Unless stated otherwise, all live-cell imaging experiments were performed using U2OS cells expressing TetR, scFv-sfGFP and PCP-mCherry-CAAX (Yan et al., 2016, Ruijtenberg et al., 2018). Cells were seeded 48h before imaging in 96-well glass bottom dishes (Matriplates, Brooks Life Science Systems) at 20-25% confluency. Cells were transfected with reporter plasmid DNA 24h before imaging. Thirty minutes before imaging, the cell culture medium was replaced with pre-warmed CO₂-independent Leibovitz's-15 medium (Gibco) and transcription of the reporters was induced by addition of doxycycline (1 µg/ml) (Sigma-Aldrich). Images were acquired using a Nikon TI inverted microscope with perfect focus system equipped with a Yokagawa CSU-X1 spinning disc, a 100x 1.49 NA objective and an iXon Ultra 897 EM-CCD camera (Andor) using Micro-Manager software (Edelstein et al., 2010) and NIS software (Nikon). During the experiment, cells were maintained at a constant temperature of 37°C. Unless stated otherwise, single Z-plane images were acquired, with the bottom of the cell in the focal plane. Camera exposure times of 500 ms were used for both GFP and mCherry, and images were acquired with an interval of 30 s, unless stated otherwise. Of note, mCherry-positive lysosomes were visible in most cells, but these could easily be distinguished from mRNAs based on fluorescence intensity and diffusion kinetics. In experiments in which untethered mRNAs were tracked, U2OS cells expressing PCP-Halo (instead of PCP-mCherry-CAAX) were used. Cells were labelled with Halo-TMR ligand (Promega) (50nM concentration for 2 h) before imaging. Single Z-plane images were acquired, with the region just below the nucleus of the cell in the focal plane. Images were taken with an interval of 15 s and camera exposure times of 500 ms were used for both GFP and TMR.

In experiments in which the intensity of green spots was measured in 3D before and after cleavage, Z-stacks were acquired for GFP. We acquired 11 slices with an inter-slice distance of 1 µm each, and used a 100 ms exposure time. For mCherry, a single Z-plane was imaged with 500 ms exposure time. Images were acquired at a 10 s time interval.

In experiments in which the intensity of red spots was measured after mRNA cleavage, low laser power (~8x lower than used for other imaging) and high exposure times (1500 ms) were used for mCherry to reduce photobleaching and increase signal-to-noise. This enabled accurate detection and measurement of mCherry foci for >300 time points.

Quantitative RT-PCR

U2OS cells stably expressing TetR or TetR, scFv-sfGFP, PCP-mCherry-CAAX (Yan et al., 2016) were seeded in 24-well plastic bottom plates at ~10% confluency 72h before harvesting of cells. When the effect of UPF1 siRNA on reporter expression was assessed, a reverse transfection with 10 nM siRNA against UPF1 was performed during seeding, and another siRNA transfection was performed 24h after seeding. 48h

after seeding, equal amounts of TPI^{WT}, TPI^{PTC160} or TPI^{PTC1} reporter constructs were co-transfected with a control plasmid, and doxycycline was added for 24h to induce transcription of the reporters. 72hr after cell seeding, RNA was isolated using RNeasy plus mini kit (Qiagen) according to manufacturer's guidelines, and cDNA was generated using Bioscript reverse transcriptase (Bioline) and Oligo-d(T) primers. qPCRs were performed using SYBR-Green Supermix (Bio-Rad) on a Bio-Rad Real time PCR machines (CFX Connect Real-Time PCR Detection System). RNA abundance of reporter mRNAs was measured using two different primer sets that amplified a ~200 nt regions upstream or downstream of the PTC. Reporter mRNA abundance was normalized to the expression of the control plasmid that was co-transfected to control for differences in transfection efficiency. The average of the two primer sets was then used as the final value for mRNA abundance.

For checking efficiencies of UPF1 and XRN1 siRNAs by qPCR, U2OS cells expressing TetR, scFv-sfGFP and PCP-mCherry-CAAX (Yan et al., 2016) were seeded in 24-wells plates. Respective siRNAs (10mM) were transfected during cell plating using reverse translation and cells were harvested 3 days after transfection. RNA isolation was performed using TRIsure (Bioline) according to manufacturer's protocol. cDNA synthesis and qPCRs were performed as described above, except that GAPDH mRNA levels were used to normalize mRNA abundance.

For determining splicing efficiency of NMD reporters, U2OS cells expressing TetR were seeded in 24-wells plates at 20-25% confluency 48h before harvesting. After 24h cells were transfected with either a kif18b reporter that contained an intron or a matched reporter in which the intron sequence was removed from the plasmid. 3h before harvesting of cells, doxycycline was added to the cell culture medium to induce transcription, and 200 µg/ml cycloheximide was added to prevent degradation of spliced transcripts by NMD. RNA was isolated using RNeasy plus mini kit (Qiagen) with on-column DNase treatment (RNase-free DNase set, Qiagen) according to manufacturer's protocol. cDNA synthesis and qPCR were performed as described above. To assess splicing efficiency, we used a primer set that amplified the reporter mRNA independent of its splicing status (total transcript) and a primer set for which one primer binds at the exon-exon junction, which only generates a PCR product when the transcript is spliced. The no-intron control reporter has the same mRNA sequence as the spliced transcript, and should therefore be amplified by both primer sets. We compared the ratio in abundance of the two amplicons (e.g. 'total' and 'spliced'), and normalized this ratio to the ratio of total and spliced amplicons obtained with the no-intron control reporter.

To ensure that the 'spliced' mRNA-specific primer set is indeed specific to spliced mRNAs, we tested the spliced mRNA-specific primer set on plasmid DNA. Amplification was tested on plasmid DNA in which the intron was present, which resembles unspliced mRNA and should thus not be amplified, and plasmid DNA in which the intron was

not present, which resembles spliced mRNA and should be efficiently amplified. Each spliced mRNA-specific primer set amplified plasmid DNA lacking an intron >500 fold more efficiently than plasmid DNA containing an intron, confirming the specificity of these primer sets.

Flow Cytometry

For analysis of splicing by flow cytometry using fluorescence splicing reporters, U2OS cells were seeded in 24-well plates at 20-25% density. 24 hr after seeding, cells were transfected with plasmids encoding the splicing reporters and 1 μ g/ml doxycycline was added to induce expression of the reporter. 48 hr after seeding, cells were harvested and analyzed for GFP and BFP expression by flow cytometry using a Cytoflex analyzer (Beckman Coulter). The ratio of BFP-to-GFP signal intensity was then determined for each cell and the average BFP/GFP ratio for the cell population was calculated for each reporter. The BFP/GFP ratio of intron-containing reporters was then normalized to the average ratio of BFP/GFP of 6 no-intron control reporters (which represent fully spliced mRNAs) to determine the splicing efficiency.

Quantitation, statistical analysis, and modelling

For detailed description of quantitation, statistical analysis and modelling, see STAR methods section online at <https://doi.org/10.1016/j.molcel.2019.05.008>.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.molcel.2019.05.008>

ACKNOWLEDGEMENTS

We thank Niels Gehring for sharing TPI plasmids (Addgene 65801 and Addgene 65804), and we would like to thank members of the Tanenbaum lab for helpful discussions. We would also like to thank Suzan Ruijtenberg, Matilde Galli, Lotte van Rijnberk, and Angela Andersen from Life Science Editors for critical reading of the manuscript. This work was financially supported by an ERC starting grant (ERC-STG 677936-RNAREG), two grants from the Netherlands Organization for Scientific Research (NWO) (ALWOP.290 and NWO/016.VIDI.189.005) and by the Howard Hughes Medical Institute through an International Research Scholar grant to M.E.T. (HHMI/IRS 55008747). All authors are supported by the Oncode Institute that is partly funded by the Dutch Cancer Society (KWF). M.V. also acknowledges support

by the gravitation program CancerGenomiCs.nl from the Netherlands Organization for Scientific Research (NWO). T.A.H. was supported by a PhD fellowship from the Boehringer Ingelheim Fonds.

AUTHOR CONTRIBUTIONS

M.E.T. and T.A.H. conceived of the project. T.A.H. and D.K. performed all the experiments. R.G.H.L. performed bioinformatics analysis under supervision of M.V. S.S. and B.V. developed computational tools. S.B. provided reagents. T.A.H., D.K. and M.E.T. analyzed the data, prepared the figures and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Belgrader, P., Cheng, J., Zhou, X., Stephenson, L. S. & Maquat, L. E. 1994. Mammalian nonsense codons can be cis effectors of nuclear mrna half-life. *Mol. Cell. Biol.*, *14*, 8219-8228.
- Boehm, V., Haberman, N., Ottens, F., Ule, J. & Gehring, N. H. 2014. 3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons. *Cell reports*, *9*, 555-568.
- Boersma, S., Deepak, K., Verhagen, B. M. P., Sonneveld, S., Grimm, J. B., Lavis, L. D. & Tanenbaum, M. E. 2018. Multi-color single molecule imaging uncovers extensive heterogeneity in mrna decoding. *Biorxiv*, <https://doi.org/10.1101/477661>.
- Buhler, M., Steiner, S., Mohn, F., Paillusson, A. & Muhlemann, O. 2006. EJC-independent degradation of nonsense immunoglobulin- μ mrna depends on 3' UTR length. *Nat. Struct. Mol. Biol.*, *13*, 462-4.
- Calvo, S. E., Pagliarini, D. J. & Mootha, V. K. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. USA*, *106*, 7507-7512.
- Chang, J. H., Xiang, S., Xiang, K., Manley, J. L. & Tong, L. 2011. Structural and biochemical studies of the 5'→3' exoribonuclease Xrn1. *Nat. Struct. Mol. Biol.*, *18*, 270-6.
- Chao, J. A., Patskovsky, Y., Almo, S. C. & Singer, R. H. 2008. Structural basis for the coevolution of a viral RNA-protein complex. *Nat. Struct. Mol. Biol.*, *15*, 103-5.
- Cheng, J., Belgrader, P., Zhou, X. & Maquat, L. E. 1994. Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mrna abundance. *Mol. Cell. Biol.*, *14*, 6317-6325.
- Cheng, J. & Maquat, L. E. 1993. Nonsense codons can reduce the abundance of nuclear mrna without affecting the abundance of pre-mrna or the half-life of cytoplasmic mrna. *Mol. Cell. Biol.*, *13*, 1892-1902.
- Dostie, J. & Dreyfuss, G. 2002. Translation is required to remove Y14 from mRNAs in the cytoplasm. *Curr. Biol.*, *12*, 1060-1067.
- Durand, S. & Lykke-Andersen, J. 2013. Nonsense-mediated mrna decay occurs during eif4f-dependent translation in human cells. *Nat. Struct. Mol. Biol.*, *20*, 702-9.
- Eberle, A. B., Lykke-Andersen, S., Mühlemann, O. & Jensen, T. H. 2009. SMG6 promotes endonucleolytic cleavage of nonsense mrna in human cells. *Nat. Struct. Mol. Biol.*, *16*, 49-55.
- Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z. & Muhlemann, O. 2008. Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *Plos Biol.*, *6*, e92.
- Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. 2010. Computer control of microscopes using μ manager. *Curr. Protoc. Mol. Biol.*, *14*.20. 1-14.20. 17.
- El-Brolosy, M. A., Kontarakis, Z., Rossi, A., Kuenne, C., Gunther, S., Fukuda, N., Kikki, K., Boezio, G. L. M., Takacs, C. M., Lai, S. L., Fukuda, R., Gerri, C., Giraldez, A. J. & Stainier, D. Y. R. 2019. Genetic compensation triggered by mutant mrna degradation. *Nature*, *568*, 193-197.
- Garcia, J. F. & Parker, R. 2015. MS2 coat proteins bound to yeast mRNAs block 5' to 3' degradation and trap mrna decay products: implications for the localization of mRNAs by MS2-MCP system. *RNA*, *21*, 1393-5.
- Gatfield, D. & Izaurralde, E. 2004. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature*, *429*, 575-8.
- He, F. & Jacobson, A. 2015. Nonsense-mediated mrna decay: degradation of defective transcripts is only

part of the story. *Annu. Rev. Genet.*, *49*, 339-366.

Horvathova, I., Voigt, F., Kotrys, A. V., Zhan, Y., Artus-Revel, C. G., Eglinger, J., Stadler, M. B., Giorgetti, L. & Chao, J. A. 2017. The dynamics of mrna turnover revealed by single-molecule imaging in single cells. *Mol. Cell*, *68*, 615-625. E9.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J. & Izaurralde, E. 2008. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA*, *14*, 2609-2617.

Ishigaki, Y., Li, X., Serin, G. & Maquat, L. E. 2001. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell*, *106*, 607-617.

Jinek, M., Coyle, S. M. & Doudna, J. A. 2011. Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay. *Mol. Cell*, *41*, 600-608.

Karousis, E. D., Nasif, S. & Mühlemann, O. 2016. Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley Interdiscip. Rev. RNA*, *7*, 661-682.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G. & Ohno, S. 2006. Binding of a novel SMG-1-Upf1-erf1-erf3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.*, *20*, 355-367.

Kim, S. H., Vieira, M., Kim, H. J., Kesawat, M. S. & Park, H. Y. 2019. MS2 Labeling of Endogenous Beta-Actin mRNA Does Not Result in Stabilization of Degradation Intermediates. *Mol. Cells*.

Kurosaki, T. & Maquat, L. E. 2016. Nonsense-mediated mRNA decay in humans at a glance. *J. Cell Sci.*, *129*, 461-467.

Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.*, *19*, 6860-6869.

Lejeune, F., Ishigaki, Y., Li, X. & Maquat, L. E. 2002. The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.*, *21*, 3536-3545.

Lindeboom, R. G., Supek, F. & Lehner, B. 2016. The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat. Genet.*, *48*, 1112-8.

Loh, B., Jonas, S. & Izaurralde, E. 2013. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.*, *27*, 2125-2138.

Lykke-Andersen, S. & Jensen, T. H. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat. Rev. Mol. Cell Biol.*, *16*, 665-77.

Maquat, L. E., Tarn, W.-Y. & Isken, O. 2010. The pioneer round of translation: features and functions. *Cell*, *142*, 368-374.

Morisaki, T., Lyon, K., Deluca, K. F., Deluca, J. G., English, B. P., Zhang, Z., Lavis, L. D., Grimm, J. B., Viswanathan, S. & Looger, L. L. 2016. Real-time quantification of single RNA translation dynamics in living cells. *Science*, *352*, 1425-1429.

Mort, M., Ivanov, D., Cooper, D. N. & Chuzhanova, N. A. 2008. A meta-analysis of nonsense mutations causing human genetic disease. *Hum. Mutat.*, *29*, 1037-1047.

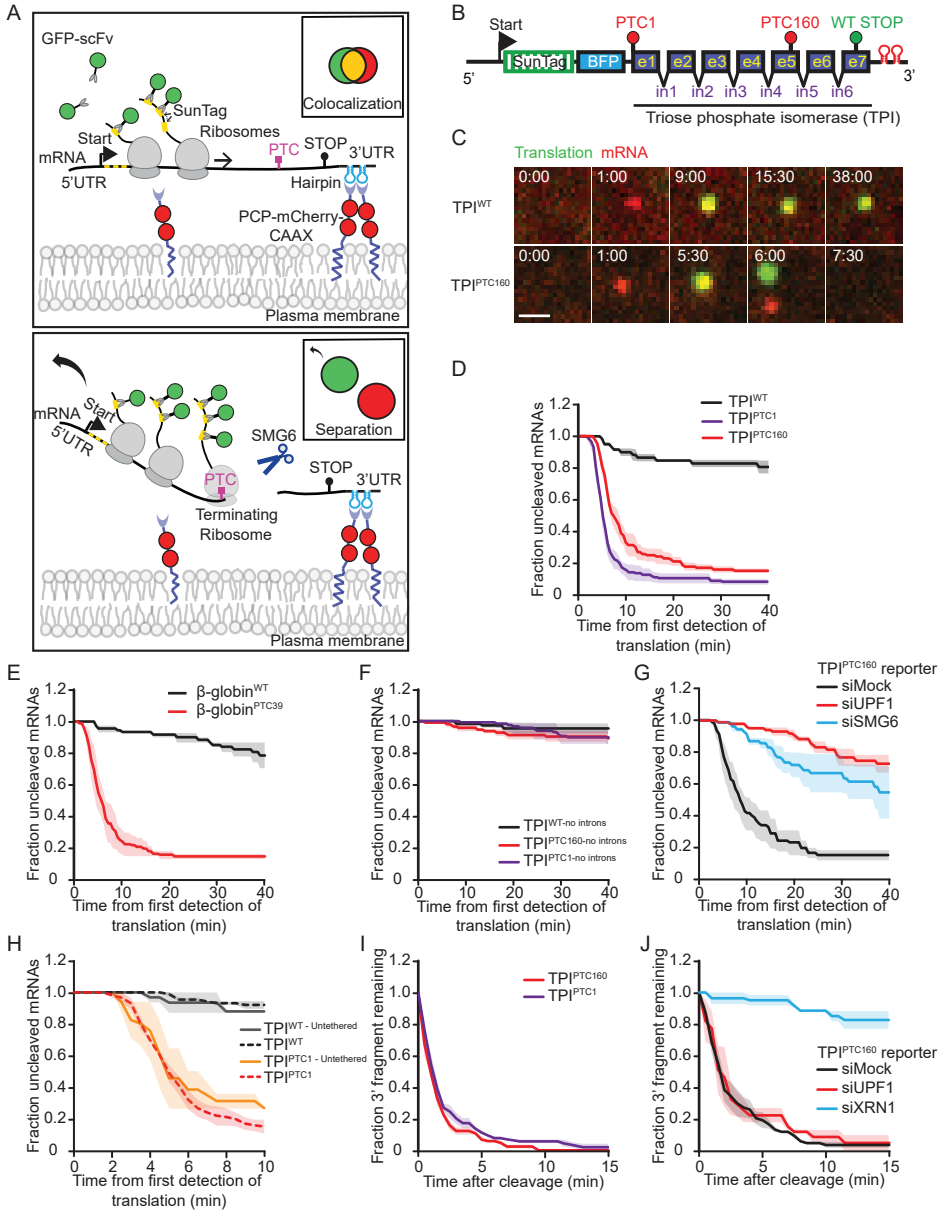
Pelechano, V., Wei, W. & Steinmetz, L. M. 2015. Widespread co-translational RNA decay reveals ribosome dynamics. *Cell*, *161*, 1400-1412.

- Pichon, X., Bastide, A., Safieddine, A., Chouaib, R., Samacoits, A., Basyuk, E., Peter, M., Mueller, F. & Bertrand, E. 2016. Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *J. Cell Biol.*, 214, 769 – 781.
- Rufener, S. C. & Mühlemann, O. 2013. Eif4e-bound mrnps are substrates for nonsense-mediated mrna decay in mammalian cells. *Nat. Struct. Mol. Biol.*, 20, 710-7.
- Ruijtenberg, S., Hoek, T. A., Yan, X. & Tanenbaum, M. E. 2018. Imaging translation dynamics of single mrna molecules in live cells. *RNA Detection. Springer*.
- Sakbarkar, M. K., Chom, V. T. & Kangueane, P. 2004. Distributions of exons and introns in the human genome. In *Silico Biol.*, 4, 387-393.
- Sato, H. & Maquat, L. E. 2009. Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin β . *Genes Dev.*, 23, 2537-2550.
- Saulière, J., Murigneux, V., Wang, Z., Marquet, E., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Crollius, H. R. & Le Hir, H. 2012. CLIP-seq of eif4aiii reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct. Mol. Biol.*, 19, 1124-31.
- Singh, G., Kucukural, A., Cenik, C., Leszyk, J. D., Shaffer, S. A., Weng, Z. & Moore, M. J. 2012. The cellular EJC interactome reveals higher-order mrnp structure and an EJC-SR protein nexus. *Cell*, 151, 750-764.
- Singh, G., Rebbapragada, I. & Lykke-Andersen, J. 2008. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mrna decay. *PLoS Biol.*, 6, e111.
- Stevens, A. 1980. Purification and characterization of a *Saccharomyces cerevisiae* exoribonuclease which yields 5'-mononucleotides by a 5' leads to 3' mode of hydrolysis. *J. Biol. Chem.*, 255, 3080-3085.
- Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, 159, 635-646.
- Tanenbaum, M. E., Stern-Ginossar, N., Weissman, J. S. & Vale, R. D. 2015. Regulation of mrna translation during mitosis. *Elife*, 4.
- Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeyer, C., Hentze, M. W. & Kulozik, A. E. 1998. Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.*, 17, 3484-3494.
- Trcek, T., Sato, H., Singer, R. H. & Maquat, L. E. 2013. Temporal and spatial characterization of nonsense-mediated mrna decay. *Genes Dev.*, 27, 541-551.
- Tutucci, E., Vera, M., Biswas, J., Garcia, J., Parker, R. & Singer, R. H. 2018. An improved MS2 system for accurate reporting of the mrna life cycle. *Nat Methods*, 15, 81-89.
- Unterholzner, L. & Izaurralde, E. 2004. SMG7 acts as a molecular link between mrna surveillance and mrna decay. *Mol. Cell*, 16, 587-596.
- Wang, C., Han, B., Zhou, R. & Zhuang, X. 2016. Real-time imaging of translation on single mrna transcripts in live cells. *Cell*, 165, 990-1001.
- Wu, B., Eliscovich, C., Yoon, Y. J. & Singer, R. H. 2016. Translation dynamics of single mrnas in live cells and neurons. *Science*, 352, 1430-1435.
- Yan, X., Hoek, T. A., Vale, R. D. & Tanenbaum, M. E. 2016. Dynamics of translation of single mrna molecules in vivo. *Cell*, 165, 976-989.
- Yanagiya, A., Suyama, E., Adachi, H., Svitkin, Y. V., Aza-Blanc, P., Imataka, H., Mikami, S., Martineau, Y.,

Ronai, Z. A. & Sonenberg, N. 2012. Translational homeostasis via the mrna cap-binding protein, eif4e. *Mol. Cell*, 46, 847-58.

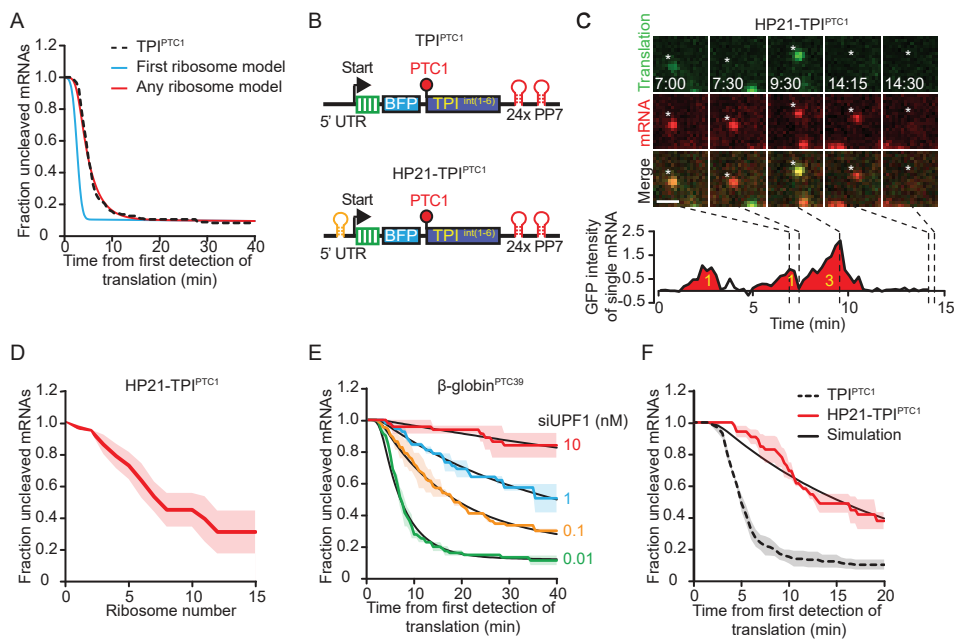
SUPPLEMENTAL FIGURES

For Table S1 go to <https://www.sciencedirect.com/science/article/pii/S1097276519303612>

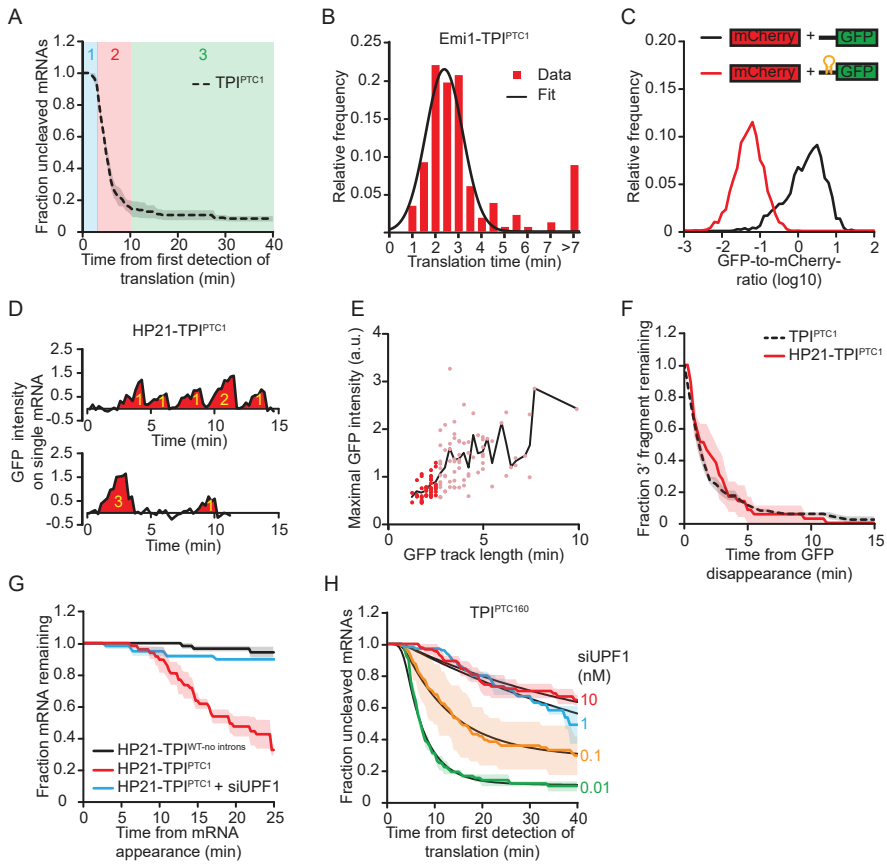


Supplemental figure S1 – related to figure 1. The NMD imaging approach faithfully recapitulates

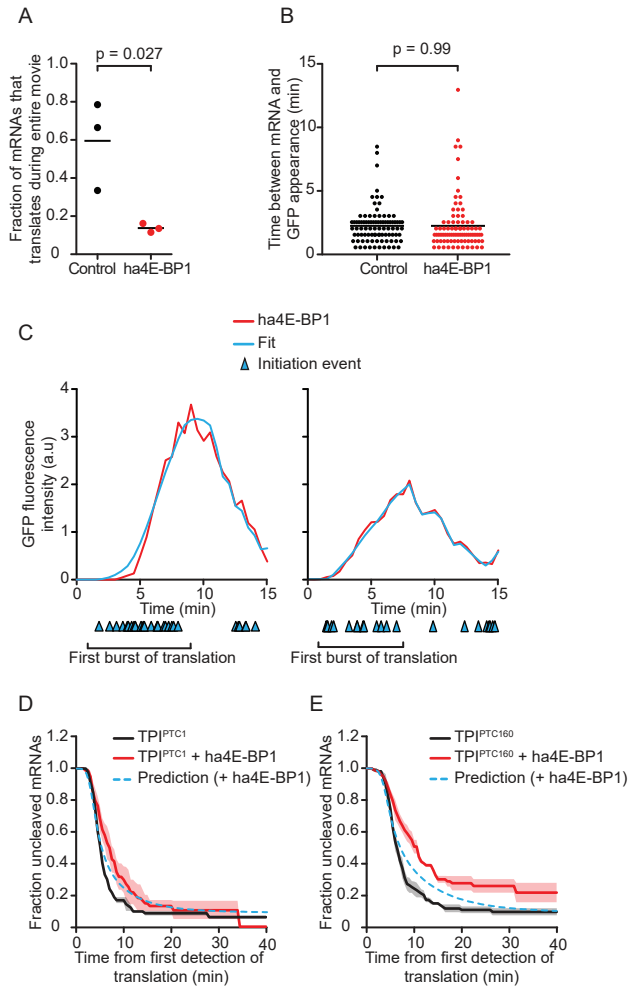
key aspects of NMD. A-C) U2OS cells (A) or U2OS cells stably expressing scFv-sfGFP and PCP-mCherry-CAAX (A-C) were transfected with indicated NMD reporter plasmid and a control plasmid (A-B) and/or with indicated siRNA (B-C). mRNA abundance was assessed by qPCR and normalized to control plasmid (A-B) or GAPDH (C). Dots represent individual experiments, and the mean value is represented by a horizontal line. (D-K) U2OS cells stably expressing scFv-sfGFP and PCP-mCherry-CAAX were transfected with indicated reporters and analyzed by time-lapse microscopy. D) The fraction of mRNAs appearing in the field of view without an associated GFP signal was quantified for indicated reporters. Note that only mRNAs were included in this analysis that showed translation at some point in their life-time. E) The time between appearance of the mRNA in the field of view and appearance of the first GFP signal was quantified for indicated reporters. F) GFP fluorescence intensities were measured over time for indicated reporter mRNAs. Traces were aligned at the moment of first detection of GFP signal. Thick lines represent average intensity of all traces, thin lines represent intensity traces of individual, representative mRNAs. G) GFP fluorescence intensities of indicated reporters were measured at the time-point immediately before cleavage for indicated reporter mRNAs. The number of ribosomes per mRNA was then determined for each mRNA (see STAR methods). Individual mRNAs (dots) and mean values (lines) are plotted. H) Z-stacks were acquired every 10 seconds, and the GFP fluorescence intensity of individual mRNAs was measured in maximum intensity projections just prior to and immediately after mRNA cleavage (indicated by horizontal dotted line). After cleavage, the GFP spot (which no longer co-localized with the mRNA (red) spot) was tracked in 3D and its intensity was measured irrespective of its position relative to the red mRNA spot (i.e. 3' mRNA fragment). Note that the GFP signal remained constant after separation of green and red foci, indicating that all ribosomes on the mRNA molecule remained together after separation of green and red foci, consistent with mRNA cleavage rather than with sequential termination of all individual ribosomes. The thick line represents the average intensity of all traces, thin lines represent intensity traces of individual, representative mRNAs. I) GFP fluorescence intensity of individual mRNAs was measured using an ROI centered on the location of the red mRNA spot (or 3' mRNA fragment after cleavage). For TPI^{WT} mRNAs, harringtonine was added 5 minutes after start of imaging to create a situation in which the GFP intensity associated with the mRNA decreased due to translation termination. GFP fluorescence intensity traces of TPI^{WT} were aligned to the moment of harringtonine addition. For TPI^{PTC160} mRNAs, GFP fluorescence intensity traces were aligned to the moment of mRNA cleavage. Note that upon cleavage of TPI^{PTC160} mRNAs, the GFP fluorescence intensity drops in a single step (as the GFP spot no longer co-localizes with the mRNA), whereas the GFP signal gradually decreased over multiple time-points in harringtonine-treated cells, indicating that red-and-green foci separation is distinct from translation termination. Thick lines represents the average intensity of all traces, thin lines represent intensity traces of individual mRNAs. J, K) The time from first detection of translation until cleavage was determined for indicated reporters. Cells in (K) were treated with UPF1 siRNA 72 hr before imaging, where indicated. II solid lines and corresponding shaded regions represent mean \pm SEM. Dotted line in (J) indicates that data is replotted from an earlier figure panel. Number of measurements for each experiment are listed in Table S1.



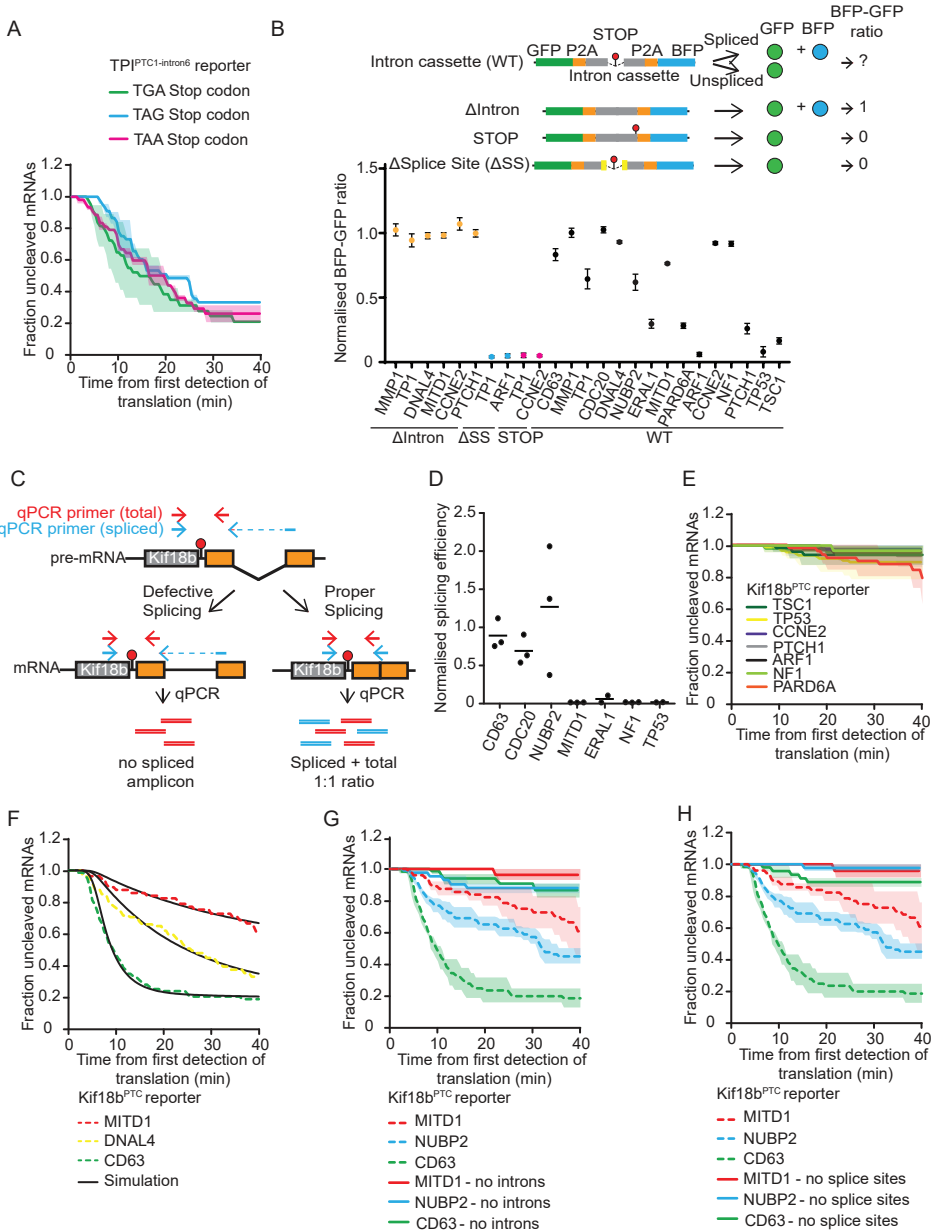
Supplemental figure S2 – related to figure 1. Additional controls to validate the NMD imaging approach. A-I) U2OS cells stably expressing scFv-sfGFP and PCP-mCherry-CAAX were transfected with indicated reporters and analyzed by time-lapse microscopy. The time from first detection of translation until cleavage was determined (A-C, H, I). A-C) Variation in cleavage efficiency between individual cells expressing TPI^{WT} (A), TPI^{PTC160} (B), or TPI^{PTC1} (C) was examined. Individual cells in which at least 4 mRNAs were observed (thin lines) and the average of all cells (thick dotted lines) are shown. Asterisk in B) indicates a cell that shows a single outlier cell that shows inefficient cleavage. D) Stochastic simulations were performed to determine the expected variation in cleavage kinetics among different single cells (see STAR methods). The cell-to-cell variation was quantified by calculating the summed squared error (SSE) between each individual cell in a simulation and the experimental average of all cells. This simulation was repeated 10,000 times to obtain a 95% confidence interval of the expected variation, which is indicated by error bars. Experimentally observed cell-to-cell variation (red dots) was similarly determined by calculating the SSE between individual cells and the experimental average of all cells. We observed a cell-to-cell variation that deviated substantially from the expected variation only for the experiments using the TPI^{PTC160} reporter. However, removal of a single outlier cell (indicated by asterisk in B)) eliminated this effect (TPI^{PTC160} – outlier cell removed). E) GFP fluorescence intensities associated with mRNAs of indicated reporters were measured 5 minutes after first detection of translation. GFP fluorescence intensities of mRNAs that would eventually be cleaved were compared to GFP intensities of mRNAs that were not cleaved over the duration of the experiment. Individual mRNAs (dots) and mean values (lines) are plotted. F, G) The time between appearance and disappearance of a GFP fluorescence signal was quantified for indicated reporters. Raw time-lapse microscopy data of cells expressing TPI^{PTC160} and TPI^{WT} reporters (Fig. 1D) (grey and black lines) was re-analyzed using similar criteria as the analysis of the images of cells expressing 5xPP7 reporter constructs (orange and red lines, see STAR methods). G) Normalization of the data in (F) was performed. For each time-point, the fraction of remaining GFP foci for TPI^{PTC160} and $TPI^{PTC160-5xPP7}$ reporters was divided by the fraction of remaining GFP foci of the corresponding control reporters (TPI^{WT} and $TPI^{WT-5xPP7}$, respectively, see STAR methods). H, I) mRNA cleavage for indicated reporters was compared in either cells transiently transfected or stably expressing the same reporter. G) Cells were transfected with XRN1 siRNA, or were mock transfected. 72 hr after transfection, cells were harvested and XRN1 mRNA abundance was determined by qPCR. Individual experiments (dots) and mean values (lines) are plotted. All solid lines and corresponding shaded regions represent mean \pm SEM. Dotted lines indicate that data is replotted from an earlier figure panel. Number of measurements for each experiment are listed in Table S1



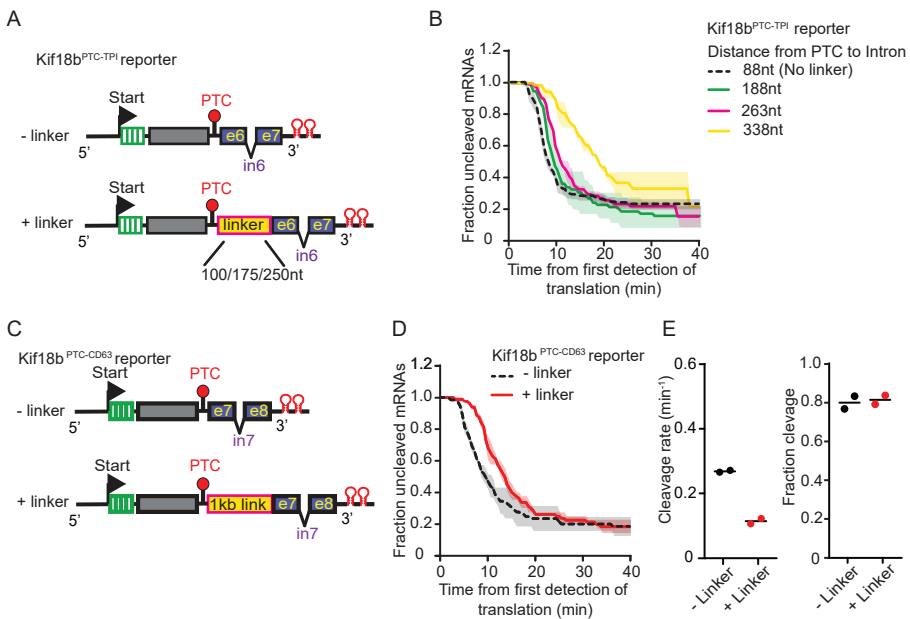
Supplemental figure S3 – related to figure 2. Multiple approaches reveal that NMD occurs with equal probability during each round of translation. A-B, D-H) U2OS cells expressing scFv-sfGFP, PCP-mCherry-CAAX and indicated reporter plasmids were analyzed by time-lapse microscopy. A) Cleavage time distribution of TPI^{PTC1} (replotted from fig. 1D) is shown. Three distinct phases in the cleavage time distribution are highlighted in different colors and numbered. B) GFP fluorescence intensity of $Emi1-TPI^{PTC1}$ mRNAs was analyzed, and the duration of peaks of GFP fluorescence were quantified. GFP peaks shorter than 3 frames or with high fluorescence intensity (>3-fold higher than the typical translation events) were excluded as they likely represent mRNAs translated by multiple ribosomes. The mean and variance of the elongation time were calculated from all traces of 4.5 minutes or less (as longer traces are likely to represent multiple ribosomes) by fitting the data with a Gaussian distribution (black line). C) U2OS cells in which two plasmids, 1) mCherry and 2) either sfGFP (black line) or HP21-sfGFP (red line), were co-transfected and analyzed by FACS. The ratio of GFP and mCherry fluorescence intensities were calculated for all mCherry positive cells (averaged data of 2 independent experiments). D) GFP intensity of HP21- TPI^{PTC1} reporter mRNAs was determined over time. Two example traces of representative mRNAs are shown, in addition to the trace shown in Fig. 2C). Red filled areas represent peaks that were called as translation events. Yellow numbers indicate the number of ribosomes that contributed to the peak. E) Maximal GFP fluorescence intensity and duration of individual GFP peaks are plotted. Dark red dots indicate GFP peaks that likely originate from single ribosomes, while pink dots indicate peaks that may originate from more than one ribosomes based on the longer track length and/or higher GFP intensity. Black line shows the average of maximal GFP intensities for different track lengths F) Quantification of the time between green and red foci separation and disappearance of the mRNA signal (TPI^{PTC1}), or between disappearance of the last GFP signal and disappearance of the mRNA signal (HP21- TPI^{PTC1}). Kinetics of mRNA disappearance are very similar for both conditions, suggesting that disappearance of the last GFP signal of HP21- TPI^{PTC1} followed by disappearance of the mRNA spot is an accurate readout for cleavage. G) Quantification of the time that indicated reporter mRNAs could be tracked. Cells were transfected with UPF1 siRNA 72h before imaging, where indicated. mRNA disappearance was rare in cells in which UPF1 was depleted by siRNA or in cells expressing a reporter lacking introns and a PTC, suggesting that mRNA disappearance is caused by NMD. H) The time from first detection of translation until mRNA cleavage was determined for indicated reporters. Cells were transfected with indicated concentrations of UPF1 siRNA for 72h. Black lines indicate the best fit from stochastic simulations based on the model in which NMD occurs with equal probability during each round of translation. All solid lines and corresponding shaded regions represent mean \pm SEM. Dashed line in (F) indicates that data is replotted from an earlier figure panel. Number of measurements for each experiment are listed in Table S1



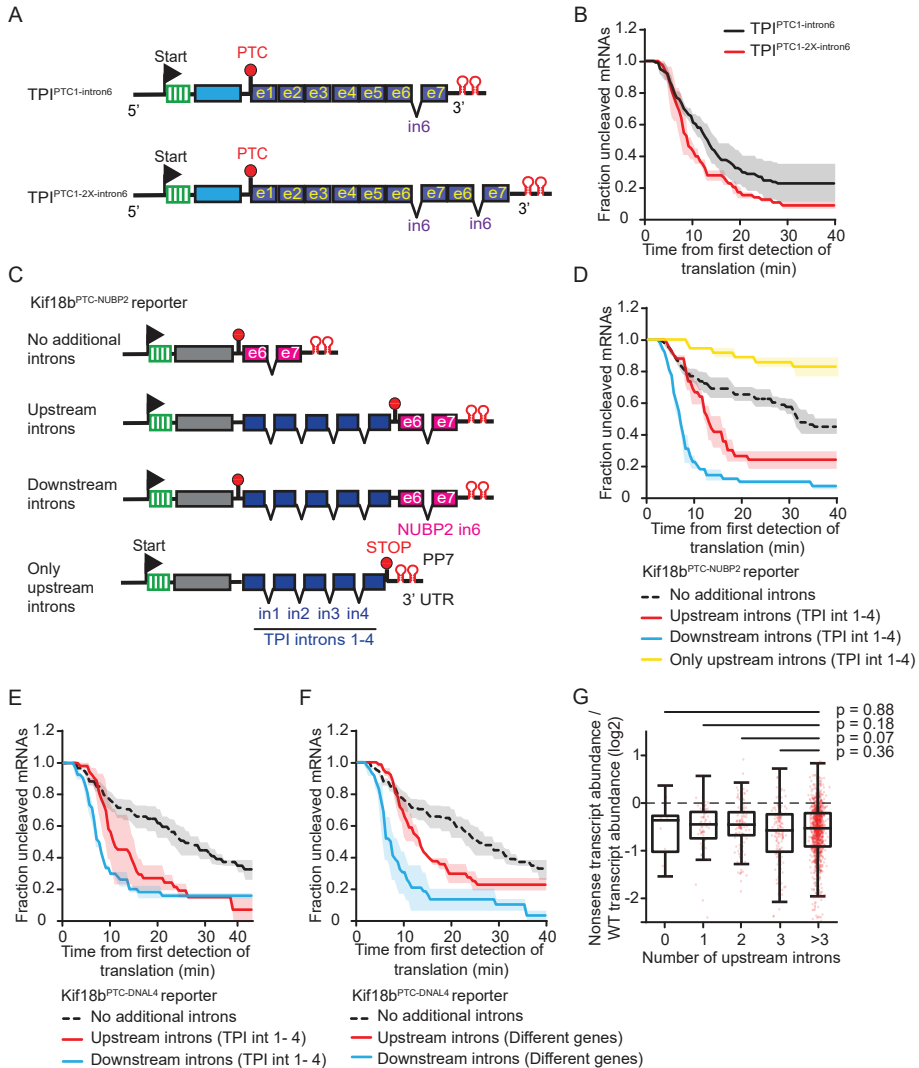
Supplemental figure S4 – related to figure 3. Kinetics of translation and NMD upon ha4E-BP1 expression. A-E) U2OS cells expressing scFv-sfGFP, PCP-mCherry-CAAX and XXXXX (A-C) or indicated reporters (D, E) were transfected with ha4E-BP1 or mock transfected and analyzed by time-lapse microscopy. A) The fraction of mRNAs is shown for which a GFP fluorescence signal could be detected for the entire duration of the movie. Only mRNAs on which the GFP fluorescence signal appeared at least 15 minutes before the end of imaging were included. Dots represent individual experiments and the mean value is represented by a horizontal line. B) The time between appearance of the mRNA in the field of view and appearance of the first GFP signal was quantified for indicated reporters. C) Two additional example traces of representative mRNAs of cells expressing ha4E-BP1 (as in Fig 3E). Red lines indicate the experimentally observed GFP intensity over time. Blue lines indicate the best fit from simulations. Blue triangles indicate translation initiation events. D-E) Quantification of time between first detection of translation and mRNA cleavage for indicated reporters with (red line) or without (black line) overexpression of ha4E-BP1. Dotted blue lines indicate the cleavage kinetics that are predicted upon ha4E-BP1 overexpression based on the reduced initiation rate in fig 3F (STAR methods). All solid lines and corresponding shaded regions represent mean \pm SEM. Number of measurements for each experiment are listed in Table S1



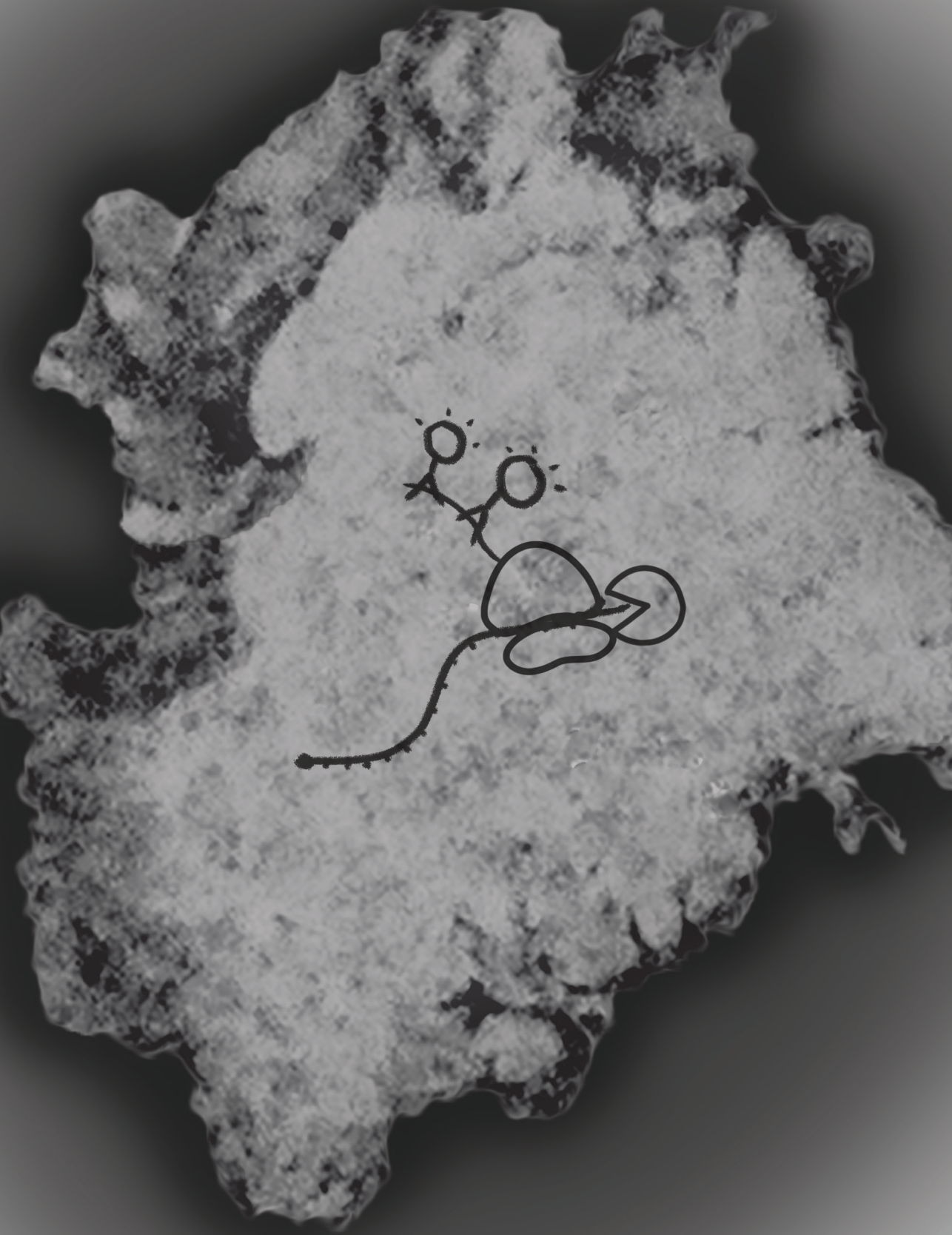
Supplemental figure S5 – related to figure 4. Splicing and decay of different NMD reporters. A, E-H), U2OS cells expressing scFv-sfGFP, PCP-mCherry-CAAX were transfected with indicated reporter constructs and analyzed by time-lapse microscopy. Time between first detection of translation and mRNA cleavage was determined.) Splicing of reporters containing a single exon-intron-exon cassettes ('intron cassettes') from different genes. Schematic of splicing reporter (top). Intron cassettes were inserted between GFP and BFP. P2A ribosome skipping sites were inserted between GFP and the intron cassette, and between the intron cassette and BFP. Three types of control reporters were generated, reporters lacking an intron, reporters lacking an intron and including an in-frame stop codon immediately upstream of the BFP and reporters in which the splice sites were mutated. B, Bottom) U2OS cells were transfected with plasmids containing indicated intron cassette reporters and fluorescence intensity of BFP and GFP was measured by FACS. Spliced mRNAs produce both GFP and BFP, while unspliced mRNAs only produce GFP. BFP-to-GFP ratio was determined for each cell and the average BFP-to-GFP ratio was calculated for all cells. Average BFP-to-GFP ratio was normalized to the average BFP-to-GFP ratio of 6 reporters lacking an intron (Δ intron). Dots and error bars represent mean \pm SEM. C) Schematic of the qPCR-based splicing assay. U2OS cells were transfected with Kif18b^{PTC} reporters containing indicated intron cassette, or a matched reporter from which the intron was removed. Abundance of the mRNA was assessed with a primer set that amplifies a product independently of the splicing status ('total'), and a primer set that only amplifies the spliced mRNA ('spliced'). The ratio between spliced and total transcript abundance was determined, and normalized to the spliced/total ratio of the matched control reporters from which the intron was removed. D) Quantification of splicing efficiencies as described in C. Dots represent individual experiments, and the mean value is represented by a horizontal line. F) Cleavage curves of three NMD reporters with distinct decay rates. Black lines indicate the best fit from stochastic simulations based on the model in which NMD occurs with equal probability during each round of translation. Il solid lines and corresponding shaded regions represent mean \pm SEM. Dashed lines in (F-H) indicate that data is replotted from an earlier figure panel. Number of measurements for each experiment are listed in Table S1



Supplemental figure S6 – related to figure 5. The PTC-to-intron distance affects the NMD decay rate. A, C) Schematic of indicated NMD reporters. B, D, E) U2OS cells expressing scFv-sfGFP, PCP-mCherry-CAAX and indicated reporter plasmids were analyzed by time-lapse microscopy. B, D) The time from first detection of translation until mRNA cleavage was determined for reporters shown in (A, C). E) The cleavage rate and fraction of mRNAs that is sensitive to NMD is presented for the data shown in (D). Each dot represents an individual experiment and lines represent the mean of 2 experiments. All solid lines and corresponding shaded regions represent mean \pm SEM. Dashed lines in (B, D) indicate that data is replotted from an earlier figure panel. Number of measurements for each experiment are listed in Table S1



Supplemental figure S7 – related to figure 6. The number and position of introns affects the NMD decay rate. A, C) Schematic of indicated NMD reporters. B, D-F) U2OS cells expressing scFv-sfGFP, PCP-mCherry-CAAX and indicated reporter plasmids were analyzed by time-lapse microscopy. The time from first detection of translation until cleavage was determined. G) Genome-wide analysis of the effect of the number of upstream introns on NMD efficiency in a large cohort of cancer samples. Only PTCs with 2 or more downstream introns were included. All solid lines and corresponding shaded regions represent mean \pm SEM. In boxplots shown in (G) boxes represent the interquartile range with the central line as median, and the whiskers extend to the extreme values after removing outliers. P values are indicated as * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) by two-tailed Mann-Whitney U tests. Dashed lines in (D-F) indicate that data is replotted from an earlier figure panel. Number of measurements for each experiment are listed in Table S1



Chapter 5

Single-Molecule Imaging of Degradation of Non-Stop mRNAs by the Exosome

Tim A. Hoek*, Iris Bally*, Marvin E Tanenbaum

** These authors contributed equally to this work*

ABSTRACT

The exosome is a multiprotein complex that degrades mRNAs without a stop codon through a mechanism called non-stop decay. Exosomal degradation is stimulated by ribosomes on the 3' end of the mRNA, but the intricacies of the interplay between ribosomes and the exosome during degradation of non-stop mRNAs remain poorly understood. Here, we have developed a method to visualize both translation and exosomal degradation of single mRNA molecules in real time. We show that exosomal degradation is not only stimulated by ribosomes on the 3' end of the mRNA, but also by ribosomes that terminate just upstream of the 3' end of the mRNA. In addition, we observe that exosomal degradation is blocked when the exosome encounters stalled, but not translating, ribosomes when degrading the mRNA, suggesting that translation is required for removal of ribosomes by the exosome. Our study produces new insights in the kinetics and mechanisms of exosomal degradation and provides a new method to study exosomal degradation of single mRNA molecules.

INTRODUCTION

mRNA degradation plays an important role in regulation of gene expression, and the stability of mRNAs varies between genes and in response to environmental cues (Romero-Santacreu et al., 2009; Schwanhäusser et al., 2011). mRNAs are normally protected from exonucleolytic degradation through 5' capping and 3' polyadenylation. The poly(A) tail of an mRNA can be removed through deadenylation, which results in degradation of the mRNA by the exosome as part of regulated mRNA turnover (Anderson and Parker, 1998; Eisen et al., 2020). In addition, exosomal degradation happens as part of mRNA quality control to degrade mRNAs that lack a stop codon ("non-stop mRNAs") in a process called non-stop decay (Frischmeyer et al., 2002; Van Hoof et al., 2002). Ribosomes translating non-stop mRNAs do not dissociate from the mRNA via translation termination, but instead keep translating until they reach the 3' end of the mRNA, and these 3' end-stalled ribosomes are thought to induce exosomal degradation of the non-stop mRNA. Non-stop mRNA can also be degraded via endonucleolytic cleavage upstream of the 3' end stalled ribosome (Arribere and Fire, 2018; Guydosh and Green, 2017; Tsuboi et al., 2012). However, exosomal decay appears to be the predominant mechanism of degradation of non-stop mRNAs as endonucleolytic cleavage is only observed when exosomal degradation is impaired (Arribere and Fire, 2018; Guydosh and Green, 2017; Tsuboi et al., 2012).

A common source of non-stop mRNAs is truncated mRNAs that arise as a consequence of endonucleolytic cleavage, for example through nonsense-mediated mRNA decay, siRNA-mediated cleavage, or cleavage induced by stalled ribosomes (Arribere and Fire, 2018; Doma and Parker, 2006; Hashimoto et al., 2017; Tsuboi et al., 2012). These non-stop mRNAs lack both a stop codon and a poly(A) tail, which causes ribosomes to translate until they reach the 3' end of the mRNA. In addition, non-stop mRNAs can also be generated through mutation of the stop codon or through premature polyadenylation, which result in mRNAs that lack a stop codon but do have a poly(A) tail. Although polyadenylated non-stop mRNAs do not have a stop codon, ribosomes might not reach the 3' end of these mRNAs as ribosomes stall when translating long polyadenine stretches (Arthur et al., 2015; Chandrasekaran et al., 2019; Koutmou et al., 2015), so it is unclear whether polyadenylated non-stop mRNAs are also targeted for non-stop decay.

Induction of non-stop decay by ribosomes at the 3' end of the mRNA involves Pelota and Hbs1. These factors are structurally similar to the eukaryotic release factors eRF1 and eRF3, but bind to the ribosome when the A-site of the ribosome does not contain mRNA and therefore preferentially bind to 3' end stalled ribosomes (Becker et al., 2011; Chen et al., 2010). Binding of Pelota and Hbs1 stimulates dissociation of the 60S ribosomal subunit from the mRNA by ABCE1 (Pisareva et al., 2011), which is required for subsequent degradation of the non-stop mRNA by the exosome (Tsuboi et al.,

2012).

Although it is not completely understood how the exosome is recruited to non-stop mRNAs after ribosome removal, some insight can be gained from structural studies on the exosome and the Ski complex, which is a co-complex of the exosome. The exosome is a barrel-like structure in which mRNA is channelled to the catalytic subunit (Liu et al., 2006), while the Ski complex binds on top of the barrel and channels mRNA into the exosome (Halbach et al., 2013). The Ski complex is normally in a ‘closed’ conformation in which the mRNA cannot pass through it and opens upon binding to the 40S subunit of ribosomes (Schmidt et al., 2016), suggesting that 3’ end-stalled ribosomes stimulate exosomal decay by binding to the Ski complex and enabling mRNA to be channelled into the exosome. The Ski complex can only bind to mRNA when the mRNA extends from the ribosome (Zinoviev et al., 2020). However, mRNA does not extend from 3’ end stalled ribosomes and can therefore not be channelled into the Ski complex and the exosome when a ribosome is bound to the 3’ end of the mRNA. Ribosomes therefore also seem to have a protective effect against degradation of non-stop mRNAs, and this protection might be relieved by removal of the ribosome via Pelota and Hbs1. Exosomal decay is thus stimulated by ribosomes through opening of the Ski complex, but also inhibited by ribosomes through protection of the 3’ end of the mRNA, and it remains to be determined how the interplay between those two contradicting effects ultimately lead to recruitment of exosomes to non-stop mRNAs.

After the exosome has been recruited to an mRNA, the mRNA is degraded from the 3’ end to the 5’ end. Exosomal degradation is a processive process, meaning that a single exosome complex consecutively removes nucleotides without dissociating from the mRNA (Dziembowski et al., 2007). As the exosome moves along the mRNA in the 3’ to 5’ direction, it can encounter ribosome that are translating the mRNA. These ribosomes need to be removed from the mRNA in order to degrade the mRNA, but limited *in vivo* data is available on what happens when exosomes collide with translating ribosomes. *In vitro* experiments suggest that the Ski complex can extract mRNA from ribosomes in 3’ to 5’ direction, and that mRNA can be extracted from ribosomes in both the pre- and post-translocation state (Zinoviev et al., 2020). This suggests that the Ski complex is not only required for initial activation of mRNA degradation, but also for degradation of the mRNA after the exosome has been recruited to the mRNA. However, these experiments were done with mRNAs translated by a single ribosome, and it is unclear how this situation relates to mRNAs translated by multiple ribosomes. Extraction in 3’ to 5’ direction results in sliding of the ribosome on the mRNA, and for monosomes this can continue until ribosomes reach the 5’ end of the mRNA (Zinoviev et al., 2020). However, when mRNAs are translated by multiple ribosomes, 3’ to 5’ extraction of mRNA from the leading ribosome should lead to collision of the leading ribosome with other ribosomes which could lead to activation of other mRNA quality control pathways (Doma and Parker, 2006). It therefore remains an open question how

the exosome and Ski complex remove ribosomes from polysomes *in vivo*.

A reason why limited *in vivo* information on exosomal degradation is available is that most current methods to study exosomal degradation *in vivo*, such as northern blotting or qPCR, rely on steady state measurements to determine RNA abundance. Although these methods work well to assess transcript stability or to detect degradation intermediates, it has remained challenging to distinguish exosome recruitment from exosomal degradation, which has made it difficult to study the exact interplay between ribosomes and exosomal decay of non-stop mRNAs.

Here, we develop a method to simultaneously visualize translation and exosomal degradation of single mRNA molecules. Translation is visualized using our previously established SunTag system, while exosomal degradation of mRNA cleavage fragments is visualized as a progressive decline in GFP fluorescence intensity that happens when the mRNA is degraded by the exosome. Using this method, we can distinguish between the time from cleavage until the start of exosomal degradation, which could represent recruitment of the exosome to the mRNA, and degradation of the mRNA that happens after exosome recruitment. We show that degradation is strongly stimulated by ribosomes in close proximity of the 3' end of the mRNA, but that this stimulation does not strictly require ribosomes to reach the 3' end of the mRNA. In addition, we find that the exosome is blocked by stalled, but not translating, ribosomes, suggesting that translation is required for extraction of mRNA from ribosomes during exosomal degradation.

RESULTS

An assay to study degradation of individual non-stop mRNAs in vivo

To visualize degradation of single non-stop mRNA molecules in live cells, we have developed a method based on our previously established SunTag system (Ruijtenberg et al., 2020; Tanenbaum et al., 2014; Yan et al., 2016). In the SunTag assay, a doxycycline inducible mRNA reporter (Figure 1A) is expressed in cells that constitutively expresses a SunTag-binding single-chain intracellular antibody which is fluorescently labelled with GFP (scFv-GFP). This mRNA reporter contains 24 repeats of a sequence that encodes the SunTag peptide at the 5' end of the mRNA coding sequence, followed by a 2.5 kb sequence coding for Kinesin-like motor protein Kif18b. As the SunTag mRNA sequence is translated, the SunTag peptides that emerge from the ribosome exit tunnel are quickly bound, and fluorescently labelled, by the scFv-GFP. Since each translating ribosome is associated with a nascent peptide chain that can be bound by up to 24 scFv-GFPs, translation of the mRNA results in a bright, green fluorescent signal (Figure 1B/C), and the intensity of this signal can be used as readout for the number of ribosomes that translate the reporter mRNA (Yan et al., 2016). The 3' UTR

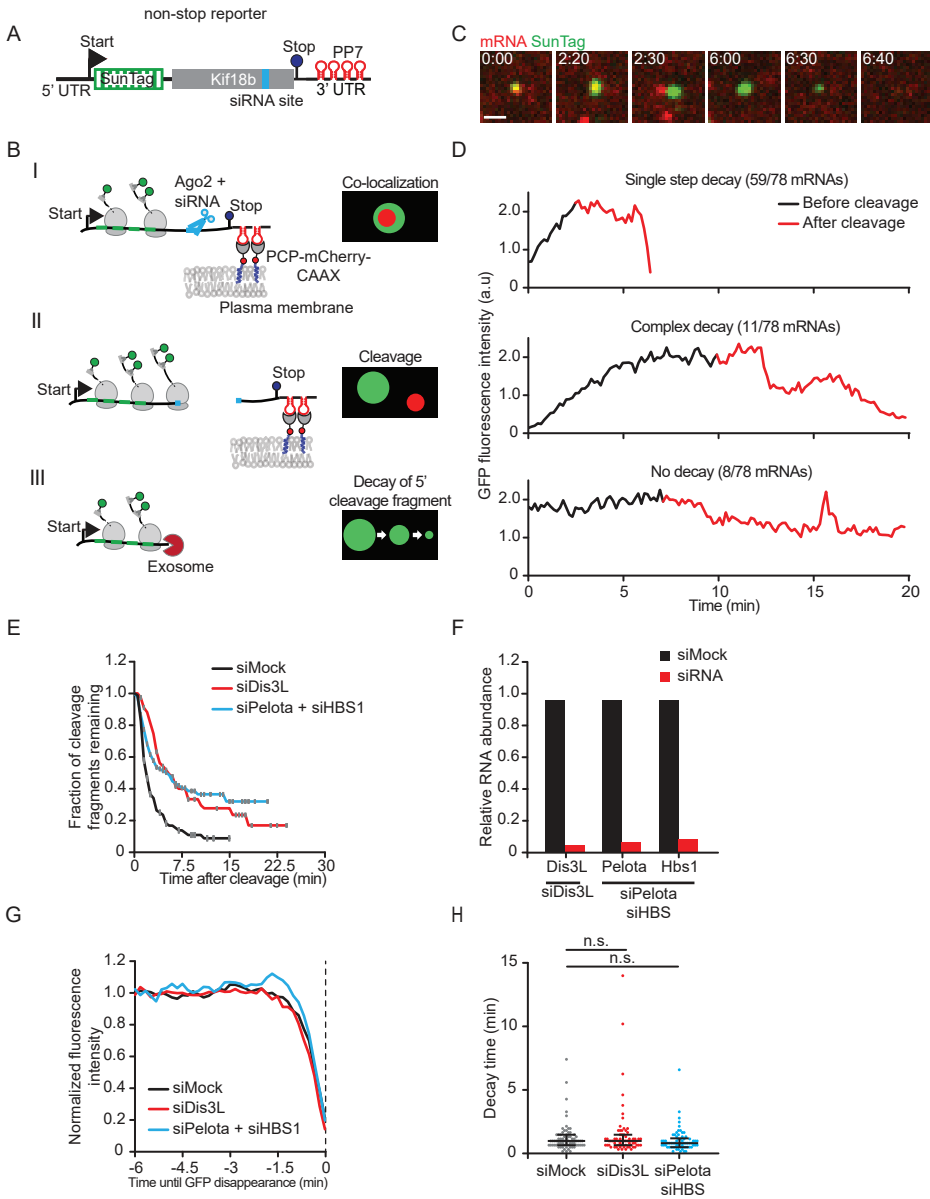


Figure 1. A new assay to visualize exosomal degradation of single mRNA molecules. A) Schematic of the non-stop reporter construct. Red loops represent 24 PCP-binding sequences, while the green box represents an array of 24 SunTag peptides. Light blue line indicates the location of cleavage by the Kif18b siRNA. Cleavage by the siRNA produces a 5' cleavage fragment that is a non-stop mRNA. B) Schematic of single-molecule imaging assay for exosomal degradation. Green and red spots in insets show nascent protein and reporter mRNA, respectively, as observed through the microscope. I) mRNA is tethered to the plasma membrane via PCP-mCherry-CAAX, resulting in a red fluorescent signal. Translation

of the mRNA reporter contains 24 PP7-hairpin structures that function as binding sites for the mCherry-tagged PP7-bacteriophage-coat protein (PCP-mCherry-CAAX) (Chao et al., 2008). Binding of the constitutively expressed PCP-mCherry-CAAX to PP7-binding sites results in a red fluorescent signal that labels the mRNA independent of its translation status (Figure 1B/C). The C-terminal CAAX motif functions as a plasma membrane anchor that tethers PCP-mCherry-CAAX, and therefore the mRNA, to the plasma membrane. Tethering facilitates imaging by confining mRNAs to a single Z-plane and limiting their diffusion. Previous experiments have shown that mRNA tethering does not detectably alter translation dynamics (Yan et al 2016). Reliable mRNA tracking is further ensured by the inducibility of the reporter using doxycycline, which allows us to control the number of mRNA molecules in a cell. Too high abundance of translating mRNAs prohibits accurate tracking, since mRNAs in close proximity of each other are hard to distinguish.

In order to make an mRNA that lacks a stop codon, we additionally transfected an siRNA in the cells that cleaves in the coding sequence of Kif18b (Figure 1A). siRNA cleavage of the reporter mRNA construct produces a 5' cleavage fragment that contains the bulk of the CDS and is associated with green fluorescent signal, and a 3' cleavage fragment that contains the PP7-binding sites and is associated with red fluorescent signal (Ruijtenberg et al., 2020). Previous studies show that cleavage of the mRNA can

Figure 1 continued... results in a green fluorescent signal associated with scFv-GFP binding to SunTag peptides, which overlaps with the red fluorescent signal. Kif18b siRNA cleaves the mRNA at the cleavage site. II) siRNA cleavage produces two cleavage fragments. The 5' fragment is a non-stop mRNA that is associated with the majority of the green signal, while the 3' fragment is associated with the red signal. Cleavage is observed as separation of the green and red spot. III) After cleavage, the 5' cleavage fragment is degraded by the exosome, which is observed as a gradual decrease in GFP fluorescence intensity. C-H) U2OS cells expressing scFv-GFP, PCP-mCherry-CAAX and the non-stop reporter were transfected with Kif18b siRNA (C-D) or transfected with Kif18b siRNA as well as indicated siRNA (E-H) and analysed by time-lapse microscopy (C-E, G, H) or by RT-qPCR (F). C) Representative images of an mRNA undergoing translation, cleavage, and a loss of GFP fluorescence caused by exosomal degradation. Scale bar, 1 μ m. D) Example traces of GFP fluorescence intensity of single mRNA molecules over time. Black lines indicate the GFP fluorescence intensity before the mRNA is cleaved, red lines indicate the GFP fluorescence intensity after the mRNA is cleaved. $t=0$ is the first frame in which the mRNA was detected. mRNAs were classified in three types of decay (single step decay, complex decay, or no decay), and the number of mRNAs in each category each is shown. E) Kaplan-Meier curve showing the time from mRNA cleavage until complete disappearance of the GFP signal. F) mRNA abundance of indicated genes and with indicated knockdowns was assessed by RT-qPCR and normalised to GAPDH mRNA abundance. G) GFP fluorescence intensity traces of mRNAs from (E) were aligned at the moment of GFP disappearance and fluorescence intensities were averaged. H) GFP fluorescence intensity traces of mRNAs from (E) were separated in a 'plateau phase' and a 'decay phase'. Duration of the 'decay phase' is shown. Each dot represents a single mRNA trace, bars indicate median and interquartile range. P values are indicated as n.s. $p>0.05$, * $p<0.05$, ** $p<0.005$, *** $p<0.0005$ by two-tailed Mann-Whitney U tests.

be detected by separation of the red and green spot (Hoek et al., 2019; Ruijtenberg et al., 2020). As siRNA cleavage happens upstream of the stop codon, the 5' cleavage fragment is a non-stop mRNA as it does not contain a stop codon anymore. As the expressed reporter mRNA creates a non-stop mRNA upon siRNA cleavage, we will refer to the reporter mRNA as the non-stop reporter. Ribosomes should be able to translate until the 3' end of this mRNA fragment, making the mRNA a substrate for non-stop decay, which should result in exosomal degradation of the mRNA and consequently disappearance of the GFP signal associated with the ribosomes translating the non-stop mRNA (Figure 1B).

To determine if we could observe degradation of 5' mRNA cleavage fragments, human U2OS cells expressing the non-stop reporter mRNA were transfected with Kif18b siRNA and visualized by spinning-disk confocal microscopy. Z-stacks were acquired for the GFP channel to be able to track 5' cleavage fragments after they move away from the plasma membrane, and images were obtained with an interval of 10 s. Most mRNAs first appeared as red spots without a green fluorescent signal, indicating that no ribosomes were associated with the mRNA when they appeared, after which the green signal gradually increased as ribosomes initiate translation. We frequently observed separation of the GFP and mCherry signal, indicating that the mRNA was cleaved by the siRNA (Figure 1C). To follow the fate of the 5' cleavage fragment, we tracked the 5' mRNA fragments containing the majority of the GFP signal and measured the GFP intensity of each 5' fragment over time (Figure 1D). For the majority of mRNAs (59/78), GFP intensity initially remained constant after cleavage, after which a rapid decline in GFP fluorescence occurred ('single step decay'). For a subset of mRNAs (11/78), the GFP spot did disappear during the time lapse imaging, but the disappearance did not occur in a single step, but instead showed multiple periods in which the GFP intensity decreased or remained constant ('complex decay'). For the third group of mRNAs (8/78 mRNAs), we could not observe disappearance of the GFP spot during the time-lapse imaging ('no decay'). We quantified the time from cleavage until complete loss of GFP fluorescence signal of the 5' mRNA cleavage fragment for all mRNAs, which showed that the median duration for GFP disappearance was 2.0 minutes (Figure 1E). As the 5' mRNA cleavage fragment does not contain a stop codon, the rapid disappearance of GFP fluorescence might be explained by degradation of the mRNA by the exosome through non-stop decay. To determine if the loss of GFP fluorescence represents non-stop decay of the 5' cleavage fragment, we used siRNA to deplete the Dis3L, the catalytic subunit of the cytoplasmic exosome, or to do a double knockdown of the non-stop decay factors Pelota and Hbs1. Knockdown efficiency was confirmed by qPCR, which showed knockdown efficiencies of 90-95% (Figure 1F). Knockdown of Dis3L resulted in a delay in the disappearance of GFP fluorescence, suggesting that the disappearance of GFP fluorescence, at least in part, represents exosomal degradation. A similar delay was observed upon knockdown of non-stop decay factors Pelota and Hbs1, suggesting

that the observed exosomal degradation is induced via non-stop decay.

Next, we wondered whether the changes in GFP fluorescence intensity could be used to gain insight into the kinetics of exosomal degradation. We aligned GFP fluorescence intensity traces at the moment of GFP disappearance (traces in which no decay occurred were excluded), and calculated the average GFP fluorescence intensity of the disappearing 5' cleavage fragments over time (Figure 1G). Without depletion of Dis3L or Pelota and Hbs1, we observed that the GFP signal disappeared in approximately one minute, suggesting that the 5' fragment is degraded within a minute once degradation has started. Surprisingly, siRNA treatment against Dis3L or Pelota and Hbs1 had no effect on the slope of the GFP signal disappearance, indicating that depletion only affects the time until the start of GFP disappearance, but not the rate of GFP disappearance itself. This result suggests that exosomal degradation is characterized by two phases: a phase in which the GFP intensity remains unchanged (from here on called 'plateau phase'), which could represent the time required for recruitment and activation of the exosome, and a phase in which the GFP signal disappears (from here on called 'decay phase'), which could represent exosomal degradation of the mRNA by the exosome. As the exosome has been proposed to degrade mRNA in a single, processive degradation step, depletion of Dis3L or Pelota and Hbs1 would prolong the plateau phase as fewer exosomes or recruitment factors are present in the cell, but would not affect the decay phase as the exosome would degrade the mRNA with equal efficiency once it has been recruited to the mRNA.

To confirm that depletion of Dis3L or Pelota and Hbs1 affects the plateau phase without affecting the decay phase, we analysed the GFP fluorescence intensity traces of single translating mRNAs and determined for each mRNA the 'decay time' (i.e. the duration of the decay phase). The start of the decay phase was determined as the first frame in which a clear decrease in GFP fluorescence could be observed. mRNAs which were not completely degraded during the time lapse imaging were excluded as they do not have a clear decay phase. This analysis revealed that the median decay time was 1.0 min (Figure 1H), similar to the decay time observed after averaging all traces (Figure 1G). Decay time was not significantly affected by depletion of Dis3L or Pelota and Hbs1 (median decay times of 1.0 and 0.83 min, $p=0.44$ and $p=0.41$, respectively), confirming that the rate of decrease in GFP fluorescence intensity is not affected by exosome abundance. It should be noted that this duration represents the time required for removal of the GFP signal, and therefore might not correspond to the time required for degradation of the complete 5' cleavage fragment, making it challenging to obtain an accurate estimate of the speed by which the exosome degrades the mRNA. However, the one minute degradation time can provide an upper limit for exosome degradation speed by assuming the entire coding sequence of the 5' cleavage fragment (which is 3904 nt) is degraded in a minute, which suggests that the upper limit for the speed by which the exosome degrades this reporter mRNA is 65 nt/s.

Exosomal degradation efficiency depends on the presence and position of the stop codon

Since exosomal decay is stimulated by binding of the ribosome to the Ski complex (Schmidt et al., 2016), we wondered whether exosomal decay strictly requires the ribosome to be stalled at the 3' end of the mRNA, or whether exosomal decay can also be stimulated by ribosomes in close proximity of the 3' end of the mRNA. In order to study the importance of 3' end-stalled ribosomes, we made three new reporters in which we introduced a stop codon either 1060, 91 or 31 nt upstream of the siRNA cleavage site (Figure 2A). After cleavage, the 5' mRNA fragments produced from these reporters are truncated mRNAs that have a stop codon 1060, 91 or 31 nt upstream of the 3' end of the mRNA, and ribosomes therefore do not reach the 3' end of the cleavage fragment but instead terminate at various distances from the 3' end (Figure 2B). Surprisingly, degradation of the 5' cleavage fragment of the 31-nt stop reporter occurred with similar kinetics as degradation of the non-stop 5' cleavage fragment (median time from cleavage until GFP disappearance of 2.0 min vs 2.0 min for the non-stop reporter, median decay time of 0.83 min vs 1.0 min for the non-stop reporter, figure 2C, 2D). In contrast, degradation of the 91- and 1060-nt stop reporters was approximately 4x slower for both reporters (median time from cleavage until GFP disappearance of 9.0 min and 8.0 min, respectively, vs 2.0 min for the non-stop reporter). The duration of the decay phase increased marginally for the 91-nt stop reporter (median decay time of 1.33 min vs 1.0 min for the non-stop reporter, $p=0.029$, Figure 2D) and remained constant for the 1060-nt stop reporter (median decay time of 1.0 min vs 1.0 min for the non-stop reporter, $p=0.29$, Figure 2D). To determine whether the observed GFP disappearance of these reporter represented exosomal decay, we depleted Dis3L in cells expressing the 31-nt and 91-nt stop reporters. Knockdown of Dis3L strongly delayed degradation of both the 31 nt-stop reporter (median time from cleavage until GFP disappearance of 8.5 min vs 2.0 min, figure 2E) and the 91-nt stop reporter (median time from cleavage until GFP disappearance of 23.5 min vs 9.0, Figure 2F), confirming that GFP disappearance represents degradation of the 5' cleavage fragment by the exosome. These data suggest that exosomal decay is not only stimulated by ribosomes at the 3' end of the mRNA, but also by ribosomes that terminate in close proximity of the 3' end of the mRNA.

Since the 31-nt stop reporter was degraded with similar kinetics as the non-stop reporter, we wondered whether the 31-nt stop reporter was degraded through the same pathway as the non-stop reporter. Exosomal degradation of non-stop mRNAs requires removal of ribosomes from the 3' end by Pelota and Hbs1. However, as ribosomes do not reach the 3' end of the 31-nt stop reporter, Pelota and Hbs1 might not be required for degradation of the 31-nt stop reporter. To address this, we depleted Pelota and Hbs1 in cells expressing the 31-nt stop reporter. Degradation of the 31 nt stop reporter was not affected by depletion of Pelota and Hbs1 (median time from cleavage until GFP disappearance of 1.5 min vs 2.0 min, Figure 2E), indicating that Pelota and

Hbs1 are required for exosomal degradation of non-stop mRNAs but not for exosomal degradation of mRNAs with a stop codon close to the 3' end of the mRNA. Together,

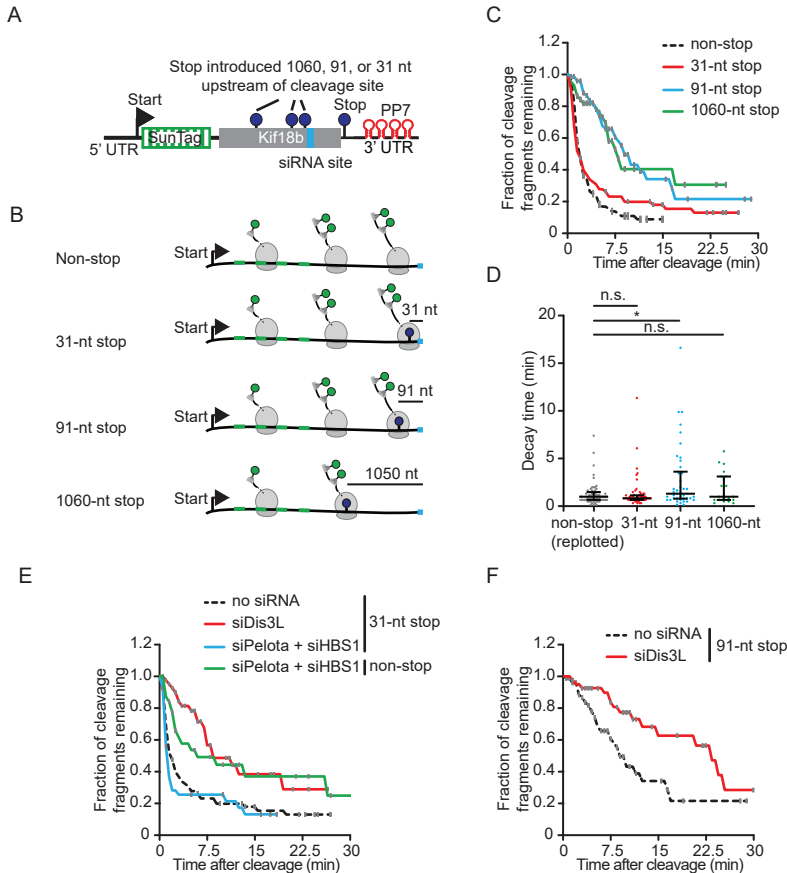


Figure 2. Stimulation of exosomal degradation by ribosomes depends on the presence and position of the stop codon. A-B) Schematic of indicated reporters (A) and 5' cleavage fragments (B) that are created upon mRNA cleavage by an siRNA targeting *Kif18b*. Stop codons were introduced 31, 91, and 1060 nt upstream of the siRNA cleavage site (A), resulting in 5' mRNA cleavage fragments with a stop codon at indicated distances from the 3' end (B). C-F) U2OS cells expressing scFv-GFP, PCP-mCherry-CAAX and indicated reporters (C-F) were transfected with a *Kif18b* siRNA (C, D) or with a *Kif18b* siRNA and indicated siRNAs (E, F) and analysed by time-lapse microscopy. C, E, F) Kaplan-Meier curves showing the time from mRNA cleavage until complete disappearance of the GFP signal. Black dashed line in (C) is replotted from figure 1E for comparison. Black dashed lines in (E) and (F) are replotted from (C) for comparison. D) GFP fluorescence intensity traces of mRNAs from (C) were separated in a 'plateau phase' and a 'decay phase'. Duration of the 'decay phase' is shown. Data from the non-stop reporter was replotted from figure 1H for comparison. Each dot represents a single mRNA trace, error bars indicate median and interquartile range. P values are indicated as n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$ by two-tailed Mann-Whitney U tests.

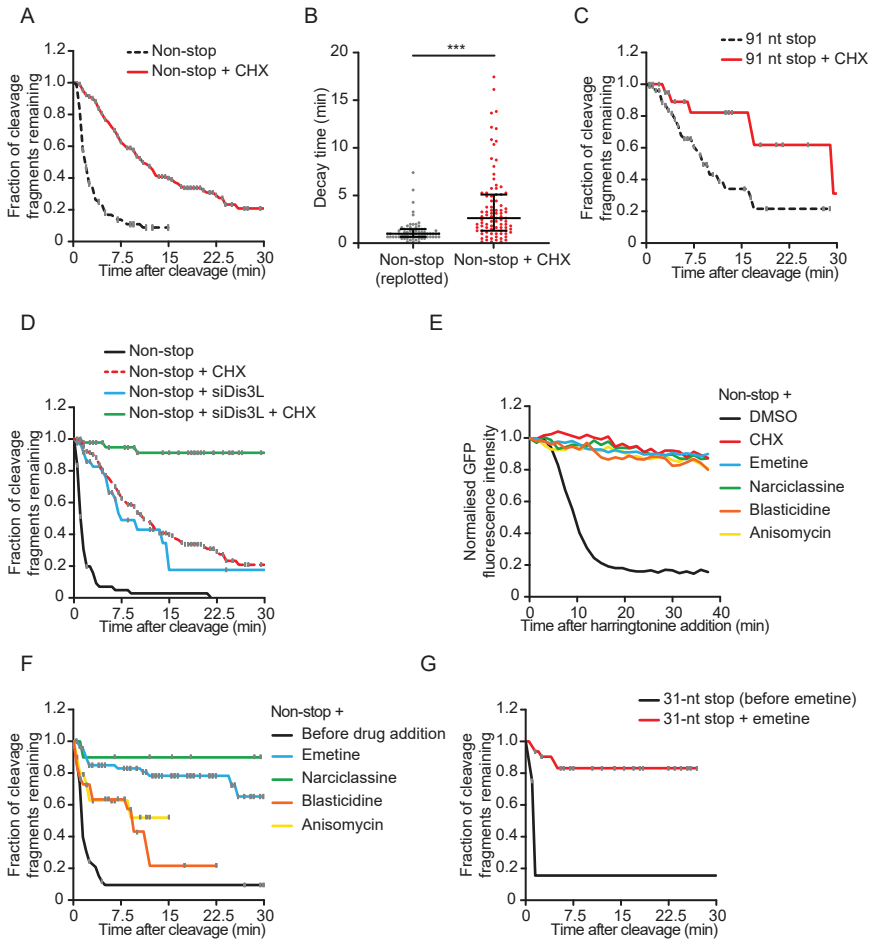


Figure 3. Blocking translation with chemical inhibitors prevents exosomal degradation. A-G) U2OS cells expressing *scFv-GFP*, *PCP-mCherry-CAAX*, and indicated reporters were transfected with *Kif18b* siRNA (A-C, F, G) or with *Kif18b* siRNA and indicated siRNA (D), treated with indicated translation elongation inhibitor, and analysed by time-lapse microscopy. A, C, D, F, G) Kaplan-Meier curves showing the time from mRNA cleavage until complete disappearance of the GFP signal. Blacked dashed lines in (A) and (C) are replotted from figure 1E and 2B, respectively, for comparison. Red dashed line in (D) is replotted from (A) for comparison. B) GFP fluorescence intensity traces of mRNAs from (A) were analysed and separated in a ‘plateau phase’ and a ‘decay phase’. Duration of the ‘decay phase’ is shown. Data from the non-stop reporter was replotted from figure 1H for comparison. Each dot represents a single mRNA trace, bars indicate median and interquartile range. P values are indicated as n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$ by two-tailed Mann-Whitney U tests. E) Quantification of the GFP fluorescence intensity of translating mRNAs after addition of harringtonine. Harringtonine was added a few minutes after the start of imaging and the GFP fluorescence intensity was normalised to the GFP fluorescence intensity at the time point of harringtonine addition.

these data show that 5' cleavage fragments both with and without a stop codon are degraded by the exosome. Exosomal degradation is stimulated by ribosomes in close proximity of the 3' end of the mRNA and does not strictly require ribosomes to stall on the 3' end of the mRNA. However, exosomal decay of non-stop mRNAs is different from decay of mRNAs with a stop codon since it depends on removal of ribosomes from the 3' end by the non-stop decay factors Pelota and Hbs1.

Exosomal degradation is blocked by stalled ribosomes

During degradation of 5' cleavage fragments, the exosome can encounter translating ribosome that must be displaced from the mRNA in order to degrade the mRNA. Studying the interplay between processive exosomal degradation and ribosomes on the mRNA has been challenging as ribosomes also play an important role in recruitment of the exosome to the mRNA (Frischmeyer et al., 2002; Van Hoof et al., 2002; Schmidt et al., 2016). However, since our assay can distinguish between the time until the start of exosomal degradation (i.e. the plateau phase) and exosomal degradation itself (i.e. the decay phase), we wanted to look in more detail into the interplay between translating ribosomes and exosomal degradation. Our data suggests that exosomal degradation of 5' cleavage fragments is not (strongly) blocked by translating ribosome as the GFP signal decreases rapidly once exosomal degradation has started. We wondered whether removal of ribosomes from the mRNA requires ribosomes to be actively translating. To this end, we stalled ribosomes on the non-stop reporter mRNA using the translation elongation inhibitor cycloheximide (CHX). Treatment of U2OS cells expressing the non-stop mRNA reporter with CHX resulted in a strong delay in degradation of the 5' cleavage fragment (median time from cleavage until GFP disappearance of 11.0 min vs 2.0 min, Figure 3A). This is consistent with previous studies that showed that degradation of non-stop mRNAs is inhibited by CHX treatment (Frischmeyer et al., 2002), and has been attributed to the requirement of ribosomes on the 3' end of the mRNA for exosome recruitment. However, analysis of the decay phase revealed that duration of the decay phase was increased by CHX treatment (median decay 2.67 min vs 1.0 min without CHX, $p < 0.0001$, Figure 3B), suggesting that CHX treatment additionally affects degradation after the exosome has been recruited to the 5' cleavage fragment. To distinguish between the effect of CHX on exosome recruitment and exosome degradation after recruitment, we treated cells expressing the 91-nt stop reporter with CHX. As our data suggests that ribosomes do not stimulate exosome recruitment on this reporter, CHX should only affect degradation of this reporter by blocking degradation after the exosome has been recruited. Addition of CHX resulted in a delay in degradation of the 91-nt stop reporter (median time from cleavage until GFP disappearance of 29.5 min vs 9.0 min, Figure 3C), suggesting that degradation by the exosome is inhibited when the exosome encounters ribosomes that are stalled by CHX. To confirm that the observed GFP disappearance in the presence

of CHX was caused by exosomal degradation, we combined CHX treatment with depletion of Dis3L. The combined treatment almost completely abolished degradation of 5' cleavage fragments, indicating that GFP disappearance was caused by exosomal decay. Interestingly, the magnitude of the effect was much stronger than expected based on the individual treatments, suggesting a synergistic effect between CHX treatment and Dis3L depletion on inhibition of exosomal degradation.

We wondered why CHX-stalled ribosomes inhibit the exosome more than translating ribosomes. One difference between CHX-stalled ribosomes and translating ribosomes is that stalled ribosomes are not going through the conformational changes associated with translation, and these conformational changes might be required for removal of the ribosome. Another possibility is that ribosomes need to be in a specific state of the translation elongation cycle to be removed (e.g. pre-translocation or post-translocation or pre-peptide bond formation), and CHX would prevent ribosome removal by blocking the ribosome in a state in which it cannot be removed. CHX binds to the E-site of the 60s ribosomal subunit and locks ribosomes in the pre-translocation state by preventing the eukaryotic elongation factor 2 (eEF2)-mediated tRNA translocation step (Schneider-Poetsch et al., 2010), which could indicate that the exosome cannot remove ribosomes stuck in the pre-translocation state. To test if removal of ribosomes by the exosome is affected by one of the conformations the ribosome adopts during translation, we tested several other translation elongation inhibitors that have different mechanisms of action for their effect on exosomal degradation. Similar to CHX, emetine locks ribosomes in the pre-translocation state, but emetine binds to the E-site on the 40s ribosomal subunit (Paulovich et al., 1993). Anisomycin, narciclasine, and blasticidin S block the peptide bond formation step and thereby stall the ribosome in the pre-peptide bond formation state. Anisomycin blocks peptide bond formation by binding in the peptidyl-transferase center (PTC) (Yusupova and Yusupov, 2017). Narciclasine also inhibits peptide bond formation, but through binding to the 60S tRNA A-site, where it hinders aminoacyl-tRNA positioning in the PTC (Fürst, 2016). Blasticidin S binds in the P-site of the large subunit, where it inhibits peptide bond formation by blocking the proper placement of the A- and P-site tRNAs at the PTC (Arenz and Wilson, 2016). We first confirmed that all translation elongation inhibitors efficiently blocked translation elongation (figure 3E, see methods). We then tested whether these inhibitors also affect decay of the non-stop reporter mRNA. All inhibitors decreased the time until GFP disappearance of the non-stop reporter (Figure 3G), and for one inhibitor (emetine), we also confirmed the inhibition of degradation in the 31-nt stop reporter mRNA (Figure 3H). These data show that the exosome can be blocked by ribosomes that are stalled in various steps of the translation cycle. However, we found that the rate of exosomal decay was highly variable between various translation elongation inhibitors (Figure 3G). Although this could be a result of a different mechanism of action between the various inhibitors, it could also be a consequence of differences in binding affinities, as no clear relationship

was observed between the binding site of the translation elongation inhibitor on the ribosome or the specific step of translation elongation that is being hindered, and the effect of the inhibitor on mRNA decay. Together, these data show that the exosome is blocked by ribosomes that are stalled by translation elongation inhibitors at various states.

DISCUSSION

A new method to study exosomal degradation of single mRNA molecules

An important strength of our new assay is that it allows us to very precisely observe the kinetics of exosomal degradation of single mRNA molecules. We demonstrate that our assay replicates defining features of non-stop decay, including the dependence on exosomal decay, a stimulation of exosomal degradation by ribosomes on the 3' end of the mRNA fragment, and a requirement for the key non-stop decay factors Hbs1 and Pelota. Using our assay, we can now differentiate between exosomal degradation and the time between cleavage and the start of exosomal degradation. This uniquely allows us to gain insight in the mechanisms behind exosome recruitment and exosome processivity, which cannot be discriminated between in bulk mRNA decay assays. In addition, single-molecule measurements can uncover different mRNA subpopulations, such as the subset of mRNA molecules that exhibit a complex decay pattern opposed to the majority of non-stop mRNA fragments that decay in a single rapid step.

Possibilities for improvement of the assay

Although the assay visualizes exosomal degradation of mRNAs, it currently uses an indirect readout for mRNA degradation by looking at disappearance of the GFP fluorescence signal that is associated with ribosomes translating the mRNA. Because of this, exosomal degradation can only be observed when mRNAs are translated, precluding insight into the mechanisms of exosomal degradation when no or few ribosomes are translating the mRNA. In addition, changes in GFP fluorescence intensities could also be caused by changes that affect ribosomes without affecting the mRNA, such as translation termination, removal of ribosomes by Pelota and Hbs1, or translation initiation of new ribosomes. Although we show for multiple reporters that the observed decrease in GFP fluorescence intensity is dependent on the exosome subunit Dis3L, the assay could be improved by directly labelling the 5' cleavage fragment, for example through insertion of MS2 hairpins that are bound by fluorescently labeled MS2-coat protein. Another advantage of labelling the mRNA this way is that it could easily be combined with tethering of the 5' cleavage fragment through addition of a CAAX motif to the MS2-coat protein. This would limit diffusion of the 5' cleavage fragment after mRNA cleavage and thereby greatly increase the tracking accuracy. In

addition, tethering would also increase the number of cells that could be visualized in each experiment, as Z-stacks are not required for observing tethered mRNAs and imaging would thus be faster. We predict that these adaptations will improve the assay to increase accuracy of observation of exosomal decay even further.

Two mechanisms of exosome recruitment

We show that non-stop mRNAs are efficiently degraded by the exosome, consistent with previous descriptions of non-stop decay (Frischmeyer et al., 2002; Van Hoof et al., 2002). Surprisingly, we observe that mRNAs with a stop codon close to the 3' end of the mRNA are degraded equally efficient by the exosome as non-stop mRNAs (Figure 2). Our data suggests that exosome recruitment to the 3' end of mRNAs is not strictly dependent on a ribosome at the 3' end of the mRNA, but instead just requires a ribosome to be in close proximity of the 3' end of the mRNA. Crystal structures of the Ski complex show how mRNA can only be channelled into the exosome when the Ski complex is bound to ribosomes (Schmidt et al., 2016), so it is plausible that ribosomes in close proximity of the 3' end of the mRNA also stimulate exosomal decay by opening of the Ski complex. The difference in degradation kinetics between the 31-nt and 91-nt stop reporter indicates that the ribosome needs to be within 31-91 nucleotides of the 3' end of the mRNA for the stimulatory effect of the ribosome to occur. Surprisingly, degradation kinetics were similar between the non-stop and 31-nt stop reporters and between the 91-nt and 1060-nt stop reporters. This suggests that stimulation of exosomal decay by ribosomes is not a gradual effect that decreases with increased distance between the ribosome and the 3' end of the mRNA, but instead is an all-or-nothing affect that only stimulates decay when the ribosome is close enough to the 3' end to the ribosome. In agreement with our data, a recent study in yeast showed that exosomal degradation is stimulated by ribosomes for 3' UTR shorter than 60 nucleotides, but requires an additional cofactor for longer 3' UTRs (Zhang et al., 2019). If ribosomes stimulate exosomal decay by simply being in proximity of the 3' end, and 3' end-stalled ribosomes prevent exosomal decay by protecting the 3' end of the mRNA, it is possible that 3' end-stalled ribosomes both induce and protect from exosomal degradation of the mRNA. However, an alternative possibility is that the 3' end stalled ribosome protects the mRNA from degradation, while degradation is stimulated by an upstream ribosome that induces degradation as soon as the 3' end-stalled ribosome is removed. Future studies could address this possibility by using reporters with various translation initiation rates to determine how ribosome abundance affects degradation kinetics.

Blocking translation prevents exosomal degradation

Consistent with previous studies, we found that blocking translation with elongation inhibitors interferes with exosomal degradation of non-stop mRNA (Figure 3)

(Frischmeyer et al., 2002). However, previous studies suggested that translation is required for the initiation of non-stop decay on non-stop mRNAs by allowing ribosomes to reach the 3' end of the mRNA. Our new assay enabled the discovery that inhibiting translation does not only influence exosomal degradation by limiting the number of end-stalled ribosomes, but also inhibits degradation by blocking the exosome when it is degrading the mRNA. Different translation elongation inhibitors that give strong translational stalling seem to have varying effects on the efficiency with which they affect exosomal degradation. The different translation elongation inhibitors stall the ribosomes in variable conformations, but no obvious correlation could be found between exosomal decay time and the mechanism of inhibition of the translation elongation inhibitors. Consistent with this, a recent study used an in-vitro reconstitution assay to show that mRNA can be extracted from ribosomes by the Ski complex in both the pre- and post-translocation state (Zinoviev et al., 2020). However, cycloheximide and emetine stall ribosomes in the pre-translocation state, yet they inhibit exosomal degradation. This suggests that extraction of mRNA from the ribosome is inhibited by ribosomes that are chemically stalled in the pre-translocation state, and thus that mRNA cannot always be extracted from the ribosome in the pre-translocation state. Further studies are required to shed light on how mRNA is extracted from ribosomes in physiological conditions.

MATERIALS & METHODS

Key resource table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	Cat# 31966021
Leibovitz's L15 medium	Gibco	Cat# 21083-027
Penicillin-Streptomycin	Gibco	Cat# 15140-122
Fetal Bovine Serum (FBS)	Sigma-Aldrich	Cat# F7524
Doxycycline	Sigma-Aldrich	Cat# D9891-1G
Opti-MEM	Sigma-Aldrich	Cat# 11058-021
FuGENE 6 Transfection Reagent	Promega	Cat# E231A
Lipofectamine RNAi-MAX	Invitrogen	Cat# 13778-075
Zeocin	Invitrogen	Cat# R25001
Bioscript Reverse Transcriptase	Bioline	Cat# Bio-27036
RNAse Inhibitor	New England Biolabs (NEB)	Cat# M0307L

Key resource table continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cycloheximide	Sigma-Aldrich	Cat# C4859
Harringtonine	Cayman Chemical	Cat# 15361-5
Emetine	Merck	Cat# 324693-250MG
Blasticidine S	bio connect	Cat# SC-204655
Narciclasine	VWR	Cat# CAYM20361-1
Anisomycin	Sigma	Cat# A9789-5MG
Critical Commercial Assays		
iQ SYBR Green SuperMix	Bio-Rad	Cat# 1708885
RNeasy mini plus kit	Qiagen	Cat# 74134
Experimental Models: Cell Lines		
Human U2OS cells	ATCC	Cat# HTB-96
Oligonucleotides		
siRNA targeting Kif18b: GGCTTTTGATGACTGTGGC	Dharmacon	N/A
siRNA-pool targeting Pelota: GAAAUAUCUCCACGUACU UUAAAUGAUUGCCGUACAA GUGUGGUACUGGAGCGCAU GCGUGGAGGCCAUCGACUU	Dharmacon	N/A
siRNA-pool targeting Hbs1: CCAGCGAUCUAUUGACAAA GGUCAUUGGUCAUGUUGAU AUGCAUGGGUCUUGGAUGA	Dharmacon	N/A
siRNA targeting Dis3L: CCAU- GUAACCGUAAGAAUA	Dharmacon	N/A
qPCR primer GAPDH forward: CACCGTCAAGGCTGAGAAC- GGG	IDT	This study
qPCR primer GAPDH reverse: GGTGAAGACGCCAGTG- GACTCC	IDT	This study
qPCR primer Pelota forward: CCGACAACAACTGCTCCT- GG	IDT	This study
qPCR primer Pelota reverse: ACAGTAGGGTCACAAAGG- GCC	IDT	This study

Key resource table continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
qPCR primer Hbs1 forward: GATGATGATCTCTACGGC- CAGTC	IDT	This study
qPCR primer Hbs1 reverse: GCTCAACGGAAGGTTTGT- CACG	IDT	This study
qPCR primer Dis3L forward: AGCTTGTCAAGCTGTGCAGC	IDT	This study
qPCR primer Dis3L reverse: GGCAGATAGCAGCATTGC- TGG	IDT	This study
Recombinant DNA		
pHR-scFv-GCN4-sfGFP-GB1- NLS-dWPRE	Addgene	Cat# 60906
pHR-PP7-2xmCherry-CAAX	Addgene	Cat# 74915
pcDNA4TO-24xGCN4_v4- kif18b-24xPP7	Addgene	Cat# 74928
pcDNA4TO-24xGCN4_v4- kif18b-31nt_stop-24xPP7	This study	This study
pcDNA4TO-24xGCN4_v4- kif18b-91nt_stop-24xPP7	This study	This study
pcDNA4TO-24xGCN4_v4- kif18b-1060nt_stop-24xPP7	This study	This study
Software and Algorithms		
ImageJ	NIH	https://imagej.nih.gov/ij/
MatLab R2012b	The Mathworks, Inc.	https://nl.mathworks.com/products/matlab.html
NIS-elements imaging software	Nikon	HC 5.11.01
Other		
96-well glass bottom imaging plates (Matriplates)	Brooks Life Science systems	Cat# MGB096-1-2-LG-L

Cell culture and siRNA & plasmid transfections

Unless indicated otherwise, cells were grown in DMEM (Gibco) supplemented with 1% Penicillin-Streptomycin (GIBCO) and 5% fetal bovine serum (Sigma-Aldrich). All transfection mixes were prepared in Opti-MEM (Sigma) using lipofectamine RNAiMAX as transfection reagent for siRNA transfections and FuGENE 6 (Promega) as transfection reagent for plasmid transfection.

For imaging experiments in which only kif18b siRNA was used, siRNA was reverse-transfected in 96-wells glass-bottom imaging plates (Matriplates, Brooks Life Science Systems) 24h prior to imaging. siRNA transfection was performed with 0.2 ul lipofectamine RNAiMAX in a total volume of 150 ul medium with a final siRNA concentration of 10 nM.

For experiments in which kif18b siRNA was combined with siRNAs targeting Pelota and Hbs1 or targeting Dis3L, U2OS cells were seeded in 24-wells plastic-bottom plates 72h prior to imaging and reverse transfected with siRNAs targeting Pelota and Hbs1 or siRNAs targeting Dis3L with 0.4 ul lipofectamine RNAiMAX in a total volume of 300 ul medium with a final siRNA concentration of 10 nM. 24h after transfection, medium was replaced by fresh DMEM. 48h after transfection, cells were seeded in 96-wells glass-bottom imaging plates and transfected with Kif18b siRNA as described above.

For experiments in which reporter constructs were transiently transfected (e.g. figure 2), U2OS cells were seeded in 96-wells glass-bottom imaging plates 72h prior to imaging. After 24h, indicated plasmids were transfected using 0.5 ul FuGENE 6 and 100-200 ng plasmid DNA per well in a total volume of 150 ul. After 48h, medium was replaced with fresh DMEM, and kif18b siRNA transfection was performed with 0.2 ul lipofectamine RNAiMAX in a total volume of 150 ul medium with a final siRNA concentration of 10 nM.

For generation of cell lines that stably express reporter constructs, U2OS cells grown in a 6-wells plate were transfected with 4 ul FuGene and 1000 ng plasmid DNA. After 24h, medium was replaced by medium containing 0.4 mg/ml Zeocin (Invitrogen) and cells were grown for 10 days to select for stable integration of the plasmid.

Microscopy

All live-cell imaging experiments were performed in U2OS cells expressing scFv-GFP, PCP-mCherry-CAAX and TetR (Boersma et al., 2019; Hoek et al., 2019; Yan et al., 2016). Cells were either seeded 24h before imaging or 72h before imaging in 96-well glass bottom plates at ~45% or ~11% confluency, respectively, as described above in 'siRNA and plasmid transfections'. Cell culture medium was replaced with pre-warmed Leibovitz's-15 medium (GIBCO) 30 minutes prior to imaging and transcription of the reporter mRNAs was induced through addition of doxycycline (1 mg/ml, Sigma-Aldrich) ~10 minutes before imaging. When translation elongation inhibitors were

included in an experiment, inhibitors were added approximately 5 minutes after the start of imaging. When translation elongation inhibitors were included together with harringtonine, translation elongation inhibitors were added at the start of imaging, and harringtonine was added approximately 5 minutes after the start of imaging.

Imaging was performed in an incubation chamber at 37°C on a Nikon TI inverted microscope with perfect focus system equipped with a Yokagawa CSU-X1 spinning disc, a 100x 1.46 NA objective and an iXon Ultra 897 EM-CCD camera (Andor) using NIS elements software (Nikon). For the GFP channel, unless stated otherwise, 9 Z-slices were acquired with a distance of 0.5 μm between each slice and an exposure time of 50 ms for each image. For the mCherry channel, a single Z-slice was acquired while focussing on the bottom of the cell with an exposure time of 500 ms. Images were acquired with an interval of 10 s between frames.

In experiments in which no kif18b siRNA was added (i.e. harringtonine runoff experiments), GFP images were acquired with a single Z-slice while focussing on the bottom of the cells with an exposure time of 500 ms for both GFP and mCherry.

Quantitative RT-PCR

U2OS cells expressing scFv-GFP, PCP-mCherry-CAAX and TetR were seeded into plastic 24-wells plate at ~11 % confluency and transfected with siRNAs targeting Pelota and Hbs1 or siRNAs targeting Dis3L as described above. After 24h, medium was replaced with fresh DMEM. After 48h, RNA was harvested using RNeasy plus mini kit (QIAGEN) according to manufacturer's guidelines. cDNA was synthesized using Bioscript reverse transcriptase kit (Bioline) and oligo-d(T) primers. qPCR reactions were prepared using SYBR-Green Supermix (Bio-Rad) with primers targeting either the depleted gene or GAPDH. Thermocycling was done on Bio-Rad Real time PCR machines (CFX Connect Real-Time PCR Detection System) and analysed using Bio-rad CFX manager 3.0 software. RNA abundance relative to GAPDH abundance was determined using the delta-delta Ct method.

Data analysis

Tracking and measuring of fluorescence intensities of translating mRNAs.

Maximum intensity projections of Z-stacks were first made in NIS elements software. Maximum intensity projections were further analysed using the previously described software package TransTrack (Boersma et al., 2019; Khuperkar et al., 2020) that runs on MatLab. Tracking was done by first automatically selecting all GFP spots and connecting them to make traces. All traces were then manually curated to ensure they accurately tracked the GFP spot in all frames. Several criteria were used to include or exclude

traces. 1) To ensure we selected intact mRNAs that tethered to the plasma membrane (which is needed for observation of mRNA cleavage), traces were only included when they overlapped with a clear mRNA spot during the start of the trace, or when the red spot appeared quickly after the green spot appeared. 2) Traces were excluded if they were located in regions of a cell that were crowded with other mRNAs, as this impairs accurate tracking. 3) Traces were excluded when they were located close to the edge of the field of view. 4) Traces were excluded if they started close to the end of the time lapse, as these traces could only be tracked for a short time and were thus not informative. Bleach correction and background subtraction were done automatically by TransTrack, and the fluorescence intensity of the spot was measured in each timepoint of the trace. In addition, for each trace we scored if and in which frame the mRNA was cleaved. Cleavage was scored when there was separation of the complete red and green fluorescent spot, as has been described previously (Hoek et al., 2019; Ruijtenberg et al., 2020). For mRNAs that were cleaved, we additionally scored the frame in which the GFP spot disappeared. When the GFP spot did not disappear, the trace was ended when 1) the end of the time-lapse was reached, 2) the GFP overlapped with other GFP spots and could therefore no longer be tracked, or 3) the GFP spot moved out of the field of view. When translation elongation inhibitors were added during imaging, we additionally scored the time point in which the inhibitor was added.

Grouping individual GFP traces into categories

GFP traces in which cleavage occurred were grouped into three categories. 1) Traces were classified as being degraded in a single step (single-step decay) when they showed a linear, continuous decrease in GFP fluorescence intensity after cleavage. The decrease could start directly after cleavage, but could also start later after the GFP intensity had remained constant for multiple frames after cleavage (plateau followed by decay). 2) Traces were classified as having complex decay if the GFP signal disappeared completely during the time lapse movie, but there was not a clear, single decay phase. This could for example happen when the GFP intensity remained constant for multiple timepoints in between two phases of decreasing GFP intensity. 3) Traces were classified as 'no decay' when the GFP signal did not disappear during the time lapse experiment. It should be noted that tracks in the last category would occasionally decrease in fluorescence intensity during the time lapse, and might therefore be partially degraded.

Time from cleavage until GFP disappearance

To determine the time from mRNA cleavage until degradation of the mRNA, we calculated for all mRNAs that were cleaved the time from cleavage until the moment of GFP disappearance or until the last frame in which the GFP spot could be tracked. We then plotted the fraction of GFP spots remaining over time using a Kaplan-Meier plot,

as this takes into account both the track length of GFP spots that disappeared and of GFP spots that did not disappear.

In experiments in which translation elongation inhibitors were added, time from cleavage until GFP disappearance was calculated in the same way, but mRNAs were only included when cleavage happened after the addition of the translation elongation inhibitor.

Aligning GFP traces at the moment of GFP disappearance

To determine the rate of GFP disappearance, we only included traces in which the GFP signal disappeared during the time lapse. We aligned the traces at the moment of GFP disappearance and calculated the average fluorescence intensity in each time point relative to the moment of GFP disappearance.

Distinguishing plateau phase from decay phase

In order to determine the duration of the decay phase of individual traces, we manually analysed each track and scored the duration of the decay phase for traces in which the GFP signal disappeared during the experiment. Duration of the decay phase was calculated as the duration between the first time point of GFP disappearance and the last time point in which the GFP signal was observed. To distinguish between GFP disappearance caused by exosomal decay and a decrease in GFP intensity as a consequence of noise, the first time point of GFP disappearance was determined manually. We assessed each GFP intensity trace and scored the first decrease in the slope of the GFP intensity curve. A change in slope was defined as a short and steep decrease in GFP fluorescence intensity or prolonged gradual decrease in GFP fluorescence intensity. Total GFP intensity needed to decrease approximately 25% in order to be classified as a decrease. However, this was variable based on the intensity and noise of the GFP spot, as changes in slope are more easy to detect in spots with high GFP fluorescence intensity and little noise. After we determined that there was a change in the slope of a curve, the first point in which this change could be detected was used as the first time point of GFP disappearance.

Determining the effect of translation elongation inhibitors on translation elongation.

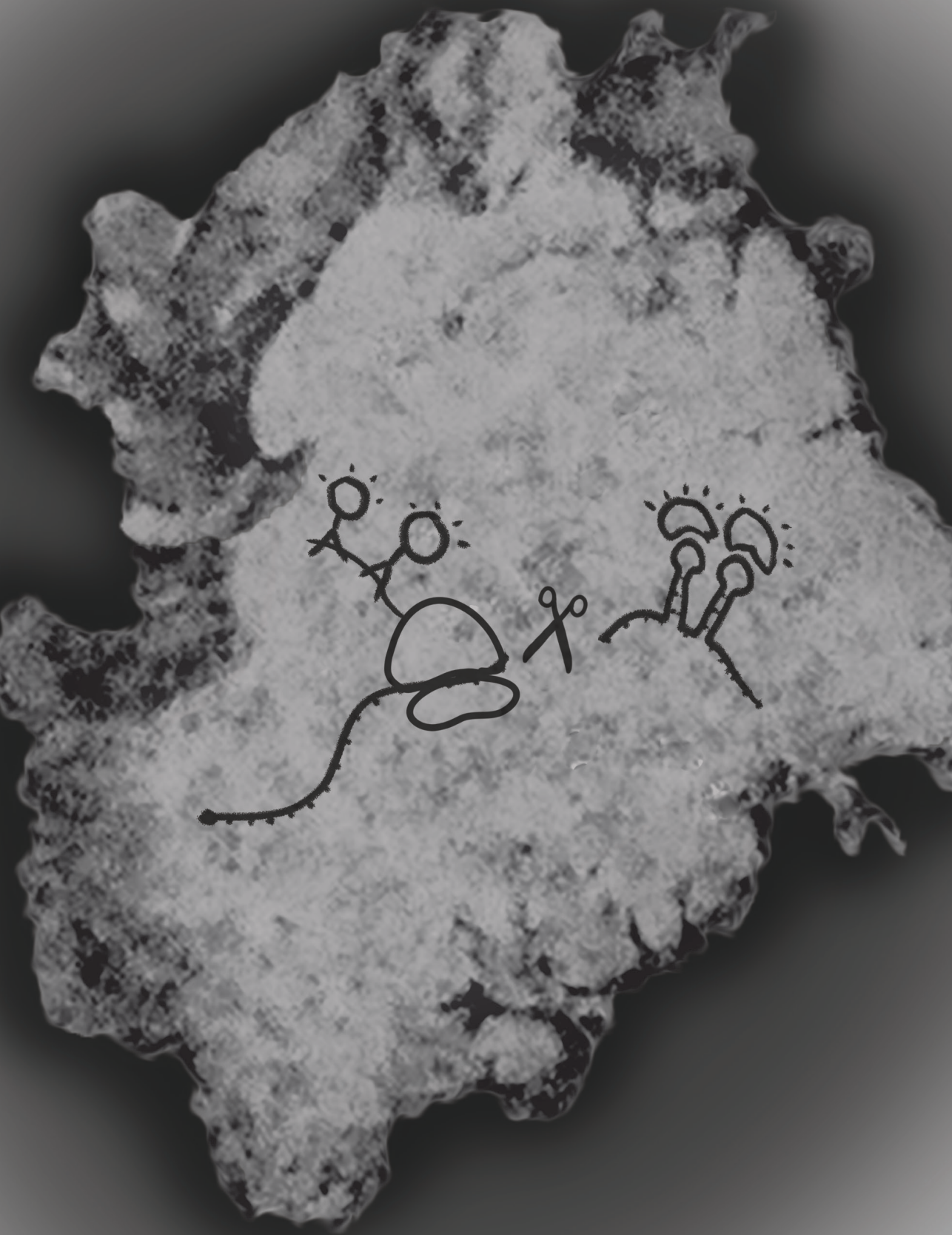
To determine if translation elongation inhibitors effectively prevent translation elongation, we performed harringtonine runoff experiments as described previously (yan 2016). Experiments were analysed in imageJ and the number and intensity of GFP spots in each cell was measured in all time frames using the spot counter plugin. The total fluorescence intensity in each frame was calculated by multiplying the mean spot intensity with the number of spots, and this intensity was normalised to the total fluorescence intensity in the frame before harringtonine addition. When

only harringtonine was added, GFP signal disappeared within ~12 minutes, consistent with previous experiments using the same reporter (Yan et al., 2016). Combination of harringtonine with translation elongation inhibitors strongly reduced translation elongation, as less than 20% of the GFP signal disappeared in 40 minutes. This indicates that translation elongation rates are strongly (>15 fold) reduced by all translation elongation inhibitors. However, it should be noted that it is difficult to determine the exact elongation rate when elongation rates are slow, as other factors that affect GFP intensity, such as noise or bleaching, dominate over changes in GFP intensity caused by the number of ribosomes/GFP molecules associated with an mRNA. Therefore, this method cannot be used to distinguish complete stalling (e.g. translation rate of 0 codons/s) from very strong stalling (e.g. translation rate of 0.01 codons/s).

REFERENCES

- Anderson, J.S.J., and Parker, R. (1998). The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SK12 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.* *17*, 1497–1506.
- Arenz, S., and Wilson, D.N. (2016). Blast from the Past: Reassessing Forgotten Translation Inhibitors, Antibiotic Selectivity, and Resistance Mechanisms to Aid Drug Development. *Mol. Cell* *61*, 3–14.
- Arribere, J.A., and Fire, A.Z. (2018). Nonsense mRNA suppression via nonstop decay. *Elife* *7*, 1–23.
- Arthur, L.L., Pavlovic-Djuranovic, S., Koutmou, K.S., Green, R., Szczesny, P., and Djuranovic, S. (2015). Translational control by lysine-encoding A-rich sequences. *Sci. Adv.* *1*, e1500154.
- Becker, T., Armache, J.-P., Jarasch, A., Anger, A.M., Villa, E., Sieber, H., Motaal, B.A., Mielke, T., Berninghausen, O., and Beckmann, R. (2011). Structure of the no-go mRNA decay complex Dom34–Hbs1 bound to a stalled 80S ribosome. *Nat. Struct. Mol. Biol.* *18*, 715–720.
- Boersma, S., Khuperkar, D., Verhagen, B.M.P., Sonneveld, S., Grimm, J.B., Lavis, L.D., and Tanenbaum, M.E. (2019). Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. *Cell* *178*, 458–472.e19.
- Chandrasekaran, V., Juszkievicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. *Nat. Struct. Mol. Biol.* *26*, 1132–1140.
- Chao, J.A., Patskovsky, Y., Almo, S.C., and Singer, R.H. (2008). Structural basis for the coevolution of a viral RNA–protein complex. *Nat. Struct. Mol. Biol.* *15*, 103–105.
- Chen, L., Muhrad, D., Hauryliuk, V., Cheng, Z., Lim, M.K., Shyp, V., Parker, R., and Song, H. (2010). Structure of the Dom34–Hbs1 complex and implications for no-go decay. *Nat. Struct. Mol. Biol.* *17*, 1233–1240.
- Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* *440*, 561–564.
- Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.* *14*, 15–22.
- Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., and Bartel, D.P. (2020). The Dynamics of Cytoplasmic mRNA Metabolism. *Mol. Cell* *77*, 786–799.e10.
- Frischmeyer, P.A., Van Hoof, A., O'Donnell, K., Guerrero, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* *295*, 2258–2261.
- Fürst, R. (2016). Narciclasine – an Amaryllidaceae Alkaloid with Potent Antitumor and Anti-Inflammatory Properties. *Planta Med.* *82*, 1389–1394.
- Guydosh, N.R., and Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. *RNA* *23*, 749–761.
- Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The yeast ski complex: Crystal structure and rna channeling to the exosome complex. *Cell* *154*, 814–826.
- Hashimoto, Y., Takahashi, M., Sakota, E., and Nakamura, Y. (2017). Nonstop-mRNA decay machinery is involved in the clearance of mRNA 5'-fragments produced by RNAi and NMD in *Drosophila melanogaster* cells. *Biochem. Biophys. Res. Commun.* *484*, 1–7.
- Hoek, T.A., Khuperkar, D., Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., Boersma, S., Vermeulen,

- M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Mol. Cell* 75.
- Van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* 295, 2262–2264.
- Khuperkar, D., Hoek, T.A., Sonneveld, S., Verhagen, B.M.P., Boersma, S., and Tanenbaum, M.E. (2020). Quantification of mRNA translation in live cells using single-molecule imaging. *Nat. Protoc.* 15, 1371–1398.
- Koutmou, K.S., Schuller, A.P., Brunelle, J.L., Radhakrishnan, A., Djuranovic, S., and Green, R. (2015). Ribosomes slide on lysine-encoding homopolymeric A stretches. *Elife* 2015, 1–18.
- Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. *Cell* 127, 1223–1237.
- Paulovich, A.G., Thompson, J.R., Larkin, J.C., Li, Z., and Woolford, J.L. (1993). Molecular genetics of cryptopleurine resistance in *Saccharomyces cerevisiae*: expression of a ribosomal protein gene family. *Genetics* 135, 719–730.
- Pisareva, V.P., Skabkin, M.A., Hellen, C.U.T., Pestova, T. V., and Pisarev, A. V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J.* 30, 1804–1817.
- Romero-Santacreu, L., Moreno, J., Perez-Ortin, J.E., and Alepuz, P. (2009). Specific and global regulation of mRNA stability during osmotic stress in *Saccharomyces cerevisiae*. *RNA* 15, 1110–1120.
- Ruijtenberg, S., Sonneveld, S., Cui, T.J., Logister, I., de Steenwinkel, D., Xiao, Y., MacRae, I.J., Joo, C., and Tanenbaum, M.E. (2020). mRNA structural dynamics shape Argonaute-target interactions. *Nat. Struct. Mol. Biol.* 27, 790–801.
- Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenuillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., et al. (2016). The cryo-EM structure of a ribosome–Ski2–Ski3–Ski8 helicase complex. *Science* 354, 1431–1433.
- Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat. Chem. Biol.* 6, 209–217.
- Schwahnhäuser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342.
- Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635–646.
- Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3' End of Aberrant mRNA. *Mol. Cell* 46, 518–529.
- Yan, X., Hoek, T.A., Vale, R.D., and Tanenbaum, M.E. (2016). Dynamics of Translation of Single mRNA Molecules in Vivo. *Cell* 165, 976–989.
- Yusupova, G., and Yusupov, M. (2017). Crystal structure of eukaryotic ribosome and its complexes with inhibitors. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160184.
- Zhang, E., Khanna, V., Dacheux, E., Namane, A., Doyen, A., Gomard, M., Turcotte, B., Jacquier, A., and Fromont-Racine, M. (2019). A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. *EMBO J.* 38, e100640.
- Zinoviev, A., Ayupov, R.K., Abaeva, I.S., Hellen, C.U.T., and Pestova, T. V. (2020). Extraction of mRNA from Stalled Ribosomes by the Ski Complex. *Mol. Cell* 77, 1340–1349.e6.



Chapter 6

Quantification of mRNA Translation in Live Cells using Single-Molecule Imaging

Deepak Khuperkar*, Tim A. Hoek*, Stijn Sonneveld, Bram M.P. Verhagen,
Sanne Boersma, Marvin E. Tanenbaum

** These authors contributed equally to this work*

Adapted from Nature Protocols (2020)

ABSTRACT

mRNA translation is a key step in gene expression. Proper regulation of translation efficiency ensures correct protein expression levels in the cell, essential for cell function. Different methods employed to study translational control in the cell rely on population-based assays that do not provide information about translational heterogeneity between cells or between mRNAs of the same gene within a cell, and generally only provide a snapshot of translation. To study translational heterogeneity and measure translation dynamics, we have developed a set of microscopy-based methods that enable visualization of translation of single mRNAs in live cells¹⁻⁴. These methods consist of a set of genetic tools, an imaging-based approach and sophisticated computational tools. Using these methods, one can investigate many new aspects of translation in single living cells, such as translation start-site selection, 3' UTR translation and translation-coupled mRNA degradation. Here, we provide detailed procedures on how to perform such experiments, including reporter design, cell line generation, image acquisition and analysis. In addition, the procedure also provides a detailed description of the image analysis pipeline and computational modelling that will enable non-experts to correctly interpret fluorescence measurements.

INTRODUCTION

Post-transcriptional regulation of gene expression tunes protein expression levels, and as such is critical for many biological processes. There are several different types of post-transcriptional gene regulation, including regulation of mRNA translation and decay. Control of gene expression through translational regulation and mRNA decay allows for rapid changes in gene expression, and these types of regulation are often used in cases where fast temporal changes in protein synthesis are required^{1,2}. Furthermore, post-transcriptional gene regulation allows for spatial control of protein synthesis, for example during early development^{3,4} and in highly polarized cells, like neurons⁵. Finally post-transcriptional gene regulation allows an additional level of RNA quality control, which ensures that only functional, full-length protein products are synthesized⁶.

mRNA translation and decay

Eukaryotic mRNA translation by the ribosome occurs through a number of well-orchestrated steps. First, the ribosome small subunit is recruited to the 5' end of the mRNA by translation initiation factors. Next, the small subunit of the ribosome scans along the 5'UTR in search of a translation initiation site, usually an AUG sequence. Once identification of the translation initiation site has occurred, the large subunit of the ribosome is recruited and translation can be initiated. The ribosome then translates the entire mRNA sequence in steps of 3 nt at a time (codons) until a termination codon is encountered. During translation termination at the stop codon, the nascent polypeptide is released and the small and large ribosomal subunits split apart and detach from the mRNA. Since each codon consists of 3 nt, each mRNA sequence contains 3 'reading frames', each of which encodes a distinct amino acid sequence. To ensure that the correct amino acid sequence is synthesized, the ribosome must initiate translation in the correct reading frame, and maintain the reading frame throughout the translation elongation phase.

The majority of mRNAs have a half-life of several hours, much shorter than the length of a cell cycle and /or lifetime of a typical cell, so mRNAs are constantly degraded in cells⁷. There are three different pathways of mRNA decay: 1) 5'-to-3' exonucleolytic decay by the exonuclease XRN1. XRN1-dependent mRNA decay is usually preceded by mRNA decapping, as the 5' mRNA cap protects mRNAs from XRN1-dependent decay. 2) 3'-to-5' exonucleolytic decay by the exosome complex, and 3) endonucleolytic cleavage of the mRNA (followed by exonucleolytic decay of the two cleavage fragments). Interestingly, mRNA degradation is frequently coupled to translation; in many cases, mRNA decay is initiated on mRNAs that are still actively translated^{8,9}. Furthermore, several quality control pathways that assess the quality of an mRNA molecule and degrade erroneous mRNAs, occur co-translationally. For example, mRNAs with a

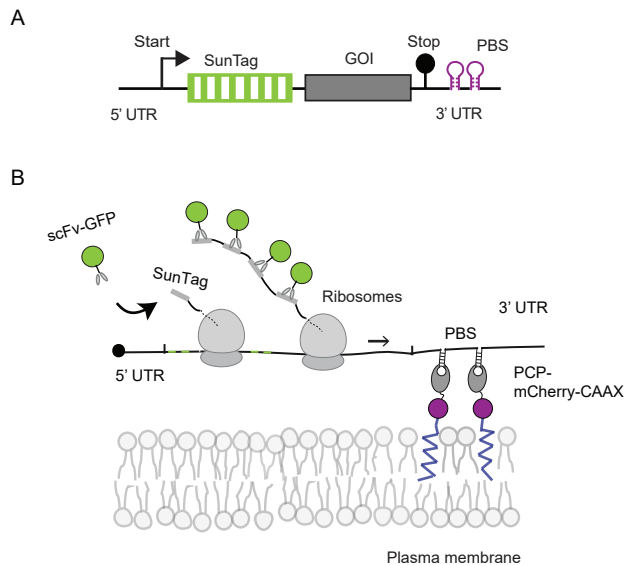


Figure 1. An assay for visualization of translation using the SunTag system. A) Schematic of a standard translation-imaging reporter with the SunTag upstream of a gene of interest (GOI). B) Schematic of the translation imaging method using the SunTag system. mRNAs are fluorescently labeled and at the same time tethered to the plasma membrane using the PP7 system through insertion of the PP7 binding sites (PBS) into the 3' UTR of the reporter mRNA. Nascent polypeptides are labeled through insertion of the SunTag sequence into the coding sequence of the reporter mRNA. As soon as the SunTag peptides emerge from the ribosome exit tunnel, they are bound by the SunTag antibody fused to GFP.

premature termination codon in their coding sequence are recognized during translation, and are degraded through a process called nonsense-mediated decay (NMD)¹⁰. Other defects in the mRNAs, such as the lack of a stop codon, can also lead to mRNA degradation during translation⁶.

The SunTag translation imaging system

We have recently established a method to visualize mRNA translation of single mRNA molecules in real-time in living cells (Fig. 1)^{11,12}. Similar methods were developed in parallel by several other groups¹³⁻¹⁶. Our method makes use of a fluorescence tagging system that we developed, called the SunTag¹⁷, which consists of an array of short linear peptides that are recognized by a single chain variable fragment (scFv) antibody fused to GFP (referred to as the scFv-GFP). To enable visualization of translation, the SunTag system is used to fluorescently label nascent polypeptides. The sequence encoding the SunTag peptide epitopes is inserted upstream of a gene of interest (GOI) and the mRNA encoding the SunTag-GOI is expressed in a cell. The synthesis of

the SunTag peptide array by ribosomes leads to recognition and binding of the co-expressed SunTag scFv-GFP to the nascent SunTag peptides, resulting in strong GFP accumulation at the site of a translating mRNA, which can be visualized as a bright green spot using fluorescence microscopy. In addition to nascent peptide labeling, individual mRNA molecules are also labeled using the PP7 system¹⁸. The PP7 system consists of the PP7 coat protein (PCP), a viral RNA binding protein that binds to a short RNA hairpin sequence (referred to as the PP7 binding site (PBS)) and is fused to a fluorescent protein (e.g. mCherry). Insertion of multiple copies (usually 24x) of the PBS into the mRNA of interest allows bright fluorescent labeling of the mRNA by co-expressed PCP-mCherry. To track mRNA molecules for extended periods of time, which is essential for understanding many aspects of post-transcriptional gene expression control, we have devised a system to tether mRNAs to plasma membrane through attachment of the PCP to the inner leaflet of the plasma membrane¹¹. Using this mRNA and translation imaging method, quantitative measurements of translation initiation and elongation can be obtained for single mRNA molecules with high spatial and temporal resolution.

Further development of the translation imaging system

Assays and genetic tools

The original SunTag translation imaging approach provided only a single read-out of mRNA translation (i.e. translation of the main open reading frame (ORF)). However, translation can be very complex; each mRNA can, in principle, be translated in three different reading frames and translation initiation can occur at different start codons, resulting in N-terminally extended or truncated proteins, or generating short peptides from upstream open reading frames (uORFs)^{19,20}. To visualize such translational complexities, we have recently developed a second, orthogonal nascent chain labeling system, called the MoonTag system²¹. In this system, translation is visualized using the MoonTag-nanobody that is fused to a fluorescent protein and binds to nascent MoonTag peptides. Combining the SunTag and MoonTag systems in a multi-color single-molecule imaging approach allows two simultaneous read-outs of translation, and allows visualization of many complex aspects of mRNA translation. For example, the MashTag (MoonTag And SunTag Hybrid Tag) contains SunTag and MoonTag peptides positioned in different reading frames, which enables monitoring of translation from two different reading frames for a single mRNA. A similar system for visualizing complex translation dynamics was developed in parallel by Lyon and colleagues²². Furthermore, we have recently established a new assay to visualize translation-coupled mRNA decay, which we have employed to study the mRNA quality control pathway nonsense-mediated mRNA decay (NMD)²³.

Computational tools

Sophisticated image analysis is key to using the translation imaging systems. Information on the number of translating ribosomes on an mRNA, the position of individual ribosomes along the mRNA, the speed at which ribosomes are moving and the rate at which ribosomes are initiating on an mRNA molecule are difficult to obtain without advanced image analysis and computational approaches. We have developed a user-friendly graphical user interface (GUI)-based image-analysis software package called ‘TransTrack’ (written in MATLAB and freely available through GitHub: github.com/TanenbaumLab/TransTrack). In addition, we have developed a computational algorithm called ‘RiboFitter’ that can use fluorescence intensity time traces of single translating mRNAs to calculate the number of ribosomes on a single mRNA molecule and to precisely determine the moment of translation initiation of each translating ribosome over time. Calibration experiments, which will be covered in this protocol, are required to determine the parameters used for running RiboFitter.

In this protocol, we will briefly summarize the general SunTag translation imaging approach, after which we will focus on three technological developments of this assay: 1) Using multi-color translation imaging to study translational heterogeneity; 2) Visualizing translation-coupled mRNA decay and 3) Quantitative image analysis using custom-developed software for mRNA tracking and quantitative analysis of translation signals.

Applications of the method and comparisons with other methods

Various methods that have been used to study mRNA decay and translational control, such as qPCR, Northern blot, luciferase assays and RNA-seq/ribosome profiling are ensemble assays where thousands of mRNAs are pooled together for analysis²⁴⁻²⁶, so these methods do not have single cell or single mRNA resolution. To overcome these limitations, additional methods have been developed previously for visualizing protein synthesis in single cells, which include methods for: 1) measuring genome-wide protein synthesis rates in single cells²⁷⁻³¹, 2) providing a snapshot of translation of single mRNAs by staining nascent polypeptides using small molecule dyes³², 3) visualizing the first round of translation³³ and 4) localizing sub-cellular sites of translation by co-localizing mRNAs and ribosomes^{34,35}. However, these previous methods do not enable real-time measurements of translation initiation and elongation dynamics of single mRNAs, and thus, provide limited insights into translation heterogeneity among different mRNA molecules. Heterogeneity among mRNA molecules can arise from differences in mRNA sequence, for example caused by alternative splicing or alternative transcription start-site usage, or by differential post-transcriptional modifications. In addition, heterogeneity among mRNA molecules can be due to distinct phases of the life cycle that different mRNAs are in. For example, all mRNA molecules are degraded

at some point in their life, but only a very small fraction of mRNAs are undergoing degradation at any given time. The translation imaging system is uniquely suited to gain insights into translation dynamics, heterogeneity and regulation at the single transcript level and is complementary to other ensemble methods to study post-transcriptional gene regulation.

The SunTag translation imaging systems can be applied to many biological questions. In recent work, we have used this system to study heterogeneity in mRNA decoding, as well as to study translation-coupled mRNA decay due to premature termination codons^{21,23}. Both types of applications will be outlined in this protocol.

Limitations

There are a number of limitations of the translation imaging method. First, the method is low throughput, limiting the number of mRNAs that can be investigated. Second, the method requires insertion of both the SunTag and PBS sequences into the mRNA of interest. These relatively large sequences could potentially interfere with endogenous regulation of the mRNA (although we have not observed any effects of these sequences in the assays which we used so far). Therefore, it is important to test whether native regulation is still observed on a reporter mRNA after inserting the SunTag and PBS sequences. Third, we have only used exogenously expressed reporter mRNAs for translation imaging so far. It is challenging to study endogenous genes using the translation imaging method, as this requires targeted integration of relatively large DNA sequences (e.g. SunTag, PBS, etc) into the native gene locus. Moreover, two rounds of targeting may be required, as the SunTag is usually inserted at the 5' end of the coding sequence, while the PBS are inserted in the 3'UTR. In addition, constitutive expression of SunTag proteins (as would occur upon endogenous gene tagging) may result in high expression levels of 'mature' SunTag protein (i.e. proteins encoded by the reporter mRNA containing the SunTag for which synthesis is completed, and which have been released from the ribosome). Since the mature SunTag proteins remain bound to the SunTag antibodies, antibodies from the cytoplasm will be sequestered by the mature proteins, limiting the amount of SunTag antibody available for nascent chain labeling (referred to as cytoplasmic antibody depletion in this protocol). Nonetheless, endogenous tagging has already been achieved for at least one gene¹⁶, providing hope that other genes can be tagged at the native locus as well.

EXPERIMENTAL DESIGN

To apply the translation imaging method described in this protocol, the following steps are required: 1) Design of mRNA reporters (Step 1 of the 'Procedure'); 2) Generation

of cell lines and expression of reporter mRNAs (Steps 2-14 of the ‘Procedure’); 3) Image acquisition (Steps 15-24 of the ‘Procedure’); 4) Image analysis (Steps 25-33 of the ‘Procedure’) and 5) Data quantification (Steps 34-46 of the ‘Procedure’) (Fig. 2). In this protocol, we will describe 5 different applications of the translation imaging system (see below). While many steps are similar for all applications, we will also highlight key differences.

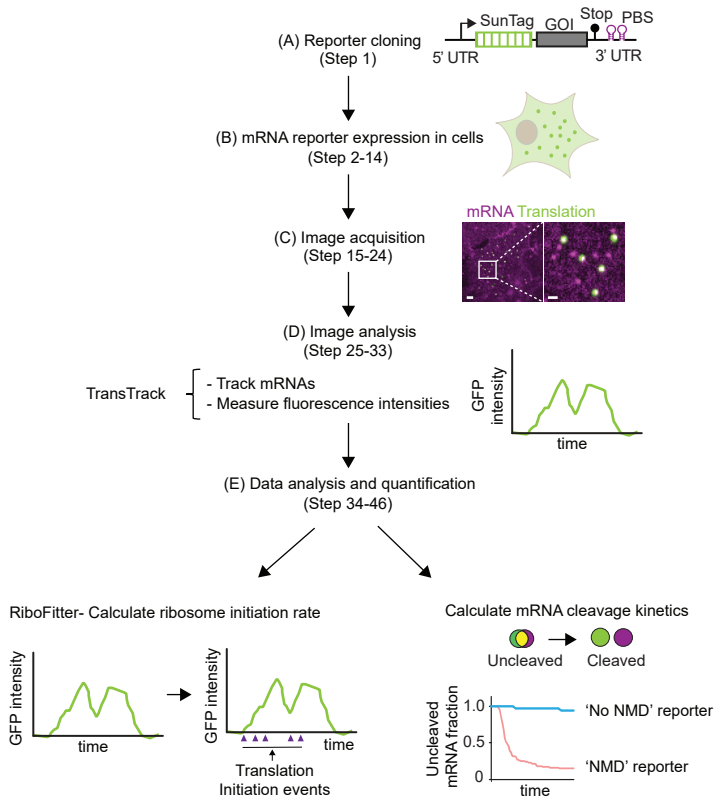


Figure 2. Overview of the experimental flow. Schematic of the workflow of the protocol. mRNA reporters are designed, cloned and expressed in mammalian cells. Next, cells are imaged and image analysis is performed; first, individual translating mRNAs are tracked. Second, fluorescence intensities of translation signals are measured over time using TransTrack to generate a fluorescence intensity time trace for each translating mRNA molecule. Further analysis on these traces is performed using a custom-built computational algorithm, RiboFitter, which determines the number of translating ribosomes and the moment of translation initiation of each ribosome for each mRNA molecule. The data obtained is then analyzed in different ways depending upon the goal of the experiment. For example, the translation initiation rate can be calculated for each reporter mRNA. In case of experiments addressing NMD, the kinetics of triggering NMD are determined by assessing frequency and timing of endonucleolytic cleavage, which is observed as a spatial separation of mRNA (mCherry) and translation (GFP) foci. Abbreviations: PBS, PP7-binding sites; UTR, untranslated region; NMD, nonsense-mediated mRNA decay; t , time.

The protocol will cover the following applications of the translation imaging system:

1. Translation imaging of a single gene of interest using the SunTag.
2. Translation imaging of two different mRNAs in the same cell.
3. Analysis of 3'UTR translation.
4. Analysis of translation start-site selection.
5. Quantifying nonsense-mediated mRNA decay.

Design of mRNA reporters (Step 1 of the 'Procedure')

In general, two key components are required for translation imaging of reporter mRNAs (Fig. 3):

1. An array of peptide epitopes (SunTag, MoonTag, or MashTag) to visualize translation, generally located at the 5' end of the coding sequence.
2. PBS in the 3'UTR (generally 24 copies) for mRNA labelling and, where applicable, tethering to the plasma membrane.

In addition, a gene of interest is often inserted downstream of the peptide array. In some cases, a specific 5' or 3' UTR is also inserted into the reporter. In cases where the 5' and 3' UTRs are not relevant to the research question, it is possible to use the default UTRs that are present in the expression vector.

We usually use a tetracycline inducible promoter to drive the expression of the reporter mRNA. Reporter mRNA expression is generally induced 15-30 min before the start of imaging. Inducible expression of the reporter mRNA is important, as constitutive expression results in a large accumulation of mature protein and cytoplasmic antibody depletion. To overcome the problems associated with an excess of mature SunTag protein, other groups have fused degrons to the protein encoded by the reporter mRNA^{14,15}. In our experience, the mature 24x SunTag peptide array bound by scFvs is highly resistant to degradation by conventional degrons, so we have chosen instead to work with an inducible promoter.

To generate cell lines stably expressing the reporter mRNA, we have introduced an antibiotic resistance gene, driven by a separate promoter, into the plasmid encoding the reporter mRNA. To generate cell lines stably expressing the reporter, transient transfection of the reporter plasmid followed by antibiotic selection is preferable over introduction of the reporter by lentivirus, because the large size and partially repetitive nature of the SunTag/MoonTag/MashTag and PP7 binding site arrays make them challenging to package into lentiviral vectors. Moreover, use of lentiviral vectors limits the choice of the 3'UTR/polyadenylation sequence. Though the basic design of the reporters (as outlined above) is the same for all the different applications of the translation imaging system, below we summarize a number of modifications to the reporter that are required for specific applications.

Translation imaging of a single gene of interest.

For imaging translation of a GOI, one can place the GOI sequence downstream of the SunTag or MoonTag peptide array followed by PBS in the 3'UTR for mRNA labeling and tracking (Fig. 3A).

Translation imaging of two different mRNAs in the same cell.

For simultaneous analysis of two different mRNAs, one can use the SunTag system for one reporter mRNA and the MoonTag system for the other reporter mRNA (Fig. 3B)²¹. We have used the PP7 system to label the mRNA of both reporters and we distinguish the two genes based solely on the translation signal. A downside of this approach is that mRNAs that do not undergo translation throughout the entire experiment cannot be assigned to either gene. To overcome this issue, one could also combine SunTag and MoonTag systems with the PP7 system for labeling of one of the two mRNAs and the MS2 labeling system to label the other mRNA^{15,36}. Combining the MS2 and PP7 system does, however, require fluorescence labeling and imaging in four different channels, and thus requires multiple dyes and/or fluorescent proteins that are suitable for imaging of weak fluorescence signals.

3' UTR translation.

Translation of the 3'UTR of mRNAs can be caused by stop codon readthrough, which is regulated during development and can cause formation of C-terminally extended proteins³⁷. In addition, translation of the 3'UTR can also be caused by re-initiation of translation after termination, which is observed, for example, when ribosome recycling is inhibited³⁸. For imaging translation of the 3'UTR, we use a translation reporter that contains the MoonTag sequence in the ORF, followed by a stop codon and the SunTag sequences in the 3'UTR (Fig. 3C). We found that the main ORF is generally translated by multiple ribosomes simultaneously, while the 3'UTR is often translated by only a single ribosome at a time, thus resulting in a dimmer fluorescence signal. Therefore, the SunTag system is used for visualizing the 3'UTR translation, as the SunTag system results in brighter fluorescence labeling than the MoonTag owing to the higher binding affinity of the SunTag scFv to the SunTag peptides²¹. The SunTag sequence can be either placed in-frame with the stop codon (to visualize stop codon read-through and translation re-initiation), or out of frame with the stop codon to visualize translation re-initiation only.

Translation start-site selection.

For studying translation start-site selection on an mRNA of interest, we developed a reading frame sensor called the MashTag (MoonTag And SunTag Hybrid Tag)²¹. The MashTag consists of both SunTag and MoonTag peptides that are fused in an

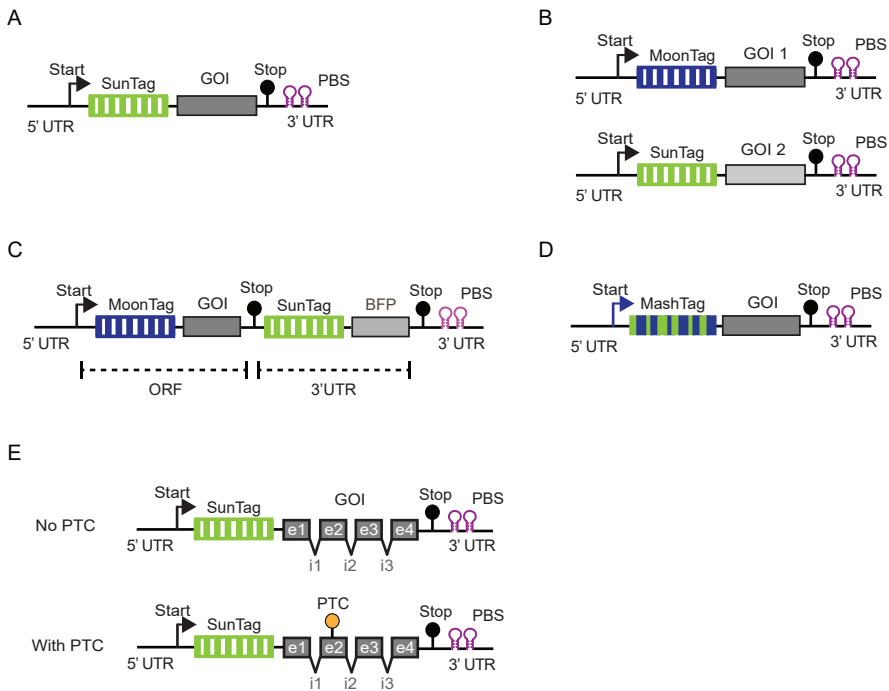


Figure 3. Reporter design. A-E) Schematics of different reporters used for translation imaging experiments. For simplicity, only 2 of 24 PBS are shown. All SunTag and MoonTag arrays contain 24 copies of the peptide array, while for the MashTag 36 copies of both the SunTag and MoonTag peptides are used to increase the fluorescence brightness. Translation start and stop codons are indicated. A) Schematic of a standard translation-imaging reporter with the SunTag upstream of a gene of interest (GOI). B) Translation of two different GOIs can be visualized simultaneously in a single cell, one GOI will be fused to the MoonTag, the other to the SunTag. Both contain PBS arrays for mRNA labeling. C) Schematic of 3'UTR translation reporter containing the MoonTag upstream of a GOI in the coding region, followed by a stop codon and SunTag sequence. BFP is added downstream of the SunTag sequence as a stuffer sequence to lengthen the coding sequence downstream of the SunTag, which increases the duration required for translation of the SunTag ORF and facilitates detection of single ribosomes translating the 3'UTR. D) Schematic of a MashTag reporter in which the MoonTag sequence is in-frame with the AUG start codon and the SunTag sequence is positioned in the -1 frame relative to the MoonTag. Stop codons are placed in both reading frames. (E) Schematic of NMD reporters that can be used to visualize endonucleolytic cleavage kinetics triggered by NMD. Two matched reporters are used, one contains a premature termination codon (PTC) within the coding region, which triggers NMD, while the other does not. Abbreviations: e, exon; i, intron.

alternating fashion, such that all the sequences encoding the MoonTag peptide epitope are positioned in the same reading frame, while all the sequences encoding the SunTag peptide are encoded in the -1 frame relative to the MoonTag peptides (Fig. 3D). The MashTag is designed such that: 1) no AUG start codons or stop codons (TGA, TAA, or

TAG) are present in any reading frame within the MashTag sequence; 2) The nucleotide sequences of individual repeats of the SunTag and MoonTag peptides are minimally repetitive (different codons were used for same amino acids in different repeats of the peptide). To study translation start-site selection for a 5'UTR sequence of interest using the MashTag, one can insert the 5' UTR sequence upstream of the MashTag in a translation reporter and assess the frequency of translation at start sites present in the two reading frames.

Analysis of NMD.

For visualizing NMD, one needs to introduce mRNA sequences in the reporter that stimulate NMD. Such sequences generally consist of a premature termination codon (PTC) and, optionally, one or more introns, at least one of which is positioned more than 55 nucleotides downstream of the stop codon. As a control, an mRNA reporter can be designed with an identical mRNA sequence but lacking a PTC or introns (Fig. 3E).

Generating cell lines and expression of reporter mRNAs (Steps 2-14 of the 'Procedure')

We have performed the majority of our experiments in U2OS cells (human osteosarcoma cells), but similar SunTag translation imaging experiments have been successfully performed in other cell types, such as HeLa cells^{14,16} and neurons^{14,15}. The main criteria to consider when choosing a suitable cell line are: 1) Is it possible to introduce the required components (e.g. SunTag antibody, reporter mRNAs, etc.) into the cells? 2) Can the cell type be imaged using high-resolution microscopy? Flat, adherent cells are generally more suitable for imaging than thick cells (e.g. oocytes) or cells that grow in suspension. However, using thick cells may be possible when applying an appropriate microscopy method (e.g. light sheet microscopy).

Preferably, cell lines are generated that stably express the SunTag and/or MoonTag antibodies, PCP-mCherry-CAAX (or PCP-mCherry, if the mRNAs should not be tethered to the plasma membrane) and the TetR transcription repressor for inducible expression³⁹ (Plasmids encoding these proteins can all be obtained from Addgene; (Addgene plasmids nos. 60907, 74925, 74926, 128603 and 128602; See also Table 1). Cell lines stably expressing these proteins can be generated through lentiviral transduction (Steps 2-13). After lentiviral transduction, cells with appropriate expression levels of fluorescently-labeled proteins are selected by flow cytometry. We generally generate monoclonal cell lines with a range of expression levels of the fluorescently-labeled SunTag/MoonTag antibodies and PCP-mCherry (-CAAX) and functionally test the cells lines by introducing a reporter mRNA and performing a translation imaging experiment. The clone with the best signal-to-noise ratio for mRNAs and translation sites fluorescence is selected for further experiments. Since optimal signal-to-noise ratio

Table 1 - plasmids from Addgene

NAME OF PLASMID	ADDGENE NUMBER	APPLICATION
Plasmids Used For Cell Line Generation		
SunTag-scFv	60907	For nascent chain labelling using SunTag
PCP-mCherry-CAAX	74925	mRNA labelling for tethered system
PCP-mCherry	74926	mRNA labelling for untethered system
MoonTag-Nb-Halo	128603	For nascent chain labelling using Halo-MoonTag-nanobody
MoonTag-Nb-GFP	128602	For nascent chain labelling using GFP-MoonTag-nanobody
Reporter Plasmids		
Standard SunTag translation imaging reporter	74928	General SunTag translation assay
MoonTag translation reporter	128604	Translation of different genes in same cell
3'UTR translation reporter	128605	3'UTR translation assay
MashTag translation reporter	128607	Translation start-site selection assay
NMD reporter (TPI-PTC160)	130698	mRNA cleavage assay
NMD control reporter (TPI-WT)	130697	mRNA cleavage assay

during imaging is very sensitive to the expression levels of both the antibodies and PCP-mCherry protein, many clones should be screened. For the SunTag antibodies, optimal expression levels are as high as possible while still allowing detection of translating mRNAs and mature SunTag proteins over the background when imaging using mature protein settings (Step 22B). For PCP-mCherry, we find that very low expression often results in the best signal-to-noise ratio. Note that PCP-mCherry can accumulate in lysosomes when generating stable cell lines, and a clone should be selected that has minimal mCherry signal in lysosomes.

Translation reporters can be either transiently transfected or stably integrated into the genome of the cells. Stable integration can be performed either by transient transfection and antibiotic selection, which results in random integration in the genome, or by targeting a specific locus in the genome (e.g. the AAVS1 locus)²¹. Transient transfection is faster, but generation of stable cell lines is easier if the reporter is imaged many times,

especially in combination with siRNAs, as we find that transfection of cells with both plasmids and siRNAs is sometimes less efficient. In our experience, mRNAs expressed from transiently transfected plasmids and stably integrated plasmids show similar translation kinetics. Also, generating monoclonal cell lines after stable integration of the reporter results in more homogeneous expression of the reporter mRNA among cells, which can be useful if the cell line will be used very frequently, or when applying other treatments that target only a subset of cells (e.g. expression of a transgene).

Imaging acquisition (Steps 15-24 of the 'Procedure')

We perform imaging experiments on a temperature-controlled spinning disc confocal microscope with a 100X NA 1.49 oil-immersion objective and an EMCCD or back-illuminated sCMOS camera to obtain high-resolution images with very sensitive detection. We prefer spinning disc confocal imaging over total internal reflection fluorescence (TIRF), *highly inclined* and laminated optical sheet (HILO) or wide field illumination due to the homogeneous illumination of spinning disc confocal imaging over the entire field-of-view. Homogeneous illumination results in homogeneous fluorescence intensity values throughout the entire image, thus facilitating image quantification. Spinning disc confocal imaging is preferred over point-scanning confocal microscopy due to the higher speed and the improved detection sensitivity when combined with an EMCCD or back-illuminated sCMOS camera. For light-sensitive applications, such as imaging single ribosomes or imaging at high time-resolution in 3D, a microscope optimized for detection sensitivity is essential. However, for imaging a highly translated mRNA labeled with the SunTag system, an epifluorescence wide field microscope or point-scanning confocal may also be sufficient, especially when mRNAs are tethered to the plasma membrane.

Image analysis (Steps 25-33 of the 'Procedure')

After image acquisition, careful image analysis is required to correctly interpret the fluorescence images. For most experiments, analysis consists of tracking individual mRNA molecules and measuring fluorescence intensities over time. In this protocol, we will describe a user-friendly software package, TransTrack, designed for mRNA tracking and intensity measurements.

Data analysis and quantification (Steps 34-46 of the 'Procedure')

After measuring fluorescence intensities over time for each mRNA molecule, additional analyses are usually required to determine translation initiation and elongation rates, as well as more sophisticated analyses, including determining the initiation rates at

multiple translation start sites or calculating the fraction of ribosomes translating the 3'UTR of an mRNA. Interpretation of fluorescence intensities is facilitated by a computational algorithm called RiboFitter, which uses fluorescence intensities of translation signals to determine the number of ribosomes translating an mRNA, as well as the moment at which each ribosome initiated translation. The first time that RiboFitter is used, several control experiments are required to determine the precise fluorescence intensity produced by a single ribosome translating the reporter mRNA, because the total translation signal associated with an mRNA translated by multiple ribosomes is a compound of multiple such single ribosome translation profiles. These control experiments will be described in the RiboFitter section under experimental procedures (Step 41 of 'Procedure'). After running RiboFitter, additional downstream analyses are often required as well, which are also described in this section. For analysis of mRNA endonucleolytic cleavage through NMD, analysis consists mainly of scoring the moment of mRNA cleavage for each mRNA molecule, although additional analyses can be performed as well.

MATERIALS

Biological materials

CAUTION The cell lines used should be checked for authenticity and regularly tested for mycoplasma contamination.

- Cell line used for imaging: Human U2OS cells obtained from ATCC (Cat# HTB-96, RRID:CVCL_0042)
- Cell line used for lentivirus generation: HEK293T cells obtained from ATCC (Cat# CRL-3216, RRID:CVCL_0063)

Reagents

- Lentiviral production plasmids from Addgene (Addgene plasmid nos. 60907, 74925, 74926, 128603 and 128602) *CAUTION* Lentiviruses can be a biological hazard, so all lentiviral work should be performed in Biosafety level-2 labs with appropriate precautions and measures for usage and disposal of lentiviral samples.
- mRNA reporter plasmids from Addgene (Addgene plasmid nos. 74928, 128604, 128605, 128607, 130698 and 130697)
- Lentiviral packaging vectors psPAX2 (Addgene # 12260) and pMD2.G (Addgene #12259)
- DMEM (High glucose with GlutaMAX supplement; Thermo Fischer Scientific, cat. no. 31966021)

- Leibovitz's L15 medium (No phenol red; Thermo Fischer Scientific, cat. no. 21083-027)
- Opti-MEM (Reduced serum medium; Sigma-Aldrich, cat. no. 11058-021)
- Fetal Bovine Serum (FBS) (Sigma-Aldrich, cat. no. F7524)
- FuGENE 6 Transfection Reagent (Promega, cat. no. E231A)
- DMSO (Dimethyl sulfoxide) (Sigma-Aldrich, cat. no. D8418-1L)
- PEI (Polyethylenimine) for plasmid transfection (Polysciences Inc, cat. no. 23966)
- Penicillin-Streptomycin (10,000 U/ml; Thermo Fischer Scientific, cat. no. 15140-122)
- Polybrene (Santa Cruz Biotechnology, cat. no. sc-134220) *CAUTION* Polybrene is harmful. Wear gloves and avoid contact with eyes and skin.
- Doxycycline (Sigma-Aldrich, cat. no. D9891-1G)
- Puromycin (Thermo Fisher Scientific, cat. no. 12122530) *CAUTION* Puromycin is harmful. Wear gloves and avoid contact with eyes and skin.
- Cycloheximide (Sigma-Aldrich, cat. no. C4859) *CAUTION* Cycloheximide is harmful. Wear gloves and avoid contact with eyes and skin.
- Harringtonine (Cayman Chemical, cat. no. 15361)
- JF⁶⁴⁶ HaloTag dye (Promega, cat. no. GA1120)
- Zeocin (Invitrogen, cat. no. R25001) *CAUTION* Zeocin is harmful. Wear gloves and avoid contact with eyes and skin.

Equipment

Cell culture

- Cell culture hood (Telstar BioVanguard, Green line)
- Nitrile gloves (Kimtech)
- Cell culture incubator (5% CO₂, humidified, 37°C; Panasonic, MCO-170AICUVH-PE)
- Serological pipettes 5 ml, 10 ml and 25 ml (Fisher Scientific, cat nos. GPN5E1, GPN10E1 and GPN25E1)
- Tissue culture plate- 6 well (Greiner bio-one CELLSTAR, cat. no. 657160)
- Tissue culture plate- 100mm (Greiner bio-one CELLSTAR, cat. no. 664160)
- Falcon 15-ml conical centrifuge tubes (Greiner bio-one CELLSTAR, cat. no. 188271)

Imaging

- 96-well glass bottom imaging plates (Matriplates) (Brooks Life Science Systems, cat. no. MGB096-1-2-LG-L)
- Confocal spinning disk microscope - Nikon TI inverted microscope with perfect

focus system equipped with a Yokogawa CSU-X1 spinning disc, a PlanApo oil immersion objective (100x 1.49 NA; Nikon) and an iXon Ultra 897 EM-CCD camera (Andor) or Prime95B camera (Photometrics). The microscope is equipped with an incubation chamber that ensures control of temperature and humidity.

- Micromanager for microscope control (Micro-Manager 1.4.22 <https://micromanager.org>)⁴⁰
- NIS-Elements Imaging Software (Nikon HC 5.11.01)

Flow Cytometry

- BD FACS Aria II

Image and data analysis software

- ImageJ (<https://imagej.nih.gov/ij/>)
- Graphpad Prism 7 (GraphPad Software Inc, <https://www.graphpad.com/scientificsoftware/prism/>)
- MATLAB R2012b The Mathworks, Inc. <https://nl.mathworks.com/products/matlab.html>)
- R 3.5.1 (R Project for Statistical Computing, <http://www.r-project.org>)
- TransTrack (MATLAB) (<https://github.com/TanenbaumLab>)
- RiboFitter (R) (<https://github.com/TanenbaumLab>)

Reagent setup

Complete DMEM

- Complete DMEM is made by adding 5% fetal bovine serum (Sigma-Aldrich, cat. no. F7524) and 1% Penicillin/Streptomycin (Thermo Fischer Scientific, cat. no. 15140-122) to DMEM (Thermo Fischer Scientific, cat. no. 31966021). The complete medium should be stored at 4°C and can be used for about 3-4 months.

PROCEDURE

Design of reporter mRNA(s) (Timing – 2-5 d)

1. Generate a reporter according to the instructions provided in ‘Experimental design’. For standard translation imaging experiments (i.e. experiments in which only translation of the main ORF is visualized), several plasmids are available from the Tanenbaum lab (Addgene #74928, #128604, #128605 and #128607, see also Table 1). Alternatively, use molecular cloning to generate appropriate reporters

suitable for your experiments. For example, for translation imaging of a gene of interest (GOI) using SunTag and PP7 systems, subclone the GOI in the standard SunTag translation reporter available on Addgene (#74928), using AgeI and EcoRV sites, which removes the GOI (KIF18B) from this plasmid.

Generation of cell lines suitable for translation imaging (Timing – 10-25 d)

CRITICAL STEP To create a cell line that can be used for translation imaging, you need to deliver the plasmids encoding fluorescently labeled proteins (e.g. SunTag-scFv, see Experimental Design) to cells by lentiviral transduction. See the ‘Experimental design’ section ‘Generating cell lines and expression of reporter mRNAs’ for a discussion on considerations for selecting a suitable cell type.

2. *Production of lentiviruses (steps 2-8)*. Seed HEK293T cells in a 6-well plate at 20-25% confluency ($\sim 3 \times 10^5$ cells per well).
3. Transfect the cells 16-24 h after plating with the lentiviral plasmid along with lentiviral packaging vectors psPAX2 (Addgene #12260) and pMD2.G (Addgene #12259) using PEI (Polyethylenimine). Firstly, mix 10 μ l PEI and 50 μ l OptiMEM per reaction.
4. Then, mix 1 μ g of lentiviral plasmid (e.g. Addgene #60907 for SunTag-scFv) with 0.4 μ g pMD2.G and 0.6 μ g psPAX2 in an Eppendorf tube. Add 60 μ l PEI + OptiMEM mix to each tube with DNA, mix by pipetting and incubate for 5-15 min at room temperature. Then, add transfection mix to cells and gently swirl the plate to mix.
5. 24 h after transfection, replace the medium with fresh medium. Check cells using a fluorescence microscope to determine transfection efficiency. 25-75% of cells should be fluorescent.
CRITICAL STEP Be gentle while changing the medium, pipet in the corner of the well, not directly onto the cells, as HEK293T cells easily detach from the plate. 20 mM HEPES can be added to the medium when cells are very dense to prevent changes in the pH of the medium, which can affect virus titers.
6. Harvest the virus-containing medium 3 days after transfection. Check cells again using a fluorescence microscope to determine virus production efficiency. All cells should be fluorescent as the newly-produced virus will have infected all non-transfected cells as well. Collect the medium into 1.5 or 2 ml tubes.
7. Spin down tubes at 3500 x g at 25^o C for 3 min. Gently remove ~ 1.5 ml of viral supernatant. Leave at least 100-200 μ l in the tube so as not to take along any cells or cell debris. The viral supernatant can also be filtered to ensure complete removal of the cell debris.
8. Use virus-containing medium immediately for infection, store at 4^o C for ~ 1 week or store virus in aliquots at -80^o C for several months.

CRITICAL STEP Aliquotting of the virus is important, as each freeze-thaw cycle can reduce the titer of the virus batch.

9. *Lentiviral transduction in U2OS cells (steps 9-13)*. Seed U2OS cells in a 6-well plate at 20-25% confluency ($\sim 2 \times 10^5$ cells per well).
10. 16-24 h after plating, add 100-500 μ l of virus-containing medium to the cells (note that the amount of virus required for efficient transduction may vary for different cell types). Make sure that the volume of cell culture medium does not exceed 1 ml. Add Polybrene to each well (10 μ g/ml).
11. Spin the 6-well plate containing the U2OS cells and virus for 90-120 minutes at 800 x g at 25^o C for increased transduction efficiency (spin-infection).
12. After spin-infection, replace the cell culture medium with fresh medium and grow the cells for at least 2 days before any further analysis or processing to obtain steady-state expression levels of the virally-transduced proteins.
13. FACS sort cells based on the expression of fluorescent proteins. Compare infected cells with uninfected control cells (that do not express any fluorescent proteins) and set gates that exclude all control cells. We recommend sorting single cells to make monoclonal cell lines, although polyclonal cell lines can also be used (see the 'Experimental design' section 'Generating cell lines and expression of reporter mRNAs').

CRITICAL STEP Ensure that expression of the fluorescently-labeled proteins is homogeneous among cells. If expression levels are variable, a correction may need to be applied during analysis of fluorescence images to correct for differences in the degree of antibody binding to the peptide epitopes.

Expression of the reporter mRNA (Timing 2-20 d)

14. For expression of reporter mRNAs using transient transfection, follow option A. For stable integration into the host cell genome, follow option B. See the 'Experimental design' section 'Generating cell lines and expression of reporter mRNAs' for a discussion of the advantages and disadvantages of these approaches.
 - A. **Transient transfection** (Timing 2 d)
 - i. Seed cells in 96-well glass bottom dish at 20-25% confluency ($\sim 2 \times 10^4$ cells per well).
 - ii. 16-24 h after plating, transfect cells using an appropriate transfection reagent according to manufacturer's guidelines. We transfect U2OS cells using FuGENE 6 according to the manufacturer's guidelines.
 - iii. Image cells 24-48 h after transfection.
 - B. **Stable integration** (Timing 12-20 d)
 - i. Seed U2OS cells in 6-well plate at 20-25% confluency ($\sim 2 \times 10^5$ cells per well).

- ii. 16-24 h after plating, transfect cells using an appropriate transfection reagent according to manufacturer's guidelines. We transfect U2OS cells using FuGENE 6 according to the manufacturer's guidelines.
- iii. 24 h after transfection, replace the medium with fresh medium and grow the cells in medium containing antibiotics (for pcDNA4/TO plasmids, we use 0.4 mg/ml Zeocin (Invitrogen)) for 10-15 days for selection.
- iv. After integration of the reporter into the genome of the cells, imaging can be performed by seeding the cells at 40-50% confluency ($\sim 4 \times 10^4$ cells per well) 16-24 h before the start of imaging.

Sample preparation for imaging (Timing 1-2 d)

15. Prepare cells as described in step 14.
16. Set the microscope temperature to 37 °C and wait until the correct temperature is reached.
CRITICAL STEP Changes in temperature can influence translation dynamics. Therefore, it is very important to ensure that temperature control is accurate during imaging.
17. Replace the cell culture medium with Leibovitz's-15 medium pre-warmed to 37°C.
CRITICAL STEP A CO₂ perfusion chamber is not required when using Leibovitz's-15 medium during imaging, but should be used when imaging in other medium (e.g. DMEM). Use cell culture medium lacking phenol red, as phenol red in the medium can reduce image quality.
CRITICAL STEP For imaging experiments in which one of the antibodies is fused to a HaloTag, label the HaloTag by incubation with 50 nM Halo^{IF646 41} in pre-warmed Leibovitz's-15 medium (37°C) for 1 h. After Halo^{IF646} ligand incubation, wash the cells three times with Leibovitz's-15 medium.
18. Add doxycycline (final concentration 1 µg/ml) (made in Leibovitz's-15 medium) to the cells to induce transcription of the reporters (See Table 2).
19. Start imaging 15-30 minutes after doxycycline addition.
CRITICAL STEP We recommend keeping a lid on the imaging plate when imaging for a long duration (> 1 h) to prevent evaporation of the medium.

Image acquisition (Timing 1-2h)

CRITICAL Image acquisition settings will vary depending on the requirements of the experiment. Here, we provide settings that we use frequently for our experiments.

20. Select x, y positions to be imaged. Representative images illustrating selection of appropriate x, y positions are shown in Fig. 4
CRITICAL STEP Selection of the best x, y positions is critical, and important

Table 2: Drugs used in the protocol

DRUG	FINAL CONCENTRATION	MODE OF ACTION	USE	STOCK MADE IN
Doxycycline	1 µg/mL	Activates Tet-on system for gene expression	To Induce transcription of reporters	Water
Puromycin	100 µg/mL	Translation inhibitor - dissociates ribosomal subunits from mRNA	To verify if GFP spots are translation sites	DMSO
Harringtonine	3 µg/mL	Translation inhibitor - blocks translocation of ribosome at initiation codon, without affecting downstream ribosomes	To estimate ribosome elongation rates on mRNA	DMSO

All the dilution for final concentration is made in Leibovitz's-15 medium

- considerations to take into account for selection of positions are explained in Box 1.
21. Find the correct focal plane using the mCherry channel.

CRITICAL STEP Because the PCP-mCherry is tethered directly to the plasma membrane, while the SunTag-scFv is tethered indirectly to the membrane (via the nascent chain, mRNA, and PCP-mCherry), the GFP signal is often located slightly above the mCherry fluorescence. Therefore, it is helpful to focus the objective slightly above the optimal z-position used to observe the PCP-mCherry signal such that both mCherry and GFP signals are observed as crisp foci. As the fluorescence intensity of GFP foci depends strongly on the precise focal plane, we recommend using an auto-focus system (e.g. Nikon Perfect Focus System) to prevent fluctuations in the focal plane during time-lapse experiments or between different experiments. In case of untethered reporter mRNAs, it is useful to focus on a region of the cytoplasm just below the nucleus, as mRNA mobility is often slightly lower in this region, facilitating long-term tracking of individual mRNAs.

22. Set appropriate laser power and exposure time for imaging. Follow option A for standard translation imaging and option B for imaging of single mature SunTag proteins (See step 41B).

CRITICAL STEP The precise laser power and exposure time will need to be optimized for each microscope and experiment.

A. Standard translation imaging

- i. For translation imaging experiments, use 500 ms exposure time for GFP, mCherry, and JF⁶⁴⁶.

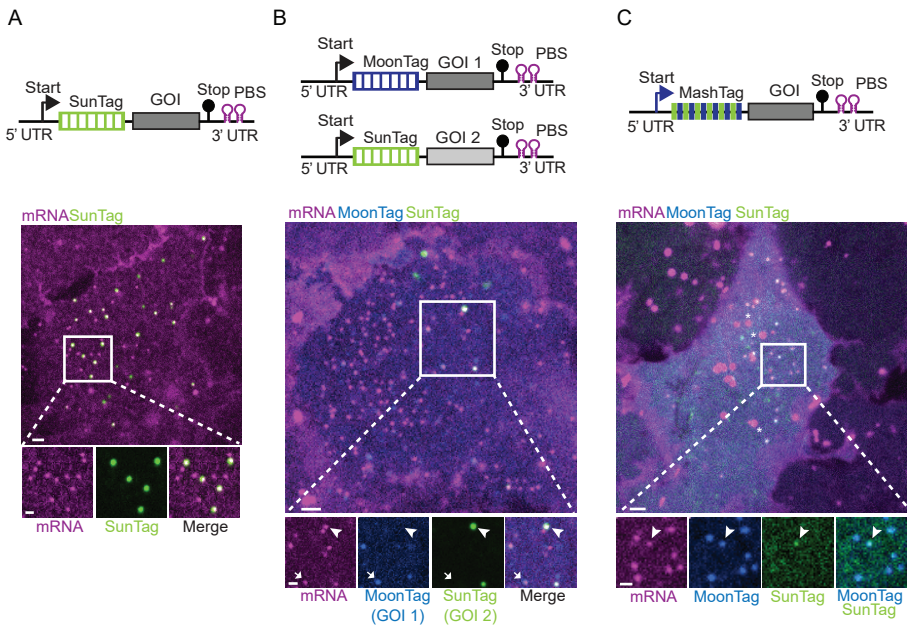


Figure 4. Representative imaging data for SunTag, MoonTag and MashTag reporters. (A-C) Schematic of indicated translation imaging reporters (top panels) and representative images of U2OS cells expressing SunTag-scFv, PCP-mCherry-CAAX (A, C) and MoonTag-nanobody (B) and indicated reporters are shown. Top images show the entire cell, and bottom images show a zoomed in view of the boxed areas of the cells. In (B), the arrowhead and arrow indicate SunTag and MoonTag translation, respectively. In (H) the arrowhead indicates mRNA with out-of-frame (SunTag frame) translation. Asterisks show fluorescent lysosomes. Scale bars, 2 μm (top images) and 1 μm (bottom images).

CRITICAL STEP Using a long exposure time results in motion blurring of the fast moving mature SunTag and/or MoonTag proteins, resulting in a more homogeneous background fluorescence, which facilitates detection of weak translation signals. Aim for a laser power that results in ~10-50% photobleaching over the duration of the entire experiment (~ 60-120 time-points). Images for different fluorescence channels can be obtained sequentially with a single camera. Good co-localization between GFP and mCherry foci is observed when images are acquired within ~1 s of each other. When available, a two-camera setup can also be used to reduce the time between images.

? TROUBLESHOOTING

Box 1. Considerations for selecting x, y positions during image acquisition.

We generally image multiple x, y positions simultaneously to acquire data from a larger number of cells in each experiment. The number of x, y positions that can be imaged simultaneously depends on a number of variables, including the time interval of imaging, the speed of the microscope and the number of channels that are imaged. For imaging SunTag reporters, we generally select 10-15 x, y positions to image simultaneously. We select positions based on the following considerations:

- a. *Mature SunTag proteins:* Presence of mature SunTag proteins diffusing in the cytoplasm is the first criteria to select cells for imaging, as the presence of mature SunTag proteins indicates that the cell is expressing the reporter plasmid. Alternatively, in the case of transient transfection, co-transfect a plasmid encoding BFP (or any other fluorescent marker) as a method to identify transfected cells. (See Troubleshooting)
- b. *Translation sites:* When mRNAs that are present from the start of the movie will be analyzed, select cells for imaging that contain ~10-40 translation sites. However, when the experiment requires visualizing the first round of translation, only mRNAs that initiate translation during the movie can be included in the analysis. In this case, start imaging earlier (~15 min after doxycycline addition) and select cells with fewer translation sites (3-10). In all cases, exclude cells with > 40 translation sites, as at high density tracking mRNAs is challenging and the translation signal will decrease because of cytoplasmic antibody depletion. (see Troubleshooting)
- c. *Fluorescent lysosomes:* PCP-mCherry can accumulate in lysosomes when stably expressed in cells (see Fig. 4F). Select cells with few fluorescent lysosomes as they interfere with mRNA tracking. In our experience, mCherry-positive lysosomes are visible in most cells at all expression levels of PCP-mCherry. Lysosomes can easily be distinguished from mRNAs based on fluorescence intensity and movement; lysosomes are generally much bigger, show very high fluorescence and move faster (and in a directed fashion) than tethered mRNAs. (see Troubleshooting)
- d. *Aggregation of mature SunTag proteins:* Avoid cells that show aggregation of mature SunTag proteins. We found that cells containing SunTag aggregates can be easily identified based on the presence of bright, largely immobile GFP-positive aggregates. Aggregation is not observed when using reporters encoding the SunTag only, but does occur when MashTag reporters are expressed at high levels. Aggregation likely occurs because of the different linker that is used in between peptide epitopes in the MashTag compared to the SunTag. While selecting x, y positions always use the lowest possible laser power and minimize the amount of time that cells are exposed to laser light to limit phototoxicity and photobleaching.

B. Mature protein imaging

- i. For measuring intensities of mature SunTag array proteins bound to SunTag-scFvs, use an exposure time of 20 ms, as this prevents motion blurring of fast-moving mature proteins. Use a laser power that is as high as possible without saturating pixel intensities on the camera.

? TROUBLESHOOTING

23. Set a time interval for the imaging experiment.

CRITICAL STEP In our experience, a time interval of 30 s allows for good tracking of tethered mRNAs and captures most translation dynamics. However, for some other applications, such as tracking untethered mRNAs or visualizing rapid translation dynamics, other time intervals are required.

24. Start image acquisition.

CRITICAL STEP When required, harringtonine (or other drugs) can be added during image acquisition. Prepare a solution containing 24 $\mu\text{g}/\text{ml}$ harringtonine in pre-warmed Leibovitz's-15 medium. Add 50 μl drug solution to the 350 μl medium and pipet up and down a few times (final concentration 3 $\mu\text{g}/\text{ml}$). Table 2 describes multiple drugs that can be used to assess translation dynamics, as well as their mechanism of action and optimal concentration for human cells.

CRITICAL STEP Be careful not to touch the microscope or the edges of the well when adding drugs, as this might cause a shift in the focal plane. Also, try to minimize the time during which the microscope chamber is open to limit changes in temperature.

? TROUBLESHOOTING

Image Analysis: Tracking single mRNAs and measuring fluorescence intensities of translating mRNAs over time with TransTrack (Timing 1-6 h)

CRITICAL We developed TransTrack, a MATLAB-based software package, to track single translating mRNAs and measure the intensity of the translation signal over time. TransTrack was developed to track membrane tethered mRNAs, but may also work for non-tethered mRNAs. Tracking of translating mRNAs is performed in a semi-automated manner, where tracks are generated automatically, but the user can manually select, reject, or adjust individual tracks. A tutorial, an example dataset, and expected outcomes of TransTrack analysis are included in Supplementary data 1.

25. Load a time-lapse movie of a single Z-plane or maximum intensity Z-projection of one or multiple x, y positions into TransTrack. Go to 'file' \rightarrow 'open'. Indicate how many fluorescence channels are present in the data, and select the folder in which the files are saved. All image files in the folder will now be loaded into TransTrack (Fig. 5)

CRITICAL STEP TransTrack only recognizes .tif file extensions. If microscopy

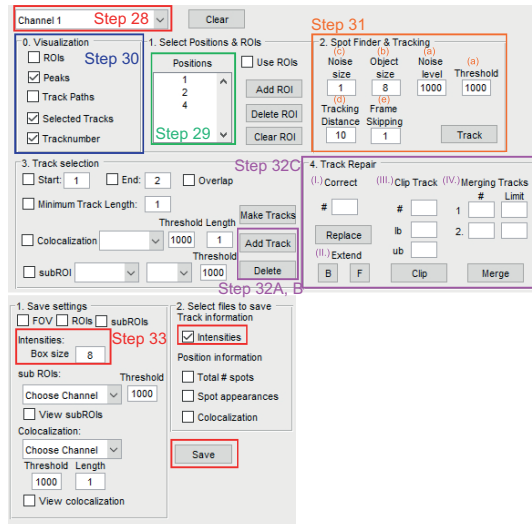


Figure 5. Graphic illustration of TransTrack software. Screenshot of the GUI of TransTrack. Option in the software are numbered in accordance with the corresponding steps in the Procedure. Different colors of boxes indicate different steps in the procedure.

images are acquired in another format (e.g. .nd2, as is the case for Nikon Elements software), files must first be converted to .tif format. This conversion can generally be performed using the image acquisition software or using ImageJ. While saving or exporting files to the .tif format, take the following considerations into account: 1) each .tif file should contain a single x, y position; 2) the .tif file should contain at most six channels and 3) the image can only consist of a single Z-plane (in case of a 3D stack of images, it is necessary to make 2D maximum intensity projections first).

26. Perform a correction for photobleaching. Select ‘analysis’ → ‘photobleach correction’ → ‘Correction’.

CRITICAL STEP The mean intensity of the entire image is determined for each channel at each time-point and an exponential function is fit to determine the rate of bleaching. Next, a correction is applied to the fluorescence intensity values of all channels in which fluorescence intensities will be measured. Note that the bleach correction will not work optimally when a large part of the image contains no signal. In this case, a manual bleach correction can be performed after obtaining intensity values from TransTrack.

27. Open the ‘single particle localization and tracking’ menu by selecting ‘analysis’ → ‘Single Channel Tracking’.
28. Select a channel for single particle localization and tracking (Fig. 5). Tracking translating mRNAs can be performed based on the mRNA signal (mCherry

Box 2. Parameters to be optimized during single particle tracing when using TransTrack

'Noise level' and 'Threshold': These parameters set the brightness threshold of particles that are detected. If some mRNAs are not detected when clicking 'Track', lower the value for these parameters. If too many spots are detected (including spots that do not represent translating mRNAs), increase the value of these parameters.

'Object size': the size of detected particles in pixels. The default size of 8 pixels is based on a pixel size of 135 nm. Adjust this value according to pixel size of your microscope so that the object size is $\sim 1 \mu\text{m}$.

'Noise size': This can generally be left at 1, but the user can experiment with different noise size values to see if a better distinction between translating mRNAs and background fluorescence foci is achieved.

'Tracking distance': The maximal distance (in pixels) that a particle can move between two consecutive frames while still being linked into a single track. The value for the tracking distance depends on the pixel size, the diffusion kinetics of an mRNA molecule and on the time interval between frames. If translation sites diffuse rapidly (e.g. when they are not tethered) or when using a long time interval between images, a higher value is required here.

'Frame skipping': A track will be ended when the particle is not detected for a selected number of time-points. When set to 1, particles are only connected when they are detected in all consecutive frames. We generally use a frame skipping value of 1, as faulty tracks can be generated when using frame skipping values larger than 1. The frame skipping value that can be used mainly depends on the particle density and brightness. At lower particle densities and brightness, frame skipping values larger than 1 may result in improved tracking.

can also be determined when the data consists of a single time-point, in which case tracks will only consist of a single value.

CRITICAL STEP In our experience, single particle tracking is more accurate when using translation signal for tracking, since the signal-to-noise in the translation channel is usually superior compared to the mRNA channel. However, the translation signal cannot be used for tracking when mRNAs are in an untranslated state part of the time. In this case, use the mRNA signal for single particle tracking.

29. When multiple x, y positions are loaded into 'TransTrack', indicate the number(s) of the x, y position(s) that will be analyzed.
30. Visualize tracks that are detected by the algorithm and tracks that are manually selected in the visualization menu by checking the boxes 'peaks', 'selected tracks' and 'track number'. Note that visualization of all tracks detected by the algorithm ('peaks' box) can require a lot of computational power and may make the software slower. If this happens, uncheck the 'peaks' box.

31. Localize and track single particles. Track particles by clicking the ‘Track’ button. This function detects particles at each time-point, and links them together in tracks. During single particle tracking, some parameters have to be optimized, which is explained in Box 2.

CRITICAL STEP More information about the algorithms used in TransTrack can be found on GitHub (github.com/TanenbaumLab/TransTrack).

32. After automated tracking has generated all the tracks, the tracks can be manually curated in a number of ways; follow instructions in option A for adding tracks, option B for deleting tracks and option C for correcting tracks.

A. Add tracks

- i. Select tracks by clicking on the ‘Add Track’ button, click on the mRNAs that should be included, and press enter.

B. Delete tracks

- i. If incorrect tracks have been manually selected, remove them by clicking on ‘delete’ and then clicking on the track that is to be removed.

C. Correct tracks:

CRITICAL Verify that each selected track contains the correct mRNA at every time-point. If a mistake is observed, the following corrective actions can be taken:

- i. *Replace an incorrectly tracked frame.* Go to frame in which the mRNA is incorrectly tracked, enter the track number at ‘#’, and click ‘replace’. Now click in the image on the correct location of the mRNA that is being tracked.
 - ii. *Extend the track.* Enter the track number at ‘#’ and click ‘B’ to extend backwards, or ‘F’ to extend forwards. The image will now go to the first frame in which the particle was not tracked anymore. Clicking on the particle will add it to the track, and move the image to the next frame. When finished extending, press enter.
 - iii. *Clip the track.* This option should be used if the track includes frames at either the start or end that should be excluded. Indicate the track number at ‘#’ (below ‘Clip Track’), and enter the lower bound (‘lb’) and upper bound (‘ub’) of the frames that should be included in the track.
 - iv. *Merge tracks.* Sometimes, a single mRNA is split into two tracks, for example when the particle moves more than the maximally allowed tracking distance. In this case, two tracks can be merged. Indicate the numbers of the tracks at ‘Merging Tracks - #’, and indicate the last frame of the first track and the first frame of the second track in ‘limit’. If there is a gap between the merged tracks, the missing frames need to be added manually.
33. Save the tracks from TransTrack by clicking on ‘File’ → ‘Save’. In the ‘save’ menu, two settings should be selected; 1) the ‘box size’, which is the pixel size of the ROI

that is used for measuring fluorescence intensities. This ROI is centered around the location of the tracked particle. We typically set a box size of $\sim 1 \mu\text{m}$. Since images on our microscope have a pixel size of 135 nm, we used a box size of 8 pixels. Ensure that all the fluorescence signal that needs to be measured is within the selected box size. 2) In the 'save' menu, select the 'Intensities' checkbox. Tracks are saved and intensities are quantified and saved by TransTrack when clicking the 'Save' button.

CRITICAL STEP The results will be stored in an Excel file which contains all track information. For each fluorescence channel, a separate sheet is generated in Excel that contains the intensities of all tracks. These intensities are background corrected and can be used for downstream analyses.

Data Analysis: Measuring translation elongation rates (Timing 2 h)

CRITICAL To measure translation elongation rates, a small molecule inhibitor of translation, harringtonine, can be used. Harringtonine stalls ribosomes on the start codon, but does not affect translating ribosomes downstream of the start codon. Thus, addition of harringtonine to cells results in a progressive loss of fluorescence signal associated with translating mRNAs, as ribosomes on the mRNA complete translation one-by-one (referred to as ribosome run-off). The speed at which the fluorescence translation signal is lost reports on the translation elongation speed. Below, we will describe how to perform ribosome run-off experiments and the analysis steps required to calculate the translation elongation speed from these experiments.

34. *Image acquisition (steps 34-35)*. Prepare cells and acquire images as described in steps 15-24. During cell selection (steps 20, 21 and Box 1), select cells that have 10-40 translation sites. Use imaging conditions as described in step 22A.
35. Five minutes after the start of imaging, add harringtonine (3 $\mu\text{g}/\text{ml}$ final concentration) as described in step 24 and Table 2.
36. *Image analysis*. Use TransTrack to measure the GFP intensity of nascent SunTag peptides associated with translating mRNAs over time. Track translating mRNAs based on the GFP channel. Make sure that the tracks continue until the entire GFP signal has disappeared; if the track stops before the GFP signal has disappeared completely, add additional time-points (step 32Ai).
CRITICAL STEP Select mRNAs that were already being translated at the start of the image acquisition. mRNAs that initiated translation shortly (1-2 min) before harringtonine addition only have ribosomes on the 5' part of the mRNA, which increases the complexity of the analysis as the translation signal first increases when ribosomes translate the SunTag sequence.
37. *Data analysis (steps 37-40)*. Normalize the intensity of each trace to the intensity of

- that trace in the frame before harringtonine addition.
38. For each translating mRNA, the total translation time is given by the time from harringtonine addition until complete disappearance of the GFP signal.
 39. Determine the elongation rate by dividing the reporter length by the total translation time.
 40. Calculate the average elongation rate of all mRNAs.
- CRITICAL STEP* Note that the translation elongation value gives the average translation elongation rate over the entire transcript. It is possible that certain sections of the mRNA are translated faster than others.

Data Analysis: Quantifying translation initiation kinetics using RiboFitter (Timing 10-180 min)

CRITICAL Generally, multiple ribosomes contribute to the total fluorescence intensity of a translating mRNA for our reporter mRNAs, and changes in fluorescence intensities are caused by a complex combination of translocating ribosomes in different positions along the mRNA. RiboFitter is a computational algorithm which deconstructs fluorescence intensity traces to determine the number of ribosomes translating an mRNA, as well as the moment at which each ribosome initiated translation. To run RiboFitter, one needs to know the intensity profile that is associated with a single ribosome that translates the mRNA. Since the peptide array is located at the 5' end of the coding sequence, the single ribosome intensity profile consists of 2 phases: 1) A phase in which the peptide epitopes are translated ('buildup phase'). During this buildup phase, the fluorescence intensity increases over time as the ribosome translates the SunTag peptides one-by-one. 2) A phase in which the ribosome translates sequences downstream of the peptide array, during which the fluorescence intensity remains constant ('plateau phase') (Fig. 6A). Dedicated experiments are required to determine the intensity profile of a single ribosome before applying RiboFitter. Below, we discuss how to perform these dedicated experiments and how to run RiboFitter software.

CRITICAL To use RiboFitter, the following input is required: 1) bleach-corrected fluorescence intensity time traces of translating mRNAs (which can be generated with TransTrack, steps 25-33), 2) the image acquisition time interval (in seconds) and 3) the intensity profile of single ribosomes (Fig. 6A). The intensity profile of single ribosome is calculated based on the buildup time (in seconds), the plateau time (in seconds) and the plateau intensity (this is the intensity associated with a single translating ribosome that has completed synthesis of the entire peptide epitope array, and is equal to the intensity of a single mature SunTag protein).

41. To calculate build up and plateau time, follow option A. To determine plateau intensity of SunTag and MoonTag peptide array, follow options B and C respectively. For running RiboFitter, follow option D.

A. Calculating the buildup and plateau time (Timing 10 min)

- i. Determine the translation elongation rate in a set of dedicated experiments (steps 34-40).
- ii. Determine the number of nucleotides of the peptide array and the number of nucleotides downstream of the peptide array.
CRITICAL STEP When analyzing stop codon readthrough-based 3' UTR translation, the length of the plateau phase depends on whether a ribosome translates the 3' UTR or not. In this case, calculate the plateau length based on a situation without 3' UTR translation and correct for 3' UTR translation after running RiboFitter (step 42B).
CRITICAL STEP Since the plateau phase starts when the last peptide has fully exited the ribosome exit tunnel, the number of nucleotides in the plateau phase equals the number of nucleotides after the peptide array minus the length of the ribosomal exit tunnel (~30 amino acids = 90 nt).
- iii. Calculate the buildup time (in seconds) by dividing the number of nucleotides of the peptide array by the elongation rate (in nucleotides per second).
- iv. Calculate the plateau time (in seconds) by dividing the number nucleotides in the plateau phase by the elongation rate (in nucleotides per seconds).

B. Calculating the plateau intensity of a SunTag peptide array (Timing 30 min)

CRITICAL The plateau intensity of a ribosome translating the SunTag peptide array is equal to the intensity of a single mature SunTag protein. The intensity of single mature proteins can be measured with distinct image acquisition settings (step 22B) in cells expressing the translation reporter. The intensity of single mature proteins will be compared to the intensity of translation sites. However, the two intensities are determined using different acquisition parameters, so a calibration step is required to convert the fluorescence intensities between the two image acquisition methods, which is described below:

- i. *Image acquisition (steps i-ii)*. Prepare cells and acquire images as described above (steps 15-24). Select 10-25 x, y positions that have 3-15 translation sites and show crisp mature protein foci (see Fig. 6B). Acquire images using mature protein imaging settings (step 22B). Imaging a single time-point suffices for this analysis.

CRITICAL STEP Perform imaging before, or shortly after doxycycline addition (<15 minutes), as fewer mature proteins have been produced at this time, which facilitates detection of individual mature proteins.

CRITICAL STEP To validate the fluorescence intensity of mature proteins, we recommend performing a parallel experiment in which mature proteins are tethered to the plasma membrane, which facilitates

- intensity measurements due to reduced diffusion of mature proteins (see Box 3).
- ii. Acquire a second image at the same x, y positions using standard translation imaging settings (step 22A). Acquire this second image as quickly as possible after acquisition of the previous image to ensure that the translation signal associated with individual mRNAs does not substantially change between the images acquired in i and ii.
 - iii. *Image analysis (steps iii-v)*. Use TransTrack to measure the intensity of mature proteins (step 25-33). Since the data contains only a single time-point, TransTrack will make ‘tracks’ of a single time-point. Manually select 10-20 mature protein spots in each cell.

CRITICAL STEP Select foci that appear as crisp spots, as these are not motion blurred and are in focus. Mature proteins that are slightly out of focus appear blurred. We found that single mature proteins can most easily be identified in regions with few SunTag proteins present (for example, the region of the cell under the nucleus or in the periphery of the cell).

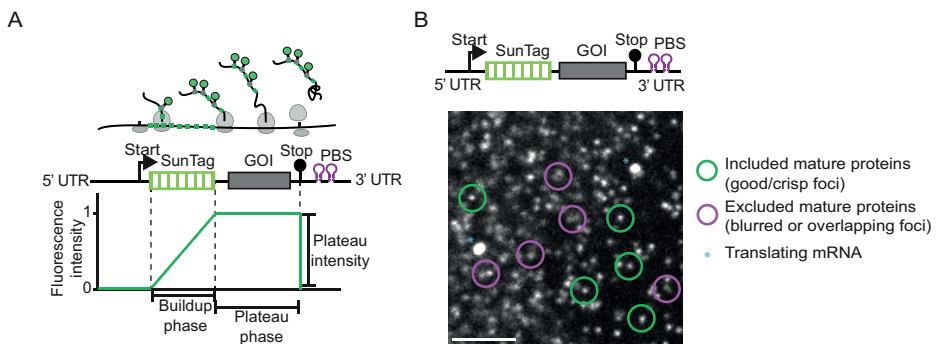


Figure 6. Fluorescence intensity profile of a ribosome translating a SunTag reporter. A) The intensity profile of a single ribosome translating a SunTag reporter consists of 3 phases: 1) a phase in which the ribosome is not associated with fluorescence. During the first phase the sequence upstream of the peptide array and the first peptide epitope is translated. 2) A buildup phase, in which the fluorescence intensity increases linearly over time. This phase starts and ends when the first and last peptide epitope have left the ribosome exit tunnel, respectively. 3) A plateau phase, in which all peptide epitopes have been synthesized and labeled with SunTag-scFv. The plateau intensity is equal to the intensity of a mature SunTag protein labeled with SunTag-scFv. After completion of the third phase, the ribosome terminates translation and releases the nascent polypeptide, resulting in rapid loss of fluorescence associated with the mRNA. B) A U2OS cell expressing SunTag-scFv, PCP-mCherry-CAAX and indicated translation reporter was imaged with mature protein settings (Step 22B(i)). Green circles show examples of mature proteins that could be included in the analysis; these foci are crisp (not motion blurred) and not overlapping with other foci. Magenta circles show examples of mature proteins that should be excluded from analysis, as they are motion blurred or overlap with other foci. Blue asterisks indicate translating mRNAs. Scale bar, 1 μm .

Box 3. An independent method to measure the fluorescence intensity of single mature proteins

It is recommended to use an independent approach to validate the brightness of mature proteins when mature proteins are measured for the first time. This is done by expressing a plasmid encoding SunTag-CAAX (Addgene #134833) instead of a translation imaging reporter. Mature SunTag-CAAX proteins made from this plasmid should have equal brightness as mature proteins expressed from a translation reporter, as they contain the same 24xSunTag peptide array. However, the mature proteins are tethered to the plasma membrane through prenylation of the CAAX domain and therefore diffuse more slowly, facilitating accurate measurements of their brightness. Measure the intensity of mature SunTag-CAAX proteins by performing step 41Bi, ii, iii, vi, vii. The mean intensity calculated in step 41Bvii should be comparable to the mean intensity calculated for mature proteins expressed from the translation reporter.

- iv. Use TransTrack to measure the intensity of the translation signal of translating mRNAs for images acquired using mature protein imaging settings (step 25-33). When saving the data (step 33), use the same 'box size' that was used in step 41Biii.
- v. Use TransTrack to measure the intensity of the translation signal of translating mRNAs for images acquired using standard translation imaging settings (step 25-33). When saving the data (step 33), use the same 'box size' that was used in step 41Biii.
- vi. *Data analysis (steps vi-viii)*. Determine the relative intensity of translation signals with the two imaging conditions by dividing the mean intensity of translation sites measured with standard imaging conditions (step 441v) by the mean intensity of translation sites measured with mature protein imaging conditions (step 41Biv).
- vii. Calculate the mean intensity of all mature proteins.
- viii. Multiply the mean intensity of a single mature protein with the calculated relative intensity of translation signals (step 41Bvi). This value is the intensity of a single mature SunTag protein when using standard imaging conditions, which represents the plateau intensity of a single ribosome translation profile (Fig. 6A).

CRITICAL STEP Although we recommend RiboFitter for analysis of the number of ribosomes per mRNA, the number of ribosomes that translate an mRNA at any given time-point can also be estimated using the fluorescence intensity of a translating mRNA (step 41Biv) and the mean intensity of a mature protein (step 41Bvii). We discuss this alternative method in detail in Box 4.

C. Calculating the plateau intensity of a MoonTag peptide array (Timing 30 min)

CRITICAL Single mature MoonTag proteins are not bright enough for reliable detection using spinning disk confocal microscopy when imaging MoonTag or MashTag reporters. Therefore, the brightness of a single mature MoonTag protein needs to be determined indirectly. This is done by calculating the number of ribosomes on a SunTag translation reporter, and comparing the SunTag reporter with a MoonTag translation reporter with identical promoter, 5'UTR, and translation start site. An assumption is made that these matched SunTag and MoonTag reporters have a similar translation initiation rate and ribosome density. Therefore, one can deduce the average number of ribosomes on the MoonTag translation reporters from the ribosome density on the SunTag reporters. The ribosome density on the MoonTag reporters, combined with the fluorescence intensity of translating MoonTag mRNAs, can be used to determine the intensity of single mature MoonTag proteins. Below, we provide a detailed protocol for these analyses. The analysis is described for MoonTag reporters, but is identical when using MashTag reporters.

- i. Acquire images and calculate the average number of ribosomes on SunTag-containing reporter mRNAs (e.g. Addgene # 74928, see Fig. 3A) as described in step 41B (Perform steps 41Bi, ii, iii, vii and see Box 4).
- ii. Calculate the average number of ribosomes on MoonTag-containing reporter mRNAs (steps ii-iii). First, divide the length of the MoonTag reporter coding sequence by the length of the SunTag reporter coding sequence.
- iii. Then, multiply the value obtained in step 41Ci by the average number of ribosomes on a SunTag-containing reporter mRNA (step 41Ci). This value is the average number of ribosomes on a MoonTag-containing reporter mRNA.
- iv. *Image acquisition (steps iv-v)*. In a parallel experiment, prepare cells expressing a MoonTag- or MashTag containing reporter (e.g. Addgene #128604 or #128607, see Fig. 3B, 3D) and acquire images as described in steps 15-24. Select 10-25 cells that have 3-15 translating mRNAs and acquire images using standard imaging settings for MoonTag (step 22A).
- v. *Image analysis*. Use TransTrack to measure the intensity of the translation signal of translating mRNA (step 25-33).
- vi. *Data analysis (steps vi-ix)*. Calculate the mean fluorescence intensity of translating mRNAs measured in step 41Cv.
- vii. Calculate the correction factor for the MoonTag reporter (see Box 4).
- viii. Divide the mean intensity of translating MoonTag mRNAs (step 41Cvi)

Box 4. An alternative method to calculate the number of ribosomes translating an mRNA at any given time-point without using RiboFitter

In this method, the number of ribosomes that translate an mRNA at any given time-point is estimated using the fluorescence intensity of a translating mRNA (step 41Biv) and the mean intensity of a mature protein (step 41Bvii). This method is somewhat faster than running RiboFitter, but it is less accurate and does not provide temporal information about translation initiation events. An assumption is made that the mRNA is homogeneously covered by ribosomes. To calculate the number of ribosomes translating the mRNA a correction is applied to account for ribosomes that haven't completed translation of the entire SunTag peptide array, and thus are associated with a lower level of fluorescence:

1. Calculate the correction factor (CF) for the intensity profile of a translating ribosome using the following equation:

$$CF = \frac{(0.5 \times L(\text{buildup})) + L(\text{plateau})}{L(\text{total})}$$

In this equation, CF is the correction factor, L(buildup) is the nucleotide length of the SunTag array (step 41Aii), L(plateau) is the number of nucleotides translated during the plateau phase (step 41Aiii), and L(total) is the nucleotide length of the complete coding sequence.

2. Divide the fluorescence intensity of a translating mRNA (step 41Biv) by the mean intensity of a mature protein (step 41Bvii).
3. Divide the value obtained in (2) by the correction factor (CF) calculated in (1) to determine the number of ribosomes translating the mRNA at a given time-point.

by the correction factor (step 41Cvii).

- ix. Divide the corrected mean intensity (step 41Cviii) by the average number of ribosomes on the reporter mRNA (step 41Ciii). This value is the intensity of a single mature MoonTag protein, which represents the plateau intensity of a single ribosome translation profile.

D. Running RiboFitter (Timing – 1-180 min)

CRITICAL Once the buildup time, plateau time and plateau intensity of a single translating ribosome are known, RiboFitter can be applied to fluorescence intensity time traces of single translating mRNAs. Two example traces, as well as the parameters required for running RiboFitter and the expected outcome are included in Supplementary Data 2.

- i. Create a comma delimited .csv file containing the bleach-corrected intensity traces. This can be done in excel by going to 'save as', and then selecting 'CSV (comma delimited)' in the 'save as type' menu. Each column should

contain one intensity trace. The first column is used to indicate the time (in frames).

CRITICAL STEP When using intensity traces of both SunTag and MoonTag, make two separate files, in which the first file contains SunTag intensities and the second file contains MoonTag intensities. Ensure that for each mRNA molecule, the SunTag and MoonTag data are in the same column in their respective files.

- ii. Open the RiboFitter script in R or RStudio. Set the correct buildup time (step 41Aiv), plateau time (step 41Av), plateau intensity (step 41B or 41C), and acquisition interval for the SunTag in lines 5-9 of the script.

CRITICAL STEP When analyzing dual-color translation imaging data, set the parameters for the MoonTag in lines 15-17 of the RiboFitter script.

- iii. Select the entire script and press the execute button in the R interface. A dialog box will appear.
- iv. Select the file containing the intensity traces of the SunTag signal. A pop-up window appears with the question whether a fit for a second channel should be included. When selecting ‘no’, RiboFitter ignores the second channel and starts fitting translation initiation events for the first channel only. If ‘yes’ is ticked, a new dialog will appear and the file containing the intensity traces of the second channel can be selected.
- v. RiboFitter will now initiate the iterative process of fitting intensity traces to identify the most probable number of ribosomes that have translated the mRNA, as well as the timing of each translation initiation event. Note that if many highly translated mRNAs are analyzed, running the script can take several hours. Once the script has finished the calculations, the following message will appear in the console window of the R interface:

```
[1] "Finished"
```

- vi. A dialog box will appear. Select a folder to save the output files in.
CRITICAL STEP As output files, RiboFitter provides graphs of intensity traces along with their optimal fits, as well as excel files with intensity values. See Box 5 for more details about these output files.

Data Analysis: Interpretation of RiboFitter data (Timing- 10 min)

CRITICAL To quantitatively compare the translation rates of 1) two different mRNA reporters (one containing the SunTag and the other containing the MoonTag), 2) a single reporter containing both the SunTag and MoonTag (e.g. the 3'UTR translation reporter) or 3) a reporter containing the MashTag, we recommend using RiboFitter to determine the exact number of translation initiation events in each channel (i.e. SunTag

or MoonTag) and performing the downstream analysis described below.

42. For analyzing experiments containing two different reporters, follow option A. For determining the fraction of ribosomes undergoing 3'UTR translation, follow option B, and for calculating the fraction of ribosomes translating in different frames on MashTag reporters, follow option C. An example RiboFitter output is shown in Fig. 7.

A. Analysis of experiments containing two different reporters.

- i. When SunTag and MoonTag are expressed simultaneously in a cell, calculate the translation initiation rate of each reporter independently by dividing the total number of ribosomes that translated the mRNA (column 2 or 5 of RiboFitter output, see Box 5) by the duration of the mRNA trace. The rate of translation initiation of both reporters can be compared directly.

B. Calculating the fraction of ribosomes translating the 3' UTR of a reporter.

- i. Determine the number of ribosomes that translated the MoonTag sequence and the SunTag sequence for each translating mRNA (column 2 and 5 of RiboFitter output, see Box 5).

CRITICAL STEP When the SunTag is placed in-frame with the stop codon and the predominant SunTag translation signal is caused by stop codon readthrough, a correction is required for the number of MoonTag ribosomes; when a ribosome reads through the stop codon, the duration of the plateau phase of ribosomes that translated the MoonTag is increased as a longer nucleotide sequence downstream of the MoonTag array is translated. Correct for this using the following steps (ii-viii).

- ii. Determine the number of ribosomes that translated the 3' UTR. This is equal to the number of ribosomes that translated the SunTag peptide array (step 42Bi).
- iii. Calculate the plateau time of a ribosome that translates the 3' UTR (step 41A).
- iv. Calculate the integrated fluorescence intensity ('FI') produced for a ribosome that only translates the annotated coding sequence, and the integrated fluorescence intensity produced for a ribosome that translates both the annotated coding and 3' UTR sequences, using the following equation:

$$FI = (0.5 \times \text{buildup time}) + \text{plateau time}$$

Note that the buildup and plateau times (without 3' UTR translation) have been calculated previously in order to run RiboFitter (step 41A). The plateau time (when 3' UTR translation occurs) should be calculated as

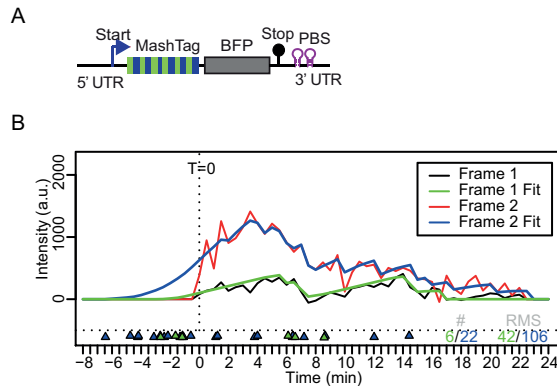


Figure 7. Example output of RiboFitter. A) Schematic of the reporter used to generate the fluorescence intensity time traces shown in (B). The reporter contains a single AUG translation initiation sequence, which is positioned in the MoonTag reading frame. B) Analysis by RiboFitter of fluorescence intensity time traces of an example mRNA of the reporter indicated in (A). Intensity time traces are shown for SunTag (black line) and MoonTag (red line) signals. Best fits obtained using RiboFitter are shown (green and blue lines for SunTag and MoonTag, respectively). Colored triangles below the traces indicate translation initiation events. Green and blue numbers and triangles represent SunTag and MoonTag, respectively. The total number of translation initiation events (#) and the root mean squared error (RMSE) (which indicates the goodness-of-fit) are indicated on the right. $T = 0$, time at the start of translation

described in step 41A.

- v. Calculate the additional integrated fluorescence intensity produced by each ribosome that translated the 3' UTR. This is given by the following equation:

$$Fl(\text{additional}) = \frac{Fl((3' \text{ UTR}) + Fl(\text{no } 3' \text{ UTR}))}{Fl(\text{no } 3' \text{ UTR})}$$

In this equation, $Fl(\text{additional})$ is the extra fluorescence produced when the 3' UTR is translated, normalized to the number of ribosomes, while $Fl(3' \text{ UTR})$ and $Fl(\text{no } 3' \text{ UTR})$ are the integrated fluorescence intensities calculated in step 42Biv.

- vi. Multiply $Fl(\text{additional})$ with the number of ribosomes that translated the 3' UTR (step 42Bii).
- vii. Subtract the value determined in step 42Bvi from the total number of ribosomes determined by RiboFitter to get the corrected number of MoonTag ribosomes.
- viii. Divide the number of SunTag ribosomes by the corrected number of MoonTag ribosomes to obtain the fraction of ribosomes that has translated the 3' UTR.

Box 5. Output files generated by RiboFitter

RiboFitter generates two output files:

Plots.pdf

Each page of this pdf file contains a plot of an experimentally-determined intensity trace and its optimal fit as determined by RiboFitter (Fig. 7). In the box below the intensity traces, the moment of every translation initiation event on the mRNA is indicated by colored triangles. The total number of translating ribosomes, as well as the goodness-of-fit (measured by the root-mean-square error (RMSE)), is displayed on the right-hand side of the box. The vertical dotted line in the graph with intensity traces indicates the start of image acquisition. Translation initiation events can also be positioned to the left of this line by RiboFitter if a translation signal was already present at the start of image acquisition. It is recommended to visually inspect whether the fits match the data well.

All_results.csv

All_results.csv is a text file containing the information listed below for the best fit obtained for the iterative fitting process. This file can be opened in Excel. Data for each translating mRNA is given in a single row and contains the following information:

Column 1: mRNA ID.

Column 2: The total number of translating ribosomes in the entire time trace for the first fluorescence channel (SunTag).

Column 3: The time of initiation (seconds) for each ribosome in the first fluorescence channel (SunTag).

Column 4: The RMSE of the fit to the data for the first fluorescence channel (SunTag).

Column 5: The total number of translating ribosomes in the entire time trace for the second fluorescence channel (MoonTag).

Column 6: The time of initiation (seconds) for each ribosome in the second fluorescence channel (MoonTag).

Column 7: The RMSE of the fit to the data for the second fluorescence channel (MoonTag).

Column 8: The time (in seconds) since the start of the intensity time trace.

Column 9: The number of ribosomes translating the mRNA in the first translation frame at each time-point.

Column 10: The number of ribosomes translating the mRNA in the second translation frame at each time-point.

- C. **Calculating the fraction of ribosomes translating either the MoonTag or SunTag reading frame on MashTag mRNAs.**
- i. Determine the number of ribosomes that translated either the MoonTag sequence or the SunTag sequence for each translating mRNA (column 2 and 5 of RiboFitter output, see Box 5).
 - ii. Divide the number of ribosomes that translated the SunTag sequence

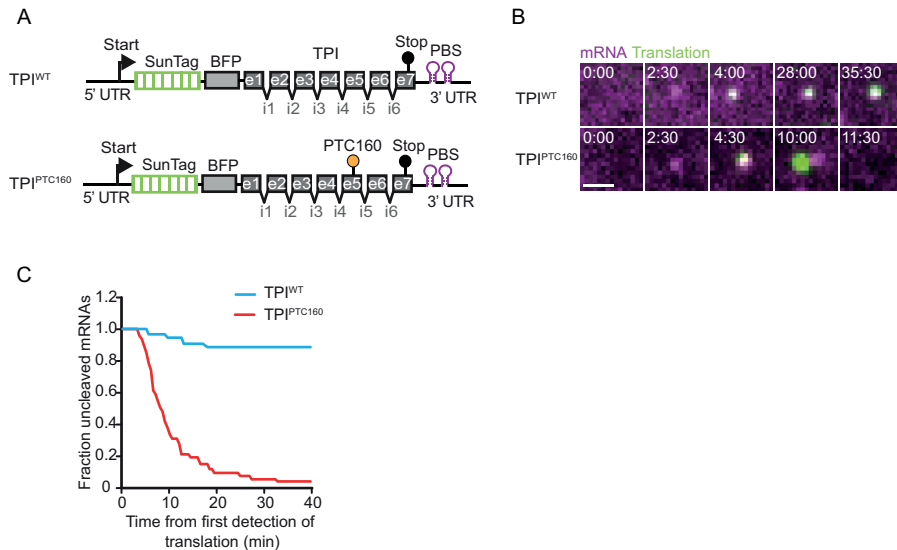


Figure 8. Representative imaging data and analyses for experiments on NMD . A) Schematic of indicated NMD reporter constructs containing the triosephosphate isomerase (TPI) gene. PTC160 indicates a PTC at amino acid 160. B) Representative images of U2OS cells expressing SunTag-scFv (green), PCP-mCherry-CAAX (magenta) and the NMD reporters indicated in (A) are shown. Scale bar, 1 μ m. Time is shown as minutes:seconds. C) The time from first detection of translation until separation of GFP and mCherry (i.e. mRNA endonucleolytic cleavage) was quantified. Note that little cleavage is observed during the first 3 min from the start of translation, which represents the time it takes for the first ribosome to reach the stop codon. Note that a fraction of mRNA molecules (~10%) appear to be insensitive to NMD-dependent mRNA cleavage. E, exon; in, intron; WT, wild type.

by the total number of ribosomes (SunTag + MoonTag) to obtain the fraction of ribosomes that translated in the SunTag reading frame.

CRITICAL STEP In the steps described above, the fraction of ribosomes translating the SunTag and MoonTag frame is calculated. In these calculations, the ribosomes translating the third frame (referred to as the ‘blank’ frame), which are not visible, are not taken into account.

Data Analysis: Analysis of translation-coupled mRNA degradation (NMD) (Timing 1-3 h)

CRITICAL We recently developed an assay to visualize endonucleolytic cleavage of single mRNA molecules using translation imaging²³. As NMD frequently results in endonucleolytic cleavage, we can observe the moment of NMD induction using this assay. Below, we provide a detailed protocol of this NMD imaging assay.

43. **Image acquisition.** Select cells with 3-10 translation sites (steps 15-24) and acquire images with standard imaging settings (step 22A).

44. *Image analysis (steps 44-45)*. For each mRNA that appears in the field-of-view during the movie, collect the following information: 1) the moment that translation signal can first be observed on the mRNA and 2) either the moment of mRNA cleavage (see Fig. 8B) or, in the case that the mRNA molecules does not undergo cleavage during the movie, the last time-point in which the mRNA can be detected/tracked.
45. *CRITICAL STEP* There are multiple possible reasons why cleavage may not be observed for a particular mRNA molecule. For example, an mRNA can move out of the field-of-view, an mRNA can no longer be reliably tracked because it crosses paths with another mRNA or the end of the movie is reached. In addition, since NMD is coupled to mRNA translation, it is critical that only mRNA molecules are included in the analysis for which all translating ribosomes are observed. Therefore, only mRNAs are analyzed for which the first round of translation could be observed. To ensure that the first round of translation is observed, only include newly-transcribed mRNAs that first appear in the field-of-view in the absence of associated translation signal²³.
46. Prevent bias in selection of mRNAs that will be analyzed. If a subset of mRNAs in the cell are ‘randomly’ selected for analysis, the selection of mRNAs can inadvertently introduce a bias, for example because mRNAs associated with a bright GFP signal are more easily detected than mRNAs associated with a dim GFP signal. To prevent an analysis bias, either analyze all mRNAs in a cell that match the criteria listed above, or, if the number of mRNAs in a cell is too large, create an ROI within a cell and analyze all mRNAs in the ROI.
47. *CRITICAL STEP* In addition to mRNA cleavage, other parameters can be scored as well, including the fluorescence intensity of the translation signal, the fraction of mRNAs that disappear during the movie (which could indicate exonucleolytic cleavage), or the time from mRNA cleavage to the disappearance of the 3’ cleavage fragments (the mCherry foci after cleavage), which represents exonucleolytic decay of the 3’ cleavage fragment by XRN1²³.
48. *Data analysis*. For each mRNA molecule, calculate the time from first detection of GFP until cleavage or until the last time-point that the mRNA could be detected/tracked (if the mRNA is not cleaved during the experiment).
49. *CRITICAL STEP* We typically represent the results as a Kaplan-Meier plot. This plot type takes into account the track length of both cleaved and uncleaved mRNAs, and shows the fraction of uncleaved mRNAs over time (relative to the start of translation). We have included a template for making a Kaplan-Meier plot in Excel in Supplementary Data 3.

TIMING

- Step 1, Design of reporter mRNA: 2-5 d
 Steps 2-13, Generation of cells lines suitable for translation imaging: ~10-25 d
 Step 14Ai-iii, Transient transfection for expression of reporter mRNA: 2 d
 Step 14Bi-iv, Stable integration of reporter plasmid: ~12-20 d
 Steps 15-19, Sample preparation for imaging: 1-2 d
 Steps 20-24, Image acquisition: 1-2 h
 Steps 25-33, Image analysis: 1-6 h
 Steps 34-40, Data Analysis - Measuring translation elongation rates: 2 h
 Step 41Ai-iv, Calculating the buildup and plateau time: 10 min
 Step 41Bi-viii, Calculating the plateau intensity of a SunTag peptide array: 30 min
 Step 41Ci-ix, Calculating the plateau intensity of a MoonTag peptide array: 30 min
 Step 41Di-vi, Running RiboFitter: 1-250 min
 Step 42, Interpretation of RiboFitter Data: 10 min
 Steps 43-46, Analysis of translation-coupled mRNA degradation (NMD): 1-3 h

TROUBLESHOOTING

Troubleshooting advice can be found in Table 3.

Table 3: Troubleshooting

STEP	PROBLEM	POSSIBLE REASON	SOLUTION
5i(a), 5iii(b)	No mature protein detected	Cells do not express the reporter	Ensure expression of reporter, for example through BFP co-transfection
	Mature proteins have low signal to noise	too high expression of SunTag scFv - mature protein not visible above background	Sort cells with lower SunTag scFv expression levels
		too low expression of SunTag scFv - not enough peptides to bind to mature protein	Sort cells with higher SunTag scFv expression levels
		Too much mature protein present in the cell	Induce expression of reporter for shorter time
		Signal below detection limit	Use higher laser power, or more sensitive camera, if available.

Table 3: Troubleshooting continued

STEP	PROBLEM	POSSIBLE REASON	SOLUTION
	Mature proteins blurred	exposure time is too long Too fast diffusion	Reduce exposure time Image in region underneath the nucleus (as diffusion is confined)
5i(b)	No translating mRNAs detected	Cell is not transfected with reporter Induction of transcription of the reporter is slow Expression of SunTag scFv is too high Too high expression of mature protein - cytoplasmic antibody depletion	Ensure expression of reporter, for example through BFP co-transfection Image cells later after doxycycline addition Sort cells with lower SunTag scFv expression levels Induce transcription for shorter time
	Unsure if foci in translation channel represent translating mRNAs	Protein aggregates can sometimes also appear as foci	Add puromycin (see Table 2): translation signal should disappear within a minute Determine whether the translation signal co-localizes with an mRNA spot
5i(c)	Are mCherry foci mRNAs or lysosomes?	Lysosomes can also appear as foci in the mRNA channel	Use lysotracker to ensure that foci do not co-localize with lysotracker
5ii	No mRNA foci are detected	Expression level of PCP-mCherry(-CAAX) is too high	Sort cells with lower PCP-mCherry expression levels
5v	Translation signal disappears during image acquisition	If background signal also disappears – high amount of photobleaching If background signal remains constant – cytoplasmic antibody depletion	Reduce laser power Select cells with fewer translating mRNAs, or sort cells with higher antibody expression

ANTICIPATED RESULTS

Fluorescence images

In this protocol, we describe several applications of the SunTag- and MoonTag-based translation imaging systems. Using this protocol, translation rates of different mRNA reporters, or of different regions of a single mRNA, can be quantitatively compared. In Fig. 4, typical images are shown of cells expressing reporters containing the SunTag and/or MoonTag peptide array. Note that some mRNA molecules are not associated with translation signal. While a subset of these mRNAs may undergo translation at a later time-point, we also usually find a subset of mRNAs that is never associated with translation signal. The reason for this is unclear, but it is possible that differences in transcription start site selection or cryptic splicing events have altered the coding sequence of these mRNAs such that the peptide epitope array is no longer present or cannot be translated. These mRNAs are usually excluded from analysis. Signal-to-noise ratio for mRNA and translation foci is often reduced if cells have incorrect expression levels of the SunTag/MoonTag antibody or PCP-mCherry(-CAAX), or if reporter mRNAs are expressed at very high levels, which results in cytoplasmic antibody depletion or aggregation of mature proteins (see Troubleshooting).

Data analysis

As discussed in this protocol, image analysis and data quantification is usually required for interpretation of fluorescence images. In many cases, image analysis consists of measurements of fluorescence intensities of translating mRNAs using TransTrack (Fig. 5), followed by analysis of ribosome number and translation initiation timing using RiboFitter (Fig. 7). We have provided a tutorial, a dataset for image analysis using TransTrack, and the expected outcome in Supplementary Data 1. To facilitate the use of RiboFitter, two example traces, fitting parameters, and the expected RiboFitter output are included in Supplementary Data 2. Typical analyses include: comparison of the translation initiation rate of two reporters or of one reporter under two different conditions, the frequency of 3' UTR translation compared to main ORF translation of the same mRNA, or the frequency of translation initiation in different translation reading frames. When assessing NMD, a comparison is typically made in the rate of NMD and the fraction of mRNA molecules sensitive to NMD for two different reporters (Fig. 8). More in depth analyses can also be performed and have been described previously^{21,23}. Because all the analyses described in this protocol quantify parameters of single mRNA molecules, one can also assess the heterogeneity between cells and between mRNAs within a cell. For example, we found that the frequency of out-of-frame translation can vary among different mRNA molecules derived from the same (reporter) gene²¹.

ACKNOWLEDGEMENTS

We would like to thank the members of the Tanenbaum lab for helpful discussions. This work was financially supported by an ERC starting grant (ERC-STG 677936-RNAREG), two grants from the Netherlands Organization for Scientific Research (NWO) (ALWOP.290 and NWO/016.VIDI.189.005), and by the Howard Hughes Medical Institute through an International Research Scholar grant to M.E.T. (HHMI/IRS 55008747). All authors are supported by the Oncode Institute that is partly funded by the Dutch Cancer Society (KWF). T.A.H. was supported by a PhD fellowship from the Boehringer Ingelheim Fonds.

AUTHOR CONTRIBUTIONS

D.K., T.A.H. and S.S. made the figures in the paper. D.K., T.A.H. and M.E.T. wrote the manuscript with input from S.S., B.V. and S.B.

COMPETING INTERESTS

The authors declare no competing interests

DATA AVAILABILITY STATEMENT

Example datasets for analysis are provided in Supplementary Data 1 and 2 and can be found online at <https://doi.org/10.1038/s41596-019-0284-x>

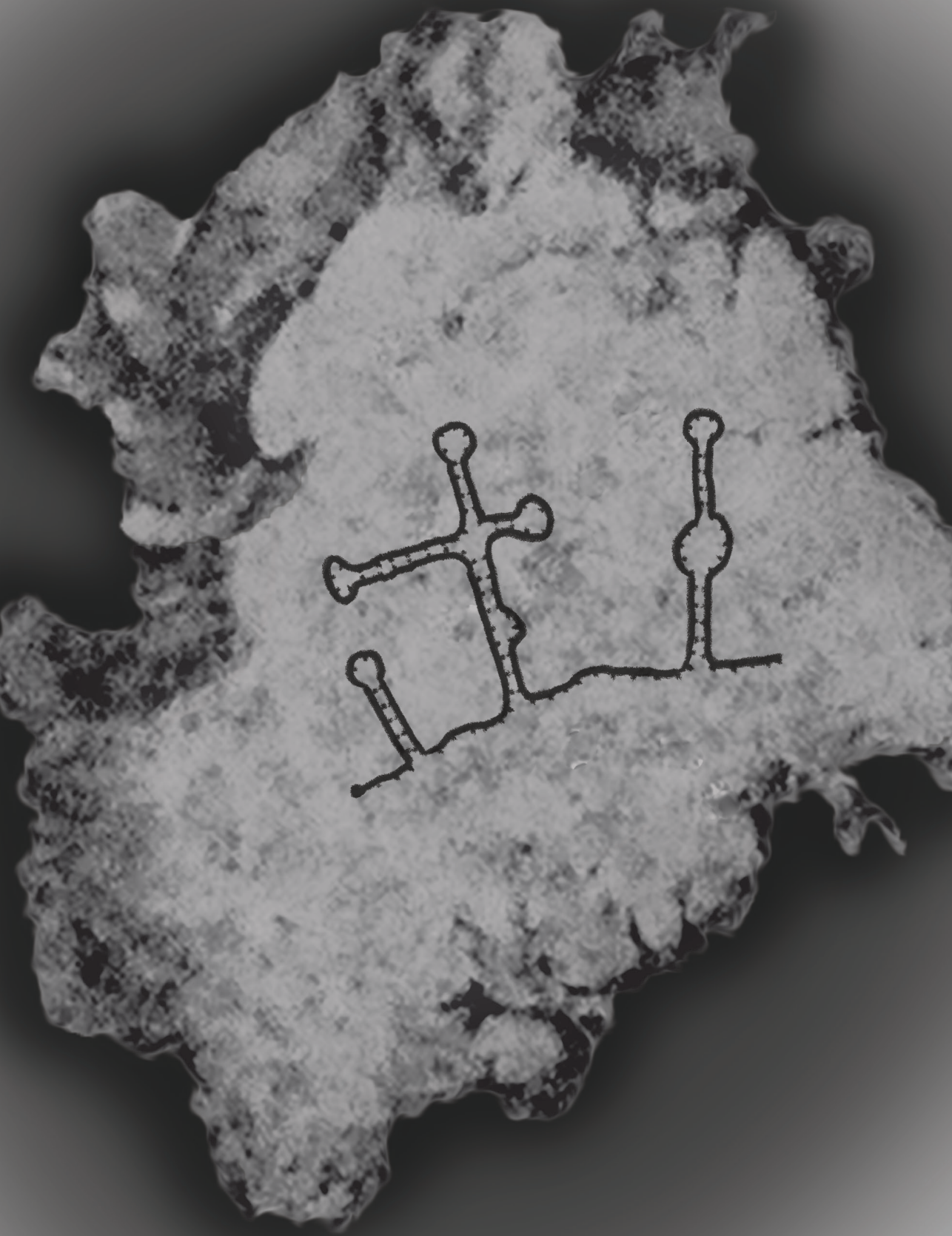
CODE AVAILABILITY STATEMENT

Analysis software (RiboFitter and TransTrack) is included in Supplementary Data 1 and 2, and also freely available through GitHub (github.com/TanenbaumLab/RiboFitter and github.com/TanenbaumLab/TransTrack, respectively). The code in this manuscript has been peer-reviewed.

REFERENCES

- 1 Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635-646, doi:10.1016/j.cell.2014.09.039 (2014).
- 2 Yan, X., Hoek, T. A., Vale, R. D. & Tanenbaum, M. E. Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* 165, 976-989, doi:10.1016/j.cell.2016.04.034 (2016).
- 3 Boersma, S. *et al.* Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. *Cell* 178, 458-472 e419, doi:10.1016/j.cell.2019.05.001 (2019).
- 4 Hoek, T. A. *et al.* Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Mol Cell*, doi:10.1016/j.molcel.2019.05.008 (2019).
- 5 Tanenbaum, M. E., Stern-Ginossar, N., Weissman, J. S. & Vale, R. D. Regulation of mRNA translation during mitosis. *eLife* 4, doi:10.7554/eLife.07957 (2015).
- 6 Trcek, T., Larson, D. R., Moldon, A., Query, C. C. & Singer, R. H. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell* 147, 1484-1497, doi:10.1016/j.cell.2011.11.051 (2011).
- 7 Ephrussi, A. & Lehmann, R. Induction of germ cell formation by oskar. *Nature* 358, 387-392, doi:10.1038/358387a0 (1992).
- 8 Lasko, P. mRNA localization and translational control in Drosophila oogenesis. *Cold Spring Harbor perspectives in biology* 4, doi:10.1101/cshperspect.a012294 (2012).
- 9 Holt, C. E. & Schuman, E. M. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* 80, 648-657, doi:10.1016/j.neuron.2013.10.036 (2013).
- 10 Shoemaker, C. J. & Green, R. Translation drives mRNA quality control. *Nature structural & molecular biology* 19, 594-601, doi:10.1038/nsmb.2301 (2012).
- 11 Schwanhauser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* 473, 337-342, doi:10.1038/nature10098 (2011).
- 12 Pelechano, V., Wei, W. & Steinmetz, L. M. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell* 161, 1400-1412, doi:10.1016/j.cell.2015.05.008 (2015).
- 13 Hu, W., Sweet, T. J., Chamnongpol, S., Baker, K. E. & Collier, J. Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature* 461, 225-229, doi:10.1038/nature08265 (2009).
- 14 He, F. & Jacobson, A. Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. *Annual review of genetics* 49, 339-366, doi:10.1146/annurev-genet-112414-054639 (2015).
- 15 Ruijtenberg, S., Hoek, T. A., Yan, X. & Tanenbaum, M. E. Imaging Translation Dynamics of Single mRNA Molecules in Live Cells. *Methods in molecular biology* 1649, 385-404, doi:10.1007/978-1-4939-7213-5_26 (2018).
- 16 Morisaki, T. *et al.* Real-time quantification of single RNA translation dynamics in living cells. *Science* 352, 1425-1429, doi:10.1126/science.aaf0899 (2016).
- 17 Wang, C., Han, B., Zhou, R. & Zhuang, X. Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell* 165, 990-1001, doi:10.1016/j.cell.2016.04.040 (2016).
- 18 Wu, B., Eliscovich, C., Yoon, Y. J. & Singer, R. H. Translation dynamics of single mRNAs in live cells

- and neurons. *Science* 352, 1430-1435, doi:10.1126/science.aaf1084 (2016).
- 19 Pichon, X. *et al.* Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *The Journal of cell biology* 214, 769-781, doi:10.1083/jcb.201605024 (2016).
 - 20 Chao, J. A., Patskovsky, Y., Almo, S. C. & Singer, R. H. Structural basis for the coevolution of a viral RNA-protein complex. *Nature structural & molecular biology* 15, 103-105, doi:10.1038/nsmb1327 (2008).
 - 21 Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E. & Baranov, P. V. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic acids research* 44, 7007-7078, doi:10.1093/nar/gkw530 (2016).
 - 22 Calvo, S. E., Pagliarini, D. J. & Mootha, V. K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proceedings of the National Academy of Sciences of the United States of America* 106, 7507-7512, doi:10.1073/pnas.0810916106 (2009).
 - 23 Lyon, K., Aguilera, L. U., Morisaki, T., Munsky, B. & Stasevich, T. J. Live-Cell Single RNA Imaging Reveals Bursts of Translational Frameshifting. *Mol Cell* 75, 172-183 e179, doi:10.1016/j.molcel.2019.05.002 (2019).
 - 24 Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2, 437-445 (1998).
 - 25 Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer control of microscopes using microManager. *Current protocols in molecular biology* Chapter 14, Unit14 20, doi:10.1002/0471142727.mb1420s92 (2010).
 - 26 Grimm, J. B. *et al.* A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nature methods* 12, 244-250, 243 p following 250, doi:10.1038/nmeth.3256 (2015).



Chapter 7

An Optimized Procedure for High-Throughput RT-qPCR-Based COVID-19 Diagnostics

Tim A. Hoek*, Peter J. Krijger*, Sanne Boersma, Mark Pieterse, Ive Logister, Martijn Bosch, Wouter L. de Laat, Marvin E. Tanenbaum

** These authors contributed equally to this work*

ABSTRACT

An important way to contain the outbreak of the current COVID-19 pandemic is through large-scale diagnostic testing. Diagnostic testing is primarily done through the RT-qPCR test, but upscaling of this test has been hampered by its laborious nature and by supply chain issues. Here, we describe a new procedure for high-throughput COVID-19 diagnostics using the gold standard RT-qPCR test. This procedure will use similar molecular biology as the current testing procedures, but throughput will be greatly increased by collecting swabs in tubes that are directly compatible with automated processing and by using a custom-designed robot that can handle >10,000 samples per day. We envision that this procedure will be an efficient way of upscaling COVID-19 diagnostic testing.

INTRODUCTION

The emergence of the novel coronavirus SARS-CoV-2 led to the outbreak of a global COVID-19 pandemic. As of August 2020, over 20 million people have been infected, leading to over 750,000 deaths (WHO, 2020). In addition, the virus has had massive social and economic impact due to lockdowns in many countries. Since no vaccine or adequate treatment against COVID-19 is available yet, the best method for controlling the pandemic that is currently available is preventing the spread of the virus. An important way to prevent spreading is through widespread testing in combination with contact tracing of infected individuals (Taipale et al., 2020). Frequent testing enables early detection of infections, thereby preventing the infected person from coming in contact with other people, while contact tracing ensures that people that have been in contact with infected individuals quarantine themselves and therefore do not spread the virus. In addition, frequent testing allows individuals that have COVID-19-like symptoms but are not infected with SARS-CoV-2 to continue working, and testing is thus very important for restarting the economy.

Although testing procedures can vary between countries or testing locations, most COVID-19 testing procedures follow roughly the same steps (Esbin et al., 2020). First, a nasopharyngeal or oropharyngeal swab is taken, which is broken at a score line and collected in a 15-ml tube containing a few ml of viral transport medium (VTM). The tube is closed and collected in a sealed biohazard bag and the bag is manually labelled with the patient's data. Sealed bags are sent to the lab, where they are manually cleaned, opened, and part of the sample is transferred to a multi-well plate containing lysis buffer. The RNA is extracted from the samples, and the presence or absence of the virus is determined using an RT-qPCR test. In the RT-qPCR test, the RNA is reverse transcribed into cDNA and the amount of viral RNA is determined by the number of rounds of PCR amplification that are required to reach a threshold value. The RT-qPCR test is highly specific, sensitive and robust for detection of SARS-CoV2, and is considered the golden standard for COVID-19 testing (Corman et al., 2020). However, the procedure is labour intensive as many steps, such as the labelling of samples or transfer of samples to multi-well plates, are done manually. It also requires expertise and specialised equipment, and scaling up testing capacity is therefore not trivial. Many countries lacked sufficient testing capacity during the early stages of the COVID-19 pandemic and still have issues with upscaling the current RT-qPCR test. Therefore, a scalable, high-throughput procedure for COVID-19 testing would be highly desirable. There are several ways in which the testing capacity could be increased. RNA extraction is a major bottleneck in the current procedure, therefore one way to increase the testing capacity is by performing the RT-qPCR directly on lysed samples without doing an RNA extraction step. Multiple studies have shown that SARS-CoV-2 RNA can be detected in lysed oro- or nasopharyngeal swabs or saliva without performing RNA extraction

(Ranoa et al., 2020; Smyrlaki et al., 2020; Vogels et al., 2020). However, lysed samples can contain a variety of contaminants that are normally removed by RNA extraction that could inhibit the RT-qPCR reaction. In addition, RNA extraction increases the RNA concentration of a sample by decreasing the volume in which it is dissolved, and skipping the extraction step thus reduces the sensitivity of the test. Testing capacity could also be improved by using diagnostic tests other than the RT-qPCR test, and several COVID-19 diagnostic tests are currently being developed. Sequencing based approaches can detect SARS-CoV-2 at very low copy numbers, but are slow, labour intensive, and expensive and are therefore not an ideal method for increasing testing capacity. Loop-mediated isothermal amplification coupled with reverse transcription (RT-LAMP) assays use a strand-displacing polymerase to amplify the genetic material of viruses (Buck et al., 2020; Dao Thi et al., 2020). Amplification happens at a single temperature and therefore does not require expensive thermocycling equipment. Although this assay is cheaper and faster than the qPCR assay, it has a higher limit of detection (more RNA copies need to be present for a positive test) and suffers from a lower sensitivity and specificity (Buck et al., 2020; Dao Thi et al., 2020). Other methods combine RT-LAMP with a new readout method, such as sequencing (LAMP-seq (Schmid-Burgk et al., 2020)) or CRISPr/Cas12 based detection of the amplicon (DETECTr (Broughton et al., 2020)). Although these methods might prove valuable, it remains to be determined whether their sensitivity matches that of the RT-qPCR assay in clinical setting and whether they can also be scaled up to high throughput. Finally, serological tests determine whether antigens against SARS-CoV-2 are present in blood and have been used to gain insight in which fraction of a population has been infected with SARS-CoV-2. However, the immune response required for these serological tests typically happens >5 days after the initial infection (Whitman et al., 2020; Woelfel et al., 2020), and this test is therefore not suitable to detect early infections. A third way to increase the testing capacity is by increasing the throughput of the current testing procedure. As many steps of the current testing procedure are done manually, the procedure could be improved by finding ways to circumvent manual handling steps, and instead doing these steps simultaneously for many samples in an automated way. A benefit of this approach is that the molecular biology underlying the test would be the same as for the current testing procedure, which facilitates implementation of a new procedure.

Here, we propose a new procedure for RT-qPCR based COVID-19 testing that includes all steps from swab collection to the result of the RT-qPCR test. The main improvement of this procedure is the collection of swabs in barcoded tubes that are directly compatible with automated sample processing. The tubes will be collected in 96-tube racks that can be further processed with minimal manual handling steps by a custom designed Tecan Fluent robot that is able to extract RNA and prepare RT-qPCR plates for >10.000 samples per day. As the full procedure has not been fully validated at

the moment of writing this thesis, this chapter will first describe the general procedure, and then focus mainly on the steps of the procedure that I optimized.

PROCEDURE DESIGN

We designed a procedure that allows automated processing and screening of oro- and nasopharyngeal swabs for the presence SARS-CoV-2 RNA (Figure 1). The procedure focusses on increasing the throughput and decreasing the handling time in order to get test results back to the patient within 12-24 hours while maintaining high selectivity and specificity. The procedure starts with swabbing of a patient using standard swabs that are currently used. The swab head is then cut with plyers and collected in a barcoded collection tube that contains a lysis buffer that inactivates the virus and preserves RNA. After cutting, plyers are cleaned in a 1% Virkon (a disinfectant that inactivates SARS-CoV-2 (Lanxess, 2020)) bath to prevent contamination between samples via the cutting process. The collection tube containing the swab head is scanned by a barcode scanner and the barcode is coupled to a unique identifier of the tested person (e.g. an email address or mobile phone application). Collection tubes are collected at the test site in barcoded 96-tube racks and send to the lab. At the lab, tube racks are first heat-inactivated to inactivate any virus on the outside or on the screw thread of the tube, centrifuged briefly, and then placed in a Tecan Fluent robot that is custom-designed for RNA extraction and RT-qPCR plate preparation. In this robot, collection tubes are automatically decapped and transferred to 384-well plates (the content of four 96-tube racks is transferred to one 384-well plate). The RNA is extracted by the robot through addition of magnetic beads followed by one isopropanol, two 70% ethanol washing steps, and elution of the RNA in water. Extracted RNA is added to plates containing an RT-qPCR mix and primers targeting the SARS-CoV-2 E-gene, and the plates are sealed by a plate sealer. The sealed plates are stored in the robot, from which they are transferred to thermocyclers that run the RT-qPCR test. Finally, RT-qPCR data automatically analysed to determine whether a test is positive or negative, and the tested person is notified of the result.

Improvements in testing throughput

Multiple aspects of this procedure, such as the use of barcoded tubes and 96-tube racks, the collection of swabs in lysis buffer, RNA extraction and RT-qPCR plate preparation by a custom-designed robot, and automated data analysis, contribute to the high throughput of this procedure. The biggest improvement in throughput obtained by collection of swabs in barcoded collection tubes that are collected in 96-tube racks and are directly compatible with robotic handling. These collection tubes do not need to be manually opened and samples do not need to be manually transferred to multi-

well plates as this is done in an automated way for 96 tubes simultaneously. In addition, because the collection tubes are barcoded they do not need to be labelled manually, but instead can quickly be scanned by a barcode scanner. Importantly, the barcode can also

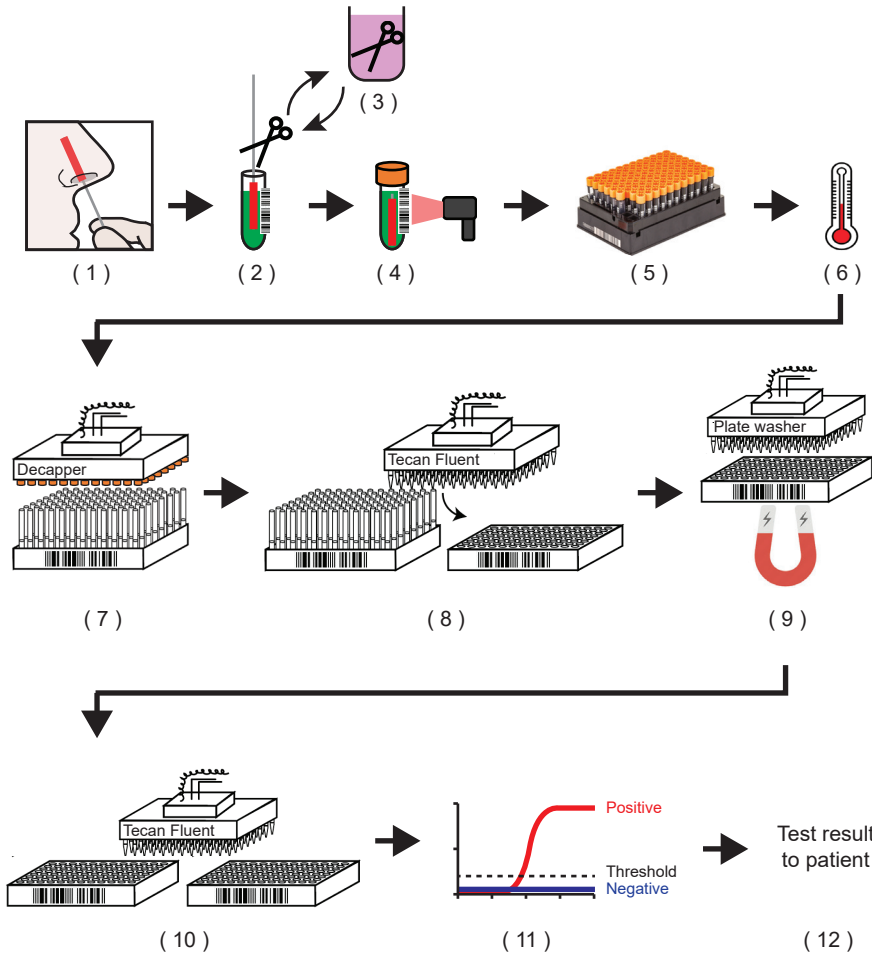


Figure 1. A new procedure for high-throughput COVID-19 diagnostic testing. Schematic overview of our new workflow for high-throughput COVID-19 diagnostic testing. 1) An oro- or nasopharyngeal swab is taken and 2) collected in a barcoded collection tube filled with lysis buffer by cutting the swab with pliers just above the swab head. 3) Pliers are disinfected in a 1% Virkon solution. 4) Tubes are capped and the barcode is scanned to connect the sample to the patient. 5) Tubes are collected in barcoded 96-tube racks and sent to the lab. 6) Racks are then heat-inactivated, and placed in a custom-designed Tecan Fluent pipetting robot (step 7-10). In this robot, 7) tubes are automatically decapped and 8) samples are transferred to 384-wells plates. 9) RNA is extracted using magnetic beads by two HydroSpeed plate washers, and 10) extracted RNA is added to RT-qPCR plates. 11) Plates are then moved to thermocyclers where the RT-qPCR is run, and 12) results of the test are sent to the patient.

be scanned from the bottom of a tube when the tube is placed in a 96-tube rack, so tubes can still easily be identified after swab collection. Because these collection tubes are small (~5 cm), swabs cannot be collected in the normal way by breaking the swab at the score line, as the broken swab would be larger than the collection tube. Therefore, swabs need to be cut directly above the swab head in order to fit in the collection tube. A second aspect of this procedure that increases throughput is the collection of swabs in lysis buffer instead of in VTM. VTM does not inactivate the virus, and samples in VTM therefore first need to be transferred to lysis buffer before they can be processed further. By collecting the samples directly in lysis buffer, we decrease the number of handling steps required and thereby speed up the procedure.

Another major improvement in the procedure is the use of a custom-designed Tecan Fluent pipetting robot that can extract RNA and prepare RT-qPCR plates for >10,000 samples per day. Although RNA extraction and RT-qPCR plate preparation in current testing procedures are also done in an automated way, other platforms are designed for flexibility rather than high-throughput. Because this robot is specifically designed for a single process, throughput can be maximized. This is achieved by designing the steps in a 'conveyor belt'-like fashion, in which multiple plates are being processed in the robot at the same time. Although the processing time of each 96-tube rack will be more than an hour, sample processing of the next 96-tube rack can start before handling of the previous 96-tube rack is completed. Although this robot has been designed and ordered and the throughput has been tested in computer simulations, it should be noted that the robot has not been delivered at the time of writing this thesis, and therefore will need to be optimized to ensure the expected throughput is achieved. Full details of the robot will become available upon publication.

A final aspect of the procedure that could increase the throughput is automated interpretation of the RT-qPCR test and notifying tested individuals of the test results. This would reduce the amount of labour required for data processing as tests do not need to be analysed by trained individuals and results are transferred automatically. However, as we have not implemented automatic data processing yet, and the procedure also works without automatic data processing, this thesis chapter will not further discuss automated data processing, but instead focus on optimization of swab collection, the molecular biology of the testing procedure, and automation steps.

RESULTS

Swab collection in barcoded collection tubes

In order for barcoded collection tubes to be compatible with automated sample transfer, a robot needs to be able to aspirate sample from the tubes without the pipet tip being physically blocked by the swab in the collection tube, and the swab should therefore

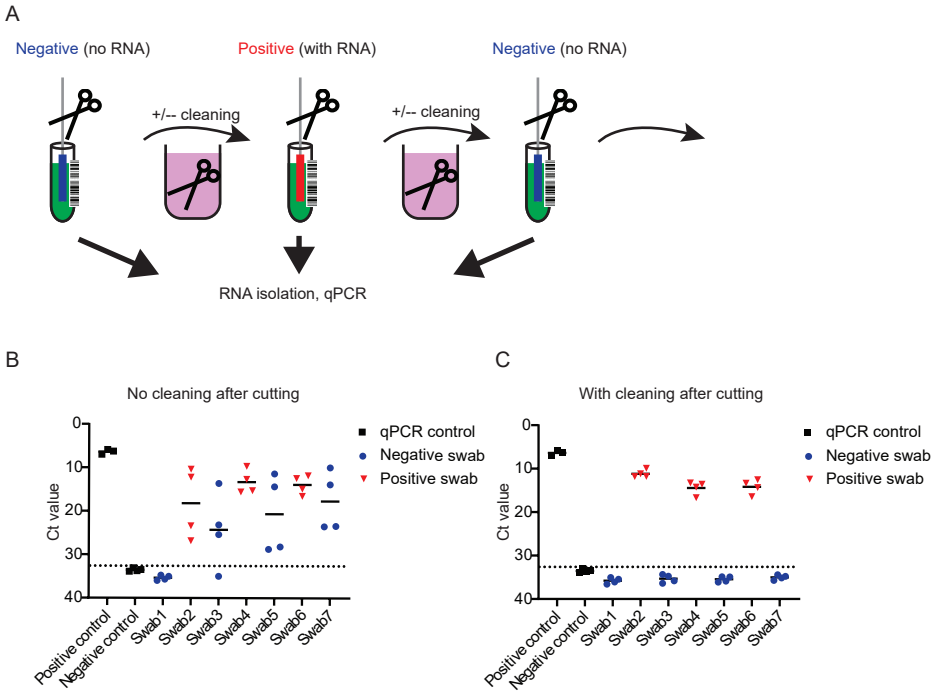


Figure 2. Swab collection without contamination. A) Schematic overview of the experimental setup. B-C) A small drop of either water (negative swabs, blue dots) or concentrated luciferase RNA in water (positive swabs, red triangles) was placed on the swab at the cutting location. Seven alternatingly negative and positive swabs were cut by pliers, and pliers were either re-used without cleaning (B) or cleaned in a 1% Virkon solution after cutting (C). After cutting, RNA was dissolved from swabs by placing the swabs in water at 55°C for 10 minutes. Dissolved RNA was directly used for cDNA synthesis using Tetro RT kit, and the abundance of luciferase RNA was determined by qPCR using IQ-SYBR green supermix. Positive and negative controls (black squares) consisted of luciferase RNA and water, respectively, that were directly used for cDNA synthesis and qPCR. $n = 2$ experiments, $n = 4$ pliers per condition.

completely be submerged in lysis buffer. The score line at which the swab is normally broken is much higher than the size of the tube, so we decided to use an approach in which the swab is cut directly above the swab head. As the cutting happens close to swab head, which could contain active virus, we wondered whether contamination between samples could occur by re-using pliers. To test this, we pipetted 5 μ l water (‘negative’) or of concentrated luciferase RNA (‘positive’) on swabs, alternatingly cut negative and positive swabs with the same pliers, and determined the amount of luciferase RNA on each swab by qPCR (figure 2A). Of note, luciferase RNA was used as we did not have access to intact SARS-CoV-2 particles and luciferase RNA was available in large quantities. Pliers were either not cleaned after cutting or cleaned by first placing them

in a 1% Virkon solution for 4 minutes, then in water for 1 minute, and finally by wiping them with a clean tissue. When no cleaning was done, the first negative swab was tested as negative in the RT-qPCR test (Ct value > 32), but subsequent positive and negative swabs were tested as positive (Ct value < 32.0, Figure 2B). This indicates that RNA was present in negative samples and thus that contamination can happen during the cutting process. However, cleaning of the pliers abolished the observed signal in negative swabs (Figure 2C), indicating that with the proposed cleaning method it is possible to cut directly above the swab head without contamination between swabs.

Heat inactivation of samples

After collecting the swabs, our procedure includes a heat inactivation step that inactivates virus on the outside and screw thread of the collection tubes, and it is important that heat inactivation does not reduce the sensitivity of the assay. To test this, we dissolved SARS-CoV-2 RNA in water or in a guanidine-thiocyanate containing lysis buffer (GITC buffer) and incubated the samples at 4°C, at 60°C for 30 minutes, or at 95°C for 5 minutes. We then spiked in control RNA to correct for RNA extraction efficiency,

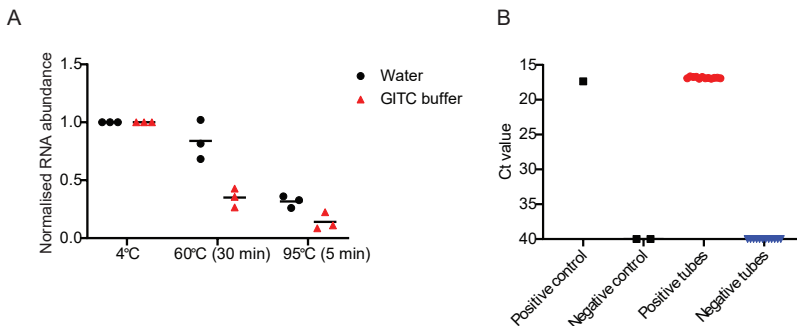


Figure 3. Heat inactivation and decapping of collection tubes. A) SARS-CoV-2 RNA in water (black circles) or in GITC buffer (red triangles) was incubated at indicated temperature for indicated duration. After heat incubation, luciferase RNA was spiked in as a control for RNA extraction, RNA was extracted using RNeasy mini spin columns and the abundance of the SARS-CoV-2 E gene and luciferase was determined by RT-qPCR. E gene abundance was normalised to luciferase abundance to control for RNA extraction, and all values were then normalised to the no heat treatment control. Each dot represents the average value of 2 RT-qPCR replicates in an independent experiment. $n = 3$ experiments. B) SARS-CoV-2 RNA in GITC buffer (positive tubes) or GITC buffer (negative tubes) was placed in barcoded collection tubes and positive and negative tubes were placed alternately in a 96-tube collection rack. Tubes were decapped and directly recapped by a FluidX decapper, after which RNA was extracted using hydrophobic Sera-Mag SpeedBeads and SARS-CoV-2 E gene abundance was determined by RT-qPCR. SARS-CoV-2 RNA was used as a positive control, water was used as a negative control. $n = 1$ experiment, 12 tubes per condition.

extracted the RNA and determined the abundance of the E gene by RT-qPCR. Heat treatment led to a clear reduction in SARS-CoV-2 RNA abundance at both 60°C and 95°, and RNA in GITC buffer was slightly more sensitive to heat treatment than RNA in water (Figure 3). Although this data suggest that a heat inactivation step reduces the sensitivity of the assay, it is possible that RNA degradation can be prevented by shorter incubation, lower temperature, or different buffer conditions, as has been suggested by others (Wang et al., 2020). Further testing should be done to determine the optimal heat-inactivation without reducing assay sensitivity.

Decapping of the collection tubes

The next step in the procedure is automated decapping of 96 tubes. Although this is a fairly simple step, in which 96 caps are removed simultaneously by a 96-tube decapper, it is important that no contamination is introduced between tubes during opening of the tube. To test if contamination happens during decapping, we filled a total of 24 collection tubes with either GITC buffer alone ('negative') or with GITC buffer with spiked in SARS-CoV-2 RNA ('positive') and alternately placed positive and negative tubes in a 96-tube rack. We briefly centrifuged the 96-tube rack before decapping, as decapping without centrifugation resulted in foaming of the lysis buffer which caused it to spill out of the tube. We then opened and closed the tubes using the automated decapper and determined SARS-CoV-2 RNA abundance by RT-qPCR. SARS-CoV-2 RNA could readily be detected in all positive tubes (Ct values = ~17), but was not amplified in any of the negative tubes (Ct value > 40, Figure 3B), indicating that decapping does not introduce contamination.

RNA extraction and cell lysis

After removing the caps from the collection tubes, samples in our procedure are transferred to 384-well plates and RNA is extracted using magnetic beads. Several types of magnetic beads are available for RNA extraction, and in order to determine which bead to use we tested RNA extraction efficiency using five different magnetic beads (SeraSil-Mag700, hydrophillic Sera-Mag SpeedBead, hydrophobic Sera-Mag SpeedBead, Xpert 300, and Xpert 600). Beads were resuspended in either poly-ethylene glycol (PEG) or isopropanol. Xpert 300 and SeraSil-Mag700 beads resuspended poorly in PEG and were therefore only resuspended in isopropanol. Beads were added to SARS-CoV-2 RNA in GITC buffer, RNA extraction was performed with one isopropanol wash step and two 70% ethanol wash steps, and the abundance of the SARS-CoV-2 E gene was determined by RT-qPCR. Most efficient RNA extraction was obtained when using hydrophobic SeraMag SpeedBeads or Xpert 300 beads resuspended in isopropanol (Figure 4A). We decided to use hydrophobic SeraMag SpeedBeads as they performed slightly better than Xpert300 beads and are from a well-established supplier

of magnetic beads, reducing the likelihood of supply chain issues when using large quantities.

Next, we wondered whether the efficiency of RNA extraction using hydrophobic SeraMag SpeedBeads beads depends on the lysis buffer that is used to lyse the virus. We tested three different lysis buffers: 1) GITC buffer, 2) GITC buffer with 1% sarkosyl, a detergent that could increase lysis efficiency, and 3) DNA/RNA shield, a commercial DNA/RNA transport medium that inactivates viruses while preserving RNA and is used as a collection buffer in some COVID-19 swabbing procedures. As DNA/RNA shield contains GITC, it might already lyse sufficiently for RNA extraction, which

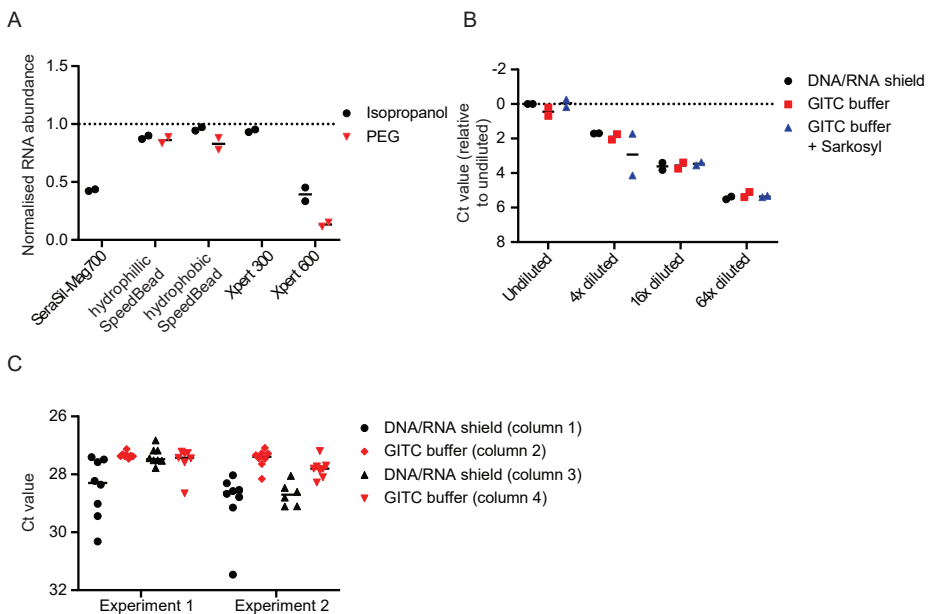


Figure 4. Efficiency of cell lysis and RNA extraction for different lysis buffers and magnetic beads. A) SARS-CoV-2 RNA in GITC buffer was extracted using indicated magnetic beads resuspended in isopropanol (black circles) or poly-ethylene glycol (PEG, red triangles). SARS-CoV-2 E gene abundance was determined by RT-qPCR and normalised to a control sample in which no RNA extraction was performed. B) Human U2Os cells were trypsinized, resuspended in PBS, and diluted in PBS with 4-fold dilutions. All samples were first mixed with DNA/RNA shield, after which indicated buffers were added. RNA was extracted using hydrophobic Sera-Mag SpeedBeads and the abundance of RNaseP was determined by RT-qPCR. Ct values were normalised to the Ct value of the undiluted DNA/RNA shield condition to control for differences in cell number between experiments. C) SARS-CoV-2 RNA was spiked in DNA/RNA shield, after which indicated buffers were added. Samples were then added to different wells of indicated columns of a 384-wells plate and RNA extraction was performed using hydrophobic Sera-Mag SpeedBeads on a HydroSpeed plate washer. SARS-CoV-2 E gene abundance was determined by RT-qPCR. (A-C) RT-qPCR was performed using TaqPath 1-Step RT-qPCR Master Mix, n=2 experiments.

would make it an ideal lysis buffer for swab collection. As we did not have access to intact virus particles, we resuspended a 4x dilution range of human U2OS cells in DNA/RNA shield and added either GITC buffer, GITC buffer with sarkosyl, or DNA/RNA shield. We then extracted RNA and measured the abundance of RNaseP by RT-qPCR as a readout for both lysis and RNA extraction efficiency. RNaseP abundance was comparable between the three buffers (Figure 4B), indicating that lysis and RNA extraction were equally efficient in RNA shield and lysis buffer. This data suggests that DNA/RNA shield could be used as both a storage and lysis buffer for swab collection. Since our RNA extraction protocol worked properly with manual pipetting, we wondered whether RNA extraction also worked efficiently when it is done in automated fashion. Therefore, we tested protocol using automated RNA extraction on a HydroSpeed plate washer. We dissolved SARS-CoV-2 RNA in either DNA/RNA shield or in DNA/RNA shield with added GITC buffer and divided each sample over sixteen different wells in two columns of a 384-wells plate to determine if RNA extraction worked properly in both buffers and if RNA extraction was effected by the position of the well in the 384-wells plate. Samples dissolved in DNA/RNA shield + GITC buffer showed comparable Ct values in all columns, indicating that RNA extraction on the plate washer was reproducible (Figure 4C). Surprisingly, samples dissolved in DNA/RNA shield had higher Ct values and were more variable than samples in DNA/RNA shield + GITC buffer, suggesting that RNA extraction on the plate washer is less efficient for samples in which no GITC buffer was added. This is likely a consequence of improper washing, as no difference between DNA/RNA shield and GITC buffer was observed with manual RNA extraction (Figure 4B). Poor washing in DNA/RNA shield might be a consequence of the higher viscosity of DNA/RNA shield compared to GITC buffer, as this decreases the speed by which beads migrate to the magnet and could therefore affect washing efficiency. Further optimization of the RNA extraction

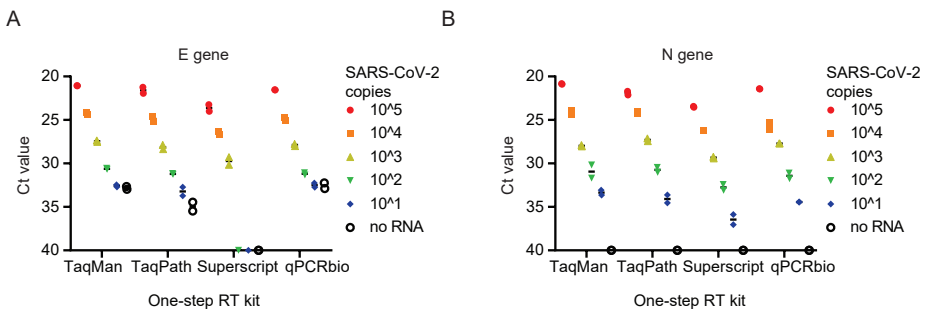


Figure 5. Comparison of commercially available one-step RT-qPCR kits. A, B) A 10-fold dilution range of SARS-CoV-2 RNA in water was made and indicated amount of copies were spiked into RT-qPCR reactions. The abundance of the SARS-CoV-2 E gene (A) and N gene (B) were determined using indicated one step RT-qPCR kits. $n = 2$ experiments

procedure should be performed to ensure efficient and reproducible RNA extraction of all samples in 384-wells plates.

One-step RT-qPCR reaction mix

The final step of the procedure is the RT-qPCR reaction on the extracted RNA. The RT-qPCR reaction is done by mixing extracted RNA, one or multiple sets of TaqMan primers and probes targeting SARS-CoV-2 genes, and an RT-qPCR mastermix that contains the buffers, nucleotides, and enzymes required for the RT-qPCR reaction. To determine which conditions to use for the RT-qPCR reaction, we tested amplification efficiency for four different RT-qPCR mastermixes (TaqMan fast virus 1-step master mix, TaqPath 1-step RT-qPCR mastermix, SuperScript III platinum one-step qRT-PCR kit, and qPCRBIO probe 1-step virus detect no-ROX) and two different TaqMan primer-probe sets (CDC assay targeting N2 gene (CDC, 2020) and Charite/Berlin assay targeting the E gene (Corman et al., 2020)). We used a 10x dilution series of SARS-CoV-2 RNA as input material, total input amounts ranging from 10^5 to 10^1 copies of SARS-CoV-2 RNA per RT-qPCR reaction. TaqMan, TaqPath and qPCRBio kits efficiently amplified SARS-CoV-2 RNA in all dilutions, while the SuperScript kit amplified efficiently with primers targeting the N-gene but not with primers targeting the E-gene. These results suggests the TaqMan, TaqPath and qPCRBIO kits could all be used for the RT-qPCR reaction, and the choice which one to use will likely depend on pricing and availability.

DISCUSSION

A new procedure for COVID-19 diagnostic testing

In this chapter, we describe a new procedure for RT-qPCR based high-throughput SARS-CoV-2 testing. Throughput of testing is greatly increase through collection of swabs in barcoded collection tubes that are directly compatible with automated sample handling by a custom-designed pipetting robot. When implemented, this will allow processing of >10.000 samples per robot per day. Importantly, the procedure uses the current gold standard RT-qPCR test for COVID-19 detection and is therefore not dependent on unvalidated procedures or reagents. This procedure could play an important role in upscaling of COVID-19 diagnostic testing.

Further optimization of the procedure is required

We show that various steps of the protocol work efficiently and do not introduce contamination between samples. For example, we show that swabs can be collected and tubes can be automatically opened without contamination and we show how

different lysis buffers, magnetic beads, and RT-qPCR reagents can be used. However, it is important to note that these steps were tested individually rather than as part of the final procedure. It will be of key importance to test whether all steps work equally well when they are performed in automated high-throughput fashion by the robot. For example, RNA extraction in DNA/RNA shield worked properly when washing was done manually, but not when washing was done using a plate washer (Figure 4), demonstrating that the washing step needs to be optimized specifically for the plate washer. Similarly, other steps of the procedure, such as swab cutting, decapping, and RT-qPCR plate preparation, will need to be tested and optimized as part of the complete procedure rather than as an individual step.

In addition, the procedure will need to be validated extensively on clinical samples. We currently did most experiments with purified SARS-CoV-2 RNA. However, clinical swabs are highly variable in the amount of SARS-CoV-2 RNA they contain and in the contaminants that are present in the swab, both of which could influence the efficiency of detection. Therefore, the entire procedure should be run using a test panel of clinical samples that have been validated with current diagnostic testing procedures.

Flexibility of the procedure

A strong upside of this procedure is that it is modular, and therefore can be used flexibly in various ways. First, although the pipeline should be used with bead-based RNA extraction, it is not dependent on a specific type of magnetic bead or on reagents from a single supplier. If supply chain issues arise or different reagents result in a higher testing sensitivity, new reagents could easily be introduced in the pipeline. Second, the pipeline is not strictly dependent on RT-qPCR based testing. New diagnostic assays for COVID19 detection are rapidly being developed, and recent studies have shown that various LAMP-based methods, such as LAMP-seq (Schmid-Burgk et al., 2020) and DETECTr (Broughton et al., 2020), can accurately detect SARS-CoV-2 RNA. If these newly developed methods are shown to have a better sensitivity than the RT-qPCR test they could easily be implemented into the procedure instead of the current RT-qPCR test. Finally, we hypothesize that the pipeline will still be useful even when COVID-19 testing is not required anymore. By changing the primers used for RT-qPCR detection, the pipeline can be modified to detect other viruses or pathogens, which could prove useful in the event of future pandemics or when other large-scale population screening is desired. Thus, we envision that this procedure will be of great use for the current COVID-19 pandemic as well as for future applications.

MATERIAL & METHODS

Contamination during swabbing

Drops of 5 μ l water or in vitro transcribed luciferase RNA in water were pipetted on nose swabs (DSM) just above the swab head. Swabs were cut with pliers just above the swab head at the droplet. Pliers were either not cleaned between cuts, or cleaned by placing them in a 1% Virkon S solution for 4 minutes, then in water for 1 minute and finally wiping them with a tissue. Swab heads were collected upside down in 20 μ l water and incubated 10 minutes at 55°C to ensure all RNA dissolved from the swab. 12 μ l RNA dissolved RNA was used for a 20 μ l cDNA synthesis reaction, and 5 μ l cDNA was used for each qPCR reaction.

RNA extraction using magnetic beads

When cells were used as input, cells were lysed by first resuspending them in 28 μ l PBS and then mixing them with 28 μ l 2x DNA/RNA shield (Zymo). After five minutes incubation, an extra lysis step was performed by adding 56 μ l 1x DNA/RNA shield, 2x GITC buffer (4M guanine thiocyanate, 55 mM Tris-HCl pH 7.6, 25 mM EDTA, 3% Triton X-100) or 2x GITC buffer with added 1.33% sarkosyl. When purified RNA was used as input, 28 μ l RNA in water was mixed with 28 μ l 2x DNA/RNA shield or 28 μ l 2x GITC buffer. After this, 0.8x volume of magnetic beads (SeraSil-Mag 700 (Cytivia) Sera-Mag SpeedBead hydrophilic (Cytivia 4515), Sera-Mag SpeedBead hydrophobic (Cytivia 6515), Xpert magnetic beads 300, or Xpert magnetic beads 600 resuspended in isopropanol or poly-ethylene glycol were added to the lysed cells or RNA and incubated for 10 minutes. Manual RNA extractions were performed with two 70% ethanol washing steps (Figure 3A, 4A/B) or one isopropanol and two 70% ethanol washing steps (Figure 3B) on a DynaMag 2 magnet (ThermoFisher). Automated extractions (Figure 4C) were performed in 384-wells plates (Greiner 781201) on a HydroSpeed plate washer (Tecan) with a magnetic separation plate for 384 well microplates (VP scientific VP 771TN-G-5A) with two 70% ethanol washing steps. In all conditions, RNA was dissolved in 20 μ l and 5 μ l RNA was used for each RT-qPCR reaction.

RT-qPCR experiments

Rt-qPCR mixes (TaqMan™ Fast Virus 1-Step Master Mix (ThermoFisher), TaqPath™ 1-Step RT-qPCR Master Mix, CG (ThermoFisher), SuperScript™ III Platinum™ One-Step qRT-PCR Kit (ThermoFisher), qPCRBIO Probe 1-Step Virus Detect No-Rox (PCR biosystems)) were prepared according to manufacturer's guidelines. RT-qPCRs reactions targeting SARS-CoV-2 E gene (E Assay_First Line Screening, IDT), SARS-CoV-2 N2 gene and RNaseP (both from 2019-nCoV RUO Kit, IDT) were performed with TaqMan primers at 400 nM and probes at 200 nM. RT-qPCR reactions targeting

luciferase (i.e. loading controls in heat inactivation experiment) were performed with primers targeting luciferase (forward primer: gcgagctgctgaacagcatg, reverse primer: agccctggtagtcggctttg) at 500 nM and with addition of 1.25 μ M EvaGreen (Biotum) in the reaction mix. Contamination tests after swab cutting were done by first separately making cDNA using tetro reverse transcriptase (BioLine) and then performing qPCR reactions using IQ-SYBR Green Supermix (Bio-Rad) according to manufacturer's guidelines with primers targeting luciferase at 500 nM.

Thermocycling was done at a Bio-Rad CFX96 connect Real-Time PCR detection system. When multiple one-step RT master mixes were compared, thermocycling conditions were unified at 2 minutes at 25°C, 15 minutes at 50°C, 3 minutes at 95°C and finally 40 cycles of 15 seconds at 95°C and 30 seconds at 60°C. When only TaqPath™ 1-Step RT-qPCR Master Mix, CG was used, thermocycling was done for 2 minutes at 25°C, 15 minutes at 50°C, 2 minutes at 95°C and finally 40 cycles of 3 seconds at 95°C and 30 seconds at 60°C. When IQ-SYBR green was used, thermocycling conditions were 3 minutes at 95°C followed by 40 cycles of 15 seconds at 95°C and 30 seconds at 60°C, followed by a melting curve from 55°C-95°C with increments of 0.5°C. Results were analysed with Bio-rad CFX manager 3.1 software and Ct values were determined with an automated threshold. Relative RNA abundances were calculated using the delta-delta Ct method for heat inactivation experiments and delta Ct method for magnetic bead comparison experiments.

SARS-CoV-2 RNA

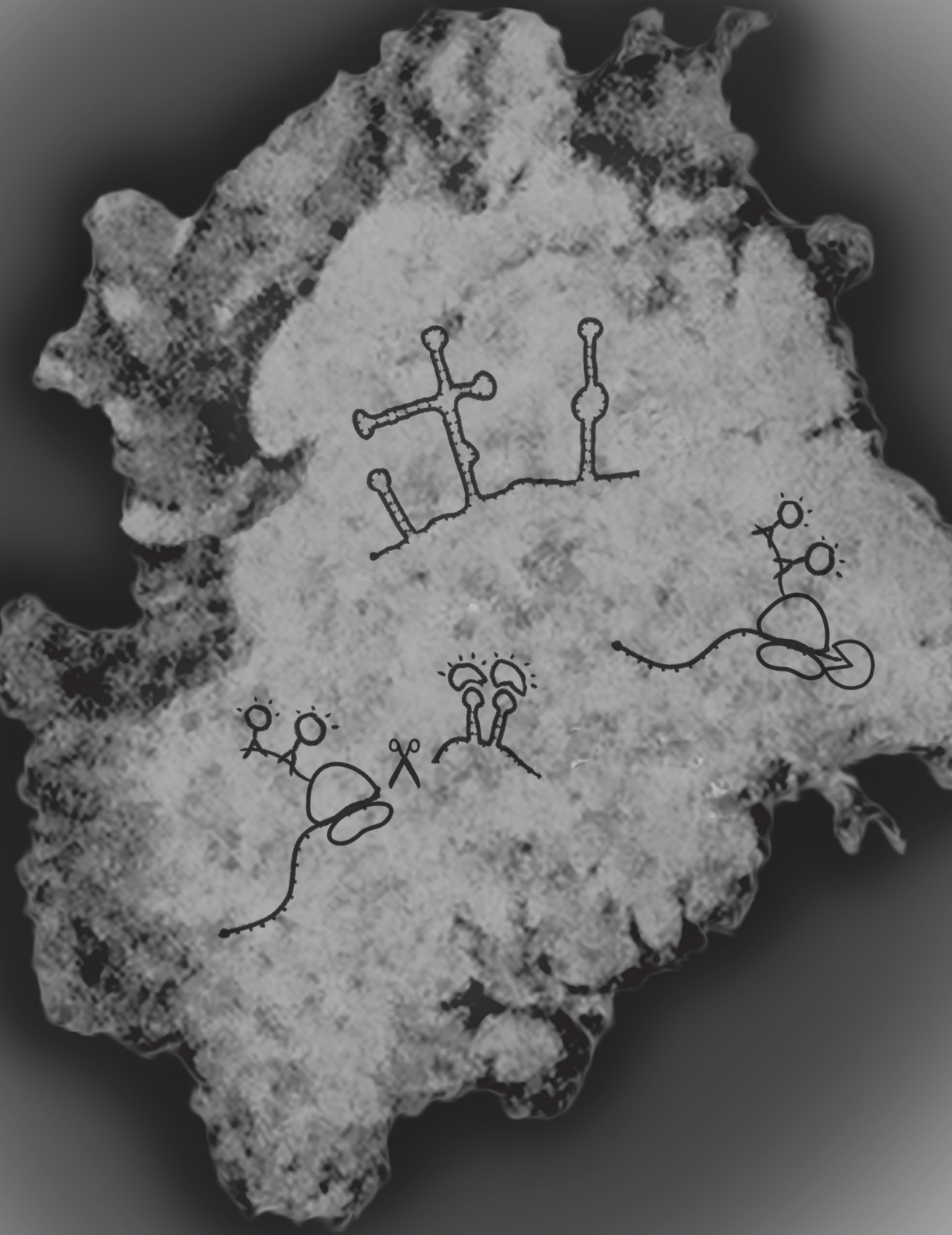
SARS-CoV-2 RNA was obtained by isolating total RNA using TRIzol extraction from African Green Monkey Vero E6 cells infected with SARS-CoV-2. Viral load was determined by RT-qPCR by comparing the extracted RNA with synthetic SARS-CoV-2 gene fragments with known concentration (Twist synthetic SARS-CoV-2 RNA control 2).

Custom-designed Tecan Fluent pipetting robot

Swabs are collected in barcoded collection tubes (FluidX 96-Format, 1.0ml External Thread, Next-Gen Jacket, Tri-Coded Tube, SopaChem), which are collected in 96-tube racks (FluidX rack 66-51020, SopaChem). The robot is a Tecan Fluent 1080 robot (Tecan) containing a FluidX IntelliXcap 96-Format automated sample tube capper/decapper (SopaChem), two HydroSpeed Plate washers (Tecan), as well as several other components required for moving plates, adding reagents, storing reagents, and storing plates during incubation times. Full design of the robot will become available upon publication.

REFERENCES

- Broughton, J.P., Deng, X., Yu, G., Fasching, C.L., Servellita, V., Singh, J., Miao, X., Streithorst, J.A., Granados, A., Sotomayor-Gonzalez, A., et al. (2020). CRISPR–Cas12-based detection of SARS-CoV-2. *Nat. Biotechnol.* *38*, 870–874.
- Buck, M.D., Poirier, E.Z., Cardoso, A., Frederico, B., Canton, J., Barrell, S., Beale, R., Byrne, R., Caidan, S., Crawford, M., et al. (2020). Standard operating procedures for SARS-CoV-2 detection by a clinical diagnostic RT-LAMP assay. MedRxiv doi: 10.1101/2020.06.29.20142430.
- CDC (2020). Real-Time RT-PCR Diagnostic Panel For Emergency Use Only. *Cdc Eua* *3*, 0.
- Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K.W., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., et al. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* *25*, 1–8.
- Dao Thi, V.L., Herbst, K., Boerner, K., Meurer, M., Kremer, L.P., Kirrmaier, D., Freistaedter, A., Papagiannidis, D., Galmozzi, C., Stanifer, M.L., et al. (2020). A colorimetric RT-LAMP assay and LAMP-sequencing for detecting SARS-CoV-2 RNA in clinical samples. *Sci. Transl. Med.* *12*, eabc7075.
- Esbin, M.N., Whitney, O.N., Chong, S., Maurer, A., Darzacq, X., and Tjian, R. (2020). Overcoming the bottleneck to widespread testing: a rapid review of nucleic acid testing approaches for COVID-19 detection. *RNA* *26*, 771–783.
- Lanxess (2020). New study: LANXESS disinfectant kills coronavirus SARS-CoV-2 in just 60 seconds.
- Ranoa, D.R.E., Holland, R.L., Alnaji, F.G., Green, K.J., Wang, L., Christopher, B., Burke, M.D., Fan, T.M., and Hergenrother, P.J. (2020). Saliva-Based Molecular Testing for SARS-CoV-2 that Bypasses RNA Extraction. *BioRxiv* doi: 10.1101/2020.04.22.056283.
- Schmid-Burgk, J.L., Li, D., Feldman, D., Slabicki, M., Borrajo, J., Strecker, J., Cleary, B., Regev, A., and Zhang, F. (2020). LAMP-Seq: Population-Scale COVID-19 Diagnostics Using a Compressed Barcode Space. *BioRxiv* doi: 10.1101/2020.04.06.025635.
- Smyrlaki, I., Ekman, M., Vondracek, M., Papanicolaou, N., Lentini, A., Aarum, J., Murdrasoli, S., Albert, J., Högberg, B., and Reinius, B. (2020). Massive and rapid COVID-19 testing is feasible by extraction-free SARS-CoV-2 RT-qPCR. *MedRxiv* doi: 10.1101/2020.04.17.20067348.
- Taipale, J., Romer, P., and Linnarsson, S. (2020). Population-scale testing can suppress the spread of COVID-19. *MedRxiv* doi: 10.1101/2020.04.27.20078329.
- Vogels, C.B., Brackney, D.E., Wang, J., Kalinich, C.C., Ott, I.M., Kudo, E., Lu, P., Venkataraman, A., Tokuyama, M., Moore, A.J., et al. (2020). SalivaDirect: Simple and sensitive molecular diagnostic test for SARS-CoV-2 surveillance One sentence summary. *MedRxiv* doi: 10.1101/2020.08.03.20167791.
- Wang, T.T., Lien, C.Z., Liu, S., and Selvaraj, P. (2020). Effective Heat Inactivation of SARS-CoV-2. *MedRxiv* doi: 10.1101/2020.04.29.20085498.
- Whitman, J.D., Hiatt, J., Mowery, C.T., Shy, B.R., Yu, R., Yamamoto, T.N., Rathore, U., Goldgof, G.M., Whitty, C., Woo, J.M., et al. (2020). Evaluation of SARS-CoV-2 serology assays reveals a range of test performance. *Nat. Biotechnol.* *38*, 1174–1183.
- WHO, W.H.O. (2020). Situation Report - 65 - Coronavirus disease 2019. *World Heal. Organ.* *2019*, 2633.
- Woelfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Mueller, M.A., Niemeyer, D., Vollmar, P., Rothe, C., Hoelscher, M., et al. (2020). Virological assessment of hospitalized cases of coronavirus disease 2019. *MedRxiv* doi: 10.1101/2020.03.05.20030502.



Chapter 8

Summarizing discussion

INTRODUCTION

In this thesis, we gained insight into the mechanisms of mRNA quality control through the development of several new methods that visualize translation and degradation of single mRNA molecules. In **chapter 2** I describe the development of SunTag translation imaging to visualize translation of individual mRNAs over time, and describe how to use this method in **chapter 3**. I discuss further development of SunTag imaging in **chapter 4**, where we combine it with visualization of mRNA degradation through nonsense-mediated mRNA decay. In addition, visualization of exonucleolytic degradation by XRN1 is discussed in **chapter 4** and by the exosome in **chapter 5**, and we describe some of these methods in more detail in **chapter 6**. Finally, we switch topic in **chapter 7**, where we try to help fight the COVID-19 pandemic by designing a new procedure for diagnostic testing.

Mechanistic insight gained from single molecule kinetics

A common theme of the research performed in this thesis is that we use methods that provide very precise kinetic measurements for individual mRNA molecules. This enables us to observe when individual mRNAs are translated, how efficiently they are translated, and when they are degraded through nonsense-mediated mRNA decay (NMD) or through exonucleases. Although knowing the exact kinetics of a process is interesting on itself, it becomes more exciting when the kinetics can be interpreted to determine the molecular mechanisms underlying the observed kinetics, and we do exactly this in this thesis on multiple occasions.

Different mechanisms of ribosome stalling

In **chapter 2**, we were able to observe how the GFP fluorescence intensity associated with translating mRNAs, changes over time. This revealed that mRNAs can dynamically switch between a translating and a non-translating state. In addition, by chemically stalling ribosomes at the start codon using harringtonine, we could observe how fast ribosomes translate the mRNA by analysing the decrease in GFP fluorescence intensity that occurs when ribosomes terminate translation. We showed that stalling could be induced by either treatment of the mRNA with the oxidizing agent 4NQO (Simms et al., 2014) or by introduction of the XBP1 stalling sequence in the mRNA (Yanagitani et al., 2011). Although the extent of stalling was comparable when assessing the average decrease in GFP fluorescence intensity of all mRNAs, observation of the decrease of GFP fluorescence of individual mRNAs revealed that the underlying mechanisms of stalling differs. 4NQO treatment resulted in a constant decrease in GFP fluorescence, indicating that all ribosomes stall to an equal extent. In contrast, introduction of the XBP1 stalling sequence in the mRNA resulted in intermingled

periods of translation and stalling, suggesting that only a subset of ribosomes stall on these mRNAs. It will be interesting to determine why such large variation exists between ribosomes translating XBP1 mRNAs. Stalling at the XBP1 stalling sequence is used to allow time for translocation of the mRNA to the ER (Yanagitani et al., 2011) and is caused by interactions of the XBP1 nascent chain with the ribosome exit tunnel (Shanmuganathan et al., 2019). Interestingly, the nascent peptide induced intermediate levels of translation arrest, as only 8 out of 20 amino acids of the stalling peptide were optimal for stalling. Intermediate stalling could explain why not all ribosomes stall on the mRNA. In addition, intermediate stalling also ensures that ribosomes do not stall permanently, and thereby allows translation to continue when the mRNA has been translocated to the ER.

Induction of NMD by terminating ribosomes is a stochastic process

In **chapter 4**, we combine observation of translation kinetics with observation of endonucleolytic cleavage through NMD. Translation is particularly relevant for NMD, as mRNA degradation by NMD is induced by ribosomes, and it has been suggested that NMD can only be induced during the first rounds of translation (Ishigaki et al., 2001; Maquat et al., 2010). As mRNAs in a cell will be in various stages of their life cycle, observing the first round of translation in bulk assays is very challenging. However, by visualizing translation of single mRNA molecules, we could quantify when cleavage occurs relative to the first round of translation for each individual mRNA. Our data shows that degradation is independent of the first round of translation. Instead, NMD can be induced by each ribosome terminating on an NMD substrate, but for efficient NMD substrates degradation does often coincide with the first round of translation. As each terminating ribosome has a certain probability of inducing NMD, we wondered what determines the chance of a ribosome to induce mRNA degradation. We found that the degradation of NMD-sensitive mRNAs is best described by a single exponential decay model, suggesting that there is a single rate-limiting step that determines whether an mRNA is degraded or not. A likely scenario is that the rate-limiting step is the formation of an interaction between the exon-junction complex (EJC) and the terminating ribosome. In this case, NMD would only be induced if termination happens when an EJC is close to the ribosome and therefore able to bind to the ribosome via the UPF proteins. In agreement with this, placing introns (i.e. EJCs) further away from the premature termination codon (PTC) reduced the NMD decay rate, likely because an EJC will be in proximity of the PTC a smaller percentage of the time when a longer nucleotide sequence is present in between the EJC and the PTC. This poses the question during which stages of translation the distance between the EJC and PTC is important for NMD efficiency. Translation termination consists of various steps (e.g. eRF1/3 recruitment, GTP hydrolysis, peptidyl hydrolysis), and it is likely that the connection between the EJC and PTC needs to happen during a specific stage of translation termi-

nation. Knowing the exact step during which NMD happens would provide valuable insight in the mechanism of NMD. One way to identify stages in which NMD is induced would be to increase the duration of the individual steps of termination, for example through mutation in eRF1/3 proteins or through addition of termination inhibitors. If the ribosome is stabilized in a conformation that is required for the rate-limiting step of NMD, more time would be available for that step to occur, which could increase the chance that a terminating ribosome induces NMD. Further experiments are needed to shed light on the rate limiting step of translation termination in NMD.

Processive exonucleolytic degradation by XRN1 and the exosome.

Using our methods, we can precisely follow mRNAs while they are processively degraded by exonucleases. In **chapter 4** we visualize degradation of 3' cleavage fragments by the exonuclease XRN1, while in **chapter 5** we observe exonucleolytic degradation of 5' cleavage fragments by the exosome. Because we can observe exonucleolytic degradation through a gradual decrease in fluorescence intensity, kinetics of degradation can be separated into the time required for recruitment of the enzyme and processive degradation of the mRNA after enzyme recruitment. From this data, we could conclude that the exosome can efficiently degrade mRNAs that are translated by multiple ribosomes. However, exosomal degradation is blocked by ribosomes that are stalled by translation inhibitors, but not by translating ribosomes, suggesting that translation might be required for ribosomes to be removed. We observed that XRN1 degrades the mRNA with high processivity, but occasionally dissociates from the mRNA. It is currently unclear why XRN1 dissociates from the mRNA. Recent studies in zebrafish suggest that XRN1-degradation intermediates play a role in regulating expression of paralogue genes (El-Brolosy et al., 2019). One possibility is that XRN1 has evolved to not be completely processive in order to allow regulation of gene expression by XRN1-degradation intermediates. Another possibility is that XRN1 processivity is affected by modifications of the mRNA. XRN1 has been shown to interact with YTHDC2, which binds to N6-methyladenosine (m6A) nucleotides. Although this specific interaction was shown to be involved in recruitment, not processivity, of XRN1, it does show how modifications of the mRNA can affect XRN1, and it is therefore possible that other modifications are responsible for dissociation of XRN1 from the mRNA.

Heterogeneity between mRNAs within a cell

As our new imaging-based methods allow us to study translation and degradation of individual mRNAs, it is also possible to measure heterogeneity between different mRNAs within a single cell. This is very useful in understanding how various factors contribute to observed phenotypes, and cannot be done with bulk detection methods. In this thesis, we discovered multiple examples of heterogeneity between mRNAs within a cell.

Heterogeneity in translation efficiency

Although we see in **chapter 2** that many mRNAs are efficiently translated for prolonged periods of time, we also detect a large population of mRNAs (~50%) that are not translated at all. This is not a consequence of changes in translation activity over time as mRNAs typically remained either translating or non-translating for the majority of the duration of a time-lapse experiment. A large fraction of mRNAs being untranslated was not a consequence of our specific mRNA, as other studies made similar observations (Pichon et al., 2016; Wu et al., 2016). However, the lack of translation activity is likely a consequence of expression of the reporter from a plasmid vector, as others observed that more mRNAs were translated when the SunTag array was knocked in on an endogenous locus (Pichon et al., 2016). Since untranslated mRNAs can be easily distinguished from translated mRNAs with our single-molecule imaging methods, they have limited impact on the conclusions that can be drawn from our experiments. However, the presence of a large fraction of untranslated mRNAs can still be important, as it suggests that plasmid-based expression can have a strong effect on observed phenotypes. In addition, a population of untranslated mRNAs can be difficult to detect using bulk methods, and could therefore have a strong impact on interpretation of experiments in which no distinction can be made between translated and untranslated mRNAs.

We also observed large heterogeneity in translational activity for mRNAs that contain a repressive 5'UTR originating from Emi1. The majority of mRNAs containing this 5'UTR were poorly translated, with one or at most a few ribosomes translating the mRNA at the same time. However, a small subset of mRNAs showed highly efficient translation for prolonged periods of time. Because of this, 2% of all reporter mRNAs were responsible for 50% of all protein produced from the reporter mRNA. There are several possible explanations for the differences in translation efficiency. One explanation is that translation of these mRNAs is regulated by RNA-binding proteins (RBPs), and that efficient translation is only possible when an RBP is bound to the mRNA. However, mRNAs did not seem to switch between an active and inactive state during time lapse imaging experiments of ~30-60 minutes, suggesting that the translational state is not caused by transient RBP binding. Another possibility is that even though all mRNAs are transcribed from the same plasmid, they have a different nucleotide sequence. This could happen as a consequence of alternative transcription start site selection or differences in mRNA processing. The exact sequence content of the mRNAs could be determined by long-read sequencing such as nanopore sequencing. Although nanopore sequencing would not reveal the translation efficiency of mRNAs, it could confirm the presence of different pools of mRNAs, which would offer an explanation for the observed heterogeneity in translation activity.

Heterogeneity between mRNAs in susceptibility to NMD

Using our single molecule NMD imaging approach, we could observe two populations of mRNA with different susceptibility to NMD. One population of mRNAs was efficiently degraded, whereas the other population appeared largely insensitive to NMD. Our modelling suggests that the NMD-insensitive population was also degraded, but with a half-life that was much longer than the half-life of the NMD-sensitive population. It should be noted that since both populations are degraded, it is not possible to determine for individual mRNAs to which population they belonged, but the presence of two populations can be inferred from the kinetics of degradation of many mRNAs. The fraction of mRNAs that was sensitive to NMD varied between different reporter mRNAs, suggesting that features of the mRNA determine which fraction can be degraded. It has been reported that EJCs are only deposited on exon-exon junctions during ~80% of all splicing events (Saulière et al., 2012). It is therefore possible that the NMD-insensitive population represents mRNA on which no EJC was loaded. In agreement with this, mRNAs containing multiple introns downstream of the PTC, which are thus more likely to have at least one EJC downstream of the PTC, had a lower fraction of uncleaved mRNAs. An exciting way to confirm this hypothesis would be to combine our NMD imaging assay with single molecule imaging of EJC proteins. This could be done by knocking in a HaloTag in the endogenous locus of an EJC protein and labelling the protein with a bright far red dye, such as JF646 (Grimm et al., 2015). Single fluorescent molecules can only be detected over background fluorescence when few fluorescent molecules are present in the cell, which can be achieved by labelling only a subset of EJC proteins with the fluorescent dye. Although most mRNAs would not be associated with a fluorescently labelled EJC when using sparse labelling, the effect of the presence of an EJC could be determined by focussing solely on the mRNAs that do have a labelled EJC. If mRNAs are insensitive because they lack an EJC, mRNAs associated with a fluorescently labelled EJC should always be sensitive to NMD, and thus not show an NMD-insensitive population.

Impact of understanding NMD on disease and therapy.

The work in this thesis expands our understanding of the molecular mechanism of NMD. Since NMD plays a role in many diseases, knowing how and when NMD degrades PTC-containing mRNAs can help in understanding the effect of NMD on disease phenotypes. In addition, knowledge of the molecular mechanism of NMD can also aid in developing therapies that modulate NMD efficiency in order to increase expression of PTC-containing genes.

A dual role of NMD in disease

NMD does not always have either a positive or negative effect on disease outcome.

Instead, whether NMD is beneficial or harmful depends on whether the proteins produced from PTC-containing mRNAs are useful or harmful. NMD can help prevent disease when production of toxic proteins is prevented. For example, NMD has a general protective function by protecting against accumulation of truncated proteins that arise as a consequence of errors in gene expression. NMD also has a strong effect on gene expression as expression levels of many unmutated genes are regulated by NMD (Mendell et al., 2004), and mutations in the NMD pathway are associated with various neurological disorders and cancer (Liu et al., 2014; Tarpey et al., 2007). In addition, NMD can also play an important role in preventing negative effects of truncated proteins caused by mutations in the DNA. For example, nonsense or frameshift mutations in BRCA1 can result in production of a truncated BRCA1 that could inhibit function of wildtype BRCA1 via a dominant-negative effect (Fan et al., 2001), but the negative effects of truncated BRCA1 are often prevented by degradation of the mRNA through NMD (Perrin-Vidoz, 2002). Finally, NMD can protect against disease by degrading exogenous mRNAs during viral infections. NMD has been shown to degrade mRNA and inhibit viral replication of a coronavirus (mouse hepatitis virus, MHV) that belongs to the same genus as SARS-CoV-2, but this effect was inhibited through repression of NMD by the MHV N-gene (Wada et al., 2018). Although the effect of NMD has not been demonstrated for SARS-CoV-2, there is large similarity between MHV and of SARS-CoV-2, and NMD could therefore even play a role in COVID-19 disease progression.

Besides protecting against disease, NMD can enhance phenotypes when it degrades mRNAs that produce partially-functional proteins. For example, mutations in the penultimate exon of dystrophin result in production of a C-terminally truncated protein that is still partially functional. However, no protein is being produced because the mRNA is degraded through NMD, and the clinical outcome of dystrophin mutations can therefore be worsened by NMD (Kerr et al., 2001). Another way in which NMD can enhance diseases is by preventing detection of tumours by the immune system. Tumours often contain many mutations and could therefore produce truncated proteins that are not present in healthy cells. If these proteins were expressed, they can act as neo-antigens that activate an immune response against the tumour. However, NMD prevents expression of truncated proteins and thereby also prevents an immune response against the tumour (Pastor et al., 2010).

Since the role of NMD in disease is not always obvious, it is important to understand the function of the protein that is affected by NMD. In addition, it is also important to know to which extent NMD affects the expression of PTC-containing transcripts. When an mRNA that would produce a (partially) functional protein is degraded by NMD, inhibition of NMD could benefit a patient by restoring production of the protein. However, not all PTC-containing transcripts are targeted for NMD (Lindeboom et al., 2016), so inhibition of NMD only makes sense when the mRNA is actually targeted

for NMD. Disease-causing mutations are highly variable as there are many genes that can be mutated and there are many locations in a gene that can be mutated to cause disease. It is therefore very useful to predict if an mRNA is targeted for NMD based on sequence alone, without the requirement for validation. Our finding in **chapter 4** that NMD efficiency is affected by the number and position of introns in a gene might help in understanding the situations in which NMD has an impact on expression levels of truncated mRNAs, and could therefore play a role in predicting when NMD inhibition could be useful.

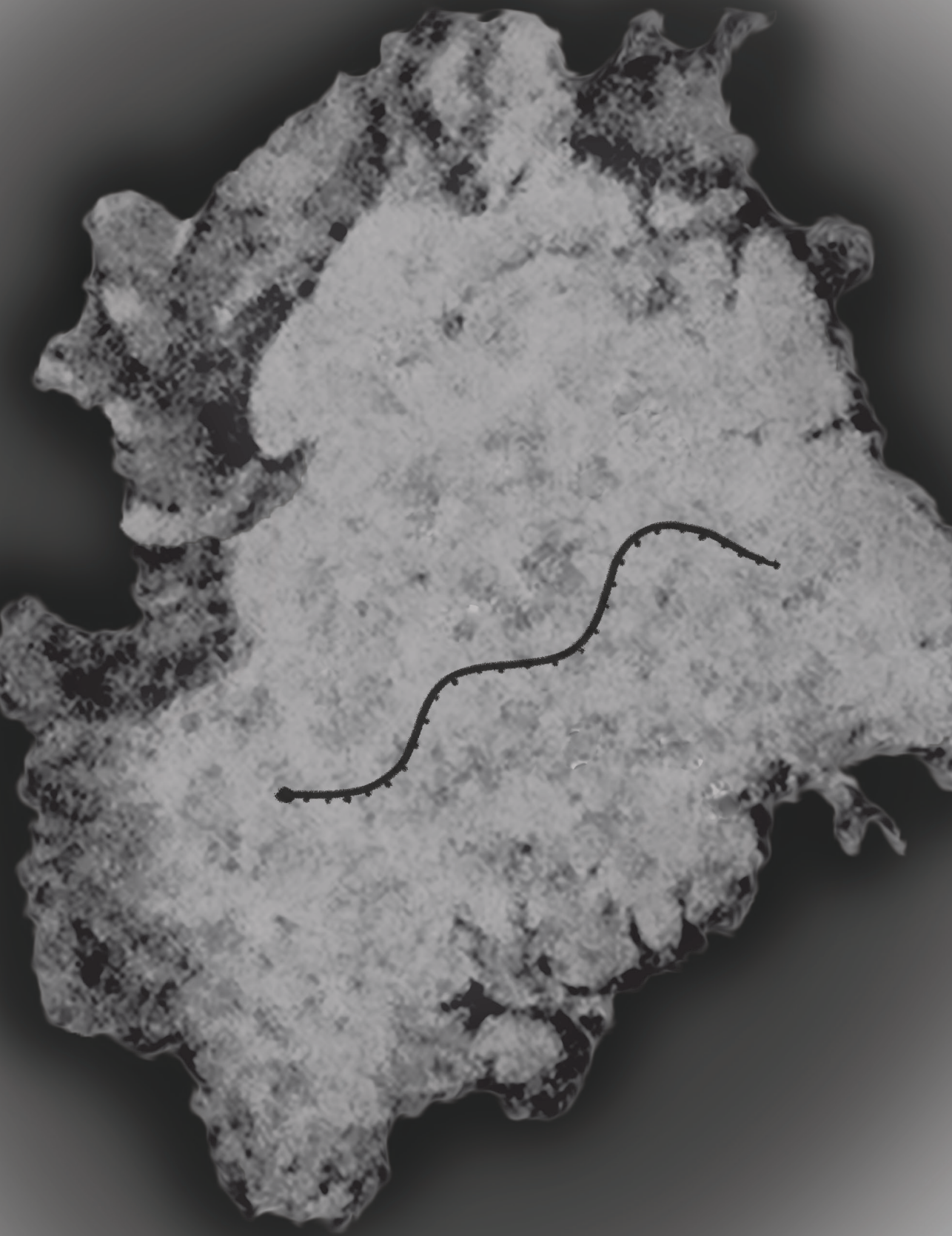
Therapeutic inhibition of NMD

When it is established that NMD negatively impacts disease outcome, NMD inhibitors could be used to restore protein expression. Several NMD inhibitors have been developed that have varying mechanisms of action. For example, the NMD inhibitor NMDi1 prevents NMD by blocking the interaction between UPF1 and SMG5 (Durand et al., 2007), the dietary compound curcumin inhibits NMD by reducing the expression of UPF1, UPF2, and UPF3 (Feng et al., 2015), and antisense oligonucleotides have been used to inhibit NMD by preventing EJC deposition on exon-exon junctions downstream of the PTC (Nomakuchi et al., 2016). Importantly, the mechanism of action of an NMD inhibitor might affect to which extent it affects protein production from a PTC-containing transcript, and understanding under which circumstances various inhibitors are most effective could therefore be useful in selecting the right NMD inhibitor.

Our finding that NMD efficiency is characterized by two parameters, namely the fraction of mRNAs that is sensitive to NMD and the decay rate of the NMD-sensitive population, could have important implications for deciding which NMD inhibitor to use. We find that the fraction of NMD-sensitive mRNA depends on the number of introns that are present downstream of a PTC, and more mRNAs are degraded when multiple downstream introns are present. As EJCs are loaded only on ~80% of exon-exon junctions (Saulière et al., 2012), this is likely because multiple downstream introns increase the likelihood that at least one EJC is loaded on the mRNA downstream of the PTC. Therefore, inhibitors that target mRNA splicing or EJC deposition, such as antisense oligonucleotides, will likely inhibit NMD by modulating the fraction of NMD-insensitive mRNAs without affecting the NMD decay rate. In contrast, we find that decay rate, or the chance that a terminating ribosome induces NMD, of the NMD-sensitive population depends on the presence of a downstream EJC, the distance between the PTC and downstream EJCs, and the expression levels of NMD factors such as UPF1 and SMG6. Therefore, inhibitors that affect expression levels of the UPF proteins, such as curcumin, or affect the interaction between EJCs and other NMD components, such as NMDi1, are likely to affect NMD by altering the NMD decay rate but not the fraction of NMD-insensitive mRNAs.

Although both decreasing the NMD decay rate and decreasing the fraction of NMD-sensitive mRNAs can be used to increase production of protein from PTC-containing transcripts, which treatment is most effective will depend on the kinetics of degradation of that transcript. In particular, the choice of treatment depends on the number of protein molecules made by the NMD-sensitive population and by the NMD-insensitive population. If the majority of protein molecules are made by NMD-insensitive mRNA molecules, the number of protein molecules will likely be more affected by drugs that affect the fraction of NMD-insensitive mRNAs than drugs that affect the decay rate. In contrast, if the NMD-sensitive mRNAs produce a large fraction of the protein molecules, a drug that affects the decay rate might be more effective. For example, the β -globin^{PTC39} reporter had a decay rate of >10% per terminating ribosome for NMD-sensitive population, while ~20% of the mRNAs were NMD insensitive. The observed decay rate suggests that <10 protein molecules are made per NMD-sensitive mRNA. Although we do not know the exact number of proteins produced from the NMD-insensitive population of mRNAs, there are on average ~2800 protein molecules per mRNA molecule in a cell, suggesting that the number of protein molecules produced by NMD-insensitive mRNAs lies in that order of magnitude. Therefore, the vast majority of β -globin protein molecules are likely to be made by the population of NMD-insensitive mRNAs. In this case, a 10-fold reduction of the NMD decay rate would have a much smaller impact on the amount of protein molecules than a 2-fold increase in the fraction of NMD-insensitive mRNAs. Our data thus predicts that NMD of β -globin^{PTC39}, or other transcripts with a high decay rate, could best be inhibited by a drug that affects the fraction of NMD-insensitive mRNAs. Of note, although the number of protein molecules could also be increased by using a very strong inhibitor of the NMD decay rate, side-effects will likely be more severe when strong inhibitors are used, so milder treatment is preferable. In contrast, if a PTC-containing transcript has a low NMD decay rate and fraction of NMD-insensitive mRNAs, it is possible that the majority of proteins are being produced from NMD-sensitive mRNAs. This could for example happen when multiple introns are present downstream of the PTC and the distance between the PTC and first downstream intron is long. We observed that weak NMD reporters were degraded with a probability of <1% per terminating ribosome, suggesting that >100 protein molecules can on average be made from NMD-sensitive mRNAs. For such a reporter, a 10-fold reduction in NMD efficiency of the NMD-sensitive mRNAs would give a very relevant increase in the number of molecules produced. In addition, it is more difficult to prevent at least one downstream EJC being loaded on the mRNA when an mRNA contains multiple introns downstream of the PTC than when only a single downstream intron is present. For such transcripts, NMD inhibitors that affect the NMD decay rate, such as NMDi1 or curcumin, are therefore likely better candidates than drugs that affect EJC loading. Our finding that NMD is characterized by an NMD decay rate and a fraction of NMD insensitive mRNAs

might therefore be useful in deciding which inhibitor to use in which condition. Further experiments should be performed to validate our predictions on the effect of NMD inhibitors on the NMD decay rate and fraction of uncleaved mRNAs and to determine whether they are indeed preferentially applicable in different situations.



Addendum

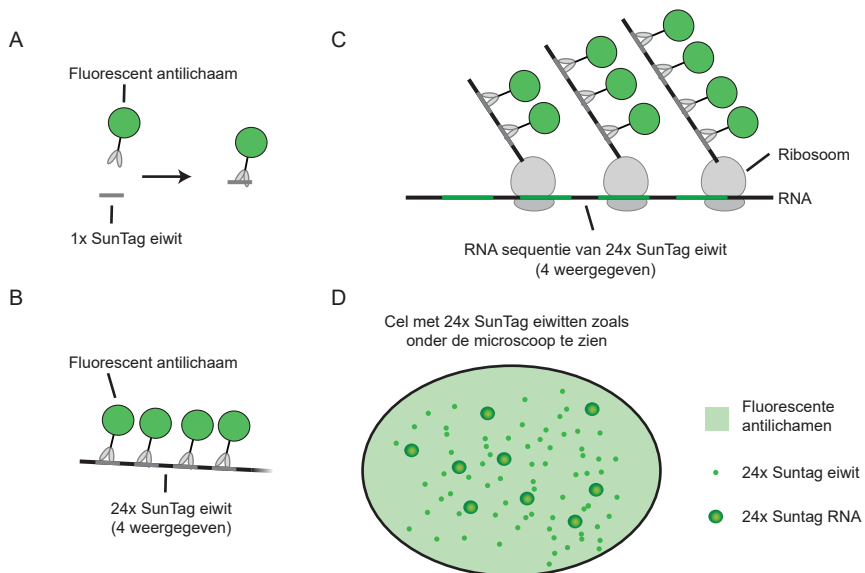
NEDERLANDSE SAMENVATTING

Het DNA in elke cel van het lichaam bevat de volledige genetische informatie die nodig is voor het vormen van een mens. Ons lichaam bestaat uit honderden verschillende typen cellen, zoals huidcellen, spiercellen, en zenuwcellen, die enorm verschillen in vorm en functie. Toch zijn al deze cellen gevormd vanuit hetzelfde DNA. Dit komt doordat niet al het DNA in alle cellen tot expressie wordt gebracht. Ons DNA bestaat uit zo'n 30.000 genen, en alleen de genen die nodig zijn voor de specifieke functie van een cel worden in die cel tot expressie gebracht. Het centrale dogma van de biologie beschrijft hoe dit proces van genexpressie werkt. Een gen wordt eerst gekopieerd naar een RNA-molecuul (dit proces heet 'transcriptie'), en daarna wordt dit RNA molecuul gebruikt als blauwdruk voor het maken van eiwitten door ribosomen (dit proces heet 'translatie'). De eiwitten zijn de functionele bouwstenen van de cel met tal van verschillende functies die ze kunnen uitvoeren nadat hun translatie voltooid is. Je zou dit proces van genexpressie in een cel kunnen vergelijken met de productie van auto's. Een autofabrikant heeft vaak verschillende auto's ontworpen, en de ontwerpen van alle auto's zijn opgeslagen in een centrale database van de fabrikant (het DNA). Om een bepaalde auto te produceren, worden kopieën van de blauwdrukken gemaakt (het RNA) en deze kopieën worden naar fabrieken (de ribosomen) gestuurd. Vervolgens worden in de fabrieken de kopieën van de blauwdrukken gebruikt om de auto's (de eiwitten) op de juiste manier in elkaar te zetten.

Bij de productie van zowel auto's als eiwitten zijn twee dingen erg belangrijk: regulatie van de hoeveelheid productie en monitoring van de kwaliteit. Voor cellen is het belangrijk dat genexpressie goed gereguleerd is zodat de juiste eiwitten aanwezig zijn in de juiste cellen. Zo hebben spiercellen veel eiwitten nodig voor het samentrekken van de spier terwijl zenuwcellen hier weinig aan hebben. Correcte genexpressie wordt in cellen bewerkstelligd door het reguleren van de hoeveelheid transcriptie en de hoeveelheid translatie. Ten tweede is het belangrijk dat er geen fouten gemaakt worden tijdens de productie van eiwitten. Er zijn verschillende soorten fouten die kunnen plaatsvinden in het proces van genexpressie. Deze fouten kunnen, afhankelijk van het soort fout en het gen of eiwit waarin de fouten worden geïntroduceerd, zeer schadelijk zijn en bijvoorbeeld leiden tot spierziektes, neurodegeneratieve aandoeningen, of kanker. Gelukkig bestaan er mechanismen die de kwaliteit van genexpressie constant monitoren. Deze kwaliteitscontrole mechanismes kunnen fouten vaak snel herkennen, en de mogelijke schadelijke effecten beperken door zowel de defecte RNAs als de defecte eiwitten af te breken. Dit proefschrift focust met name op de regulatie van translatie en de kwaliteitscontrole die plaatsvindt tijdens translatie.

Hoewel er al veel bekend is over hoe genexpressie gereguleerd wordt en hoe fouten in genexpressie door kwaliteitscontrole mechanismes herkend worden, is er ook nog een hoop dat we niet weten. Een belangrijke reden dat we nog niet alles weten is dat veel

huidige technieken niet ver genoeg kunnen ‘inzoomen’ op een proces, en daardoor niet precies genoeg zien wat er gebeurt tijdens de kwaliteitscontrole. De meeste technieken die gebruikt worden om translatie te onderzoeken kijken niet naar de translatie van individuele moleculen, maar kijken naar de gemiddelde hoeveelheid translatie van alle moleculen in cel of zelfs van duizenden cellen. In onze autovergelijking zou je inzicht willen krijgen in de efficiëntie van productie door te berekenen hoeveel auto’s er in totaal in elke fabriek gemaakt worden. Dit is zeker informatief, maar er kan nog veel meer



Figuur 1. Het SunTag translatie visualisatie systeem. A) Het SunTag systeem bestaat uit twee onderdelen: een SunTag eiwit en een fluorescente antilichaam. Het SunTag eiwit heeft een unieke structuur die normaal gesproken niet in humane cellen voor komt. Dit SunTag eiwit kan worden gebonden door een antilichaam dat gekoppeld is aan een groen fluorescent eiwit (GFP). B) door 24 SunTag eiwitten te fuseren ontstaat een groter eiwit (24xSunTag eiwit) dat tot 24 fluorescente antilichamen kan binden. C) Een 24xSunTag RNA bevat de RNA sequentie die codeert voor de productie van 24xSunTag eiwitten. Als ribosomen het RNA transleren worden 24xSunTag eiwitten geproduceerd. De fluorescente antilichamen kunnen al binden terwijl deze 24xSunTag eiwitten gemaakt worden. Omdat het RNA door meerdere ribosomen tegelijk wordt getransleerd, en al deze ribosomen tot 24 fluorescente antilichamen kunnen binden, rekruteert het 24xSunTag RNA nog veel meer fluorescente antilichamen dan de afzonderlijke 24xSunTag eiwitten. D) Er zijn drie verschillende groepen van fluorescente antilichamen te zien als je een SunTag cel met een fluorescentiemicroscopie bekijkt. Ten eerste zijn er de fluorescente antilichamen die niet aan een SunTag eiwit zijn gebonden. De cel bevat heel veel van deze losse fluorescente antilichamen die allemaal niet erg fel zijn. Daardoor zijn de ongebonden antilichamen niet te zien als individuele stippen, maar als een diffuus groen signaal. Ten tweede zijn er de 24xSunTag eiwitten die te zien zijn als zwakke groene stippen. Ten derde zijn er de 24x SunTag RNAs. Aangezien de RNAs meestal door meerdere ribosomen tegelijk getransleerd worden zijn deze RNAs zichtbaar als hele felle groene stippen.

informatie verzameld worden door in de fabriek te observeren hoe de auto's worden gemaakt. In dit proefschrift hebben we aantal methodes ontwikkeld die het mogelijk maken om translatie en kwaliteitscontrole van individuele RNA moleculen zichtbaar te maken. Met behulp van microscopie kunnen we deze RNA moleculen uren lang volgen, zien wanneer en hoe efficiënt translatie van deze moleculen verloopt, en observeren wanneer defecte RNAs worden afgebroken door kwaliteitscontrolemechanismes.

De basis van de verschillende methodes die we ontwikkelen bestaat uit het zichtbaar maken van translatie met behulp van het SunTag translatie visualisatie systeem. Het SunTag systeem bestaat uit twee componenten, namelijk een SunTag eiwit en een fluorescent antilichaam (figuur 1A), die het samen mogelijk maken om translatie van individuele RNA moleculen zichtbaar te maken. Het SunTag eiwit is een heel klein eiwit met een unieke structuur die normaal gesproken niet in menselijke cellen voorkomt. Het antilichaam is een ander eiwit dat specifiek aan het SunTag eiwit bindt en niet aan andere eiwitten in humane cellen. Dit antilichaam is gekoppeld aan een fluorescent molecuul (green fluorescent protein, of GFP) dat groen oplicht onder de microscoop als er licht op geschoten wordt. Er zijn veel van deze fluorescente antilichamen in een cel, wat zorgt voor een diffuus groen signaal dat egaal over een cel verspreid is (Figuur 1D). Als een fluorescent antilichaam bindt aan een SunTag eiwit wordt het groene signaal naar de plek van het SunTag eiwit gerekruteerd. Dit geeft op zichzelf nog weinig voordeel, omdat het SunTag eiwit nu net zo fel is als het fluorescente antilichaam. Echter, omdat het SunTag eiwit erg klein is kunnen meerdere SunTag eiwitten (tot wel 24) aan elkaar vast gemaakt worden om één groot 24x SunTag eiwit te maken. Aan zo'n 24x SunTag eiwit kunnen ook 24 fluorescente antilichamen binden, waardoor op één plek veel meer fluorescente antilichamen zijn dan in de rest van de cel. Hierdoor is een 24x SunTag eiwit onder de microscoop zichtbaar als een groene stip die feller is dan de rest van de cel (Figuur 1B, D). De fluorescente antilichamen kunnen ook al aan het SunTag eiwit binden terwijl het eiwit gemaakt wordt door ribosomen in het proces van translatie. Echter, translatie van een RNA gebeurt meestal niet door 1 ribosoom, maar wordt door meerdere ribosomen tegelijk gedaan (vergelijk dit met een lopende band waar aan meerdere auto's tegelijk kan worden gewerkt). Iedere ribosoom die een 24x SunTag RNA transleert maakt een 24x SunTag eiwit. Als er veel ribosomen tegelijk een 24x SunTag RNA transleren, zorgt dit voor een clustering van 24x SunTag eiwitten die nog verbonden zijn aan het RNA, waardoor heel veel fluorescent antilichamen gerekruteerd worden (Figuur 1C). Het fluorescente signaal van een 24x SunTag RNA is hierdoor onder de microscoop te zien als een groene stip die nog veel feller is dan de stippen van 24x SunTag eiwitten. (Figuur 1D). Deze methode van SunTag labelen maakt het mogelijk om individuele RNA moleculen in een cel te herkennen. Een bijkomend voordeel is dat de felheid van de stip informatie geeft hoe actief translatie van dat RNA op dat moment is. Hoe feller de stip, hoe meer ribosomen op dat moment het RNA aan het transleren zijn. Met fluorescentiemicroscopie is het mogelijk om levende

cellen gedurende minuten, uren, of zelfs dagen te filmen om te observeren hoe de felheid van deze stippen verandert met de tijd. Deze data kunnen dan gebruikt worden om bijvoorbeeld te achterhalen of translatie verandert over tijd en of er verschil is in translatie tussen verschillende RNA moleculen. Op deze manier gebruiken we SunTag translatie visualisatie om inzicht te krijgen in de mechanismes die efficiëntie van translatie reguleren. In latere hoofdstukken breiden we de methode steeds verder uit om naast translatie ook andere processen te visualiseren.

In **hoofdstuk 2** hebben we de SunTag translatie visualisatie methode ontwikkeld. Als eerste doen we meerdere experimenten om te laten zien dat de felle groene stippen inderdaad getransleerde RNAs zijn. We laten ook zien dat het toevoegen van de SunTag eiwitten en fluorescente antilichamen geen effect heeft op de efficiëntie van translatie, wat betekent dat de resultaten die we met onze methode krijgen betrouwbaar zijn. Vervolgens onderzoeken we hoeveel ribosomen tegelijkertijd een RNA molecuul transleren en hoe snel ieder ribosoom een RNA transleert. We zien dat de snelheid waarmee een ribosoom een RNA transleert afhangt van het RNA dat getransleerd wordt, en dat deze snelheid lager wordt als het RNA beschadigd is. Daarnaast observeren we dat de mate waarin een RNA molecuul getransleerd wordt kan veranderen met de tijd. Dit lijkt een alles-of-niets proces te zijn, waarin translatie of erg actief is, of helemaal uit lijkt te staan. We zien ook dat niet alle RNA moleculen even actief getransleerd worden, zelfs als deze RNA moleculen afkomstig zijn van hetzelfde DNA. Zo zien we voor een bepaalde RNA sequentie dat een kleine subgroep van de moleculen 50x actiever wordt getransleerd dan de andere RNA moleculen. Samengevat maakt deze nieuwe methode het voor het eerst mogelijk om naar de translatie van individuele RNA moleculen te kijken, en dit heeft veel inzicht gegeven in de efficiëntie en variabiliteit van translatie.

In **hoofdstuk 3** gaan we dieper in op wat er nodig is om het SunTag translatie visualisatie systeem op te zetten. We beschrijven hierin wat voor cellen er nodig zijn, wat voor microscoop gebruikt kan worden, en hoe microscopie data geïnterpreteerd kan worden om inzicht te krijgen in de dynamiek van translatie. Dit hoofdstuk bevat geen nieuwe data, maar is gericht op onderzoeksgroepen die dit systeem willen gebruiken maar er nog geen ervaring mee hebben.

Er zijn verschillende situaties waarin fouten worden gemaakt in translatie, en een veelvoorkomende fout is het stoppen met translatie voordat een volledig eiwit is gemaakt. In **hoofdstuk 4** ontwikkelen we onze methode verder om deze situatie te visualiseren. Onvolledige eiwitten kunnen erg schadelijk zijn voor de cel als ze nog wel een deel van hun functies kunnen uitvoeren. Vergelijk dit met een onvolledige auto: het is veel veiliger om geen auto te gebruiken dan om in een auto zonder remmen te rijden. Onvolledige eiwitten kunnen ontstaan wanneer er stop-sequenties, of stopcodons, te vroeg (=niet aan het einde van het RNA) in het RNA terecht zijn gekomen. Te vroege stopcodons ontstaan onder andere door mutaties in het DNA en dit soort mutaties zijn verantwoordelijk voor zo'n 20% van alle genetische ziektes! Als een RNA met een

te vroeg stopcodon getransleerd wordt, kan het RNA herkend worden door een RNA kwaliteitscontrole mechanisme genaamd nonsense-mediated RNA decay (afgekort NMD). Herkenning door NMD zorgt ervoor dat het RNA in tweeën wordt geknipt en daarna verder wordt afgebroken. In **hoofdstuk 4** ontwikkelen we de SunTag translatie visualisatie methode verder om ook het afbreken van RNAs door NMD zichtbaar te visualiseren. Om dit te doen voegen we, naast het groene signaal van het fluorescent antilichaam, een rood fluorescent signaal toe dat verbonden is aan het uiteinde van de RNA moleculen. Omdat beide fluorescente signalen aan hetzelfde RNA molecuul verbonden zijn, overlappen de rode en de groene stip normaal gesproken altijd. Echter, als RNAs een te vroeg stopcodon hebben en in tweeën worden geknipt door NMD kunnen de rode en de groene stip van elkaar weg bewegen. Omdat dit normaal gesproken nooit gebeurt, kunnen we hiermee precies zien waar en wanneer RNA moleculen afgebroken worden door NMD. Het kunnen observeren van NMD van individuele RNA moleculen heeft belangrijke nieuwe inzichten opgeleverd. Zo ontdekten we dat niet elk ribosoom dat stopt op een te vroeg stopcodon zorgt voor afbraak van het RNA. Dit komt niet doordat ribosomen verschillend zijn, maar komt doordat elke ribosoom maar een kleine kans heeft om afbraak te induceren. Hoe groot die kans is hangt af van eigenschappen van het RNA molecuul. Op RNAs die zeer efficiënt herkend worden door NMD was de kans dat een ribosoom afbraak induceerde ongeveer 10%. Aangezien een RNA gemiddeld door zo'n 1000 tot 10.000 ribosomen getransleerd wordt zorgt een 10% kans nog steeds voor hele snelle afbraak van het RNA. Op minder efficiënte substraten was die kans minder dan 1%, terwijl RNAs zonder een te vroeg stopcodon praktisch nooit door NMD werden afgebroken. Daarnaast zagen we voor zowel efficiënte als inefficiënte NMD-substraten dat een subset van de RNA moleculen helemaal niet afgebroken werd. Dit lijkt te maken te hebben met eiwitten die aan RNA binden wanneer het RNA gemaakt wordt tijdens transcriptie. Onze observaties geven nieuw inzicht verschillende aspecten die relevant zijn voor de efficiëntie van NMD, en kunnen op den duur mogelijk belangrijk zijn bij het bepalen van de juiste medicijnen bij het behandelen van ziektes waarbij NMD een rol speelt.

In **hoofdstuk 5** gebruiken we SunTag translatie visualisatie om naar een ander kwaliteitscontrole mechanisme te kijken, namelijk naar non-stop decay (NSD). Non-stop decay is mechanisme dat zoekt naar RNAs die helemaal geen stopcodon hebben, en waardoor ribosomen niet op een goede manier stoppen met translatie. De eiwitten die hierdoor gemaakt worden hebben extra toevoegingen aan het eind van het eiwit die schadelijk kunnen zijn. Vergelijk dit met een auto die af is maar waar daarna nog willekeurige onderdelen aan de motor toegevoegd worden, dit kan ervoor zorgen dat de motor niet meer werkt. Non-stop decay wordt geactiveerd wanneer ribosomen die een RNA transleren helemaal bij het einde van het RNA komen en dit zorgt ervoor dat het RNA wordt afgebroken door het exosoom, een eiwit dat als een soort pacman het RNA vanaf het uiteinde afbreekt (of opeet). Als een SunTag RNA wordt afgebroken door

een exosoom zorgt dit ervoor dat SunTag eiwitten (en bijbehorende ribosomen) één voor één van het RNA afvallen. Onder de microscoop ziet dit eruit als een geleidelijke afname in de felheid van de groene stip. We ontdekten dat deze afbraak niet alleen gebeurt wanneer een ribosoom het einde van het RNA bereikt, maar ook als een ribosoom dicht in de buurt komt van het uiteinde van het RNA. In dat geval zijn er andere eiwitten nodig om de afbraak op te starten, maar de uiteindelijke afbraak gebeurt op dezelfde manier als wanneer ribosomen wel het einde van het RNA bereiken.

In **hoofdstuk 6** beschrijven we in meer detail wat er nodig is om de verdere ontwikkelingen van SunTag translatie visualisatie uit te voeren. We gaan hier ook dieper in op hoe nieuwe software gebruikt kan worden om de analyse van microscopie data te automatiseren zodat de maximale hoeveelheid informatie uit ieder experiment gehaald kan worden.

In **hoofdstuk 7** werken we voor het eerst niet met de SunTag, maar beschrijven we een nieuwe methode voor diagnostische COVID-19 testen die we aan het ontwikkelen zijn. Deze methode maakt gebruik van de standaard PCR test, maar gebruikt een nieuw protocol om het RNA van het virus met behulp van magnetische beads te isoleren. Meerdere optimalisaties in deze methode zorgen ervoor dat de testcapaciteit drastisch hoger is dan bij huidige methodes. De verwerking van de testen wordt gedaan door een nieuwe robot die we speciaal hebben ontworpen voor COVID-19 diagnostiek en die tot 20.000 testen per dag kan verwerken. Daarnaast verzamelen we de swabs (de stokjes die de neus in gaan bij het testen) in nieuwe buizen die direct compatibel zijn met de robot, waardoor het menselijk handelen wordt geminimaliseerd. In het hoofdstuk beschrijven we hoe we verschillende componenten van de methode testen, zoals de magnetische beads die we gebruiken, de buffer waarin we de swabs opvangen, en de beste manier om swabs te verzamelen zonder contaminatie te introduceren. Sinds het schrijven van hoofdstuk 7 is de robot ook daadwerkelijk geleverd en wordt de toepasbaarheid van de methode in de praktijk getest.

Samengenomen ontwikkelen we in dit proefschrift verschillende methodes om translatie en kwaliteitscontrole van RNA zichtbaar te maken. Deze methodes maken het voor het eerst mogelijk om te bestuderen hoe translatie van individuele RNA moleculen varieert met de tijd, wat voor verschillen in translatie er zijn tussen RNA moleculen in een cel. Daarnaast is het met deze methodes ook mogelijk om de interactie tussen translatie en kwaliteitscontrole zichtbaar te maken door heel precies te observeren wanneer een RNA wordt afgebroken. In onze experimenten vergelijken we vaak fysiologische en pathologische omstandigheden waardoor we het effect van verschillende omstandigheden op translatie en kwaliteitscontrole kunnen bestuderen. Dit onderzoek draagt bij aan de kennis van de moleculaire processen die een rol spelen bij verschillende ziektes, en verder onderzoek kan op den duur leiden tot betere behandeling van ziektes aan de hand van deze kennis.

CURRICULUM VITAE

Tim Andreas Hoek werd op 11 februari 1992 in Utrecht geboren. Hij behaalde zijn diploma aan het Utrechts Stedelijk Gymnasium te Utrecht met de profielen Natuur & Techniek en Natuur & Gezondheid. Daarna heeft hij een bachelor biomedische wetenschappen gedaan aan de Universiteit Utrecht. In 2013 begon hij aan zijn master Cancer, Stem Cells & Developmental Biology aan de Graduate School of Life Sciences van dezelfde universiteit. In het kader van deze Master heeft Tim een stage van 9 maanden gedaan in het lab van Prof. Dr. Geert Kops. Hier onderzocht hij de rol van het eiwit BUB1 bij het mitotische checkpoint van de celcyclus. Daarna heeft hij een stage van 8 maanden gedaan in het lab van Prof. Dr. Jeff Gore bij het Massachusetts Institute of Technology (MIT) in Cambridge, MA. Hier bestudeerde hij hoe de mate van samenwerking tussen verschillende microben af hangt van de hoeveelheid voedingsstoffen in de omgeving. Nadat de master was afgerond met de distinctie Cum Laude is Tim in 2015 begonnen als PhD student in het lab van Dr. Marvin Tanenbaum bij het Hubrecht Instituut in Utrecht. Hij is hierin financieel gesteund door een fellowship van het Boehringer Ingelheim Fonds. De resultaten verkregen in dit promotietraject staan beschreven in dit proefschrift.

LIST OF PUBLICATIONS

Yan, X., **Hoek, T.A.**, Vale, R.D., and Tanenbaum, M.E. (2016). Dynamics of Translation of Single mRNA Molecules in Vivo. *Cell* 165, 976–989.

Ruijtenberg, S., **Hoek, T.A.**, Yan, X., and Tanenbaum, M.E. (2018). Imaging translation dynamics of single mRNA molecules in live cells. *Methods in molecular biology* 1649, 384-404.

Hoek, T.A.*, Khuperkar, D.*, Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., Boersma, S., Vermeulen, M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Mol. Cell* 75.

Khuperkar, D.*, **Hoek, T.A.***, Sonneveld, S., Verhagen, B.M.P., Boersma, S., and Tanenbaum, M.E. (2020). Quantification of mRNA translation in live cells using single-molecule imaging. *Nat. Protoc.* 15, 1371–1398.

* Equal contribution

DANKWOORD

Na meer dan 5 jaar zit het er dan toch echt op en ga ik mijn PhD verdedigen. Het is een mooie tijd geweest waarin veel mensen hebben bijgedragen aan het plezier dat ik in mijn PhD heb beleefd, waar ik jullie allemaal voor wil bedanken.

Marvin, ik weet nog goed hoe ons online sollicitatiegesprek meer dan 3 uur duurde omdat we maar niet uitgepraat raakten over allerlei wetenschappelijke en niet wetenschappelijke onderwerpen. Jouw enthousiasme voor de wetenschap was erg aanstekelijk, en onze wekelijkse (en soms bijna dagelijkse) meetings waarin we over alle mogelijke scenario's konden brainstormen hoorden voor mij tot de leukste en leerzaamste momenten van mijn PhD. Ik ben blij dat ik vanaf het begin af aan bij het lab heb gezeten en heb kunnen zien en bijdragen aan hoe het lab zich over de jaren heeft ontwikkeld. Ik waardeer het ook dat je naast wetenschap altijd bezig bent om een goede sfeer in het lab te creëren en nog dicht genoeg bij ons staat om af en toe een 1 april grap uit te halen (of te ontvangen).

Wouter, bedankt dat je mijn promotor bent. Ik ben blij dat we elkaar in mijn laatste jaar regelmatig hebben gesproken voor het ontwikkelen van the Beast. **Geert, Alexander, Jacques, Lucas, and Michiel**, thank you all for being in my reading committee. **Geert**, hoewel ik mijn PhD niet in jouw lab ben gaan doen heb ik altijd het gevoel gehad alsof ik er nog een beetje bij hoorde, en ik heb jouw betrokkenheid bij mijn ontwikkeling erg gewaardeerd.

Current and former Tanenbaum lab members, **Deepak, Stijn, Sanne, Ive, Bram, Max, Iris, Huib, Rupa, Lucas, Micha, Dhanushika, Sora, Lenno, Suzan**, thank you all for being such great colleagues and making it fun to go to work every day. As a group you are smart, critical, and, above all, fun. My PhD wouldn't have been the same without all of you and all the fun activities we did. **Deepak**, it's been amazing to have you around as a colleague and friend. I really enjoyed your positive attitude and endless optimism, how we could always talk about everything, playing endless random games (header high score, coffee pong, foosball), going to borrels together, watching soccer, and teaching you Dutch pickup lines (is het de of het magazine?). Best of luck in your next post-doc, be it Europe or US, where I would really like to come visit you. **Sanne**, je bent een geweldige wetenschapper, en het was altijd een genot om met je te brainstormen over projecten, soms zelf samen te werken (die ene proef in negen maanden), elkaar te ondersteunen na een feestje, en klaagsessies te houden over wat er allemaal wel niet vervelend is aan een PhD. Ik ben blij dat het me gelukt is voor je te promoveren, zodat je me zelfs als je straks de nobelprijs wint nog steeds als de senior PhD zal zien. **Stijn**, bedankt voor al die jaren samen werken, klimmen (deze zomer een keer buiten?), borrelen, discussiëren (veel nieuwe inzichten in politiek gekregen), en semi-correcte opmerkingen (Poetin van het jaar). **Max**, your eagerness to learn has always made it very enjoyable to work together with you (so no need to apologize for

asking questions). It's been great to go climbing together for the last years, and maybe we can go on a snowboarding trip next year? **Iris** (buurtje), het is mooi om te zien hoe jij steeds zelfstandiger wordt als wetenschapper. Ik vond het heel leuk om samen met je te werken aan het non-stop decay project (wie weet ooit nog gepubliceerd?) en om samen naar Parijs te gaan (ondanks Li). **Rupa**, I really loved your energy (especially when parties went late) and thanks for making Christmas a bit more fun with the Secret Santa in this time of lockdown. **Bram**, bedankt voor je gezelligheid op het lab en daarbuiten (ook Zweedse Bram bedankt). Als ik weer eens echt fit ben durf ik misschien wel een keertje met je te gaan hardlopen. **Ive**, bedankt voor alle gezelligheid en hulp in het lab. Ik ben nog steeds onder de indruk van hoe je me keihard hebt ingemaakt met planken! **Huib**, altijd relaxed en vrolijk op het lab (maar toch 1 keer boos geweest!). Bedankt dat je altijd voor me klaar stond als ik met iemand koffie wou drinken, zelfs als je kopje nog vol was. **Lucas**, onuitputtelijke bron van interessante (of misschien zelfs nutteloze?) weetjes. Ik ben blij dat ik door jou voor het eerst ben gaan mountainbiken. **Micha**, even though you joined only recently, you seem to have integrated flawlessly and I feel like we know each other quite well already. Keep up the progress in climbing and you will be the best of us in no time. **Dhanushika** and **Sora**, I did not get to spend a great deal of time with you, but I'm sure you will have a great time in the Netherlands and in our lab. **Lenno** en **Suzan**, het was ontzettend leuk om samen met jullie in het lab te beginnen. Het voelde altijd als een klein hecht team en het was raar om de laatste periode zonder jullie af te maken. Mijn voormalige studenten, **Linda** en **Lotte**, ook jullie bedankt. Ik wens jullie het beste en hoop dat jullie een mooie plek hebben gevonden om je (wetenschappelijke) carrière voort te zetten

Verder zijn er een hoop mensen met wie ik tijdens mijn PhD heb samengewerkt, die ik graag wil bedanken. **Peter**, ik heb met veel plezier met je samengewerkt aan het COVID-project, en ben erg blij dat we het tot een mooi einde hebben kunnen krijgen. Ik denk niet dat ik ooit zo veel met iemand gebeld heb als met jou (gelukkig is mijn vriendin niet snel jaloers)! **Martijn**, zonder jouw expertise was het nooit gelukt om ook maar in de buurt van de throughput van The Beast te komen. **Jeroen**, **Lieven**, **René**, het was leuk om bij jullie te zien hoe de COVID-diagnostiek eraan toe gaat en om te zien hoe de robot steeds verder wordt ontwikkeld. **Sanne**, **Thomas** (gelukkig is het niet de LAMP assay geworden), **Mark**, **Bram**, **Ive**, **Pim**, ook jullie bedankt voor jullie hulp. Ik hoop dat de robot volledig operationeel is tegen de tijd dat jullie dit boekje zien. **Gautam** en **Imre**, ik ben heel blij dat er ondanks de vertragingen nog steeds progressie wordt gemaakt in het project. **Rik**, ik ben blij dat je met jouw kennis van NMD ons hebt kunnen helpen met het maken van een mooi paper. Succes in Cambridge. **Jelmer**, bedankt voor je hulp bij het project van **Linda** om te kijken naar het effect van schimmel supernatanten op NMD. **Erdal**, **Rogier**, **Maikel**, en nu ook **Timo**, ik heb misschien wat meer contact gehad met Nikon dan ik gewild zou hebben, maar het was altijd een prettige samenwerking. **Anko** en **John**, bedankt jullie altijd snel konden helpen met

microscopie- (en andere) problemen, zelfs als jullie gigantisch druk waren met andere dingen. **Litha**, jouw hulp heeft een hoop stress voorkomen door te zorgen dat alles soepel verliep bij het maken van mijn thesis. Verder ook bedankt aan de **technische dienst, civiele dienst, receptie, IT, mediabereiding, en personeelszaken!**

I also want to thank everyone from the **Boehringer Ingelheim Fonds**, both other fellows and staff. It has been a great experience to be part of a fellowship with such involved people, and it has been awesome to see most of you even outside of the BIF meetings in the Brussels retreat.

Een groot deel van wat mijn PhD zo leuk heeft gemaakt komt voort uit de sociale activiteiten. De vele borrels, kerstdiners, PhD-retreats en vakanties met jullie samen hebben het een hele mooie tijd gemaakt. Ik wil iedereen bedanken die bij deze activiteiten aanwezig zijn geweest, ook als je niet bij naam wordt genoemd.

Caro, vanaf dat we samen in de PV zaten hadden we meteen een goede klik. Bedankt voor de goede tijden tijdens borrels (vooral de dronken photoshoots), koffie (als werk even wat veel werd), hardloopsessies (blij dat ik toch nét iets sneller was op de halve marathon, ook al was jij daarna altijd sneller), en vakanties. Jouw aanwezigheid heeft mijn PhD ontzettend veel leuker gemaakt. Ik ben heel blij dat ik bij jouw verdediging paranimf kon zijn en dat je nu (als het lukt) uit Boston overkomt om ook bij mijn verdediging als paranimf aanwezig te zijn (hopelijk de dag erna weer net zo brak als na jouw feest)! **Lotte B, Sven, Sasja, Caro, Maria, Martha**, het was een genot om met jullie in de PV te zitten, bedankt voor een mooie tijd! **Lotte B**, ook heel erg bedankt bij het helpen vormgeven van dit boekje. Zonder jou het had het er niet zo mooi uit gezien! **Bas C**, altijd in voor een biertje zodra de klok op vrijdag 5 uur sloeg. Gelukkig ben ik minstens één keer naar de pubquiz gekomen. **Colinda**, altijd tijd om even koffie te drinken of om te klieren op borrels. Veel succes als groepsleider in Leuven. **Erik**, leider van alle spelletjes of pubquizen, het is nooit saai met jou erbij. **Christa**, bedank voor alle leuke gesprekken, ik waardeer je openheid. **Dennis**, bedankt voor de vele potjes tafelvoetbal (hopelijk win je ooit nog een keer). **Laura**, it was always fun to joke around together with you. **Annabel**, bedankt dat je samen met mij paranimf bent geweest. **Corina** (Krhkah!), bedankt voor je gezelligheid. **Banafsheh**, bedankt voor de diepe gesprekken en ik wens je alle geluk met Mina. **Reinier**, bedankt voor je eindeloze energie en enthousiasme. **Ajit**, bedankt voor de leuke borrels en feestjes. **Bas de W**, bedankt voor de goede drankspelletjes (race cage!). **Timo**, bedankt voor de gezelligheid en tegenwoordig ook hulp bij de microscopen bij Nikon. **Wim**, bedankt voor het organiseren van het zaalvoetbal. **Anne**, bedankt voor de gezelligheid en het oprichten van de Burping appgroep. **Sonja**, ik snap nog steeds niet hoe je me hebt verslagen in de WK-poule. **Richard**, altijd lekker kunnen discussiëren over voetbal. **Gaby, Wouter, Bas M, Kim, Margit, Carlos, Nannette, Jens, Wessel**, thank you all for the fun times we had.

Everyone who joined to Mallorca: **Sjoerd, Kim, Ramada, Silke, Ana, Joana, Deepak,**

Stijn, Iris Bram, Caro, Rupa, Lotte, thanks for a great holiday! I greatly enjoyed the swimming, drinking games (Cheers to the governor), partying, poledancing, flyboarding (**Ramada**), people arriving a day before and sleeping on the beach (**Iris**), and all the great moments we had.

I am very happy that we also had a skiing trip to Winterberg with a great group of people. **Deepak, Stijn, Ajit, Anne, Lotte, Samy, Marjolein, Claudia, Bana, Sjoerd, Sanne, Bram, Ator, Bas**, thanks for an amazing time. Lots of snowboarding, partying (in a very interesting club), something that resembles dance moves (**Stijn** and **Sjoerd**), another kind of move (**Deepak**), snowboard lessons (**Lotte** and **Ator**), and fun all around!

I also had the pleasure to join to multiple trips to the Ardennes. These were amazing weekends with amazing people, **Erik** (kampleider), **Wim en Bas C** (bedankt voor de organisatie en pubquizen), **Reinier, Clement, Spiros, Margit, Rob** (the source), **Caro** en **Christa** (hardlopen in het grauwe weer), **Lotte, Roxanne, Annabel, Kim, Wessel, Dennis, Anne, Jens** (always in for soccer), **Ramada, Bas de W, Laura, Lotte B, Colinda, Ajit, Sonja, Stijn, Sanne** (sorry dat ik je uit de boom heb geholpen/geschopt), and many more people. Thank you all!

Ook buiten het Hubrecht zijn er mensen die veel voor me betekenen hebben tijdens mijn PhD die ik graag wil bedanken. **Ruben**, al vrienden vanaf de basisschool, en wat hebben we veel samen meegemaakt. Samen reizen naar Vietnam (en **Na** ontmoet) en wandelen in de Alpen (jammer dat je die laatste top niet meer op wilde), voetballen, Age of Empires spelen, adtje kratje, chillen, en praten over van alles en nog wat. Ik ben blij dat we al zo lang goede vrienden zijn en dat je als paranimf naast me zal staan bij mijn verdediging. **Tim**, mijn mede PhD'er buiten werk. Bedankt voor onze goede gesprekken, grappige momenten (Perpetuum mobile!), dronken avonden (heey, mooie steen) en voor de vele potjes voetbal in de lockdown (altijd kontje kick bij **Ruben**). **Daan**, altijd in voor nieuwe uitdagingen (samen een marathon gelopen!), gekke weddenschappen (20 euro als je nu acht gooit) en goede gesprekken. Bedankt voor mooie tijden. **Paul** (KAPUTMACHEN!), ik kan altijd goede diepe gesprekken met jou hebben en dat waardeer ik enorm! Nu nog even aan je conditie werken zodat je bij de volgende wandeltocht niet als eerste afvalt. Samen met jullie heb ik het geluk gehad om met een geweldige groep vrienden in een voetbalteam te zitten bij het mooie VV Sterrenwijk. **Frank, Marijn, Wally, Max N, Max R, Jeroen, Lennart, Stash, Mark, Bas O, Bas vH, Vincent vS, Elmo, Merijn, Roel, Wouter B, Camille, Bram, Gijs, Vincent, Sven, Hugo, Mike**, ook jullie bedankt. Het is een genot om met jullie allemaal in het team te zitten of te hebben gezeten, en ik ben blij dat we na jaren van tegenslag nu een team hebben dat regelmatig wint en geen moeite meer heeft om 11 spelers bij elkaar te rapen.

Dan mijn jaarclub, **Guido, Casper, Boris, Clinton, en Douwe**. Wat ben ik blij dat we na zoveel jaar nog steeds zo'n hechte groep zijn. Onze jaarclubavonden waren de

afgelopen jaren heel fijn om werk even los te laten. Ik heb altijd erg genoten van onze gesprekken en jaarclubweekenden en kijken al uit naar het volgende weekend!

Audrey, mijn maatje uit Boston. Bedankt voor de leuke skypesessies en vakanties die we samen hebben doorgebracht. Het was mooi onze PhDs ongeveer tegelijk te beginnen en te eindigen, alsof we samen het proces doormaakten. Ik kijk er naar uit om je snel weer te zien en als het mogelijk is zelfs samen te reizen! **Margo**, bedankt voor de leuke momenten van samen klimmen, borrellen, en de klimvakanties op bezoek bij **Thom** (geweldige klimgids, volgende keer ga ik weer mee!). **Ruben van H**, bedankt voor de leuke borrels en klimesessies. Laten we snel weer een biertje gaan doen!

Elly, Sil, Anisa, en Wies, ik ben blij dat ik jullie via **Lotte** heb leren kennen en ook steeds hechter met jullie wordt. Ik ben dankbaar voor de gezellige etentjes, de vele spelletjes (**Elly** nu ook fan van Dominion), en de leuke uitjes (slapen in het Archeon!) die we samen hebben gedaan, en hoop dat we binnenkort ook met z'n allen op vakantie kunnen naar Zwitserland.

Wouter, wat leuk dat je ook op de Notebomenlaan bent komen wonen en nu ook meevoetbalt. Laten we vaak samen biertjes blijven doen, en geniet nog even van je bitcoin-bubbel. **Teun**, ik heb een aantal hele mooie jaren met jou samengewoond, waarin we veel hebben gelachen, gekloot, gepraat, en strategieën voor de beste berichtjes hebben bedacht. Bedankt! **Henk, Anna, Bart, Carla, Anne, Lisa**, bedankt voor de gezellige verjaardagen, sinterklaasavonden, en kerstdiners!

Pap, Mam, Marijn, Jorinde, Laura, en Koen. Bedankt dat jullie zo'n geweldige en gezellige familie zijn! **Jorinde**, zusje, het is mooi om te zien hoe je steeds volwassener wordt en steeds meer je plek lijkt te vinden (en nu zelfs verstandig met geld om gaat!). Ik ben heel blij om jou zo goed te kennen en veel met je om te gaan, van samen op vakantie gaan (volgende keer niet ziek worden!), naar feestjes, of gewoon op bezoek in Amsterdam en kletsen onderaan de pier over alles wat ons dwars zit. Nu ook met **Koen** erbij, gezellig en gelukkig geeft hij je voldoende tegengas. **Marijn**, wat ben ik blij jou als broer te hebben. Zo veel samen gedaan, samen in het voetbal team gezeten, samen met jou en met **Laura** naar Lowlands en op vakantie geweest, en altijd goede gesprekken onder het genot van een wandelingetje. Ik kijk er naar uit om binnenkort oom te worden en wens jullie veel succes met de nieuwe fase in het leven! **Pap en mam**, bedankt dat jullie er altijd voor me zijn. Het heeft mij veel vertrouwen gegeven om te weten dat jullie me altijd steunen. Jullie oprechte interesse maakt het altijd leuk om langs te komen (zelfs al wil ik soms snel weer door) of dingen met jullie samen te doen. Ik kan me geen betere ouders voorstellen dan jullie!

Lotte, leukerd, wat hebben we al veel samen meegemaakt de afgelopen jaren! Samen skydiven, Lowlands, kamperen in Slovenië, roadtrippen in Amerika, samenwonen, **Lizzy**, straks Mexico, en nog veel meer andere leuke dingen. Ik ben dankbaar dat je er altijd voor me bent! Mijn PhD zou niet hetzelfde zijn geweest zonder jou en ik kijk uit naar ons volgende avontuur samen.

