Year: 2024

Investigating re-initiation mechanisms in vivo and in vitro

Meurs Romane

Meurs Romane, 2024, Investigating re-initiation mechanisms in vivo and in vitro

Originally published at : Thesis, University of Lausanne
Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_8C17B5F64E643

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Center for Integrative Genomics

Investigating re-initiation mechanisms *in vivo* and *in vitro*

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Romane Meurs

Master en Biologie Moléculaire et Cellulaire et Biochimie
de l'Université Libre de Bruxelles, Belgique.

Jury

Prof. Thierry Pedrazzini, Président
Prof. Olivier Staub, Président
Prof. David Gatfield, Directeur de thèse
Prof. Stefanie Jonas, Experte
Prof. Aurelio Teleman, Expert

Lausanne 2024

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Center for Integrative Genomics

Investigating re-initiation mechanisms *in vivo* and *in vitro*

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Romane Meurs

Master en Biologie Moléculaire et Cellulaire et Biochimie
de l'Université Libre de Bruxelles, Belgique.

Jury

Prof. Thierry Pedrazzini, Président
Prof. Olivier Staub, Président
Prof. David Gatfield, Directeur de thèse
Prof. Stefanie Jonas, Experte
Prof. Aurelio Teleman, Expert

Lausanne 2024

UNIL | Université de Lausanne
Faculté de biologie et de médecine

Ecole Doctorale
Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Olivier	**Staub**
Directeur·trice de thèse	Monsieur	Prof.	David	**Gatfield**
Expert·e·s	Madame	Prof.	Stefanie	**Jonas**
	Monsieur	Prof.	Aurelio	**Teleman**

le Conseil de Faculté autorise l'impression de la thèse de

Romane Adèle Meurs

Master en Biochimie et biologie moléculaire et cellulaire, Université Libre de Bruxelles, Belgique

intitulée

Investigating re-initiation mechanisms *in vivo* and *in vitro*

Lausanne, le 22 août 2024

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Olivier Staub

Table of Contents

Acknowledgments

I would like to thank all the people who contributed scientifically or non-scientifically to my PhD thesis. This work would have never been possible without your essential support.

I would first like to express my gratitude to my supervisor, David Gatfield, for his trust and availability. You progressively giving me more autonomy without ever leaving me to flounder alone. You also motivated me everytime I stopped believing in me or in the project so thank you for that as well.

I also thank my thesis committee: Thierry Pedrazzini, Stefanie Jonas, and Aurelio Teleman for their time and feedback throughout the PhD process. Your inputs contributed significantly to the success of my PhD and to the article.

I would like to thank the lab members as well, without who I would have never made it to the end of the PhD. Thank you for the laughs, the lunchtimes, the out-of-lab activities, the scientific support and the friendly atmosphere. Thank you Georgia, Sebastien, Hima, Anupam, more than colleagues, you truly became my labmates and certainly my friends. Thank you also to Lisa, Mara, Gizem, Angelica, Enes for your precious help in the lab and for the chats. Special thanks to the former bioinformaticians, Rene Dreos and Bulak Arpat, for their patience and their crucial help for the analysis of my sequencing results. Finally, thank you to my officemates Lukas and Virginie and all the visitors who distracted and entertained me during the writing of my thesis.

Thank you also to the scientists who directly contributed to the project. Thank you Evan Karousis for sharing your protocol and teaching me how to prepare the translation-competent lysates. Thank you Adrian Bothe who produced several recombinant proteins for my *in vitro* assays. Finally, thank you Michael Taschner for your expertise that permitted to efficiently and rapidly produce the recombinant GADD34.

I would like to thank all the people who, behind the scenes, made my life easier within the CIG by simplifying my research, including the people who work for the facilities, the magasin, the secretaries, etc.

Importantly, I warmly thank my friends and family for supporting me and making this period of my life smoother. My Swiss friends for bringing me around Lausanne and Switzerland and for all the fun activities we did together. My flatmates, specially my friend Sophia, for the chill times at home and in EPFL. My CIG friends, particularly Carlos, Erfan and Maria for the relaxing and adventurous moments shared together. My Belgian friends and my family for visiting me in Switzerland and welcoming me during my numerous trips to Belgium. Thank you for the incredible memories that we now share together.

Finally, I would like to thank Flavio, my amazing partner, who accompanied me for the last 2.5 years of this journey. Your smile, your humor, your tenderness and your unconditional support contributed to the success of this work.

I unfortunately could bot include all your names in the article but you would all deserve it. You were all truly essential to this journey.

Abstract

Cis-regulatory mRNA elements are RNA-encoded features that directly regulate the expression of the transcripts on which they are located. During my PhD, I worked on two *cis*-regulatory elements that control translation efficiency through ribosomal fluxes, upstream Open Reading Frames (uORFs) and frameshifting sites, with a main focus on uORFs which I first introduce in this thesis manuscript.

uORFs are regulatory elements located in the 5' untranslated region (5' UTR) of mRNAs that inhibit translation of the downstream protein-coding sequences (CDS) by capturing initiating ribosomes. Protein expression of uORF-containing transcripts is made possible by leaky scanning and re-initiation, an enigmatic process that allows the terminating ribosome to stay on the transcript after uORFs translation and re-scan the downstream sequence to initiate on the next start site. Re-initiation is regulated by the non-canonical translation factors MCTS1-DENR as well as by canonical initiation factors that have not yet been formally identified. To investigate the function of *trans*-acting factors in re-initiation, I developed a cell-free re-initiation assay based on HeLa lysates. By comparing *in vivo* and *in vitro* re-initiation assays on two DENR-regulated reporters, *Asb8* and *Klhdc8a*, I validated the faithful recapitulation of leaky scanning and MCTS1-DENR-dependent re-initiation *in vitro*. I then investigated the importance of active eIF2 availability for re-initiation and observed that the phosphorylation of eIF2α only mildly affected the re-initiation rate of my reporters. Using my cell-free system and ribosome profiling data, I determined that the MCTS1-DENR homolog, eIF2D, did not regulate expression of uORF-containing transcripts but played a distinct role in cells, displaying only a weak re-initiation activity *in vitro*. I also observed that the previously identified MCTS1 retrocopy, MCTS2, can interact with DENR *in vivo*, and that the heterodimer MCTS2-DENR exhibited re-initiation activity *in vitro*. These results demonstrate that my assay could be used to examine the role of uORF features and canonical initiation factors in re-initiation, opening many possibilities for further investigations.

In the second part of the thesis, I describe the collaborative work on characterizing the -1 programmed frameshifting event on the SARS-CoV-2 RNA. The ribosomal frameshifting is essential for the expression of the viral proteins encoded by the out-of-frame ORF1b. By combining the structural analysis of two different ribosomal frameshifting stages with *in vivo* reporter assays and ribosome profiling, the later being where I contributed, we elucidated the

mechanism of SARS-CoV-2 frameshifting. Additionally, the usage of an inhibitory compound on infected cells confirmed that viral replication could be limited by targeting the frameshifting event.

In summary, this work has significantly contributed to the understanding of two important mechanisms of translation regulation.

Résumé

Les éléments *cis*-régulateurs des ARNm sont des séquences d'ARN qui régulent directement l'expression des transcripts sur lesquels elles se trouvent. Lors de mon doctorat, j'ai étudié deux éléments *cis*-régulateurs qui limitent l'efficacité de traduction en controllant le flux de ribosomes : les upstream Open Reading Frames (uORFs) et les sites de changement de cadre. Ma recherche s'est principalement portée sur les uORFs que je vais introduire en premier dans ce manuscript de thèse.

Les uORF sont des éléments régulateurs situés dans la region 5' non traduite (5' UTR) des ARNm et qui inhibent la traduction de sa séquence codante (CDS) en capturant les ribosomes initiateurs. L'expression des transcripts qui contiennent des uORFs est conditionnée par le leaky scanning et la réinitiation, un processus énigmatique qui permet aux ribosomes en fin de traduction des uORFs de rester sur le transcript et de rescanner la sequence en aval afin d'initier au prochain site de démarrage de la traduction. La réinitiation est régulée par les facteurs de traduction non-conventionnels MCTS1-DENR ainsi que par d'autres facteurs d'initiation conventionnels qui n'ont pas encore été formellement identifiés. Afin d'étudier la fonction de certains facteurs dans la réinitiation, j'ai dévéloppé un systeme de traduction *in vitro* basé sur la production de lysats de cellules HeLa. En comparant l'essai de réinitiation *in vitro* avec celui *in vivo* de deux raporteurs DENR-dépendants, *Asb8* et *Klhdc8a*, j'ai pu valider la fidèle reproduction des processus de leaky scanning et de réinitiation dans mon système *in vitro*. En jouant avec le taux de phosphorylation d'eIF2α, j'ai observé que l'effet de la phosphorylation sur l'efficacité de réinitiation était faible, remettant ainsi en question l'importance de la disponibilité d'eIF2 sur la réinitiation. La combinaison de mon essai *in vitro* avec la technique de profilage des ribosomes m'a permis de déterminer qu'eIF2D, un homologue de MCTS1-DENR, n'était pas un facteur de régulation des transcripts contenant des uORF, mais qu'il jouait plutot un role distinct dans la cellule. Cependant, j'ai pu observer une faible activité de réinitiation *in vitro* pour eIF2D, uniquement dans des cellules déplétées de DENR. J'ai également observé que MCTS2, une rétrocopie de MCTS1, pouvait intéragir avec DENR et ainsi promouvoir la réinitiation. Ces resultats démontrent que cet essai *in vitro* pourrait être utilisé dans des projets futurs afin d'étudier le role de certains facteurs d'initiation de la traduction et de différentes characteristiques des uORFs dans la réinitiation.

Dans la seconde partie de ma thèse, je décris le travail collaboratif mené sur le changement de cadre de lecture programmé du ribosome qui a lieu sur le genome du SARS-CoV-2. Ce changement de cadre de lecture est essentiel pour l'expression des protéines virales encodées par l'ORF1b qui se trouve dans un autre cadre de lecture que l'ORF1a. En combinant l'analyse structurale de deux états diffférents du ribosome lors du changement de cadre de lecture avec des essais de traduction *in vivo* et du profilage de ribosomes, nous avons élucidé le mécanisme de changement de cadre de lecture du SARS-CoV-2. De plus, l'usage d'une molécule inhibitrice sur des cellules contaminées a confirmé la plausabilité de bloquer la réplication des virus en ciblant l'évenement de changement de cadre de lecture.

En conclusion, ce travail de thèse a significativement contribué à l'étude de deux mécanismes fondamentaux de la régulation de la traduction.

Introduction

1. Eukaryotic translation

Ribosomes are the molecular machines that decode the mRNA templates into peptides. This essential process, called cellular translation, is highly energetically demanding for the cell. Indeed, mammalian cells can dedicate up to an estimated 60% of their energy to synthesize the 1 to 10 million ribosomes necessary to produce the proteins that compose 50% of their dry mass (Neurohr and Amon 2020; Harper and Bennett 2016). The amount of proteins produced is critical for the cell as their over-expression or reduction can have a huge impact on cellular homeostasis. For example, the over-expression of several oncogenes due to mutations of the uORFs within their 5' UTRs is associated with malignancy (see section 2) (Schulz et al. 2018; Barbosa, Peixeiro, and Romão 2013).

It was first assumed that transcription of genes to mRNA was the main regulatory process of gene expression and that the amount of mRNA was a good proxy for protein abundance. In the past 20 years, it has become evident that a large part of gene expression regulation takes place during the translation step. Indeed, several studies showed that the mRNA abundance does not consistently correlate with the amount of proteins produced (Gygi et al. 1999; Schwanhäusser et al. 2011), while ribosomal occupancy on mRNAs correlates better with absolute protein abundance, highlighting the importance of post-transcriptional regulation (T.-Y. Liu et al. 2017).

1.1. Translation initiation

The central importance of translation on cellular physiology and its high energetic cost have led to the emergence of numerous regulatory mechanisms, most of which take place during the initiation step. This makes initiation the most complicated step of translation with more than ten initiation factors involved, some of which are multiple subunit complexes.

Briefly, the 40S ribosomal subunit first interacts with the initiation factors eIF1A, eIF1, and eIF3 then secondly with eIF5 and the eIF2-GTP-Met-tRNA$_i^{Met}$ complex (the functions of these initiation factors are summarized in Table 1). This 43S pre-initiation complex is then recruited to the m7G cap of the mRNAs by the eIF4F complex, composed of the cap-binding protein eIF4E, the scaffold protein eIF4G and the RNA helicase eIF4A to form the 48S initiation complex (figure 1). By interacting with polyA-binding proteins (PABPs), the eIF4F complex is stabilized and the mRNA has been proposed to take a special closed-loop conformation that enhances initiation, mRNA stability and facilitates a new round of initiation on the transcript by close proximity with the recycled 40S. The formed 48S initiation complex then scans the so-called 5' untranslated region (UTR) of the mRNA in an open conformation. During scanning, the helicases eIF4A, together with the factor eIF4B or eIF4H, unwind the mRNA secondary structures. Scanning continues until recognition of a start codon (AUG or non-AUG), corresponding to the Met-tRNA$_i^{Met}$ present in the P-site of the small ribosomal subunit, which triggers the hydrolysis of the eIF2-bound GTP by eIF2 after its activation by eIF5. (Jackson, Hellen, and Pestova 2010; Brito Querido et al. 2020; Kumar, Hellen, and Pestova 2016; Passmore et al. 2007; Algire, Maag, and Lorsch 2005).

The start codon recognition efficiency is influenced by several factors including the start codon identity, the flanking sequences and secondary structures downstream of the AUG. The near-cognate codons are recognized by the 48S scanning complex thanks to the increased

flexibility of the base-pairing between the 3rd base of the codon and the 1st of the anticodon which allows selection of alternative codons including CUG, GUG and UUG with a reduced efficiency (Brito Querido, Díaz-López, and Ramakrishnan 2024; Andreev et al. 2022; Cao and Slavoff 2020). The second feature that can enhance start codon recognition is the flanking sequence of the AUG known as the Kozak sequence. The optimal context CRCCaugG (where R is a purine) increases initiation efficiency thanks to the interaction of the nucleotides with eIF1A, eIF2α and the 18S ribosomal mRNA (Kozak 1984a; 1981; 1984b). Finally, RNA secondary structures downstream of an AUG can slow down the scanning complex progression and enhance initiation on this particular start codon (Xiang et al. 2023; Kozak 1990; Wang et al. 2022).

Table 1: Functions of the 43S pre-initiation complex factors (reviewed in (Brito Querido, Díaz-López, and Ramakrishnan 2024).

Factors	Function	References
eIF1A	- Induces and stabilizes the open conformation - Participates in accurate recognition of the start codon	Pestova and Kolupaeva 2002; Passmore et al. 2007; Fekete et al. 2007
eIF1	- Induces and stabilizes the open conformation - Participates in accurate recognition of the start codon - prevents premature eIF5-induced hydrolysis of eIF2-bound GTP and Pi release	Pestova and Kolupaeva 2002; Thakur and Hinnebusch 2018; Passmore et al. 2007
eIF3	- Assembly of the 43S complex - Recruitment of the 43S complex to the cap associated to eIF4F - Start-site selection - Stabilizes the binding of the ternary complex eIF2–GTP–Met-tRNA$_i^{Met}$ - Interacts with ABCE1 to promote ribosome recycling	Brito Querido et al. 2020; Aitken et al. 2016; Hinnebusch 2014; D. J. Young and Guydosh 2019; Kumar, Hellen, and Pestova 2016; She, Luo, and Weissman 2023
eIF5	- Promotes GTP hydrolysis by eIF2	Das, Ghosh, and Maitra 2001; Paulin et al. 2001
eIF2	- Mediates ribosomal recruitment of Met-tRNA$_i^{Met}$	Jackson, Hellen, and Pestova 2010; Dorris, Erickson, and Hannig 1995

Upon start codon recognition, eIF1 is displaced and the release of the Pi produced by the hydrolysis of the eIF2-bound GTP triggers the removal of eIF2-GDP and eIF5. The restructuring of the initiation complex induces the closed conformation of the ribosome and provokes the recruitment of eIF5B-GTP, a factor responsible for the binding of the 60S ribosomal large subunit to the 40S to create the 80S. The hydrolysis of the eIF5B-bound GTP into GDP+Pi provokes the concomitant release of eIF5B and eIF1A and the disposition of the aminoacyled end of the Met-tRNA$_i^{Met}$ in the peptidyl transfer center of the 60S subunit. This last event ends the translation initiation step and initiates elongation (Shirokikh and Preiss 2018; Lapointe et al. 2022; B. Y. Huang and Fernández 2020; Nanda et al. 2013; Hashem and Frank 2018).

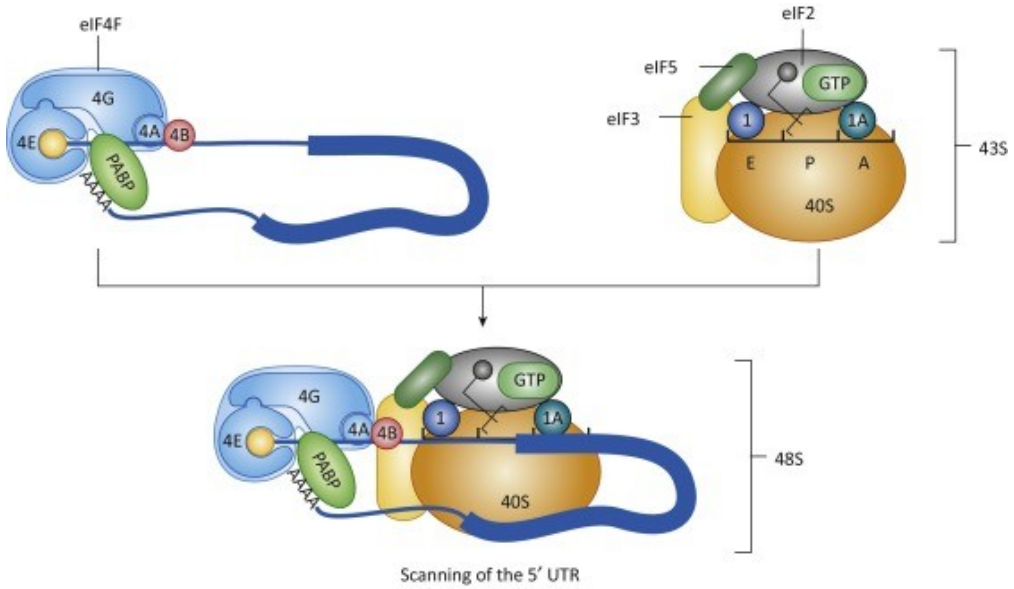


Figure 1 - Schematic of 43S pre-initiation complex and eIF4F complex assembly into the 48 initiation complex (Chu et al. 2016).

1.1.1 eIF2α regulation upon stress

Due to the high energy consumption of translation, translation initiation must be regulated in order to avoid waste of energy or accumulation of unwanted or harmful proteins. Typically amino acid deprivation, endoplasmic reticulum stress due to accumulation of unfolded proteins, detection of heavy metals or dsRNA, a sign of viral infection, trigger activation of specific kinases (GCN2, PERK, PKR and HRI1). These kinases will then phosphorylate the α subunit of eIF2 which leads to a global translation inhibition, but still allows expression of a specific subset of genes involved in the cellular strategy to recover from stress. For instance, eIF2α phosphorylation is known to increase expression of ATF4 and ATF5, two transcription factors inducing expression of a subset of genes involved in appropriate response to the stress (Watatani et al. 2008; Lu, Harding, and Ron 2004; Vattem and Wek 2004).

To return to pre-stress conditions, Growth Arrest and DNA Damage-inducible protein 34 (GADD34) will target eIF2α and recruit the Protein Phosphatase 1 (PP1) to induce dephosphorylation of eIF2α, which releases translation inhibition (Rojas, Vasconcelos, and Dever 2015; Hicks et al. 2023).

1.2. Translation elongation

At the beginning of the elongation phase, most of the initiation factors are already released from the ribosome as a result of the start codon recognition. However, it has been observed that eIF3, eIF4E and eIF4G remain associated with early elongating ribosomes before being detached (J. Bohlen, Fenzl, et al. 2020; Mohammad et al. 2017).

During elongation, eEF1A-GTP interacts with aminoacyl-tRNAs to bring them into the ribosomal A-site. If the anticodon of the charged tRNA corresponds to the codon of the mRNA, the GTP is hydrolyzed and the eEF1A-GDP complex is released. The factor eEF1B

participates to the recycling of eEF1A-GDP by promoting the exchange of GDP to a GTP. The catalysis of the peptide bond formation by the ribosome and eIF5A creates a hybrid state where the acceptor arm of the aminoacyl-tRNA shifts from the A-site to the P-site while the acceptor arm of the tRNA attached to the nascent chain is pushed from the P-site to the E-site. On top of its role in the peptide bond formation, eIF5a promotes translation efficiency of polyproline residues and is involved in initiation by regulating start codon selection (Dever, Dinman, and Green 2018; Xu, Liu, and Song 2022; Gutierrez et al. 2013).

eEF2-GTP binds to the ribosomal A-site to promote faithful translocation of the full tRNAs to the P-site and the E-site thanks to a post-translationally modified histidine called diphthamide. Indeed, removal of the diphthamide amino acid reveales its essential role to prevent programmed -1 ribosomal frameshifting in mammalian cells (S. Liu et al. 2012). The elongation cycle ends with the eviction of the deacylated tRNA from the E-site.

1.3. Translation termination and recycling

Translation termination is triggered by the entry of a stop codon (UAG, UAA or UGA) into the A-site of the ribosome. Stop codons do not have matching aminoacylated-tRNAs and are instead recognized by the release factor eRF1. eRF1 together with ABCE1 induces the release of the synthesized peptide from the P-site through a GTP hydrolysis stimulated by the GTPase eRF3. The two ribosomal subunits are then split and recycled by ABCE1 while the mRNA and deacylated tRNA are released from the 40S subunit (Hellen 2018; Brito Querido, Díaz-López, and Ramakrishnan 2024). It has been shown in *in vivo* studies in yeast and mostly *in vitro* assays in mammals that the removal of the tRNA from the P-site is promoted by the heterodimer MCTS1-DENR and its homolog eIF2D (D. J. Young et al. 2018; J. Bohlen, Harbrecht, et al. 2020; Skabkin et al. 2010).

1.4. Contribution of my PhD research to understanding the role of RNA regulatory elements and translation factors in translation

During my PhD, I worked on two particular RNA features that regulate translation of downstream ORFs: uORFs and frameshifting sites. My main project focused on understanding the mysterious mechanism of re-initiation that allows the ribosome to reach the CDS after uORFs translation and how MCTS1-DENR promotes this process. A preprint version of the article has been prepared and will soon be submitted to a journal and to the server bioRxiv. Few months after the beginning of my thesis, the global pandemic disrupted the course of my doctorate and gave me the opportunity to participate in a project on the programmed frameshifting event that takes place on the SARS-CoV-2 RNA genome. That successful collaborative project concluded with the publication of the results in 2021 (Bhatt et al. 2021).

In this introduction, I will first talk about uORFs and the re-initiation process that occupied most of my thesis work. I will then present the main outcomes of this project that I was in the process of publishing at the time of the submission of the thesis manuscript. In the second part, I will briefly introduce the usage of the cell translation machinery by SARS-CoV-2 to express its genome and focus on the ribosomal frameshifting event. Finally, I will summarize the results of the published collaborative work.

2. uORFs

Upstream open reading frames (uORFs) are short mRNA elements (typically shorter than 30 codons in mammals) found within the 5' 'untranslated' regions (5' UTRs) of eukaryotic transcripts. They are characterized by a start codon and an in frame stop codon but do not generally code for functional peptides (Wethmar 2014; Gunišová et al. 2017). Overlapping uORFs represent a specific category of uORFs that run into the CDS region but in a different frame. Almost 50% of the human transcriptome contains one or multiple uORFs within their 5' UTR whereas only ~13% of transcripts have uORFs in yeast (Calvo, Pagliarini, and Mootha 2009; Wethmar 2014). The regulatory mechanisms of uORF-containing transcripts appear to differ partially between human and yeast, and I will focus on mammalian uORFs in this introduction.

Establishment of ribosome profiling experiments, consisting in sequencing of the ribosomal footprints to monitor ribosomal occupancy, revealed that many uORFs were translated and conserved in mouse and human (Ingolia, Lareau, and Weissman 2011; Bazzini et al. 2014; S. Lee et al. 2012). Translated uORFs inhibit expression of the CDS as most of the ribosomes are recycled at the uORF stop codon and do not reach the start codon of the CDS. Several studies showed that the amino acid sequences of most uORF-encoded peptides are not conserved and that most synthesized peptides are quickly degraded in the cell, indicating that uORF-dependent regulation is driven by their translation rather than by the produced peptide (Johnstone, Bazzini, and Giraldez 2016; Iacono, Mignone, and Pesole 2005; Plaza, Menschaert, and Payre 2017; H. Zhang et al. 2021), with some exceptions of active microproteins (Chen et al. 2020; Wright et al. 2022).

By inhibiting the translation of downstream ORFs, uORFs participate in tuning the translation efficiency of CDS under physiological and stress conditions. Indeed, it has been observed that

deregulation of uORF-dependent translation is responsible for several disease phenotypes such as cancer and diabetes (Schulz et al. 2018; Occhi et al. 2013; A. E. von Bohlen et al. 2017; Hornig et al. 2016; Barbosa, Peixeiro, and Romão 2013; D. S. M. Lee et al. 2021).

CDS expression of uORF-containing transcripts results from the combination of two translational and one translation-coupled mechanism: leaky scanning, re-initiation and nonsense-mediated decay.

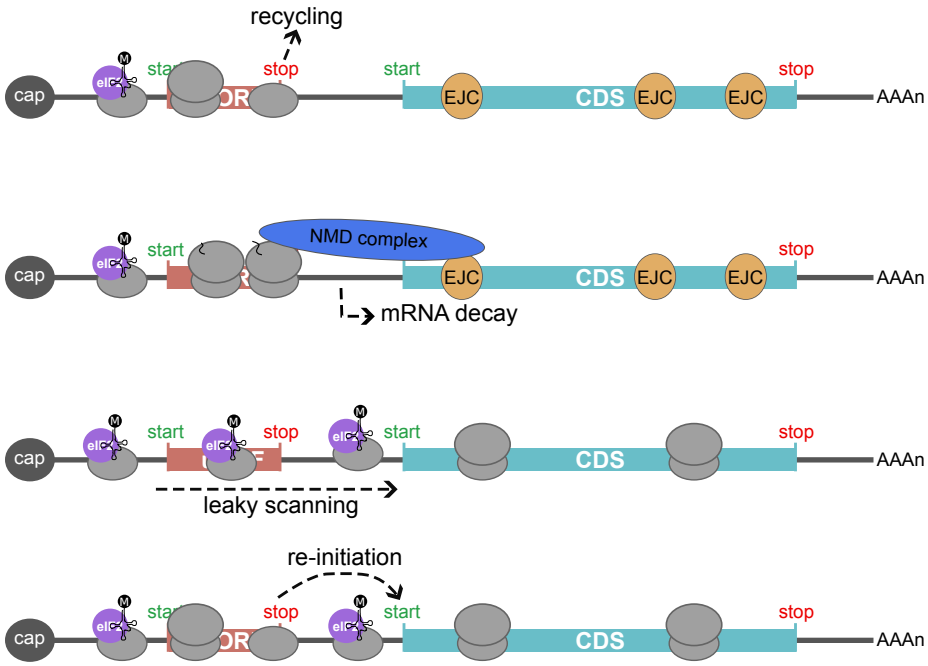


Figure 2 - Schematics of ribosomal fluxes on the 5' UTR of uORF-containing transcripts. The ribosome can be fully recycled once it reaches the uORF stop codon, inhibiting completely translation of the CDS. The terminating ribosome and the EJCs found on the CDS can also be recognized by the NMD factors which will trigger mRNA decay. The pre-initiation scanning complex can also bypass the uORF start codon by leaky scanning and initiate at the CDS start codon. Finally, if the uORF is translated, the terminating ribosome can also only be partially recycled at the uORF stop codon, letting the 40S reach the CDS start codon through re-initiation.

2.1. Nonsense-mediated decay

On top of inhibiting CDS expression by deviating initiation ribosomes, uORFs can also influence the stability of the transcripts through activation of nonsense-mediated decay (NMD). Briefly, NMD is a surveillance mechanism that recognizes and destabilizes transcripts with premature termination codon (PTC) within their coding sequence. The canonical trigger of NMD in mammals are exon-junction complexes (EJCs) that are deposited during splicing and are removed by the ribosomes during elongation. If a ribosome terminates on a PTC, the NMD complex will recognize a downstream EJC and interact with it, leading to the induction of a cleavage and further decay of the transcript (Hug, Longman, and Cáceres 2016).

If a uORF is translated by the first loaded ribosome, downstream EJCs on the CDS are not removed by elongating ribosomes and still interact with the transcript when the ribosome terminates on the uORF stop codon. This situation is similar to the presence of a PTC and should trigger NMD and destabilize the transcript. Interestingly, not all uORF-containing transcripts are sensitive to NMD to the same extent. First, weak start site uORFs should be less affected by NMD than strong start site uORFs as the first loaded ribosome can bypass the weak uORF start site to translate the CDS and remove the EJCs. It has been observed that only AUG uORFs indeed destabilize the transcripts, enhanced by the stalling of ribosomes upstream of the terminating ribosome (Arribere and Gilbert 2013; Gaba, Jacobson, and Sachs 2005; Colombo et al. 2017). Some papers showed that PABP can interact with eRF3, which enhances termination and decreases NMD rate for PTCs close to an AUG, a criteria met by shorter uORFs (Peixeiro et al. 2012; P. V. Ivanov et al. 2008). It has also been proposed that re-initiation could protect the transcript from NMD, but the published results are inconsistent (Russell et al. 2023; J Zhang and Maquat 1997; Romão et al. 2000).

2.2. Leaky scanning

During scanning of the 5’ UTR, some 48S ribosomal complexes bypass the uORF start codon and directly initiate at the next start site in a process called leaky scanning. The leaky scanning rate depends on the strength of the uORF start site context and on the abundance of canonical initiation factors implicated in start codon selection such as eIF1 and eIF5 (J. Bohlen, Roiuk, et al. 2023; Shestakova et al. 2023; Kozak 2002).

EIF4G2 is a non-canonical initiation factor, homolog to the scaffold factor eIF4G1. Unlike eIF4G1, it lacks the ability to interact with eIF4E and PABP but binds to eIF4A, eIF3 and eIF2 (Sugiyama et al. 2017; Bryant et al. 2018). Thus, eIF4G2 can be directly recruited to the pre-initiation scanning complex, independently of the cap-binding factor eIF4E. It was recently discovered that eIF4G2 interacts with the PRRC2 proteins (PRRC2A, PRRC2B and PRRC2C) to contribute to efficient translation of uORF-containing transcripts by promoting leaky-scanning (Weber et al. 2022; Shestakova et al. 2023; Sugiyama et al. 2017; J. Bohlen, Roiuk, et al. 2023; Smirnova et al. 2022).

In the case of overlapping uORFs, the ribosome cannot re-initiate as backward movements of ribosomes (3′ to 5′) during scanning are rare and short-lived (Wang et al. 2022; Kozak 2001). As a consequence, transcripts with overlapping uORFs are fully reliant on leaky scanning for expression of their CDS.

2.3. Re-initiation

Classically, ribosomes that translate uORFs are recycled at the uORF stop codon following the procedure described in 1.3. However, it appears that some ribosomes are only partially recycled, allowing a re-scanning complex to form and scan the downstream sequence. This process, called re-initiation, leads to the recognition of a downstream start site and expression of the CDS (Gunišová et al. 2017; Wethmar 2014).

The translation factors involved in the re-scanning and re-initiation processes are still uncertain. The recruitment of a new Met-tRNA$_i^{Met}$ by eIF2 (or possibly by alternative non-canonical factors) is necessary for re-initiation on the CDS. When eIF2 is inactivated by phosphorylation, the elevated ratio of eIF2α-p/eIF2α reduces the amount of available eIF2α-Met-tRNA$_i^{Met}$ and increases its recruitment time to the re-initiation complex. The intercistronic distance between the uORF stop codon and the CDS start codon then becomes a limiting feature for expression of the CDS. This effect is well-documented for *Atf4* and *Atf5* transcripts, containing of a short uORF upstream to an overlapping uORF that strongly inhibits expression of the CDS under physiological conditions. Upon stress, the re-scanning ribosomes bypass the overlapping uORF start codon and initiate at the CDS start site, which leads to high translation of ATF4 and ATF5 (Vattem and Wek 2004; Watatani et al. 2008, 5; Gunišová et al. 2017). Logically, this mechanism should also apply to other uORF-containing transcripts and positively or negatively regulate their expression in stressed cells.

Beside eIF2, eIF1, eIF1A and eIF5 could be part of the re-scanning complex as absence of these factors would reduce the fidelity of start codon selection (I. P. Ivanov et al. 2017; Fekete et al. 2007; Passmore et al. 2007). But one can imagine that the relatively short distance between the uORF and the CDS would decrease the chances of finding a non-CDS weak start

site. Interestingly, an eIF1-independent re-initiation event would promote re-initiation on suboptimal CDS start site (Gunišová et al. 2017).

While it was first assumed that the re-initiation factors are recruited directly to the re-scanning complex, another hypothesis proposes that some initiation factors could remain attached to the elongation complex during translation of the short uORFs. Indeed, it was recently observed that eIF4G1, eIF4E and eIF3 (or at least the eIF3e and eIF3b subunits) remain associated with elongating ribosomes for a short period (with a decay of < 60 codons) before being released (J. Bohlen, Fenzl, et al. 2020; Mohammad et al. 2017; Lin et al. 2020). In this regard, uORFs encoding 1aa long peptides (also called start-stop uORFs) should have an advantage and undergo more efficient re-initiation.

Concerning the recycling step at the uORF stop codon, it has been proposed that two non-canonical translation factors, MCTS1-DENR and its homolog eIF2D, are necessary for the removal of tRNAs from the 40S P-site to promote re-initiation (see sections 2.3.1 and 2.3.3).

Beside the regulatory effects of translation factors, some uORF-containing transcript features have been associated with higher or lower re-initiation efficiency. First, uORF secondary structures as well as uORF length decrease re-initiation, probably because a longer elongation duration increases the risk of loosing the initiation factors that are still interacting with the elongating ribosomes (Kozak 2001; Russell et al. 2023). Intercistronic distance between the uORF stop codon and the CDS start codon may also influence the time for recruitment of initiation factors to the re-scanning complex and therefore the chances of re-initiation (Kozak 1987).

2.3.1 MCTS1-DENR-dependent re-initiation

MCTS1 (Malignant T-Cell Amplified Sequence 1, previously referred to as MCT-1) and DENR (Density-regulated protein, previously named DRP) are two non-canonical translation factors that interact together to form a heterodimer. They are co-dependent on each other for stability as depletion of one of the two factors destabilizes the other one at the protein level (Schleich et al. 2017; 2014).

MCTS1 is composed of a DUF1947 domain that binds to the 40S and a PUA domain, a widespread domain that typically binds dsRNA, that interacts with the CCA tail of tRNAs. DENR contains a SWIB/MDM2 to interact with the acceptor arm of tRNAs and a SUI1 domain analogous to eIF1 that binds to the small ribosomal subunit (figure 3) (Weisser et al. 2017; Skabkin et al. 2010; Lomakin, Dmitriev, and Steitz 2018). Together, MCTS1 and DENR can promote recruitment of aminoacylated-tRNAs to the P-site of the small ribosomal subunit as well as release of deacylated-tRNAs after ABCE1-mediated dissociation of post-termination ribosomes *in vitro* (Skabkin et al. 2010). *In vivo*, TMA20-TMA22, the yeast orthologs of MCTS1-DENR, is involved in the 40S subunit recycling step and inhibits re-initiation (D. J. Young et al. 2018) while the heterodimer promotes re-initiation in *Drosophila* and mammals (Schleich et al. 2014; 2017). MCTS1-DENR was also shown to promote Repeat-associated non-AUG (RAN) translation (Green et al. 2022).

It appears that only a small portion of uORF-containing transcripts undergo MCTS1-DENR-dependent re-initiation (J. Bohlen, Harbrecht, et al. 2020; Castelo-Szekely et al. 2019). At first, short start-stop uORFs were identified as specific MCTS1-DENR targets (Schleich et al. 2017). However, it was shown by a former PhD student in the lab that DENR was also required for the CDS translation of transcripts containing longer uORFs (Castelo-Szekely et al. 2019). She determined that DENR-dependent transcripts are enriched for AUG uORFs

with strong Kozak contexts, have a rather low GC content and are enriched for high number of uORFs. Longer intercistronic distance between the uORF and the CDS was also associated with MCTS1-DENR-dependent re-initiation. This last result may be explained by the time needed to recruit the new translation factors involved in start site recognition (Castelo-Szekely et al. 2019). To uncover specificity of MCTS1-DENR regulated transcripts, Bohlen et al. investigated the influence of uORF stop codon context with the intent to reconcile both MCTS1-DENR re-initiation and recycling functions. They found that certain uORF penultimate codons are overrepresented in DENR-regulated transcripts and proposed that MCTS1-DENR participate in recycling of some deacylated-tRNA that are more strongly binding to the P-site of the 40S thus promoting downstream re-initiation (J. Bohlen, Harbrecht, et al. 2020). This interesting hypothesis deserves further investigation.

MCTS1 was first described as an oncogene that promotes cell cycle progression (Prosniak et al. 1998). Expression of its partner, DENR, was initially shown to increase with cell density (Deyo, Chiao, and Tainsky 1998). A direct link has been recently made between MCTS1-DENR and the cell cycle. Transient phosphorylation of DENR by Cyclin B/CDK1 and Cyclin A/CDK2 during early mitosis increases its stability and therefore promotes translation of uORF-containing transcripts, some of which are involved in the cell cycle (Clemm von Hohenberg et al. 2022).

While MCTS1-DENR has long been presented has an obligate heterodimer, it was recently observed that *Mcts2* (chromosome2 in mouse and chromosome 20 in human), an intronless retrogene copy of *Mcts1* (X chromosome), was expressed in mouse cell lines and tissues (McCole et al. 2011; Zhijun et al. 2014; Wood et al. 2007). However, MCTS2 can only partially rescue *Mcts1* knock-out re-initiation defect in HeLa cells and it does not seem to be

expressed in several human cell types, suggesting that MCTS2 may not be non-functional (J. Bohlen, Zhou, et al. 2023).

2.3.2 Clinical relevance of MCTS1-DENR-dependent translation

MCTS1-DENR controls expression of important proteins involved in key cellular processes. Among others, MCTS1-DENR regulates translation efficiency of oncogenes such as *a-Raf*, *c-Raf*, *Atf4* and *Cdk4* through their re-initiation activity (J. Bohlen, Harbrecht, et al. 2020; Vasudevan et al. 2020). MCTS1 itself is an oncogene that contributes to malignancy and tumor development (Hsu, Shi, and Gartenhaus 2005; Shi et al. 2003; Levenson et al. 2005; Dierov et al. 1999; Mazan-Mamczarz et al. 2009; Z. Huang et al. 2021; Gao et al. 2021).

DENR also participates in neuronal development as misregulation of DENR is associated with neuronal diseases. Two reported missense mutations in the *Denr* sequence disrupt its translation function, which is detrimental to the development of cortical neurons and causes autism spectrum disorders (Haas 2016; D. J. Young, Meydan, and Guydosh 2021). Furthermore, Parkinson's disease in several patients was associated with low expression of DENR (Wu et al. 2024).

Another syndrome, the Mendelian susceptibility to mycobacterial disease (MSMD), has recently been linked to mutations in the *Mcts1* gene in several independent patients. All these mutations were shutting down expression of MCTS1 and, hence, protein biosynthesis from MCTS1-DENR-regulated transcripts (J. Bohlen, Zhou, et al. 2023). Interestingly, full knock-out of DENR but not of MCTS1 is embryonically lethal in mice as shown by the International Mouse Phenotyping Consortium (IMPC), suggesting that DENR could have an alternative partner, MCTS2 for instance, that could compensate for the lack of MCTS1 or that DENR could play an additional role in cells.

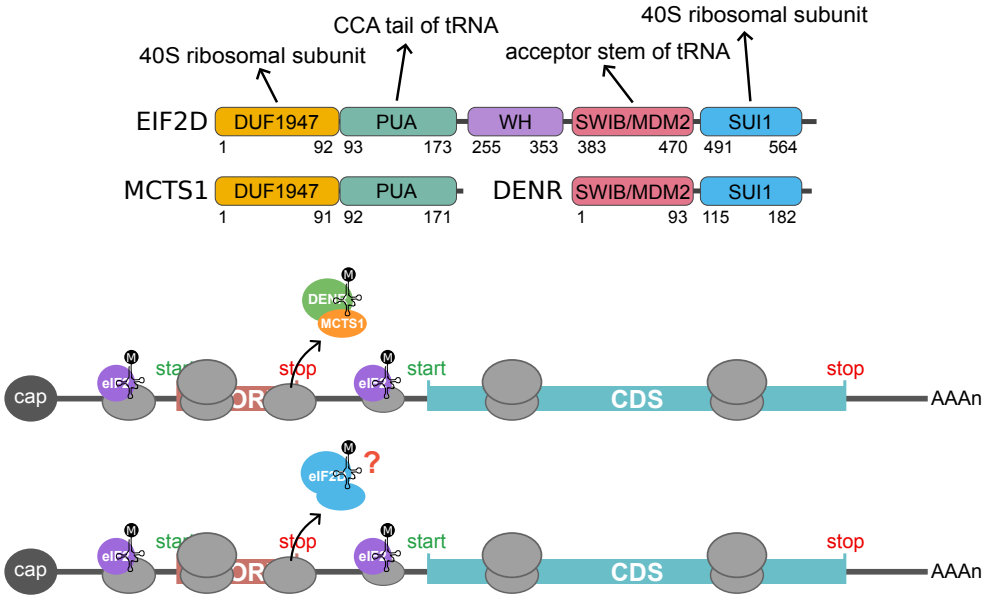


Figure 3 - Schematics of the domains of MTCS1, DENR and eIF2D and of MCTS1-DENR-dependent re-initiation. MCTS1-DENR would promote re-initiation by recycling the deacylated tRNA of more strongly binding tRNAs from the P-site of the 40S. The role of eIF2D in re-initiation requires further investigations.

2.3.3 eIF2D

eIF2D is a homolog of the heterodimer MCTS1-DENR encoded by a single peptide (figure 3). It was first named Ligatin (LGTN), which led to a confusion between eIF2D and the actual Ligatin, a membrane trafficking receptor for phosphoglycoproteins (Jakoi and Marchase 1979; Marchase et al. 1982; Aravind and Koonin 1999).

In parallel to MCTS1-DENR, Skabkin et al. and Dmitriev et al. identified eIF2D as a non-canonical translation factor capable of promoting the recruitment, including at near-cognate codons, and release of charged and uncharged tRNAs from the P-site of the 40S *in vitro*, suggesting that eIF2D could act as a recycling factor or alternative tRNA recruiter (Skabkin et al. 2010; Dmitriev et al. 2010). Structures of eIF2D revealed that it interacts with the ribosome in a similar fashion as MCTS1-DENR. They also uncovered that eIF2D does not

directly recruit the tRNA to the P-site but instead affects ribosomal conformation, allowing tRNA-small subunit interactions. Furthermore, they observed that the position of eIF2D on the 40S subunit prevents the recruitment of a 60S subunit, in line with an initiation or 40S recycling function (Weisser et al. 2017; Vaidya 2017). In yeast, its ortholog TMA64 was first described as a recycling factor, like the heterodimer TMA20-TMA22 (D. J. Young et al. 2018). However, it was then observed, using a single *Tma64* KO strain, that it plays at best a minor role in recycling (D. J. Young, Meydan, and Guydosh 2021). In *Drosophila* and human cells, eIF2D is considered as a re-initiation factor, similarly to MCTS1-DENR, mostly due to their structural resemblance (Weisser et al. 2017). Indeed, eIF2D also regulates ATF4 expression in these two model organisms (J. Bohlen, Harbrecht, et al. 2020; Vasudevan et al. 2020), but it does not promote translation of an MCTS1-DENR-dependent start-stop uORF reporter, suggesting that it could regulate other uORF-containing transcripts (Schleich et al. 2017). Moreover, a case of eIF2D-dependent re-initiation on a subgenomic bicistronic calcivirus mRNA constructs has been studied by Zinoviev et al. *in vitro* (Zinoviev, Hellen, and Pestova 2015).

Based on the *in vitro* assays of Skabkin et al. and Dmitriev et al., it was suggested that eIF2D could act as a GTP-Met-tRNA$_i^{Met}$ recruitment factor for near-cognate start codons. Supporting this hypothesis, eIF2D has been shown to be also involved in RAN translation of toxic poly-GA dipeptide repeats in *C. elegans* and mammalian cells, but this result could not be reproduced in D*rosophila* (Sonobe 2021; Green et al. 2022). However, eIF2D depletion does not affect recognition of near-cognate codons in a human cell line (Ichihara et al. 2021). Instead, the canonical initiation factor eIF2 seems to be the main regulatory factor of these non-canonical initiation events. Following this idea, other laboratories have been interested in the function of eIF2D in the translation of viruses transcripts. Indeed, expression of viral

proteins often depends on non-canonical initiation event when eIF2 is inactivated, such as IRES-dependent initiation or non-AUG initiation. While one group concluded on an eIF2D-dependent expression of the enteroviral genome (H. Kim et al. 2023), others observed that eIF2D was not able to replace eIF2 for initition at non-AUG codons or IRES on different viral genomes (Sanz, González Almela, and Carrasco 2017; Sanz et al. 2019; González-Almela et al. 2018).

Finally, eIF2D was shown to inhibit *Hoxa3* initiation via its 5' UTR translation inhibitory element (TIE) in mammalian cells, proposing a novel translation inhibitory function for eIF2D (Alghoul et al. 2021).

In short, the currently available data are inconsistent and do not allow us to conclude on a defined role for eIF2D, suggesting that further investigations should be conducted to uncover its function.

2.4. Context of my doctoral project

When I started my PhD, the article of the former PhD student Violeta Castelo-Szekely about DENR-dependent re-initiation was just published. Based on her ribosome profiling results and on the modest literature at the time about eIF2D, I was interested in understanding whether eIF2D and MCTS1-DENR have the same functions and regulate translation of the same transcripts. The strategy used was based on the ribosome profiling of *Eif2d* knock-down vs. *Denr* knock-down cells to identify and compare eIF2D and DENR target transcripts. Analysis of the ribosome profiling data revealed that eIF2D and DENR depletion had differential effects on translation. Interestingly, eIF2D-depletion modified abundance of thousands of transcripts while only influencing translation efficiency of less than a hundred mostly uORF-free transcripts. I therefore attempted to investigate eIF2D function by exploring the identified presumed targets and the phenotype of *Eif2d*-depleted cells but was unsuccessful. One strategy consisted in performing RNA-seq of *Eif2d* knock-down cells vs. WT to understand if eIF2D was regulating transcription or RNA stability of the target mRNAs and concluded that the stability was most affected (data not shown).

In parallel to this first project, I initiated a side project to set up an *in vitro* translation assay in order to understand ribosome fluxes on the 5' UTR of selected uORF-containing reporters by measuring leaky scanning rate, re-initiation rate, DENR-dependent re-initiation rate and uORF inhibition. The first attempts were achieved using commercial rabbit reticulocyte lysates and gave promising results but the main limitation of this set-up was that initiation is mainly cap-independent in this cell-free system. The next strategy consisted in producing translation-competent HeLa lysates from WT and *Denr* knock-out cells (cells kindly provided by my committee expert Aurelio Teleman after the mid-thesis discussion). The *in vitro* lysates were produced in-house using a protocol newly developed by Lukas-Adrian Gurzeler and

Evangelos Karousis from Oliver Mühlemann's lab and shared with me before publication. After some adjustments, the assay gave promising results and this second project became my main project for the last two years of my PhD.

2.5. An in vitro assay of MCTS1-DENR-dependent re-initiation and ribosome profiling uncover the activity of MCTS2 and distinct function of eIF2D. (Meurs et al, 2024): main results

In order to investigate the role of different mRNA features and factors on re-initiation, we attempted to set-up an *in vitro* translation assay to measure leaky scanning and re-initiation rate on uORF-containing transcripts. To design reporters whose expression was re-initiation-susceptible, we selected two DENR-dependent transcripts from a ribosome profiling experiment, *Asb8* (start-stop uORF) and *Klhdc8a* (13aa long uORF), which both contain a translated and conserved AUG uORF in mouse and human cells (Meurs et al., figure 1.F and supplementary figure 1.B). *In vivo* reporter assays using the *Klhdc8a* 5' UTR revealed that its expression was mostly re-initiation-dependent (with 19% of ribosomes that reached the CDS thanks to re-initiation) with very low leaky scanning rate (2%). By using *Denr* KO HeLa cells, we could also evaluate the part of re-initiation that was DENR-dependent, which was of ~85% in this set-up (Meurs et al., figure 1.I and 1.K).

Using homemade HeLa cell-free extracts, we reproduced these results *in vitro* with very similar outcomes, with slightly reduced re-initiation rate (11%; Meurs et al., figure 2.E and 2.F). To check if mutations of the uORF penultimate codon would influence the re-initiation rate or/and the DENR-dependence, I substituted the uORF penultimate codon of *Klhdc8a* for an enriched or a depleted codon of DENR-regulated transcripts according to Bohlen at al. While there was no difference in the case of the enriched codon, the depleted codon construct displayed a reduced DENR-dependence but an unaffected re-initiation rate (Meurs et al., figure 2.I).

In vitro, the *Asb8* reporter displayed a high leaky scanning rate (58%) due to the rather weak Kozak context of start-stop uORFs and a high translation rate (29%, figure 2.G and 2.H).

Insertion of a 2nd codon to the *Asb8* 1aa uORF did not reduce the re-initiation rate but the TGT penultimate codon (slightly enriched in DENR-targets) reduced DENR-dependence (Meurs et al., figure2.J).

In order to investigate the effect of eIF2α phosphorylation on re-initiation, we carried out the *in vitro* assay in the absence of GADD34, an eIF2α dephosphorylation factor that we use in all the *in vitro* reactions to prevent global inhibition of translation. We observed non significant increase of leaky scanning and reduction of re-initiation for both *Klhdc8a* (figure 3.A and 3.B) and *Asb8* reporters (Meurs et al., figure 3.C and 3.D), suggesting that the strong phosphorylation of eIF2α only mildly affected the recruitment of a new GTP-Met-tRNA$_i^{Met}$. We repeated the experiment in HeLa WT and *Denr* KO cells treated with tunicamycin and transduced with the *Klhdc8a* reporters. *In vivo*, *Klhdc8a* reporters displayed a weak but significant reduction of re-initiation rate (Meurs et al., supplementary figure 3.B). Of note, DENR depletion did not decrease translation of the endogenous transcripts and *Atf4* and *Atf5* *in vivo* reporters under physiological or stress conditions but provoked an accumulation of footprints on the stop codon of uORF1, consistent with the accumulation of ribosomes due to a recycling defect (Meurs et al., supplementary figure 5 and figure 4).

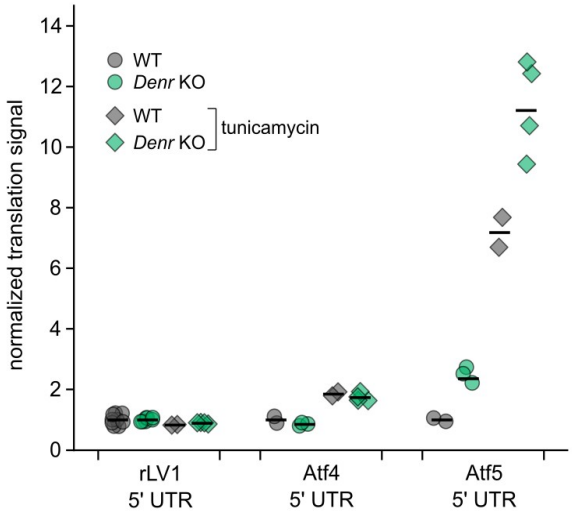


Figure 4 - An *in vivo* dual glo luciferase reporter assay of Atf4 and Atf5 5' UTR reporters in WT and Denr KO HeLa cells treated for 14 h with 1ug/ml tunicamycin or untreated confirmed the induction of the two reporters and revealed that their expression was not reduced upon DENR depletion.

Due to its resemblance with MCTS1-DENR, the non-canonical translation factor eIF2D is considered as a re-initiation factor. However, as described in the introduction, its function remains uncertain as current studies reached different conclusions. To identify eIF2D targets and define if eIF2D depletion had the same effect as DENR depletion on translation, I performed a ribosome profiling experiment on *Eif2d* vs. *Denr* knock-down NIH/3T3 cells. 73 transcripts displayed reduced CDS translation efficiency upon eIF2D depletion while 223 where affected by DENR depletion (Meurs et al., figure 4.C and 1.B). eIF2D-dependent transcripts were distinct from the ones regulated by DENR (supplementary figure 4.A) and were not enriched for translated uORFs (Meurs et al., figure 4.D).

Using our *in vitro* re-initiation assay, we observed a reduction in the re-initiation rate only on *Klhdc8a* 5' UTR in *Eif2d* knock-out HeLa extracts but could not rescue it with recombinant eIF2D (Meurs et al., figure 4.H). Surprisingly, addition of recombinant eIF2D could partially rescue re-initiation in *Denr* knock-out (Meurs et al., figure 4.I) and in double *Denr/Eif2d*

knock-out extracts (figure 4.J). These results suggest that eIF2D is not per se an efficient re-initiation factor *in vivo* but it can promote re-initiation *in vitro* when DENR is absent.

I used again the *in vitro* system to investigate if eIF2D could replace the canonical factor eIF2 to recruit a new GTP-Met-tRNA$_i^{Met}$ during re-scanning, The concomitant depletion of eIF2D and phosphorylation of eIF2α did not further affect expression of *Klhdc8a* reporters while addition of recombinant eIF2D did not rescue the re-initiation signal, suggesting that eIF2D is not acting as an alternative GTP-Met-tRNA$_i^{Met}$ recruiting factor during re-initiation (Meurs et al., figure 4.K).

To test the hypothesis that MCTS2, the paralog of MCTS1, is expressed and can interact with DENR to fulfill MCTS1-DENR re-initiation function, we preformed a co-immunoprecipitation mass spectrometry experiment using an endogenously tagged version of DENR and identified MCTS2 as an *in vivo* partner of DENR (Meurs et al., figure 5.D). We used our *in vitro* re-initiation assay to check if recombinant MCTS2-DENR promoted expression of the uORF reporters in DENR-MCTS1-depleted extracts and observed that MCTS2-DENR had a similar re-initiation activity to MCTS1-DENR (Meurs et al., figure 5.H and 5.I).

3. Translation of the SARS-CoV-2 genome

3.1. Use of cellular translation machinery by viruses

Viruses are small infectious particles that are fully dependent on a host cell for their reproduction. As a consequence, viruses rely on the translation machinery of the host cell in order to produce the viral proteins. To do so, they developed various strategies to inhibit the expression of host cell proteins, with the dual benefit of preventing the cell's antiviral responses to the viral infection and increasing the translation efficiency of its own ORFs. Furthermore, the viral genome length being limited by the capsid size, viruses use a wide range of methods, some of which are specific to viruses, to optimize the quantity of proteins produced from a minimum number of nucleotides. Two examples of such methods are the use of polyproteins and ribosomal frameshifting in order to translate out of frame ORFs (Rozman, Fisher, and Stern-Ginossar 2023; Karousis, Schubert, and Ban 2024).

3.2. The coronavirus family

Coronaviruses are enveloped single plus-strand RNA genome viruses (+ssRNA) that cause respiratory and digestive tract infections in mammals and birds (Karousis, Schubert, and Ban 2024; V'kovski et al. 2021). The prefix *corona* comes from the latin word which means crown. The name coronavirus was given by June Almeida and David Tyrrell who were the first to observe these crown-shape human viruses under an electron microscope in 1968 (figure 5, Tyrrell and Fielder 2002).

Three deadly coronaviruses for humans, the SARS-CoV-1, the MERS, and the SARS-CoV-2, were responsible for outbreaks in the past 25 years. Among them, the novel SARS-CoV-2

virus was responsible for the global COVID-19 pandemic that started in 2019 (V'kovski et al. 2021). The lack of antiviral treatment against coronaviruses has had a substantial impact on the management of the health crisis and revealed the need for an increased comprehension of the coronaviruses in order to develop effective drugs against them. For this reason, the SARS-CoV-2 has been extensively studied in the past 4 years by molecular biologists and biochemists.

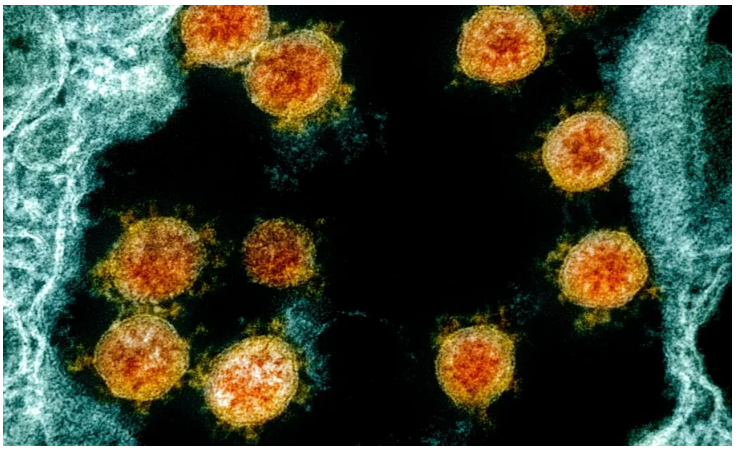


Figure 5 - Electron microscopy image of SARS-CoV-2 particles acquired and color-enhanced by the National Institute of Allergy and Infectious Diseases (NIAID) Integrated Research Facility in Fort Detrick, Md (NIAID Integrated Research Facility 2020). The spike proteins on the surface of the viral envelope are responsible for the crown shape that gave the name idea of coronaviruses to the researchers.

3.3. Translation of the coronavirus SARS-CoV-2 RNA genome

The SARS-CoV-2 genome is composed of a capped and polyadenylated single stranded RNA. Shortly after injection of the SARS-CoV-2 genome into the host cell, the first viral proteins, encoded by ORF1, are produced by cellular ribosomes from the sense RNA strand thanks to its resemblance with cellular mRNAs. ORF1 is composed of two sub-ORFs, ORF1a and ORF1b, that encode for a long polyprotein that will be further processed by a viral protease (NSP3) into 16 different proteins (NSP1 to 16). The NSP proteins are involved in different processes including replication of the coronavirus genome and inhibition of host mRNAs translation (Karousis, Schubert, and Ban 2024; D. Kim et al. 2020).

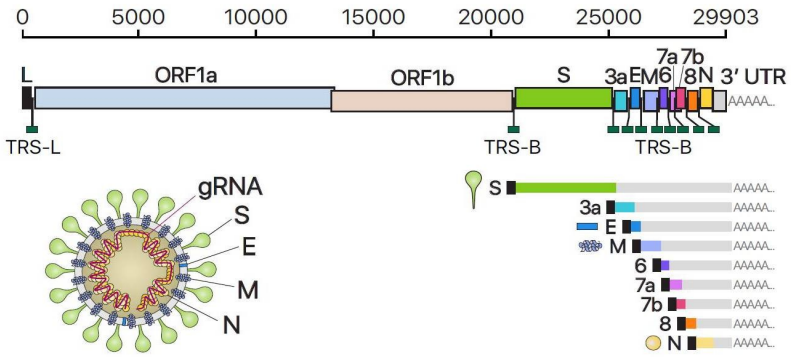


Figure 6 - Schematic of the ORFs encoded by the SARS-CoV-2 genome. A frameshifting event is taking place between the ORF1a and ORF1b, which allow translation of the RNA-dependent RNA polymerase. The other ORFs are sub-products of the genomic RNA. (D. Kim et al. 2020)

Among the NSP proteins encoded by ORF1a, NSP1 binds to the mRNA channel of the 40S ribosomal subunit to inhibit translation of the host mRNAs. Concomitantly, it promotes translation of its own genomic mRNAs in a 5’ UTR-dependent manner by interfering with accommodation of NSP1 into the decoding center (Schubert et al. 2020; Thoms et al. 2020).

The structural proteins, including the nucleocapsid (N), membrane (M), envelope (E) and spike (S) proteins, are secondarily produced from positive-strand subgenomic RNAs (sgRNAs), themselves generated from an intermediate negative-strand RNA synthesized by the viral RNA replicative machinery (V’kovski et al. 2021; Eriani and Martin 2022). In total, 10 different highly translated viral mRNAs are found in the infected cells, of which 9 are sgRNAs (figure 6). They are composed of different size 5’ UTRs which contain different regulatory elements including some inhibitory uORFs that regulate the translation rate of the protein coding ORFs (Condé et al. 2022; Finkel et al. 2021).

3.3.1 Ribosomal frameshifting on the SARS-CoV-2 genome

The ORF1a is the first ORF to be translated from the frame 0 after the viral infection. The downstream ORF1b being encoded on frame -1, the longer version of ORF1 is only expressed following a -1 programmed ribosomal frameshifting event that takes place between ORF1a and ORF1b. As a result, fewer protein molecules encoded by ORF1b than ORF1a are produced (around 50% less ORF1b), which contribute to the stoichiometric control of viral proteins (Karousis, Schubert, and Ban 2024; Finkel et al. 2021). The frameshifting event is conserved among the coronavirus family. Two important RNA elements are responsible for the frameshifting on coronavirus genomes: a slippery site and a downstream pseudoknot structure (Plant and Dinman 2008). From studies on several different virus types, we know that the mechanism can significantly differ from one virus to another (Champagne et al. 2022). Interestingly, no natural pseudoknot-dependent -1 programmed frameshifting event has been observed so far on human transcriptome, making frameshifting an interesting process to target with drugs in order to block the viral propagation without affecting cellular processes. For these reasons, the examination of SARS-CoV-2 frameshifting could provide new insights into the specificity of this event on its genome and could serve as a model to design potential remedies.

Of note, it has been recently observed that MCTS1-DENR promotes ribosomal frameshifting on the SARS-CoV-2 genome by promoting recycling of post-terminating ribosome on the stop codon of the ORF1a, allowing trailing ribosomes to reach the slippery site and frameshift to translate the ORF1b (Rehfeld 2023).

3.4. Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome (Bhatt, Scaiola et al, 2021): main results

The goal of this project initiated by the lab of Nenad Ban and John Atkins was to capture the structure of the frameshifting ribosome between the ORF1a and the ORF1b of the SARS-CoV-2 single stranded RNA genome. As explained in the introduction section, the programmed frameshifting event is necessary for expression of ORF1b, downstream of ORF1a, and to precisely control the stoichiometry of viral proteins.

The rabbit reticulocyte lysate (RRL) system chosen for this study successfully reproduced the frameshifting event as shown by the 0 frame and -1 frame products observed in a Western Blot experiment (Bhatt et al., figure 1.B). In order to capture the ribosome at different stages of frameshifting, a mutant transcript was generated with a UAA stop codon introduced in the slippery site and the *in vitro* reaction was supplemented with a mutant version of the eukaryotic release factor 1 [eRF1 (AAQ)]. Using a cryo-EM approach, the Ban lab captured the structure of a terminating ribosome positioned on the stop codon and a high resolution structure (2.2-Å) of a translating ribosome on the slippery site, slightly upstream of the stop codon. The second structure showed that a pseudoknot directly downstream to the slippery site was slowing down the ribosome progression. A monosome and disome footprint profiling experiment that I carried out in the same set-up as the structural experiments confirmed the slow down of the ribosome on the slippery site and revealed that the ribosomal pause led to the collision of the subsequent ribosome in a similar fashion for both WT and mutant transcripts (Bhatt et al., figure 3). Conversely to the second structure, the one captured on the stop codon did not contain a clear pseudoknot structure at the entrance to the mRNA channel, in agreement with the hypothesis that the ribosome successfully unfolded the pseudoknot structure after the pausing.

The Ban lab next wanted to investigate the contribution of different elements to frameshifting efficiency. Mutation of some segments of the pseudoknot were found to strongly reduce the frameshifting efficiency, including segments interacting with 18S rRNA (Bhatt et al., figure 2.C). It was also shown that a stop codon included within the pseudoknot on the 0 frame served to prevent the trailing ribosome from continuing translation on this frame and to let the pseudoknot fold again for the next elongating ribosomes. Indeed, mutation of the stop codon, without affecting the structure of the pseudoknot, decreased the frameshifting efficiency by 50% while reduction of ribosomal loading by mutation of the start codon partially rescued this effect (Bhatt et al., figure 2.D). Finally, the potential role of the nascent peptide chain in the regulation of the ribosomal frameshifting was investigated. Indeed, the nascent peptide interacts with some nucleotides of the 28S rRNA within the ribosomal exit tunnel (Bhatt et al., figure 4.B) and replacement of the entire peptide sequence increased the frameshifting efficiency by 35% (Bhatt et al., figure 4.F).

Finally, treatment of SARS-CoV-2 infected VeroE6 cells with merafloxacin, a compound that specifically inhibits ribosomal frameshifting, significantly reduced viral proliferation, proving that drugs targeting frameshifting could be developed and used to treat patients suffering from illnesses caused by coronaviruses.

Discussion

1. RNA *cis*-regulatory elements orchestrate expression of mammalian and viral proteins

RNAs have emerged as key regulators of their own translation by modulating ribosome fluxes (Xue and Barna 2015; Madhugiri et al. 2016). The challenge is now to measure those fluxes to better understand how RNA regulatory elements work and what are the consequences of their disruption. In the case of diseases promoted by *cis*-regulatory elements, such as viral infections and cancers, it is even crucial to identify them to prevent or enhance their activity and thus provide new remedies. Several techniques are now used in order to study ribosomal occupancy and movement on transcripts. Two important ones are the ribosome profiling and translation reporter assay. The first one provides an overview of ribosome occupancy on the transcriptome and on individual transcripts under chosen conditions or at specific time points and, when coupled with RNA-seq, gives a measure of translation efficiency (Ingolia, Hussmann, and Weissman 2019). The second one is a technique that gives information about translation of a single reporter that can be modified according to the RNA element of interest. An interesting variant of the reporter assay is the usage of cellular extracts, which give a simplified and controllable system, independent of nuclear regulatory events (Rakotondrafara and Hentze 2011). As part of my thesis, the combination of ribosome profiling, *in vivo* and *in vitro* translation reporter assays successfully answered some of the questions raised in the introduction section.

2. Re-initiation regulation by canonical and non-canonical translation factors

During my thesis, I mainly focused on uORFs for which usage of both ribosome profiling and translation reporter assays partially unraveled their regulatory mechanisms. Nevertheless, as mentioned in the introduction, several questions remain unanswered regarding the translation of transcripts subject to uORF regulation. I contributed to the uORF field through my work on the *cis*-regulatory RNA elements and the factors involved in the re-initiation process. During my investigations, I combined the classically used techniques for the study of translation but I also set up a cell-free re-initiation assay that faithfully reproduced re-initiation. Using homemade HeLa lysates, I measured the ribosomal fluxes on the 5' UTRs of two reporters, *Klhdc8a* and *Asb8*, which contain different configurations of uORFs. Based on this reporter assay, I evaluated the importance of the uORF penultimate codon identity in DENR-dependent re-initiation as well as the participation of eIF2, eIF2D and MCTS2 in re-initiation. Finally, by comparing ribosome occupancy on *Eif2d* and *Denr* knock-down cells, I investigated eIF2D translation function.

2.1. The specificity of DENR-dependent re-initiation

As described in the introduction, the non-canonical translation factors MCTS1-DENR promote re-initiation on specific uORF-containing transcripts in mammalian cells. The specificity of DENR targets has been investigated by a former PhD student of the lab, Violeta Castelo-Szekely, and by Aurelio Teleman's group. Violeta showed that DENR also regulates uORFs that are longer than 6 nucleotides (start-stop uORFs) while Bohlen et al. reported that

DENR targets were enriched for specific penultimate codons (J. Bohlen, Harbrecht, et al. 2020; Castelo-Szekely et al. 2019).

In my ribosome profiling data, I observed that the uORF length distribution of DENR targets is different, yet not significantly, from the length distribution of all translated uORFs. Indeed, DENR-regulated transcripts seemed to be slightly enriched for start-stop uORFs, suggesting that uORF length or translation elongation influences DENR-dependent regulation. Of note, lengthening of *Asb8* uORF from 1 aa to 2 aa affected DENR-dependence, even in the case of the penultimate GCG (highly enriched in DENR targets in Bohlen et al.) in my *in vitro* re-initiation assay. Therefore, it seems that translation elongation on uORFs or presence of an AUG penultimate codon impacts DENR function. This last hypothesis is in line with the observed enrichment of AUG penultimate codons upstream to the CDS stop codon of transcripts with 40S accumulation upon *Denr* knock-out (J. Bohlen, Harbrecht, et al. 2020). Yet, is seems that only a small portion of these start-stop uORF-containing transcripts displays reduced DENR-dependence, implying that the presence of an AUG in the P-site is not sufficient to explain the recruitment of DENR.

Using my *Asb8* and *Klhdc8a* reporters, I observed that only some of the tested penultimate codons diminished DENR-dependence compared to the WT one, in line with the published data from the Teleman group. These codons are recognized by specific tRNAs that are associated with different amino acids. It seems that re-initiation is less DENR-dependent when the uORF of *Klhdc8a* ends with the penultimate codon GTC (depleted in DENR-dependent uORFs) than when its contains the WT penultimate codon GTG, although they both code for a valine. This result suggests that the critical feature for DENR-dependence may be the tRNAs anti-codon identity rather than the associated amino acids. The *in vitro* assay also showed that the identity of the penultimate codon influenced specifically DENR-

dependence without affecting the re-initiation rate *per se*, confirming that re-initiation is not necessarily DENR-dependent. Regarding the relatively low number of DENR targets identified with ribosome profiling (223 in mouse from my RPF-seq, 517 in human in Bohlen at al.) and the high number of transcripts with translated uORFs (3430 in mouse from my RPF-seq), it is tempting to conclude that most re-initiation events are DENR-independent. DENR can then serve as an additional layer of regulation for some transcripts with specific uORF penultimate codons involved in key cellular processes such as cell cycle specific genes. Indeed, as mentioned in the introduction, DENR abundance itself is cell cycle- and cell density-dependent (Clemm von Hohenberg et al. 2022; Deyo, Chiao, and Tainsky 1998).

As observed in my dataset, the most enriched type of translated uORFs are start-stop uORFs (25% in mouse from my article) and uORF frequency rapidly decreases with uORF length (Meurs et al., figure 1.E), which may indicate that uORFs undergo a certain selection pressure to stay particularly short. Indeed, a strong anticorrelation between a conservation score and the uORF length has been observed in *Drosophila* and human (H. Zhang et al. 2021). Several hypotheses are raised in order to explain this observation. First, longer uORFs are obviously more likely to be disrupted by frameshifts or stop codons. Another hypotheses stipulates that translation of non-elongating uORFs limits the loss of initiation factors, such as eIF3, and promotes re-initiation (Schleich et al. 2017). Indeed, the uORF length is inversely correlated with re-initiation efficiency (Kozak 2001). We observed this effect in our re-initiation assays as *Asb8* 5’ UTR (1aa long uORF) displayed a higher re-initiation rate (29%) than *Klhdc8a* 5’ UTR (13aa long uORF, 11%). However, the re-initiation rate was similar for the start-stop and the 2 aa long *Asb8* uORFs in my *in vitro* re-initiation assay, showing that only the uORF length and not the entry into the translation elongation phase influences re-initiation efficiency. Another hypothesis proposes that the very short peptides, and even more so for the

start-stop uORFs that are only coding for a methionine, cost less energy to be translated and therefor the balance of the regulatory outcome vs. energy cost is optimized (H. Zhang et al. 2021). The shorter uORF peptides are also faster to degrade and are less likely to have any effect on other cellular processes. Another possible explanation is that the U at the 4th position of the start site of start-stop uORFs keeps the Kozak context score low to limit the inhibitory effect of these uORFs. Indeed, the expression of my *in vitro* start-stop uORF reporter *Asb8* is mainly driven by leaky scanning, with 58% of ribosomes that bypass the uORF start site to reach the CDS. Finally, as explained in the introduction, the close proximity between the start and stop codons may protect short uORFs from NMD, which favors expression of the protein (Peixeiro et al. 2012; P. V. Ivanov et al. 2008). Although the reason for the enrichment of short uORFs remains unclear, a combination of several of these mechanisms is most likely involved in the conservation mechanism.

2.2. MCTS2 interacts with DENR to promote re-initiation

The partners MCTS1 and DENR work together to promote re-initiation. They stabilize each other by forming a heterodimer, as depletion of one of the two reduces abundance of the other one. However, depletion of MCTS1 does not always lead to the same phenotype as DENR depletion. As explained in the introduction, depletion of DENR is prenatally lethal while *Mcts1* knock-out mice are viable and harbor only minor phenotypical defects as reported in the International Mouse Phenotyping Consortium (IMPC) database. Furthermore, *Mcts1* knock-down did not have the same disruptive effect as *Denr* knock-down on the migration and development of cortical neurons in the mouse brain (Haas 2016). These differences indicate that DENR might have another partner to interact with in order to promote re-initiation. Our results show that MCTS2, the retrogene copy of MCTS1, is expressed in

NIH/3T3 cells and interacts with DENR to promote re-initiation. Like for MCTS1-DENR, the interaction of DENR with MCTS2 should stabilize both factors. This could be verified in the future by comparing the abundance of DENR upon MCTS1 depletion to abundance of MCTS1 upon DENR delpetion as MCTS1 knock-down should have less effect on DENR abundance than DENR depletion on MCTS1 abundance.

Other pieces of evidence exist to support the idea that MCTS2 is expressed and active. Some mutations on MCTS2 are conserved among rodents and primates (4 mutations out of 11 for rodents and 9 for primates compared to MCTS1) suggesting that a certain selection pressure may exist on *Mcts2* to keep a similar amino acid sequence. Another article also reached the conclusion that all the *Mcts* family members are under evolutionary pressure (McCole et al. 2011).

Of note, Bohlen et al. did not find RNA-seq reads mapping specifically to *Mcts2* in their human samples (J. Bohlen, Zhou, et al. 2023), suggesting that *Mcts2* is not expressed in all cell types and that MCTS2-DENR-dependent re-initiation only exists in specific cell types or under certain conditions. Furthermore, *Mcts2* is expressed in several but not all organs during embryonic development (Zhijun et al. 2014). This *Mcts2* embryonic expression could explain why *Mcts1* knock-out mice survive into adulthood while *Denr* knock-out mice die pre-nataly.

2.3. eIF2α phosphorylation only mildly affects re-initiation and leaky scanning on uORF-containing transcripts

Protein expression from several uORF-containing transcripts is increased following eIF2α inactivation during the integrated stress response (S. K. Young et al. 2015; Y.-Y. Lee, Cevallos, and Jan 2009; Amiri et al. 2024). *Atf4* and *Atf5* are two well-studied representatives of this category of transcripts. Under physiological conditions, the overlapping uORFs of *Atf4*

and *Atf5* are translated, which severely inhibits expression of the CDS. The induction mechanism of these two transcripts under stress conditions relies on the bypassing of the overlapping uORFs during re-initiation between the upstream uORFs and the CDS start codon. To explain this phenomenon, several publications have hypothesized that the low availability of active eIF2-GTP-Met-tRNA$_i^{Met}$ complexes increases its recruitment time to the re-scanning ribosome, which forces ribosome to re-initiate at the CDS start codon (Vattem and Wek 2004; Watatani et al. 2008; Zhou et al. 2008). Using my *in vitro* re-initiation assay, I discovered that the re-initiation and leaky scanning rates of my *Asb8* and *Klhdc8a* reporters were only mildly reduced by the phosphorylation of EIF2α. It seems therefore that the strong re-initiation delay observed for *Atf4* and *Atf5* transcripts does not apply to all uORF-containing transcripts. Furthermore, some uORF-regulatory effects are specific to some stresses, thus inducing eIF2α is not sufficient for inhibiting or promoting translation of uORF-containing transcripts (Akulich et al. 2019; Jun Zhang and Shi 2024). A recent paper showed that ATF4 enhances expression of eIF1 which in turn increases stringency of start site selection. Upon stress, the scanning ribosomes are thus more likely to bypass uORFs with weak start sites and initiate at the next CDS, which could contribute to the expression of uORF-regulated transcripts (Amiri et al. 2024).

The stronger reduction in re-initiation efficiency of *Asb8* uORF can be explained by the short intercistronic distance between its uORF and the CDS start site (17 nt). Interestingly, the same intercistronic distance is slightly shorter for *Klhdc8a* (73 nt) than the distance between the uORF and the overlapping uORF for *Atf4* (87 nt), implying that re-initiation on *Atf4* should be even less sensitive to the phosphorylation of EIF2α. This could be tested by comparing the signal of *Atf4* reporters to the signal of our two other reporters in our *in vitro* re-initiation

system. To evaluate the length effect, one could also shorten the intercistronic distance of the *Klhdc8a* and *Asb8* reporters to measure the effect on re-initiation rate.

Several studies showed that the *Atf4* transcript undergoes MCTS1-DENR-dependent re-initiation in *Drosophila* and human cells (J. Bohlen, Harbrecht, et al. 2020; Vasudevan et al. 2020). In yeast however, GCN4 is not regulated by the MCTS1-DENR ortholog (TMA20-TMA22) (Gaikwad et al. 2021). I did not observe a decrease in *Atf4* and *Atf5* reporter expression upon DENR depletion in HeLa cells nor on the endogenous transcripts in NIH/3T3 cells. But I observed an accumulation of footprints at the non-overlapping uORF stop codon, indicating a recycling defect associated with the depletion of DENR.

2.4. eIF2D is not a re-initiation factor under physiological conditions

eIF2D is a non-canonical initiation factor with a tRNA recruitment and release activity *in vitro*. Its function and its targets remain unclear. Like DENR, it is considered to be a re-initiation factor in animals. However, validating experiments are missing in order to confirm its re-initiation function. Moreover, it it not known if eIF2D regulates the same class of uORF-containing transcripts or if they have different targets. The ribosome profiling experiment on *Eif2d* knock-down cells revealed that eIF2D-regulated transcripts were not enriched with translated uORFs. Furthermore, only one transcript (*Nrf1*) had a reduced translation efficiency in both *Denr* and *Eif2d* knock-down cells. Taken together, these results show that eIF2D regulates a different set of transcripts than DENR and it does not promote re-initiation under physiological conditions *in vivo*. Still, the results of the *in vitro* re-initiation assay indicate that eIF2D is able to promote some level of re-initiation when DENR is depleted, in line with the observation that the knock-out of the yeast eIF2D ortholog, TMA64, enhances the recycling defect of TMA20 (MCTS1) or TMA22 (DENR) depletion but does not

affect recycling on its own (D. J. Young, Meydan, and Guydosh 2021). In conclusion, these results indicate that eIF2D is involved in another translation related process but can inefficiently promote re-initiation if MCTS1 or DENR are absent.

If eIF2D is not involved in re-initiation, what is its main function? To answer this question, we first attempted to validate some eIF2D candidate transcripts identified in the ribosome profiling analysis with an *in vivo* reporter assay. We were unable to observe a 5' UTR-dependent eIF2D regulation of the selected transcripts, suggesting that eIF2D indirectly controlled expression of these transcripts or that it regulated them in a 5' UTR-independent manner. Moreover, my analysis of the eIF2D-regulated transcripts did not reveal any specific feature within their sequence that would explain their eIF2D-dependence. Yet, many of the affected transcripts seemed to be particularly abundant and were enriched for mRNAs involved in translation and ribosome biogenesis. I also observed that eIF2D depletion greatly affected RNA abundance of hundreds of transcripts. Like the transcripts with a reduced translation efficiency, those with a changed abundance were also enriched for highly abundant transcripts and coding for translation-associated proteins. These transcripts were also involved in the regulation of the cell cycle. Indeed, this massive change in transcript abundance is associated with an accumulation of cells in the G1 phase of the cell cycle in *Eif2d* knock-down cells. A link between the arrest of the cell cycle at the G1/S1 phases checkpoint and reduction of ribosomal protein expression has been made and could reconcile our observations (Derenzini 2017). However, one limitation of the depletion technique used for this experiment is that a 4 day selection of the transduced cells is required, leaving time for secondary effects to occur. Based on the ribosome profiling data alone it is therefore not possible to conclude on a direct interaction between eIF2D and the transcripts affected by the knock-down. It is also not possible to know if the reduced expression of ribosomal proteins is

responsible for the cell cycle arrest or if it is the cell cycle arrest that provokes a reduction of ribosomal protein expression.

In vitro, eIF2D can recruit a Met-tRNA$_i^{Met}$ independently of eIF2 and the hydrolysis of a GTP molecule (Skabkin et al. 2010). Based on this observation, it has been proposed that eIF2D could promote initiation on unconventional start sites or under conditions where eIF2 is inactivated, e.g. when eIF2α is phosphorylated (Wek, Anthony, and Staschke 2023). Most of the investigations about eIF2D failed to observe such non-canonical initiation functions *in vivo* (Green et al. 2022; Ichihara et al. 2021; Sanz, González Almela, and Carrasco 2017; Sanz et al. 2019; She, Luo, and Weissman 2023). Using my *in vitro* set-up, I tested the initiation function of eIF2D in eIF2α phosphorylated extracts. The knock-down of eIF2D did not further impact re-initiation and leaky scanning rates on *Klhdc8a* reporters. Furthermore, addition of a large amount of recombinant eIF2D did not increase expression of the reporters in extracts treated or not treated with recombinant GADD34. In line with the previous findings, these results suggest that eIF2D cannot replace eIF2 for the recruitment of a new Met-tRNA$_i^{Met}$ during re-initiation and leaky scanning.

Among the identified eIF2D targets, *Hoxa3* has been shown to be inhibited by eIF2D through its translation inhibitory element, which contains a uORF (Alghoul et al. 2021). In my ribosome profiling data, *Hoxa3* had a low expression level but had an almost significantly increased translation efficiency. However, only 5 transcripts had a significant reduced translation efficiency upon eIF2D depletion and none of them contained a translated uORF, implying that the regulation of *Hoxa3* by eIF2D is rather a specific mechanism than a general one.

Several other experiments could be carried out in order to identify eIF2D targets and function. For instance, a selective ribosome profiling experiment could be performed to identify direct

eIF2D target transcripts and the sites where eIF2D-bound ribosomes interact with those transcripts.

3. Expression of the complete SARS-CoV-2 genome is conditioned by a programmed and highly regulated frameshifting event

Other techniques can be used to study ribosomal translation events such are structural analysis. Among those techniques, Bhatt et al. used cryo-EM to characterize the frameshifting event on the SARS-CoV-2 genome. They chose to work with an *in vitro* rabbit reticulocyte lysate (RRL) system supplemented with a harmless SARS-CoV-2 reporter. As explained above, *in vitro* systems have several advantages such as a high translation activity and the uncoupling from nuclear events, but the RRL system has the disadvantages to be produced from specialized cells in which some translation processes and pathways are deficient (Pisareva et al. 2015).

Using cryo-EM, they successfully obtained two different structures that are representative of two different stages of the frameshifting. However, this technique provides a very static *in vitro* view of cellular events and had to be completed with some *in vivo* techniques such as reporter assays and ribosome profiling, in where I became involved. Indeed, the ribosome profiling experiment exhibited the ribosomal pausing on the slippery site of the ORF1, confirming that the translating ribosome was stalled due to the downstream pseudoknot structure. One challenge of this project was to combine an *in vitro* translation assay with the generation of ribosomal footprints on a single mRNA reporter. This work showed that it was possible to generate such libraries after some adjustments. Another challenge we encountered

was to digest the strong secondary structures of the SARS-CoV-2 reporter that were contaminating the ribosomal footprints. We successfully removed those by performing a double digestion with MNase and RNase 1.

Finally, this project allowed a precise characterization of the SARS-CoV-2 mRNA-specific frameshifting event. This characterization can potentially be used for drug design as inhibition of the frameshifting should only impact expression of viral proteins. An independent study also identified merafloxacin as a compound that can inhibit viral replication by blocking the ribosomal frameshifting event (Sun et al. 2021). Investigations are still ongoing in order to identify more specific molecules to target the frameshifting event in order to avoid secondary effects on human cells (Yang et al. 2023).

References

Aitken, Colin Echeverría, Petra Beznosková, Vladislava Vlčkova, Wen-Ling Chiu, Fujun Zhou, Leoš Shivaya Valášek, Alan G Hinnebusch, and Jon R Lorsch. 2016. “Eukaryotic Translation Initiation Factor 3 Plays Distinct Roles at the MRNA Entry and Exit Channels of the Ribosomal Preinitiation Complex.” *ELife* 5 (October):e20934. https://doi.org/10.7554/eLife.20934.

Akulich, Kseniya A., Pavel G. Sinitcyn, Desislava S. Makeeva, Dmitry E. Andreev, Ilya M. Terenin, Aleksandra S. Anisimova, Ivan N. Shatsky, and Sergey E. Dmitriev. 2019. “A Novel UORF-Based Regulatory Mechanism Controls Translation of the Human MDM2 and EIF2D MRNAs during Stress.” *Biochimie* 157 (February):92–101. https://doi.org/10.1016/j.biochi.2018.11.005.

Alghoul, Fatima, Schaeffer Laure, Gilbert Eriani, and Franck Martin. 2021. “Translation Inhibitory Elements from Hoxa3 and Hoxa11 MRNAs Use UORFs for Translation Inhibition.” *ELife* 10 (June):e66369. https://doi.org/10.7554/eLife.66369.

Algire, Mikkel A., David Maag, and Jon R. Lorsch. 2005. “Pi Release from EIF2, Not GTP Hydrolysis, Is the Step Controlled by Start-Site Selection during Eukaryotic Translation Initiation.” *Molecular Cell* 20 (2): 251–62. https://doi.org/10.1016/j.molcel.2005.09.008.

Amiri, Mehdi, Stephen J. Kiniry, Anthony P. Possemato, Niaz Mahmood, Tayebeh Basiri, Catherine R. Dufour, Negar Tabatabaei, et al. 2024. “Impact of EIF2α Phosphorylation on the Translational Landscape of Mouse Embryonic Stem Cells.” *Cell Reports* 43 (1). https://doi.org/10.1016/j.celrep.2023.113615.

Andreev, Dmitry E., Gary Loughran, Alla D. Fedorova, Maria S. Mikhaylova, Ivan N. Shatsky, and Pavel V. Baranov. 2022. “Non-AUG Translation Initiation in Mammals.” *Genome Biology* 23 (May):111. https://doi.org/10.1186/s13059-022-02674-2.

Aravind, L., and Eugene V. Koonin. 1999. “Novel Predicted RNA-Binding Domains Associated with the Translation Machinery.” *Journal of Molecular Evolution* 48 (3): 291–302. https://doi.org/10.1007/PL00006472.

Arribere, Joshua A., and Wendy V. Gilbert. 2013. “Roles for Transcript Leaders in Translation and MRNA Decay Revealed by Transcript Leader Sequencing.” *Genome Research* 23 (6): 977–87. https://doi.org/10.1101/gr.150342.112.

Barbosa, Cristina, Isabel Peixeiro, and Luísa Romão. 2013. “Gene Expression Regulation by Upstream Open Reading Frames and Human Disease.” *PLOS Genetics* 9 (8): e1003529. https://doi.org/10.1371/journal.pgen.1003529.

Bazzini, Ariel A, Timothy G Johnstone, Romain Christiano, Sebastian D Mackowiak, Benedikt Obermayer, Elizabeth S Fleming, Charles E Vejnar, et al. 2014. “Identification of Small ORFs in Vertebrates Using Ribosome Footprinting and Evolutionary Conservation.” *The EMBO Journal* 33 (9): 981–93. https://doi.org/10.1002/embj.201488411.

Bhatt, Pramod R., Alain Scaiola, Gary Loughran, Marc Leibundgut, Annika Kratzel, Romane Meurs, René Dreos, et al. 2021. “Structural Basis of Ribosomal Frameshifting during Translation of the SARS-CoV-2 RNA Genome.” *Science (New York, N.Y.)* 372 (6548): 1306–13. https://doi.org/10.1126/science.abf3546.

Bohlen, Anna E. von, Johann Böhm, Ramona Pop, Diana S. Johnson, John Tolmie, Ralf Stücker, Deborah Morris-Rosendahl, and Gerd Scherer. 2017. “A Mutation Creating an Upstream Initiation Codon in the SOX9 5′ UTR Causes Acampomelic Campomelic Dysplasia.” *Molecular Genetics & Genomic Medicine* 5 (3): 261–68. https://doi.org/10.1002/mgg3.282.

Bohlen, Jonathan, Kai Fenzl, Günter Kramer, Bernd Bukau, and Aurelio A. Teleman. 2020. “Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells.” *Molecular Cell* 79 (4): 561-574.e5. https://doi.org/10.1016/j.molcel.2020.06.005.

Bohlen, Jonathan, Liza Harbrecht, Saioa Blanco, Katharina Clemm von Hohenberg, Kai Fenzl, Günter Kramer, Bernd Bukau, and Aurelio A. Teleman. 2020. “DENR Promotes Translation Reinitiation via Ribosome Recycling to Drive Expression of Oncogenes Including ATF4.” *Nature Communications* 11 (1): 4676. https://doi.org/10.1038/s41467-020-18452-2.

Bohlen, Jonathan, Mykola Roiuk, Marilena Neff, and Aurelio A Teleman. 2023. “PRRC2 Proteins Impact Translation Initiation by Promoting Leaky Scanning.” *Nucleic Acids Research* 51 (7): 3391–3409. https://doi.org/10.1093/nar/gkad135.

Bohlen, Jonathan, Qinhua Zhou, Quentin Philippot, Masato Ogishi, Darawan Rinchai, Tea Nieminen, Simin Seyedpour, et al. 2023. “Human MCTS1-Dependent Translation of JAK2 Is Essential for IFN-γ Immunity to MycobacteriaHuman MCTS1-Dependent Translation of JAK2 Is Essential for IFN-γ Immunity to Mycobacteria.” *Cell* 186 (23): 5114-5134.e27. https://doi.org/10.1016/j.cell.2023.09.024.

Brito Querido, Jailson, Irene Díaz-López, and V. Ramakrishnan. 2024. “The Molecular Basis of Translation Initiation and Its Regulation in Eukaryotes.” *Nature Reviews Molecular Cell Biology* 25 (3): 168–86. https://doi.org/10.1038/s41580-023-00624-9.

Brito Querido, Jailson, Masaaki Sokabe, Sebastian Kraatz, Yuliya Gordiyenko, J. Mark Skehel, Christopher S. Fraser, and V. Ramakrishnan. 2020. “Structure of a Human 48S Translational Initiation Complex.” *Science* 369 (6508): 1220–27. https://doi.org/10.1126/science.aba4904.

Bryant, Jeffrey D., Michael C. Brown, Mikhail I. Dobrikov, Elena Y. Dobrikova, Sarah L. Gemberling, Qing Zhang, and Matthias Gromeier. 2018. “Regulation of Hypoxia-Inducible Factor 1α during Hypoxia by DAP5-Induced Translation of PHD2.” *Molecular and Cellular Biology* 38 (11): e00647-17. https://doi.org/10.1128/MCB.00647-17.

Calvo, Sarah E., David J. Pagliarini, and Vamsi K. Mootha. 2009. “Upstream Open Reading Frames Cause Widespread Reduction of Protein Expression and Are Polymorphic among Humans.” *Proceedings of the National Academy of Sciences* 106 (18): 7507–12. https://doi.org/10.1073/pnas.0810916106.

Cao, Xiongwen, and Sarah A. Slavoff. 2020. “Non-AUG Start Codons: Expanding and Regulating the Small and Alternative ORFeome.” *Experimental Cell Research* 391 (1): 111973. https://doi.org/10.1016/j.yexcr.2020.111973.

Castelo-Szekely, Violeta, Mara De Matos, Marina Tusup, Steve Pascolo, Jernej Ule, and David Gatfield. 2019. “Charting DENR-Dependent Translation Reinitiation Uncovers Predictive UORF Features and Links to Circadian Timekeeping via Clock.” *Nucleic Acids Research* 47 (1): 5193–5209. https://doi.org/10.1093/nar/gkz261.

Champagne, Julien, Kelly Mordente, Remco Nagel, and Reuven Agami. 2022. “Slippy-Sloppy Translation: A Tale of Programmed and Induced-Ribosomal Frameshifting.” *Trends in Genetics* 38 (11): 1123–33. https://doi.org/10.1016/j.tig.2022.05.009.

Chen, Jin, Andreas-David Brunner, J. Zachery Cogan, James K. Nuñez, Alexander P. Fields, Britt Adamson, Daniel N. Itzhak, et al. 2020. “Pervasive Functional Translation of Noncanonical Human Open Reading Frames.” *Science (New York, N.Y.)* 367 (6482): 1140–46. https://doi.org/10.1126/science.aay0262.

Chu, Jennifer, Marie Cargnello, Ivan Topisirovic, and Jerry Pelletier. 2016. “Translation Initiation Factors: Reprogramming Protein Synthesis in Cancer.” *Trends in Cell Biology* 26 (12): 918–33. https://doi.org/10.1016/j.tcb.2016.06.005.

Clemm von Hohenberg, Katharina, Sandra Müller, Sibylle Schleich, Matthias Meister, Jonathan Bohlen, Thomas G. Hofmann, and Aurelio A. Teleman. 2022. “Cyclin B/CDK1 and Cyclin A/CDK2 Phosphorylate DENR to Promote Mitotic Protein Translation and Faithful Cell Division.” *Nature Communications* 13 (1): 668. https://doi.org/10.1038/s41467-022-28265-0.

Colombo, Martino, Evangelos D. Karousis, Joël Bourquin, Rémy Bruggmann, and Oliver Mühlemann. 2017. “Transcriptome-Wide Identification of NMD-Targeted Human MRNAs Reveals Extensive Redundancy between SMG6- and SMG7-Mediated Degradation Pathways.” *RNA* 23 (2): 189. https://doi.org/10.1261/rna.059055.116.

Condé, Lionel, Omran Allatif, Théophile Ohlmann, and Sylvain de Breyne. 2022. “Translation of SARS-CoV-2 GRNA Is Extremely Efficient and Competitive despite a High Degree of Secondary Structures and the Presence of an UORF.” *Viruses* 14 (7): 1505. https://doi.org/10.3390/v14071505.

Das, S., R. Ghosh, and U. Maitra. 2001. “Eukaryotic Translation Initiation Factor 5 Functions as a GTPase-Activating Protein.” *The Journal of Biological Chemistry* 276 (9): 6720–26. https://doi.org/10.1074/jbc.M008863200.

Derenzini, Massimo, Montanaro, Lorenzo, Trerè, Davide. 2017. “Ribosome Biogenesis and Cancer.” *Acta Histochemica* 119 (3): 190-7. https://doi.org/10.1016/j.acthis.2017.01.009

Dever, Thomas E., Jonathan D. Dinman, and Rachel Green. 2018. “Translation Elongation and Recoding in Eukaryotes.” *Cold Spring Harbor Perspectives in Biology* 10 (8): a032649. https://doi.org/10.1101/cshperspect.a032649.

Deyo, Jeff E., Paul J. Chiao, and Michael A. Tainsky. 1998. “Drp, a Novel Protein Expressed at High Cell Density but Not During Growth Arrest.” *DNA and Cell Biology* 17 (5): 437–47. https://doi.org/10.1089/dna.1998.17.437.

Dierov, J., M. Prosniak, G. Gallia, and R.b. Gartenhaus. 1999. “Increased G1 Cyclin/Cdk Activity in Cells Overexpressing the Candidate Oncogene, MCT-1.” *Journal of Cellular Biochemistry* 74 (4): 544–50. https://doi.org/10.1002/(SICI)1097-4644(19990915)74:4<544::AID-JCB4>3.0.CO;2-4.

Dmitriev, Sergey E., Ilya M. Terenin, Dmitri E. Andreev, Pavel A. Ivanov, Jacov E. Dunaevsky, William C. Merrick, and Ivan N. Shatsky. 2010. “GTP-Independent TRNA Delivery to the Ribosomal P-Site by a Novel Eukaryotic Translation Factor.” *Journal of Biological Chemistry* 285 (35): 26779–87. https://doi.org/10.1074/jbc.M110.119693.

Dorris, D R, F L Erickson, and E M Hannig. 1995. “Mutations in GCD11, the Structural Gene for EIF-2 Gamma in Yeast, Alter Translational Regulation of GCN4 and the Selection of the Start Site for Protein Synthesis.” *The EMBO Journal* 14 (10): 2239–49.

Eriani, Gilbert, and Franck Martin. 2022. “Viral and Cellular Translation during SARS-CoV-2 Infection.” *FEBS Open Bio* 12 (9): 1584–1601. https://doi.org/10.1002/2211-5463.13413.

Fekete, Christie A, Sarah F Mitchell, Vera A Cherkasova, Drew Applefield, Mikkel A Algire, David Maag, Adesh K Saini, Jon R Lorsch, and Alan G Hinnebusch. 2007. “N- and C-Terminal Residues of EIF1A Have Opposing Effects on the Fidelity of Start Codon Selection.” *The EMBO Journal* 26 (6): 1602–14. https://doi.org/10.1038/sj.emboj.7601613.

Finkel, Yaara, Orel Mizrahi, Aharon Nachshon, Shira Weingarten-Gabbay, David Morgenstern, Yfat Yahalom-Ronen, Hadas Tamir, et al. 2021. “The Coding Capacity of SARS-CoV-2.” *Nature* 589 (7840): 125–30. https://doi.org/10.1038/s41586-020-2739-1.

Gaba, Anthony, Allan Jacobson, and Matthew S. Sachs. 2005. “Ribosome Occupancy of the Yeast *CPA1* Upstream Open Reading Frame Termination Codon Modulates Nonsense-Mediated MRNA Decay.” *Molecular Cell* 20 (3): 449–60. https://doi.org/10.1016/j.molcel.2005.09.019.

Gaikwad, Swati, Fardin Ghobakhlou, David J Young, Jyothsna Visweswaraiah, Hongen Zhang, and Alan G Hinnebusch. 2021. “Reprogramming of Translation in Yeast Cells Impaired for Ribosome Recycling Favors Short, Efficiently Translated MRNAs.” *ELife* 10 (March):e64283. https://doi.org/10.7554/eLife.64283.

Gao, Cun, Rui Dong, Yongmeng Li, Jinghui Liang, and Hui Tian. 2021. “MCTS1 Promotes the Development of Lung Adenocarcinoma by Regulating E2F1 Expression.” *Oncology Letters* 22 (1): 531. https://doi.org/10.3892/ol.2021.12792.

González-Almela, Esther, Hugh Williams, Miguel A. Sanz, and Luis Carrasco. 2018. “The Initiation Factors EIF2, EIF2A, EIF2D, EIF4A, and EIF4G Are Not Involved in Translation Driven by Hepatitis C Virus IRES in Human Cells.” *Frontiers in Microbiology* 9:207. https://doi.org/10.3389/fmicb.2018.00207.

Green, Katelyn M, Shannon L Miller, Indranil Malik, and Peter K Todd. 2022. “Non-Canonical Initiation Factors Modulate Repeat-Associated Non-AUG Translation.” *Human Molecular Genetics* 31 (15): 2521–34. https://doi.org/10.1093/hmg/ddac021.

Gunišová, Stanislava, Vladislava Hronová, Mahabub Pasha Mohammad, Alan G Hinnebusch, and Leoš Shivaya Valášek. 2017. “Please Do Not Recycle! Translation Reinitiation in Microbes and Higher Eukaryotes.” *FEMS Microbiology Reviews* 42 (2): 165–92. https://doi.org/10.1093/femsre/fux059.

Gutierrez, Erik, Byung-Sik Shin, Christopher J. Woolstenhulme, Joo-Ran Kim, Preeti Saini, Allen R. Buskirk, and Thomas E. Dever. 2013. “EIF5A Promotes Translation of Polyproline Motifs.” *Molecular Cell* 51 (1): 35–45. https://doi.org/10.1016/j.molcel.2013.04.021.

Gygi, Steven P., Yvan Rochon, B. Robert Franza, and Ruedi Aebersold. 1999. “Correlation between Protein and MRNA Abundance in Yeast.” *Molecular and Cellular Biology* 19 (3): 1720–30.

Haas, Matilda A. 2016. “De Novo Mutations in DENR Disrupt Neuronal Development and Link Congenital Neurological Disorders to Faulty MRNA Translation Re-Initiation.”

Harper, J. Wade, and Eric J. Bennett. 2016. “Proteome Complexity and the Forces That Drive Proteome Imbalance.” *Nature* 537 (7620): 328–38. https://doi.org/10.1038/nature19947.

Hashem, Yaser, and Joachim Frank. 2018. “The Jigsaw Puzzle of MRNA Translation Initiation in Eukaryotes: A Decade of Structures Unraveling the Mechanics of the Process.” *Annual Review of Biophysics* 47 (May):125–51. https://doi.org/10.1146/annurev-biophys-070816-034034.

Hellen, Christopher U.T. 2018. “Translation Termination and Ribosome Recycling in Eukaryotes.” *Cold Spring Harbor Perspectives in Biology* 10 (10): a032656. https://doi.org/10.1101/cshperspect.a032656.

Hicks, Danielle, Krithika Giresh, Lisa A. Wrischnik, and Douglas C. Weiser. 2023. “The PPP1R15 Family of EIF2-Alpha Phosphatase Targeting Subunits (GADD34 and CReP).” *International Journal of Molecular Sciences* 24 (24): 17321. https://doi.org/10.3390/ijms242417321.

Hinnebusch, Alan G. 2014. “The Scanning Mechanism of Eukaryotic Translation Initiation.” *Annual Review of Biochemistry* 83:779–812. https://doi.org/10.1146/annurev-biochem-060713-035802.

Hornig, Nadine C., Carine de Beaufort, Friederike Denzer, Martine Cools, Martin Wabitsch, Martin Ukat, Alexandra E. Kulle, et al. 2016. “A Recurrent Germline Mutation in the 5’UTR of the Androgen Receptor Causes Complete Androgen Insensitivity by Activating Aberrant UORF Translation.” *PloS One* 11 (4): e0154158. https://doi.org/10.1371/journal.pone.0154158.

Hsu, Hsin-Ling, Bo Shi, and Ronald B. Gartenhaus. 2005. “The MCT-1 Oncogene Product Impairs Cell Cycle Checkpoint Control and Transforms Human Mammary Epithelial Cells.” *Oncogene* 24 (31): 4956–64. https://doi.org/10.1038/sj.onc.1208680.

Huang, Bridget Y., and Israel S. Fernández. 2020. “Long-Range Interdomain Communications in EIF5B Regulate GTP Hydrolysis and Translation Initiation.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (3): 1429–37. https://doi.org/10.1073/pnas.1916436117.

Huang, Zhexun, Qiao Su, Wuguo Li, Hui Ren, Huiqiang Huang, and Anxun Wang. 2021. “MCTS1 Promotes Invasion and Metastasis of Oral Cancer by Modifying the EMT Process.” *Annals of Translational Medicine* 9 (12): 997. https://doi.org/10.21037/atm-21-2361.

Hug, Nele, Dasa Longman, and Javier F. Cáceres. 2016. “Mechanism and Regulation of the Nonsense-Mediated Decay Pathway.” *Nucleic Acids Research* 44 (4): 1483–95. https://doi.org/10.1093/nar/gkw010.

Iacono, Michele, Flavio Mignone, and Graziano Pesole. 2005. “UAUG and UORFs in Human and Rodent 5′untranslated MRNAs.” *Gene* 349 (April):97–105. https://doi.org/10.1016/j.gene.2004.11.041.

Ichihara, Kazuya, Akinobu Matsumoto, Hiroshi Nishida, Yuki Kito, Hideyuki Shimizu, Yuichi Shichino, Shintaro Iwasaki, Koshi Imami, Yasushi Ishihama, and Keiichi I Nakayama. 2021. “Combinatorial Analysis of Translation Dynamics Reveals EIF2 Dependence of Translation Initiation at Near-Cognate Codons.” *Nucleic Acids Research* 49 (13): 7298–7317. https://doi.org/10.1093/nar/gkab549.

Ingolia, Nicholas T., Jeffrey A. Hussmann, and Jonathan S. Weissman. 2019. “Ribosome Profiling: Global Views of Translation.” *Cold Spring Harbor Perspectives in Biology* 11 (5): a032698. https://doi.org/10.1101/cshperspect.a032698.

Ingolia, Nicholas T., Liana F. Lareau, and Jonathan S. Weissman. 2011. “Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity of Mammalian Proteomes.” *Cell* 147 (4): 789–802. https://doi.org/10.1016/j.cell.2011.10.002.

Ivanov, Ivaylo P., Jiajie Wei, Stephen Z. Caster, Kristina M. Smith, Audrey M. Michel, Ying Zhang, Andrew E. Firth, et al. 2017. “Translation Initiation from Conserved Non-AUG Codons Provides Additional Layers of Regulation and Coding Capacity.” *MBio* 8 (3): e00844-17. https://doi.org/10.1128/mBio.00844-17.

Ivanov, Pavel V, Niels H Gehring, Joachim B Kunz, Matthias W Hentze, and Andreas E Kulozik. 2008. “Interactions between UPF1, ERFs, PABP and the Exon Junction Complex Suggest an Integrated Model for Mammalian NMD Pathways.” *The EMBO Journal* 27 (5): 736–47. https://doi.org/10.1038/emboj.2008.17.

Jackson, Richard J., Christopher U. T. Hellen, and Tatyana V. Pestova. 2010. “The Mechanism of Eukaryotic Translation Initiation and Principles of Its Regulation.” *Nature Reviews. Molecular Cell Biology* 11 (2): 113–27. https://doi.org/10.1038/nrm2838.

Jakoi, E. R., and R. B. Marchase. 1979. “Ligatin from Embryonic Chick Neural Retina.” *The Journal of Cell Biology* 80 (3): 642–50. https://doi.org/10.1083/jcb.80.3.642.

Johnstone, Timothy G., Ariel A. Bazzini, and Antonio J. Giraldez. 2016. “Upstream ORFs Are Prevalent Translational Repressors in Vertebrates.” *The EMBO Journal* 35 (7): 706–23. https://doi.org/10.15252/embj.201592759.

Karousis, Evangelos D., Katharina Schubert, and Nenad Ban. 2024. “Coronavirus Takeover of Host Cell Translation and Intracellular Antiviral Response: A Molecular Perspective.” *The EMBO Journal* 43 (2): 151–67. https://doi.org/10.1038/s44318-023-00019-8.

Kim, Dongwan, Joo-Yeon Lee, Jeong-Sun Yang, Jun Won Kim, V. Narry Kim, and Hyeshik Chang. 2020. “The Architecture of SARS-CoV-2 Transcriptome.” *Cell* 181 (4): 914-921.e10. https://doi.org/10.1016/j.cell.2020.04.011.

Kim, Hyejeong, David Aponte-Diaz, Mohamad S. Sotoudegan, Djoshkun Shengjuler, Jamie J. Arnold, and Craig E. Cameron. 2023. “The Enterovirus Genome Can Be Translated in an IRES-Independent Manner That Requires the Initiation Factors EIF2A/EIF2D.” *PLoS Biology* 21 (1): e3001693. https://doi.org/10.1371/journal.pbio.3001693.

Kozak, M. 1981. “Possible Role of Flanking Nucleotides in Recognition of the AUG Initiator Codon by Eukaryotic Ribosomes.” *Nucleic Acids Research* 9 (20): 5233–52. https://doi.org/10.1093/nar/9.20.5233.

Kozak, M. 1984a. “Compilation and Analysis of Sequences Upstream from the Translational Start Site in Eukaryotic MRNAs.” *Nucleic Acids Research* 12 (2): 857–72.

Kozak, M. 1984b. “Point Mutations Close to the AUG Initiator Codon Affect the Efficiency of Translation of Rat Preproinsulin in Vivo.” *Nature* 308 (5956): 241–46. https://doi.org/10.1038/308241a0.

Kozak, M. 1987. “Effects of Intercistronic Length on the Efficiency of Reinitiation by Eucaryotic Ribosomes.” *Molecular and Cellular Biology* 7 (10): 3438. https://doi.org/10.1128/mcb.7.10.3438.

Kozak, M. 1990. “Downstream Secondary Structure Facilitates Recognition of Initiator Codons by Eukaryotic Ribosomes.” *Proceedings of the National Academy of Sciences* 87 (21): 8301–5. https://doi.org/10.1073/pnas.87.21.8301.

Kozak, M. 2001. “Constraints on Reinitiation of Translation in Mammals.” *Nucleic Acids Research* 29 (24): 5226–32.

Kozak, M. 2002. “Pushing the Limits of the Scanning Mechanism for Initiation of Translation.” *Gene* 299 (1): 1–34. https://doi.org/10.1016/S0378-1119(02)01056-9.

Kumar, Parimal, Christopher U. T. Hellen, and Tatyana V. Pestova. 2016. “Toward the Mechanism of EIF4F-Mediated Ribosomal Attachment to Mammalian Capped MRNAs.” *Genes & Development* 30 (13): 1573–88. https://doi.org/10.1101/gad.282418.116.

Lapointe, Christopher P., Rosslyn Grosely, Masaaki Sokabe, Carlos Alvarado, Jinfan Wang, Elizabeth Montabana, Nancy Villa, et al. 2022. “EIF5B and EIF1A Reorient Initiator TRNA to Allow Ribosomal Subunit Joining.” *Nature* 607 (7917): 185–90. https://doi.org/10.1038/s41586-022-04858-z.

Lee, David S. M., Joseph Park, Andrew Kromer, Aris Baras, Daniel J. Rader, Marylyn D. Ritchie, Louis R. Ghanem, and Yoseph Barash. 2021. “Disrupting Upstream Translation in MRNAs Is Associated with Human Disease.” *Nature Communications* 12 (1): 1515. https://doi.org/10.1038/s41467-021-21812-1.

Lee, Sooncheol, Botao Liu, Soohyun Lee, Sheng-Xiong Huang, Ben Shen, and Shu-Bing Qian. 2012. “Global Mapping of Translation Initiation Sites in Mammalian Cells at Single-Nucleotide Resolution.” *Proceedings of the National Academy of Sciences* 109 (37): E2424–32. https://doi.org/10.1073/pnas.1207846109.

Lee, Yun-Young, Randal C. Cevallos, and Eric Jan. 2009. “An Upstream Open Reading Frame Regulates Translation of GADD34 during Cellular Stresses That Induce EIF2alpha Phosphorylation.” *The Journal of Biological Chemistry* 284 (11): 6661–73. https://doi.org/10.1074/jbc.M806735200.

Levenson, Anait S., Kenneth E. Thurn, Laura A. Simons, Dorina Veliceasa, Jennifer Jarrett, Clodia Osipo, V. Craig Jordan, Olga V. Volpert, Robert L. Satcher, and Ronald B. Gartenhaus. 2005. “MCT-1 Oncogene Contributes to Increased in Vivo Tumorigenicity of MCF7 Cells by Promotion of Angiogenesis and Inhibition of Apoptosis.” *Cancer Research* 65 (23): 10651–56. https://doi.org/10.1158/0008-5472.CAN-05-0845.

Lin, Yingying, Fajin Li, Linlu Huang, Christine Polte, Haoran Duan, Jianhuo Fang, Li Sun, et al. 2020. “EIF3 Associates with 80S Ribosomes to Promote Translation Elongation, Mitochondrial Homeostasis, and Muscle Health.” *Molecular Cell* 79 (4): 575-587.e7. https://doi.org/10.1016/j.molcel.2020.06.003.

Liu, Shihui, Christopher Bachran, Pradeep Gupta, Sharmina Miller-Randolph, Hailun Wang, Devorah Crown, Yi Zhang, et al. 2012. “Diphthamide Modification on Eukaryotic Elongation Factor 2 Is Needed to Assure Fidelity of MRNA Translation and Mouse Development.” *Proceedings of the National Academy of Sciences* 109 (34): 13817–22. https://doi.org/10.1073/pnas.1206933109.

Liu, Tzu-Yu, Hector H. Huang, Diamond Wheeler, Yichen Xu, James A. Wells, Yun S. Song, and Arun P. Wiita. 2017. “Time-Resolved Proteomics Extends Ribosome Profiling-Based Measurements of Protein Synthesis Dynamics.” *Cell Systems* 4 (6): 636-644.e9. https://doi.org/10.1016/j.cels.2017.05.001.

Lomakin, Ivan B., Sergey E. Dmitriev, and Thomas A. Steitz. 2018. “Crystal Structure of the DENR-MCT-1 Complex Revealed Zinc-Binding Site Essential for Heterodimer Formation.” *Proceedings of the National Academy of Sciences* 116 (2): 528–33. https://doi.org/10.1073/pnas.1809688116.

Lu, Phoebe D., Heather P. Harding, and David Ron. 2004. “Translation Reinitiation at Alternative Open Reading Frames Regulates Gene Expression in an Integrated Stress Response.” *The Journal of Cell Biology* 167 (1): 27–33. https://doi.org/10.1083/jcb.200408003.

Madhugiri, R., M. Fricke, M. Marz, and J. Ziebuhr. 2016. “Coronavirus Cis-Acting RNA Elements.” *Advances in Virus Research* 96:127–63. https://doi.org/10.1016/bs.aivir.2016.08.007.

Marchase, Richard B., Lillian A. Koro, Charles M. Kelly, and David R. McClay. 1982. “Retinal Ligatin Recognizes Glycoproteins Bearing Oligosaccharides Terminating in Phosphodiester-Linked Glucose.” *Cell* 28 (4): 813–20. https://doi.org/10.1016/0092-8674(82)90060-5.

Mazan-Mamczarz, Krystyna, Patrick Hagner, Bojie Dai, Sharon Corl, Zhenqui Liu, and Ron B. Gartenhaus. 2009. “Targeted Suppression of MCT-1 Attenuates the Malignant Phenotype through a Translational Mechanism.” *Leukemia Research* 33 (3): 474–82. https://doi.org/10.1016/j.leukres.2008.08.012.

McCole, Ruth B, Noeleen B Loughran, Mandeep Chahal, Luis P Fernandes, Roland G Roberts, Franca Fraternali, Mary J O’Connell, and Rebecca J Oakey. 2011. “A Case-by-Case Evolutionary Analysis of Four Imprinted Retrogenes.” *Evolution; International Journal of Organic Evolution* 65 (5): 1413–27. https://doi.org/10.1111/j.1558-5646.2010.01213.x.

Mohammad, Mahabub Pasha, Vanda Munzarová Pondělíčková, Jakub Zeman, Stanislava Gunišová, and Leoš Shivaya Valášek. 2017. “In Vivo Evidence That EIF3 Stays Bound to Ribosomes Elongating and Terminating on Short Upstream ORFs to Promote Reinitiation.” *Nucleic Acids Research* 45 (5): 2658–74. https://doi.org/10.1093/nar/gkx049.

Nanda, Jagpreet S., Adesh K. Saini, Antonio M. Muñoz, Alan G. Hinnebusch, and Jon R. Lorsch. 2013. “Coordinated Movements of Eukaryotic Translation Initiation Factors EIF1, EIF1A, and EIF5 Trigger Phosphate Release from EIF2 in Response to Start Codon Recognition by the Ribosomal Preinitiation Complex.” *The Journal of Biological Chemistry* 288 (8): 5316–29. https://doi.org/10.1074/jbc.M112.440693.

Neurohr, Gabriel E., and Angelika Amon. 2020. “Relevance and Regulation of Cell Density.” *Trends in Cell Biology* 30 (3): 213–25. https://doi.org/10.1016/j.tcb.2019.12.006.

NIAID Integrated Research Facility. 2020. “Transmission Electron Micrograph of SARS-CoV-2 Virus Particles, Isolated from a Patient.” https://www.nih.gov/sites/default/files/styles/floated_media_breakpoint-large/public/news-events/research-matters/2023/20230613-covid.jpg?itok=SiLHuKM-&timestamp=1686667220.

Occhi, Gianluca, Daniela Regazzo, Giampaolo Trivellin, Francesca Boaretto, Denis Ciato, Sara Bobisse, Sergio Ferasin, et al. 2013. “A Novel Mutation in the Upstream Open Reading Frame of the CDKN1B Gene Causes a MEN4 Phenotype.” *PLoS Genetics* 9 (3): e1003350. https://doi.org/10.1371/journal.pgen.1003350.

Passmore, Lori A., T. Martin Schmeing, David Maag, Drew J. Applefield, Michael G. Acker, Mikkel A. Algire, Jon R. Lorsch, and V. Ramakrishnan. 2007. “The Eukaryotic Translation Initiation Factors EIF1 and EIF1A Induce an Open Conformation of the 40S Ribosome.” *Molecular Cell* 26 (1): 41–50. https://doi.org/10.1016/j.molcel.2007.03.018.

Paulin, F. E., L. E. Campbell, K. O’Brien, J. Loughlin, and C. G. Proud. 2001. “Eukaryotic Translation Initiation Factor 5 (EIF5) Acts as a Classical GTPase-Activator Protein.” *Current Biology: CB* 11 (1): 55–59. https://doi.org/10.1016/s0960-9822(00)00025-7.

Peixeiro, Isabel, Ângela Inácio, Cristina Barbosa, Ana Luísa Silva, Stephen A. Liebhaber, and Luísa Romão. 2012. “Interaction of PABPC1 with the Translation Initiation Complex Is Critical to the NMD Resistance of AUG-Proximal Nonsense Mutations.” *Nucleic Acids Research* 40 (3): 1160–73. https://doi.org/10.1093/nar/gkr820.

Pestova, Tatyana V., and Victoria G. Kolupaeva. 2002. “The Roles of Individual Eukaryotic Translation Initiation Factors in Ribosomal Scanning and Initiation Codon Selection.” *Genes & Development* 16 (22): 2906–22. https://doi.org/10.1101/gad.1020902.

Pisareva, Vera P., Ilham A. Muslimov, Andrew Tcherepanov, and Andrey V. Pisarev. 2015. “Characterization of Novel Ribosome-Associated Endoribonuclease SLFN14 from Rabbit Reticulocytes.” *Biochemistry* 54 (21): 3286–3301. https://doi.org/10.1021/acs.biochem.5b00302.

Plant, Ewan P., and Jonathan D. Dinman. 2008. “The Role of Programmed-1 Ribosomal Frameshifting in Coronavirus Propagation.” *Frontiers in Bioscience : A Journal and Virtual Library* 13 (May):4873–81. https://doi.org/10.2741/3046.

Plaza, Serge, Menschaert, Gerben, and Payre, Francois. 2017. “In Search of Lost Small Peptides.” *Annual Review of Cell and Developmental Biology* 33 (1):391-416. 10.1146/annurev-cellbio-100616-060516.

Prosniak, Misha, Jamil Dierov, Kenji Okami, Brian Tilton, Brad Jameson, Bassel E. Sawaya, and Ronald B. Gartenhaus. 1998. “A Novel Candidate Oncogene, MCT-1, Is Involved in Cell Cycle Progression.” *Cancer Research* 58 (19): 4233–37.

Rakotondrafara, Aurélie M., and Matthias W. Hentze. 2011. “An Efficient Factor-Depleted Mammalian in Vitro Translation System.” *Nature Protocols* 6 (5): 563–71. https://doi.org/10.1038/nprot.2011.314.

Rehfeld, Frederick. 2023. “CRISPR Screening Reveals a Dependency on Ribosome Recycling for Efficient SARS-CoV-2 Programmed Ribosomal Frameshifting and Viral Replication.” *Cell Reports* 42 (2): 112076. https://doi.org/10.1016/j.celrep.2023.112076.

Rojas, Margarito, Gabriel Vasconcelos, and Thomas E. Dever. 2015. “An EIF2α-Binding Motif in Protein Phosphatase 1 Subunit GADD34 and Its Viral Orthologs Is Required to Promote Dephosphorylation of EIF2α.” *Proceedings of the National Academy of Sciences* 112 (27): E3466–75. https://doi.org/10.1073/pnas.1501557112.

Romão, Luísa, Ângela Inácio, Susana Santos, Madalena Ávila, Paula Faustino, Paula Pacheco, and João Lavinha. 2000. “Nonsense Mutations in the Human β-Globin Gene Lead to Unexpected Levels of Cytoplasmic MRNA Accumulation.” *Blood* 96 (8): 2895–2901. https://doi.org/10.1182/blood.V96.8.2895.

Rozman, Batsheva, Tal Fisher, and Noam Stern-Ginossar. 2023. “Translation—A Tug of War during Viral Infection.” *Molecular Cell* 83 (3): 481–95. https://doi.org/10.1016/j.molcel.2022.10.012.

Russell, Paul J., Jacob A. Slivka, Elaina P. Boyle, Arthur H.M. Burghes, and Michael G. Kearse. 2023. “Translation Reinitiation after UORFs Does Not Fully Protect MRNAs from Nonsense-Mediated Decay.” *RNA* 29 (6): 735–44. https://doi.org/10.1261/rna.079525.122.

Sanz, Miguel Angel, Esther González Almela, Manuel García-Moreno, Ana Isabel Marina, and Luis Carrasco. 2019. “A Viral RNA Motif Involved in Signaling the Initiation of Translation on Non-AUG Codons.” *RNA (New York, N.Y.)* 25 (4): 431–52. https://doi.org/10.1261/rna.068858.118.

Sanz, Miguel Angel, Esther González Almela, and Luis Carrasco. 2017. “Translation of Sindbis Subgenomic MRNA Is Independent of EIF2, EIF2A and EIF2D.” *Scientific Reports* 7 (February):43876. https://doi.org/10.1038/srep43876.

Schleich, Sibylle, Julieta M. Acevedo, Katharina Clemm Von Hohenberg, and Aurelio A. Teleman. 2017. “Identification of Transcripts with Short StuORFs as Targets for DENR•MCTS1-Dependent Translation in Human Cells.” *Scientific Reports* 7 (1): 3722. https://doi.org/10.1038/s41598-017-03949-6.

Schleich, Sibylle, Katrin Strassburger, Philipp Christoph Janiesch, Tatyana Koledachkina, Katharine K. Miller, Katharina Haneke, Yong-Sheng Cheng, et al. 2014. “DENR–MCT-1 Promotes Translation Re-Initiation Downstream of UORFs to Control Tissue Growth.” *Nature* 512 (7513): 208–12. https://doi.org/10.1038/nature13401.

Schubert, Katharina, Evangelos D. Karousis, Ahmad Jomaa, Alain Scaiola, Blanca Echeverria, Lukas-Adrian Gurzeler, Marc Leibundgut, Volker Thiel, Oliver Mühlemann, and Nenad Ban. 2020. “SARS-CoV-2 Nsp1 Binds the Ribosomal MRNA Channel to Inhibit Translation.” *Nature Structural & Molecular Biology* 27 (10): 959–66. https://doi.org/10.1038/s41594-020-0511-8.

Schulz, Julia, Nancy Mah, Martin Neuenschwander, Tabea Kischka, Richard Ratei, Peter M. Schlag, Esmeralda Castaños-Vélez, et al. 2018. “Loss-of-Function UORF Mutations in Human Malignancies.” *Scientific Reports* 8 (February):2395. https://doi.org/10.1038/s41598-018-19201-8.

Schwanhäusser, Björn, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. 2011. “Global Quantification of Mammalian Gene Expression Control.” *Nature* 473 (7347): 337–42. https://doi.org/10.1038/nature10098.

She, Richard, Jingchuan Luo, and Jonathan S Weissman. 2023. “Translational Fidelity Screens in Mammalian Cells Reveal EIF3 and EIF4G2 as Regulators of Start Codon Selectivity.” *Nucleic Acids Research* 51 (12): 6355–69. https://doi.org/10.1093/nar/gkad329.

Shestakova, Ekaterina D., Roman S. Tumbinsky, Dmitri E. Andreev, Fedor N. Rozov, Ivan N. Shatsky, and Ilya M. Terenin. 2023. “The Roles of EIF4G2 in Leaky Scanning and Reinitiation on the Human Dual-Coding POLG MRNA.” *International Journal of Molecular Sciences* 24 (24): 17149. https://doi.org/10.3390/ijms242417149.

Shi, Bo, Hsin-Ling Hsu, Andy M. Evens, Leo I. Gordon, and Ronald B. Gartenhaus. 2003. “Expression of the Candidate *MCT-1* Oncogene in B- and T-Cell Lymphoid Malignancies.” *Blood* 102 (1): 297–302. https://doi.org/10.1182/blood-2002-11-3486.

Shirokikh, Nikolay E., and Thomas Preiss. 2018. “Translation Initiation by Cap-Dependent Ribosome Recruitment: Recent Insights and Open Questions.” *Wiley Interdisciplinary Reviews. RNA* 9 (4): e1473. https://doi.org/10.1002/wrna.1473.

Skabkin, M. A., O. V. Skabkina, V. Dhote, A. A. Komar, C. U. T. Hellen, and T. V. Pestova. 2010. “Activities of Ligatin and MCT-1/DENR in Eukaryotic Translation Initiation and Ribosomal Recycling.” *Genes & Development* 24 (16): 1787–1801. https://doi.org/10.1101/gad.1957510.

Smirnova, Victoria V, Ekaterina D Shestakova, Daria S Nogina, Polina A Mishchenko, Tatiana A Prikazchikova, Timofei S Zatsepin, Ivan V Kulakovskiy, Ivan N Shatsky, and Ilya M Terenin. 2022. “Ribosomal Leaky Scanning through a Translated UORF Requires EIF4G2.” *Nucleic Acids Research* 50 (2): 1111–27. https://doi.org/10.1093/nar/gkab1286.

Sonobe, Yoshifumi, Aburas, Jihad, Krishnan, Gopinath, Fleming, Andrew C., Ghadge, Ghanashyam, Islam, Priota, Warren, Eleanor C., Gu, Yuanzheng, Kankel, Mark W., Brown, André E. X., Kiskinis, Evangelos, Gendron, Tania F., Gao, Fen-Biao. Roos, Raymond P. and Kratsios, Paschalis. 2021. “A C. Elegans Model of C9orf72-Associated ALS/FTD Uncovers a Conserved Role for EIF2D in RAN Translation.” *Nature Communications* 12 (1): 6025 https://doi.org/10.1038/s41467-021-26303-x

Sugiyama, Hayami, Kazutoshi Takahashi, Takuya Yamamoto, Mio Iwasaki, Megumi Narita, Masahiro Nakamura, Tim A. Rand, Masato Nakagawa, Akira Watanabe, and Shinya Yamanaka. 2017. “Nat1 Promotes Translation of Specific Proteins That Induce Differentiation of Mouse Embryonic Stem Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (2): 340–45. https://doi.org/10.1073/pnas.1617234114.

Sun, Yu, Laura Abriola, Rachel O. Niederer, Savannah F. Pedersen, Mia M. Alfajaro, Valter Silva Monteiro, Craig B. Wilen, et al. 2021. “Restriction of SARS-CoV-2 Replication by Targeting Programmed −1 Ribosomal Frameshifting.” *Proceedings of the National Academy of Sciences* 118 (26): e2023051118. https://doi.org/10.1073/pnas.2023051118.

Thakur, Anil, and Alan G. Hinnebusch. 2018. “EIF1 Loop 2 Interactions with Met-TRNAi Control the Accuracy of Start Codon Selection by the Scanning Preinitiation Complex.” *Proceedings of the National Academy of Sciences* 115 (18): E4159–68. https://doi.org/10.1073/pnas.1800938115.

Thoms, Matthias, Robert Buschauer, Michael Ameismeier, Lennart Koepke, Timo Denk, Maximilian Hirschenberger, Hanna Kratzat, et al. 2020. “Structural Basis for Translational Shutdown and Immune Evasion by the Nsp1 Protein of SARS-CoV-2.” *Science (New York, N.Y.)* 369 (6508): 1249–55. https://doi.org/10.1126/science.abc8665.

Tyrrell, David Arthur John, and Michael Fielder. 2002. *Cold Wars: The Fight Against the Common Cold*. Oxford University Press.

Vaidya. Anand T.,Lomakin, Ivan B., Joseph, Newlyn N., Dmitriev, Sergey E., Steitz, Thomas A. 2017. “Crystal Structure of the C-Terminal Domain of Human EIF2D and Its Implications on Eukaryotic Translation Initiation.” *Nature Communications* 429 (18): 2765-71. https://doi.org/10.1016/j.jmb.2017.07.015.

Vasudevan, Deepika, Sarah D. Neuman, Amy Yang, Lea Lough, Brian Brown, Arash Bashirullah, Timothy Cardozo, and Hyung Don Ryoo. 2020. “Translational Induction of ATF4 during Integrated Stress Response Requires Noncanonical Initiation Factors EIF2D and DENR.” *Nature Communications* 11 (1): 4677. https://doi.org/10.1038/s41467-020-18453-1.

Vattem, Krishna M., and Ronald C. Wek. 2004. “Reinitiation Involving Upstream ORFs Regulates ATF4 MRNA Translation in Mammalian Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (31): 11269–74. https://doi.org/10.1073/pnas.0400541101.

V’kovski, Philip, Annika Kratzel, Silvio Steiner, Hanspeter Stalder, and Volker Thiel. 2021. “Coronavirus Biology and Replication: Implications for SARS-CoV-2.” *Nature Reviews. Microbiology* 19 (3): 155–70. https://doi.org/10.1038/s41579-020-00468-6.

Wang, Jinfan, Byung-Sik Shin, Carlos Alvarado, Joo-Ran Kim, Jonathan Bohlen, Thomas E. Dever, and Joseph D. Puglisi. 2022. “Rapid 40S Scanning and Its Regulation by MRNA Structure during Eukaryotic Translation Initiation.” *Cell* 185 (24): 4474-4487.e17. https://doi.org/10.1016/j.cell.2022.10.005.

Watatani, Yujiro, Kenji Ichikawa, Noriko Nakanishi, Maki Fujimoto, Hitoshi Takeda, Natsumi Kimura, Hidenori Hirose, Shigeru Takahashi, and Yuji Takahashi. 2008. “Stress-Induced Translation of ATF5 MRNA Is Regulated by the 5′-Untranslated Region *.” *Journal of Biological Chemistry* 283 (5): 2543–53. https://doi.org/10.1074/jbc.M707781200.

Weber, Ramona, Leon Kleemann, Insa Hirschberg, Min-Yi Chung, Eugene Valkov, and Cátia Igreja. 2022. “DAP5 Enables Main ORF Translation on MRNAs with Structured and UORF-Containing 5′ Leaders.” *Nature Communications* 13 (December):7510. https://doi.org/10.1038/s41467-022-35019-5.

Weisser, Melanie, Tanja Schäfer, Marc Leibundgut, Daniel Böhringer, Christopher Herbert Stanley Aylett, and Nenad Ban. 2017. “Structural and Functional Insights into Human Re-Initiation Complexes.” *Molecular Cell* 67 (3): 447-456.e7. https://doi.org/10.1016/j.molcel.2017.06.032.

Wek, Ronald C., Tracy G. Anthony, and Kirk A. Staschke. 2023. “Surviving and Adapting to Stress: Translational Control and the Integrated Stress Response.” *Antioxidants & Redox Signaling* 39 (4–6): 351–73. https://doi.org/10.1089/ars.2022.0123.

Wethmar, Klaus. 2014. “The Regulatory Potential of Upstream Open Reading Frames in Eukaryotic Gene Expression.” *Wiley Interdisciplinary Reviews. RNA* 5 (6): 765–78. https://doi.org/10.1002/wrna.1245.

Wood, Andrew J., Roland G. Roberts, David Monk, Gudrun E. Moore, Reiner Schulz, and Rebecca J. Oakey. 2007. “A Screen for Retrotransposed Imprinted Genes Reveals an Association between X Chromosome Homology and Maternal Germ-Line Methylation.” *PLOS Genetics* 3 (2): e20. https://doi.org/10.1371/journal.pgen.0030020.

Wright, Bradley W., Zixin Yi, Jonathan S. Weissman, and Jin Chen. 2022. “The Dark Proteome: Translation from Noncanonical Open Reading Frames.” *Trends in Cell Biology* 32 (3): 243–58. https://doi.org/10.1016/j.tcb.2021.10.010.

Wu, Jiecong, Wenqi Wu, Ping Jiang, Yuhao Xu, and Ming Yu. 2024. “Identification of SV2C and DENR as Key Biomarkers for Parkinson’s Disease Based on Bioinformatics, Machine Learning, and Experimental Verification.” *Journal of Molecular Neuroscience* 74 (1): 6. https://doi.org/10.1007/s12031-023-02182-3.

Xiang, Yezi, Wenze Huang, Lianmei Tan, Tianyuan Chen, Yang He, Patrick S. Irving, Kevin M. Weeks, Qiangfeng Cliff Zhang, and Xinnian Dong. 2023. “Pervasive Downstream RNA Hairpins Dynamically Dictate Start-Codon Selection.” *Nature* 621 (7978): 423–30. https://doi.org/10.1038/s41586-023-06500-y.

Xu, Benjin, Ling Liu, and Guangtao Song. 2022. “Functions and Regulation of Translation Elongation Factors.” *Frontiers in Molecular Biosciences* 8 (January):816398. https://doi.org/10.3389/fmolb.2021.816398.

Xue, Shifeng, and Maria Barna. 2015. “Cis-Regulatory RNA Elements That Regulate Specialized Ribosome Activity.” *RNA Biology* 12 (10): 1083–87. https://doi.org/10.1080/15476286.2015.1085149.

Yang, Mo, Feyisola P. Olatunji, Curran Rhodes, Sumirtha Balaratnam, Kara Dunne-Dombrink, Srinath Seshadri, Xiao Liang, et al. 2023. “Discovery of Small Molecules Targeting the Frameshifting Element RNA in SARS-CoV-2 Viral Genome.” *ACS Medicinal Chemistry Letters* 14 (6): 757–65. https://doi.org/10.1021/acsmedchemlett.3c00051.

Young, David J., and Nicholas R. Guydosh. 2019. “Hcr1/EIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo.” *Cell Reports* 28 (1): 39-50.e4. https://doi.org/10.1016/j.celrep.2019.05.111.

Young, David J., Desislava S. Makeeva, Fan Zhang, Aleksandra S. Anisimova, Elena A. Stolboushkina, Fardin Ghobakhlou, Ivan N. Shatsky, Sergey E. Dmitriev, Alan G. Hinnebusch, and Nicholas R. Guydosh. 2018. “Tma64/EIF2D, Tma20/MCT-1, and Tma22/DENR Recycle Post-Termination 40S Subunits In Vivo.” *Molecular Cell* 71 (5): 761-774.e5. https://doi.org/10.1016/j.molcel.2018.07.028.

Young, David J., Sezen Meydan, and Nicholas R. Guydosh. 2021. “40S Ribosome Profiling Reveals Distinct Roles for Tma20/Tma22 (MCT-1/DENR) and Tma64 (EIF2D) in 40S Subunit Recycling.” *Nature Communications* 12 (1): 2976. https://doi.org/10.1038/s41467-021-23223-8.

Young, Sara K., Jeffrey A. Willy, Cheng Wu, Matthew S. Sachs, and Ronald C. Wek. 2015. “Ribosome Reinitiation Directs Gene-Specific Translation and Regulates the Integrated Stress Response.” *The Journal of Biological Chemistry* 290 (47): 28257–71. https://doi.org/10.1074/jbc.M115.693184.

Zhang, Hong, Yirong Wang, Xinkai Wu, Xiaolu Tang, Changcheng Wu, and Jian Lu. 2021. “Determinants of Genome-Wide Distribution and Evolution of UORFs in Eukaryotes.” *Nature Communications* 12 (1): 1076. https://doi.org/10.1038/s41467-021-21394-y.

Zhang, J, and L E Maquat. 1997. “Evidence That Translation Reinitiation Abrogates Nonsense-Mediated MRNA Decay in Mammalian Cells.” *The EMBO Journal* 16 (4): 826–33. https://doi.org/10.1093/emboj/16.4.826.

Zhang, Jun, and Yuguang Shi. 2024. “An Upstream Open Reading Frame (5′-UORF) Links Oxidative Stress to Translational Control of ALCAT1 through Phosphorylation of EIF2α.” *Free Radical Biology and Medicine* 214 (March):129–36. https://doi.org/10.1016/j.freeradbiomed.2024.02.015.

Zhijun, H., H. Zhengbin, Z. Fengwei, H. Hongjuan, Y. Shihuan, and W. Qiong. 2014. “Spatiotemporal Expression of Retrogene–Host Pair *Mcts2/H13* in Mouse Embryo, and *Mcts2* Has No Influence on *H13* Transcription Pattern in NIH/3T3 Cells.” *Acta Histochemica* 116 (2): 312–18. https://doi.org/10.1016/j.acthis.2013.08.008.

Zhou, Donghui, L. Reddy Palam, Li Jiang, Jana Narasimhan, Kirk A. Staschke, and Ronald C. Wek. 2008. “Phosphorylation of EIF2 Directs *ATF5* Translational Control in Response to Diverse Stress Conditions*.” *Journal of Biological Chemistry* 283 (11): 7064–73. https://doi.org/10.1074/jbc.M708530200.

Zinoviev, Alexandra, Christopher U.T. Hellen, and Tatyana V. Pestova. 2015. “Multiple Mechanisms of Reinitiation on Bicistronic Calicivirus MRNAs.” *Molecular Cell* 57 (6): 1059–73. https://doi.org/10.1016/j.molcel.2015.01.039.

Research articles

(1) **Romane Meurs**, Mara de Matos, Nicolas Guex, David Gatfield. *An in vitro assay of MCTS1-DENR-dependent re-initiation and ribosome profiling uncover the activity of MCTS2 and distinct function of eIF2D.* In preparation (2024)

(2) Pramod R. Bhatt, Alain Scaiola, Gary Loughran, Marc Leibundgut, Annika Kratzel, **Romane Meurs**, René Dreos, Kate M. O'Connor, Angus McMillan, Jeffrey W. Bode, Volker Thiel, David Gatfield, John F. Atkins, Nenad Ban. *Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome.* Science (2021) DOI: 10.1126/science.abf3546.

An in vitro assay of MCTS1-DENR-dependent re-initiation and ribosome profiling uncover the activity of MCTS2 and distinct function of eIF2D.

Romane Meurs, Mara De Matos, Nicolas Guex, David Gatfield

In preperation

Contribution to the study

David Gatfield and I designed this study. I performed all the experiments presented in this article manuscript with minor exceptions. Some of the molecular clonings were carried out by Mara De Matos. The ribosome profiling libraries were sequenced by the Lausanne Genomic Technologies Facility (GTF) and the mass spectrometry experiment was performed by the Protein Analysis Facility (PAF). The HeLa *Denr*, *Eif2d* and *Denr/Eif2d* double KO cell lines were kindly provided by the lab of Aurelio Teleman and the recombinant MCTS1, DENR, MCTS2 and eIF2D were generated by Adrian Bothe from Nenad Ban's lab. I carried out amost all the data analyses presented in this study, including the bioinformatic analyses for which I received the valuable assistance of Bulak Arpat and René Dreos. Nicolas Guex performed the structural analysis. I designed all the figures with the major contribution of

David Gatfield and valuable inputs from the lab members. I wrote the paper draft which was proofread and significantly improved by David Gatfield.

Contribution of the Pr. Aurelio Teleman, expert in the thesis committee, to the study

Following the mid-thesis exam discussion, Aurelio Teleman kindly donated the *Denr*, *Eif2d* and *Denr/Eif2d* double KO HeLa cell lines, previously generated in his lab, for the PhD project. He did not further contribute to the experiments, analyses and interpretation performed using these cell lines. We believe that this contribution will not influence his impartiality during the final thesis exam.

An *in vitro* assay of MCTS1-DENR-dependent re-initiation and ribosome profiling uncover the activity of MCTS2 and distinct function of eIF2D.

Romane Meurs[1], Mara De Matos[1], Nicolas Guex[2], David Gatfield[1*]

Affiliations:

[1] Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

[2] Bioinformatics Competence Center, University of Lausanne, 1015 Lausanne, Switzerland

Email addresses: romane.meurs@unil.ch, mara.dematos@unil.ch, david.gatfield@unil.ch

Abstract

Ribosomes scanning from the mRNA 5’ cap to the start codon may initiate at upstream open reading frames (uORFs), decreasing protein biosynthesis. Termination at a uORF can lead to re-initiation, where the 40S subunit resumes scanning and initiates another translation event downstream. In mammals, the noncanonical translation factors MCTS1-DENR participate in re-initiation at specific uORFs, but knowledge of other *trans*-acting factors and uORF features influencing re-initiation is limited. Here, we describe a cell-free re-initiation assay using HeLa cell lysates. Comparing *in vivo* and *in vitro* re-initiation activities on uORF-containing model reporters, we validate that MCTS1-DENR-dependent re-initiation is accurately recapitulated *in vitro*. Using this system and ribosome profiling in cultured cells, we found that knockdown of the homolog eIF2D causes widespread gene expression deregulation unrelated to uORF translation, suggesting distinct functions from MCTS1-DENR. Additionally, we identified MCTS2, encoded by a retrogene copy of *Mcts1*, as an alternative DENR partner that promotes re-initiation *in vitro*. Our findings on re-initiation and the new assay provide valuable insights and tools for future research on uORF features and *trans*-acting factors.

Introduction

Upstream open reading frames (uORFs) are abundant, short regulatory elements found within the 5' UTRs of an estimated 50% of human transcripts [1]. uORF translation has long been recognised as a mechanism to tune the translation efficiency (TE) of the protein coding sequence (CDS) under physiological and stress conditions [2-5], and several cases of misregulated uORF-dependent translational control have been linked to human diseases including cancer [6-9]. Because uORFs intercept scanning ribosomes and reduce the probability that they engage in initiation downstream, they are generally inhibitory to protein biosynthesis from the CDS. Two main mechanisms are implicated in ensuring that CDS translation can take place nevertheless: first, in a process called leaky scanning, the 43S preinitiation complex can scan across the uORF start codon, thus ignoring the uORF; second, the ribosome can translate the uORF, terminate at its stop codon, and then engage in the process of re-initiation, where the small ribosomal subunit resumes scanning to a downstream start codon [1]. In contrast to canonical initiation, for which many implicated initiation factors and mechanistic details have been identified over the last decades, our understanding of re-initiation has

remained incomplete. Thus, large knowledge gaps exist regarding the *trans*-acting protein machinery and to what extent it is shared with canonical initiation, and the *cis*-acting mRNA and uORF sequence requirements that predispose to re-initiation.

An peculiar case of re-initiation-competent uORFs that has been studied contains only a single amino acid – i.e., the start codon – followed immediately by a stop codon (termed '1aa uORF' or 'start-stop uORF' in the following) [10, 11]. In regular initiation, the initiator-tRNA is positioned in the ribosomal P-site, and on 1aa uORFs the stop codon is thus already placed in the A-site, ready for termination and without the ribosome ever entering into elongation. Hence, and as shown for eIF3 [12]. some canonical initiation factors from the uORF initiation event may likely still be available on the ribosome for subsequent re-initiation. Whether and under which circumstances eIFs remain on or rejoin the ribosome in the case of longer, elongating uORFs in order to become competent for a second round of scanning has remained unclear.

The heterodimeric, non-canonical initiation factor MCTS1-DENR has been identified as the first, re-initiation-specific translation factor with conserved activity in *Drosophila*, mouse and human [10, 11, 13-16]. *In vitro*, MCTS1-DENR shows the double activity of being able to recruit and release tRNA from the ribosomal P-site [13, 17], and the yeast orthologues, Tma20-Tma22, appear to mainly function in post-termination tRNA removal and 40S ribosome recycling rather than as promoters of re-initiation [18]. In animal model systems, DENR loss-of-function followed by ribosome profiling has been used to identify CDS translation efficiency changes indicative of altered re-initiation, thus identifying likely direct MCTS1-DENR targets [10, 11, 16, 19]. These studies revealed that only a small fraction of all uORF-containing transcripts undergo MCTS1-DENR-dependent re-initiation and, while start-stop uORFs are highly represented among targets, longer uORFs were frequently associated with MCTS1-DENR-dependent re-initiation as well. As a first hint towards the selectivity of MCTS1-DENR for certain uORFs, an enrichment for specific penultimate uORF codons was observed among targets in HeLa cells [19], leading to the hypothesis that MCTS1-DENR removes

empty P-site tRNA from the terminated ribosome and that the efficiency of this process may be tRNA-specific.

The protein eIF2D is a close homologue of MCTS1-DENR, containing the same domains encoded on a single polypeptide. Like MCTS1-DENR, eIF2D (then still under the name LGTN, for Ligatin) was initially identified as a recycling and initiation factor based on its *in vitro* activities [13], and the yeast orthologue, Tma64, has been noted for its role in 40S subunit recycling, too [18]. Due to these similarities, eIF2D is frequently referred to as a re-initiation factor as well, and evidence for comparable functions of MCTS1-DENR and eIF2D on uORFs come from studies of *Atf4* regulation in *Drosophila* and in HeLa cells [19, 20]. However, eIF2D does not regulate an MCTS1-DENR-dependent synthetic model start-stop uORF *in vivo* [11], suggesting that the two factors may be active on distinct sets of uORFs and transcripts. Furthermore, it has recently been reported that the single knockout of yeast Tma64/*Eif2d* does not recapitulate the 40S recycling defect observed in the single knockout of Tma22/*Denr* [21]. Therefore, it would appear that the involvement of eIF2D in re-initiation and/or recycling still needs to be formally established.

In this study, we combined ribosome profiling, *in vivo* reporter assays and experiments using a cell-free translation system to investigate MCTS1-DENR- and eIF2D-dependent re-initiation. Using the *in vivo* and *in vitro* re-initiation assays, we measured the fluxes of ribosomes on the 5' UTRs of two identified DENR-dependent model transcripts, *Klhdc8a* and *Asb8*, and addressed the effect of penultimate codon identity. Ribosome profiling from DENR and eIF2D loss-of-function cells pointed at different and distinct roles for the two factors in gene expression regulation and, importantly, indicated that eIF2D does not act as a general uORF re-initiation factor. Finally, we identified MCTS2 as an alternative DENR heterodimerisation partner *in vivo* whose activity in re-initiation we analysed and validated in our *in vitro* assay.

Results

Upstream open reading frame-containing transcripts *Klhdc8a* and *Asb8* are regulated by DENR *in vivo* and suitable for model reporter design.

To gain insights into mechanism and targets of DENR-dependent re-initiation, we first identified uORF-containing transcripts for which CDS translation was altered after shRNA-mediated knock-down of *Denr* in murine NIH/3T3 cells (**Figure 1A**). To this end, we carried out ribosome profiling (Ribo-seq) and RNA-seq from *Denr* shRNA-treated cells, as well as from two different control-treated cells, exposed either to scramble (Scr) shRNA or to non-functional (nf) *Eif2d* shRNA that did not cause any knockdown (**Figure 1A**). From the obtained data, we calculated translation efficiencies (TEs, ratio CDS footprints to RNA) transcriptome-wide, thus identifying 223 and 6 transcripts, respectively, with significantly reduced and increased TE (applying FDR-corrected p<0.1) (**Figure 1B;** Supplementary Table S1). The strong bias for reduced TEs was in line with our expectations for direct targets of a factor facilitating re-initiation on the CDS after uORF translation, and with previous findings [16, 19]; we further noted that the annotated 5’ UTRs of transcripts with reduced TE were significantly longer than would be expected from the transcriptome-wide distribution of 5' UTR lengths (**Figure 1C**), compatible with increased uORF content. Indeed, annotation of translated uORFs from the footprint data showed strong enrichment among the DENR-responsive transcripts (**Figure 1D**). The lengths of the identified uORFs showed a broad distribution, with 1aa/start-stop uORFs being the most abundant species but many longer, elongating uORFs found on the DENR-responsive transcripts as well (**Figure 1E**). We selected two transcripts to design re-initiation reporters, representing an elongating uORF (*Klhdc8a, Kelch domain containing 8A*; 13 amino acid uORF) and a start-stop uORF (*Asb8, Ankyrin repeat and SOCS box-containing 8*) (**Figure 1B**). Both mRNAs presented significantly reduced TE upon *Denr* knockdown (decrease by ~70% for *Klhdc8a* and by ~30% for *Asb8*; **Figure 1G**) and, moreover, their 5' UTRs contained only one strongly translated, AUG-initiating uORF (**Figure 1F** and **Supplementary Figure S1A, B**). Furthermore, the uORFs are conserved in mammals (**Supplementary Figure S1C, D**), including in humans, and *Denr* KO in

human HeLa cells indeed led to a strong downregulation of endogenous KLHDC8A protein as evaluated by immunoblot analysis using specific antibodies (**Figure 1J**). The lack of functional antibodies against mouse or human ASB8 precluded a similar analysis for this protein.

We first used the *Klhdc8a* 5' UTR (and the first codons of its CDS, fused in-frame to firefly luciferase) to establish a set of three different reporter genes that would allow us to estimate re-initiation and leaky scanning rates (**Figure 1H**). Thus, the *WT uORF* reporter contained the full, unmodified *Klhdc8a* 5' UTR. Mutating the uORF start codon (*no uORF* reporter) allowed us to quantify luciferase signal when no inhibitory uORF is present. A third reporter contained a mutated uORF stop codon, which extended the uORF coding sequence all the way into the luciferase CDS, but in a different frame (*overlapping uORF* reporter). Thus, this reporter would give luciferase signal only through leaky scanning of ribosomes across the uORF start codon. After reporter delivery into HeLa cells (via lentiviral constructs that also contained a Renilla luciferase reporter for internal normalisation across samples), re-initiation signal can be extracted from the individual reporter readouts as shown in **Figure 1I**. The relative reduction of signal from *no uORF* to *WT uORF* reporters thus represents the overall inhibitory activity of the uORF, with remaining signal corresponding to the sum of leaky scanning- and re-initiation-mediated reporter translation. By further subtracting the *overlapping uORF* reporter signal that derives from leaky scanning, re-initiation is quantified. Using WT and *Denr* KO HeLa cells (**Figure 1J**), this assay indicated strong inhibition by the *Klhdc8a* uORF, and remaining signal that was largely DENR-dependent, as well as very low leaky scanning rates (**Figure 1I**). This data allowed us to propose a model of how ribosomal fluxes occur on the *Klhdc8a* reporter transcript *in vivo* (**Figure 1K**). Our analyses suggest that ~98% of ribosomes initiate on the uORF and only ~2% undergo leaky scanning. Of the uORF ribosomes about 19% re-initiate on the main CDS, a process that is largely DENR-dependent.

We concluded that translation of endogenous and reporter-encoded KLHDC8A was highly re-initiation-dependent and that our system based on three reporters, together with the use of WT vs. *Denr* KO cells, allowed the quantification of DENR-dependent re-initiation *in vivo*.

An *in vitro* re-initiation assay recapitulates regulation and ribosome fluxes observed *in vivo*.

Next, we set out to establish an *in vitro* assay that could recapitulate re-initiation. To this end, we prepared translation-competent extracts from HeLa cells, essentially adapting a recently reported protocol based on detergent-free cell lysis applying reproducible shearing forces via a dual centrifugation step [22] (**Figure 2A**). Under the gentle lysis conditions used, a sizeable proportion of cells remained intact (**Figure 2B**, pellet fraction), yet we found the supernatant to reproducibly recover highly translation-competent lysate that also contained the three proteins of interest for our study (MCTS1, DENR, eIF2D) (**Figure 2B**). Initial tests revealed that during the *in vitro* translation reaction, eIF2α became phosphorylated at Ser51 (**Figure 2C**), a modification that is associated with decreased initiation rates leading to low translational output. As previously reported [22, 23], the addition of recombinant phosphatase GADD34Δ1-240 greatly counteracted Ser51 phosphorylation (**Figure 2C**). Indeed, in the absence of GADD34Δ1-240, luciferase signals that were obtained in *in vitro* translation assays for various *in vitro* transcribed, capped and polyadenylated reporter RNAs were >50% reduced as compared to reporter signals in the presence of recombinant GADD34Δ1-240 (**Figure 2D**). As default condition in our experiments, we hence added recombinant GADD34Δ1-240 to all reactions. Finally, we further optimised the *Klhdc8a*-based reporters by shortening the 5' UTR from its original ~650 nt down to ~180 nt, retaining the larger uORF environment yet removing far upstream 5' UTR sequence (**Supplementary Figure S2A**). These short *Klhdc8a* 5' UTR reporters showed strongly increased signals in our assays (**Supplementary Figure S2B**), in line with the notion that very long 5’ UTRs are generally inhibitory to translation.

First, we carried out the *in vitro* translation assays on the set of *Klhdc8a* reporters (no uORF, WT uORF, overlapping uORF), using HeLa extracts prepared from both wild-type and *Denr* KO cells. These experiments revealed clear uORF- and DENR-dependence (**Figure 2E**, left side; identical outcomes for “long *Klhdc8a* 5' UTR” reporter, see **Supplementary Figure S2C**), closely resembling our observations *in vivo* (**Figure 1I**). The calculated ribosomal fluxes showed excellent correspondence as well (**Figure 2F**). In particular, the uORF strongly impeded CDS expression with

very low rate of leaky scanning (1%). Re-initiation-mediated reporter expression (11%) was quantitatively slightly lower than *in vivo* (19%, see **Figure 1K**), yet remained highly DENR-dependent (i.e., in *Denr* KO extracts the signal from the 'WT uORF' reporter was close to the leaky scanning signal from the 'overlapping uORF' reporter). Importantly, this lack of re-initiation in *Denr* KO extracts could be fully rescued through complementation with recombinant MCTS1-DENR (**Figure 2E**, right), confirming the direct involvement of these proteins in re-initiation *in vitro*. Note also that in *Denr* KO, MCTS1 is co-depleted (**Figure 1J**) as a result of its instability in the absence of its binding partner (as previously observed in [19, 24]), making a rescue with both proteins necessary. We concluded that our *in vitro* assays showed very good correspondence with re-initiation parameters observed *in vivo*.

Next, we assayed the start-stop uORF-containing *Asb8* reporter set (**Figure 2G**). These experiments revealed that this uORF was rather permissive to leaky scanning, likely due to poor Kozak context of its start codon; in particular, given that all stop codons (TGA, TAA, TAG) begin with a T nucleotide, start-stop uORFs by default have a (suboptimal) T in Kozak sequence +4 position. The lower efficiency of inhibition by the *Asb8* uORF (as compared to the *Klhdc8a* uORF) was in line with the translation efficiencies measured *in vivo* by Ribo-seq (**Figure 1G**). According to our calculations (**Figure 2H**), 58% of ribosomes underwent leaky scanning and 42% engaged in uORF translation. About two-thirds of uORF termination events led to re-initiation at the CDS and the remaining one-third did not (29% vs. 13%). Re-initiation was mostly MCTS1-DENR-dependent, as judged from the signals obtained in *Denr* KO extracts and from recombinant MCTS1-DENR rescue (**Figure 2G**).

MCTS1-DENR selectivity for certain uORFs has been linked to specific penultimate codons [19]. The *Klhdc8a* uORF terminates with the sequence '… GTG(Val) TAG(Stop)', yet penultimate GTG was not observed as specifically associated with DENR-dependence previously [19]. In order to test penultimate codon influence in our assay, we tested reporters in which the GTG(Val) was mutated to GTC(Val) (among the reported significantly depleted penultimate codons in DENR-dependent uORFs [19]) and to GCG(Ala) (most strongly and significantly enriched in DENR-dependent uORFs

according to [19]) (**Figure 2I**). We found that the inhibitory impact of the uORF and re-initiation rates were only weakly affected by these codon changes; in particular, the enriched penultimate codon GCG did not cause higher re-initiation rates – if anything, re-initiation was lower (non-significant) than in wild-type codon constructs and it still remained strongly DENR-dependent. Interestingly, in the depleted penultimate codon case, GTC(Val), DENR-dependence of re-initiation was reduced, yet overall re-initiation capacity still intact. This observation is compatible with the idea that the depleted penultimate codon is able to partially decouple the re-initiation process from its dependence on these particular re-initiation factors.

Next, we examined *Asb8* reporter variants, converting the 1aa uORF to a 2aa uORF (and thus from start-stop to elongating uORF) and testing different penultimate codons. Thus, we inserted either a TGT(Cys), which is (according to [19]) the only significantly enriched penultimate codon starting with 'T' (hence preserving the +4 Kozak context of the start codon), or GCG(Ala), which is a highly enriched codon and additionally improves the Kozak context for uORF initiation (**Figure 2J**). In the *in vitro* translation assays, GCG insertion indeed led to strongly decreased leaky scanning rates, converting the weakly translated 1aa uORF to a strongly translated 2aa uORF (**Figure 2J**, right), compatible with the improved Kozak consensus associated with the +4 position 'G'. Lengthening the uORF with TGT(Cys) had little consequences on leaky scanning and overall re-initiation rate, yet DENR-dependence appeared reduced (**Figure 2J**, middle). We concluded that re-initiation rates were not *per se* different between start-stop and elongating uORFs, but the DENR-dependence of re-initiation was modulated, which is in agreement with the reported preferential MCTS1-DENR regulation of start-stop uORFs [11, 19].

Phospho-eIF2*α* levels only mildly affect re-initiation rates of *Klhdc8a* and *Asb8* reporters.

Phosphorylation of eIF2α is key to the integrated stress response (ISR) that lowers bulk translation initiation and promotes preferential translation of stress-related mRNAs through a mechanism that involves uORFs, re-initiation and start codon selection [25]. The current model deduced from studying transcripts such as *Atf4/Gcn4* suggests that after uORF termination, the re-scanning 40S subunit

becomes rapidly re-initiation-competent when non-phosphorylated (active) eIF2α is abundant. This leads to translation of a relatively close (87 nt) downstream, second uORF rather than the CDS that is more distant (184 nt). However, when p-eIF2α levels are high, the probability of recruiting active eIF2α in time for the second uORF is lower, and re-initiation will more often occur further downstream, on the CDS. Recent work has highlighted the importance of both MCTS1-DENR and eIF2D for *Atf4* CDS translation [19, 20]. Despite these links, whether eIF2α activity is a rate-limiting factor for MCTS1-DENR-mediated re-initiation *per se*, and to what extent the reported effects are a consequence of global initiation decrease, remains unclear. We reasoned that our well-controlled *in vitro* system would allow us to address such questions.

We thus carried out our assays in the presence vs. absence of GADD34Δ1-240, which led to low and high Ser51 phosphorylation levels of eIF2α, respectively (**Figure 2C**). The level of Ser51 phosphorylation occurring in the absence of GADD34Δ1-240 was even higher than that seen after a standard ISR induction protocol by tunicamycin treatment of cells (**Supplementary Figure S3A**). For the *Klhdc8a* reporters, absolute signals were significantly reduced in the absence of GADD34Δ1-240 as expected (analogous to experiments in **Figure 2D**), yet after normalisation to the (likewise reduced) Renilla luciferase signal, the observed levels of leaky scanning (slightly up) and uORF-mediated inhibition (slightly down), the re-initiation rate (slightly down) and its DENR-dependence (up), and the calculated ribosomal fluxes were highly similar between both conditions (**Figure 3A, B**). We concluded that in the *Klhdc8a* reporter *in vitro* experiments, the availability of non-phosphorylated eIF2α was not *per se* a strongly rate-limiting step for re-initiation. On the *Asb8* reporter – where the intercistronic distance between uORF and CDS is very short (17 nt, as compared to 73 nt in *Klhdc8a*) – we found overall more variability in the experimental outcomes in the absence of GADD34Δ1-240 (**Figure 3C, D**). We observed a tendency to increased leaky scanning and reduced re-initiation (both non-significant). In an *in vivo* experiment on the *Klhdc8a* reporter set, combined with tunicamycin treatment, we also observed a weak reduction in re-initiation rate (**Supplementary Figure S3B, C**) and validated that under these conditions the translation of positive control reporters carrying the 5'

UTRs of *Atf4* and *Atf5* upstream of luciferase were indeed upregulated (**Supplementary Figure S3D**). Taken together, these results suggest that eIF2α levels are only moderately limiting for re-initiation.

Limited evidence for a role of eIF2D in re-initiation *in vivo* and *in vitro*, pointing to uORF-independent cellular functions in gene expression regulation.

Next, we extended our analyses to the close MCTS1-DENR homologue, eIF2D, which contains the same domains encoded as a single polypeptide, linked by an additional WH domain (**Figure 4A**). The high similarity between the proteins, in addition to the findings that the yeast orthologues show shared activities, has led to the assumption that the homology in mammals extends to their function. To test this hypothesis, we downregulated eIF2D using specific shRNAs (**Figure 1A**) and carried out Ribo-seq and RNA-seq identically to the experiment in *Denr*-depleted cells shown in **Figure 1B**. For a first comparison between the two genotypes, i.e. *Eif2d*-depleted and *Denr*-depleted, we evaluated in a transcriptome-wide fashion the relative ribosome distribution between mRNA 5' UTR and CDS sequences. This measure is expected to increase when a re-initiation factor is depleted because ribosomes are selectively lost from CDS sequences but will still translate uORFs when re-initiation in inhibited [16]. Indeed, the distribution of the changes seen for 5' UTR-to-CDS footprint ratios between *Denr* shRNA-treated cells vs. control cells was strongly shifted to the positive (**Figure 4B**), indicating globally detectable ribosome redistribution in the absence of MCTS1-DENR, consistent with previous findings [16]. By contrast, the shift was considerably smaller in the cells that had been treated with *Eif2d* shRNAs (**Figure 4B**). Next, we carried out an analysis identical to that applied to *Denr* knockdown cells in **Figure 1B**, to determine the effects induced by *Eif2d* depletion at the level of CDS translation efficiencies (**Figure 4C**). This analysis, too, revealed strikingly different transcriptome-wide effects between the homologues. We thus detected much fewer transcripts with a decreased TE upon *Eif2d* knockdown (n=78, with FDR-corrected p<0.1; Supplementary Table S2) (**Figure 4C**) as compared to *Denr* knockdown cells (n=223) (**Figure 1C**), and the overlap between the two sets was minimal (**Supplementary Figure S4A, B**). Moreover, the mRNAs with low TE in *Eif2d*-depleted cells tended to have short 5' UTRs (**Supplementary Figure S4C**) and they were not enriched for translated

uORFs (**Figure 4D**). We also noted that many of the translationally affected transcripts were highly abundant, including >25 mRNAs encoding for ribosomal proteins and general translation machinery (**Supplementary Figure S4D, E**). Particularly strong consequences of eIF2D loss-of-function were evident for RNA abundances, with a large number of mRNAs significantly up (n = 460) and downregulated (n = 617) (**Figure 4C**) and, here, abundant transcripts were overrepresented as well (**Supplementary Figure S4F**) and associated with cellular processes such as the cell cycle and translation (**Supplementary Figure S4G, H**). Alterations to the cell cycle could indeed be further validated as a robust phenotype following eIF2D loss-of-function (**Supplementary Figure S4I**). Taken together, these findings suggested that, first, the knockdown of *Eif2d* was phenotypically highly consequential as judged by the widespread gene expression reprogramming that it induced. However, second, it would seem that the primary activity of eIF2D may not lie in regulating the re-initiation of uORF-containing transcripts.

To further investigate the above observations on potential eIF2D targets, we selected several candidate transcripts to test if their TE regulation was 5' UTR-mediated. Of these, *Cenpa*, *Ndc80*, *Med23* and *Rps20* were picked as candidate mRNAs with low TE from **Figure 4C**. We also included *Hoxa3*, which has been identified as a *bona fide* eIF2D-regulated mRNA in a previous study [26]. According to the published findings, eIF2D is implicated in uORF-mediated inhibition, rather than in re-initiation, of *Hoxa3* CDS translation. Indeed, our *in vivo* data confirmed *Hoxa3* CDS TE up-regulation upon *Eif2d* depletion (**Figure 4E**; note that *Hoxa3* is very lowly expressed in our cells and did not feature among the significantly changed TE transcripts in **Figure 4C**, but the differences were statistically significant in the direct comparison of TEs shown in **Figure 4E**). *In vivo* luciferase reporter assays using the 5' UTRs of the selected endogenous candidate transcripts did not reveal significant regulation by eIF2D (**Figure 4F**). For the *Hoxa3* 5' UTR, downregulation of reporter output in the absence of eIF2D was observable, in principle in support of eIF2D re-initiation activity. We also analysed *Klhdc8*a (endogenous and 5' UTR reporter) and *Asb8* (endogenous), but we found no

indication that eIF2D was involved in re-initiation for these transcripts *in vivo* (**Figure 4E, F; Supplementary Figure S4**).

We next tested the outcome of eIF2D loss-of-function and rescue in our *in vitro* assay. In lysates from *Eif2d* KO HeLa cells (**Figure 1J**), no difference in re-initiation activity or leaky scanning was detectable for the *Asb8* reporter (**Figure 4G**). By contrast, for the *Klhdc8a* reporter, re-initiation rates in *Eif2d* KO lysates were reduced by about half (**Figure 4H**). However, neither recombinant eIF2D nor recombinant MCTS1-DENR were able to fully rescue the low signal obtained with the WT uORF reporter in *Eif2d* KO lysates (**Figure 4H**). These observations were most compatible with the model that lower re-initiation signal was an indirect effect of eIF2D loss, e.g. due to secondary effects related to the widespread gene expression reprogramming seen in **Figure 4C**. Thus, we would not expect indirect effects to be rescuable *in vitro*, whereas a direct role in the re-initiation process itself should be rescued by adding the needed re-initiation factor as a recombinant protein.

In *Eif2d* knockout extracts, MCTS1-DENR is abundant and may mask the re-initiation function of eIF2D. To address this possibility, we carried out two experiments. First, using re-initiation-deficient *Denr* KO extracts and recombinant eIF2D, we were indeed able to partially rescue re-initiation activity (**Figure 4I**), although not as efficiently as with recombinant MCTS1-DENR (compare with **Figure 2E**). Second, we used *Denr* + *Eif2d* double knockout cell lysates. *In vitro* translation reactions in these extracts showed reduced re-initiation activity (**Figure 4J**, left) that was very similar to *Denr* single knockouts (**Figure 2E**). Addition of recombinant MCTS1-DENR showed the expected rescue (**Figure 4J**, right) and, interestingly, also recombinant eIF2D partially rescued under in these conditions, i.e. in the absence of MCTS1-DENR (**Figure 4J**, middle). Taken together, these results indicate that eIF2D is likely not an essential component of the re-initiation machinery *in vivo* (**Figure 4C**) and *in vitro* (**Figure 4H**). However, when MCTS1-DENR is absent, adding high concentrations of eIF2D (0.5 µM) can partially compensate for MCTS1-DENR loss. Thus, it is a probable scenario that eIF2D has in principle retained some shared activity with MCTS1-DENR, but its main cellular functions seem to lie outside of uORF re-initiation.

Finally, it is also possible that eIF2D is active in re-initiation in a conditional fashion; in particular upon induction of the integrated stress response that leads to unavailability of eIF2α, eIF2D may be able to deliver initiator-tRNA to the 40S ribosome. We addressed this possibility by carrying out *in vitro* translation in *Eif2d* knockout extracts (without/with recombinant eIF2D rescue) in the absence of GADD34Δ1-240 (**Figure 4K**). Also under these conditions, we observed the (most likely) indirect/secondary effects of eIF2D loss that could not be rescued by recombinant eIF2D, and the experiment revealed no further indications for a specific re-initiation role of eIF2D.

Identification of MCTS2 as a functional DENR heterodimerization partner.

It has been observed that MCTS1 and DENR are mutually dependent on each other for protein stability [19, 24], giving rise to the idea that they act as obligate heterodimers. However, closer inspection of the available cell culture data would suggest that the depletion of MCTS1 in *Denr* KO cells is stronger than the depletion of DENR in *Mcts1* KOs [19]; moreover, *in vivo* phenotypes in mouse neuronal migration are stronger upon DENR loss than MCTS1 loss [27]. These findings could indicate that DENR is indeed an obligatory component of the heterodimer, but that it can interact with and be stabilised by alternative heterodimerization partners as well. To obtain insights into this hypothesis, we sought to purify DENR-interacting proteins by a co-immunoprecipitation and mass-spectrometry approach. Using CRISPR gene editing, we modified the endogenous *Denr* locus in NIH/3T3 cells such that it expressed DENR protein with a C-terminal tag suitable for immunoprecipitation via a 3xFLAG sequence as well as for degron-mediated protein depletion via the dTAG / dTAG-13 degrader system [28] (**Figure 5A**). Genotyping indicated successful modification of both alleles in selected cell clones (**Figure 5B**) and immunoblot analysis validated the functionality of the degron system (**Figure 5C**). We then analysed by mass-spectrometry the proteins enriched in FLAG-IP performed on extracts from control-treated vs. dTAG-13-treated cells (**Figure 5D**). The two most strongly enriched co-purifying proteins were MCTS1 and, moreover, MCTS2. This finding indicates that DENR assembles into two distinct heterodimeric complexes, MCTS1-DENR and MCTS2-DENR. *Mcts2* is a paralog and retrogene copy of *Mcts1* that has been noted for its

transcription from an imprinted gene region in mouse and humans, with predominant expression from the paternal allele [29]. A recent study on human *Mcts1* deficiency and associated disease phenotypes has speculated on redundancy with *Mcts2*, yet based on reporter assays and other analyses, it came to the conclusion that this paralog was largely non-functional [30]. By contrast, our interaction data collected from endogenous proteins suggests that both MCTS paralogs form *bona fide* heterodimers with DENR. The *Mcts1* and *Mcts2* coding sequences differ at 81 individual nucleotide positions that are distributed quite evenly across the overall CDS of 546 nt (**Supplementary Figure S5A**), allowing to unambiguously assign RNA-seq reads and Ribo-seq footprints to the two paralogous sequences. We could thus compare the expression of the *Mcts* paralogs, which indicated higher levels of *Mcts2* than *Mcts1* in our cells (**Figure 5E**). These findings strongly indicate that MCTS2 is a relevant interaction partner of DENR besides MCTS1.

Human and mouse MCTS1 protein sequences are identical apart from one leucine to isoleucine substitution at amino acid position 30 (**Figure 5F**). In MCTS2, divergence to MCTS1 can be found at 14 positions when aligning mouse, rat, human and macaque protein sequences, with 4 substitutions common to MCTS2 across species and other changes specific to the analysed rodents or primates (**Figure 5F**). Using available structural data from *in vitro* reconstituted MCTS1-DENR in complex with the 40S ribosome, initiator-tRNA and a viral model RNA [31], we found that MCTS2-specific amino acid changes affected positions that mapped outside of the interaction surface between MCTS1-DENR and also outside of that between MCTS1-DENR and the 40S ribosomal subunit (**Figure 5G**). Hence, the capacity of MCTS2 vs. MCTS1 to interact with DENR, and of MCTS2-DENR vs. MCTS1-DENR to interact with 40S can be expected to be comparable. However, the substitutions – of which in particular the change at amino acid position 75 from glutamate (negative charge) in MCTS1 to lysine (positive charge) in MCTS2 (**Figure 5F**) would have the potential to significantly alter protein properties – could affect additional, so far unknown interactions taking place at this solvent-exposed surface within the complex (**Figure 5G**). To gain insights into whether MCTS1-DENR and MCTS2-DENR could redundantly act in re-initiation, we tested both as recombinant proteins on

Klhdc8a and *Asb8* reporters. These experiments revealed that both paralogs were able to rescue *Denr* KO extracts in a comparable fashion (**Figure 5H, I**). Taken together, we concluded that *Mcts2*, although previously categorised as a pseudogene [32], is expressed to relevant levels and produces protein that assembles with DENR into protein complexes *in vivo*. *In vitro*, the paralogous complexes show similar activity on our two re-initiation model substrates, indicating redundancy in activity at least for certain substrates.

Discussion

Translation re-initiation is a process that is at odds with several principles of canonical initiation, termination and post-termination subunit recycling and that remains mechanistically particularly poorly understood. The MCTS1-DENR heterodimer is so far the only reported *trans*-acting factor that is specific to re-initiation, i.e. it does not appear to be essential for general translation as well – quite in contrast to several other proteins that have also been linked to re-initiation in mammals, such as components of eIF3 [33-35] or eIF4 [35]. For these general eIFs, the analysis of specific re-initiation effects *in vivo*, for example through loss-of-function experiments, is challenging, as they cannot easily be deconvoluted from the abundant global translational alterations and their secondary effects. A defined *in vitro* assay that faithfully recapitulates and quantifies re-initiation gives the possibility of dissecting this process specifically, allowing for targeted, better interpretable experiments. Therefore, we view the development of the *in vitro* re-initiation assay presented in this study as an important methodological advance that will allow the characterisation of *cis*-acting mRNA/uORF sequence requirements and *trans*-acting protein machinery under defined conditions. Our careful benchmarking of the assay, including the use of different reporter constructs and variants, the demonstration of its dependence on MCTS1-DENR using knockout lysates and rescue with recombinant proteins, and the calculation of ribosomal fluxes, demonstrate the suitability of our system for detailed future mechanistic dissection of re-initiation. For example, a strategy to test the requirement for general eIFs in re-initiation could consist in producing HeLa *in vitro* translation lysates devoid of the protein of interest (e.g. through acute degron-mediated depletion, just before harvest of the HeLa cells), and to

drive uORF translation on the *Klhdc8a* reporter through IRES sequences (e.g. the "factorless" Cricket Paralysis Virus IRES [36]) that bypass the requirement for the particular eIF of interest. Thus, it should be possible to query eIF requirement specifically in re-initiation. Other attractive future applications for our *in vitro* assay include structural biology approaches (i.e., cryo-EM on re-initiation complexes) or compound screens (i.e., testing for specific inhibitors of re-initiation).

In addition to the methodological advancement, our study provides important new biological insights, notably regarding MCTS1-DENR and the paralogous MCTS2, as well as on the role that eIF2D plays in gene expression regulation. In particular our observations on eIF2D are surprising, as the *in vivo* and *in vitro* evidence that we provide indicates that the main functions of this protein may not lie in uORF re-initiation after all. In retrospect, the case of eIF2D carrying re-initiation activity in mammals *in vivo* was always heavily built on circumstantial evidence. This evidence was, mainly, the clear homology to MCTS1-DENR, a validated mammalian re-initiation factor; the *in vitro* ability shared by both factors to act on 40S subunits and deliver Met-tRNA$^{Met}_{i}$ to the empty P-site ([13]; using an experimental model where 40S is in complex with HCV IRES-like mRNA); an *in vitro* role of eIF2D for another viral element-dependent translational shuttling phenomenon that resembles re-initiation [37]; the shared *in vitro* activity to promote the release of deacylated tRNA and mRNA from recycled 40S subunits [13]; the high structural similarities of the proteins in reconstituted complexes together with the 40S ribosome [24, 31, 38]; and, finally, *in vivo* evidence in yeast that the orthologous proteins have shared functions which, however, do not lie in uORF re-initiation but in post-termination tRNA removal and 40S ribosome recycling [18] – events from which re-initiation can be envisaged to have subsequently evolved in higher eukaryotes. Finally, two recent studies have observed that eIF2D loss-of-function affected the expression of ATF4 in the integrated stress response in *Drosophila* [20] and of reporter constructs that carried the *Atf4* 5' UTR (or several other 5' UTRs containing uORFs) in HeLa cells [19]. These effects were exacerbated by simultaneous DENR loss-of-function, which was interpreted as partially overlapping re-initiation activities of eIF2D and MCTS1-DENR. However, whether the eIF2D-dependent effects indeed relied on the mRNAs' uORFs, or whether they could be

mechanistically distinct, was not specifically addressed. Notably, evidence has been accumulating that eIF2D plays roles separate from MCTS1-DENR [21, 26, 39, 40], and the data presented in our study supports this idea as well.

Our work, based on the transcriptome-wide analysis of how RNA abundances and translation are altered upon eIF2D depletion, provides evidence that eIF2D is an important regulator of cellular gene expression, as judged from the massive reprogramming at the transcriptome level with hundreds of mRNAs up- and down-regulated and a cell cycle phenotype. However, minor changes occur at the translational level and, moreover, the affected mRNAs are not enriched for uORFs, which points to a main cellular role that does not involve re-initiation and therefore does not overlap with that of MCTS1-DENR. The specific analysis of *Atf4* mRNA in our datasets also supports this model, as we observe increased footprint abundance on the *Atf4* uORF1 stop codon in *Denr* shRNA-treated cells (in line with ribosomal stalling when deacylated tRNA removal is impaired [19]), but not so in *Eif2d* shRNA-treated cells whose profiles are identical to control cells (**Supplementary Figure S6**). Still, since our *in vitro* experiments also reveal that high concentrations of eIF2D can partially compensate for MCTS1-DENR loss, it is quite likely that eIF2D has retained some capacity to act as a re-initiation factor. Whether this activity comes to bear under certain conditions or ins specific cell types *in vivo* will require careful further investigations. Probably the most revealing experiment to elucidate the biological function and mechanism by which eIF2D acts, and to distinguish direct from indirect targets in our profiling analyses, would be to identify the mRNAs and sequence elements that eIF2D physically interacts with *in vivo*, for example through selective 40S footprinting [35]. Our efforts in this direction have, unfortunately, not yet been successful.

That MCTS2 can act as a *bona fide*, functionally active DENR interaction partner that is expressed to significant levels in the NIH/3T3 cells used for our study, is an important finding as it potentially explains some of the enigmatic published discrepancies between *Mcts1* vs. *Denr* knockout phenotypes. Thus, neuronal migration has been reported to be more strongly perturbed upon DENR loss-of-function than MCTS1 loss-of-function [27]. Moreover, according to the International Mouse

Phenotyping Consortium (IMPC; https://www.mousephenotype.org/), *Denr* knockout leads to preweaning lethality with complete penetrance, yet *Mcts1* knockouts are viable, with relatively mild reported phenotypes (mainly affecting eye morphology in early adult animals). The consortium did not phenotype knockouts of *Mcts2*, likely due to its pseudogene assignment. The high *Mcts2* RNA and footprint expression observed in our study could be a particularity of NIH/3T3 cells. Of note, our pilot experiments in mouse embryonic stem cells (mESCs) have recovered MCTS2 as an *in vivo* interactor of DENR as well (data not shown), and re-analysis of our previous Ribo-seq data from mouse liver [41] and kidney [42] indicates that MCTS2 is co-expressed alongside MCTS1 in these organs *in vivo*, with relative proportions of MCTS1 vs. MCTS2 biosynthesis different across these organs (**Supplementary Figure S7**). Therefore, we propose that MCTS2 is a broadly expressed paralog. In the future, it will be particularly interesting to explore whether MCTS1-DENR and MCTS2-DENR heterodimers have redundant functions or whether the identity of the MCTS paralog has functional consequences for the activity of the heterodimer. Finally, while Bohlen et al. recently reported that *Mcts2* expression is undetectable in various cells of the myeloid lineage (human whole blood cells, THP-1 cells, T-cell blasts) [30], our findings from mouse warrant a more thorough analysis across human tissue and cell types.

Material and methods

Cloning

To generate *in vivo* dual luciferase reporter plasmids, we amplified the 5' UTRs and first 5 or 10 (depending on construct) CDS codons (in frame with firefly luciferase CDS) of the selected transcripts by PCR from mouse cDNA or genomic DNA and cloned them into the BamHI restriction site of pLV1 dual luciferase reporter plasmid with mutated firefly CDS start, essentially as previously described [16]. Sequences of the primers used for the PCR are reported in Table 1. All constructs were validated by Sanger sequencing.

In vitro translation reporters were generated by PCR amplification of *Asb8* or *Klhdc8a* 5' UTRs together with the 6 first CDS codons (in frame with the firefly luciferase CDS, lacking additional ATG

start codon), from mouse cDNA, genomic DNA, or the corresponding pLV1 *in vivo* reporter constructs. The PCR products were then cloned into the BamHI restriction site of luciferase T7 control DNA plasmid (Cat. No. L4821, Promega). Modifications of uORFs or other mutations were performed using the Q5 site-directed mutagenesis kit (Cat. No. E0554S, New England Biolabs) and validated by Sanger sequencing. Sequences of primers for PCRs and mutagenesis are reported in Table 2.

Plasmids for MCTS1, DENR and eIF2D protein expression have been reported [31] and cloning of human *Mcts2* cDNA for recombinant expression was analogous to that of *Mcts1*; briefly, a fragment encoding the first 154 amino acids of MCTS2 (i.e., the part carrying substitutions with regard to MCTS1) was PCR-amplified from HeLa genomic DNA with primers reported in Table 3 and cloned into the XhoI and EcoRV restriction sites of the pLIC C-terminally 6x-His-TEV-tagged MCTS1-expression vector [31]. Correct sequence was validated by Sanger sequencing.

For endogenous tagging of DENR, a gRNA (GTCATCCACCTAGAAGACACAGG) targeting a downstream sequence of the *Denr* stop codon was inserted into vectors pCas9-2A-mCherry [43] and pRRE200 [44] using oligos described in Table 4. 5µM of oligos were annealed in annealing buffer (100µM Tris-HCl pH7.5, 500mM NaCl and 10mM EDTA) by cooling from 98°C to 20°C with a ramp of 0.02°C per second. Annealed oligos were diluted 10 times then ligated into pCas9-2A-mCherry BsaI restriction site or pRRE200 AatII and SacI restriction sites. Colony PCRs using the reverse gRNA oligo and a pCas9-2A-mCherry or pRRE200 forward primer was carried out to check for correct insertion of the annealed oligos and the cloned plasmids validated by Sanger sequencing.

A pBluescript II SK(-) vector was ordered as a synthetic clone to contain the knock-in cassette 3xFLAG-NeonGreen-dTAG. *Denr* 5' and 3' homology arms were amplified from mouse genomic DNA by PCR using primers reported in Table 4 and ligated in the pBluescript II SK(-) first digested with SpeI and BamHI for the 5' homology arm then with NdeI and KpnI for the 3’ homology arm. Positive clones were identified by colony PCR using PCR homology arm primers and by Sanger sequencing.

Table 1. Sequences of oligonucleotide primers used for the *in vivo* reporters.

Transcripts	Primer sequences for cloning in prLV1
Klhdc8a	Forward: aaaggatccGTCTCCGACCCTGTAGACACTGCAG
	Reverse: tttggatccactagtATTGGGCACTTCCATGGCAGCCCGG
Klhdc8a start mut	Forward: GGCAGGGAGCCCGTGCAGCGCGgtaccGAGGCTGAGAGAGGGGACGCGCC
	Reverse: GGCGCGTCCCCTCTCTCAGCCTCggtacCGCGCTGCACGGGCTCCCTGCC
Klhdc8a stop mut	Forward: CGGCGCCGTGcacCCCGCGGGCC
	Reverse: GTCGGCGCGTCCCCTCTCTC
Med23	Forward: aaaggatccAGAGCAAGAGAGAGCGGCG
	Reverse: tttggatccactagtGCTCTGCAGTTGCGTCTCCATC
Rps20	Forward: aaaggatccGTTCCGGGTCACAAAGCACC
	Reverse: tttggatccactagtGGGCGTCTTTCCGGTATCTTTAAATG
Hoxa3	Forward: aaaggatccAGGACAATTCGTCTCTTGGGC
	Reverse: tttggatccactagtTGAGCTGTCGTAGTAGGTCGC
Cenpa	Forward: aaaggatccATTGGCTTCAGACCTTTATTCTCATTGG
	Reverse: tttggatccactagtTGGGGTCTGCGGTTTGCGAC
Ndc80	Forward: aaaggatccTGACGTCAGCGCGGGCG
	Reverse: tttggatccactagtACCACAGGTGGAAACTGAACTGCG

Table 2. Sequences of oligonucleotide primers used for the *in vitro* reporters.

Transcripts	Primer sequences for cloning in pT7luc
Klhdc8a start mut	Forward: ggtGAAGACGCCAAAAACATAAAG
	Reverse: ttgaATCCTGACATTGGGCACT
Klhdc8a stop mut	Forward: CGGCGCCGTGcacCCCGCGGGCC

	Reverse: GTCGGCGCGTCCCCTCTCTC
Klhdc8a deletion 499bp 5' UTR	Forward: GGTGTCAGCGGGGCAGCAAGCTTGGGAGGCCTTA
	Reverse: TAAGGCCTCCCAAGCTTGCTGCCCCGCTGACACC
Klhdc8a penultimate codon GTC (WT & start mut)	Forward: GACCGGCGCCGTCTAGCCCGCGG
	Reverse: GGCGCG TCCCCTCTCTCAGC
Klhdc8a penultimate codon GTC (stop mut)	Forward: GACCGGCGCCGTCCACCCCGCGG
	Reverse: GGCGCGTCCCCTCTCTCAGCCTCC
Klhdc8a penultimate codon GCG (WT & start mut)	Forward: GACCGGCGCCGCGTAGCCCGCGG
	Reverse: GGCGCGTCCCCTCTCTCAGCC
Klhdc8a penultimate codon GCG (stop mut)	Forward: GACCGGCGCCGCGCACCCCGCGG
	Reverse: GGCGCGTCCCCTCTCTCAGCCTCC
Asb8	Forward: aaaggatccTTCTGCTTCCGGGTCACGCC
	Reverse: tttggatcctCCACATGCTGGAACTCATCAAGG
Asb8 start mut	Forward: CGTTTGGAGCaccTGAACACACTCTGAGCCTTGATGAG
	Reverse: AACCGCGGGTCAGATCGG

Table 3. Sequences of oligonucleotide primers used for cloning the pLIC C-terminally 6x-His-TEV-tagged *Mcts2* vector.

Transcripts	Primer sequences for cloning in pLIC
hsMcts2	Forward: AAACTCGAGTTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATAGATCATGTTCAA GAAGTTTGATGAAAAGG
	Reverse: AAAGATATCTTCTGCAGACATCTTCATGAC

Table 4. Sequences of oligonucleotides used for the cloning of *Denr* gRNA and homology arms into the pCas9-2A-mCherry, pRRE200 and pBluescript II SK(-).

Plasmids	Primers type	Oligo sequences
pCas9-2A-mCherry	Annealing oligos	Forward: caccgGTCATCCACCTAGAAGACAC
		Reverse: aaacGTGTCTTCTAGGTGGATGACc
	Colony PCR	Forward: GCCTATTTCCCATGATTCCTTCA
pRRE200	annealing oligos	Forward: GTCGTCATCCACCTAGAAGACACAGGTGAACGT
		Reverse: aaacGTGTCTTCTAGGTGGATGACc
	Colony PCR	Forward: CCACTTTGCCTTTCTCTCCACAG
pBluescript II SK(-)	To PCR 5' homology arm	Forward: aaaactagtATTTAAAATGTAGATGAAATTTCAAAATGAAG
		Reverse: aaaggatccCTTCTTCACTTCTCCAAGG
	To PCR 3' homology arm	Forward: aaaattaatTGATCCTGAAGTCATGTGTTTCTAACTG
		Reverse: aaaggtaccTGCACGAGGCCCTGGGTTC
	Q5 mutagenesis	Forward: GTTCTTTCACgTGTGTCTTCTAG
		Reverse: TTCATGAATTGTGGCCAAAAC

Cell culture and lentiviral transduction

NIH/3T3, HEK293FT and HeLa cells were cultured under standard conditions (DMEM; 10% FCS, 1% penicillin/streptomycin, all from Invitrogen; 37°C; 5% CO_2). The HeLa *Denr* knockout (KO) cells have been published (Reference); *Eif2d* KO and *Denr* / *Eif2d* double KO cells were kindly provided by Aurelio Teleman, DKFZ Heidelberg.

Lentiviral transduction

Lentiviral shRNA plasmids targeting *Eif2d* were purchased from Sigma (Cat. No. TRCN0000119842 for non-functional (nf) shRNA, TRCN0000119844 for shRNA1 and TRCN0000119845 for shRNA2). The *Denr* ("*Denr* shRNA2") and Scramble shRNA plasmids have been described previously [16].

Production of lentiviral particles in HEK293FT cells using envelope pMD2.G and packaging psPAX2 plasmids and viral transduction of NIH/3T3 cells were performed following published protocols [16, 45], with puromycin selection at 5µg/ml for 4 days for shRNA transduced cells.

Western Blot

Total protein extracts from NIH/3T3 and HeLa cells were prepared according to the NUN method [46]. For analysis of phosphorylated proteins, PhosSTOP™ (Cat. No. 4906845001, Roche) was included in the NUN buffer. Antibodies used were rabbit anti-KLHDC8A 1:5000 (ab235419, Abcam), rabbit anti-VINCULIN 1:5000 (ab129002, Abcam) mouse anti-U2AF65 1:5000 (U4758, Sigma), rabbit anti-EIF2D 1:750 (12840-1-AP, Proteintech), rabbit anti-DENR 1:5000 (ab108221, Abcam), rabbit anti-phospho-eIF2alpha (Ser51) 1:1000 (#9721S, Cell Signaling), mouse anti-FLAG 1:10000 (M2F3165, Sigma), rabbit anti-RPL23 (ab112587, Abcam), mouse anti-beta-Tubulin 1:5000 (T5201, Sigma) and secondaries goat anti-guinea pig-HRP 1:10000 (SA00001-12, Proteintech), anti-rabbit-HRP (W4011, Promega) and anti-mouse-HRP 1:10000 (S3721, Promega). The guinea-pig anti-MCTS1 antibodies (used at 1:500) have been described (Schleich et al. 2014). Blots were visualized with the Fusion Fx7 system.

Translation-competent extract preparation

The overall protocol for extract preparation was an adaptation of that published by Gurzeler et al., 2022 [22]. Briefly, confluent 15-cm dishes of HeLa cells were washed once with PBS at 37°C, then trypsinized with 0.05% Trypsin-EDTA (Cat. No. 25300054, Gibco™). Harvested cells were collected by centrifugation at 500 x *g* for 5 min at 4°C. The cell pellet was washed twice with ice-cold PBS and centrifugation at 200 x *g* for 5 min at 4°C. The cell pellet was resuspended in 1:1 lysis buffer (10mM HEPES pH7.3, 10mM potassium acetate, 500µM $MgCl_2$, 5mM DTT and 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail (Cat. No. 4693132001, Roche)). Cells were lysed by dual centrifugation at 500 RPM for 4 min at −5°C using a ZentriMix 380 R centrifuge (Hettich AG) with a 3206 rotor and 3209 adapters. The resulting lysate was then centrifuged in a table-top centrifuge at 13,000 x g at 4°C

for 10min and the supernatant was aliquoted and snap-frozen in liquid nitrogen; storage at -80°C for up to 6 months.

Expression and purification of recombinant proteins

Recombinant EIF2D, MCTS1 and DENR protein purification has been described [31]. Recombinant MCTS2 was expressed and purified identically to MCTS1. GADD34Δ240 expression plasmid has been described [22]; for protein preparation, bacterial strain BL-21 (DE-3) (pLysS) was transformed with the plasmid, grown in 1000 ml Luria broth at 37°C until the OD600 reached 0.4-0.6. After induction (1mM IPTG) bacteria were cultured for another 4h and cells were harvested. Expression of recombinant GADD34 Δ240 was validated on Coomassie gel. The cell pellet was resuspended in 50 ml lysis/wash buffer (50mM Tris HCl pH 7.5, 300mM NaCl, 5% glycerol, 25mM imidazole), lysed by sonication with 240 x 1s pulses at 40% amplitude on ice and centrifuged at 18,000 x g for 20min at 4°C. The supernatant was mixed with 1ml HisPur Ni-NTA washed resin slurry (Cat. No. 88222, Thermo Fisher Scientific) for at least 2h at 4°C under continuous rotation. Unbound proteins and background were removed by washing the resin 2x with 50ml of lysis/wash buffer, then bound proteins were eluted 2x with 3ml elution buffer (lysis/wash buffer supplemented with 400mM imidazole). The proteins were dialyzed in protein reconstitution buffer (30mM NaCl, 5mM HEPES pH 7.3) overnight using a Slide-A-Lyzer Dialysis Cassette (MWCO 10K) (Cat. No. 88400TS, Thermo Fisher Scientific) at 4°C under agitation. Protein concentration was quantified with a nanodrop device.

***In vitro* transcription**

Reporter DNA templates were PCR amplified before transcription for 2h using the MEGAscript® T7 Transcription Kit protocol (AM1334, Invitrogen) (forward primer: GCGCGTTGGCCGATTCATTAATGC; reverse primer: T_{100}GGGAGCTCGCCCCCTCGGAG). Capping was carried out co-transcriptionally using the CleanCap Reagent AG (3' Ome) (N-7413, TriLink) and the Yeast Inorganic Pyrophosphatase (Cat. No. M2403S, New England BioLabs)

following the CleanCap protocol. The capped RNAs were treated for 15min at 37°C with 0.14 U/µl Turbo DNase (AM2238, Invitrogen), then purified using the Monarch RNA Cleanup Kit (T2040L, New England Biolabs). Quality and size of the produced RNAs were assessed using an Agilent Fragment Analyzer.

***In vitro* translation assay**

6.5µl of translation-competent extracts were freshly supplemented with 0.4mM amino acids (L4461, Promega), 15mM HEPES pH 7.3, 6mM Creatine Phosphate, 102ng/ml Creatine kinase, 28mM potassium acetate, 1mM $MgCl_2$, 24mM KCl, 1U/µl RNasin® Plus Ribonuclease Inhibitor (N2611, Promega) and 2fmol/µl of Renilla luciferase reporter in a final volume of 12.5µl prior to each *in vitro* translation assay. Unless specified, 16ng/µl of recombinant GADD34Δ240 was included to the supplemented extracts. 30fm/µl of firefly luciferase reporter were used per reaction. The *in vitro* translation assays were carried out at 37°C for 50min. Luminescence activity was quantified using the Dual-Glo® Luciferase Assay System (E2940, Promega) and a Tecan Safire2™ plate reader.

***In vivo* dual luciferase assay**

NIH/3T3 cells were first stably transduced with the lentiviral dual luciferase reporters containing the 5' UTRs to be assayed (upstream of firefly luciferase) and Renilla luciferase (internal normalisation), bidirectionally driven by the *Pgk1* promoter. After establishing the reporter-expressing cell population and several passages to ensure stable expression, cells were transduced with the shRNA lentiviruses and selected on puromycin for 4 days before lysis in 1x passive lysis buffer (E1941, Promega) and dual luciferase readout using the Dual-Glo® Luciferase Assay System (E2940, Promega) and a Tecan Safire2™ plate reader.

Ribosome profiling and library preparation

One 15 cm dish of confluent NIH/3T3 cells was used per replicate. Ribosome profiling and parallel RNA-seq were performed in triplicate for *Denr* shRNA2, *Eif2d* shRNA1, *Eif2d* shRNA2, non-functional (nf) *Eif2d* shRNA and Scr shRNA transduced cells, closely following our previously reported protocol [16] with minor modifications as described below. Cells were washed in cold PBS

(without cycloheximide), harvested by scraping down in small volume of PBS (without cycloheximide), pelleted by brief centrifugation, then flash frozen. Cell lysates were prepared as described (with cycloheximide) [16]. For the generation of ribosome-protected fragments, cell lysates were treated with 5 units RNaseI (ART-Seq, Epicenter/Illumina) per OD260, and monosomes were purified on MicroSpin S-400 columns (GE Healthcare). For both Ribo- and RNA-seq, 2µg of RNA was used for rRNA depletion following the RiboCop rRNA Depletion Kit V1.2 (Cat. No. 037.24, Lexogen). Finally, libraries were amplified using 10-14 PCR cycles, and sequenced (HiSeq2500 platform).

Sequencing pre-processing, alignment, quantification

Sequencing reads were processed as described previously [47]. Briefly, sequencing adapters were trimmed using Cutadapt [48] and the reads were size filtered by a custom Python script (26-35 nt for RPF and 21-60nt for RNA). The reads were then aligned subsequently to mouse rRNA, human rRNA, mouse tRNA and mouse cDNA from Ensembl mouse database release 100 using Bowtie2 [49]. In order to estimate expressed isoforms, the RNA-seq reads were mapped in parallel to the mouse genome using STARmapper version 2.5.3a [50] and processed with StringTie [51] to measure FPKM for each transcript. A custom Python script classified the transcripts into single expressed isoform or multiple transcript isoforms. The mapped reads were counted on 5' UTR or CDS of the protein coding genes using a custom Python script. The location of the putative A-site of RPFs were estimated as 5' position +16 for reads shorter than or equal to 31 nt and +17 for reads longer than 31 nt. Transcripts for which there were less than 10 counts in total of all samples were filtered out of the further analysis. Read counts were normalized using the upper quartile normalization method [52] from edgeR [52] then transformed to transcripts per million (TPM) to normalize for transcript length. Translation efficiencies were calculated as the $\log_2$-transformed ratio of normalized CDS footprint counts to normalized CDS RNA counts and averaged over replicates. All analyses were carried out on R version 4.2.1.

Differential translation efficiency analysis

Ribosome profiling sequencing data were analysed using the deltaTE method [53] in order to assess significant changes in translation efficiency and RNA abundance. Significance threshold was set to FDR < 0.1.

uORF annotation and analyses

For uORF annotation, transcripts were used for which reads could be unambiguously assigned to the 5' UTR, i.e. transcripts that are the only protein-coding isoform expressed (single expressed isoform) (N = 6841). In order to include as many transcripts as possible for the annotations, genes with multiple protein coding isoforms were also considered if all expressed isoforms had the same CDS start, in which case the transcript with the longest 5' UTR was selected (N = 2425). For the selected transcripts, uORFs were annotated and considered as translated with the following criteria: (i) they started with AUG, CUG, GUG or UUG, (ii) they had an in-frame stop codon within the 5' UTR, and (iii) they had a coverage of at least 33%. When several potential uORFs were overlapping, the one with the highest coverage (read count/uORF length) was considered.

Cell cycle monitoring

The cell proliferation assay was performed using the Click-iT® EdU Flow Cytometry Assay Kits (C10424, Invitrogen) following the vendor's protocol with some modifications. Briefly, NIH/3T3 cells transduced with scramble shRNA, nf *Eif2d* shRNA, *Denr* shRNA and *Eif2d* shRNA1 and shRNA2 were seeded into 6-cm plates under puromycin selection. After 4 days, the cells were treated for 1h with 10µM of EdU, washed with PBS, trypsinized and collected into FACS tubes. The cells were then centrifuged at 1200 x *g* for 5 min and resuspended in PBS. The cells were counted, and all conditions were adjusted to the same number of cells. After washing with 3ml of 1% BSA in PBS, the cells were resuspended in 100µl fixative buffer and incubated overnight at 4°C. Another wash with 1% BSA in PBS followed by a 15 min incubation in 100µl of 1x saponin-based permeabilization and wash reagent. 200µl of Click-iT reaction cocktail were added to each tube and the cells were incubated for

30min at room temperature. After a wash in 3ml of 1% BSA in PBS, the cells were resuspended in 100µl of 1x saponin-based permeabilization and wash reagent complemented with 5µl of 20mg/ml RNaseA and 2µl of 1mg/ml propidium iodide and incubated for30 min at room temperature. Fluorescence was measured using Accuri C6 FACS machine.

MCTS1-DENR structural models

The structure of the human ribosome in complex with MCTS1-DENR (pdb entry 5vyc) can be downloaded as 6 bundles each containing a hetero 36-mer complex. We started to load bundle 2 in Swiss-PdbViewer and kept only chain A (18S ribosomal RNA). We then loaded bundle 1 in a new layer and used the function “select residues close to another layer” to identify and retain only chains with at least one residue within 10A of the 18S ribosomal RNA (chains H, M, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, 0, and 1). We repeated the operation with bundle 6, for which we retained only chains Q, T and W. We then loaded the crystal structure of the human 40S-eIF2D-re-initiation complex (pdb entry 5oa3) and selected the first 218 residues of the ribosomal protein S2 (chain C of 5oa3 and chain s of 5vyc:bundle 1, present in both structures) to bring both structures in the same referential using the “fit molecules from selection” function with backbone atoms (rmsd 0.95A for 872 atoms). We then loaded the crystal structure of MCTS1-DENR complex (pdb entry 6MS4), in which a larger fragment of DENR has been resolved than in 5vyc and superposed it onto the DENR region resolved in pdb entry 5vyc.

Generation of endogenously tagged cell lines

C-terminal tagging of endogenous DENR with 3xFLAG-NeonGreen-dTAG in NIH/3T3 cells was carroed pit using CRISPR-Cas9 homology-directed repair. 10^5 NIH/3T3 cells were seeded in a 6-well plate 24h prior to transfection. 0.9µg of the pBLU donor plasmid (containing 400bp 5' and 3' homology arms directly upstream and downstream of *Denr* stop codon) as well as 0.5µg of the pCas9-gRNA and 0.1µg of the control pRRE reporter were transfected into NIH/3T3 using Lipofectamine™ 3000 (L3000008, Invitrogen). Medium was changed after 24h and cells were FACS sorted based on

mCherry and EGPF expression in a 10cm dish 48h post-transfection. Cells were counted and diluted to have 1 cell per well of a 96-well plate. When cell clones became visible, they were preselected based on NeonGreen expression, then split into a 96-well plate for genotyping and a 12-well plate for maintenance. For genomic DNA extraction, cells were washed with PBS then lysed in 20µl of DNA lysis buffer (10mM Tris-HCl pH8, 0.5mM EDTA, 0.5% Triton X-100, proteinase K 0.5mg/ml). After incubation for 1h at 55°C, cooling on ice and further incubation 10min at 95°C to inactivate proteinase K, this prep was used to check the insertion of the cassette into the *Denr* locus by PCR using primers upstream of the 5' and downstream of the 3' homology arms.

Co-immunoprecipitation

Triplicate 15cm plates of 80% confluent NIH/3T3 DENR-3xFLAG-NeonGreen-dTAG cells were treated with 500nM dTAG-13 or control-treated for 5h prior to harvesting. After washing with PBS, cells were scraped in 1ml PBS (4°C). Cells were collected by centrifugation at 500 x g for 5min at 4°C and lysed in 400µl IP buffer (50mM Tris-HCl pH7.5, 120mM NaCl, 1% NP-40, 0.5mM DTT, 10mM MgCl2, 1x PhosSTOP™ (4906837001, Roche), 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail (Cat. No. 4693132001, Roche), 0.04U/µl RNasin® Plus Ribonuclease Inhibitor (CN2611, Promega)) for 30 min on ice with mixing by pipetting every 10min. Lysed cells were centrifuged for 7min at 5000 x g at 4°C and the supernatant was collected. Proteins were quantified using Pierce™ BCA Protein Assay kit (Cat. No. 23225, Thermo Fisher Scientific) and 500µl of 2mg/ml proteins were incubated with washed anti-FLAG magnetic beads (Cat. No. M8823, Millipore) at 4°C for 3h on rotating wheel. The flow-through was kept for Western Blot and beads were washed twice in wash buffer (20mM Tris-HCl pH 7.5, 100mM NaCl, 0.5mM DTT, 10mM $MgCl_2$, 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail (Cat. No. 4693132001, Roche), 0.04U/µl RNasin® Plus Ribonuclease Inhibitor (N2611, Promega)) with change of tubes between the 1st and the 2nd wash. Proteins were eluted in 80µl elution buffer (NH_4OH pH 11-12) and incubated for 15min at 4°C. Elution step was repeated twice and eluates were pooled, frozen and dried in a SpeedVac system.

Mass-spectrometry and data analysis

Samples were digested following a modified version of the iST method (named miST method). Briefly, dried material was redissolved in 50µl miST lysis buffer (1% Sodium deoxycholate, 100mM Tris pH 8.6, 10mM DTT), heated 5min at 95°C and diluted 1:1 (v:v) with water. Reduced disulfides were alkylated by adding ¼ vol. of 160mM chloroacetamide (32mM final) and incubating for 45min at RT in the dark. Samples were adjusted to 3mM EDTA and digested with 1.0µg Trypsin/LysC mix (Promega #V5073) for 1h at 37°C, followed by a second 1h digestion with an additional 0.5µg of proteases. To remove sodium deoxycholate, two sample volumes of isopropanol containing 1% TFA were added to the digests, and the samples were desalted on a strong cation exchange (SCX) plate (Oasis MCX; Waters Corp., Milford, MA) by centrifugation. After washing with isopropanol/1%TFA, peptides were eluted in 200µl of 80% MeCN, 19% water, 1% (v/v) ammonia, and dried by centrifugal evaporation. Tryptic peptide mixtures were injected on an Ultimate RSLC 3000 nanoHPLC system interfaced via a nanospray Flex source to a high resolution Orbitrap Exploris 480 mass spectrometer (Thermo Fisher, Bremen, Germany). Peptides were loaded onto a trapping microcolumn Acclaim PepMap100 C18 (20mm x 100µm ID, 5µm, Dionex) before separation on a C18 custom packed column (75µm ID × 45cm, 1.8µm particles, Reprosil Pur, Dr. Maisch), using a gradient from 4% to 90% acetonitrile in 0.1% formic acid for peptide separation (total time: 140min). Full MS survey scans were performed at 120,000 resolution. A data-dependent acquisition method controlled by Xcalibur software (Thermo Fisher Scientific) was used that optimised the number of precursors selected (“top speed”) of charge 2+ to 5+ while maintaining a fixed scan cycle of 2s. Peptides were fragmented by higher energy collision dissociation (HCD) with a normalized energy of 30% at 15’000 resolution. The window for precursor isolation was of 1.6m/z units around the precursor and selected fragments were excluded for 60s from further analysis. Data files were analysed with MaxQuant 1.6.14.0 incorporating the Andromeda search engine Cysteine carbamidomethylation was selected as fixed modification while methionine oxidation and protein N-terminal acetylation were specified as variable modifications. The sequence databases used for searching were the mouse *(Mus musculus)* reference proteome based on the UniProt database (www.uniprot.org, version of April 6th, 2021, containing

55’341 sequences RefProt _Mus_musculus_20210604.fasta) supplemented with the sequences of common contaminants, and a database of mammalian immunoglobulin sequences (https://www.imgt.org version of 20th Nov. 2018, 2243 entries). Mass tolerance was 4.5ppm on precursors (after recalibration) and 20ppm on MS/MS fragments. Both peptide and protein identifications were filtered at 1% FDR relative to hits against a decoy database built by reversing protein sequences. All subsequent analyses were done with the Perseus software package (version 1.6.15.0). Contaminant proteins were removed, and intensity iBAQ values were $\log_2$-transformed. After assignment to groups, only proteins quantified in at least 3 samples of one group were kept. After missing values imputation (based on normal distribution using Perseus default parameters), *t*-tests were carried out among all conditions, with permutation-based FDR correction for multiple testing (Q-value threshold <0.05). Log_2FC of non-treated Denr-3xFLAG-NeonGreen-dTAG vs. dTAG treated Denr-3xFLAG-NeonGreen-dTAG was calculated and p-value was corrected for multiple testing by False discovery rate (significance threshold was set to FDR < 0.05).

Acknowledgements

We thank the Lausanne Genomic Technologies Facility for library sequencing, the Lausanne Protein Analysis Facility for mass spectrometry-based proteomics work, A. Teleman for the *Denr*, *Eif2d* and *Denr Eif2d* double KO HeLa cell lines, Evan Karousis for help with the initial setup of the *in vitro* translation assay and for the GADD34-expression plasmid, and Michael Taschner for his help with preparation of recombinant GADD34. We thank members of the Gatfield lab for critical comments on the manuscript.

Funding

This work was supported by Swiss National Science Foundation (SNSF) NCCR RNA & Disease Phase III funding (grant number 205601) and by project grant number 212423 (to D.G.).

Author contributions

Conceptualization: D.G., R.M. Methodology: R.M., M.D.M. Investigation: R.M., M.D.M. Visualization: R.M., N.G., D.G. Supervision: D.G. Writing original draft: R.M., D.G.

Competing interests

The authors declare that they have no competing interests.

Data and script availability

Sequencing data have been deposited at GEO (GSE263991) and mass spectrometry data have been submitted to PRIDE (PXD051482).

References

1. Wethmar K: **The regulatory potential of upstream open reading frames in eukaryotic gene expression.** *Wiley Interdiscip Rev RNA* 2014, **5:**765-778.
2. Jousse C, Bruhat A, Carraro V, Urano F, Ferrara M, Ron D, Fafournoux P: **Inhibition of CHOP translation by a peptide encoded by an open reading frame localized in the chop 5'UTR.** *Nucleic Acids Res* 2001, **29:**4341-4351.
3. Lee YY, Cevallos RC, Jan E: **An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation.** *J Biol Chem* 2009, **284:**6661-6673.
4. Hinnebusch AG, Ivanov IP, Sonenberg N: **Translational control by 5'-untranslated regions of eukaryotic mRNAs.** *Science* 2016, **352:**1413-1416.
5. Clemm von Hohenberg K, Muller S, Schleich S, Meister M, Bohlen J, Hofmann TG, Teleman AA: **Cyclin B/CDK1 and Cyclin A/CDK2 phosphorylate DENR to promote mitotic protein translation and faithful cell division.** *Nat Commun* 2022, **13:**668.
6. Schulz J, Mah N, Neuenschwander M, Kischka T, Ratei R, Schlag PM, Castanos-Velez E, Fichtner I, Tunn PU, Denkert C, et al: **Loss-of-function uORF mutations in human malignancies.** *Sci Rep* 2018, **8:**2395.
7. Occhi G, Regazzo D, Trivellin G, Boaretto F, Ciato D, Bobisse S, Ferasin S, Cetani F, Pardi E, Korbonits M, et al: **A novel mutation in the upstream open reading frame of the CDKN1B gene causes a MEN4 phenotype.** *PLoS Genet* 2013, **9:**e1003350.
8. von Bohlen AE, Bohm J, Pop R, Johnson DS, Tolmie J, Stucker R, Morris-Rosendahl D, Scherer G: **A mutation creating an upstream initiation codon in the SOX9 5' UTR causes acampomelic campomelic dysplasia.** *Mol Genet Genomic Med* 2017, **5:**261-268.
9. Hornig NC, de Beaufort C, Denzer F, Cools M, Wabitsch M, Ukat M, Kulle AE, Schweikert HU, Werner R, Hiort O, et al: **A Recurrent Germline Mutation in the 5'UTR of the Androgen Receptor Causes Complete Androgen Insensitivity by Activating Aberrant uORF Translation.** *PLoS One* 2016, **11:**e0154158.
10. Schleich S, Strassburger K, Janiesch PC, Koledachkina T, Miller KK, Haneke K, Cheng YS, Kuchler K, Stoecklin G, Duncan KE, Teleman AA: **DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth.** *Nature* 2014, **512:**208-212.
11. Schleich S, Acevedo JM, Clemm von Hohenberg K, Teleman AA: **Identification of transcripts with short stuORFs as targets for DENR*MCTS1-dependent translation in human cells.** *Sci Rep* 2017, **7:**3722.
12. Mohammad MP, Munzarova Pondelickova V, Zeman J, Gunisova S, Valasek LS: **In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation.** *Nucleic Acids Res* 2017, **45:**2658-2674.

13. Skabkin MA, Skabkina OV, Dhote V, Komar AA, Hellen CU, Pestova TV: **Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling.** *Genes Dev* 2010, **24:**1787-1801.
14. Skabkin MA, Skabkina OV, Hellen CU, Pestova TV: **Reinitiation and other unconventional posttermination events during eukaryotic translation.** *Mol Cell* 2013, **51:**249-264.
15. Zlotorynski E: **Translation. DENR-MCT1 reinitiates translation.** *Nat Rev Mol Cell Biol* 2014, **15:**498.
16. Castelo-Szekely V, De Matos M, Tusup M, Pascolo S, Ule J, Gatfield D: **Charting DENR-dependent translation reinitiation uncovers predictive uORF features and links to circadian timekeeping via Clock.** *Nucleic Acids Res* 2019, **47:**5193-5209.
17. Dmitriev SE, Terenin IM, Andreev DE, Ivanov PA, Dunaevsky JE, Merrick WC, Shatsky IN: **GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor.** *J Biol Chem* 2010, **285:**26779-26787.
18. Young DJ, Makeeva DS, Zhang F, Anisimova AS, Stolboushkina EA, Ghobakhlou F, Shatsky IN, Dmitriev SE, Hinnebusch AG, Guydosh NR: **Tma64/eIF2D, Tma20/MCT-1, and Tma22/DENR Recycle Post-termination 40S Subunits In Vivo.** *Mol Cell* 2018, **71:**761-774 e765.
19. Bohlen J, Harbrecht L, Blanco S, Clemm von Hohenberg K, Fenzl K, Kramer G, Bukau B, Teleman AA: **DENR promotes translation reinitiation via ribosome recycling to drive expression of oncogenes including ATF4.** *Nat Commun* 2020, **11:**4676.
20. Vasudevan D, Neuman SD, Yang A, Lough L, Brown B, Bashirullah A, Cardozo T, Ryoo HD: **Translational induction of ATF4 during integrated stress response requires noncanonical initiation factors eIF2D and DENR.** *Nat Commun* 2020, **11:**4677.
21. Young DJ, Meydan S, Guydosh NR: **40S ribosome profiling reveals distinct roles for Tma20/Tma22 (MCT-1/DENR) and Tma64 (eIF2D) in 40S subunit recycling.** *Nat Commun* 2021, **12:**2976.
22. Gurzeler LA, Ziegelmuller J, Muhlemann O, Karousis ED: **Production of human translation-competent lysates using dual centrifugation.** *RNA Biol* 2022, **19:**78-88.
23. Mikami S, Kobayashi T, Imataka H: **Cell-free protein synthesis systems with extracts from cultured human cells.** *Methods Mol Biol* 2010, **607:**43-52.
24. Ahmed YL, Schleich S, Bohlen J, Mandel N, Simon B, Sinning I, Teleman AA: **DENR-MCTS1 heterodimerization and tRNA recruitment are required for translation reinitiation.** *PLoS Biol* 2018, **16:**e2005160.
25. Dever TE, Ivanov IP, Hinnebusch AG: **Translational regulation by uORFs and start codon selection stringency.** *Genes Dev* 2023, **37:**474-489.
26. Alghoul F, Laure S, Eriani G, Martin F: **Translation inhibitory elements from Hoxa3 and Hoxa11 mRNAs use uORFs for translation inhibition.** *Elife* 2021, **10**.
27. Haas MA, Ngo L, Li SS, Schleich S, Qu Z, Vanyai HK, Cullen HD, Cardona-Alberich A, Gladwyn-Ng IE, Pagnamenta AT, et al: **De Novo Mutations in DENR Disrupt Neuronal Development and Link Congenital Neurological Disorders to Faulty mRNA Translation Re-initiation.** *Cell Rep* 2016, **15:**2251-2265.
28. Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A, et al: **The dTAG system for immediate and target-specific protein degradation.** *Nat Chem Biol* 2018, **14:**431-441.
29. Wood AJ, Roberts RG, Monk D, Moore GE, Schulz R, Oakey RJ: **A screen for retrotransposed imprinted genes reveals an association between X chromosome homology and maternal germ-line methylation.** *PLoS Genet* 2007, **3:**e20.

30. Bohlen J, Zhou Q, Philippot Q, Ogishi M, Rinchai D, Nieminen T, Seyedpour S, Parvaneh N, Rezaei N, Yazdanpanah N, et al: **Human MCTS1-dependent translation of JAK2 is essential for IFN-gamma immunity to mycobacteria.** *Cell* 2023, **186:**5114-5134 e5127.
31. Weisser M, Schafer T, Leibundgut M, Bohringer D, Aylett CHS, Ban N: **Structural and Functional Insights into Human Re-initiation Complexes.** *Mol Cell* 2017, **67:**447-456 e447.
32. Nandi S, Shi B, Perreault J, Gartenhaus RB: **Characterization of the MCT-1 pseudogene: identification and implication of its location in a highly amplified region of chromosome 20.** *Biochim Biophys Acta* 2006, **1759:**234-239.
33. Hronova V, Mohammad MP, Wagner S, Panek J, Gunisova S, Zeman J, Poncova K, Valasek LS: **Does eIF3 promote reinitiation after translation of short upstream ORFs also in mammalian cells?** *RNA Biol* 2017, **14:**1660-1667.
34. Shu XE, Mao Y, Jia L, Qian SB: **Dynamic eIF3a O-GlcNAcylation controls translation reinitiation during nutrient stress.** *Nat Chem Biol* 2022, **18:**134-141.
35. Bohlen J, Fenzl K, Kramer G, Bukau B, Teleman AA: **Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells.** *Mol Cell* 2020, **79:**561-574 e565.
36. Pestova TV, Hellen CU: **Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA.** *Genes Dev* 2003, **17:**181-186.
37. Zinoviev A, Hellen CUT, Pestova TV: **Multiple mechanisms of reinitiation on bicistronic calicivirus mRNAs.** *Mol Cell* 2015, **57:**1059-1073.
38. Lomakin IB, Stolboushkina EA, Vaidya AT, Zhao C, Garber MB, Dmitriev SE, Steitz TA: **Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1.** *Cell Rep* 2017, **20:**521-528.
39. Rehfeld F, Eitson JL, Ohlson MB, Chang TC, Schoggins JW, Mendell JT: **CRISPR screening reveals a dependency on ribosome recycling for efficient SARS-CoV-2 programmed ribosomal frameshifting and viral replication.** *Cell Rep* 2023, **42:**112076.
40. Sonobe Y, Aburas J, Krishnan G, Fleming AC, Ghadge G, Islam P, Warren EC, Gu Y, Kankel MW, Brown AEX, et al: **A C. elegans model of C9orf72-associated ALS/FTD uncovers a conserved role for eIF2D in RAN translation.** *Nat Commun* 2021, **12:**6025.
41. Janich P, Arpat AB, Castelo-Szekely V, Lopes M, Gatfield D: **Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames.** *Genome Res* 2015, **25:**1848-1859.
42. Castelo-Szekely V, Arpat AB, Janich P, Gatfield D: **Translational contributions to tissue specificity in rhythmic and constitutive gene expression.** *Genome Biol* 2017, **18:**116.
43. Knuckles P, Carl SH, Musheev M, Niehrs C, Wenger A, Buhler M: **RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding.** *Nat Struct Mol Biol* 2017, **24:**561-569.
44. Flemr M, Buhler M: **Single-Step Generation of Conditional Knockout Mouse Embryonic Stem Cells.** *Cell Rep* 2015, **12:**709-716.
45. Salmon P, Trono D: **Production and titration of lentiviral vectors.** *Curr Protoc Hum Genet* 2007, **Chapter 12:**Unit 12 10.
46. Lavery DJ, Schibler U: **Circadian transcription of the cholesterol 7 alpha hydroxylase gene may involve the liver-enriched bZIP protein DBP.** *Genes Dev* 1993, **7:**1871-1884.
47. Arpat AB, Liechti A, De Matos M, Dreos R, Janich P, Gatfield D: **Transcriptome-wide sites of collided ribosomes reveal principles of translational pausing.** *Genome Res* 2020, **30:**985-999.

48. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** *EMBnetjournal* 2011, **17:**10-12.
49. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat Methods* 2012, **9:**357-359.
50. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013, **29:**15-21.
51. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL: **StringTie enables improved reconstruction of a transcriptome from RNA-seq reads.** *Nat Biotechnol* 2015, **33:**290-295.
52. Robinson MD, Oshlack A: **A scaling normalization method for differential expression analysis of RNA-seq data.** *Genome Biol* 2010, **11:**R25.
53. Chothani S, Adami E, Ouyang JF, Viswanathan S, Hubner N, Cook SA, Schafer S, Rackham OJL: **deltaTE: Detection of Translationally Regulated Genes by Integrative Analysis of Ribo-seq and RNA-seq Data.** *Curr Protoc Mol Biol* 2019, **129:**e108.

Figure Legends

Figure 1. DENR is a re-initiation factor that promotes translation of hundreds of uORF-containing transcripts, including *Klhdc8a* and *Asb8*.

(A) Immunoblot analysis of NIH/3T3 cells confirms DENR and eIF2D depletion upon lentiviral transduction with *Denr* or *Eif2d* shRNAs as compared to scramble (Scr) shRNA and *Eif2d* non-functional (nf) shRNA. β-tubulin serves as loading control.

(B) Scatter plot of changes between *Denr* shRNA and control shRNA treated cells for RNA abundances vs. translation efficiencies (RPF counts/RNA counts), based on ribosome profiling and RNA-seq experiments. Evaluation is based on *Denr* shRNA (triplicates) and two independent control shRNAs (triplicates each). mRNAs with significant change for TE are represented in green and mRNAs with significant change for RNA abundance are represented in olive (adjusted p-values < 0.10, Wald test followed by FDR adjustment). Position of *Asb8* and *Klhdc8a* are indicated.

(C) Violin plot showing 5' UTR lengths of all expressed transcripts (n=9203, grey) vs. transcripts with lower TE upon *Denr* depletion (n=221, green). DENR-responsive genes have longer 5' UTRs (median = 260 nt) than overall expressed transcripts (median = 153 nt) (p = 1.12e-14, Kolmogorov–Smirnov test).

(D) Analysis of proportion of transcripts with at least one translated uORF. The transcripts with lower TE upon *Denr* depletion are enriched for translated uORFs. For all expressed transcripts (n = 9266), 3430 carry a translated uORF, whereas 3370 have a uORF sequence that is non-translated and 2465 have no uORF at all. In *Denr* shRNA cells, transcripts with significantly reduced TE (n=179 after removal of transcripts with ambiguous/several expressed 5' UTRs), 156 have translated uORFs vs. 13 non-translated uORFs and 10 no uORF (p-value < 1e-5, Fisher's exact test).

(E) The uORF length distribution of transcripts with reduced TE in *Denr* shRNA transduced cells (n=522) is similar to that of all expressed transcripts (n=7811) (p-value = 0.3717, Kolmogorov-Smirnov test).

(F) Mapped footprint A-sites of *Klhdc8a* and *Asb8* transcripts in control and *Denr* knock-down cells. Read counts were normalized to library depth by subsampling and replicates were merged for increased coverage.

(G) Quantification of translation efficiencies on the coding sequence for *Klhdc8a* and *Asb8* in control and *Denr*-depleted cells (*Klhdc8a* p = 0.01 and *Asb8* p = 2.81e-3, two-tailed unpaired *t*-test)

(H) Schematic of the *Klhdc8a* 5' UTR reporters used for the *in vivo* and *in vitro* translation assays. The 'no uORF reporter' serves to evaluate the regulation by the 5' UTR in the absence of a uORF (100% signal). The 'overlapping uORF reporter' can only produce luciferase signal through leaky scanning. Comparison of the WT reporter, containing the relevant uORF, with the two former reporters allows to

specifically calculate re-initiation. Not depicted: all plasmids also express Renilla luciferase for internal normalization.

(I) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after transduction in WT and *Denr* KO HeLa cells. The canonical initiation signal of the 'no uORF reporter' was set to 100% in WT and *Denr* KO cells individually. uORF inhibition, re-initiation and leaky scanning are calculated as indicated by the shaded boxes (significance calculated using two-tailed unpaired *t*-test).

(J) Western Blot analysis of WT, *Denr* KO and *Eif2d* KO HeLa cells validates depletion of KLHDC8A in the absence of DENR. U2AF65 serves as loading control.

(K) Schematic representation of ribosomal fluxes on *Klhdc8a* 5' UTR estimated from the results shown in panel (I).

Figure 2. *In vitro* translation of *Klhdc8a* and *Asb8* reporters recapitulates regulation *in vivo*.

(A) Outline of the preparation of HeLa translation-competent extracts and of the capped RNA reporters for the *in vitro* translation assays. The RNA for Renilla reporter was consistently used in all *in vitro* assays for normalization purposes.

(B) Immunoblot analysis of proteins extracted from HeLa cells, i.e. whole cell lysate, supernatant and pellet post-dual-centrifugation and translation-competent extract post-translation. MCTS1, DENR, eIF2D and RPL23 are consistently recovered in the translation-competent extracts.

(C) Immunoblot analysis of whole HeLa cell lysate and translation-competent extracts after *in vitro* translation reveals increased eIF2α-p51 during the *in vitro* translation reaction. β-tubulin serves as loading control.

(D) Raw luminescence signals of firefly and Renilla luciferase reporters after *in vitro* translation in HeLa translation-competent extracts complemented with or without GADD34Δ1-240 validates reduced translation efficiency upon eIF2α phosphorylation.

(E) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Denr* KO HeLa lysates (left) and rescue assay with 0.5µM recombinant MCTS1-DENR (right); 'no uORF reporter' signal was set to 100% and uORF inhibition, re-initiation and leaky scanning were calculated as indicated by the shaded boxes (significance calculated using two-tailed unpaired *t*-test).

(F) Schematic representation of ribosomal fluxes on *Klhdc8a* 5' UTR reporter estimated from the results shown in panel (E).

(G) As in (E), but for the *Asb8* reporters.

(F) As in (F) for the *Asb8* 5' UTR in *vitro* data shown in (G).

(I) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters with mutated uORF penultimate codon (GTG to GTC or GCG) reveals the effect of penultimate codon identity on DENR-dependence.

(J) Normalized luminescence signal (firefly/Renilla) of the *Asb8* reporters with inserted codon (TGT or GCG) between the uORF start and stop codons indicates that lengthening of the uORF decreases DENR-dependence.

Figure 3. eIF2α phosphorylation status moderately affects re-initiation rate of start-stop uORF and longer uORF reporters.

(A) Normalized luminescence signal of *Klhdc8a* reporters after *in vitro* translation in WT and *Denr* KO HeLa extracts complemented with 16ng/µl recombinant GADD34Δ1-240 (left) or without GADD34Δ1-240 (right) (significance calculated using two-tailed unpaired *t*-test).

(B) Schematic representation of ribosomal fluxes on *Klhdc8a* 5' UTR *in vitro* estimated from the results shown in panel (A). Quantifications in presence of GADD34Δ1-240 are shown in black and without GADD34Δ1-240 in blue.

(C) As in (A) for *Asb8* reporters.

(D) As in (B) for *Asb8* reporter data shown in (C).

Figure 4. No evidence for eIF2D uORF re-initiation activity *in vivo* but low activity *in vitro*.

(A) Schematic representation of human MCTS1-DENR and eIF2D proteins, with domains indicated. The relevant amino acids positions defining the domains are specified.

(B) Distribution of ribosomes on 5' UTR relative to CDS in *Denr* knock-down (in green) or *Eif2d* knock-down (in blue) vs. control cells (n = 4242). *Denr* knock-down cells show a strong redistribution of ribosomes from CDS to 5' UTR whereas the effect is milder but still significant for *Eif2d* knock-down cells (p-values calculated using Wilcoxon signed rank test).

(C) Scatter plot of changes between *Eif2d* shRNA and control cells for RNA abundances vs. translation efficiencies (RPF counts/RNA counts), based on ribosome profiling and RNA-seq experiments. Evaluation is based on two independent *Eif2d* shRNAs (triplicates each) and two independent control shRNAs (triplicates each). mRNAs with significant change on TE were represented in blue and mRNAs with significant change on RNA abundance were represented olive green (adjusted p-values < 0.10, Wald test followed by FDR adjustment). Positions of *Asb8*, *Klhdc8a*, *Hoxa3*, *Med23*, *Rps20*, *Ndc80* and *Cenpa* are indicated.

(D) Analysis of proportion of transcripts with at least one translated uORF. The transcripts with lower TE upon *Eif2d* depletion are not enriched for translated uORFs. For all expressed transcripts (n = 9266), 3430 carry a translated uORF, whereas 3370 have a uORF sequence that is non-translated and 2465 have no uORF at all. In *Eif2d* shRNA cells, transcripts with significantly reduced TE (n=58 after removal of transcripts with ambiguous/several expressed 5' UTRs), 19 have translated uORFs vs. 19 non-translated uORFs and 20 no uORF (p-value = 0.59, Fisher’s exact test).

(E) Quantification of translation efficiencies on the coding sequence of selected transcripts with lower TE upon *Eif2d* depletion in control and *Eif2d*-depleted cells (p-values calculated using a two-tailed unpaired *t*-test).
(F) Normalized luminescence signal (firefly/Renilla) of lentivirally transduced reporters with 5' UTRs of selected transcripts with lower TE upon *Eif2d* depletion (significance calculated using two-tailed unpaired *t*-test). The signal of individual reporter expressed in WT cells was set to 100%.
(G) Normalized luminescence signal (firefly/Renilla) of the *Asb8* reporters after *in vitro* translation in WT and *Eif2d* KO HeLa extracts (p-value calculated using two-tailed unpaired *t*-test).
(H) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Eif2d* KO HeLa extracts (left), supplemented with 0.5µM recombinant eIF2D (middle) or supplemented with 0.5µM of recombinant MCTS1-DENR (right, p-values calculated using two-tailed unpaired *t*-test).
(I) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Denr* KO HeLa extracts (left) or supplemented with 0.5µM recombinant eIF2D (right); p-value calculated using two-tailed unpaired *t*-test.
(J) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Denr* / *Eif2d* double KO HeLa extracts (left), supplemented with 0.5µM recombinant eIF2D (middle) or supplemented with 0.5µM recombinant MCTS1-DENR (right); p-values calculated using two-tailed unpaired *t*-test.
(K) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Eif2d* KO HeLa extracts treated with GADD34Δ1-240 (left), without GADD34 (middle), or without GADD34 and complemented with 0.5µM eIF2D (right); p-values calculated using two-tailed unpaired *t*-test.

Figure 5. MCTS2 interacts with DENR *in vivo* and promotes re-initiation *in vitro*.
(A) Schematic representation of endogenous DENR tagged with 3xFLAG-NeonGreen-dTAG. Treatment with the dTAG-13 ligand targets DENR-3xFLAG-NeonGreen-dTAG to proteasomal degradation.
(B) Genotyping PCR analysis of *Denr*-tagged clones validates the insertion of the 3xFLAG-NeonGreen-dTAG cassette upstream of *Denr* stop codon. Tagged *Denr* amplicon has a length of 2067 bp and WT *Denr* of 885 bp.
(C) Western Blot analysis of NIH/3T3 DENR-3xFLAG-NeonGreen-dTAG cells treated for 7 or 14h with 500nM dTAG-13 shows complete depletion of DENR upon treatment. Asterisk indicates a non-specific band that migrates at a similar height of the tagged protein.
(D) Mass spectrometry analysis of proteins co-immunoprecipitated with endogenously tagged DENR in non-treated vs. dTAG-13-treated cells reveals the interaction of MCTS2 with DENR. Only DENR

and significant hits enriched in non-treated cells are shown (p-values calculated using two-tailed unpaired *t*-test followed by correction for multiple testing).

(E) Normalized RNA abundances and ribosome footprints of *Denr*, *Mcts1* and *Mcts2* in control and *Denr* depleted NIH/3T3 cells (p-values calculated using two-tailed unpaired *t*-test).

(F) Alignment of human, rhesus macaque, mouse and rat MCTS1 and MCTS2 amino acid sequences. The amino acids changing relative to mouse MCTS1 are highlighted with colors according to their side-chain chemistry.

(G) Structural model of MCTS1-DENR (turquoise/green) in interaction with the 40S ribosomal subunit (grey), Met-tRNA$^{Met}_{i}$ (magenta) and an IRES mRNA (yellow). The amino acids that undergo changes in MCTS2 are indicated in red in the structure.

(H) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Denr* KO HeLa lysates (left), substituted with 0.5µM of recombinant DENR, MCTS1-DENR or MCTS2-DENR as indicated (p-values calculated using two-tailed unpaired *t*-test).

(I) As in (H), for the *Asb8* reporters.

Supplementary Figure Legends

Supplementary Figure 1. *Klhdc8a* and *Asb8* both contain a single AUG uORF that is conserved across mammals.

(A) Nucleotide sequences of mouse *Klhdc8a* 5' UTR and beginning of CDS. Position and amino acid sequence of uORF and CDS are shown and highlighted in colors.

(B) As in (A), but for *Asb8* 5' UTR and beginning of CDS.

(C) Sequence alignments of the 3' portion of the 5' UTR of *Klhdc8a* from rat, mouse, rabbit, gorilla, human, pig and lion show good conservation of their uORF sequences.

(D) As in (C) for *Asb8*.

Supplementary Figure 2. Shortening of *Klhdc8a* 5' UTR increases the raw luciferase signal without affecting the re-initiation rate.

(A) Schematic representation of the short and long 5' UTR of *Klhdc8a* and the whole 5' UTR of *Asb8* WT reporters used in this study compared to the mouse *Atf4* transcript. uORF lengths, upstream and downstream sequence lengths and GC contents are indicated.

(B) Raw luminescence signal of long vs. short 5' UTRs of *Klhdc8a* reporters after *in vitro* translation in HeLa WT lysate shows higher signal upon shortening of the 5' UTR.

(C) Normalized luminescence signal (firefly/Renilla) of the *Klhdc8a* reporters after *in vitro* translation in WT and *Denr* KO HeLa lysates of the long 5' UTR reporters (left) and the short 5' UTR reporters (right, significance calculated using two-tailed unpaired *t*-test)

Supplementary Figure 3. Effect of tunicamycin treatment on *Klhdc8a* reporter *in vivo*.

(A) Western blot analysis of whole cell protein extracts or translation-competent extracts before and after *in vitro* translation, non-treated or treated with 1µg/ml tunicamycin or with 16ng/µl GADD34Δ1-240, reveals that eIF2α-p51 phosphorylation level increase is stronger in *in vitro* translation reactions without GADD34Δ1-240 than after tunicamycin-mediated ISR induction.

(B) Normalized luminescence signal (firefly/Renilla) of *Klhdc8a* reporters after transduction in WT and *Denr* KO cells non-treated (left) and treated with tunicamycin (right).

(C) Schematic representation of ribosomal fluxes on *Klhdc8a* 5' UTR estimated from the results shown in panel (B). Values measured in non-treated cells are shown in black and those quantified from tunicamycin-treated cells in blue.

(D) Normalized *in vivo* translation signal (firefly/Renilla) of *Atf4* and *Atf5* reporters measured in transduced HeLa WT and tunicamycin-treated cells confirms the increased translation of the two reporters upon eIF2α phosphorylation.

Supplementary Figure 4. Ribosome profiling and *in vivo* dual luciferase assay reveals differential effects of *Denr* and *Eif2d* knock-down.

(A) Venn diagram of transcripts with reduced TE upon *Denr* or *Eif2d* knock-down shows that only one transcript overlaps between the two groups.

(B) Correlation analysis between translation efficiency in *Denr* and *Eif2d* knock-down cells shows a weak but significant correlation (Pearson correlation = 0.1339, p-value < 2.2e-16).

(C) Violin plot of 5' UTR lengths of all expressed transcripts (n=9203, grey) vs. transcripts with lower TE upon *Denr* depletion (n=221, green) and *Eif2d* depletion (n=73, blue). eIF2D-responsive genes tend to have a shorter 5' UTR (median = 110 nt) than overall expressed transcripts (median = 153 nt) (p = 0.09, Kolmogorov–Smirnov test).

(D) Cumulative distributions of transcript per kilobase million (TPM) of all expressed transcripts, transcripts with decreased TE upon *Eif2d* knock-down and transcripts with decreased TE upon *Denr* knock-down show a strong enrichment for highly abundant transcripts among eIF2D-responsive mRNAs (p-value calculated using the Kolmogorov–Smirnov test).

(E) Gene Ontology analysis of transcripts with decreased TE upon *Eif2d* depletion (p-values adjusted for multiple tests using FDR).

(F) Cumulative distributions of TPM of all expressed transcripts, transcripts with changed abundance upon *Eif2d* knock-down and transcripts with changed abundance upon *Denr* knock-down show an enrichment for highly abundant transcripts among eIF2D-responsive mRNAs (p-value calculated using the Kolmogorov–Smirnov test).

(G) Gene Ontology analysis of transcripts with increased abundance upon *Eif2d* depletion (p-values adjusted for multiple tests using FDR).

(H) Gene Ontology analysis of transcripts with decreased abundance upon *Eif2d* depletion (p-values adjusted for multiple tests using FDR).

(I) Cell proliferation analysis of NIH/3T3 cells transduced with scramble shRNA, nf *Eif2d* shRNA, *Denr* shRNA and *Eif2d* shRNAs shows that a larger fraction of *Eif2d* knock-down cells are found in the G1 phase of the cell cycle than is the case across control cells or *Denr*-depleted cells (p-values calculated using two-tailed unpaired *t*-test).

(J) Normalized luminescence signals (firefly/Renilla) of *Klhdc8a* reporters after transduction in HeLa WT and *Eif2d* KO cells show no change in re-initiation rates upon *Eif2d* depletion (p-values calculated using two-tailed unpaired *t*-test).

Supplementary Figure 5. Alignment of mouse *Mcts1* and *Mcts2* coding sequences.

(A) Nucleotide alignment of *Mcts1* and *Mcts2* coding sequences from *Mus musculus*.

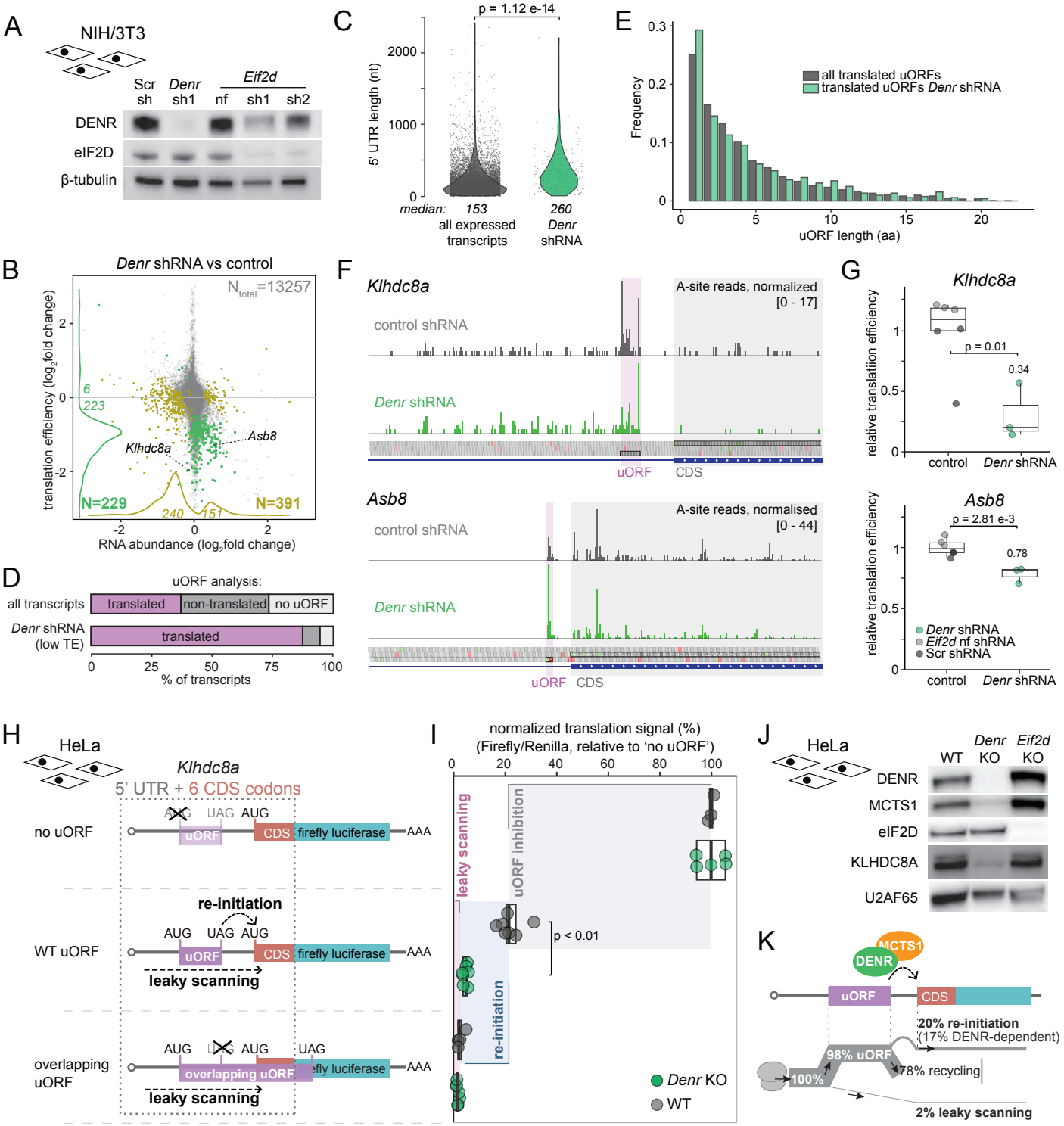
Supplementary Figure 6. DENR depletion leads to an accumulation of ribosomes at the uORF stop codon on the endogenous *Atf4* transcript.

(A) Mapped footprint A-sites of mouse *Atf4* transcripts in control, *Eif2d* and *Denr* knock-down cells. Read numbers were normalized to library depth by subsampling and replicates were merged for increased coverage.

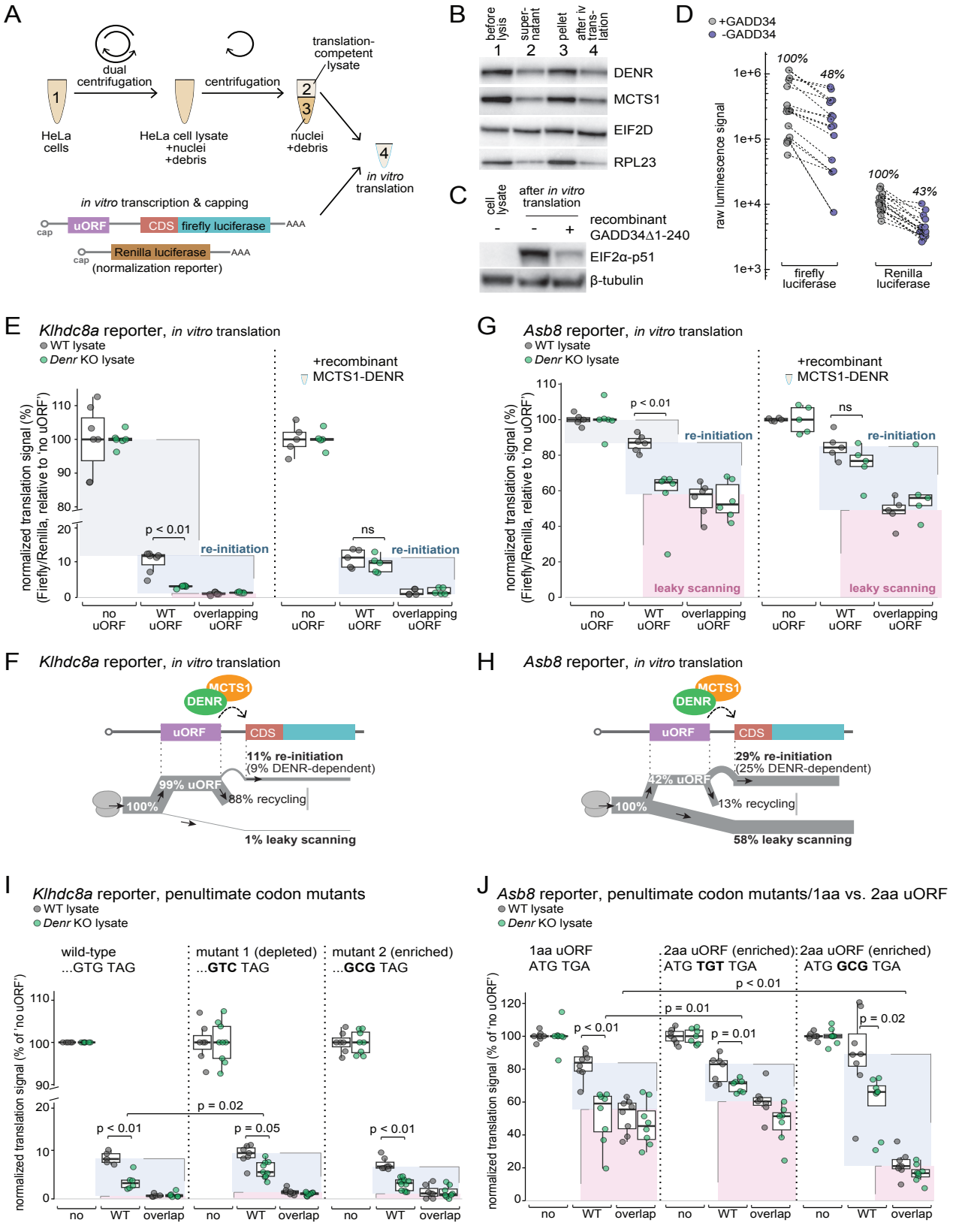
Supplementary Figure 7. *Mcts2* expression in mouse liver and kidney.

(A) RPKM of *Denr*, *Mcts1* and *Mcts2* around-the-clock in mouse liver and kidney confirms expression of *Mcts2* in adult mice. Liver and kidney RNA-seq and Ribo-seq data were taken from previously published studies [41, 42].

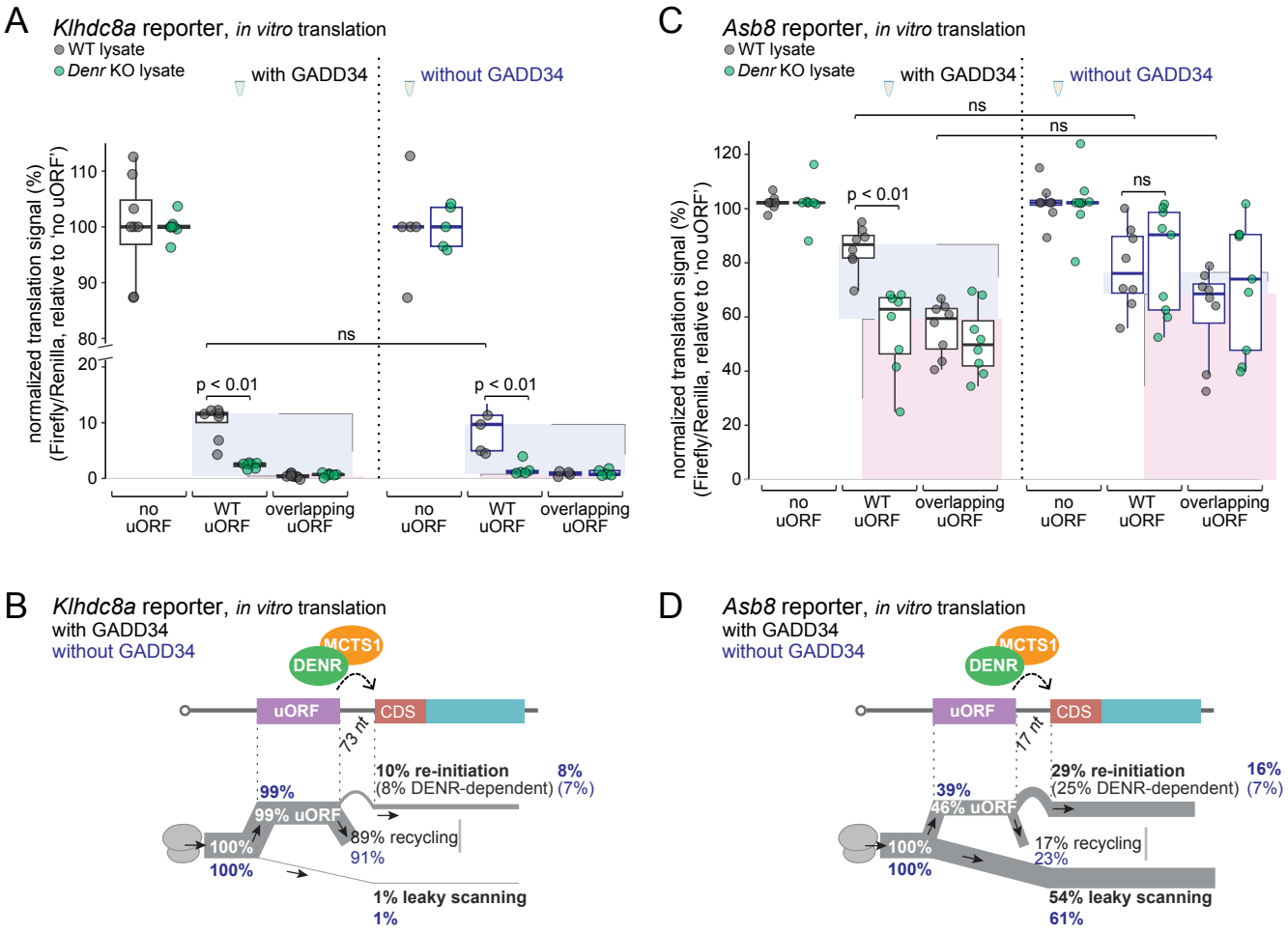
Meurs et al., Figure 1



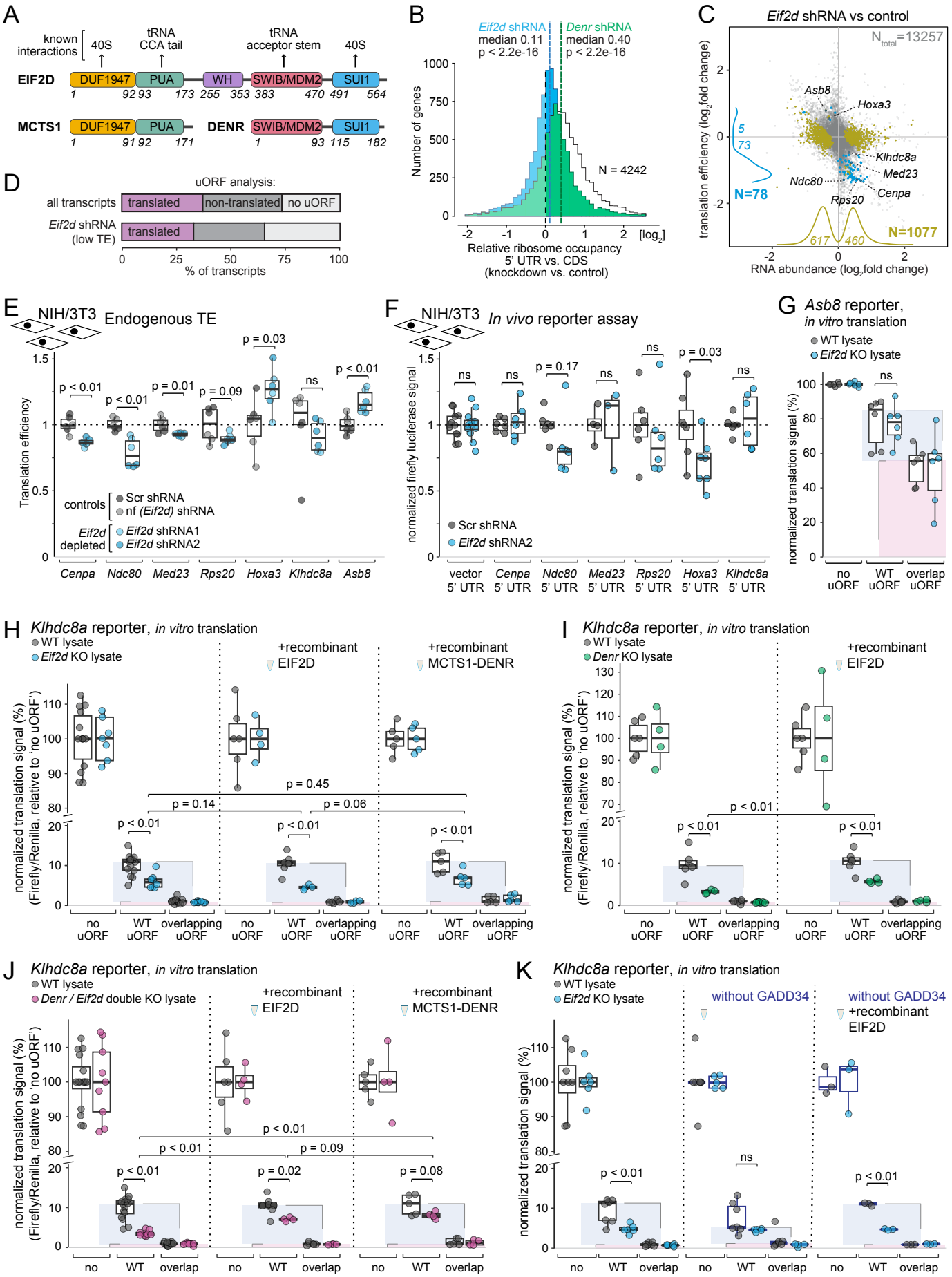
Meurs et al., Figure 2



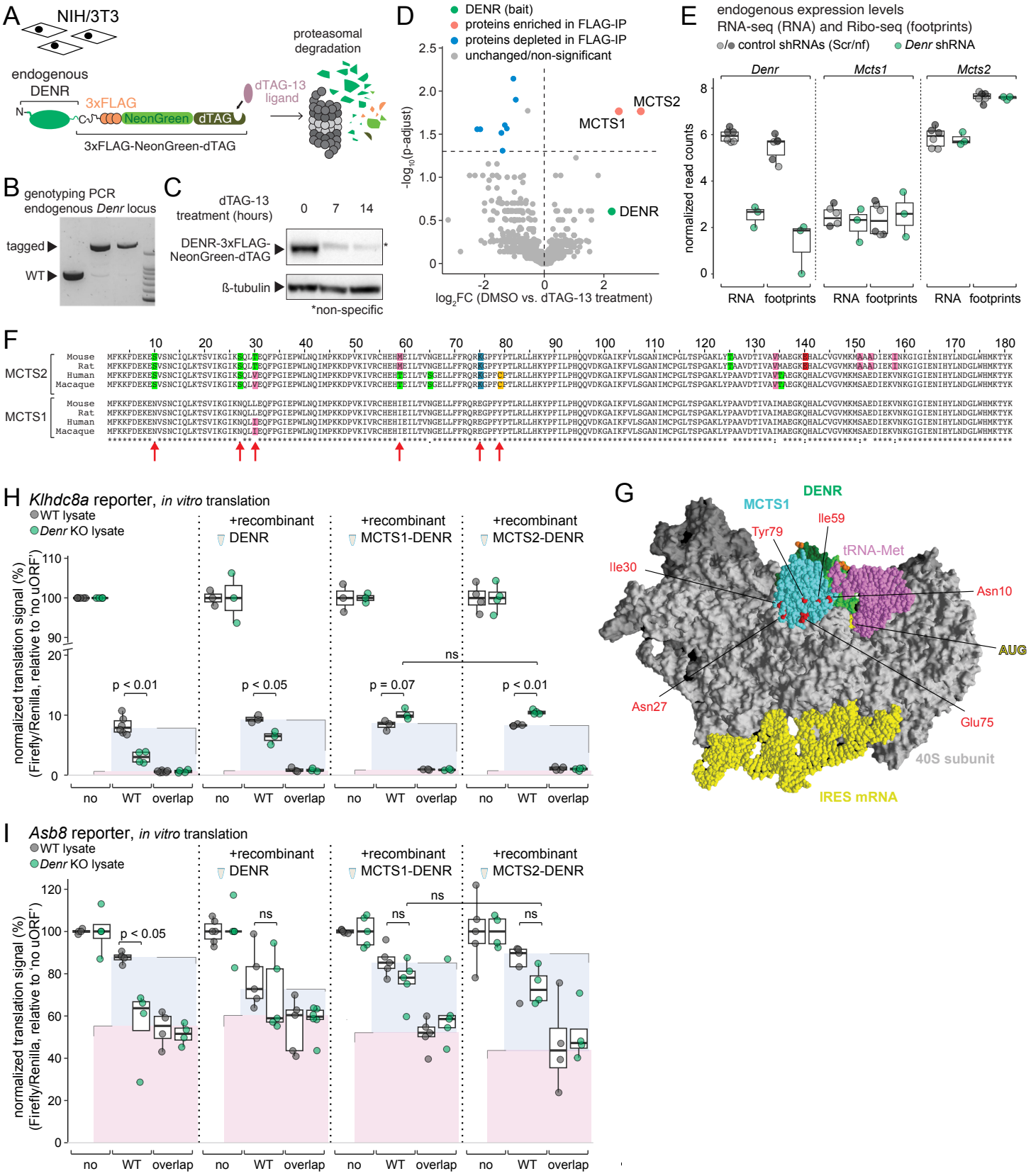
Meurs et al., Figure 3



Meurs et al., Figure 4



Meurs et al., Figure 5



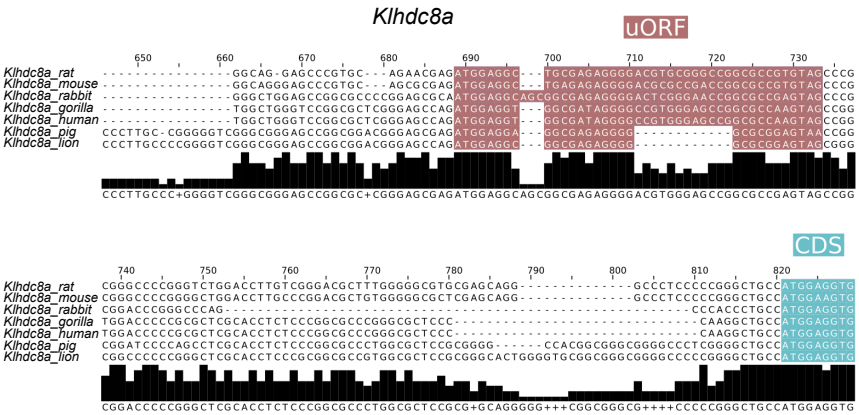
A mouse *Klhdc8a* 5' UTR and beginning of CDS (ENSMUST00000046071)

```
  1 GTCTCCGACCCTGTAGACACTGCAGGCCCTGCCCTCGCTCGCCCTGGTCGCCCGGGGCTG   60
 61 ACGCACAGGCCACAGCCTCCGCCTCGGGGCGCGTGGGCTCCTGCCCAGGTGTGGAAGCTG  120
121 CGGGGAGGCGGTGCGGGGAACCCGCGTCCTCCGGCCGGAGCTCGGCTTCTGGCGACCGCT  180
181 CTCCTGCGCGGGTGACGGCGGCGTGCGGGCTGGGAAGGGCAGGGAGCGCTGTCCTCCGCT  240
241 CCGCGGCTGCCAGGCGGACCGTGAAAGCTTCGCGCGGCTGTAGGCGCCGCCCGGGAGCTA  300
301 GTGCTGGGCGAGCGCGGCGCCCACGCGCGGTGGGGTGGGCTCCGAGCTCCGGGGGACCGG  360
361 GCGGGGGCGGGGCGCATCCTGCAAAGTTGACCAAGGGTGAGCTACCGCCGGCGCCCTGCT  420
421 GCCCTAGATTCGCCAGGTTCTTTCCGCGGAGCCAGGCGTGGACGTCGCTGCCCCGCTGAC  480
481 ACCCAAGGGACCCGGAGAAGCCAGCGGGCGTCGGCAGGGAGCCCGTGCAGCGCGAGATGG  540
                                                      uORF -M--
541 AGGCTGAGAGAGGGGACGCGCCGACCGGCGCCGTGTAGCCCGCGGGCCCCGGGGCTGGAC  600
    E--A--E--R--G--D--A--P--T--G--A--V--*-
601 CTTGCCCGGACGCTGTGGGGGCGCTCGAGCAGGGCCCTCCCCCGGGCTGCCATGGAAGTG  660
                                                      CDS -M--E--V-
661 CCCAATGTCAAGGACTTTCAGTGGAAGCGCCTTGCTCCACTGCCTAGCCGCCGGGTCTAT  720
    -P--N--V--K--D--F--Q--W--K--R--L--A--P--L--P--S--R--R--V--Y-
```

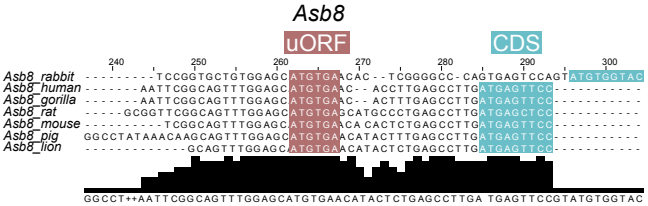
B mouse *Asb8* 5' UTR and beginning of CDS (ENSMUST00000143400)

```
  1 AGAATATCATGTCGCTCAGAGGCTCTGACCATGCAGCACCCAGAGAGCACAAGGGGTCTG   60
 61 GGCCAAGAGGGAGGGGAATTTCCTGGAGAAGCACTGAAGATCATGGGAAGCAGGTGCAGA  120
121 GGCCCAGGAGCGTGGCTGATCCAGTCCAATTCCAGGCAGTTTGGAGCATGTGAACACACT  180
                                                      uORF -M--*-
181 CTGAGCCTTGATGAGTTCCAGCATGTGGTATATTATGCAGAGCATTCAGAGCAAATACTC  240
          CDS -M--S--S--S--M--W--Y--I--M--Q--S--I--Q--S--K--Y--S
```

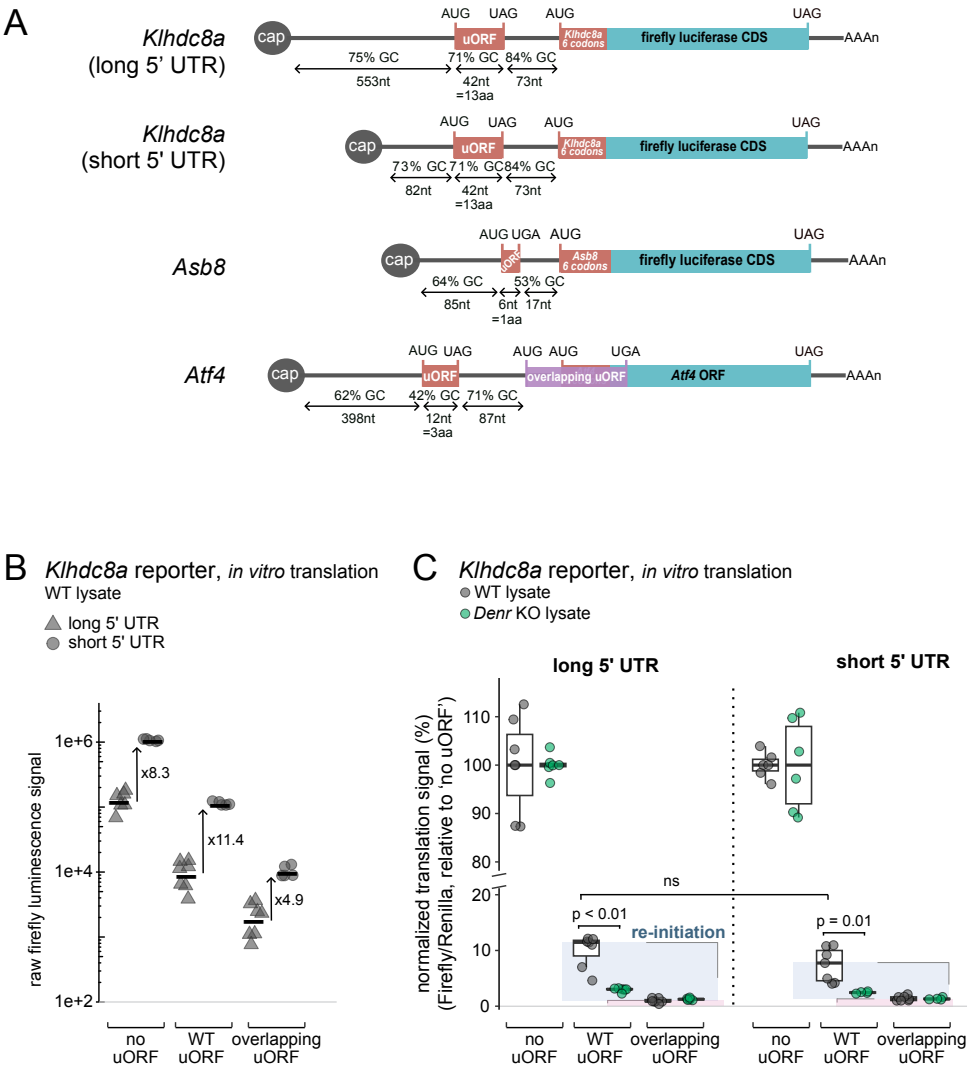
C



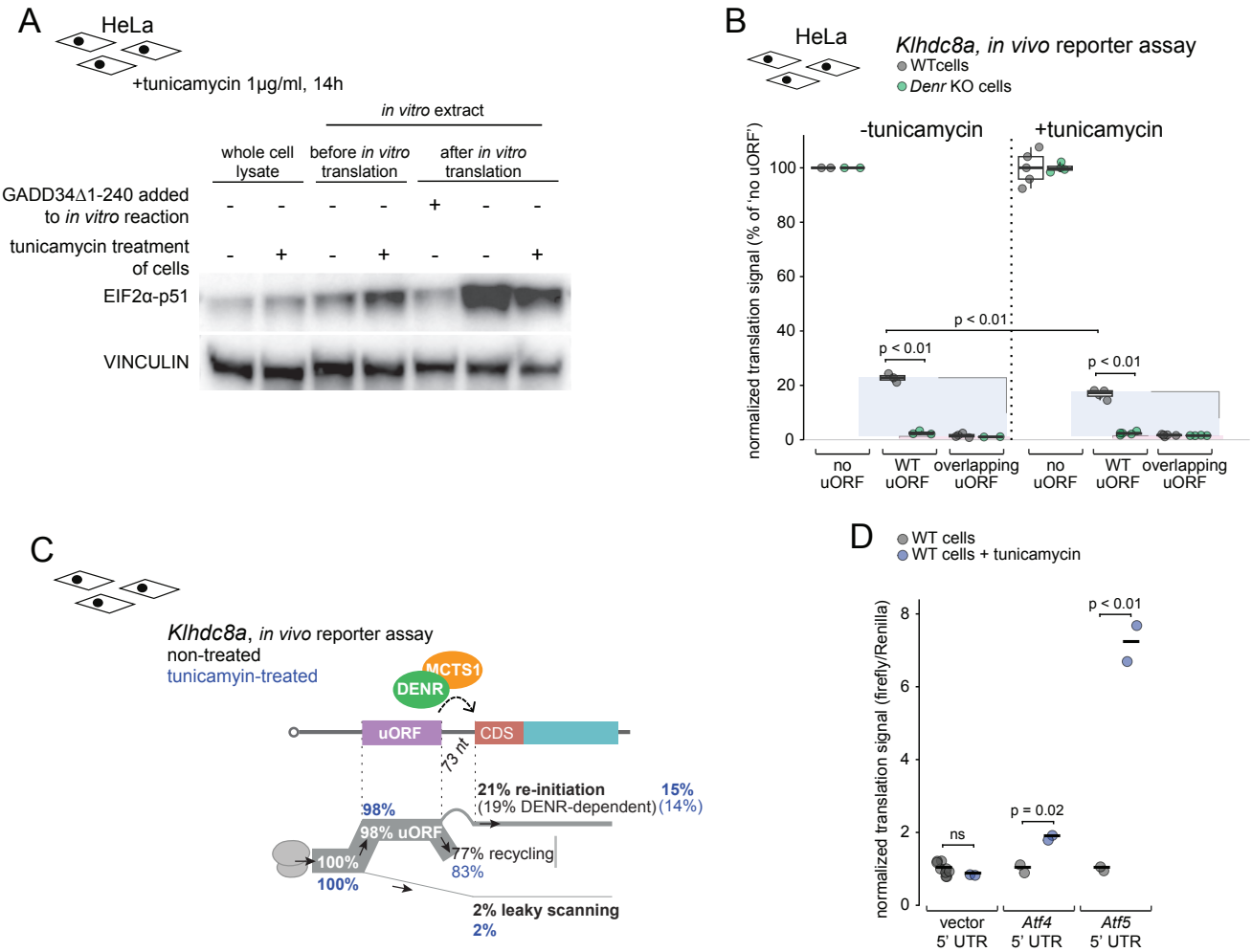
D



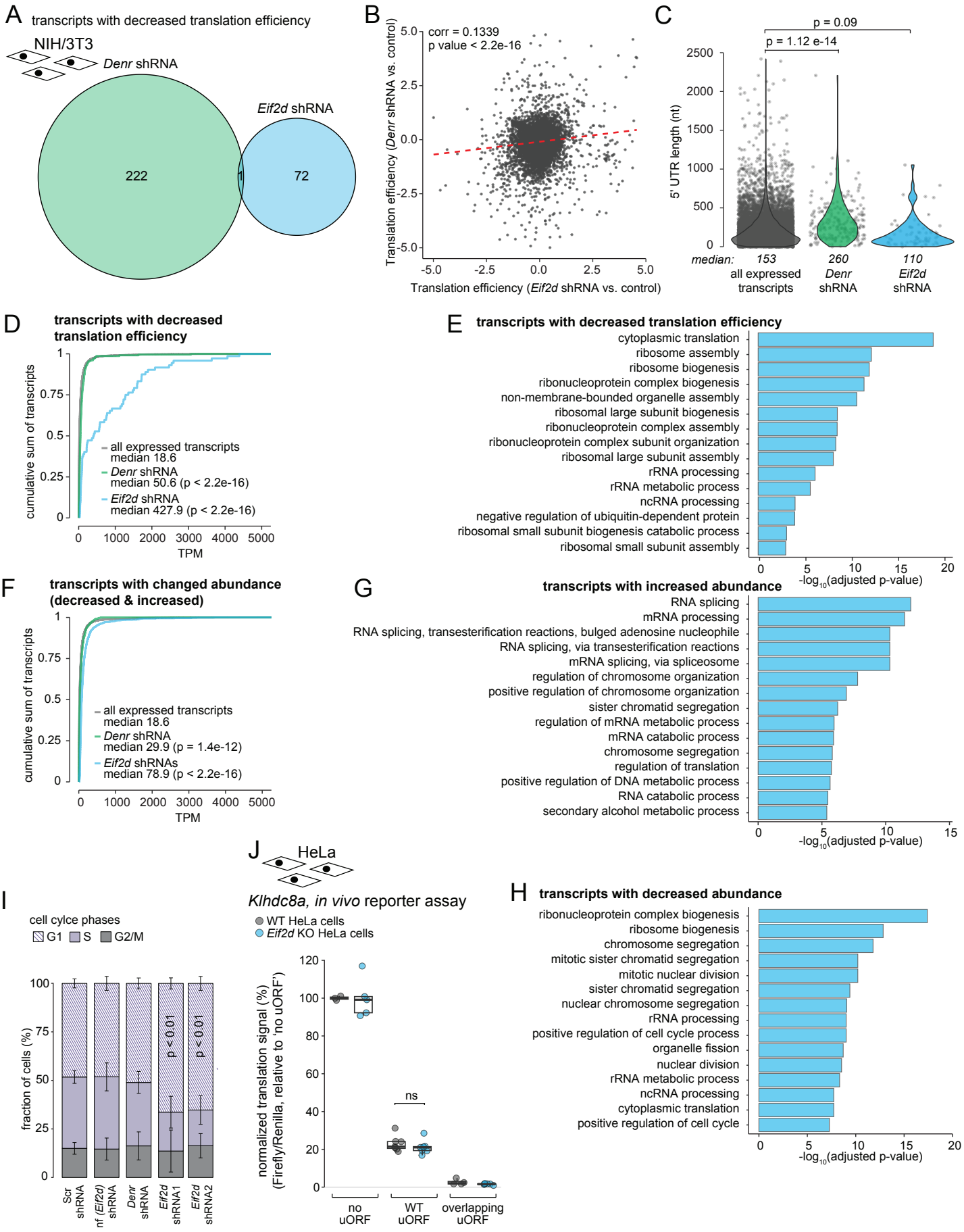
Meurs et al., Supplementary Figure S2



Meurs et al., Supplementary Figure S3



Meurs et al., Supplementary Figure S4

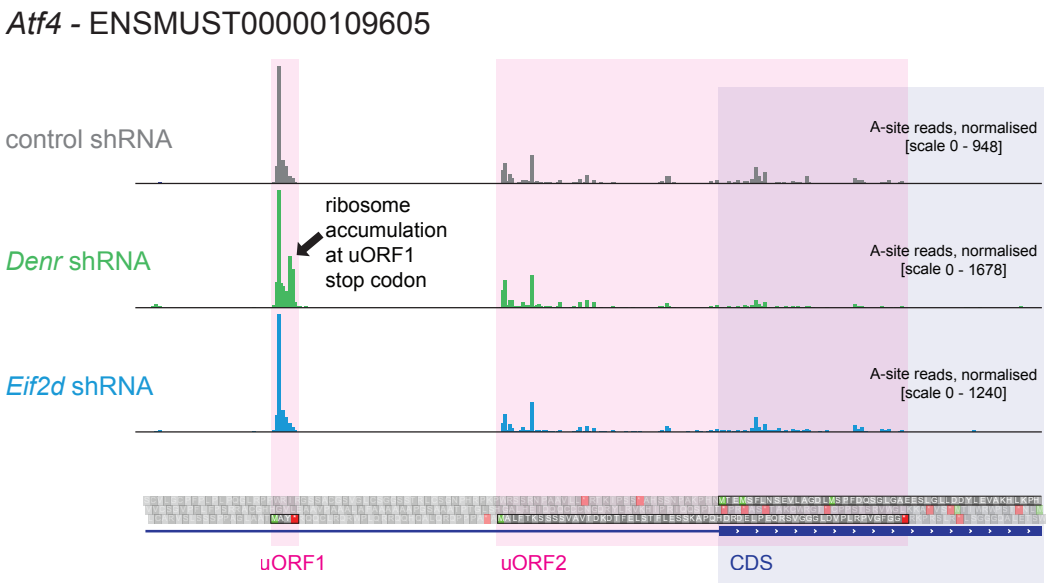


A Alignment of mouse *Mcts1* and *Mcts2* coding sequences:

```
MmMcts1_CDS      ATGTTCAAGAAATTTGATGAAAAAGAAAATGTGTCCAACTGCATCCAGTTGAAAACCTCG        60
MmMcts2_CDS      ATGTTCAAGAAATTTGACGAGAAGGAAAGTGTGTCCAACTGCATCCAACTGAAAACTTCC        60
                 ***************** ** ** **** ******************  ******* **

MmMcts1_CDS      GTTATTAAGGGTATTAAAAATCAATTGCTAGAGCAATTTCCAGGTATTGAACCATGGCTT       120
MmMcts2_CDS      GTTATTAAGGGTATTAAGAGCCAACTGACTGAGCAGTTTCCAGGTATCGAGCCGTGGCTT       120
                 ***************** *  *** **   ***** *********** ** ** ******

MmMcts1_CDS      AATCAAATCATGCCTAAGAAAGACCCTGTGAAAATTGTCCGATGCCATGAACACATAGAA       180
MmMcts2_CDS      AATCAAATCATGCCTAAGAAAGATCCCGTCAAAATAGTGAGATGCCATGAACACATGGAA       180
                 ********************** ** ** ***** **  ***************** ***

MmMcts1_CDS      ATCCTTACAGTAAATGGAGAATTACTGTTTTTTAGACAAAGAGAAGGGCCTTTTTATCCA       240
MmMcts2_CDS      ATCCTTACAGTCAACGGAGAATTACTGTTTTTCAGGCAGAGAAAAGGACCTTTTTATCCA       240
                 *********** ** ***************** ** ** *** **** ************

MmMcts1_CDS      ACTTTAAGATTACTTCATAAATATCCTTTTATCTTGCCACATCAGCAGGTTGATAAAGGA       300
MmMcts2_CDS      ACGCTAAGACTACTTCACAAATACCCGTTTATCCTGCCACACCAGCAGGTCGACAAAGGA       300
                 **  ***** ******* ***** ** ****** ******* ******** ** ******

MmMcts1_CDS      GCCATCAAATTTGTACTCAGTGGAGCAAATATCATGTGTCCTGGCTTAACTTCTCCCGGA       360
MmMcts2_CDS      GCCATCAAATTTGTGCTCAGTGGTGCAAATATCATGTGCCCGGGTTTAACGTCTCCTGGA       360
                 ************** ******** ************** ** ** ***** ***** ***

MmMcts1_CDS      GCTAAGCTTTATCCTGCTGCAGTAGATACTATTGTTGCAATCATGGCAGAAGGAAAACAA       420
MmMcts2_CDS      GCGAAGCTCTACACTGCTGCAGTAGATACCATCGTGGCGGTCATGGCAGAGGGGAAAGAG       420
                 ** ***** **  ***************** ** ** **  ********** ** *** *

MmMcts1_CDS      CATGCTTTATGTGTGGGTGTCATGAAGATGTCTGCAGAAGATATTGAGAAAGTAAACAAA       480
MmMcts2_CDS      CATGCCCTGTGTGTCGGAGTCATGAAGATGGCTGCAGCAGACATTGAGAAAATCAACAAG       480
                 *****  * ***** ** ************ ****** *** ********* * *****

MmMcts1_CDS      GGAATTGGCATTGAAAATATCCATTATCTAAATGATGGTCTGTGGCATATGAAGACATAT       540
MmMcts2_CDS      GGGATCGGCATTGAGAATATCCATTATCTAAATGACGGGCTGTGGCACATGAAGACATAT       540
                 ** ** ******** ******************** ** ******** ************

MmMcts1_CDS      AAATGA      546
MmMcts2_CDS      AAGTGA      546
                 ** ***
```

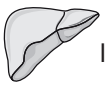
Meurs et al., Supplementary Figure S6

A



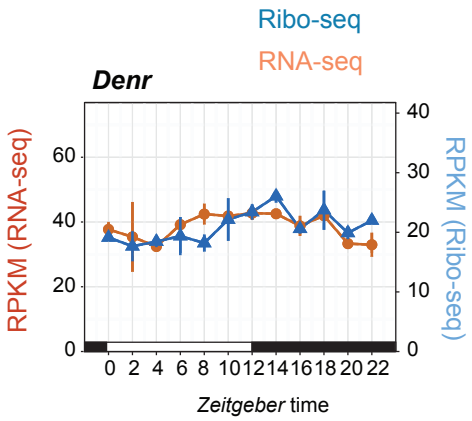
Meurs et al., Supplementary Figure S7

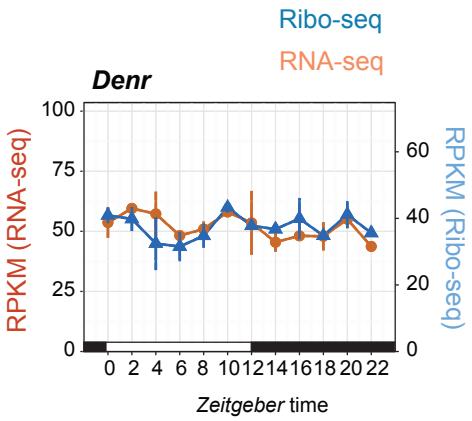
A

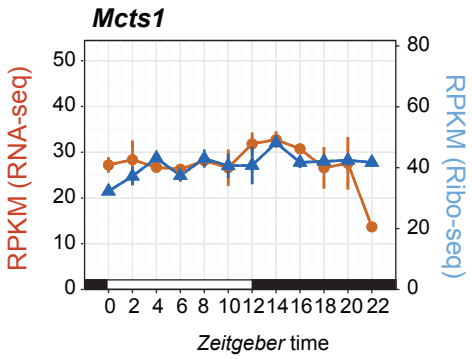


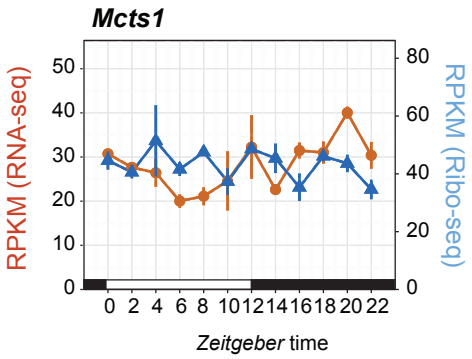
liver - Janich et al., 2015

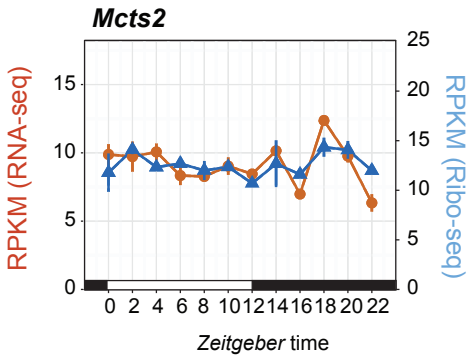
kidney - Castelo-Szekely et al., 2017

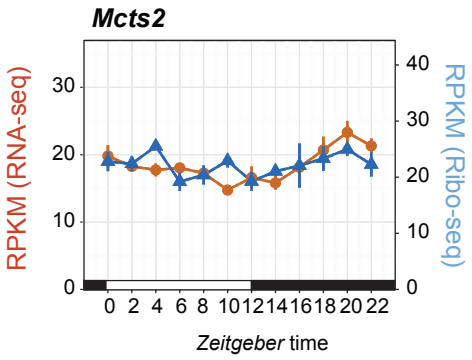












Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome.

Pramod R. Bhatt, Alain Scaiola, Gary Loughran, Marc Leibundgut, Annika Kratzel, **Romane Meurs**, René Dreos, Kate M. O'Connor, Angus McMillan, Jeffrey W. Bode, Volker Thiel, David Gatfield, John F. Atkins, Nenad Ban.

Science (2021)

Contribution to the study

The study was mostly designed and performed by Pramod R. Bhatt and other members of the Nenad Ban and John Atkins labs. I carried out the ribosome profiling experiment from rabbit reticulocyte lysate (figure 3) and participated in the analysis of the results together with René Dreos and David Gatfield. I also contributed to the proofreading of the manuscript.

RESEARCH ARTICLE

CORONAVIRUS

Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome

Pramod R. Bhatt[1,2,3]†, Alain Scaiola[1]†, Gary Loughran[2], Marc Leibundgut[1], Annika Kratzel[4,5,6], Romane Meurs[7], René Dreos[7], Kate M. O'Connor[2], Angus McMillan[8], Jeffrey W. Bode[8], Volker Thiel[4,5], David Gatfield[7], John F. Atkins[2,3,9]*, Nenad Ban[1]*

Programmed ribosomal frameshifting is a key event during translation of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA genome that allows synthesis of the viral RNA-dependent RNA polymerase and downstream proteins. Here, we present the cryo–electron microscopy structure of a translating mammalian ribosome primed for frameshifting on the viral RNA. The viral RNA adopts a pseudoknot structure that lodges at the entry to the ribosomal messenger RNA (mRNA) channel to generate tension in the mRNA and promote frameshifting, whereas the nascent viral polyprotein forms distinct interactions with the ribosomal tunnel. Biochemical experiments validate the structural observations and reveal mechanistic and regulatory features that influence frameshifting efficiency. Finally, we compare compounds previously shown to reduce frameshifting with respect to their ability to inhibit SARS-CoV-2 replication, establishing coronavirus frameshifting as a target for antiviral intervention.

Ribosomal frameshifting, a process during which the reading frame of translation is changed at the junction between open reading frames (ORFs) 1a and 1b, is one of the key events during translation of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) positive-sense single-stranded RNA genome. This programmed –1 translational frameshifting is conserved in all coronaviruses and is necessary for the synthesis of viral RNA-dependent RNA polymerase (RdRp or Nsp12) and downstream viral nonstructural proteins that encode core enzymatic functions involved in capping of viral RNA, RNA modification and processing, and RNA proofreading (*1*). Although the translational machinery typically prevents frameshifting as a potential source of one of the most disruptive errors in translation (*2*, *3*), many viruses rely on programmed ribosomal frameshifting to expand and fine-tune the repertoire and stoichiometry of expressed proteins (*4*).

[1]Department of Biology, Institute of Molecular Biology and Biophysics, ETH Zurich, Zurich, Switzerland. [2]School of Biochemistry and Cell Biology, University College Cork, Cork T12 XF62, Ireland. [3]School of Microbiology, University College Cork, Cork T12 K8AF, Ireland. [4]Institute of Virology and Immunology, University of Bern, Bern, Switzerland. [5]Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland. [6]Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland. [7]Center for Integrative Genomics, Génopode, University of Lausanne, 1015 Lausanne, Switzerland. [8]Laboratorium für Organische Chemie, Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, Switzerland. [9]MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK.
***Corresponding author. Email: j.atkins@ucc.ie (J.F.A.); ban@mol.biol.ethz.ch (N.B.)**
†These authors contributed equally to this work.

Programmed –1 frameshifting in SARS-related coronaviruses occurs at the slippery sequence U_UUA_AAC in the context of a 3′ stimulatory RNA sequence that was predicted to form a three-stemmed pseudoknot structure (*5*) and, in parallel, was independently tested by our lab and others (*6*–*8*). The frameshifting occurs with high efficiency (25 to 75%), depending on the system used (*6*, *7*, *9*–*11*), and changes the reading frame to UUU_AAA_C (*12*) (Fig. 1A). Consequently, two viral polyproteins are synthesized: one encoded by ORF1a when frameshifting does not take place, and ORF1ab, which is expressed as a result of frameshifting. Translation of ORF1a produces polyprotein 1a, which ends with Nsp10 followed by the short Nsp11. Conversely, when the frameshift occurs, the polyprotein 1ab is generated, which contains almost 2700 additional amino acids and in which the viral RdRp, Nsp12, is produced after Nsp10 as a consequence of translation in the –1 frame. A putative secondary structure element in the viral RNA that forms a loop upstream of the shift site has been proposed to play an attenuating role in frameshifting and is referred to as the 5′ attenuator loop (*8*). Maintaining the precise level of coronavirus frameshifting efficiency is crucial for viral infectivity, as evidenced by the fact that mutation of a single nucleotide in the frameshifting region of the SARS-CoV-1 RNA results in a concomitant abrogation of viral replication (*13*). Therefore, the importance of three-stemmed pseudoknot-dependent –1 ribosomal frameshifting for the propagation of SARS-related coronaviruses, a process that has not been seen to occur on any endogenous human transcript in human cells, presents itself as an opportune drug target with minimal tolerance for drug-resistant mutations.

Because of its importance in the life cycle of many important viruses and coronaviruses in particular, programmed frameshifting has been extensively studied using a range of structural and functional approaches (*4*). The structure of a 3′ stimulatory pseudoknot in isolation or in context of the viral genome has been proposed recently by various groups using techniques that include molecular dynamics, nuclease mapping, in vivo selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE), nuclear magnetic resonance (NMR), and cryo–electron microscopy (cryo-EM) (*7*, *14*–*17*). Furthermore, a ribosomal complex with a frameshift stimulatory pseudoknot from the avian infectious bronchitis virus was reported at low resolution (*18*). Here, to provide a structural and mechanistic description of the events during ribosomal frameshifting, we investigated mammalian ribosomes captured in distinct functional states during translation of a region of SARS-CoV-2 genomic RNA where –1 programmed frameshifting occurs.

Structure determination of a frameshifting-primed ribosomal complex

We captured a 0 frame, preframeshift ribosomal complex by introducing a stop codon in place of the second codon of the slippery site (U_UUA_AAC to U_UUA_UAA) (Fig. 1A) and adding mutant eukaryotic release factor 1 [eRF1 (AAQ)] that is unable to release the nascent polypeptide. Translating complexes were prepared in an in vitro translation reaction using an in-house–generated rabbit reticulocyte lysate (RRL) system that supported efficient frameshifting in the previously reported range of around 50% (*19*) according to dual luciferase experiments (see methods). The ribosomes were programmed with mRNA encoding an affinity tag and harboring a region of the SARS-CoV-2 genome that encodes proteins Nsp10 (C terminus), Nsp11, and most of Nsp12. Western blotting showed that when using the wild-type (WT) RNA template, frameshifting was efficient, whereas the stop codon mutation prevented frameshifting and led to ribosome pausing. This effect was further enhanced when eRF1 (AAQ) was present in excess over endogenous WT eRF1 (Fig. 1B).

The cryo-EM three-dimensional (3D) reconstruction of ribosome–nascent chain complexes affinity-purified from the reactions supplemented with eRF1 (AAQ) revealed two distinct ribosomal complexes captured in the process of translating the slippery sequence (figs. S1 and S2). One represented a termination complex that contained the ATP-binding cassette

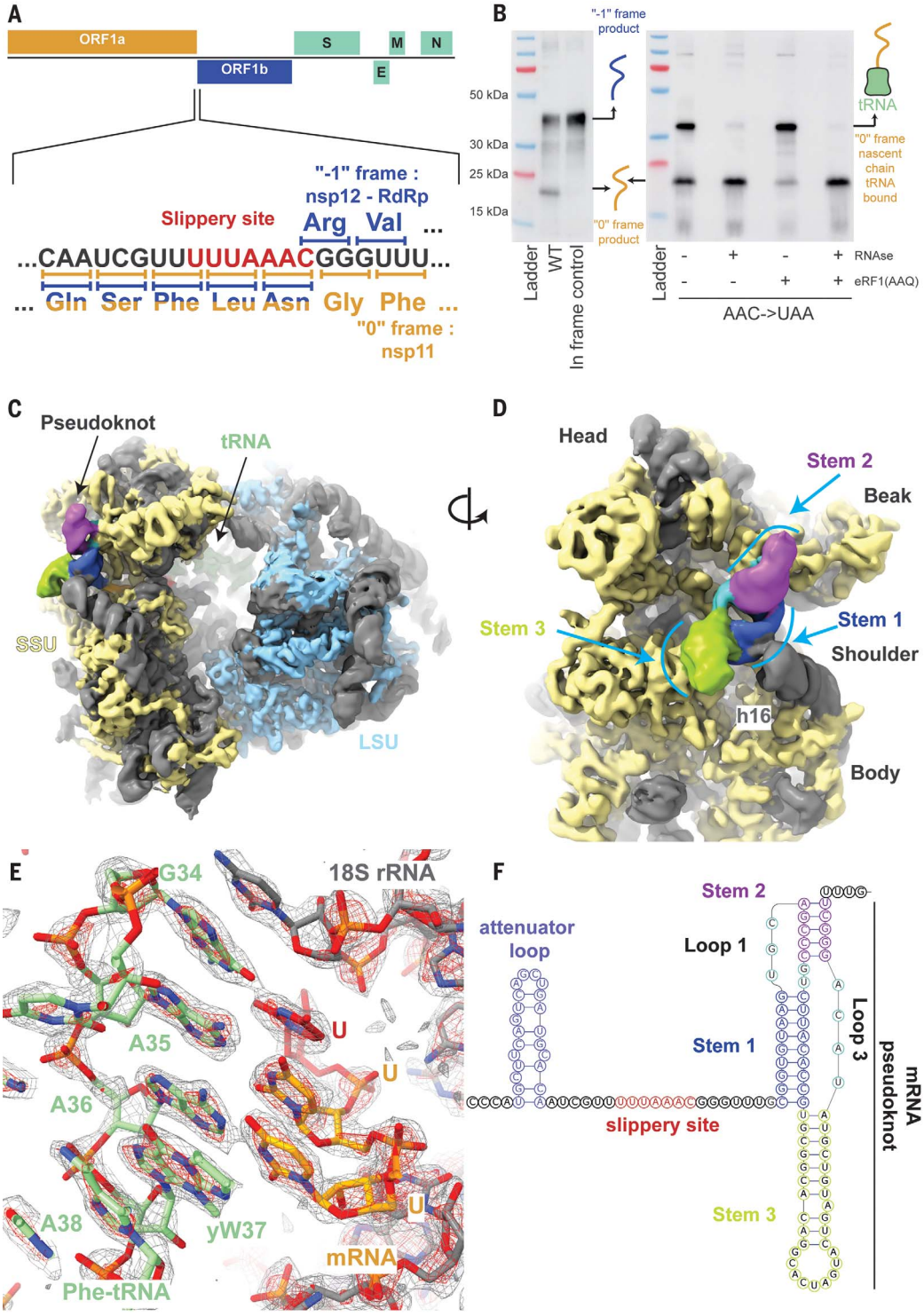


Fig. 1. The SARS-CoV-2 pseudoknot interacts with the ribosome and pauses translation upstream of the slippery site. (**A**) Schematic of the SARS-CoV-2 main ORF. In the close-up view of the frameshift event, codons and corresponding amino acids are shown. During −1 frameshifting, the slippery site codons UUA (Leu) and AAC (Asn) are the last codons decoded in the 0 frame. Upon −1 frameshifting of the AAC codon to AAA, translation resumes at the CGG (Arg) triplet, where elongation proceeds uninterrupted to produce full-length Nsp12. (**B**) In vitro translation reaction depicting pausing at the frameshift site, as shown with Western blotting. Efficient frameshifting is observed for the WT template, consistent with our dual luciferase assays (see methods). Samples for cryo-EM originally intended to be trapped by dominant negative eRF1 (AAQ) show a tRNA-bound pause in proximity of the frameshift site. The tRNA-associated band is lost upon RNase treatment. Reactions without added eRF1 (AAQ) produce a similarly paused product. (**C**) Overview of the density low-pass filtered to 6 Å with the pseudoknot found close to the entry of the mRNA channel on the small subunit (SSU). The SSU proteins are colored in yellow, the large subunit (LSU) proteins in blue, and the rRNA in gray. The pseudoknot is colored according to its secondary structure as in (F), and the P-site tRNA is colored in green. (**D**) Close-up view of the pseudoknot from the solvent-exposed side of the SSU. Helix h16 of the 18S rRNA interacts with the base of Stem 1. Unpaired loop-forming nucleotides are colored in cyan. (**E**) P-site codon-anticodon interactions reveal a Phe (UUU) codon interacting with tRNA(Phe). yW37, wybutosine at position 37. (**F**) Schematic of the revised secondary structure elements in the pseudoknot necessary for −1 programmed ribosomal frameshifting, with different functional regions labeled and colored accordingly.

transporter 1 (ABCE1), which is known to be involved in termination and recycling together with mutant eRF1 interacting with the stop codon (fig. S3). The second reconstruction resolved translating 80*S* ribosomes containing bound P- and E-site tRNAs (fig. S2). This reconstruction at 2.2-Å resolution allowed us to build the most accurate structure of a mammalian 80*S* ribosome so far and directly visualize many protein and virtually all rRNA modifications identified for the human ribosome based on quantitative mass spectrometry and as interpreted in a recent human ribosome structure (*20*, *21*), consistent with the complete conservation of all modified residues between rabbit and human ribosomal RNAs (rRNAs) (figs. S4 and S5; and tables S1 to S3). Importantly, this reconstruction also featured additional density at the entrance to the mRNA channel suggestive of a structured RNA, which, after focused classification, revealed a prominent density for

a complete 3′ frameshifting stimulatory pseudoknot at the entry of the mRNA channel on the 40*S* subunit (Fig. 1, C and D). The resolution of this reconstruction ranged from 2.4 Å at the core of the ribosome to ~7 Å at the periphery, where the most flexible regions of the pseudoknot are located (figs. S2 and S6). Based on the high-resolution maps that allowed visualization of the codon-anticodon interactions and modifications in the tRNA (Fig. 1E and fig. S6, A and B), we could unequivocally determine that a Phe-tRNA(Phe) was bound at

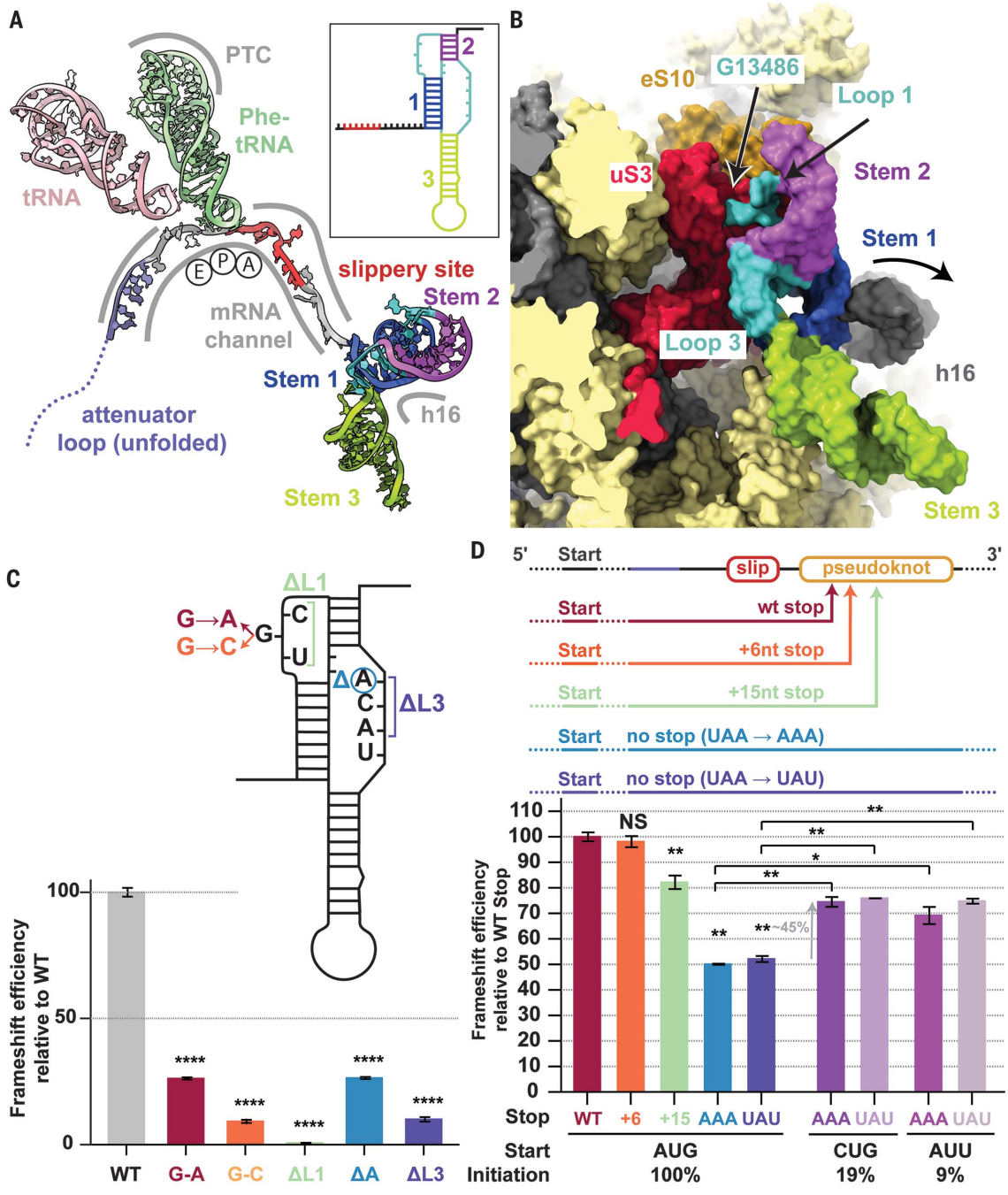


Fig. 2. Critical features of the ribosome-bound pseudoknot. (**A**) Overview of the frameshift-primed state. The stimulatory pseudoknot pauses the ribosome at the penultimate codon (UUU) of the slippery site (red), with P-sites (green) and E-sites (pink) occupied by tRNAs and an empty A-site awaiting decoding in the nonrotated state. The length of the spacer region (gray) is critical for exact positioning of the pseudoknot as the spacer exerts tension at the entry of the mRNA channel (fig. S6C). The inset shows a secondary structure depiction of the frameshift-stimulating pseudoknot colored accordingly. PTC, peptidyl transferase center. (**B**) The backbone of Loop 1 (UGC) (cyan) of the pseudoknot interacts with the N-terminal domain of uS3 (red) and the C-terminal tail of eS10 (orange). mRNA residue G13486 is flipped out and interacts with uS3 (fig. S6D). (**C**) Mutagenesis experiments using dual luciferase assays in HEK293T cells indicate that the G13486 interaction is specific. Mutation of G13486 to other residues leads to a marked reduction in frameshifting efficiency, and deletion of Loop 1 (ΔL1) completely abolishes frameshifting. Similarly, deletion of a single nucleotide (A13537) in Loop 2 reduces frameshifting, whereas deletion of the entire loop (ΔL2) abolishes frameshifting. Normalized (Firefly-Renilla) luciferase activities were calculated for each construct as a percentage of their individual normalized in-frame controls. Data are presented as mean values ± standard deviations of three biological replicates (sets of translation reactions) averaged after three measurements, with error bars representing standard deviations. ****P < 0.0001 by Student's two-tailed t test. (**D**) Mutagenesis experiments using dual luciferase reporter assays in HEK293T cells show that the position of the 0 frame stop codon influences frameshifting. When leaving the pseudoknot unaltered, an incremental increase in the distance of the 0 frame stop codon from the frameshift site leads to a concomitant decrease in frameshifting levels. Loss of the stop codon in the 0 frame leads to a sharp decline in frameshifting levels. This reduction is rescued by ~45% upon decreasing ribosome loading levels by implementing weaker initiation codons. The graph is normalized relative to the WT frameshifting of 25%. Mutations and complementary mutations are shown in fig. S8. Error bars represent standard deviation. NS, not significant; *P < 0.1; and **P < 0.01 by Student's two-tailed t test.

the P-site (*22*). The mRNA does not adopt any unusual structure in the A-site of the ribosome as was observed for the HIV-1 frameshifting sequence visualized on the bacterial ribosome (*23*). This implied that the ribosome is paused by the downstream pseudoknot located at the entrance to the mRNA channel such that the P-site tRNA interacts with the UUU codon just prior to the first codon, UUA, of the slippery site (Fig. 2A).

The pseudoknot causes ribosomal pausing prior to −1 frameshifting

The observation that the pseudoknot acts as an obstacle to slow down translation as the ribosome approaches the slippery site is mechanistically reasonable. Because the pseudoknot is a stable structural element in the mRNA, it will resist unfolding and consequently generate a back-pull on the viral RNA, resulting in an increased chance of −1 frameshifting as the tRNAs are translocated. A pause in translocation at a codon that precedes the slippery site, characterized by a >10 times longer occupancy prior to the slippage event, was observed in an analogous case of heptanucleotide −1 frameshifting on the bacterial *dnaX* gene using single-molecule experiments (*24*). According to this model, it would be anticipated that a further round of translocation results in unwinding of Stem 1 of the downstream stimulatory

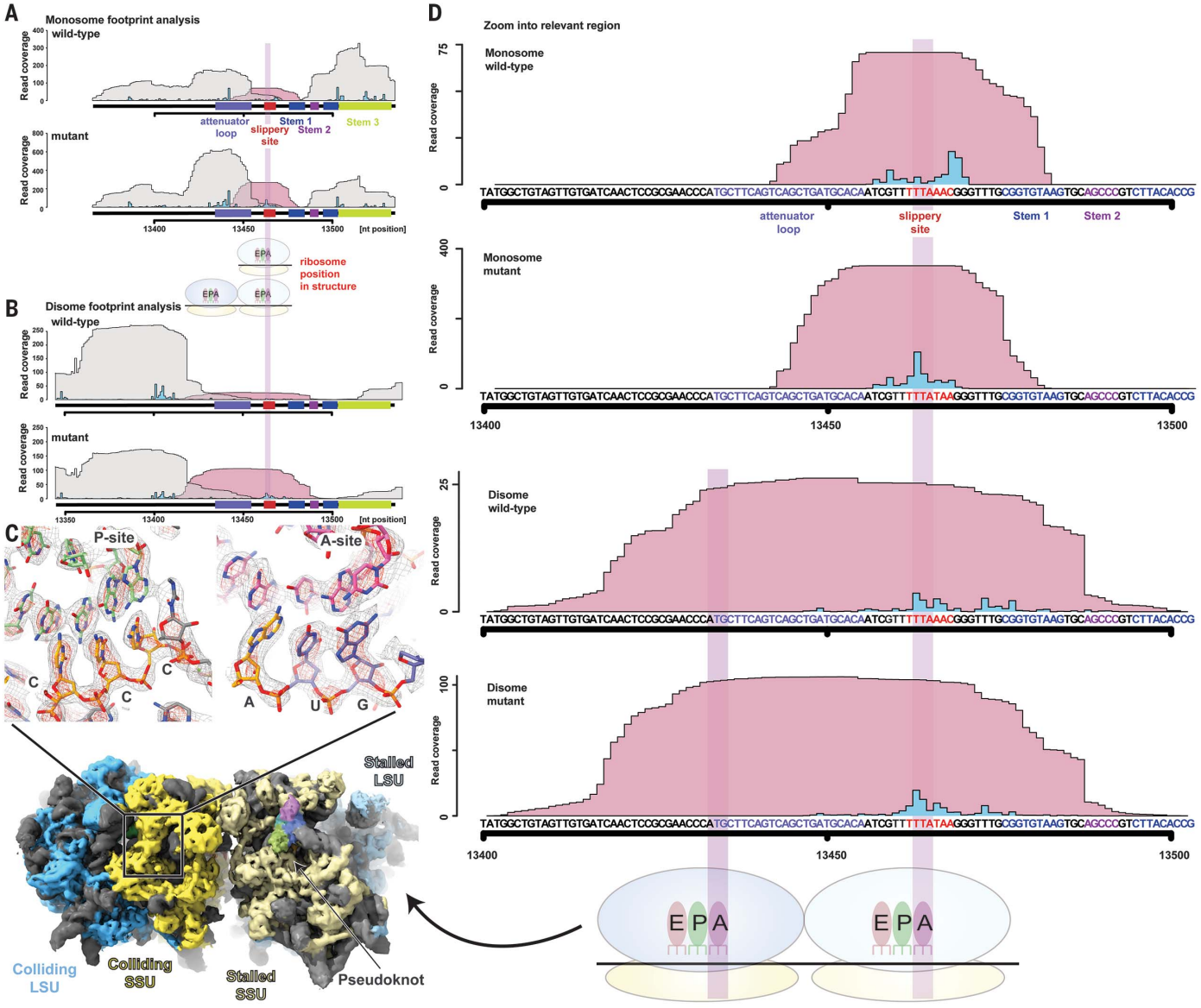


Fig. 3. Pseudoknot-mediated pause occurs prior to the −1 frameshifting event. (**A** and **B**) Footprint coverage for WT and mutant constructs determined by monosome-selective (A) and disome-selective (B) ribosome profiling. Pileup of reads from the indicated areas are plotted separately for reads that overlap (pink) or do not overlap (gray) the frameshift site (indicated by red bar below the *x* axis). The predicted A-sites of the ribosomes that give rise to the footprints are depicted as blue peaks. A-site predictions were carried out as described in the supplementary materials. (**C**) In high-resolution cryo-EM reconstructions of disomes at the frameshift site, the P- and A-sites of the trailing ribosome show occupancy of CCC and AUG codons, respectively, corresponding to the positions estimated by disome profiling. Disome maps were calculated by separately refining the orientational parameters for each ribosome. (**D**) Magnification of the frameshift region from (A) and (B) reveals that monosome profiles show transient occupancy in the vicinity of the frameshift site, whereas disome profiles, which are indicative of strong pause sites, show a similarly enhanced occupancy at the first codon (UUA) of the frameshift site in both WT and mutant constructs. A-site codons of the leading and trailing ribosome are highlighted with a translucent bar and correspond to those seen in the disome structure in (C).

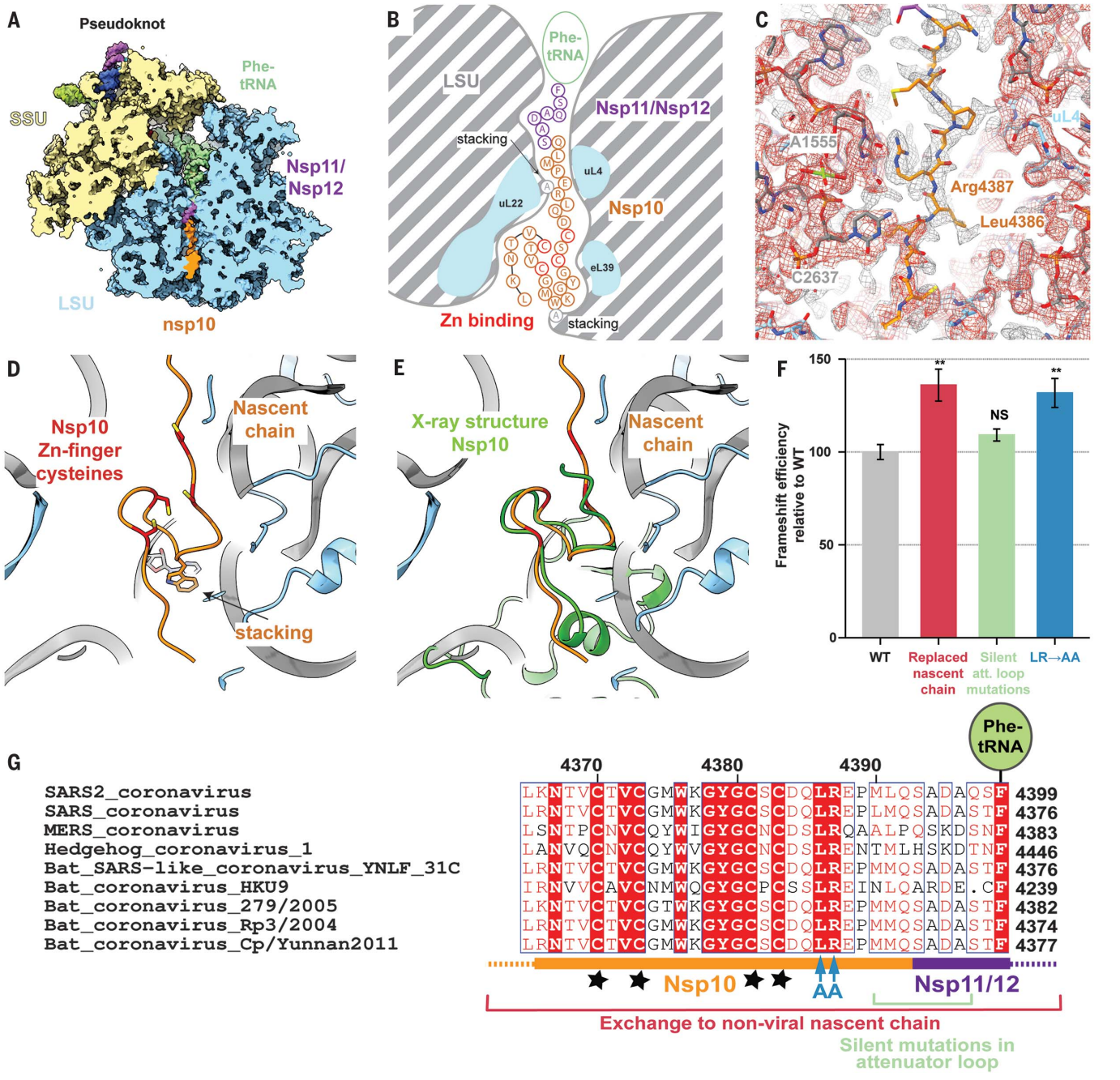


Fig. 4. The nascent viral polypeptide cotranslationally folds and specifically interacts with the ribosomal tunnel. (**A**) Cross-section of the pseudoknot-paused ribosome structure showing the exit tunnel. The nascent C terminus of Nsp10 (orange) and the N terminus of Nsp11/Nsp12 (purple) are visible from the PTC to the periphery of the ribosome exit tunnel (LSU in blue). (**B**) Schematic representation of the path of nascent peptide along the exit tunnel. Arg^{4387} stacks with 28*S* rRNA residue A^{1555} at the constriction site. Further down, where the tunnel widens, the C-terminal zinc finger domain of Nsp10 folds cotranslationally, with Trp^{4376} stacking on A^{2261} of 28*S* rRNA. (**C**) A well-ordered density is visible for Arg^{4387} of Nsp10 as it stacks onto A^{1555} of 28*S* rRNA at the constriction site and is stabilized by Leu^{4386}. The structure is shown within the cryo-EM map contoured at two different levels (gray and red). (**D** and **E**) The overlay of the cotranslationally folded zinc finger domain with the crystal structure of Nsp10 [green, PDB 2FYG (*37*)] reveals the structural similarity. (**F**) Probing the role of nascent chain interactions with the ribosome exit tunnel using an RRL in vitro system. Mutations of the interacting residues were tested for their effect on frameshifting shown in comparison to the WT frameshifting (41% frameshifting was normalized to 100%). Replacement of the entire nascent chain with an unrelated sequence leads to a 35% relative increase in frameshifting, which is only in part due to the loss of the 5′ attenuator loop. Interactions around the constriction site likely serve to attenuate frameshifting, because replacement of the interacting Arg^{4387} and stabilizing Leu^{4386} (LR) with Ala (AA) increases frameshifting by 30%. Error bars represent standard deviation. NS, not significant, and $^{**}P < 0.01$ by Student's two-tailed *t* test. (**G**) Alignment of SARS2 with closely related sequences of other coronaviruses highlighting the conservation of the mutated residues [colored as in (F)]. The shown sequence stretch encompasses the C-terminal zinc finger domain of Nsp10 (orange) and parts of Nsp11/Nsp12 (purple) visible in our reconstruction. Nascent-chain residues Leu^{4386} and Arg^{4387} that interact with the ribosomal exit tunnel are strictly conserved, whereas the conservation of neighboring residues is lower. Stars represent the four cysteines of the Nsp10 zinc finger. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

pseudoknot structure. Consistently, in our structure of the eRF1 (AAQ)–bound ribosome that advanced one codon further along the mRNA, no clear secondary structure is visible at the entrance to the mRNA channel because the mRNA now becomes disordered at this position (figs. S1 and S3, A and B).

To investigate the slowdown of translation on the WT slippery sequence, we performed disome footprint profiling, a method that identifies translational pause sites through the analysis of transitory ribosome collisions (*25*–*27*) (see methods). Notably, recent studies using conventional ribosome profiling methodology reported a lack in monosome footprint coverage across the frameshifting region on the SARS-CoV-2 RNA (*11*, *28*), possibly because ribosomes in this area became trapped in temporary collisions. Moreover, the highly structured pseudoknot at the entry to the mRNA channel would likely preclude efficient trimming by ribonuclease I (RNase I), the enzyme used for footprint generation, further reducing efficient monosome footprint capture. Using a modified nuclease treatment protocol (see methods) that recovered monosome footprints from the frameshift region (Fig. 3, A and C), our experiments revealed that ribosome collisions occur as a result of ribosomal pausing at the same position that is observed in the structure of the pseudoknot-engaged ribosome (Fig. 3, B and D). Apparently, although the base substitutions creating a stop codon in the 3′ adjacent slippery site did not change the features of pausing, they increased the dwell time of the ribosomes at the pause site sufficiently to allow visualization in the cryo-EM experiment.

The results of our disome profiling experiments prompted us to structurally investigate disomes by cryo-EM. We were able to visualize the pseudoknot-paused ribosome followed by a closely trailing ribosome. Upon focused refinement, we obtained a high-resolution (3.1 Å) structure of the trailing ribosome in a rotated state (fig. S1). In congruence with our estimated positioning of the ribosomes in disome profiling (Fig. 3D), the purine-pyrimidine pattern of codon-anticodon pairs in the structure of the colliding ribosome revealed that the pause occurs with CCC and AUG triplets in the P- and A-sites, respectively (Fig. 3C).

The SARS-CoV-2 RNA pseudoknot specifically interacts with ribosomal proteins and 18*S* rRNA

The intermediate local resolution (5 to 7 Å) of the cryo-EM map in the area of the pseudoknot allowed us to visualize the overall fold of the RNA and readjust its previously predicted secondary structure (*14*–*17*, *19*) (Fig. 1, C, D, and F). The stimulatory pseudoknot forms an H-type pseudoknot with Stem 1 and Stem 2 coaxially stacked on top of each other to form a quasi-continuous helix, whereas Stem 3 stands out almost perpendicular to this plane (Figs. 1D and 2B). This corkscrew-like formation provides a bulky and well-structured obstacle wedged at the mRNA entry channel, which has the potential to resist unwinding by the helicase activity of the ribosome and generate tension on the upstream mRNA up to the decoding center. Stem 1 of the pseudoknot forms a 9–base pair helix that is GC rich at the bottom (Fig. 1F). The penultimate nucleotides of the "spacer region" before Stem 1 are located at the mRNA entry tunnel, where they interact with several basic residues in the C-terminal domain of uS3 on one side and are supported by uS5 from the other, with an additional weak contact contributed by the C-terminal end of eS30. uS3 and eS30 are primary components of the ribosome helicase, and uS5 has been proposed to be a component of the ribosomal helicase processivity clamp at the mRNA entry site (*29*, *30*). The observed distance between the P-site UUU codon and Stem 1 of the pseudoknot underscores the critical dependence of the frameshifting efficiency on the length of the spacer region (*31*). Translocation to the next codon would place the frameshifting codon UUA into the P-site, with a simultaneous increase in the tension of the mRNA and unwinding of the GC-rich base of Stem 1 upon entering the mRNA entry channel, comparable to the situation when the ribosome proceeds to the engineered stop codon, as observed in our eRF1 (AAQ)–stalled structure (fig. S3).

The pseudoknot structure also reveals a hitherto unobserved and possibly unappreciated role for the distal site of the mRNA entrance channel in helicase activity. Although mRNA unwinding studies outside the mRNA entrance channel have so far implicated only a helix in the C-terminal domain of uS3 (*32*), we noticed that Loop 1 of the pseudoknot contacts the N-terminal domain of uS3 as well as the C-terminal tail of eS10 (Fig. 2B and fig. S6D), whereas the flipped-out base G13486 in this loop forms specific interactions (Fig. 2B). Furthermore, because the pseudoknot is located at the entry to the mRNA channel, helix h16 of the 18*S* rRNA is noticeably pushed outward owing to a direct contact with the minor groove of Stem 1 (Fig. 2B and fig. S7A). Because the pseudoknot wedges between the head and the body of the small ribosomal subunit, it would restrict their relative motions that need to take place during translocation. This is consistent with the studies on dynamics of coronavirus frameshifting, which revealed that the mechanism of −1 frameshifting involves restriction of small subunit head motion (*33*).

The structure also reveals another key aspect of the architecture of the pseudoknot as the ribosome encounters it. The start of the pseudoknot is shifted relative to the predicted secondary structure (*14*–*17*, *19*) by two nucleotides. The two opposed nucleotides, which were assumed to base pair with Stem 1, are actually forming the start of Stem 3 by pairing with bases predicted to be in the single-stranded linker 2 (Fig. 1F and fig. S7, B and C). Our cryo-EM density reveals that Loop 3 accommodates a total of four nucleotides, three of which were originally attributed to Stem 2. Thus, we observe that Loop 3 is shifted and expanded relative to the initially predicted secondary structures (*14*–*17*, *19*).

To functionally support our structural findings and confirm the nature and specificity of the pseudoknot interactions, we performed structure-guided mutagenesis experiments using dual luciferase reporter assays in human embryonic kidney (HEK) 293T cells (see methods) and monitored the frameshifting efficiency relative to the WT (Fig. 2C). Mutation of G13486 of Loop 1 to another purine reduced the frameshifting efficiency to 30% of the WT level, and mutation of this base to a pyrimidine further reduced frameshifting to 15%. As expected from our structural data, deletions of the nucleotides of the spacer regions also had a deteriorating effect on frameshifting. Loss of Loop 1 entirely abolished frameshifting. Deletion of a single nucleotide of Loop 3 in agreement with its proposed role in forming the base-pairing interactions diminished the frameshifting rate to 25% of the WT level. Loss of the entire Loop 3 reduced frameshifting to 10% of WT levels.

Frameshifting efficiency depends on the position of the 0 frame stop codon

In SARS-CoV-2, the 0 frame stop codon is located five codons downstream of the frameshift site and is a constituent of Stem 1. The placement of the stop codon in such proximity to the frameshift site is a common feature in coronaviruses, and its presence in a critical region of the stimulatory pseudoknot prompted us to probe the effect of the distance of the 0 frame stop codon on frameshifting. To this end, knowledge of the 3D structure of the pseudoknot helped us to confidently manipulate the stop codon without hampering pseudoknot formation. We introduced mutations to incrementally extend the stop codon from the WT position and to completely remove the occurrence of a stop codon in the 0 frame (Fig. 2D and fig. S8). Whereas introducing a stop codon six nucleotides downstream of the WT position only marginally decreased the frameshifting rate (98% of WT), a stronger attenuation was observed when the distance of the stop codon was increased to 15 nucleotides from the WT stop (80% of WT). Finally, removal of the stop codon by two different point mutations led to a reduction of frameshifting efficiency to 50% of WT levels. To test whether reduced ribosomal loading rescues the effect

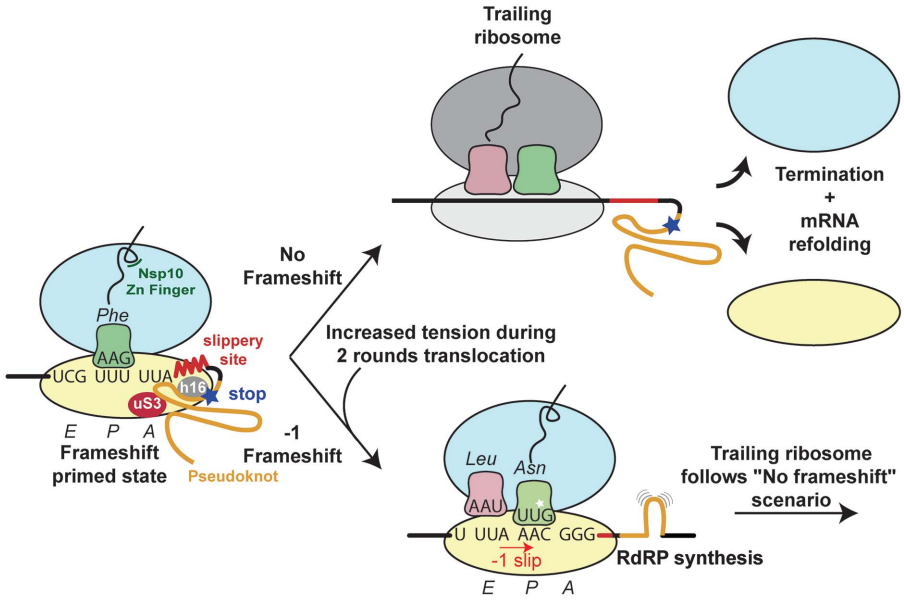


Fig. 5. Structure-based model for −1 programmed frameshifting in coronaviruses and its regulation. The observed interactions between the pseudoknot and the ribosome prime the system for frameshifting. The features of the pseudoknot and the interactions between the nascent chain and the ribosomal tunnel play a role in the efficiency of frameshifting. The efficiency of frameshifting is increased by the presence of a stop codon near the frameshifting site. Ribosomes that progress beyond the frameshifting site in the 0 frame quickly terminate and disassemble, thereby increasing the chances that the pseudoknot will refold before it is encountered by the closely trailing ribosome. The trailing ribosome in turn encounters the pseudoknot, which increases the possibility of undergoing −1 frameshifting.

of stop codon removal, we analyzed the frameshifting efficiency in the context of weaker initiation codons such as CUG and AUU (Fig. 2D). These constructs led to a 45% rescue of the reduction in frameshifting compared with stop codon mutants initiating at an AUG start.

Taken together, these observations suggest that the stop codon position plays an important role in maintaining optimum frameshift efficiency. We propose that the stop codon serves to prevent the closely trailing ribosome from encountering a viral RNA that was unfolded by the leading ribosome. In this case, upon encountering a stop codon, termination and subunit disassembly will occur, which will provide an opportunity for the pseudoknot to refold without the constraints of the mRNA channel (see Conclusions). According to this model, although the WT stop codon will make the frameshifting efficiency less sensitive to ribosome loading in the "no-frameshifting" scenario, the frameshifting events that occur after a −1 frameshift will nevertheless be more likely when the ribosomes are spaced further apart. Our measurements of the efficiency of frameshifting for the WT sequence in the context of different rates of translation initiation are in agreement with this hypothesis (fig. S9). This mechanism, consistent with our biochemical data, increases the efficiency of frameshifting to the levels required by SARS-CoV-2 and may be used by viruses in general when high-efficiency frameshifting is required.

Nascent chain forms specific interactions with the ribosomal exit tunnel

Notably, in the reconstruction of the paused translating ribosome, the nascent chain that corresponds to the viral polyprotein was visible along the entire length of the ribosomal exit tunnel (Fig. 4A). The density corresponds to the C-terminal region of Nsp10, which is the activator of the viral proofreading exonuclease and N7-methyltransferase Nsp14 (*34*, *35*), and then (depending on the frameshifting event) continues as either the viral RNA-dependent RNA polymerase Nsp12 (*6*) or as protein Nsp11, whose function is still unknown (Figs. 1A and 4B). The nascent chain makes several specific interactions with the ribosomal tunnel, one of which is at the constriction site where Arg^{4387} of Nsp10 interacts with A^{1555} of the 28*S* rRNA [corresponding to A^{1600} in humans, numbering according to PDB 6EK0 (*36*)] and is stabilized by the preceding Leu^{4386} (Fig. 4C). Notably, these two amino acids are highly conserved across multiple coronaviruses (Fig. 4G), although they are located in the unstructured C-terminal region of Nsp10 and therefore considered not to be important for the fold of the protein (*37*).

Further down the tunnel, the C-terminal end of Nsp10 adopts a partially folded zinc finger motif (Fig. 4, D and E), which, upon superposition, reveals similarity with the corresponding fully folded C-terminal domain previously observed in the crystal structure of SARS-CoV-1 Nsp10 (*37*). Trp^{4376}, which is located between the two pairs of cysteines that form the zinc finger, stacks with A^{2261} (A^{2418}), an interaction that might serve to promote the change of nascent chain direction and facilitate folding of the zinc finger at the end of the exit tunnel. Cotranslational events, such as insertion of a transmembrane domain at the exit of the ribosomal tunnel, were shown to promote −1 ribosomal frameshifting in alphaviruses (*38*).

To investigate whether the observed contacts between the nascent chain and the ribosomal tunnel are specific and whether these interactions and cotranslational folding of Nsp10 might play a role in modulating the frameshifting process, we used our dual luciferase reporter assay to measure the frameshifting efficiency of WT and mutant nascent chain sequence constructs. Because our measurements in HEK293T cells did not reveal an appreciable change of frameshift efficiency, we carried out the same experiments in vitro using RRL to monitor the effects in a single mRNA setup. Replacement of the entire nascent chain with an unrelated sequence leads to a 35% increase in frameshifting (Fig. 4F). Importantly, this effect was provoked by the change in peptide sequence and not simply by the loss of the 5′ attenuator loop, given that a reporter containing silent attenuator loop mutations resulted in only a slight increase in frameshifting (Fig. 4F). Mutation of the Leu^{4386} and Arg^{4387} to alanine led to a considerable (30%) increase in frameshifting (Fig. 4, F and G), implying that these nascent chain interactions with the ribosomal exit tunnel play an important role in regulating frameshifting levels, possibly mechanistically akin to the well-studied SecM stalling system in bacteria (*39*), where it was shown that cotranslational folding and the translocon-induced mechanical force can rescue the stall induced by interactions between the nascent chain and the ribosomal tunnel (*40*). These observations also suggest that any cellular nascent chain factors (*41*, *42*) might influence the rate of frameshifting.

Inhibition of viral replication by a compound that targets the SARS-CoV-2 pseudoknot

The sensitivity of the coronavirus to the finely controlled frameshifting levels (*13*) may present an opportunity to develop compounds that interfere with the frameshifting process and thus inhibit replication of the virus. Using computational modeling and reporter assays, compounds that have been predicted to bind the pseudoknot and inhibit SARS-CoV-2 frameshifting were described (*19*, *43*) but never tested with respect to their ability to inhibit viral replication. Furthermore, the fluoroquinolone compound merafloxacin was recently reported to also inhibit −1 frameshifting efficiency of SARS-CoV-2 and other betacoronaviruses (*44*). To demonstrate that the inhibition of frameshifting is a plausible strategy for drug development, we compared two of the previously described compounds with respect to their ability to reduce viral levels in infected African green monkey VeroE6 cells (fig. S10 and methods). Our experiments demonstrate that merafloxacin is a better candidate compound because it showed a concentration-dependent inhibition of frameshifting, whereas, contrary to earlier reports (*19*, *43*), the small-molecule ligand MTDB did not specifically inhibit frameshifting under our experimental conditions (fig. S10). The two compounds showed no cellular toxicity and resulted in a three to four orders of magnitude reduction of SARS-CoV-2 titer, with a half-maximal inhibitory concentration (IC_{50}) of 48 μM for MTDB and an order of magnitude higher efficacy for merafloxacin, with an IC_{50} of 4.3 μM (fig. S10). Because MTDB did not appear to affect frameshifting in our reporter construct experiments in vitro and in vivo, it is possible that it inhibits SARS-CoV-2 replication by a different mechanism. Although the potency range for these compounds is not what would be expected from potential drug candidates, it nevertheless provides a starting point for high-throughput screening and establishes that frameshifting is a viable target for therapeutic intervention against SARS-CoV-2.

Conclusions

Our results provide a mechanistic description of frameshifting that occurs during translation of the SARS-CoV-2 genome and reveal the features that may be exploited by the virus to finely control the stoichiometry of viral proteins at different stages of infection (Fig. 5). Interfering with the frameshifting process at the level of nascent chain interactions with the ribosomal tunnel, at the level of RNA folding that leads to the formation of the frameshift stimulatory pseudoknot, or to perturb the interactions between the pseudoknot and the mRNA channel represent viable strategies in our search for new drugs against SARS-CoV-2, the virus that is currently causing the global COVID-19 pandemic. Our results will also be useful for understanding the mechanism of programmed ribosomal −1 frameshifting (*4*), including that used by many other medically important viruses.

REFERENCES AND NOTES

1. P. V'kovski, A. Kratzel, S. Steiner, H. Stalder, V. Thiel, *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
2. J. Parker, *Microbiol. Rev.* **53**, 273–298 (1989).
3. J. M. Ogle, A. P. Carter, V. Ramakrishnan, *Trends Biochem. Sci.* **28**, 259–266 (2003).
4. J. F. Atkins, G. Loughran, P. R. Bhatt, A. E. Firth, P. V. Baranov, *Nucleic Acids Res.* **44**, 7007–7078 (2016).
5. F. Dos Ramos, M. Carrasco, T. Doyle, I. Brierley, *Biochem. Soc. Trans.* **32**, 1081–1083 (2004).
6. P. V. Baranov *et al.*, *Virology* **332**, 498–510 (2005).
7. E. P. Plant *et al.*, *PLOS Biol.* **3**, e172 (2005).
8. M. C. Su, C. T. Chang, C. H. Chu, C. H. Tsai, K. Y. Chang, *Nucleic Acids Res.* **33**, 4265–4275 (2005).
9. I. Brierley, P. Digard, S. C. Inglis, *Cell* **57**, 537–547 (1989).
10. N. Irigoyen *et al.*, *PLOS Pathog.* **12**, e1005473 (2016).
11. Y. Finkel *et al.*, *Nature* **589**, 125–130 (2021).
12. V. Thiel *et al.*, *J. Gen. Virol.* **84**, 2305–2315 (2003).
13. E. P. Plant, R. Rakauskaite, D. R. Taylor, J. D. Dinman, *J. Virol.* **84**, 4330–4340 (2010).
14. R. Rangan *et al.*, *RNA* **26**, 937–959 (2020).
15. K. Zhang *et al.*, Cryo-electron microscopy and exploratory antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. bioRxiv 2020.07.18.209270 [Preprint]. 20 July 2020; https://doi.org/10.1101/2020.07.18.209270.
16. T. C. T. Lan *et al.*, Insights into the secondary structural ensembles of the full SARS-CoV-2 RNA genome in infected cells. bioRxiv 2020.06.29.178343 [Preprint]. 19 February 2021; https://doi.org/10.1101/2020.06.29.178343.
17. N. C. Huston *et al.*, *Mol. Cell* **81**, 584–598.e5 (2021).
18. O. Namy, S. J. Moran, D. I. Stuart, R. J. C. Gilbert, I. Brierley, *Nature* **441**, 244–247 (2006).
19. J. A. Kelly *et al.*, *J. Biol. Chem.* **295**, 10741–10748 (2020).
20. M. Taoka *et al.*, *Nucleic Acids Res.* **46**, 9289–9298 (2018).
21. W. Li, S. T. L. Chang, F. R. Ward, J. H. D. Cate, *Nat. Commun.* **11**, 4941 (2020).
22. G. Keith, G. Dirheimer, The primary structure of rabbit, calf and bovine liver tRNAPhe*Biochim. Biophys. Acta* **517**, 133–149 (1978).
23. C. Bao *et al.*, *eLife* **9**, e55799 (2020).
24. J. Choi, S. O'Loughlin, J. F. Atkins, J. D. Puglisi, *Sci. Adv.* **6**, eaax6969 (2020).
25. P. Han *et al.*, *Cell Rep.* **31**, 107610 (2020).
26. A. B. Arpat *et al.*, *Genome Res.* **30**, 985–999 (2020).
27. S. Meydan, N. R. Guydosh, *Mol. Cell* **79**, 588–602.e6 (2020).
28. M. Puray-Chavez *et al.*, The translational landscape of SARS-CoV-2 and infected cells. bioRxiv 2020.11.03.367516 [Preprint]. 16 November 2020; https://doi.org/10.1101/2020.11.03.367516.
29. J. Rabl, M. Leibundgut, S. F. Ataide, A. Haag, N. Ban, *Science* **331**, 730–736 (2011).
30. S. Takyar, R. P. Hickerson, H. F. Noller, *Cell* **120**, 49–58 (2005).
31. Z. Lin, R. J. C. Gilbert, I. Brierley, *Nucleic Acids Res.* **40**, 8674–8689 (2012).
32. H. Amiri, H. F. Noller, *RNA* **25**, 364–375 (2019).
33. N. Caliskan, V. I. Katunin, R. Belardinelli, F. Peske, M. V. Rodnina, *Cell* **157**, 1619–1631 (2014).
34. M. Bouvet *et al.*, *J. Biol. Chem.* **289**, 25783–25796 (2014).
35. E. C. Smith *et al.*, *J. Virol.* **89**, 6418–6426 (2015).
36. S. K. Natchiar, A. G. Myasnikov, H. Kratzat, I. Hazemann, B. P. Klaholz, *Nature* **551**, 472–477 (2017).
37. J. S. Joseph *et al.*, *J. Virol.* **80**, 7894–7901 (2006).
38. H. R. Harrington *et al.*, *J. Biol. Chem.* **295**, 6798–6808 (2020).
39. H. Nakatogawa, K. Ito, *Cell* **108**, 629–636 (2002).
40. D. H. Goldman *et al.*, *Science* **348**, 457–460 (2015).
41. G. Kramer, D. Boehringer, N. Ban, B. Bukau, *Nat. Struct. Mol. Biol.* **16**, 589–597 (2009).
42. K. Döring *et al.*, *Cell* **170**, 298–311.e20 (2017).
43. S. J. Park, Y. G. Kim, H. J. Park, *J. Am. Chem. Soc.* **133**, 10094–10100 (2011).

ACKNOWLEDGMENTS

We thank A. Jomaa for advice on cryo-EM data processing, N. Casey for support with technical resources, and A. Picenoni for help with the grid preparation. We are indebted to the ETH scientific center for optical and electron microscopy (ScopeM) for access to electron microscopes and, in particular, to M. Peterek and D. Boehringer. We thank J. Guo for helpful discussions and for providing the merafloxacin compound. **Funding:** This work was supported by the Swiss National Science Foundation (SNSF) to N.B. and V.T. (grants 31003A_182341 and 310030_173085) through the National Center of Excellence in RNA and Disease (project funding 182880) to N.B., V.T., and D.G.; by the Federal Ministry of Education and Research, Germany (BMBF; grant RAPID, #01KI1723A) to A.K. and V.T.; and by the Irish Research Council Advanced Laureate (IRCLA/2019/74) to J.F.A., who thanks G. F. Fitzgerald for funds from Carbery Group Ltd. **Author contributions:** P.R.B. and N.B. initiated the project and, together with J.F.A., designed the experiments. P.R.B. carried out biochemical experiments and sample preparation. A.S. and P.R.B. prepared grids. A.S. carried out data collection and calculated EM maps. A.S. and M.L. performed molecular model building and refinement. A.S., M.L., P.R.B., and N.B. interpreted the structure. P.R.B., A.S., M.L., and N.B. drafted the manuscript. A.S. prepared the figures. G.L. and K.M.O. performed dual luciferase reporter assays. A.M. prepared and purified the drug-like ligand MTDB. A.K. and V.T. formulated experiments to test viral infected cells. A.K. carried out testing of ligand on viral-infected cells. R.M., R.D., and D.G. designed, carried out, analyzed, and interpreted ribosome profiling experiments. All authors contributed to the final version of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The structure and cryo-EM map for the high-resolution reconstruction is available in the Protein Data Bank (PDB) as PDB ID 7O7Y and in the Electron Microscopy Data Bank (EMDB) as EMD-12756, respectively. The structure refined into the further classified particle set reconstruction and the corresponding maps are available as PDB-7O7Z and EMD-12757, respectively. The structure of the ribosome bound with eRF1 (AAQ) and ABCE1 reconstruction and the corresponding maps are available as PDB-7O80 and EMD-12758, respectively. The structure of the colliding ribosome reconstruction and the corresponding maps are available as PDB-7O81 and EMD-12759, respectively. The structure of the composite map of the disome is available in the EMDB as EMD-12760. Ribosome profiling data have been deposited at GSE167421.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6548/1306/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S4
References (*44*–*78*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

20 October 2020; resubmitted 24 February 2021
Accepted 7 May 2021
Published online 13 May 2021
10.1126/science.abf3546

Supplementary Materials for

Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome

Pramod R. Bhatt[1,2†], Alain Scaiola[1†], Gary Loughran[2], Marc Leibundgut[1], Annika Kratzel[3], Angus McMillan[4], Kate M. O' Connor[2], Jeffrey W. Bode[4], Volker Thiel[3,5], John F. Atkins[2,6*], Nenad Ban[1*]

[1] Department of Biology, Institute of Molecular Biology and Biophysics, ETH Zurich, Zurich, Switzerland

[2] Schools of Biochemistry and Microbiology, University College Cork, Cork, Ireland, T12 YT57

[3] Institute of Virology and Immunology, University of Bern, Bern, Switzerland

[4] Laboratorium für Organische Chemie, Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, Switzerland.

[5] Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland

[6] Department of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330, USA.

† These authors contributed equally to this work

* Corresponding authors. Email: j.atkins@ucc.ie, ban@mol.biol.ethz.ch

This PDF file includes:

Materials and Methods

Generation of DNA templates and *in vitro* transcription of the mRNA

Constructs used for testing and generating ribosomal complexes *in vitro* contain the following elements: A linear DNA template containing a T7 promoter and Kozak sequence upstream of a sequence encoding an N-terminal 3x FLAG-tag followed by a 290 nt linker, Wild type (WT) or mutant SARS-CoV-2 frameshift context and the majority of *nsp12* was synthesized as a gene fragment by Genscript. The SARS-CoV-2 WT/mutant frameshift context sequences in turn consist of the following elements in sequential order: The final 115 nucleotides of sequence encoding the C-terminal end of Nsp10; the coding sequence of Nsp11/Nsp12 (which includes the so-called 'attenuator loop') up to the Phe 'UUU' codon that just precedes the frameshift site; WT ('UUA AAC') or mutant ('UUA UAA') frameshift sites; the WT frameshift stimulatory pseudoknot sequence; 540 nucleotides of sequence encoding Nsp12 after the frameshift site.

The entire fragment was PCR amplified from opposite ends with complementary primers using Q5 DNA Polymerase (NEB) and purified using the QIAquick PCR Purification kit (Qiagen). *In vitro* transcription was performed using the Ampliscribe T7 Flash system (Lucigen), followed by lithium chloride precipitation to purify the synthesized RNA.

Sequence of synthesized DNA (Gene Universal, USA) used to test WT SARS-CoV-2 frameshifting efficiency in RRL (Frameshift context underlined):

GATATTCTAAGTGACGTTACGCTAGTGATCGCCGTAATACGACTCACTATAGGCAACAACAA
CAAACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGACTAC
AAAGACCACGACGGTGATTATAAAGATCACGACATCGATTACAAGGACGACGACGACAAGT
CCAAGGAGCCGCTTCGGCCACGGTGCCGCCCCATCAACGCCACCCTGGCTGTGGAGAAGGA
GGGCTGCCCCGTGTGCATCACCGTCAACACCACCATCTGTGCCGGCTACTGCCCCACCGCAA
CCCGCGTGCTGCAGGGGGTCCTGCCGGCCCTGCCTCAGGTGGTGTGCAACTACCGCCGGTCC
GTAACCCACCGTATTCTTACCGTTCCGATTGCCCAAGATCAAGTGGGCGCATACTATCAGCA
ACCAGGTCAGCAGAACGCCACCTGGATTGTGCCACCAGGGCCAACTTGTGCTAATGACCCTG
TGGGTTTTACACTTAAAAACACAGTCTGTACCGTCTGCGGTATGTGGAAAGGTTATGGCTGT
AGTTGTGATCAACTCCGCGAACCCATGCTTCAGTCAGCTGATGCACAATCGTTTTTAAACGG
GTTTGCGGTGTAAGTGCAGCCCGTCTTACACCGTGCGGCACAGGCACTAGTACTGATGTCGT
ATACAGGGCTTTTGACATCTACAATGATAAAGTAGCTGGTTTTGCTAAATTCCTAAAAACTA
ATTGTTGTCGCTTCCAAGAAAAGGACGAAGATGACAATTTAATTGATTCTTACTTTGTAGTTA
AGAGACACACTTTCTCTAACTACCAACATGAAGAAACAATTTATAATTTACTTAAGGATTGT
CCAGCTGTTGCTAAACATGACTTCTTTAAGTTTAGAATAGACGGTGACATGGTACCACATAT
ATCACGTCAACGTCTTACTAAATACACAATGGCAGACCTCGTCTATGCTTTAAGGCATTTTG
ATGAAGGTAATTGTGACACATTAAAAGAAATACTTGTCACATACAATTGTTGTGATGATGAT

TATTTCAATAAAAAGGACTGGTATGATTTTGTAGAAAACCCAGATATATTACGCGTATACGC
CAACTTAGGTGAACGTGTACGCCAAGCTTTG

Sequence of synthesized DNA (Gene Universal, USA) used to test 0 frame stop codon mutation of SARS-CoV-2 frameshifting efficiency in RRL (mutant frameshift context underlined):

GATATTCTAAGTGACGTTACGCTAGTGATCGCCGTAATACGACTCACTATAGGCAACAACAACAAACA
TTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGACTACAAAGACCACGAC
GGTGATTATAAAGATCACGACATCGATTACAAGGACGACGACGACAAGTCCAAGGAGCCGCTTCGGC
CACGGTGCCGCCCCATCAACGCCACCCTGGCTGTGGAGAAGGAGGGCTGCCCCGTGTGCATCACCGTC
AACACCACCATCTGTGCCGGCTACTGCCCCACCGCAACCCGCGTGCTGCAGGGGGTCCTGCCGGCCCT
GCCTCAGGTGGTGTGCAACTACCGCCGGTCCGTAACCCACCGTATTCTTACCGTTCCGATTGCCCAAGA
TCAAGTGGGCGCATACTATCAGCAACCAGGTCAGCAGAACGCCACCTGGATTGTGCCACCAGGGCCA
ACTTGTGCTAATGACCCTGTGGGTTTTACACTTAAAAACACAGTCTGTACCGTCTGCGGTATGTGGAAA
GGTTATGGCTGTAGTTGTGATCAACTCCGCGAACCCATGCTTCAGTCAGCTGATGCACAATCGTTTTTA
TAAGGGTTTGCGGTGTAAGTGCAGCCCGTCTTACACCGTGCGGCACAGGCACTAGTACTGATGTCGTA
TACAGGGCTTTTGACATCTACAATGATAAAGTAGCTGGTTTTGCTAAATTCCTAAAAACTAATTGTTGT
CGCTTCCAAGAAAAGGACGAAGATGACAATTTAATTGATTCTTACTTTGTAGTTAAGAGACACACTTT
CTCTAACTACCAACATGAAGAAACAATTTATAATTTACTTAAGGATTGTCCAGCTGTTGCTAAACATGA
CTTCTTTAAGTTTAGAATAGACGGTGACATGGTACCACATATATCACGTCAACGTCTTACTAAATACAC
AATGGCAGACCTCGTCTATGCTTTAAGGCATTTTGATGAAGGTAATTGTGACACATTAAAAGAAATAC
TTGTCACATACAATTGTTGTGATGATGATTATTTCAATAAAAAGGACTGGTATGATTTTGTAGAAAACC
CAGATATATTACGCGTATACGCCAACTTAGGTGAACGTGTACGCCAAGCTTTGTAATAGTCATAGAGG
AT

Purification of eRF1(AAQ)

Mutant eRF1(AAQ) placed in a vector containing an N-terminal 6x His-tag followed by a GST-tag and TEV protease cleavage site was expressed from a pET-24 d(+) vector. The plasmid was transformed into *E. coli* BL21 (DE3) cells under Kanamycin selection, and cells were grown in 2× YT medium at 37 °C. At an optical density at 600 nm (OD600) of 0.8, cultures were shifted to 18 °C and induced with IPTG added to a final concentration of 1 mM. After 16 h, cells were collected by centrifugation, resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 8.0) and lysed using a cell disrupter (Constant Systems). The lysate was cleared by centrifugation for 30 min at 48'000 x g and loaded onto a batch-binding column of Ni-NTA beads (Agarose Bead Technologies). After binding at 4 °C on a nutator, beads were washed with 10 column volumes of lysis buffer containing 45 mM imidazole, and proteins were eluted in 1.5 column volumes of lysis buffer containing 300 mM imidazole. Eluted proteins were dialyzed overnight in the presence of TEV protease (produced in-house) and passed over Ni-NTA beads. Tag-free eRF1(AAQ) was collected in the flowthrough, and loaded on to a Superdex 75 16/60

column in storage buffer (50 mM HEPES, pH 7.4, 150 mM KOAc, 5 mM $Mg(OAc)_2$, 10 mM imidazole, 10% glycerol, 1 mM DTT).

In vitro translation reaction and RNC purification

Translationally active RRL was generated from untreated rabbit reticulocyte lysate (Green Hectares, USA) as described in (*38*) with certain modifications. The final translation reaction (per 20 µL) contained 11.4 µL activated nuclease-treated and activated RRL, 200 ng/µL mRNA, 1.2 mM $MgCl_2$, 0.5 mM DTT, 0.4 mM spermidine and eRF1(AAQ) at a final concentration of 3.5 µM to trap translating ribosomes. Reactions were incubated for 25 min at 32 °C. 4 mL translation reactions were chilled on ice for 10 mins to halt translation, and HEPES KOH pH 7.5 was added to a final concentration of 50 mM. Chilled translation reactions were directly incubated with 400 µL of packed anti-FLAG M2 beads (Sigma) for 2 h at 4 °C with gentle mixing. The beads were then washed with 4 ml of 50 mM HEPES, pH 7.4, 100 mM KOAc, 5 mM $Mg(OAc)_2$, 0.1% Triton X-100, 1 mM DTT; 4 ml of 50 mM HEPES, pH 7.4, 250 mM KOAc, 5mM $Mg(OAc)_2$, 0.5% Triton X-100, 1 mM DTT and 6 ml of RNC buffer (50mM HEPES, pH 7.4, 100 mM KOAc, 5 mM $Mg(OAc)_2$, 1mM DTT). RNCs were eluted after 4 sequential 10 min incubations at room temperature in RNC buffer that contained 0.2 mg/mL 3x FLAG peptide (Sigma). The elutions were combined and centrifuged at 186,000 x g at 4 °C for 2 h in a TLA 55 rotor (Beckman Coulter). Supernatant was discarded and the pellet was resuspended in RNC buffer at a concentration of 80 nM. At each step of translation reactions and purification, aliquots were taken to perform Western blots using an anti-FLAG antibody (Sigma Cat. No. A8592).

Cryo-electron microscopy, sample preparation and data collection

To prepare cryo-EM grids using a Vitrobot (ThermoFisher Scientific), 5 µL of sample were applied to Quantifoil R2/2 holey carbon copper grids (Quantifoil Micro Tool), which were covered with a sheet of continuous carbon and glow-discharged for 15 s at 15 mA using a easiGlow Discharge cleaning system (Pelco) beforehand. The sample was incubated for 30 s in the Vitrobot chamber, which was kept at 4 °C and 100% humidity. The excess sample was then blotted for 6 to 10 s and immediately plunged into a mixture of ethane:propane (1:2). Grids were loaded into a Titan Krios cryo-transmision electron microscope (ThermoFisher Scientific) operating at 300 kV, and data collection was performed using a GIF Quantum LS energy filter (Gatan)-mounted K3 direct electron camera (Gatan) in counting and super-resolution mode. The microscope was used with a 81'000 x nominal magnification, resulting in a physical pixel size of 1.08 Å/pixel and therefore a super-resolution pixel of 0.54 Å/pixel. Both data collections were set up with the EPU program (Thermo Fisher Scientific), taking advantage of the Aberration-Free Image Shift (AFIS) mode. 40-frame movies were recorded with a total dose of 60 $e^- * Å^{-2}$, and the defocus was set to change between -0.6 and -3 µm with 0.3 µm increments. The energy filter slit width was set to 20 eV during exposure.

Cryo-electron microscopy data processing (fig. S1)

Approximately 10'000 movies were processed with MotionCor2 (*39*) to dose-weight the radiation damage, apply the gain correction, correct for motion during the exposure, and bin twice the super-resolution micrographs. GCTF (*40*) was then used to estimate the Contrast Transfer Functions (CTF) of the motion-corrected and dose-weighted micrographs.

Based on both the quality of the micrographs and their respective CTF, 9'580 micrographs were selected for further processing. 1'897'486 particles were picked using the Laplacian-of-Gaussian-based method in Relion3.1(*41*). Those particles were extracted using Relion3.1 and then imported into cryoSPARC v2 (*42*). After 2D classification using cryoSPARC v2 to select good particles, they were first refined into a 3D structure, followed by a two-step classification using cryoSPARC v2 heterogeneous refinement and a variability analysis (*43*). Out of those classifications, 695'501 particles containing a visible pseudoknot were selected for final refinement, resulting in a reconstruction with an overall resolution of 2.2 Å (fig. S1).

Those particles were then further classified in Relion3.1 using a mask around the pseudoknot, which resulted in a class containing 171'706 particles with the three stems of the pseudoknot well visible. Those particles were refined in cryoSPARC v2 to reach an overall resolution of 2.4 Å (fig. S1).

General model building

As a starting model, the small subunit head, body and the large subunit of the 3.3Å structure of a rabbit ribosomal elongation complex (PDB 5LZS (*44*)) were rigid body docked into the 2.2 Å cryo-EM map using CHIMERA (*45*). While the ribosomal proteins could be readily readjusted in COOT (*46*), obvious discrepancies in the rRNA sequences lead us to retrieve updated full-length sequences for the 28S and 5.8S rRNAs, the 18S rRNA and the 5S rRNA from the *Oryctolagus cuniculus* Transcriptome Shotgun Assembly (TSA) at the GeneBank (entries GBCN01009604.1, GBCT01000564.1 and GBCM01014045.1, respectively). The model was revised and completed using a sharpened 2.2 Å cryo-EM map in combination with the further classified 2.4Å cryo-EM map. While the further classified map revealed the features of the COVID-19 mRNA with the pseudoknot, the nascent chain, and tails of peripheral proteins and rRNA expansion segments, the higher resolved map showed clear additional densities for the rRNA and tRNA modifications, bound spermine/spermidine ligands, ions such as hexa-coordinated Mg^{2+}, and the ordered solvent (fig. S3). The strictly octahedrally coordinated magnesium ions were built according to their star-shaped appearance, while the nature of numerous other ions remained unassigned due to incomplete or asymmetric coordination shells. Automated water picking was performed using the phenix.douse tool recently developed for ordered solvent building into high-resolution cryo-EM maps and implemented in PHENIX (*47*), followed by manual examination.

Building of the modified rRNA and tRNA(Phe) residues

In an attempt to make use of the modified rRNA residues described in a cryo-EM structure of the human ribosome (PDB 6EK0 (*19*)), we superimposed the human rRNA onto our structure. However, it turned out that the modification

pattern disagreed with our high-resolution EM-map to a considerable extent. In contrast, when we used for model building the full set of human rRNA modifications recently established by quantitative mass spectroscopy, we observed that the modification pattern coincided with our maps remarkably well, supporting the existence of eukaryotic-typical rather than human-specific rRNA modifications (*18*). Although the difference between pseudouridines and uridines cannot be directly established from the overall shape of the base, we saw a water molecule bound to the N1 atom of the pseudouridine ring in many positions, while in uridine such a hydrogen bond does not exist. Therefore, apart from all modifications comprising the easily distinguishable additional groups, we decided to also include the complete set of pseudouridines in our model, which may serve as a template for other mammalian high-resolution structures in the future. For modelling of the P-site tRNA(Phe), we followed the sequence and modification pattern of mammalian tRNA(Phe) determined by Keith and Dirheimer (*20*), which could be built without any obvious discrepancy into our maps. As the E-site tRNA was less-well resolved and its identity could not be established from the maps, we modeled a generic tRNA and left the residues unassigned.

Real space refinement

The completed structure was refined for five cycles into the higher resolution (including ordered solvent) or further classified (excluding waters not involved in Mg^{2+} hexa-coordination) maps, respectively, using real space refinement in PHENIX version 1.18 (*47*). Custom restraints for modified rRNA residues were generated using PRODRG (*48*) or exported from COOT (*46*), and protein secondary structure and Ramachandran as well as RNA base pair and stacking restraints were applied throughout to maintain good model geometry also in less-well-ordered areas at the periphery of the 80S maps. Remaining discrepancies between the model and the maps were detected and corrected using real space difference maps, followed by two additional cycles of refinement as described above. The structures were validated using MOLPROBITY (*49*). The resulting models display excellent geometries and map correlations (table S1), and the resolution of the model *vs.* map FSCs at a value of 0.5 coincide well with the resolution determined between the map half-sets at the FSC=0.143 criterion (fig. S2B).

Alignment of SARS2 with sequences of related coronaviruses

The sequences for the alignments shown in Fig. 4G were obtained from the UniProt database (https://www.uniprot.org/) and correspond to the polyprotein translated from ORF1a before the frameshift event takes place. The following sequences were used: P0DTC1 (SARS2), P0C6T7 (Bat coronavirus Rp3), P0C6U8 (SARS), P0C6F5 (Bat coronavirus 279/2005), A0A0K1Z0N1 (Bat SARS-like coronavirus YNLF_31C), R9QTH2 (Bat coronavirus Cp/Yunnan2011), P0C6T6 (Bat coronavirus HKU9), K9N638 (MERS), U5LR11 (*Betacoronavirus Erinaceus*/VMC/DEU/2012). The sequences were aligned with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (*50*) and visualized with ESPript (http://espript.ibcp.fr) (*51*).

Structure-guided mutagenesis experiments

Dual luciferase expression constructs were generated by either 1-step or 2-step PCR using primer sequences outlined in table S2, during which 5' *Xho*I and 3' *Bam*HI restriction sites were incorporated. PCR amplicons were digested with *Xho*I / *Bam*HI and cloned into a *Psp*XI / *Bgl*II -digested pSGDlucV3.0 vector (Addgene 119760, (*52*)).

HEK293T cells (ATCC) were maintained in DMEM supplemented with 10% FBS, 1 mM L-glutamine and antibiotics. HEK293T cells were transfected with Lipofectamine 2000 reagent (Invitrogen) using the 1-day protocol, in which suspended cells are added directly to the DNA complexes in half-area 96-well plates. To each well, 25 ng of each plasmid plus 0.2 µl Lipofectamine 2000 in 25 µl Opti-Mem (Gibco) were added. The transfecting DNA complexes in each well were incubated with 4×10^4 cells suspended in 50 µl DMEM + 10% FBS at 37 °C in 5% CO_2 for 24 h.

For the *in vitro* translation reactions, Plasmid DNAs (50 ng) were used as templates in 5 µl reactions of the RRL TNT® T7 Quick Coupled Transcription/Translation system (Promega) supplemented with 1 mM methionine (25 µM final concentration). Reactions were incubated at 30 °C for 90 min.

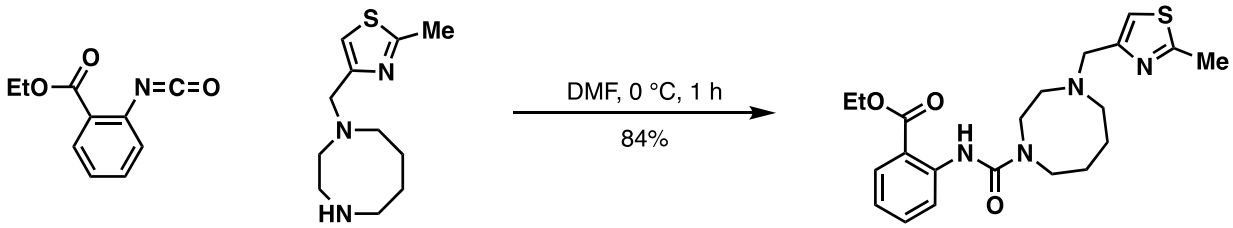
For the Dual Luciferase Assay, relative light units were measured on a Veritas Microplate Luminometer with two injectors (Turner Biosystems). Transfected cells were lysed in 15 µl of 1 × passive lysis buffer (PLB), and light emission was measured following injection of 50 µl of either Renilla or firefly luciferase substrate (*53*). For *in vitro* translation reactions, 45 µl of 1 × PLB was added to each 5 µl reaction, and luciferase activities were assayed from 10 µl.

Frameshifting efficiencies (% frameshifting) were determined by calculating relative luciferase activities (firefly/Renilla) from test constructs and dividing by relative luciferase activities from replicate wells of matched in-frame control constructs. The number of biological replicates for each experiment are indicated in each figure legend. For all Boxplots center line medians are shown; box limits indicate the 25th and 75th percentiles as determined by the R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots.

Synthesis of ethyl 2-(4-((2-methylthiazol-4-yl)methyl)-1,4-diazocane-1-carboxamido)benzoate (MTDB)

The RNA pseudoknot-binding ligand ethyl 2-(4-((2-methylthiazol-4-yl)methyl)-1,4-diazocane-1-carboxamido)benzoate was prepared from commercially obtained materials using the following procedure. Under a nitrogen atmosphere, 4-((1,4-diazocan-1-yl)methyl)-2-methylthiazole (110 mg, 0.51 mmol) was charged to a 10 mL oven dried round bottom flask equipped with magnetic stirring. Anhydrous DMF (2.0 mL) was added and the solution cooled to 0 °C using an ice bath. Ethyl 2-isocyanatobenzoate (130 mg, 0.70 mmol) was dissolved in anhydrous DMF (2.0 mL) and added to the reaction dropwise. The reaction was stirred at 0 °C for 1 h then diluted with water (40 mL). The product was extracted in ethyl acetate (4 x 15 mL). The combined organic phases were washed with water (4 x 20 mL) and brine (2 x 20 mL) then dried over an excess of anhydrous magnesium sulfate. The resulting solution was concentrated under reduced pressure to afford the crude product which was purified by flash silica chromatography (100% ethyl acetate) to afford the title compound as a light-yellow oil (170 mg, 0.43 mmol, 84% yield). The product

was characterized by NMR spectroscopy (1H, 13C, COSY, HSQC, HMBC on a Bruker AV-400 and -500 MHz spectrometer), IR spectroscopy (thin film, Jasco FT/IR-4100) and HRMS (measured by the mass spectrometry service of the ETH Zürich Laboratorium für Organische Chemie on a Bruker Daltonics maXis ESI-QTOR spectrometer). The ligand was stored in a sealed vial at 4 °C.



Testing the effect of MTDB on SARS-CoV-2 infected cells

VeroE6 cells (kindly provided by Doreen Muth, Marcel Müller and Christian Drosten, Charité, Berlin, Germany) were propagated in Dulbecco's modified EMEM (DMEM), supplemented with 10% heat inactivated fetal bovine serum, 1% nonessential amino acids, 100 µg/mL of streptomycin and 100 IU/mL of penicillin, and 15 mM of HEPES. Cells were maintained at 37 °C in a humidified incubator with 5% CO_2. Cells were infected using SARS-CoV-2 (SARS-CoV-2/München-1.1/2020/929, kindly provided by Daniela Niemeyer, Marcel Müller and Christian Drosten, Charité, Berlin, Germany).

Antiviral analysis

Cells were plated to 20,000 cells per 96 well 24 h prior to infection. Cells were infected with SARS-CoV-2 at a multiplicity of infection (MOI) of 0.01 for 1.5 h at 37 °C and washed 3 times with PBS. MTDB (or respective volumes of its 20 % DMSO/100 mM HCl solvent) was added to cells in following concentrations: 0 µM, 5 µM, 10 µM, 20 µM, 50 µM, 100 µM, 150 µM and 200 µM. 24 hours post infection virus-containing supernatant was serially diluted, and the 50% tissue culture infectious dose ($TCID_{50}$) per milliliter was displayed using Crystal Violet and calculated by the Spearman-Kärber algorithm after 72 h as described (*54*). Cytotoxic effects of MTDB or its 20 % DMSO/100 mM HCl solvent were monitored using CytoTox 96® Non-Radioactive Cytotoxicity Assay (Promega).

Immunofluorescence analysis

VeroE6 cells were fixated with 4% formalin. Cells were permeabilized in PBS supplemented with 50 mM NH_4Cl, 0.1% (w/v) saponin and 2% (w/v) bovine serum albumin. Cells were immunostained with mouse monoclonal antibody against dsRNA (SCICONS, clone J2). Alexa-Fluor 488-labeled donkey-anti mouse IgG (H+L) (JacksonImmuno) was used as secondary antibody. Images were acquired using an EVOS FL Auto 2 Imaging System, using 10x objective, processed using Fiji software packages (*55*) and assembled using the FigureJ plugin (*56*).

Figure generation

All density and structure representations were generated using UCSF ChimeraX (*57*) Local Resolution estimate were performed within cryoSPARC v2 (*42*), which uses an implementation of BlocRes (*58*).

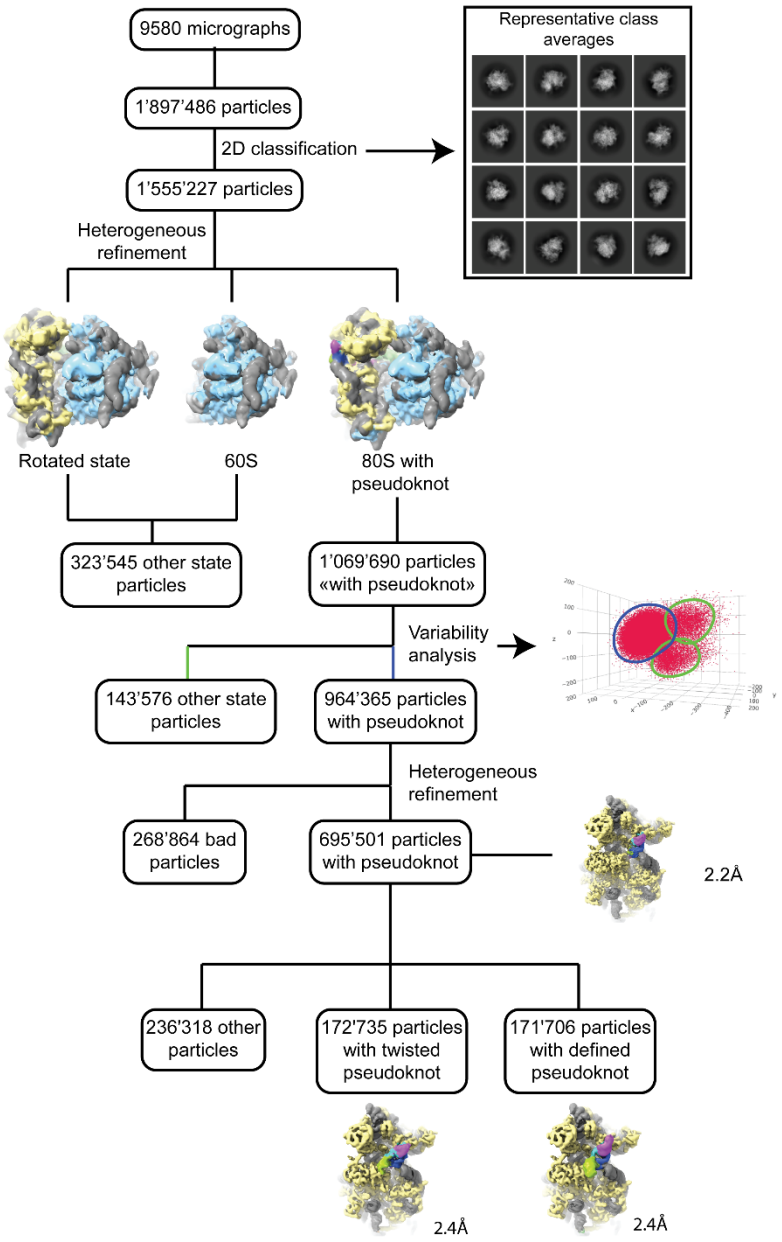


Fig. S1 Cryo-EM data processing workflow.

From 9580 micrographs selected for their quality and the quality of their Contrast Transfer Function (CTF), 1'897'486 particles were picked and extracted using Relion3.1 (*41*). The particles were then imported into cryoSPARC v2 (*42*) and processed to remove bad particles, 60S ribosomal subunit and 80S ribosomes in other states. 695'501 selected particles were refined using cryoSPARC v2 to reach an overall resolution of 2.2 Å (fig. S2A-B), corresponding to the Nyquist frequency for our collection pixel size. Further classification using Relion3.1 allowed the separation of a good fractions of particles (171'706) with a clearly defined pseudoknot, which were then refined in cryoSPARC v2 to an overall resolution of 2.4 Å (fig. S2B-C).

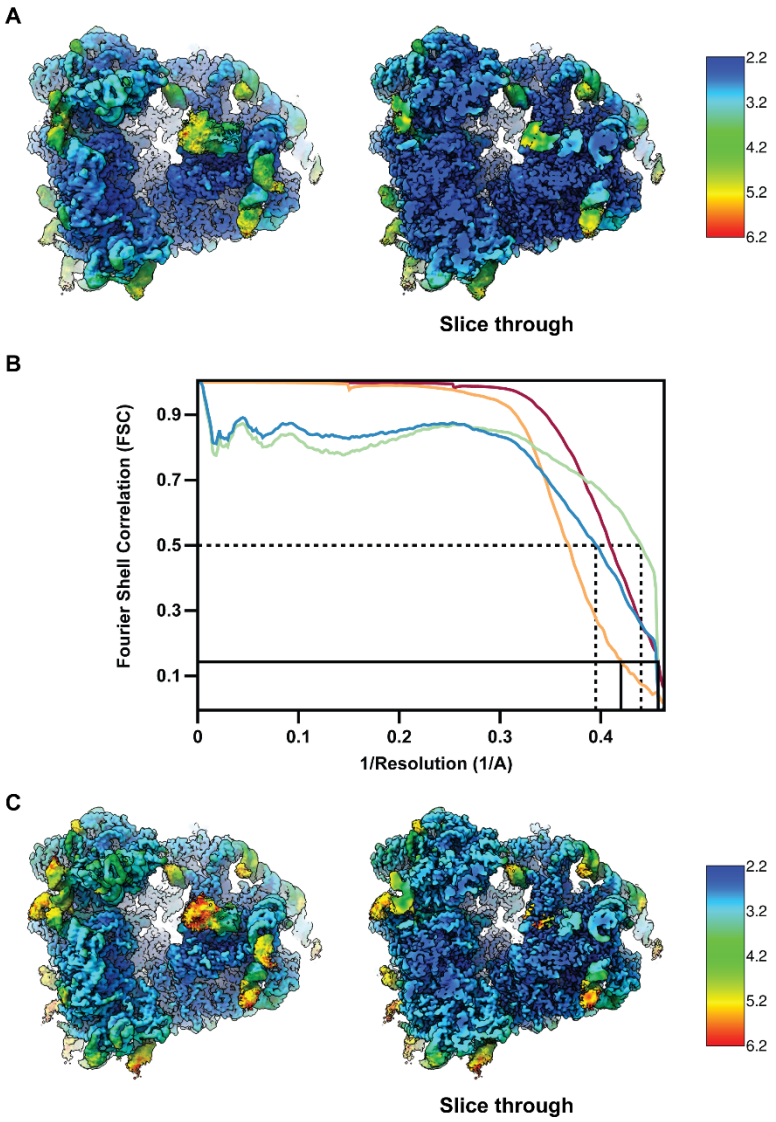


Fig. S2 Local resolution estimate and cryo-EM FSC.
(**A**) Local resolution heat map of the high resolution cryo-EM reconstruction with a slice through the density on the right. The local resolution was calculated with the cryoSPARC v2 (*42*) implementation of BlocRes (*58*). The local resolution varies from 2.2Å in the center of the ribosome to roughly 5 to 6 Å for the flexible rRNA expansion and protein segments at the periphery. (**B**) Fourier shell correlations (FSCs) between masked half maps for the high resolution cryo-EM map (red) and the reconstruction of the further classified set of particles (orange). FSCs were also calculated between maps and models for the high resolution (green) and the further classified set (blue). The similar values for the obtained resolutions of map-versus-map and map-versus-model FSCs indicates the absence of overfitting. (**C**). Similar to panel A, but for the reconstruction from further classified set of particles. The resolution varies from roughly 2.5 Å in the center to 5-6 Å for the peripheral rRNA expansion and protein segments and the pseudoknot.

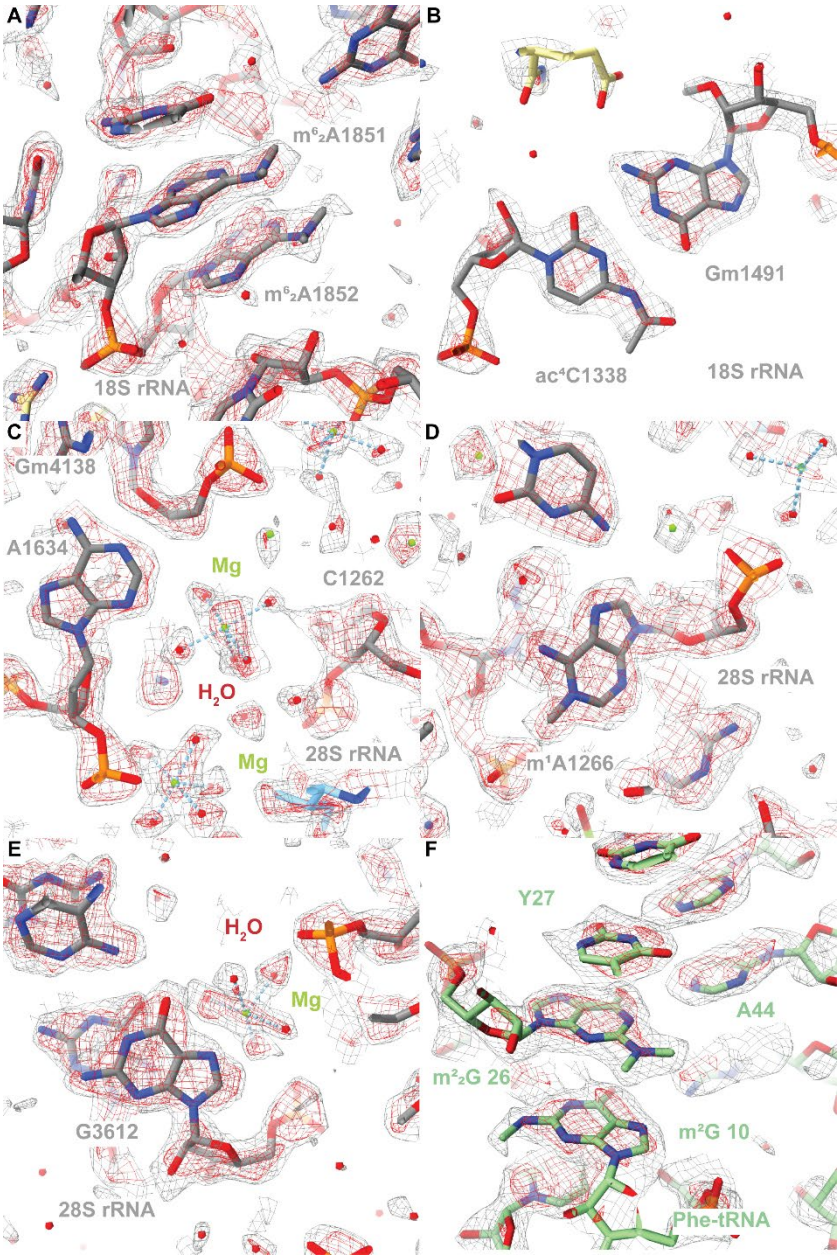


Fig. S3 Close-up view of rRNA modification and quality of the cryo-EM reconstruction.
The high resolution cryo-EM map is shown at two contour levels as grey and red mesh. (**A**) Close-up view of two N^6,N^6-dimethyladenosine found at position 1851 and 1852 of the SSU 18S rRNA. (**B**). Close-up view of two modified residues base-pairing in the SSU 18S rRNA: N^4-acetylcytidine 1338 and O^2-methylguanosine 1491. (**C**) Close-up view of coordinated magnesium ions found in the core of the LSU, for which clear density for the water molecules can be observed. (**D**) Close-up view of the 1-methyladenosine 1266 of the LSU 28S rRNA (**E**) Close-up view of G3612 of the 28S rRNA which was previously reported as 7-propylguanosine 3880 (PDB 6EK0 (*19*)). (**F**) Close-up view of the Phe-tRNA modifications pseudouridine 27, N^2,N^2-dimethylguanosine 26, and N^2-methylguanosine 10.

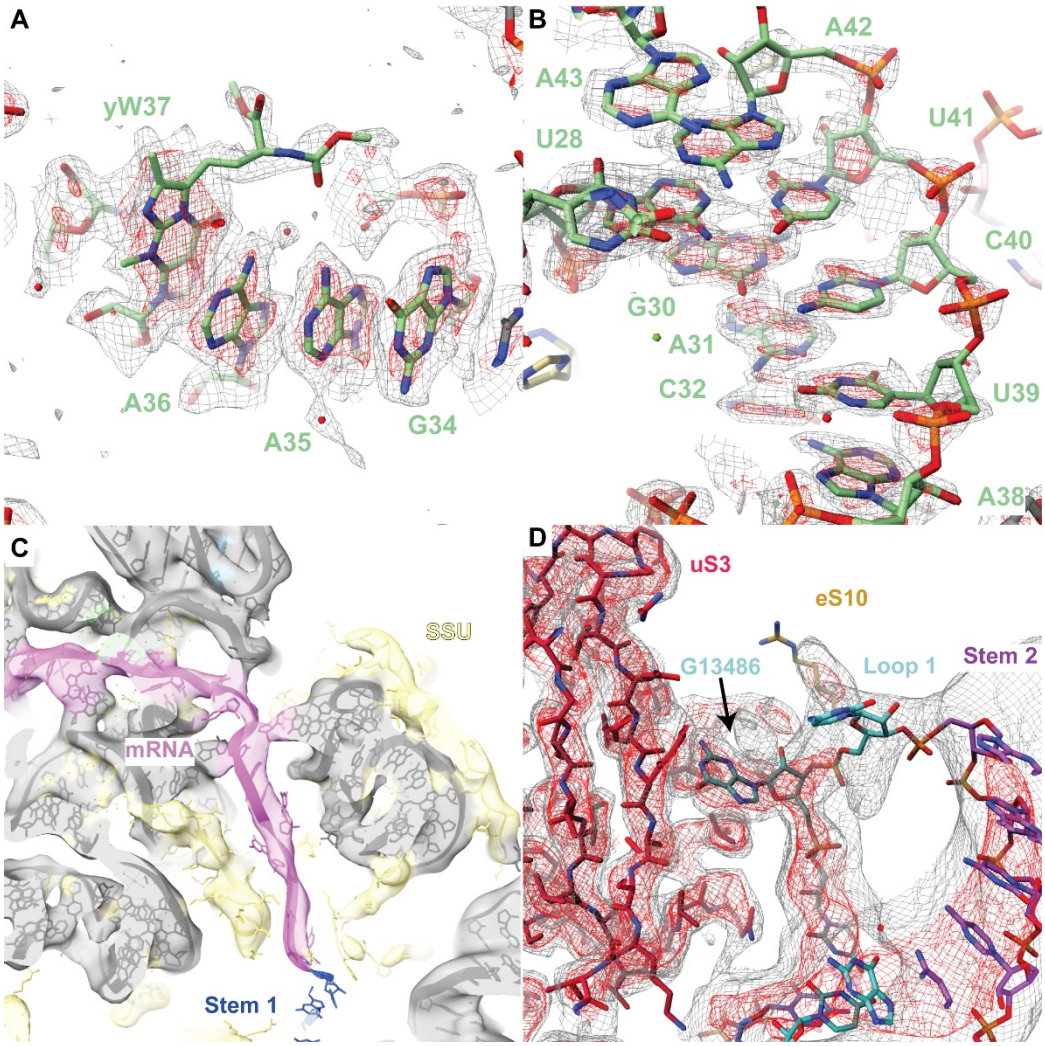


Fig. S4 Close-up view of features in cryoEM maps.

Densites shown in panel (A), (B) and (D) are shown at two contour levels as grey and red mesh. (**A**) Close-up view of the anticodon loop of the Phe-tRNA found in the P-site. The hyper-modified wybutosine in position 37 can be clearly identified in the density of the high-resolution cryo-EM reconstruction. (**B**) Close-up view of the anticodon stem-loop shown in the high-resolution cryo-EM map. The tRNA in the P-site could be unambiguously identified as Phe-tRNA based on the purine-pyrimidine pattern of the codon-anticodon (Fig. 1E), its modification pattern (fig. S3), the anticodon stem loop shown here, and the attached amino acid residues of the nascent chain (**C**) Close-up view of the mRNA density seen in the cryo-EM map after further classification. A continuous density can be seen from the P-site codon up to the stem of the pseudoknot, allowing us to set the pseudoknot registry start, which is shifted relative to the previously reported secondary structure diagram (*13*). (**D**) Close-up view of the loop 1 of the pseudoknot, where residue G13486 can be seen being flipped out to contact uS3. The density shown corresponds to the cryo-EM map classified for the pseudoknot.

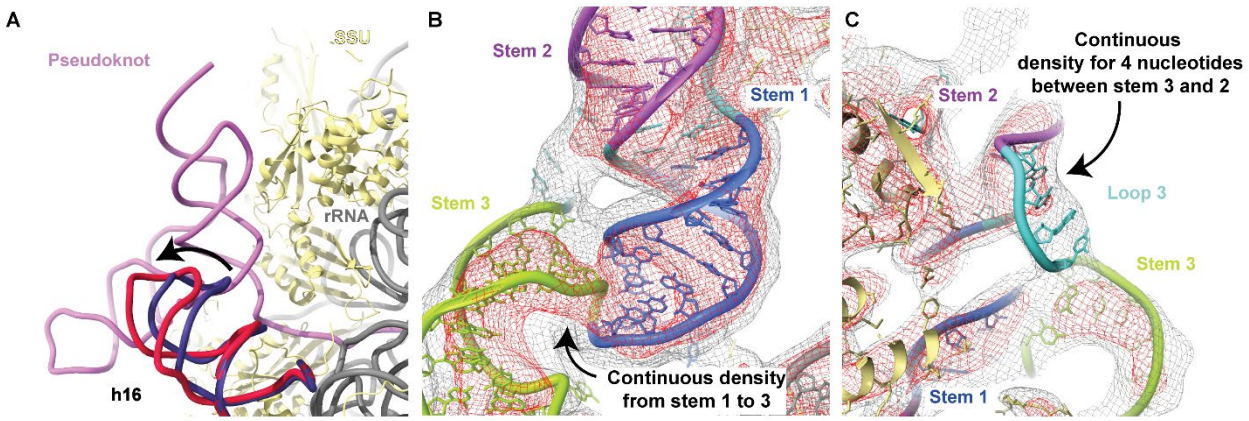


Fig. S5 Structure of the pseudoknot and interaction with h16.

(**A**) Close-up view of the pseudoknot (pink) bound to the SSU (yellow proteins and grey rRNA). The pseudoknot can be seen interacting with 18S rRNA helix h16 (red), pushing it outwards compared to translating ribosome (purple, PDB 6SGC (*59*). (**B**) Close-up view of the pseudoknot to highlight the transition between the Stems 1 and 3. The direct connection requires adjustments relative to the previously reported secondary structure diagram and agrees with the altered base-pairing pattern observed for Stem 1 (fig. S4C). (**C**) Close-up view of the other side of the pseudoknot relative to panel B. Loop 3 is well visible and corresponds in length to the newly defined four unpaired nucleotides, in agreement with our new secondary structure diagram.

Table S1.EM data collection and structure refinement statistics.

	80S – SARS2 pseudoknot (higher resolution map) (EMD-YYYY, PDB XXXX)	80S – SARS2 pseudoknot (further classified set map) (EMD-WWWW, PDB ZZZZ)
Data collection and processing		
Magnification	81'000x (nominal)	
Voltage (kV)	300	
Electron exposure (e–/Å2)	60	
Defocus range (μm)	0.6-3	
Pixel size (Å)	1.08 (super-resolution pixel at 0.54Å/pixel)	
Initial particle images (no.)	1'897'486	
Final particle images (no.)	695'501	171'706
Map resolution at FSC=0.143 (Å)	2.2	2.4
Refinement		
Model resolution at FSC=0.5 (Å)	2.3	2.7
CCmask	0.81	0.82
Map sharpening B factor (Å2)	– 75.8	– 71.7
Model composition		
Non-hydrogen atoms	237'931	
Protein residues	12'093	
RNA residues (modified)	6050 (220)	
Ligands: Zn2+/Mg2+/other ions/SPM/SPD	8/382/284/4/22	
Waters (strictly coordinated to Mg2+)	10'155 (1874)	- (1874)
B factors min/max/mean (Å2)		
Protein	1/50/25	7/57/27
RNA	1/78/33	13/87/37
Ligand	3/56/21	9/66/23
Water	1/104/17	-
R.m.s. deviations		
Bond lengths (Å)	0.002	0.002
Bond angles (°)	0.667	0.659
Validation		
MolProbity score	1.6	1.6
Clashscore	5.88	5.77
Poor rotamers (%)	1.81	2.08
Protein		
EM Ringer score	3.97	3.55
Ramachandran plot		
Favored (%)	96.63	97.8
Allowed (%)	2.33	2.15
Disallowed (%)	0.04	0.04
RNA		
Pucker outliers (%)	0.10	0.10
Bond outliers (%)	0.29	0.29
Angle outliers (%)	0.03	0.03
Suite outliers (%)	14.5	14.7

Table S2. List of primers used for mutagenesis experiments

Primer name	Primer sequence (5'-3')
SARS CoV2 WT S XhoI	ATAACTCGAGACCAACTTGTGCTAATGACCCTGTG
SARS CoV2 WT AS BamHI	ATAAGGATCCATTGTAGATGTCAAAAGCCC
G-A S	GCGGTGTAAGTACAGCCCGTCTTACAC
G-A AS	GTGTAAGACGGGCTGTACTTACACCGC
G-C S	GCGGTGTAAGTCCAGCCCGTCTTACAC
G-C AS	GTGTAAGACGGGCTGGACTTACACCGC
del L1 S	GGGTTTGCGGTGTAAGAGCCCGTCTTACACCG
del L1 AS	ACGGTGTAAGACGGGCTCTTACACCGCAAACCC
Bulged A deletion AS	ATAAGGATCCGATTGTAGATGTCAAAAGCCCGTATACGACATCAGTAC
del L3 AS	TGTAGATGTCAAAAGCCCATACGACATCAGTAC
plus 6 stop S	GCGGTGAAAGTGTAGCCCGTCTTTCAC
plus 6 stop AS	GTGAAAGACGGGCTACACTTTCACCGC
plus 15 stop S	GGGTTTGCGGTGTTAGTGCAGCCCGTCTAACACCGTGCGGC
plus 15 stop AS	TGCCGCACGGTGTTAGACGGGCTGCACTAACACCGCAAACCC
UAA to AAA S	GGGTTTGCGGTGAAAGTGCAGCCCGTCTTTCACCGTGCGGC
UAA to AAA AS	TGCCGCACGGTGAAAGACGGGCTGCACTTTCACCGCAAACCC
UAA to UAU S	GGGTTTGCGGTGTATGTGCAGCCCGTCATACACCGTGCGGC
UAA to UAU AS	TGCCGCACGGTGTATGACGGGCTGCACATACACCGCAAACCC
LR to AA S	TGTAGTTGTGATCAAGCCGCCGAACCCATGCTTCAGTCAGCT
LR to AA AS	AGCTGACTGAAGCATGGGTTCGGCGGCTTGATCACAACTACA
Syn changes to attenuator S	CTCCGCGAACCCATGTTACAATCAGCAGATGCACAATCG
Syn changes to attenuator AS	CGATTGTGCATCTGCTGATTGTAACATGGGTTCGCGGAG