

## RESEARCH ARTICLE

## Translational buffering by ribosome stalling in upstream open reading frames

Ty A. Bottorff<sup>1,2</sup>, Heungwon Park<sup>1</sup>, Adam P. Geballe<sup>3\*</sup>, Arvind Rasi Subramaniam<sup>1,2\*</sup>

**1** Basic Sciences Division and Computational Biology Program of the Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, Washington, United States of America, **2** Biological Physics, Structure and Design Graduate Program, University of Washington, Seattle, Washington, United States of America, **3** Human Biology and Clinical Research Divisions, Fred Hutchinson Cancer Center, Seattle, Washington, United States of America

\* [ageballe@fredhutch.org](mailto:ageballe@fredhutch.org) (APG); [rasi@fredhutch.org](mailto:rasi@fredhutch.org) (ARS)

## Abstract

Upstream open reading frames (uORFs) are present in over half of all human mRNAs. uORFs can potentially regulate the translation of downstream open reading frames through several mechanisms: siphoning away scanning ribosomes, regulating re-initiation, and allowing interactions between scanning and elongating ribosomes. However, the consequences of these different mechanisms for the regulation of protein expression remain incompletely understood. Here, we performed systematic measurements on the uORF-containing 5' UTR of the cytomegaloviral *UL4* mRNA to test alternative models of uORF-mediated regulation in human cells. We find that a terminal diproline-dependent elongating ribosome stall in the *UL4* uORF prevents decreases in main ORF protein expression when ribosome loading onto the mRNA is reduced. This uORF-mediated buffering is insensitive to the location of the ribosome stall along the uORF. Computational kinetic modeling based on our measurements suggests that scanning ribosomes dissociate rather than queue when they collide with stalled elongating ribosomes within the *UL4* uORF. We identify several human uORFs that repress main ORF protein expression via a similar terminal diproline motif. We propose that ribosome stalls in uORFs provide a general mechanism for buffering against reductions in main ORF translation during stress and developmental transitions.

## OPEN ACCESS

**Citation:** Bottorff TA, Park H, Geballe AP, Subramaniam AR (2022) Translational buffering by ribosome stalling in upstream open reading frames. *PLoS Genet* 18(10): e1010460. <https://doi.org/10.1371/journal.pgen.1010460>

**Editor:** Gregory S. Barsh, HudsonAlpha Institute for Biotechnology, UNITED STATES

**Received:** September 7, 2022

**Accepted:** October 4, 2022

**Published:** October 31, 2022

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1010460>

**Copyright:** © 2022 Bottorff et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data and code associated with the manuscript is publicly available at [https://github.com/rasilab/bottorff\\_2022](https://github.com/rasilab/bottorff_2022).

**Funding:** This work was supported by grants from the National Institute of General Medical Sciences

## Author summary

Life requires proteins for nearly all functions. mRNA molecules relay information from the DNA code to protein-making molecular machines called ribosomes. Ribosomes load onto mRNA molecules and translate sections of the code, termed open reading frames, to make proteins. Cells' needs for proteins change depending on the cell type and growth environment. Thus, protein synthesis is a highly regulated process. One way for cells to regulate protein synthesis is to vary the rate at which ribosomes load onto mRNA molecules. Unexpectedly, we found that increasing the rate of ribosome loading onto mRNA molecules can decrease protein synthesis. This unexpected result can arise in mRNA molecules that have multiple open reading frames from which multiple distinct proteins are

(GM119835, to ARS), the National Science Foundation (1846521, to ARS) and the National Institute of Allergy and Infectious Diseases (AI156152, to APG), and a National Science Foundation graduate fellowship (2140004, to TAB). This research was supported by the Genomics Shared Resource of the Fred Hutch/University of Washington Cancer Consortium (P30 CA015704) and Fred Hutch Scientific Computing (NIH grants S10-OD-020069 and S10-OD-028685). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

made. If there are sequences that stall ribosomes within the first encountered open reading frame on an mRNA molecule, then we find that increasing ribosome loading can decrease protein synthesis at the second open reading frame. Our findings can be explained if trailing ribosomes that collide with a stalled ribosome within the first encountered open reading frame dissociate from, rather than queue on, the mRNA molecule. Our findings have implications for stress-responsive mRNA molecules whose second open reading frames are preferentially translated during cellular stress when the ribosome loading rate is reduced.

## Introduction

About half of human mRNAs have at least one upstream open reading frame (uORF) in their 5' untranslated region [1–3]. Ribosome profiling studies estimate that at least twenty percent of these uORFs are actively translated [4,5]. uORFs can regulate gene expression via the biological activity of the uORF peptide, but they also often *cis*-regulate translation of the downstream main ORF [6,7]. Despite having poor initiation sequence contexts, many eukaryotic uORFs repress main ORF translation [1,3,4,7–11]. uORF mutations are implicated in several human diseases via changes to main ORF translation [12,13]. For example, uORF mutations in oncogenes and tumor suppressors can act as driver mutations in cancer [14,15].

uORFs can regulate translation via a variety of molecular mechanisms. uORFs can constitutively repress translation by siphoning away scanning ribosomes from initiating at downstream main ORFs. Multiple uORFs can interact together to regulate the re-initiation frequency at the main ORF. For example, uORFs in the *S. cerevisiae* *GCN4* mRNA and the homologous human *ATF4* mRNA render main ORF translation sensitive to cellular levels of the eIF2 $\alpha$ -GTP-tRNA<sub>Met</sub> ternary complex [16,17]. Although the initiation rate usually limits translation [18–20], inefficient elongation or termination on uORFs can also regulate protein expression by preventing scanning ribosomes from reaching the main ORF [21–26]. Inefficient elongation can be driven by the nascent uORF peptide [27,28], poorly translated codons in the uORF [29,30], or small molecules such as amino acids or polyamines [23,24]. Further, interactions between scanning and elongating ribosomes on uORFs may cause dissociation of scanning ribosomes or enhanced initiation at start codons [23,31,32].

Despite the plethora of proposed uORF regulatory mechanisms, their implications for the regulation of protein expression are not clear. For example, are some uORF regulatory mechanisms more effective than others at repressing protein expression across a wide range of biochemical parameters? How do uORFs alter the response of main ORF translation to changes in cellular and environmental conditions? Answering these questions requires a joint accounting of how the different steps of translation, such as initiation, scanning, and elongation, together influence the overall rates of uORF and main ORF translation. Since it is not straightforward to monitor the rates of individual steps of translation [33], indirect measurements of protein expression are often necessary to infer the underlying mechanism of uORF-mediated regulation. Such inference requires rigorous kinetic models of uORF regulation that make testable experimental predictions for the effects of genetic mutations on protein expression.

Computational kinetic modeling has been widely used to study mechanisms of translational control [34]. Quantitative modeling of uORF translation has been used to support the regulated re-initiation model for the *GCN4* mRNA [35–37]. A computational model predicted that elongating ribosomes can dislodge leading scanning ribosomes on uORFs and confer stress resistance to protein expression [38]. However, these models have not been compared against

alternative models of uORF regulation that predict queuing or dissociation of scanning ribosomes upon collision with paused elongating ribosomes [21,23]. A critical barrier to such comparison has been the lack of a computational framework for the specification and simulation of different kinetic models of uORF-mediated translational regulation. Such a computational framework is necessary for the identification of unique experimental signatures of each proposed model and for their comparison with experimental measurements. Even though simulation code has been made available in many computational studies of mRNA translation [18,38], it is often highly tailored for specific models and cannot be easily modified to consider alternative regulatory mechanisms.

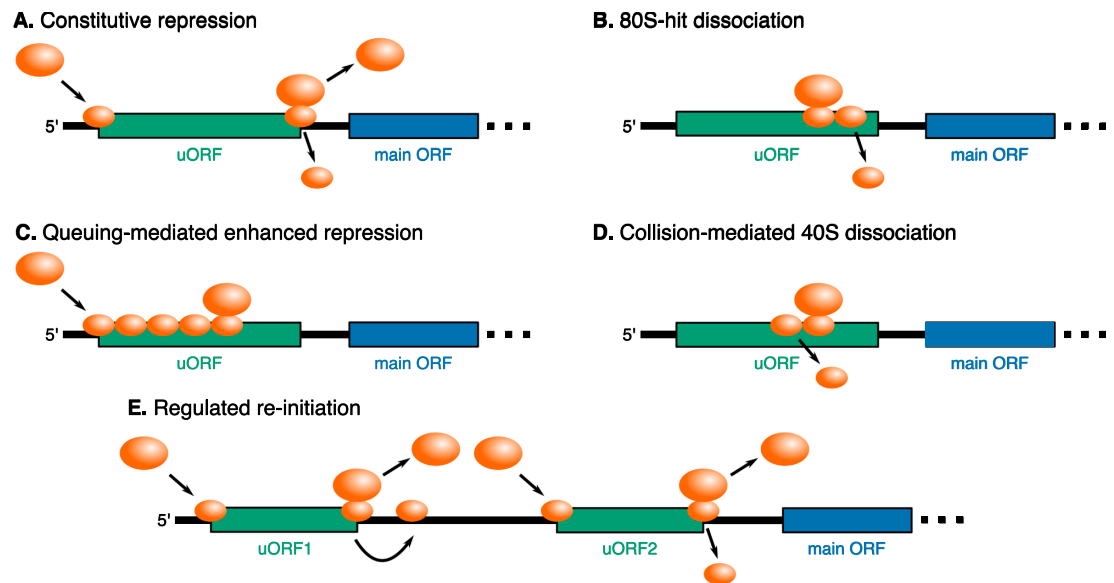
Here, we use experimental measurements on the well-studied uORF-containing 5' UTR of the human cytomegaloviral *UL4* mRNA to test different kinetic models of uORF-mediated translational control [21]. The second uORF (henceforth uORF2) in the *UL4* 5' UTR contains a terminal diproline motif that stalls 80S ribosomes by disrupting peptidyl transferase center activity [27,28]. For systematic model comparisons, we rely on a recent computational framework that allows easy specification and efficient simulation of arbitrary kinetic models of translational control [39]. Using this experimentally-integrated modeling approach, we find that the presence of 80S stalls in uORF2 of *UL4* 5' UTR confers resistance (henceforth called buffering) of main ORF translation to reduced ribosome loading on the mRNA. Modeling suggests that collisions of scanning ribosomes with the stalled 80S ribosome confer this buffering behavior. Experimental variation of the distance between the uORF2 start codon and the elongating ribosome stall supports a kinetic model in which scanning ribosomes dissociate rather than queue upon colliding with the 80S stall. We also identify several human uORFs that have repressive terminal diproline motifs similar to the *UL4* uORF2 80S stall. We propose that ribosome stalls in uORFs enable buffering of main ORF protein expression against reduced ribosome loading across cellular and environmental transitions. Together, our results illustrate the value of experimentally-integrated kinetic modeling for the comparison of different uORF regulatory mechanisms and the identification of novel experimental signatures from complex molecular interactions.

## Results

### Models of uORF regulation of main ORF translation

We surveyed five previously proposed models of uORF regulation of main ORF translation (Fig 1). We tested these models using a combination of computational modeling and experimental reporter assays. In the constitutive repression model [9] (Fig 1A), uORFs siphon away scanning ribosomes from the main ORF since re-initiation is usually infrequent [40–43]. In the 80S-hit dissociation model [38] (Fig 1B), elongating ribosomes that hit downstream scanning ribosomes cause the 3' scanning ribosomes to dissociate from the mRNA. In the queuing-mediated enhanced repression model [23] (Fig 1C), a stalled elongating ribosome within the uORF allows upstream scanning ribosomes to queue in the 5' region. This queuing can bias scanning ribosomes to initiate translation at the uORF start codon rather than leaky scan past it. In the collision-mediated 40S dissociation model [31,32] (Fig 1D), scanning ribosomes instead dissociate if they collide with a 3' stalled elongating ribosome.

Lastly, in the regulated re-initiation model [16,44,45] (Fig 1E), for example in the *GCN4* (*S. cerevisiae* homolog of human *ATF4*) mRNA, translation of the first uORF is followed by re-initiation at either a second downstream uORF or the main ORF depending on the stress status of the cell. After termination at the first uORF stop codon, scanning ribosomes must reacquire a new eIF2 $\alpha$ -GTP-tRNA<sub>Met</sub> ternary complex (TC) before re-initiating at a downstream start codon. The time to reacquire a new TC correlates with the proportion of phosphorylated



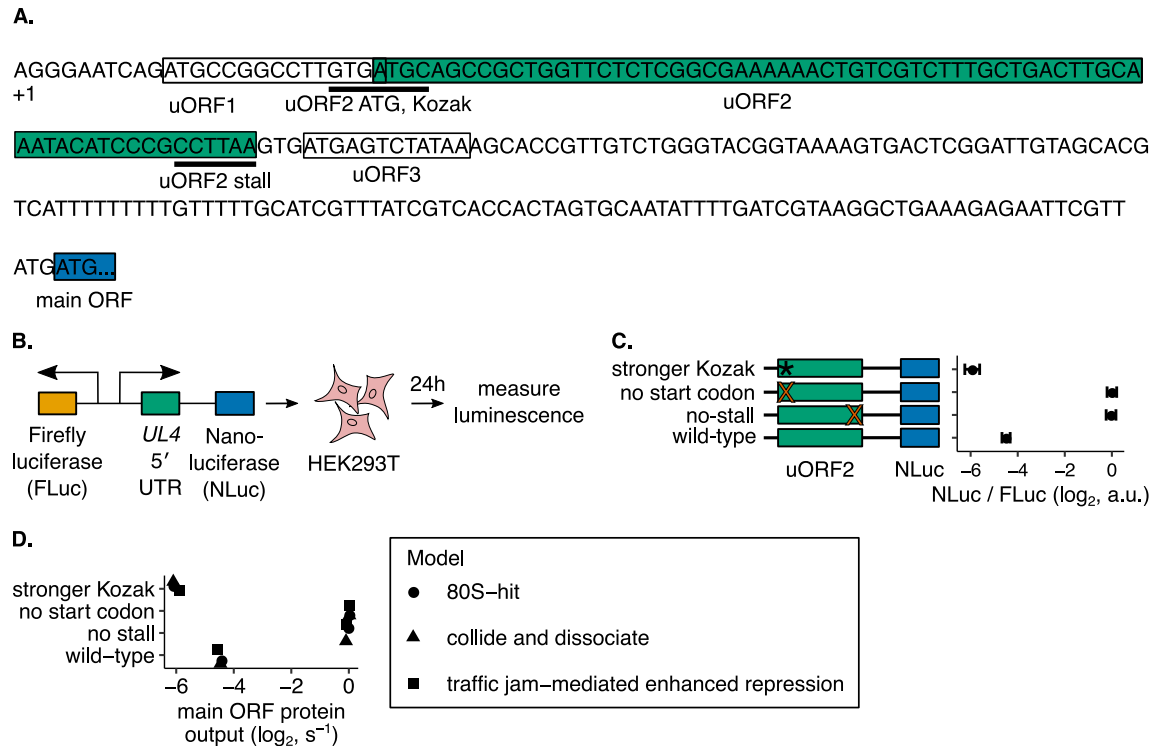
**Fig 1. Models of uORF regulation considered in this study.** (A) *Constitutive repression*. The uORF constitutively siphons away a proportion of scanning ribosomes from the main ORF. (B) *80S-hit dissociation*. Elongating ribosomes that collide with 3' scanning ribosomes cause the leading scanning ribosome to dissociate from the mRNA. (C) *Queuing-mediated enhanced repression*. Scanning or elongating ribosomes form a queue behind a 3' stalled elongating ribosome. If the queue correctly positions a scanning ribosome at the uORF start codon, then the proportion of scanning ribosomes that initiate translation at the uORF start codon increases. (D) *Collision-mediated 40S dissociation*. Scanning ribosomes that collide with a 3' stalled ribosome dissociate from the mRNA. (E) *Regulated re-initiation*. Ribosomes initiate translation at the first uORF start codon, and scanning continues after termination at the stop codon of the first uORF. Ribosomes re-initiate at the main ORF start codon or the second downstream uORF start codon when phosphorylated eIF2 $\alpha$  levels are high or low, respectively. The schematic is depicted in a low phosphorylated eIF2 $\alpha$  state.

<https://doi.org/10.1371/journal.pgen.1010460.g001>

eIF2 $\alpha$ . Therefore, when cells are not stressed and the proportion of phosphorylated eIF2 $\alpha$  is lower, translation of the first uORF is followed by re-initiation at the second downstream uORF start codon. Alternatively, when cells are stressed and the proportion of phosphorylated eIF2 $\alpha$  is higher, translation of the first uORF is instead followed by re-initiation at the main ORF start codon.

## Experimental system for testing different models of uORF-mediated translational regulation

To differentiate between proposed models of uORF regulation (Fig 1), we used the well-studied human cytomegaloviral *ULA* uORF2 [31] as an experimental model (Fig 2A). uORF2 represses main ORF translation via an elongating ribosome stall that is dependent on the uORF2 peptide sequence [21] (Fig 2A, irrelevant uORFs boxed in white, key uORF2 boxed in green). The two C-terminal proline residues, regardless of codon usage, in uORF2 are necessary for the elongating ribosome stall [32]. These residues are poor substrates for nucleophilic attack to generate a peptide bond and also reorient the ribosomal peptidyl transferase center to reduce termination activity [28]. Termination activity is further reduced through interactions between the uORF2 nascent peptide and the GGQ motif within eRF1 [46]. Even though the A-site of the uORF2-stalled ribosome is occupied by a stop codon, we refer to it as an elongating ribosome stall since they are functionally equivalent for the purposes of this study. This terminology is also inclusive of elongation stalls within other uORFs [22,23,47–51]. The 5' leader region preceding the *ULA* coding sequence contains two other uORFs besides uORF2. uORF1



**Fig 2. An experimental and computational platform for assessing uORF-mediated regulation of main ORF translation.** (A) The 236 nt 5' UTR of *UL4* mRNA from human cytomegalovirus contains 3 uORFs. The terminal proline and stop codons of uORF2 at which the P- and A-sites of the stalled ribosome are positioned are highlighted as uORF2 stall. (B) A dual-luciferase reporter system for measuring 5' UTR repressiveness in HEK293T cells. FLuc signal serves as an internal control for transfection efficiency. (C) The reporter system recapitulates the known elongating ribosome stall-dependent repression of protein expression by the *UL4* uORF2 [21]. The indicated mutations improve the uORF2 Kozak context (ACCATGG instead of GTGATGC), remove the start codon (ACC instead of ATG), or remove the elongating ribosome stall by mutating the terminal proline codon to an alanine codon (GCT instead of CCT). Error bars show standard error of mean NLuc / FLuc ratios over 3 biological replicates. Data are normalized to the no-uORF start codon control. (D) Computationally predicted uORF regulation in the 80S-hit dissociation, queuing-mediated enhanced repression, and collision-mediated 40S dissociation models. Data are normalized to the no-uORF start codon control. The parameter combination that best recapitulated the control behavior in Fig 2C is displayed in Table 1. Error bars for simulated data are smaller than data markers.

<https://doi.org/10.1371/journal.pgen.1010460.g002>

slightly reduces uORF2 repressiveness by siphoning scanning ribosomes away from uORF2, and uORF3 is irrelevant for repression [31].

We inserted the uORF2-containing *UL4* leader sequence into a dual-luciferase reporter system (Fig 2B) in which nanoluciferase (NLuc) signal provides a readout of uORF2 repressiveness and firefly luciferase (FLuc) signal serves as normalization for transfection efficiency. This experimental platform can detect differences in luciferase activity over a 1000-fold range (S1 Fig). We confirmed that uORF2 repressiveness depends on its translation and the terminal diproline-dependent elongating ribosome stall (Fig 2C). Near-cognate start codons within uORF2 do not contribute to the uORF2 repressiveness (S1 Fig). We used this *UL4*-based luciferase reporter to quantitatively dissect the kinetics of uORF2-mediated translational regulation.

We complemented our experimental measurements with computational kinetic modeling of proposed models of uORF regulation (Fig 1). We aimed to find unique modeling predictions that would allow us to experimentally distinguish between the different models of uORF regulation. We specified the kinetics of each of the proposed models of uORF regulation using PySB, a framework for compact specification of rule-based models [58]. We then expanded the model into the BioNetGen modeling language syntax [59] and inferred a reaction

dependency graph for efficient simulation [39]. Next, we stochastically simulated the models using an agent-based Gillespie algorithm implemented in NFSim [60]. The molecules and reactions within the kinetic model are shown in Fig 3A and 3B, respectively, and are described in detail in the Materials and Methods section. We experimentally tested predictions from this computational modeling and used the results to refine our model specifications. This iterative cycle of experimental testing and computational modeling constituted our platform for differentiating between proposed uORF regulatory models.

To derive estimates for unknown parameters ('This work' in Table 1), we first calibrated our computational models to our reporter measurements on wild-type or mutant uORF2 (Fig 2C). We did not fit the constitutive repression and regulated re-initiation models (Fig 1A and 1E) to our reporter measurements (Fig 2C) since these models cannot account for the critical role of the *UL4* uORF2 elongating ribosome stall in regulating main ORF translation in single uORF-containing mRNAs. We used previously generated estimates for kinetic parameters not directly inferred in our work (Table 1).





Simulations of the queuing-mediated enhanced repression (Fig 1C) and collision-mediated 40S dissociation (Fig 1D) models readily recapitulate measurements of NLuc protein output from wild-type and mutant *UL4* reporters (Fig 2D, triangles and squares). The 80S-hit dissociation model (Fig 1B), modified to include an elongating ribosome stall within the uORF, also recapitulates the reporter measurements (Fig 2D, circles). However, this modified 80S-hit dissociation model requires the difference between the stronger Kozak and wild-type uORF initiation fractions to be quite large (80% vs. 2% compared to 50% vs. 10% for 2 other models mentioned above, Table 1). The derived ribosome loading rates ( $\sim 0.02/s$  for all three of these models (Fig 1B–1D)) are in line with literature estimates [52–54]. The re-initiation fractions derived here (50–70%, Table 1) are within the range of measured values across mRNAs with different sequence features [40–43]. A complete description of the derivation of model parameters can be found in the Materials and Methods section.

### Computational modeling predicts that different models of uORF regulation have unique parameters important for buffering

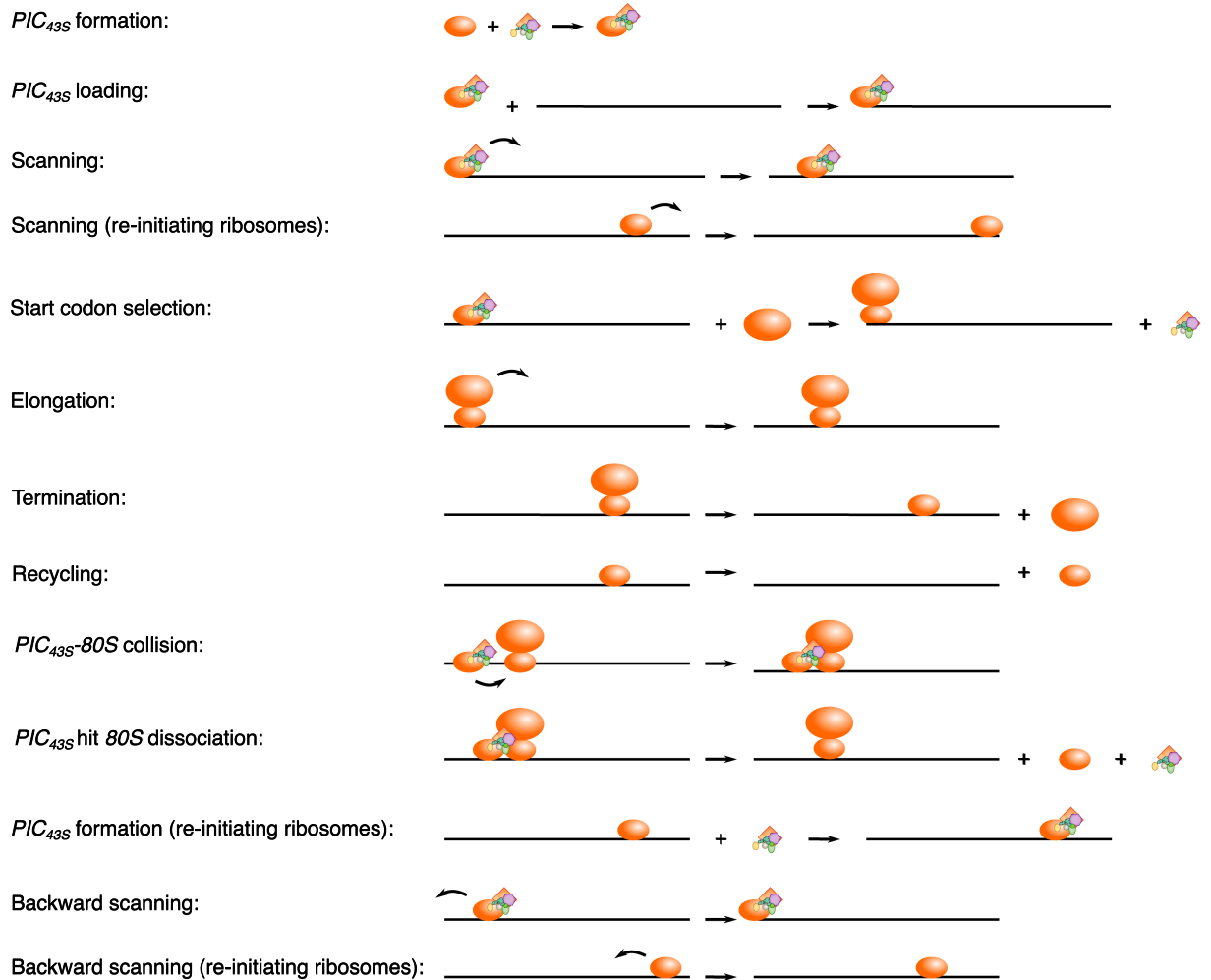
Following calibration of our computational models to recapitulate experimental data, we used these models to predict how translation would be perturbed upon varying other kinetic parameters. While many kinetic parameters could be varied to help distinguish between proposed models of uORF regulation (Fig 1), we honed in on the rate of ribosome loading onto the mRNA for two key reasons. Firstly, this rate is reduced endogenously in response to a variety of cellular and environmental signals. Amino acid deprivation, ribosome collisions, dsRNA viral infection, unfolded proteins, and heme deprivation are sensed by one of the four eIF2 $\alpha$  kinases (GCN2, PKR, PERK, and HRI) to reduce TC concentration [61–64]. A reduction in the concentration of eIF2 $\alpha$ -containing TCs reduces the rate of ribosome loading. Viral infection also leads to reduced ribosome loading via interferon-induced proteins with tetratricopeptide repeats (IFITs) [65]. Cellular stress also reduces ribosome loading via inhibition of mTOR and sequestration of eIF4E by hypophosphorylated 4EBP [66]. Secondly, translated repressive uORFs are enriched in transcripts buffered against reduced ribosome loading [67–70]. Therefore, we were particularly interested in varying this ribosome loading rate to investigate if and how uORFs provide this buffering across various proposed models. For each of the five surveyed models of uORF regulation (Fig 1), we investigated what uORF parameter combinations, if any, allow buffering against reduced ribosome loading rates.

We use the term 'buffer' to describe the observation of main ORF protein output decreasing less than expected, or even increasing, with reduced ribosome loading in comparison to the

**A. Molecules in the kinetic model**

| Molecule                              | Component   | State values or bond partners  |
|---------------------------------------|---|--|
| Small ribosomal subunit ( $R_{40S}$ ) | <i>isbi tc</i><br><br><i>t</i> <i>a</i> <i>l</i> | bond to large ribosomal subunit<br>bond to TC<br>bond to trailing ribosome<br>bond to leading ribosome<br>bond to mRNA codon at position $i$ ( $c_i$ ) |
| Large ribosomal subunit ( $R_{60S}$ ) | <br><i>isbi</i>                                  | bond to small ribosomal subunit  |
| Ternary complex (TC)                  | <br><i>ssu</i>                                   | bond to small ribosomal subunit  |
| mRNA ( $M$ )                          |    | initiation footprint ( <i>start</i> )<br>codon at position $i$ ( $c_i$ )   |

**B. Reactions in the kinetic model**



**Fig 3. Modeling workflow.** (A) Molecules have components each of which has state values or binding sites for other molecules (called *bond* in BioNetGen). Components are abbreviated in parentheses by how they are referenced in the model specification. For example, the mRNA ( $M$ ) initiation footprint ( $c_1$  to  $c_n$ , where  $n$  is equal to the ribosome footprint size in nt) can either be *clear* of ribosomes, and therefore free for a  $PIC_{43S}$  loading reaction to occur, or *blocked* by a ribosome, preventing this reaction. (B) Visual representations of the reactions in the kinetic model. Re-initiation necessitates several additional reactions.  $PIC_{43S}$  formation ( $R_{40S}$  binding TC) can occur if the  $R_{40S}$  is bound to the mRNA; this TC re-binding is required for start codon selection competence.  $R_{40S}$  molecules can scan forward or backward. Some reactions in the kinetic model, such as different types of collision and dissociation reactions, are not depicted here.

<https://doi.org/10.1371/journal.pgen.1010460.g003>

**Table 1. Parameter ranges and fit values for modeling.**

| Parameter   | Value range | Fit value (80S-hit dissociation) | Fit value (queuing-mediated enhanced repression) | Fit value (collision-mediated 40S dissociation) | Reference         |
|---|-------------|----------------------------------|--|---|-------------------|
| $k_{cap\ bind}$ ( $s^{-1}$ )                      | 0.02–0.06   | 0.016                            | 0.023  | 0.025   | This work [52–54] |
| $k_{scan}$ (nt/s)                                 | 1–10        | 5                                | 5  | 5   | [38,55]           |
| $k_{start\ uORF2\ WT}$ ( $s^{-1}$ )               | unknown     | 0.1                              | 0.5  | 0.5   | This work         |
| WT uORF2 initiation (%)                           | unknown     | 2                                | 10   | 10  | This work         |
| $k_{start\ uORF2\ strong\ Kozak}$ ( $s^{-1}$ )    | unknown     | 20                               | 5  | 5   | This work         |
| strong Kozak uORF2 initiation (%)                 | unknown     | 80                               | 50   | 50  | This work         |
| $k_{elong}$ (codons/s)                            | 3–10        | 5                                | 5  | 5   | [52–54,56]        |
| $k_{elong\ stall}$ (codons/s)                     | 0.001       | 0.001                            | 0.001  | 0.001   | [57]              |
| $k_{terminate}$ ( $s^{-1}$ )                      | 0.5–5       | 1                                | 1  | 1   | [56]              |
| $k_{terminated\ ssu\ recycle\ uORF}$ ( $s^{-1}$ ) | unknown     | 2                                | 5  | 5   | This work         |
| Re-initiation (%)                                 | unknown     | 75                               | 50   | 50  | This work         |
| $k_{dissociate}$ ( $s^{-1}$ )                     | unknown     | 2                                | 0  | 2   | This work         |
| uORF length (codons)                              | 21          | 21                               | 21   | 21  | [32]              |

<https://doi.org/10.1371/journal.pgen.1010460.t001>

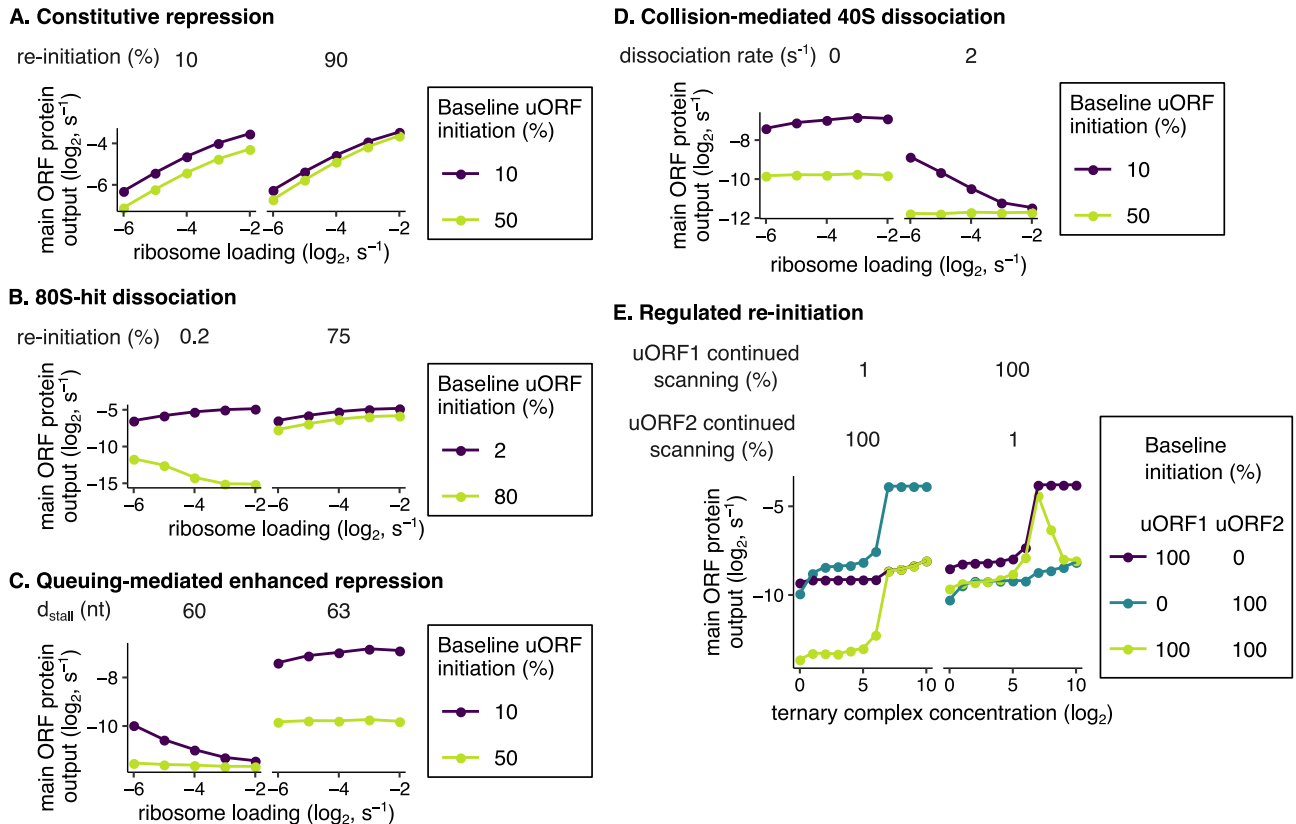
constitutive repression model (Fig 1A). The constitutive repression model (Fig 1A) has no buffering (Fig 4A) since its repression is independent of the ribosome loading rate. Buffering requires an interaction between ribosome loading and the degree of translational repression. We use buffering as an overarching term that encompasses both resistance and preferred translation. Resistance refers to a decrease in main ORF protein output to a lower extent than in the constitutive repression model when ribosome loading is reduced. Preferred translation refers to increased main ORF protein output when ribosome loading is reduced.

The 80S-hit dissociation model (Fig 1B) displays buffering (Fig 4B, left panel, yellow-green line) in agreement with previous work [38]. This behavior arises because the number of 5' elongating ribosomes that collide with scanning ribosomes correlates with the ribosome loading rate. However, buffering requires strong uORF initiation, minimal re-initiation, and a long uORF (Fig 4B, left panel, yellow-green line, S2A Fig) as observed previously [38]. These observations can be rationalized as follows. Strong uORF initiation generates sufficient elongating ribosomes that hit and knock off 3' scanning ribosomes. Minimal re-initiation prevents the many uORF-translating ribosomes from also translating the main ORF. Longer uORFs offer more time for elongating ribosomes to catch up, hit, and knock off 3' scanning ribosomes. However, most eukaryotic uORFs only weakly initiate translation and are short [1,3,4,8–11,31]. *UL4* uORF2 is 22 codons long, and we estimate re-initiation to be frequent (Table 1). Accordingly, buffering is no longer observed (S2B Fig) in the 80S-hit dissociation model when parameters (Table 1) derived from control *UL4* variants (Fig 2C) are used.

The queuing-mediated enhanced repression model [23] (Fig 1C) displays buffering behavior (Fig 4C, left panel, purple line) since the number of scanning ribosomes that initiate translation at the uORF is dependent on the rate of ribosome loading. In this model, reduced ribosome loading decreases the average queue length of ribosomes behind the elongation stall and, thus, also the fraction of ribosomes that initiate at the uORF2 start codon (S2C Fig, left). Unlike the 80S-hit dissociation model (Fig 1B), weakly initiating uORFs, such as *UL4* uORF2, still confer buffering in the queuing-mediated enhanced repression model (Fig 4C, left panel, purple line).

In the queuing-mediated enhanced repression model, enhanced uORF initiation and, therefore, buffering are sensitive to the distance between the uORF start codon and the





**Fig 4. Kinetic modeling predicts translational buffering by uORFs.** Buffering refers to a smaller than expected decrease (small positive slope), or even increase (negative slope), in main ORF protein output with reduced ribosome loading. (A) The constitutive repression model, without an elongating ribosome stall, has no buffering behavior. uORFs simply siphon away scanning ribosomes from the main ORF. (B) Buffering in the 80S-hit dissociation model depends on uORF initiation and re-initiation frequencies [38]. For buffering to occur in this model, uORFs must initiate well enough to have elongating ribosomes hit 3' scanning ribosomes (yellow-green line). uORFs must also not continue scanning at high frequencies following termination (left panel); frequent continuation of scanning coupled with high uORF initiation allows many scanning ribosomes to make it to the main ORF. Buffering occurs better for longer uORFs that have more time for elongating ribosomes to hit 3' scanning ribosomes (S2A Fig, yellow-green line). Here, the uORF is 100 codons long. The dissociation rate is 200s<sup>-1</sup>, so 99% of scanning ribosomes hit by 5' elongating ribosomes dissociate rather than continue scanning. The scanning rate is 2 nucleotides/s, and the elongation rate is 2 codons/s. There is no elongating ribosome stall in this model. (C) Buffering in the queuing-mediated enhanced repression model depends on *d<sub>stall</sub>*: the distance between the uORF start codon and elongating ribosome stall. In this model, uORF initiation can increase above baseline with increased ribosome loading when *d<sub>stall</sub>* is an integer multiple of the ribosome footprint (30 nt, left panel). When this condition is met, buffering occurs. For *d<sub>stall</sub>* values of 60 and 63 nt, the uORF length is 21 and 22 codons, respectively. (D) Buffering in the collision-mediated 40S dissociation model depends on the dissociation rate. Here, *d<sub>stall</sub>* is 63 nt; with a low dissociation rate, this model reduces to the queuing-mediated enhanced repression model. (E) Buffering in the regulated re-initiation model depends on uORF initiation and continued scanning frequencies. For buffering to occur, several conditions must be met. At least 2 uORFs are required, both of which must be well-translated (yellow-green line). Continued scanning following termination at the first uORF must be frequent, and continued scanning following termination at the second downstream uORF must be rare (right panel). The second downstream uORF is 3 codons long. There is no elongating ribosome stall in this model. uORFs are located 25 nt from the 5' cap. 99% of scanning ribosomes that make it to the main ORF will initiate translation; 1% will leaky scan. Unless otherwise stated, parameters (Table 1) obtained from calibrating models to reporter measurements on wild-type or mutant uORF2 (Fig 2C) are used here. Ribosome loading is the *k<sub>cap bind</sub>* rate for non-regulated re-initiation models. We model changes in ribosome loading via changes in *k<sub>cap bind</sub>* as that rate is easier to match to *in vivo* estimates of ribosome loading. However, buffering in the regulated re-initiation model is dependent on an eIF2α phosphorylation mechanism; we instead vary the number of ternary complexes in this model. Error bars of simulated data are smaller than data markers.

<https://doi.org/10.1371/journal.pgen.1010460.g004>

elongating ribosome stall (*d<sub>stall</sub>*) (S2C Fig and Fig 4C, left vs. right panels). This sensitivity arises because *d<sub>stall</sub>* determines if the P-site of a queued scanning ribosome is correctly positioned at the uORF start codon to productively increase uORF initiation (S2C Fig, left). In the idealized case of homogeneously sized ribosomes (30 nt footprints [56,71]) and strict 5'-3' scanning, *d<sub>stall</sub>* must be an integer multiple of 30 nt for buffering to occur. This strong

dependence of buffering on  $d_{stall}$  is relaxed when backward scanning [41,72–74] occurs with a high rate (S2D Fig, middle). To simplify our modeling interpretations, we considered *UL4* uORF2, that is 22 codons long, to be 21 codons so that a queue behind the terminating ribosome stall positions a scanning ribosome's P-site exactly at the start codon.

The collision-mediated 40S dissociation model (Fig 1D) displays buffering (Fig 4D, right panel, purple line) because the number of scanning ribosomes that collide with 3' stalled ribosomes depends on the rate of ribosome loading. Buffering in this model requires the collision-induced 40S dissociation rate to be somewhat fast (Fig 4D, right vs. left panels, S2E Fig, teal and yellow-green lines). If this rate is too low (for example, 0 in Fig 4D, left panel), this model reduces to the queuing-mediated enhanced repression model (Fig 1C). With an appreciable dissociation rate, the collision-mediated 40S dissociation model is not sensitive to the distance between the stall and the start codon (S2F Fig, purple lines). As in the queuing model (Fig 1C), weakly initiating uORFs, such as *UL4* uORF2, can still confer buffering (Fig 4D, right panel, purple line) in the collision-mediated 40S dissociation model (Fig 1D). This effect arises because, unlike in the 80S-hit dissociation model (Fig 1B), the elongation stall is now rate-limiting for main ORF translation. Therefore, an elongation stall in the collision-mediated 40S dissociation model or in the queuing-mediated enhanced repression model with permissive  $d_{stall}$  spacing imparts buffering.

In the regulated re-initiation model (Fig 1E), buffering is observed (Fig 4E, right panel, yellow-green line) because termination at the first uORF stop codon is followed by re-initiation at either the second downstream uORF or the main ORF depending on the ternary complex concentration. Buffering in the regulated re-initiation model (Fig 1E) depends on the initiation efficiency and continued scanning fraction (fraction of terminating but non-recycling ribosomes) of the two uORFs (Fig 4E). Continued scanning following termination at the first uORF must be frequent while continued scanning following termination at the second downstream uORF must be rare. Higher ternary complex concentrations bias towards initiation at the second downstream uORF (S3A Fig). Reductions in ternary complex concentrations bias towards main ORF initiation; therefore main ORF translation can increase with decreased ribosome loading.

As such, our computational results provide the first systematic dissection of different mechanisms of uORF-mediated regulation (Fig 1) and enable their comparison with experimental measurements below.

### ***UL4* uORF2 buffers against reductions in main ORF protein output from reduced ribosome loading in an elongating ribosome stall-dependent manner**

We next tested whether the computational predictions of uORF-mediated buffering (Fig 4) can be observed experimentally with *UL4* uORF2. To this end, we experimentally varied the rate of ribosome loading and measured effects on main ORF protein output using our reporter system (Fig 2B). Since the no-stall uORF2 variants have similar protein expression to the no-start uORF2 variants (Fig 2C), luciferase signal from the no-stall uORF2 variants provides a readout of the ribosome loading rate. If buffering were absent, then we would expect NLuc translation to be reduced equally between the no-stall and wild-type variants when ribosome loading is reduced.

We used three strategies to vary the rate of ribosome loading. We first used stem-loops near the 5' cap to reduce the rate of 43S-cap binding without affecting mRNA stability (Fig 5A) [75,76]. We varied the degree to which ribosome loading is reduced by altering the GC content of the stem-loops; generally, higher GC content stem-loops are more stable and therefore

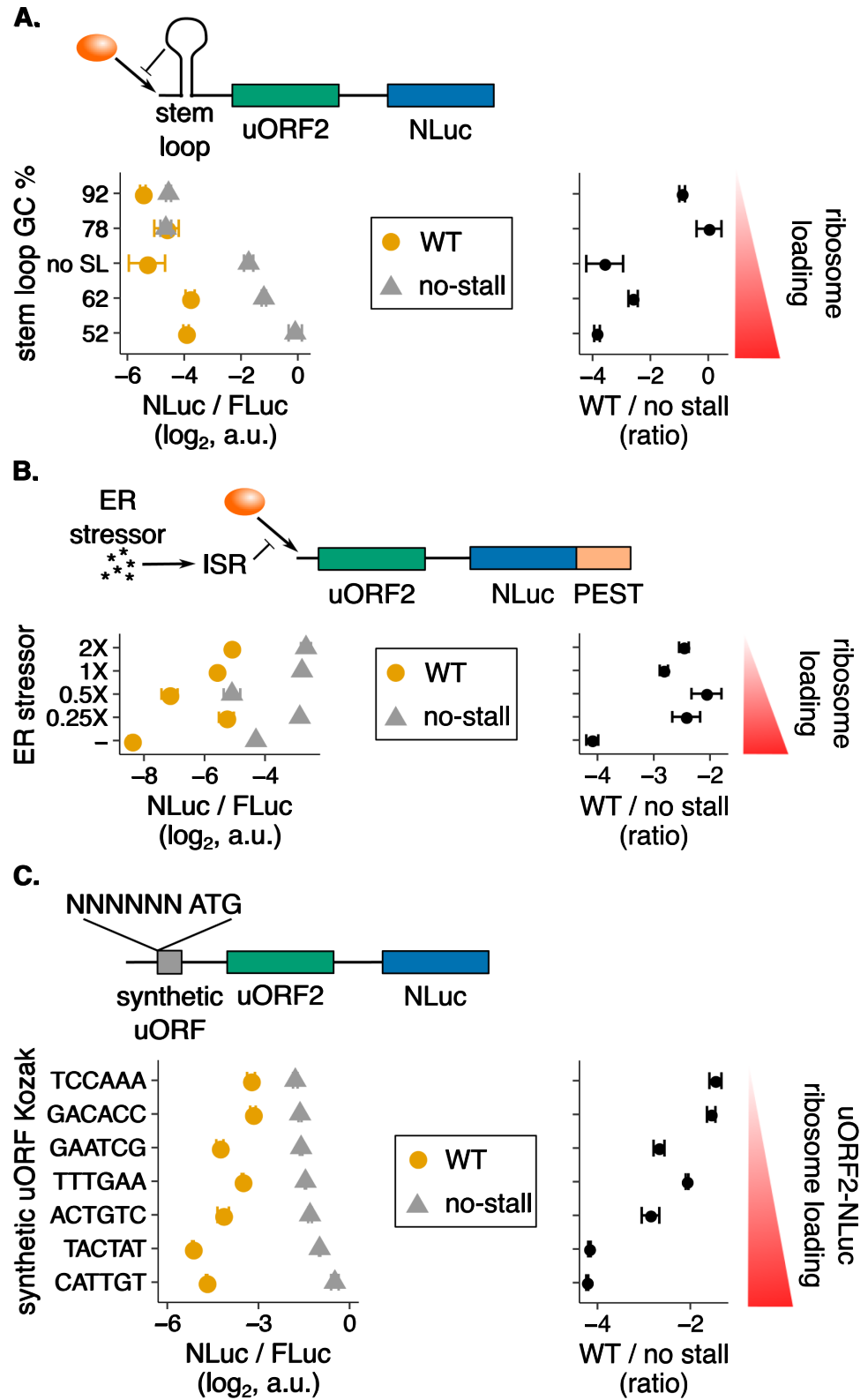
cause greater reductions in ribosome loading. We observe that NLuc signal decreases less with reduced ribosome loading for the wild-type *UL4* reporter in comparison to the no-stall *UL4* variant (Fig 5A, left panel, yellow vs. gray circles). Therefore, NLuc protein output is resistant to stem-loop-mediated reduction in ribosome loading. When the wild-type data are normalized by the no-stall data, NLuc translation negatively correlates with ribosome loading, indicative of buffering against reduced ribosome loading by wild-type uORF2 (Fig 5A, right panel).

We also reduced ribosome loading with the drug thapsigargin, which induces the integrated stress response (ISR) by triggering ER stress (Fig 5B) [79]. We added a PEST tag [78] to increase the turnover of the NLuc protein to more accurately measure changes in main ORF translation during drug treatment. NLuc protein output from the wild-type *UL4* reporter decreases less in comparison to the no-stall control upon thapsigargin treatment (Fig 5B, left panel, yellow vs. gray circles), indicative of resistance. Again, when the wild-type data are normalized by the no-stall data, we observe that NLuc translation negatively correlates with ribosome loading, indicative of buffering against reduced ribosome loading by wild-type uORF2 (Fig 5B, right panel).

Finally, we added a short, synthetic uORF, 5' to the *UL4* uORF2, to siphon scanning ribosomes away from uORF2 (Fig 5C). We varied the degree of ribosome siphoning by varying the Kozak context of the synthetic uORF, which in turn determines the rate of ribosome loading onto the uORF2-NLuc portion of the mRNA. Here, we observe that more NLuc is produced from the wild-type *UL4* reporter as scanning ribosomes are increasingly siphoned off by improving the Kozak context of the synthetic uORF (Fig 5C, left panel, yellow circles). While resistance is observed with the other strategies of reduced ribosome loading (Fig 5A and 5B, left panels), preferred translation is observed here (Fig 5C, left panel), perhaps because ribosome loading is reduced at the scanning step instead of at the cap-binding step. Similar to the other two strategies, when the wild-type data are normalized by the no-stall data, NLuc translation negatively correlates with ribosome loading, indicative of buffering against reduced ribosome loading by wild-type uORF2 (Fig 5C, right panel).

### Distance between the start codon and the stall does not systematically regulate uORF repressiveness or buffering

Given our experimental data demonstrating uORF2-mediated buffering of *UL4* reporters (Fig 5), we narrowed our focus from the five surveyed models (Fig 1) to the two (Fig 1C and 1D) most relevant for *UL4* uORF2: the queuing-mediated enhanced repression (Fig 1C) and collision-mediated 40S dissociation (Fig 1D) models. These two models are computationally predicted to confer buffering in an elongating ribosome stall-dependent manner without needing multiple uORFs (Fig 4C and 4D). To differentiate between these models, we turned to our computational modeling prediction that, only in the queuing-mediated enhanced repression model (Fig 1C), main ORF protein output is sensitive to the distance between the uORF start codon and the elongating ribosome stall (Fig 4C). Our computational modeling of the queuing-mediated enhanced repression model (Fig 6A, yellow-green line) predicts two broadly spaced clusters of main ORF protein output. Protein output from the main ORF is repressed when the start codon-stall distance is an integer multiple of the ribosome size. Protein output from the main ORF is high when the start codon-stall distance is not an integer multiple of the ribosome size. In contrast, the collision-mediated 40S dissociation model (Fig 1D) predicts a much lower effect of  $d_{stall}$  on uORF repressiveness (Fig 6A, left panel, purple line). The residual effect of  $d_{stall}$  on uORF repressiveness (Fig 6A, left panel, purple line) in the collision-mediated 40S dissociation model (Fig 1D) arises because the dissociation rate is low enough to allow rare queuing.



**Fig 5. The human cytomegaloviral uORF2 buffers against reductions in main ORF protein output.** The human cytomegaloviral *UL4* uORF2 is used in the dual-luciferase assay (Fig 2B) in conjunction with three experimental strategies to reduce ribosome loading. (A) Ribosome loading is reduced using stem-loops [76] with the indicated GC percentages. All stem-loops are positioned 8 nt from the 5' cap and have the same predicted stability of -30 kcal/mol.

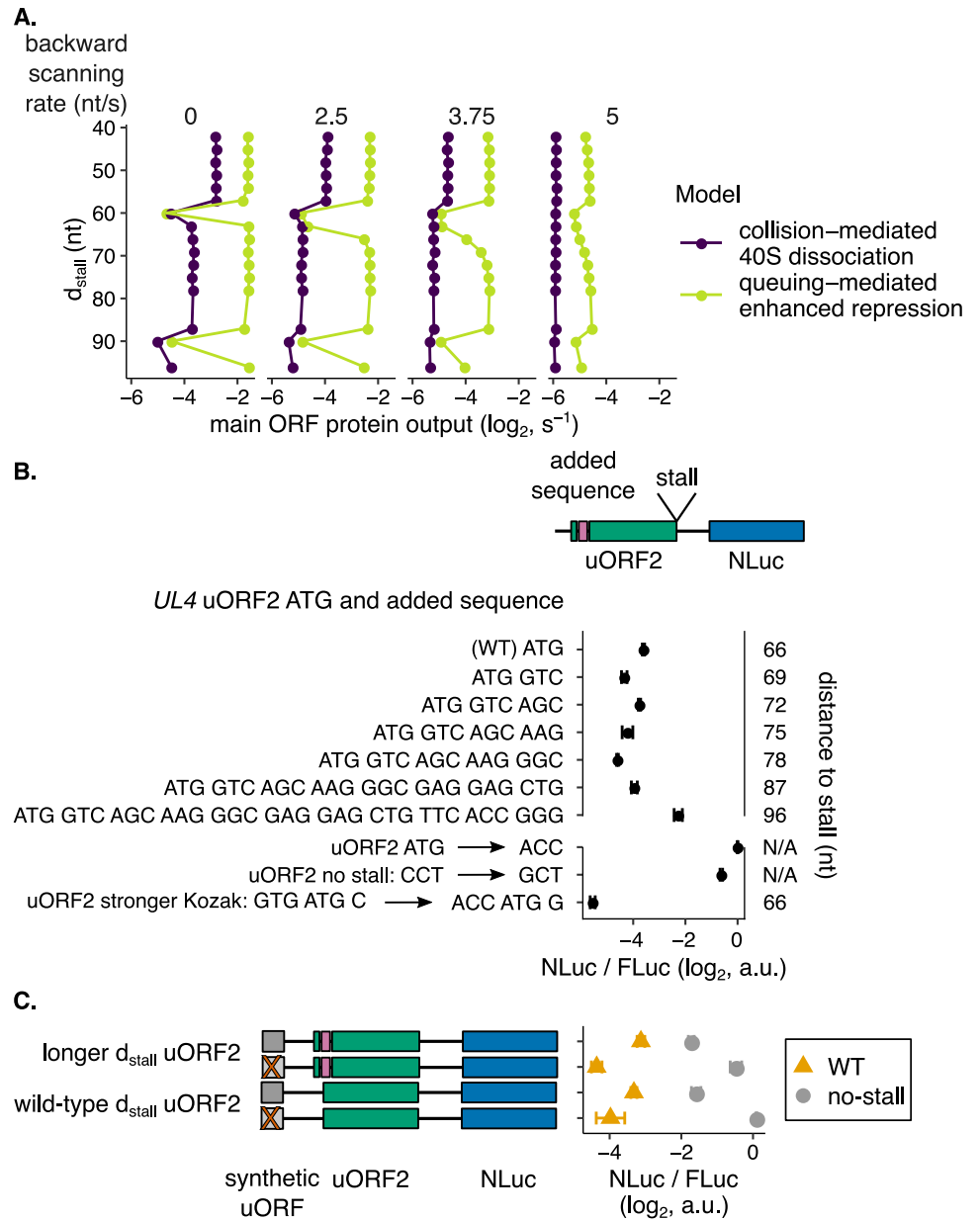
The no-stem-loop construct has a CAA repeat instead of a stem-loop. The 5' UTR is 287 nt long. Data are normalized to a no-uORF start codon control without a stem-loop. (B) Ribosome loading is reduced using the drug thapsigargin (1X = 1  $\mu$ M) [77], which induces the integrated stress response (ISR) by triggering ER stress. NLuc has a C-terminal PEST tag to turnover [78] of protein produced prior to the 6-hour drug treatment. The 5' UTR is 236 nt long. Data are normalized to a no-uORF start codon control without a PEST tag. Error bars show standard error of mean NLuc / FLuc ratios over 4 biological replicates. (C) Ribosome loading onto the uORF2-NLuc portion of the transcript is reduced using a 5' synthetic uORF: ATG GGG TAG. The synthetic uORF Kozak is varied to alter ribosome loading. The variants are vertically ordered by the no-stall means. The 5' UTR is 262 nt long. Data are normalized to a no-uORF start codon control without a synthetic uORF. Right panels in A,B, C show wild-type (WT) mean values normalized by the corresponding no-stall values. The no-stall uORF2 mutants lack their terminal diproline motifs (P22A mutation). Unless stated otherwise, error bars show standard error of mean NLuc / FLuc ratios over 3 biological replicates.

<https://doi.org/10.1371/journal.pgen.1010460.g005>

Backward scanning is predicted to diminish the periodicity in main ORF translation with varying  $d_{stall}$  lengths in both models (Fig 6A). However, backward scanning occurring as fast as forward scanning ( $\sim 5$  nucleotides/s) is required to abolish the periodicity in the queuing model (Fig 6A, right panel, yellow-green line). While there are estimates of how far ribosomes can backward scan [72–74], we are not aware of any backward scanning rate estimates. It is unlikely that the rate of backward scanning approaches the rate of forward scanning (5 nucleotides/s here) given the 5'-3' directionality of scanning. Slower backward scanning ( $\sim 3.75$  nucleotides/s) is sufficient to abolish periodicity in the collision-mediated 40S dissociation model (Fig 6A, middle panel, purple line). This effect is not surprising given that the presence of periodicity in the latter model arises from rare queuing behavior. Therefore, our computational predictions of greater periodicity in main ORF translation across varied  $d_{stall}$  in the queuing model hold even with backward scanning.

We then experimentally varied the distance between the start codon and the stall of the *UL4* uORF by adding codons to the 5' end of uORF2. With *EYFP* donor sequences, we observe less than 2-fold changes in translational regulation (Fig 6B, top 7 rows) with no systematic trend with variations in uORF2 length, which is inconsistent with computational modeling predictions of the queuing-mediated enhanced repression model (Fig 6A, left panel). We observe similar results with a different donor sequence (S4 Fig, top 7 rows), confirming the generality of the observed repression with changes in uORF2 length. With both donor sequences, the longest uORF mutants are less repressive, but this effect may be due to decreased elongating ribosome stall strength. In these cases, the longer nascent peptides can extend out of the exit tunnel and can be bound by additional factors [27,28] or cotranslationally fold to exert a pulling force [80] to relieve the stall. Thus, in summary, varying the length of the *UL4* uORF2 stall does not match computational predictions for sensitivity of main ORF repression to  $d_{stall}$  in the queuing-mediated enhanced repression model (Fig 1C) and better supports the collision-mediated 40S dissociation model (Fig 1D).

In the queuing-mediated enhanced repression model (Fig 1C), buffering is uniquely predicted to be sensitive to the distance between the uORF start codon and the elongating ribosome stall (Fig 4C). We, therefore, asked whether or not buffering would still be experimentally observed with a disruption in this distance. Using our synthetic uORF method of reducing ribosome loading (Fig 6C), we observe that a 6 nt longer  $d_{stall}$  uORF still buffers against reduced ribosome loading (Fig 6C, top two rows compared to bottom two rows). Since backward scanning of 15–17 nt has been observed [72–74], one would expect that buffering would still be predicted in the queuing model even with an increase in  $d_{stall}$  of 6 nt. However, our computational modeling predicts that even very fast backward scanning does not restore buffering when  $d_{stall}$  is disrupted by 6 nt (S2D Fig, right). Thus, our experimental data does not match computational predictions of buffering sensitivity to  $d_{stall}$  in the queuing-mediated



**Fig 6. Changes to the distance between the human cytomegaloviral uORF2 start codon and the elongating ribosome stall do not change repressiveness or buffering, consistent with the collision-mediated 40S dissociation model.** (A) Computational modeling predicts greater changes in uORF repressiveness with changes in  $d_{stall}$  in the queuing-mediated enhanced repression model. Fast backward scanning abolishes this periodicity.  $d_{stall}$  refers to the distance between the start codon and the elongating ribosome stall. As backward scanning increases in rate (moving right along panels), the collision-mediated enhanced repression model loses periodicity (middle panel, purple line) before the queuing-mediated enhanced repression model (right panel, yellow-green line). Parameters that best recapitulated reporter measurements on wild-type or mutant uORF2 (Fig 2C and Table 1) are used here. The forward scanning rate is 5 nucleotides/s. Data are normalized to a no-uORF start codon control. Error bars of simulated data are smaller than data markers. (B) Experimentally varying the distance between the human cytomegaloviral uORF2 start codon and the elongating ribosome stall does not systematically affect its repression of main ORF protein output. The human cytomegaloviral *UL4* uORF2 is used in the dual-luciferase assay (Fig 2B) in conjunction with various length inserts from the N-terminus of the *EYFP* main ORF. The *EYFP* main ORF sequence is inserted directly 3' to the uORF2 start codon. The added sequence increases the distance between the uORF2 start codon and the elongating ribosome stall. The bottom three controls improve the uORF2 Kozak context, remove the start codon, and remove the elongating ribosome stall. Error bars show standard error of mean NLuc / FLuc ratios over 3 biological replicates. Data are normalized to the no-uORF start codon control. (C) Experimentally varying the human cytomegaloviral uORF2

$d_{stall}$  does not strongly regulate the capacity of buffering against reductions in main ORF protein output. Ribosome loading is reduced with a 5' synthetic uORF: ATG GGG TAG. The no-stall uORF2 mutants lack their terminal diproline motifs (P22A mutation). No-synthetic uORF mutants (ATG to AAG) are depicted by transparent, gray bars with red Xs and have a higher relative ribosome loading rate onto the uORF2-NLuc portion of the transcript. The distance between the uORF2 start codon and the elongating ribosome stall is varied as indicated by adding 6 nt, GTC AGC, from the N-terminus of the *EYFP* main ORF. Data are normalized to a no-uORF start codon control without a synthetic uORF.

<https://doi.org/10.1371/journal.pgen.1010460.g006>

enhanced repression model (Fig 1C) but is consistent with the collision-mediated 40S dissociation model (Fig 1D).

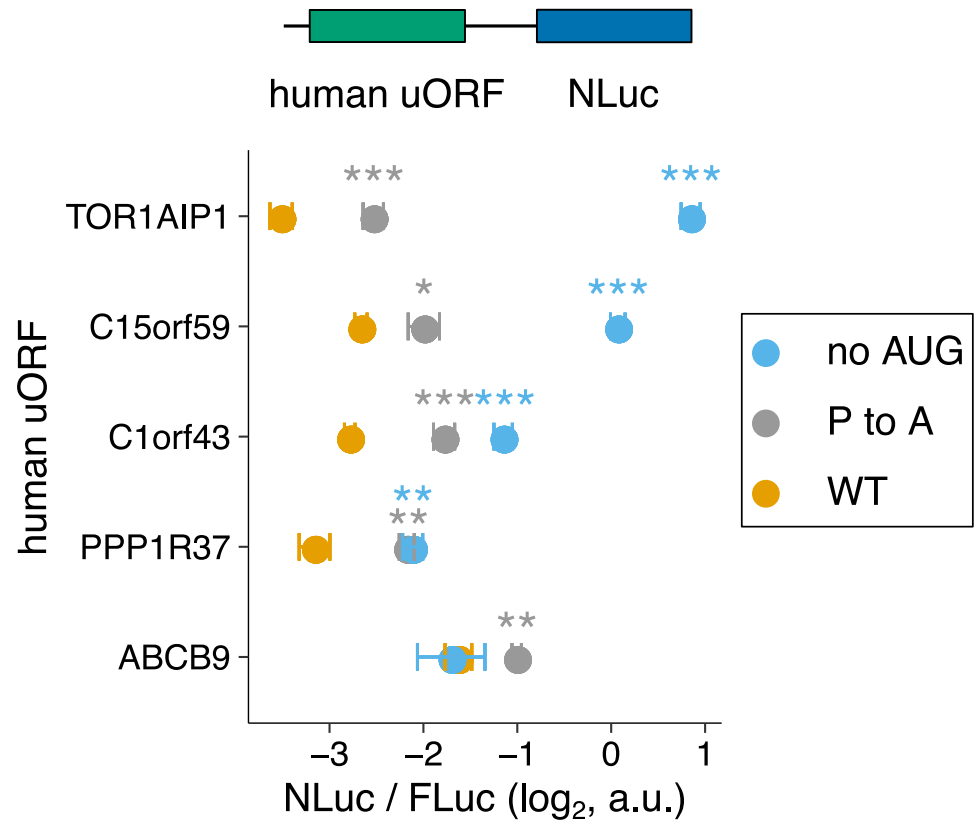
### Several human uORFs have repressive terminal diproline motifs

Given that the elongating ribosome stall in the human cytomegaloviral *UL4* uORF2 is dependent on a terminal diproline motif, we asked whether there are other human uORFs similarly ending in diproline motifs that are also repressive. We searched for such uORFs in three databases: a comprehensive database of ORFs in induced pluripotent stem cells and human foreskin fibroblasts with 1,517 uORFs [6], a database integrated from *de novo* transcriptome assembly and ribosome profiling with 3,577 uORFs [81], and a database of proteins less than 100 residues in size derived from literature mining, ribosome profiling, and mass spectrometry with 1,080 uORFs [82]. We identified several human transcripts with terminal diproline-containing uORFs: *C1orf43*, *C15orf59*, *TOR1AIP1*, *PPP1R37*, and *ABCB9*. We replaced *UL4* uORF2 in our reporter (Fig 2B) with these human uORFs. We mutated the terminal proline codon to alanine codon as well as the start codon of these human uORFs and measured the effects of these mutations on NLuc protein output relative to the wild-type uORFs. While many of the tested uORFs are repressive (Fig 7, yellow vs. blue), unlike the human cytomegaloviral uORF2, these human uORFs still repress NLuc protein output without their terminal diproline motif (Fig 7, gray vs. blue), indicating additional contributions to translational repression from other residues in the nascent peptide and due to siphoning of scanning ribosomes at the start codon.

## Discussion

In this study, we use a combination of computational modeling and experimental reporter measurements to dissect the kinetics of uORF-mediated translational regulation of the *UL4* mRNA of human cytomegalovirus. We find that the elongating ribosome stall in *UL4* uORF2 buffers against reductions in main ORF protein output due to reduced ribosome loading (Fig 4). Using an experimentally-integrated modeling approach, we differentiate between models of regulation that can explain this observation. Our computational framework allows easy specification and efficient simulation of several previously proposed kinetic models of uORF regulation (Fig 1). While uORFs are enriched in stress-resistant transcripts, not all uORFs provide buffering [67]. We can predict which models of uORF regulation allow buffering and which parameters are key for buffering in each model (Fig 4). To our knowledge, our work is the first systematic investigation of what uORF metrics impart buffering in each kinetic model of uORF regulation.

uORFs are generally thought to simply siphon away scanning ribosomes from main ORFs, but this simple behavior in the constitutive repression model (Fig 1A) is not predicted to provide buffering (Fig 4C) [67–70]. Instead, we find that 5' UTRs containing one (or some combination) of the following enable buffering of main ORF translation: scanning ribosome dissociation due to 80S hits from the 5' end (Fig 1B), a single uORF with an elongating



**Fig 7. Several human uORFs have repressive terminal diproline motifs.** Terminal diproline motif-containing human uORFs are used in the dual-luciferase assay (Fig 2B). The terminal proline codon in each uORF is mutated to an alanine codon in the P to A mutant. Start codons are mutated to ACC for the no-AUG mutants. P values comparing the indicated mutants to the wild-type are from a two sample t-test: \* (0.01 < P < 0.05), \*\* (0.001 < P < 0.01), \*\*\* (P < 0.001). Error bars show standard error of mean NLuc / FLuc ratios over 5 biological replicates. Data are normalized to a no-*UL4*-uORF2 start codon control.

<https://doi.org/10.1371/journal.pgen.1010460.g007>

ribosome stall (Fig 1C and 1D), or multiple uORFs acting through the regulated re-initiation model (Fig 1E).

Long, well-initiating uORFs that do not re-initiate well allow buffering (Fig 4B, left panel, yellow-green line, S2 Fig panel A, yellow-green line) in the 80S-hit model (Fig 1B), but these requirements are at odds with the typically short and poorly initiating nature of known uORFs [1,3,4,8–11]. Indeed, when we use parameters specific to *UL4* uORF2 for the 80S-hit model (Table 1), namely that uORF2 initiates poorly, re-initiates well, and is not very long, buffering is no longer predicted (S2B Fig).

Computational predictions from the regulated re-initiation model (Fig 1E) agree (Figs 4E and S3A) with previous work [35,36] showing that buffering requires: 1) two well-translated uORFs and 2) frequent and rare continued scanning after termination at uORFs 1 and 2, respectively. Since 30% of human transcripts contain multiple uORFs, some of these might enable buffering by the regulated re-initiation model. However, about 25% of human transcripts only have one uORF [2] and cannot provide buffering under this model.

We narrowed our focus to the two models (Fig 1C and 1D) that are most pertinent to *UL4* uORF2. Both the queuing-mediated enhanced repression (Fig 1C) and collision-mediated 40S dissociation (Fig 1D) models are predicted to allow buffering (Fig 4C and 4D) with weakly initiating uORFs and elongating ribosome stalls. Both of these models require only a single uORF



for buffering (Fig 4C and 4D). Computational modeling not only predicts this buffering behavior but also allows us to differentiate between these two models. We predict that the queuing-mediated enhanced repression model (Fig 1C) is uniquely sensitive to the distance between the uORF start codon and elongating ribosome stall (Fig 6A, yellow-green line, Fig 4C, purple lines, S2F Fig, purple lines). We experimentally vary this distance and do not find any systematic changes in either main ORF protein output (Fig 6B) or buffering (Fig 6C). Based on our results, we propose that scanning ribosomes dissociate rather than queue when encountering a 3' stalled elongating ribosome on uORF2 of *UL4* mRNA.

Scanning ribosomes have been predicted to dissociate upon encountering stable secondary structures on the mRNA [83]. Collisions between scanning ribosomes and their subsequent dissociation have also been proposed in a model of initiation quality control [84]. This dissociation could serve to maintain the free pool of 40S ribosomal subunits while still allowing regulation of main ORF translation. Collisions between scanning and elongating ribosomes and subsequent quality control are not well understood; what we describe as scanning ribosome dissociation here may be rescued by a quality control pathway.

Although our data from *UL4* uORF2 does not support the queuing-mediated enhanced repression model (Fig 1C) [23], this model might describe translation kinetics on other mRNAs. Translation from near-cognate start codons is resistant to cycloheximide, perhaps due to queuing-mediated enhanced initiation, but sensitive to reductions in ribosome loading [85]. Loss of eIF5A, which helps paused ribosomes continue elongation, increases 5' UTR translation on human mRNAs with pause sites proximal to the start codon, perhaps also through queuing-mediated enhanced initiation [86]. There is also evidence of queuing-enhanced uORF initiation in the 23 nt long *Neurospora crassa* arginine attenuator peptide [87] as well as in transcripts with secondary structure near and 3' to start codons [88]. Additional sequence elements in the mRNA might determine whether scanning ribosome collisions result in queuing or dissociation. Small subunit profiling data [89] from human uORFs that have conserved amino acid-dependent elongating ribosome stalls do not show evidence of scanning ribosome queues (S5A Fig), consistent with the collision-mediated 40S-dissociation model. However, subtle queues might not be observed given low read counts arising from insufficient capture of small ribosomal subunits in these experiments.

In our modeling, we assume homogenous footprint lengths of 30 nt for both scanning and elongating ribosomes. Even though heterogeneously sized footprints have been observed for small ribosomal subunits [89–91] and elongating ribosomes [92,93], our modeling of homogenous footprint length is appropriate for the following reasons. Firstly, with respect to the small ribosomal subunit footprints, crosslinking of associated eIFs is thought to be the main source of length heterogeneity [89,90], and homogenous 30 nt footprints are observed in the absence of crosslinking [90]. Secondly, in the context of the strong, minutes-long *UL4* uORF2 elongating ribosome stall [57], collided ribosomes, if they do not dissociate, will wait for long periods of time in a queue relative to normal scanning or elongating ribosomes, during which associated eIFs likely dissociate [90]. Thirdly, a sizable fraction of mRNAs exhibit cap-tethered translation in which eIFs must dissociate from ribosomes before new cap-binding events, and therefore collisions, can occur [90]. Elongating ribosome footprint heterogeneity is much less drastic than that observed for scanning ribosomes and likely arises from different conformational states such as empty or occupied A sites [92,93]. While different elongating ribosome footprints arise from differences in mRNA accessibility to nucleases, it is unclear whether the distance between two collided ribosomes changes across different ribosome conformations.

In addition to the *UL4* viral uORF studied here, several human uORFs are known to contain an elongating ribosome stall [22,23,26,47]. Apart from terminal diprolines, other motifs such as Arg-X-Lys at E-P-A sites [94] or specific dipeptides such as Gly-Ile, Asp-Ile, Gly-Asp

[95] can also cause elongation stalls. There are a variety of other mechanisms that may reduce the rate of elongation, such as mRNA stem-loops and G-quadruplexes [96], low tRNA availability [97,98], or interactions between the nascent peptide and the ribosome [99,100]. uORFs are often short [3] and may therefore be better poised to stall ribosomes since the nascent peptides might not be accessible to co-translational factors that pull the nascent peptide out of the ribosome [101,102]. Thus, a key role for elongating ribosome stalls in uORFs might be to enable buffering. While few uORF stalls have been mechanistically characterized [1], other elongating ribosome stall-containing uORFs, such as the ones in *MTR* [47] and *AMD1* [103] mRNAs, might enable buffering; the elongating ribosome stall-containing uORFs in *AZIN1* [23], *PPP1R15A* (*GADD34*) [104], and *DDIT3* (*CHOP*) [22] have already been shown to enable buffering. Conversely, uORFs in several single uORF transcripts known to buffer against stress, such as *SLC35A4*, *C19orf48*, and *IFRD1* [67], might act through elongating ribosome stalls.

The computational models considered here can be readily extended to incorporate more complex mechanisms of translational control. For example, in our models, initiation proceeds via a cap-severed mechanism in which multiple scanning ribosomes can be present in the 5' UTR at the same time. If we were to model cap-tethered initiation, strong uORF elongating ribosome stalls would eventually sever this connection, similar to how the cap-eIF-ribosome connection is severed during the usually longer translation of main ORFs [90,105,106]. It will also be interesting to consider the effect of cellular stress-reduced elongation rates [107] and increased re-initiation [108], both of which might regulate uORF-mediated buffering, as well as elongating ribosome dissociation through known quality control pathways [39,84,109–114]. Translation heterogeneity among isogenic mRNAs has been observed in several single-molecule translation studies [33,52–54,115]. This heterogeneity may arise from variability in intrasite RNA modifications [116], RNA binding protein occupancy, or RNA localization. We do not capture these sources of heterogeneity in our modeling since the observables in our simulations are averaged over long simulated time scales and used to predict only bulk experimental measurements. However, the models studied here can readily be extended through compartmentalized and state-dependent reaction rates [59] to account for the different sources of heterogeneity observed in single-molecule studies.

## Materials and methods

### Plasmid construction

The parent cloning vector was created as follows. A commercial vector (Promega pGL3) with ampicillin resistance was used to clone NLuc and FLuc. NLuc expression is driven by a CMV promoter. FLuc expression is driven in the opposite direction within the plasmid and serves as an internal transfection control. The human cytomegaloviral *UL4* 5' UTR was PCR amplified from HCMV genomic DNA. To create mutant 5' UTR versions of the parent pGL3-FLuc-NLuc vector, the vector was digested with KpnI/EcoRI unless otherwise noted. 1 or 2 PCR-amplified fragments with 20–30 bp homology arms were then cloned using isothermal assembly [117]. The stem-loop [76] 5' UTR mutants were cloned as follows. The stem-loops were ordered as oligonucleotides with overhangs for ligation into ClaI and NotI sites. The oligonucleotides were annealed and used in PCR reactions to add CMV homology arms. An AAVS1 parent vector was digested with ClaI and NotI. These stem-loops were then inserted into the ClaI/NotI restriction digested parent vector by isothermal assembly [117]. The stem-loops were then PCR amplified off of this plasmid and inserted into the pGL3-Fluc-*UL4*-5'-UTR-NLuc parent vector described above. The several tested human uORFs were PCR amplified from human genomic DNA and inserted into a PstI/EcoRI digested parent. The inserted sequences were confirmed by Sanger sequencing. Kozak context and stall codon mutations were

introduced in the PCR primers used for amplifying inserts before isothermal assembly. Standard molecular biology procedures were used for all other plasmid cloning steps [118]. S1 Table lists the plasmids described in this study. Key plasmid maps are available at [https://github.com/rasilab/bottorff\\_2022](https://github.com/rasilab/bottorff_2022) as SnapGene.dna files. Plasmids will be sent upon request.

## Cell culture

HEK293T cells were cultured in Dulbecco's modified Eagle medium (DMEM 1X, with 4.5 g/L D-glucose, + L-glutamine,—sodium pyruvate, Gibco 11965–092) and passaged using 0.25% trypsin in EDTA (Gibco 25200–056).

## Dual-luciferase reporter assay

Plasmid constructs were PEI or Lipofectamine 3000 (Invitrogen, L3000-008) transiently transfected into HEK293T cells for 12–16h in 96 well plates. After the 12–16h transfection, the ~110  $\mu$ L media was removed and replaced with 20  $\mu$ L media per well. The Promega dual-luciferase kit was used. Cells were lysed with 20  $\mu$ L ONE-Glo EX Luciferase Assay Reagent per well for three minutes to measure firefly (*Photinus pyralis*) luciferase activity. Then, 20  $\mu$ L NanoDLR Stop & Glo Reagent was added per well for 10 minutes to quench the FLuc signal and provide the furimazine substrate needed to measure NLuc luciferase activity. FLuc activity serves as an internal control for transfection efficiency, and NLuc activity provides a readout of 5' UTR regulation of NLuc translation.

## Kinetic modeling

We specify our kinetic models using the PySB interface [58] to the BioNetGen modeling language [59] (Fig 3). The Python script is parsed by BioNetGen into a .bngl file and converted into an xml file for use as input to the agent-based stochastic simulator Nfsim [60].

## Molecules

Our kinetic models of eukaryotic translational control describe the interactions between 3 molecule types: mRNA, ribosome (composed of separate large and small subunits), and ternary complex. Here, we describe these molecules' components, states, and binding sites (Fig 3A). mRNA molecules have the following components: 5' end and codon sites ( $c_i$ ). The mRNA 5' end can either be free of (*clear*) or occupied with a ribosome (*blocked*). The mRNA 5' end must be clear for a 43S to bind, which leaves the 5' end blocked until the ribosome scans (or elongates) sufficiently 3' downstream. The mRNA codon sites serve as binding sites for the ribosome A site. Small ribosomal subunits have the following components: inter-subunit binding interface (*isbi*), ternary complex contact (*tc*), 5' side (*t* for trailing), 3' side (*l* for leading), and A site (*a*). The inter-subunit binding interface site allows interactions between small and large ribosomal subunits; large ribosomal subunits also have the inter-subunit binding interface (*isbi*) components. The 5' and 3' side sites serve as binding sites for other ribosomes during collisions (5' or 3' side). The A site serves as a binding site for the mRNA. Both scanning and elongating ribosomes have mRNA footprints of 10 codons in our simulations based on mammalian ribosome profiling data [56,71]. Ternary complex molecules have a single component *ssusite* that serves as a binding site for the small ribosomal subunit.

## Reactions

We describe here each type of kinetic reaction in our models of eukaryotic translational control (Fig 3B). We use a syntax similar to that of BioNetGen [59] to illustrate the kinetic

reactions. We scale TC and ribosome subunit numbers (100 each) to the single mRNA present in the simulation. Simulation of a single mRNA over several rounds of translation is sufficient to infer steady state translation dynamics.

**Initiation: PIC (43S) formation.** Small ribosomal subunits must bind TCs to form pre-initiation complexes (PICs, 43Ss) before loading onto mRNAs. We assume that PIC formation is irreversible. PIC formation is not rate-limiting in our simulations; we set the rate of 43S-cap binding ( $k_{cap\ bind}$ ) to be rate-limiting and to a total rate (independent of [43S]) to match the overall initiation rate to that of cellular estimates. Therefore, we arbitrarily set the second-order PIC formation rate (40S-TC binding rate,  $k_{ssu\ tc\ bind}$ ) to  $0.01 * TC^{-1} * SSU^{-1}$  such that 100 40S-TC binding events occur per second, which is much higher than the 43S-cap binding rate.

**Initiation: PIC (43S) loading onto mRNA.** We model ribosome footprints at 30 nt following mammalian ribosome profiling data [56,71]. Therefore, PIC loading can occur when the 5' most 30 nucleotides (nt) of the mRNA are not bound to any ribosome. The rate at which PICs load onto the 5' end of the mRNA,  $k_{cap\ bind}$ , is varied over a 100-fold range from the maximum ribosome loading rate, 0.125/s, based on single-molecule estimations in human cells [52]. PICs can load onto the mRNA when a ribosome footprint-sized region at the 5' mRNA end is free of ribosomes. PIC loading results in the 5' end being blocked until this ribosome scans or elongates past a ribosome footprint from the 5' cap. We assume that PIC loading is irreversible.

**Initiation: Scanning and start codon selection.** The scanning rate is 5 nucleotides/s following estimates in a mammalian cell-free translation system [55] and a previous computational study [38]. Small ribosomal subunit A sites must be positioned exactly over start codons to initiate translation. The uORF start codon is 25 nt from the 5' cap. We vary the rate at which this start codon selection occurs at the uORF in our modeling. Start codon selection releases the TC bound to the small ribosomal subunit. We assume that TC is regenerated instantaneously. The start codon selection rate divided by the sum of this start codon selection rate, the scanning rate, and the backward scanning rate equals the baseline initiating fraction. This calculation of the baseline initiating fraction will underestimate the initiating fraction in the case of correctly positioned 3' ribosome queues (as in the queuing-mediated enhanced repression model). We assume that start codon selection is irreversible.

**Elongation.** Elongation results in the ribosome A site moving from codon  $c_i$  to codon  $c_{i+1}$ . The rate of elongation is set to 5 codons/s following single-molecule method and ribosome profiling estimates in mammalian cells of 3–18 codons/s [52–54,56,119]. Elongation may only proceed if there is no occluding 3' ribosome; in other words, elongation may only proceed from codon  $c_i$  to codon  $c_{i+1}$  if the next 3' ribosome's A site is bound to a codon no more 5' than  $c_{i+1}$ . The elongation rate at the stall within the uORF is set to 0.001/s [57].

**Termination, continued scanning, and re-initiation.** Termination results in the dissociation of the large ribosomal subunit, but the small ribosomal subunit may continue scanning and subsequently re-initiate if a new TC is acquired before the next start codon is encountered. The termination rate is set to 1/s given that ribosome density tends to be higher at stop codons than within ORFs [56,92]. The recycling rate of terminated small ribosomal subunits after uORF translation is varied to model the effect of varied continued scanning after uORFs on the regulation of main ORF translation. The scanning rate divided by the sum of the scanning rate and this recycling rate equals the continued scanning fraction.

**Collisions and dissociations.** A collision between two ribosomes requires them to be separated by exactly one ribosome footprint in distance on the mRNA and results in binding between the 5' side of the leading (3' most) ribosome and the 3' side of the trailing (5' most) ribosome. Abortive (premature) termination of ribosomes results in their dissociation from

the mRNA and any collided ribosomes they are bound to. Different models have different non-zero dissociation rates. For instance in the 80S-hit model, the following rates are equal and non-zero:  $k_{scan\ term\ 5\ hit\ 80s}$ ,  $k_{scan\ term\ both\ hit\ 80s\ 80s}$ ,  $k_{scan\ term\ both\ hit\ 80s\ 40s}$ . These rates relate to the dissociation of scanning ribosomes upon collisions with a 5' elongating ribosome. Both hit refers to collisions with ribosomes on both sides. In the collision-mediated 40S dissociation model, the following rates are equal and non-zero:  $k_{scan\ term\ 3\ hit\ 40s}$ ,  $k_{scan\ term\ 3\ hit\ 80s}$ ,  $k_{scan\ term\ both\ hit\ 40s\ 40s}$ ,  $k_{scan\ term\ both\ hit\ 40s\ 80s}$ ,  $k_{scan\ term\ both\ hit\ 80s\ 40s}$ ,  $k_{scan\ term\ both\ hit\ 80s\ 80s}$ . These rates relate to the dissociation of scanning ribosomes upon collisions with a 3' scanning or elongating ribosome. The in vivo abortive termination rates of scanning ribosomes are not known. Small ribosomal subunits that make it to the 3' end of the mRNA through leaky scanning of all (u)ORFs always dissociate.

### Model calibration to reporter measurements

We derive the  $k_{cap\ bind}$  rates by spline interpolation of computationally modeled protein output fit to experimental data (Fig 2C). We minimized the root mean square error between modeled protein output across variations in these parameters and the experimental data.

### Human uORF search

We import uORF lists from several databases [6,81,82]. The SmProt database [82] includes 3162 uORFs from ribosome profiling data, which we filter down first to 1080 uORFs after filtering for aligned matches, available Kozak context, near-cognate start codons, and non-duplicates. Two of these uORFs end in diproline motifs, including *C1orf43*. Another database is a set of high confidence ORFs derived from ribosome profiling of human-induced pluripotent stem cells (iPSCs) or foreskin fibroblast cells (HFFs) and was downloaded from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4720255/bin/NIHMS741295-supplement-3.csv> [6]. This database includes 1517 high confidence (ORF-RATER score > 0.8) uORFs from either iPSCs or HFFs, which we filter down to 3 that end in diproline motifs, including *ABCB9*, *C1orf43*, and *TOR1AIP1*. The third database derives from HEK293T, HeLa, and K562 cells using ribosome profiling and was downloaded from [https://static-content.springer.com/esm/art%3A10.1038%2Fs41589-019-0425-0/MediaObjects/41589\\_2019\\_425\\_MOESM3\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1038%2Fs41589-019-0425-0/MediaObjects/41589_2019_425_MOESM3_ESM.xlsx) [81]. This database includes 3577 uORFs which we filter down to 3 that end in diproline motifs and that are less than 60 codons in length for ease of cloning, including *ABCB9*, *C15orf59*, and *PPP1R37*.

### Supporting information

**S1 Fig.** (A) The indicated mutations not present in Fig 2C (the no NLuc start codon and no uORF2 near-cognate start codons) remove two adjacent NLuc ATG codons (ATGATG to ACCACC) and remove 4 uORF2 near-cognate start codons (CTG to CTA or TTG to TTA, red bars), respectively. The no NLuc start codon mutant abolishes NLuc signal, and the no uORF2 near-cognate start codons mutant does not greatly affect uORF2 repressiveness. Error bars show standard error of mean NLuc / FLuc ratios over 3 biological replicates. Data are normalized to the no-uORF start codon control. (B) Raw FLuc and NLuc signals for indicated mutations from Fig 2C. Mock refers to transfection of no plasmid. Multiple data points indicate biological replicates. (EPS)

**S2 Fig.** (A) Buffering in the 80S-hit dissociation model is affected by uORF length. Re-initiation is 0.2%. uORF initiation is 80%. (B) Buffering in the 80S-hit dissociation model is lost

with control matched parameters. Buffering in the 80S-hit dissociation model requires strong uORF initiation and rare re-initiation (Fig 4B, left panel, yellow-green line) and is stronger with longer uORFs (S2A Fig, yellow-green line). However, we estimate re-initiation to be frequent (Table 1) following calibration of our modeling (Fig 2D) to reporter measurements on wild-type or mutant uORF2 (Fig 2C). uORF initiation is 2%. When the elongating ribosome stall is present,  $d_{stall}$  is 63 nt to prevent reduction to the queuing-mediated enhanced repression model. (C) Queuing-mediated enhanced uORF initiation is sensitive to  $d_{stall}$ . As the rate of ribosome loading increases, the average queue size increases and allows enhanced uORF initiation only when  $d_{stall}$  equals an integer multiple of the ribosome footprint (30 nt). (D) Backward scanning only relaxes the dependence of buffering on  $d_{stall}$  in the queuing-mediated enhanced repression model when  $d_{stall}$  is close to an integer multiple of the ribosome footprint (30 nt). The forward scanning rate is 5 nucleotides/s. For  $d_{stall}$  values of 60, 63, 66 nt, the uORF length is 21, 22, 23 codons, respectively. (E) Buffering in the collision-mediated 40S dissociation model occurs even with a rather low dissociation rate. Here,  $d_{stall}$  is 63 nt. (F) Buffering in the collision-mediated enhanced repression model (Fig 1D) is insensitive to  $d_{stall}$ . All rates and labels are identical to Fig 4 unless otherwise specified. Error bars of simulated data are smaller than data markers.

(EPS)

**S3 Fig.** (A) Initiation at the second downstream uORF is dependent on high ternary complex concentration. Initiation at the first uORF is 100%. Continued scanning fractions at both uORFs are 100%. Following termination at the first uORF, initiation at the second downstream uORF depends on if a new ternary complex has been acquired since termination at the first uORF. Only when ternary complex concentration is high does this real uORF2 initiation fraction approach the predicted fraction. (B) With an elongating ribosome stall, the 80S-hit dissociation model acquires  $d_{stall}$ -dependent buffering similar to that in the queuing-mediated enhanced repression model (Fig 4C). Re-initiation is 0.2%. All rates and labels are identical to Fig 4 unless otherwise specified. Error bars of simulated data are smaller than data markers.

(EPS)

**S4 Fig.** Experimentally increasing the distance between the human cytomegaloviral uORF2 start codon and elongating ribosome stall using FLAG donor sequence. The human cytomegaloviral *UL4* uORF2 is used in the dual-luciferase assay (Fig 2B) in conjunction with various length inserts from the N-terminus of the *FLAG* main ORF. The *FLAG* main ORF sequence is inserted directly 3' to the uORF2 start codon. The added sequence increases the distance between the uORF2 start codon and elongating ribosome stall. The bottom two controls improve the uORF2 Kozak context and remove the start codon. Error bars show standard error of mean NLuc / FLuc ratios over 3 biological replicates. Data are normalized to a no-uORF start codon control.

(EPS)

**S5 Fig.** Ribosome density within elongation stall-containing human uORFs. (A) Small ribosomal subunit (TCP-seq [89]) coverage data. (B) Elongating ribosome (Ribo-seq, A site global aggregate) coverage data. *AMD1*, *AZIN1*, *DDIT3* (*CHOP*), *MTR* and *PPP1R15A* (*GADD34*) uORF amino acid sequences are MAGDIS, IPPKKRRRFRTRFGPLSHGELSDQVYNYPEGL-GEVLYREQDFNAEPPWEPS, MLKMSGWQRQSQNSWNLRRRECSRRKCFIHHHT, MSRRPPLPVFSWVLFRAVPRRLWPRVSGC, and MNALASLTVRTCDRFWQTE-PALLPPG, respectively. Elongation stall locations are marked with red arrows. Coverage data were downloaded from GWIPS [120].

(EPS)

**S1 Table. List of plasmids used in this study.**  
(CSV)

## Acknowledgments

We thank members of the Subramaniam lab, the Basic Sciences Division, and the Computational Biology Program at Fred Hutch for assistance with the project and discussions and feedback on the manuscript. The computations described here were performed on the Fred Hutch Cancer Research Center computing cluster.

## Author Contributions

**Conceptualization:** Ty A. Bottorff, Adam P. Geballe, Arvind Rasi Subramaniam.

**Formal analysis:** Ty A. Bottorff, Arvind Rasi Subramaniam.

**Funding acquisition:** Arvind Rasi Subramaniam.

**Investigation:** Ty A. Bottorff, Arvind Rasi Subramaniam.

**Methodology:** Heungwon Park, Adam P. Geballe, Arvind Rasi Subramaniam.

**Project administration:** Adam P. Geballe, Arvind Rasi Subramaniam.

**Supervision:** Adam P. Geballe, Arvind Rasi Subramaniam.

**Writing – original draft:** Ty A. Bottorff, Arvind Rasi Subramaniam.

**Writing – review & editing:** Ty A. Bottorff, Adam P. Geballe, Arvind Rasi Subramaniam.

## References

1. Wethmar K. The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA*. 5(6):765–78. <https://doi.org/10.1002/wrna.1245> PMID: 24995549
2. Ye Y, Liang Y, Yu Q, Hu L, Li H, Zhang Z, et al. Analysis of human upstream open reading frames and impact on gene expression. *Hum Genet*. 2015 Jun; 134(6):605–12. <https://doi.org/10.1007/s00439-015-1544-7> PMID: 25800702
3. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci U S A*. 2009 May 5; 106(18):7507–12. <https://doi.org/10.1073/pnas.0810916106> PMID: 19372376
4. Johnstone TG, Bazzini AA, Giraldez AJ. Upstream ORFs are prevalent translational repressors in vertebrates. *The EMBO Journal* [Internet]. 2016 Apr 1 [cited 2020 Nov 20]; 35(7):706–23. Available from: <https://www.embopress.org/doi/full/10.15252/embj.201592759> PMID: 26896445
5. Rodriguez CM, Chun SY, Mills RE, Todd PK. Translation of upstream open reading frames in a model of neuronal differentiation. *BMC Genomics* [Internet]. 2019 May 20 [cited 2020 Nov 20]; 20(1):391. Available from: <https://doi.org/10.1186/s12864-019-5775-1> PMID: 31109297
6. Chen F, Wu P, Deng S, Zhang H, Hou Y, Hu Z, et al. Dissimilation of synonymous codon usage bias in virus–host coevolution due to translational selection. *Nature Ecology & Evolution* [Internet]. 2020 Apr [cited 2020 Apr 16]; 4(4):589–600. Available from: <http://www.nature.com/articles/s41559-020-1124-7> <https://doi.org/10.1038/s41559-020-1124-7> PMID: 32123323
7. Zhang H, Wang Y, Wu X, Tang X, Wu C, Lu J. Determinants of genome-wide distribution and evolution of uORFs in eukaryotes. *Nature Communications* [Internet]. 2021 Feb 17 [cited 2021 Feb 23]; 12(1):1076. Available from: <https://www.nature.com/articles/s41467-021-21394-y> <https://doi.org/10.1038/s41467-021-21394-y> PMID: 33597535
8. Giess A, Cleuren YNT, Tjeldnes H, Krause M, Bizuayehu TT, Hiensch S, et al. Deconstructing the individual steps of vertebrate translation initiation. *bioRxiv* [Internet]. 2019 Oct 21 [cited 2021 May 17]; 811810. Available from: <https://www.biorxiv.org/content/10.1101/811810v1>
9. Chew GL, Pauli A, Schier AF. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nature Communications* [Internet]. 2016 May 24 [cited 2020 Nov 20]; 7

- (1):11663. Available from: <https://www.nature.com/articles/ncomms11663> <https://doi.org/10.1038/ncomms11663> PMID: 27216465
10. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* [Internet]. 1987 Oct 26 [cited 2020 Nov 20]; 15(20):8125–48. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC306349/> <https://doi.org/10.1093/nar/15.20.8125> PMID: 3313277
  11. Dvir S, Velten L, Sharon E, Zeevi D, Carey LB, Weinberger A, et al. Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *PNAS* [Internet]. 2013 Jul 23 [cited 2020 Nov 21]; 110(30):E2792–801. Available from: <https://www.pnas.org/content/110/30/E2792> <https://doi.org/10.1073/pnas.1222534110> PMID: 23832786
  12. Lee DSM, Park J, Kromer A, Baras A, Rader DJ, Ritchie MD, et al. Disrupting upstream translation in mRNAs is associated with human disease. *Nat Commun* [Internet]. 2021 Dec [cited 2021 Aug 17]; 12(1):1515. Available from: <http://www.nature.com/articles/s41467-021-21812-1> <https://doi.org/10.1038/s41467-021-21812-1> PMID: 33750777
  13. Jürgens L, Manske F, Hubert E, Kischka T, Flötotto L, Klaas O, et al. Somatic Functional Deletions of Upstream Open Reading Frame-Associated Initiation and Termination Codons in Human Cancer. *Bio-medicines*. 2021 May 29; 9(6):618. <https://doi.org/10.3390/biomedicines9060618> PMID: 34072580
  14. Schulz J, Mah N, Neuenschwander M, Kischka T, Ratei R, Schlag PM, et al. Loss-of-function uORF mutations in human malignancies. *Sci Rep*. 2018 Feb 5; 8(1):2395. <https://doi.org/10.1038/s41598-018-19201-8> PMID: 29402903
  15. Smith RCL, Kanellos G, Vlahov N, Alexandrou C, Willis AE, Knight JRP, et al. Translation initiation in cancer at a glance. *J Cell Sci* [Internet]. 2021 Jan 1 [cited 2021 Jan 19]; 134(1). Available from: <https://jcs.biologists.org/content/134/1/jcs248476> <https://doi.org/10.1242/jcs.248476> PMID: 33441326
  16. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proceedings of the National Academy of Sciences* [Internet]. 2004 Aug 3 [cited 2021 Jun 18]; 101(31):11269–74. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.0400541101> PMID: 15277680
  17. Dever TE, Feng L, Wek RC, Cigan AM, Donahue TF, Hinnebusch AG. Phosphorylation of initiation factor 2 alpha by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. *Cell*. 1992 Feb 7; 68(3):585–96. [https://doi.org/10.1016/0092-8674\(92\)90193-g](https://doi.org/10.1016/0092-8674(92)90193-g) PMID: 1739968
  18. Shah P, Ding Y, Niemczyk M, Kudla G, Plotkin JB. Rate-limiting steps in yeast protein translation. *Cell*. 2013 Jun 20; 153(7):1589–601. <https://doi.org/10.1016/j.cell.2013.05.049> PMID: 23791185
  19. Gandin V, Miluzio A, Barbieri AM, Beugnet A, Kiyokawa H, Marchisio PC, et al. Eukaryotic initiation factor 6 is rate-limiting in translation, growth and transformation. *Nature* [Internet]. 2008 Oct [cited 2021 Nov 16]; 455(7213):684–8. Available from: <https://www.nature.com/articles/nature07267> <https://doi.org/10.1038/nature07267> PMID: 18784653
  20. Sharma AK, Sormanni P, Ahmed N, Ciryam P, Friedrich UA, Kramer G, et al. A chemical kinetic basis for measuring translation initiation and elongation rates from ribosome profiling data. *PLoS Comput Biol* [Internet]. 2019 May 23 [cited 2021 Jan 21]; 15(5). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6559674/> <https://doi.org/10.1371/journal.pcbi.1007070> PMID: 31120880
  21. Cao J, Geballe AP. Coding sequence-dependent ribosomal arrest at termination of translation. *Molecular and Cellular Biology* [Internet]. 1996 Feb 1 [cited 2020 Nov 20]; 16(2):603–8. Available from: <https://mcb.asm.org/content/16/2/603> <https://doi.org/10.1128/MCB.16.2.603> PMID: 8552088
  22. Young SK, Palam LR, Wu C, Sachs MS, Wek RC. Ribosome Elongation Stall Directs Gene-specific Translation in the Integrated Stress Response. *Journal of Biological Chemistry* [Internet]. 2016 Mar [cited 2021 Jun 22]; 291(12):6546–58. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0021925820443443> <https://doi.org/10.1074/jbc.M115.705640> PMID: 26817837
  23. Ivanov IP, Shin BS, Loughran G, Tzani I, Young-Baird SK, Cao C, et al. Polyamine Control of Translation Elongation Regulates Start Site Selection on the Antizyme Inhibitor mRNA via Ribosome Queuing. *Mol Cell* [Internet]. 2018 Apr 19 [cited 2020 Nov 20]; 70(2):254–264.e6. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5916843/>
  24. Wei J, Wu C, Sachs MS. The arginine attenuator peptide interferes with the ribosome peptidyl transferase center. *Mol Cell Biol*. 2012 Jul; 32(13):2396–406. <https://doi.org/10.1128/MCB.00136-12> PMID: 22508989
  25. Wang Z, Gaba A, Sachs MS. A Highly Conserved Mechanism of Regulated Ribosome Stalling Mediated by Fungal Arginine Attenuator Peptides That Appears Independent of the Charging Status of Arginyl-tRNAs \*. *Journal of Biological Chemistry* [Internet]. 1999 Dec 31 [cited 2021 Nov 10]; 274(53):37565–74. Available from: [https://www.jbc.org/article/S0021-9258\(19\)52930-1/abstract](https://www.jbc.org/article/S0021-9258(19)52930-1/abstract) <https://doi.org/10.1074/jbc.274.53.37565> PMID: 10608810



26. Lin Y, May G, Kready H, Nazzaro L, Mao M, Spealman P, et al. Impacts of uORF codon identity and position on translation regulation. 2019 Oct 31 [cited 2020 Nov 20]; Available from: /articles/journal\_contribution/Impacts\_of\_uORF\_codon\_identity\_and\_position\_on\_translation\_regulation/10110866/1
27. Bhushan S, Meyer H, Starosta AL, Becker T, Mielke T, Berninghausen O, et al. Structural Basis for Translational Stalling by Human Cytomegalovirus and Fungal Arginine Attenuator Peptide. *Molecular Cell* [Internet]. 2010 Oct 8 [cited 2020 Nov 16]; 40(1):138–46. Available from: <http://www.sciencedirect.com/science/article/pii/S1097276510007112> <https://doi.org/10.1016/j.molcel.2010.09.009> PMID: 20932481
28. Wilson DN, Arenz S, Beckmann R. Translation regulation via nascent polypeptide-mediated ribosome stalling. *Current Opinion in Structural Biology* [Internet]. 2016 Apr 1 [cited 2021 Nov 8]; 37:123–33. Available from: <https://www.sciencedirect.com/science/article/pii/S0959440X16000117> <https://doi.org/10.1016/j.sbi.2016.01.008> PMID: 26859868
29. Meijer HA, Thomas AAM. Ribosomes stalling on uORF1 in the *Xenopus* Cx41 5' UTR inhibit downstream translation initiation. *Nucleic Acids Res*. 2003 Jun 15; 31(12):3174–84. <https://doi.org/10.1093/nar/gkg429> PMID: 12799445
30. Young SK, Baird TD, Wek RC. Translation Regulation of the Glutamyl-prolyl-tRNA Synthetase Gene EPRS through Bypass of Upstream Open Reading Frames with Noncanonical Initiation Codons. *Journal of Biological Chemistry* [Internet]. 2016 May [cited 2021 Jun 30]; 291(20):10824–35. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0021925820401127> <https://doi.org/10.1074/jbc.M116.722256> PMID: 27002157
31. Cao J, Geballe AP. Translational inhibition by a human cytomegalovirus upstream open reading frame despite inefficient utilization of its AUG codon. *J Virol*. 1995 Feb; 69(2):1030–6. <https://doi.org/10.1128/JVI.69.2.1030-1036.1995> PMID: 7815480
32. Degnin C, Schleiss M, Cao J, Geballe A. Translational inhibition mediated by a short upstream open reading frame in the human cytomegalovirus gpUL4 (gp48) transcript. *Journal of virology*. 1993; <https://doi.org/10.1128/JVI.67.9.5514-5521.1993> PMID: 8394459
33. Boersma S, Khuperkar D, Verhagen BMP, Sonneveld S, Grimm JB, Lavis LD, et al. Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. *Cell*. 2019 Jul 11; 178(2):458–472.e19. <https://doi.org/10.1016/j.cell.2019.05.001> PMID: 31178119
34. Zur H, Tuller T. Predictive biophysical modeling and understanding of the dynamics of mRNA translation and its evolution. *Nucleic Acids Res*. 2016 Nov 2; 44(19):9031–49. <https://doi.org/10.1093/nar/gkw764> PMID: 27591251
35. Abastado JP, Miller PF, Hinnebusch AG. A quantitative model for translational control of the GCN4 gene of *Saccharomyces cerevisiae*. *New Biol*. 1991 May; 3(5):511–24. PMID: 1883814
36. You T, Stansfield I, Romano MC, Brown AJ, Coghill GM. Analysing GCN4 translational control in yeast by stochastic chemical kinetics modelling and simulation. *BMC Syst Biol* [Internet]. 2011 Dec [cited 2021 Oct 26]; 5(1):131. Available from: <https://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-5-131> PMID: 21851603
37. Marasco ONJM Roussel MR, Thakor N. Probabilistic models of uORF-mediated ATF4 translation control. *Mathematical Biosciences* [Internet]. 2021 Dec 6 [cited 2021 Dec 14]; 108762. Available from: <https://www.sciencedirect.com/science/article/pii/S0025556421001607>
38. Andreev DE, Arnold M, Kiniry SJ, Loughran G, Michel AM, Rachinskii D, et al. TASEP modelling provides a parsimonious explanation for the ability of a single uORF to derepress translation during the integrated stress response. Sonenberg N, editor. *eLife* [Internet]. 2018 Jun 22 [cited 2020 Nov 20]; 7:e32563. Available from: <https://doi.org/10.7554/eLife.32563> PMID: 29932418
39. Park H, Subramaniam AR. Inverted translational control of eukaryotic gene expression by ribosome collisions. *PLoS Biol*. 2019 Sep; 17(9):e3000396. <https://doi.org/10.1371/journal.pbio.3000396> PMID: 31532761
40. Poyry TAA. What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes & Development* [Internet]. 2004 Jan 1 [cited 2021 Nov 8]; 18(1):62–75. Available from: <http://www.genesdev.org/cgi/doi/10.1101/gad.276504> PMID: 14701882
41. Kozak M. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res* [Internet]. 2001 Dec 15 [cited 2021 Nov 8]; 29(24):5226–32. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC97554/> <https://doi.org/10.1093/nar/29.24.5226> PMID: 11812856
42. Luukkonen BG, Tan W, Schwartz S. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J Virol* [Internet]. 1995 Jul [cited 2021 Nov 8]; 69(7):4086–94. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC189143/> <https://doi.org/10.1128/JVI.69.7.4086-4094.1995> PMID: 7769666

43. Kozak M. Effects of Intercistronic Length on the Efficiency of Reinitiation by Eucaryotic Ribosomes. *MOL CELL BIOL*. 1987; 7:8. <https://doi.org/10.1128/mcb.7.10.3438-3445.1987> PMID: 3683388
44. Grant CM, Hinnebusch AG. Effect of sequence context at stop codons on efficiency of reinitiation in GCN4 translational control. *Mol Cell Biol* [Internet]. 1994 Jan [cited 2020 Nov 20]; 14(1):606–18. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC358410/> <https://doi.org/10.1128/mcb.14.1.606-618.1994> PMID: 8264629
45. Hinnebusch AG. Translational Regulation of Yeast GCN4: A WINDOW ON FACTORS THAT CONTROL INITIATOR-tRNA BINDING TO THE RIBOSOME \*. *Journal of Biological Chemistry* [Internet]. 1997 Aug 29 [cited 2021 Nov 8]; 272(35):21661–4. Available from: [https://www.jbc.org/article/S0021-9258\(19\)65604-8/abstract](https://www.jbc.org/article/S0021-9258(19)65604-8/abstract)
46. Janzen DM, Frolova L, Geballe AP. Inhibition of translation termination mediated by an interaction of eukaryotic release factor 1 with a nascent peptidyl-tRNA. *Mol Cell Biol*. 2002 Dec; 22(24):8562–70. <https://doi.org/10.1128/MCB.22.24.8562-8570.2002> PMID: 12446775
47. Col B, Oltean S, Banerjee R. Translational Regulation of Human Methionine Synthase by Upstream Open Reading Frames. *Biochim Biophys Acta* [Internet]. 2007 [cited 2021 Jan 19]; 1769(9–10):532–40. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2682437/> <https://doi.org/10.1016/j.bbaexp.2007.06.003> PMID: 17683808
48. Yamashita Y, Takamatsu S, Glasbrenner M, Becker T, Naito S, Beckmann R. Sucrose sensing through nascent peptide-mediated ribosome stalling at the stop codon of Arabidopsis bZIP11 uORF2. *FEBS Letters* [Internet]. 2017 [cited 2021 Nov 18]; 591(9):1266–77. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/1873-3468.12634> PMID: 28369795
49. Tanaka M, Sotta N, Yamazumi Y, Yamashita Y, Miwa K, Murota K, et al. The Minimum Open Reading Frame, AUG-Stop, Induces Boron-Dependent Ribosome Stalling and mRNA Degradation. *The Plant Cell* [Internet]. 2016 Nov 1 [cited 2021 Nov 18]; 28(11):2830–49. Available from: <https://doi.org/10.1105/tpc.16.00481> PMID: 27760805
50. Hou C, Lee W, Chou H, Chen A, Chou S, Chen H. Global Analysis of Truncated RNA Ends Reveals New Insights into Ribosome Stalling in Plants. *The Plant Cell* [Internet]. 2016 Oct 1 [cited 2021 Nov 18]; 28(10):2398–416. Available from: <https://doi.org/10.1105/tpc.16.00295> PMID: 27742800
51. Uchiyama-Kadokura N, Murakami K, Takemoto M, Koyanagi N, Murota K, Naito S, et al. Polyamine-Responsive Ribosomal Arrest at the Stop Codon of an Upstream Open Reading Frame of the Ado-MetDC1 Gene Triggers Nonsense-Mediated mRNA Decay in Arabidopsis thaliana. *Plant and Cell Physiology* [Internet]. 2014 Sep 1 [cited 2021 Nov 18]; 55(9):1556–67. Available from: <https://doi.org/10.1093/pcp/pcu086> PMID: 24929422
52. Yan X, Hoek Tim A, Vale Ronald D, Tanenbaum Marvin E. Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* [Internet]. 2016 May 5 [cited 2021 Jan 21]; 165(4):976–89. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4889334/> <https://doi.org/10.1016/j.cell.2016.04.034> PMID: 27153498
53. Morisaki T, Lyon K, DeLuca KF, DeLuca JG, English BP, Zhang Z, et al. Real-time quantification of single RNA translation dynamics in living cells. *Science* [Internet]. 2016 Jun 17 [cited 2021 Aug 30]; 352(6292):1425–9. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.aaf0899> PMID: 27313040
54. Wu B, Eliscovich C, Yoon YJ, Singer RH. Translation dynamics of single mRNAs in live cells and neurons. *Science* [Internet]. 2016 Jun 17 [cited 2021 Jan 21]; 352(6292):1430–5. Available from: <https://science.sciencemag.org/content/352/6292/1430> <https://doi.org/10.1126/science.aaf1084> PMID: 27313041
55. Vassilenko KS, Alekhina OM, Dmitriev SE, Shatsky IN, Spirin AS. Unidirectional constant rate motion of the ribosomal scanning particle during eukaryotic translation initiation. *Nucleic Acids Res*. 2011 Jul; 39(13):5555–67. <https://doi.org/10.1093/nar/gkr147> PMID: 21415006
56. Ingolia NT, Lareau LF, Weissman JS. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* [Internet]. 2011 Nov 11 [cited 2020 Nov 20]; 147(4):789–802. Available from: <https://doi.org/10.1016/j.cell.2011.10.002> PMID: 22056041
57. Cao J, Geballe AP. Ribosomal release without peptidyl tRNA hydrolysis at translation termination in a eukaryotic system. *RNA* [Internet]. 1998 Feb [cited 2021 Jan 28]; 4(2):181–8. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1369606/> PMID: 9570317
58. Lopez CF, Muhlich JL, Bachman JA, Sorger PK. Programming biological models in Python using PySB. *Molecular Systems Biology* [Internet]. 2013 Jan 1 [cited 2020 Nov 21]; 9(1). Available from: <https://www.embopress.org/doi/abs/10.1038/msb.2013.1> PMID: 23423320
59. Harris LA, Hogg JS, Tapia JJ, Sekar JAP, Gupta S, Korsunsky I, et al. BioNetGen 2.2: advances in rule-based modeling. *Bioinformatics* [Internet]. 2016 Nov 1 [cited 2021 Jan 5]; 32(21):3366–8. Available from: <https://doi.org/10.1093/bioinformatics/btw469> PMID: 27402907

60. Sneddon MW, Faeder JR, Emonet T. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nat Methods*. 2011 Feb; 8(2):177–83. <https://doi.org/10.1038/nmeth.1546> PMID: 21186362
61. Costa-Mattioli M, Walter P. The integrated stress response: From mechanism to disease. *Science* [Internet]. 2020 Apr 24 [cited 2021 Jun 22];368(6489). Available from: <https://science.sciencemag.org/content/368/6489/eaat5314> <https://doi.org/10.1126/science.aat5314> PMID: 32327570
62. Pakos-Zebrucka K, Koryga I, Mnich K, Lujic M, Samali A, Gorman AM. The integrated stress response. *EMBO Rep* [Internet]. 2016 Oct [cited 2021 Oct 20]; 17(10):1374–95. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5048378/> <https://doi.org/10.15252/embr.201642195> PMID: 27629041
63. Wu CCC, Peterson A, Zinshteyn B, Regot S, Green R. Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell* [Internet]. 2020 Jul 23 [cited 2020 Nov 20]; 182(2):404–416. e14. Available from: <http://www.sciencedirect.com/science/article/pii/S0092867420306930> <https://doi.org/10.1016/j.cell.2020.06.006> PMID: 32610081
64. Darnell AM, Subramaniam AR, O'Shea EK. Translational Control through Differential Ribosome Pausing during Amino Acid Limitation in Mammalian Cells. *Molecular Cell* [Internet]. 2018 Jul 19 [cited 2018 Sep 7]; 71(2):229–243.e11. Available from: <http://www.sciencedirect.com/science/article/pii/S1097276518305161> <https://doi.org/10.1016/j.molcel.2018.06.041> PMID: 30029003
65. Vladimer GI, Gónna MW, Superti-Furga G. IFITs: Emerging Roles as Key Anti-Viral Proteins. *Front Immunol* [Internet]. 2014 Mar 10 [cited 2021 Oct 20]; 5:94. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3948006/>
66. Spriggs KA, Bushell M, Willis AE. Translational Regulation of Gene Expression during Conditions of Cell Stress. *Molecular Cell* [Internet]. 2010 Oct 22 [cited 2021 Oct 20]; 40(2):228–37. Available from: <https://www.sciencedirect.com/science/article/pii/S1097276510007562> <https://doi.org/10.1016/j.molcel.2010.09.028> PMID: 20965418
67. Andreev DE, O'Connor PB, Fahey C, Kenny EM, Terenin IM, Dmitriev SE, et al. Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Sonenberg N, editor. eLife* [Internet]. 2015 Jan 26 [cited 2020 Nov 20]; 4:e03971. Available from: <https://doi.org/10.7554/eLife.03971> PMID: 25621764
68. Sidrauski C, McGeachy AM, Ingolia NT, Walter P. The small molecule ISRIB reverses the effects of eIF2 $\alpha$  phosphorylation on translation and stress granule assembly. *Ron D, editor. eLife* [Internet]. 2015 Feb 26 [cited 2021 Feb 23]; 4:e05033. Available from: <https://doi.org/10.7554/eLife.05033> PMID: 25719440
69. Baird TD, Palam LR, Fusakio ME, Willy JA, Davis CM, McClintick JN, et al. Selective mRNA translation during eIF2 phosphorylation induces expression of IBTK $\alpha$ . *Mol Biol Cell*. 2014 May; 25(10):1686–97.
70. Wei S, Guo W, Qian Y, Xiang J, Liu K, Gao XJ, et al. Ribosome profiling reveals translational remodeling in cancer cells in response to zinc oxide nanoparticles. *Aging* [Internet]. 2021 Oct 15 [cited 2021 Dec 15]; 13(19):23119–32. Available from: <https://www.aging-us.com/lookup/doi/10.18632/aging.203606> PMID: 34620733
71. Jackson R, Standart N. The awesome power of ribosome profiling. *RNA* [Internet]. 2015 Apr [cited 2021 Jan 21]; 21(4):652–4. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4371319/> <https://doi.org/10.1261/rna.049908.115> PMID: 25780177
72. Matsuda D, Dreher TW. Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *RNA* [Internet]. 2006 Jul [cited 2021 Nov 11]; 12(7):1338–49. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1484435/> <https://doi.org/10.1261/rna.67906> PMID: 16682564
73. Gu Y, Mao Y, Jia L, Dong L, Qian SB. Bi-directional ribosome scanning controls the stringency of start codon selection. *Nat Commun*. 2021 Nov 15; 12(1):6604. <https://doi.org/10.1038/s41467-021-26923-3> PMID: 34782646
74. Li K, Kong J, Zhang S, Zhao T, Qian W. Distance-dependent inhibition of translation initiation by downstream out-of-frame AUGs reveals that ribosome scans in a Brownian ratchet process [Internet]. *Molecular Biology*; 2021 Nov [cited 2021 Nov 18]. Available from: <http://biorxiv.org/lookup/doi/10.1101/2021.11.01.466764>
75. Kozak M. Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol Cell Biol* [Internet]. 1989 Nov [cited 2020 Nov 20]; 9(11):5073–80. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC363659/> <https://doi.org/10.1128/mcb.9.11.5073-5080.1989> PMID: 2601709
76. Babendure JR, Babendure JL, Ding JH, Tsien RY. Control of mammalian translation by mRNA structure near caps. *RNA* [Internet]. 2006 May [cited 2020 Nov 20]; 12(5):851–61. Available from: <https://>

- [www.ncbi.nlm.nih.gov/pmc/articles/PMC1440912/](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1440912/) <https://doi.org/10.1261/rna.2309906> PMID: 16540693
77. Cencic R, Hall DR, Robert F, Du Y, Min J, Li L, et al. Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. *PNAS* [Internet]. 2011 Jan 18 [cited 2021 Dec 18]; 108(3):1046–51. Available from: <https://www.pnas.org/content/108/3/1046> <https://doi.org/10.1073/pnas.1011477108> PMID: 21191102
  78. Rechsteiner M, Rogers SW. PEST sequences and regulation by proteolysis. *Trends in Biochemical Sciences* [Internet]. 1996 Jul 1 [cited 2021 Jul 1]; 21(7):267–71. Available from: <https://www.sciencedirect.com/science/article/pii/S0968000496100311> PMID: 8755249
  79. Guan BJ, van Hoef V, Jobava R, Elroy-Stein O, Valasek LS, Cargnello M, et al. A Unique ISR Program Determines Cellular Responses to Chronic Stress. *Mol Cell* [Internet]. 2017 Dec 7 [cited 2022 May 18]; 68(5):885–900.e6. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5730339/> <https://doi.org/10.1016/j.molcel.2017.11.007> PMID: 29220654
  80. Nilsson OB, Hedman R, Marino J, Wickles S, Bischoff L, Johansson M, et al. Cotranslational Protein Folding inside the Ribosome Exit Tunnel. *Cell Reports* [Internet]. 2015 Sep 8 [cited 2022 May 3]; 12(10):1533–40. Available from: <https://www.sciencedirect.com/science/article/pii/S2211124715008554> <https://doi.org/10.1016/j.celrep.2015.07.065> PMID: 26321634
  81. Martinez TF, Chu Q, Donaldson C, Tan D, Shokhirev MN, Saghatelian A. Accurate annotation of human protein-coding small open reading frames. *Nature Chemical Biology* [Internet]. 2020 Apr [cited 2020 Dec 29]; 16(4):458–68. Available from: <https://www.nature.com/articles/s41589-019-0425-0> <https://doi.org/10.1038/s41589-019-0425-0> PMID: 31819274
  82. Hao Y, Zhang L, Niu Y, Cai T, Luo J, He S, et al. SmProt: a database of small proteins encoded by annotated coding and non-coding RNA loci. *Brief Bioinform* [Internet]. 2018 Jul 20 [cited 2019 Nov 28]; 19(4):636–43. Available from: <https://academic.oup.com/bib/article/19/4/636/2962821> <https://doi.org/10.1093/bib/bbx005> PMID: 28137767
  83. Neumann T, Tuller T. Modeling the ribosomal small subunit dynamic in *Saccharomyces cerevisiae* based on TCP-seq data. *Nucleic Acids Research* [Internet]. 2022 Jan 31 [cited 2022 Feb 7]; gkac021. Available from: <https://doi.org/10.1093/nar/gkac021> PMID: 35100399
  84. Garshott DM, An H, Sundaramoorthy E, Leonard M, Vicary A, Harper JW, et al. iRQC, a surveillance pathway for 40S ribosomal quality control during mRNA translation initiation. *bioRxiv* [Internet]. 2021 Apr 21 [cited 2021 May 24]; 2021.04.20.440649. Available from: <https://www.biorxiv.org/content/10.1101/2021.04.20.440649v1> PMID: 34469731
  85. Kearse MG, Goldman DH, Choi J, Nwaezeapu C, Liang D, Green KM, et al. Ribosome queuing enables non-AUG translation to be resistant to multiple protein synthesis inhibitors. *Genes Dev*. 2019 Jul 1; 33(13–14):871–85. <https://doi.org/10.1101/gad.324715.119> PMID: 31171704
  86. Manjunath H, Zhang H, Rehfeld F, Han J, Chang TC, Mendell JT. Suppression of Ribosomal Pausing by eIF5A Is Necessary to Maintain the Fidelity of Start Codon Selection. *Cell Reports* [Internet]. 2019 Dec 3 [cited 2020 Nov 20]; 29(10):3134–3146.e6. Available from: <http://www.sciencedirect.com/science/article/pii/S2211124719314627> <https://doi.org/10.1016/j.celrep.2019.10.129> PMID: 31801078
  87. Gaba A, Wang H, Fortune T, Qu X. Smart-ORF: a single-molecule method for accessing ribosome dynamics in both upstream and main open reading frames. *Nucleic Acids Research* [Internet]. 2020 Dec 16 [cited 2020 Dec 21]; (gkaa1185). Available from: <https://doi.org/10.1093/nar/gkaa1185>
  88. Circumstances Kozak M. and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol Cell Biol* [Internet]. 1989 Nov [cited 2020 Nov 20]; 9(11):5134–42. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC363665/>
  89. Wagner S, Herrmannová A, Hronová V, Gunišová S, Sen ND, Hannan RD, et al. Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes. *Mol Cell*. 2020 Aug 20; 79(4):546–560.e7. <https://doi.org/10.1016/j.molcel.2020.06.004> PMID: 32589964
  90. Bohlen J, Fenzl K, Kramer G, Bukau B, Teleman AA. Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. *Molecular Cell* [Internet]. 2020 Aug 20 [cited 2020 Nov 20]; 79(4):561–574.e5. Available from: <http://www.sciencedirect.com/science/article/pii/S1097276520303907> <https://doi.org/10.1016/j.molcel.2020.06.005> PMID: 32589966
  91. 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* [Internet]. 2021 May 20 [cited 2022 Jul 28]; 12(1):2976. Available from: <https://www.nature.com/articles/s41467-021-23223-8> <https://doi.org/10.1038/s41467-021-23223-8> PMID: 34016977
  92. Wu CCC, Zinshteyn B, Wehner KA, Green R. High-Resolution Ribosome Profiling Defines Discrete Ribosome Elongation States and Translational Regulation during Cellular Stress. *Molecular Cell*

- [Internet]. 2019 Mar 7 [cited 2021 Jun 28]; 73(5):959–970.e5. Available from: <https://www.sciencedirect.com/science/article/pii/S1097276518310633> <https://doi.org/10.1016/j.molcel.2018.12.009> PMID: 30686592
93. Lareau LF, Hite DH, Hogan GJ, Brown PO. Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *eLife* [Internet]. 2014 May 9 [cited 2022 Jul 28]; 3:e01257. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4052883/> <https://doi.org/10.7554/eLife.01257> PMID: 24842990
  94. Han P, Shichino Y, Schneider-Poetsch T, Mito M, Hashimoto S, Udagawa T, et al. Genome-wide Survey of Ribosome Collision. *Cell Reports* [Internet]. 2020 May 5 [cited 2020 Dec 29]; 31(5):107610. Available from: <http://www.sciencedirect.com/science/article/pii/S2211124720305593> <https://doi.org/10.1016/j.celrep.2020.107610> PMID: 32375038
  95. 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* [Internet]. 2020 Jul [cited 2021 Jun 18]; 30(7):985–99. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.257741.119> PMID: 32703885
  96. Tuck AC, Rankova A, Arpat AB, Liechti LA, Hess D, Iesmantavicius V, et al. Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. *Molecular Cell* [Internet]. 2020 Mar 19 [cited 2020 Mar 25]; 77(6):1222–1236.e13. Available from: <http://www.sciencedirect.com/science/article/pii/S1097276520300071> <https://doi.org/10.1016/j.molcel.2020.01.007> PMID: 32048998
  97. Rubio A, Ghosh S, Mülleder M, Ralser M, Mata J. Ribosome profiling reveals ribosome stalling on tryptophan codons and ribosome queuing upon oxidative stress in fission yeast. *Nucleic Acids Res*. 2021 Jan 11; 49(1):383–99. <https://doi.org/10.1093/nar/gkaa1180> PMID: 33313903
  98. Rooijers K, Loayza-Puch F, Nijtmans LG, Agami R. Ribosome profiling reveals features of normal and disease-associated mitochondrial translation. *Nat Commun* [Internet]. 2013 Dec 3 [cited 2022 Aug 12]; 4(1):2886. Available from: <https://www.nature.com/articles/ncomms3886> <https://doi.org/10.1038/ncomms3886> PMID: 24301020
  99. Ramu H, Vázquez-Laslop N, Klepacki D, Dai Q, Piccirilli J, Micura R, et al. Nascent peptide in the ribosome exit tunnel affects functional properties of the A-site of the peptidyl transferase center. *Mol Cell*. 2011 Feb 4; 41(3):321–30. <https://doi.org/10.1016/j.molcel.2010.12.031> PMID: 21292164
  100. Dimitrova LN, Kuroha K, Tatematsu T, Inada T. Nascent Peptide-dependent Translation Arrest Leads to Not4p-mediated Protein Degradation by the Proteasome. *J Biol Chem* [Internet]. 2009 Apr 17 [cited 2022 Aug 12]; 284(16):10343–52. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2667721/> <https://doi.org/10.1074/jbc.M808840200> PMID: 19204001
  101. Gamerding M, Kobayashi K, Wallisch A, Kreft SG, Sailer C, Schlömer R, et al. Early Scanning of Nascent Polypeptides inside the Ribosomal Tunnel by NAC. *Mol Cell*. 2019 Sep 5; 75(5):996–1006.e8. <https://doi.org/10.1016/j.molcel.2019.06.030> PMID: 31377116
  102. Weber R, Chung MY, Keskeny C, Zinnall U, Landthaler M, Valkov E, et al. 4EHP and GIGYF1/2 Mediate Translation-Coupled Messenger RNA Decay. *Cell Reports* [Internet]. 2020 Oct 13 [cited 2020 Oct 14]; 33(2):108262. Available from: <http://www.sciencedirect.com/science/article/pii/S2211124720312511> <https://doi.org/10.1016/j.celrep.2020.108262> PMID: 33053355
  103. Law GL, Raney A, Heusner C, Morris DR. Polyamine regulation of ribosome pausing at the upstream open reading frame of S-adenosylmethionine decarboxylase. *J Biol Chem*. 2001 Oct 12; 276(41):38036–43. <https://doi.org/10.1074/jbc.M105944200> PMID: 11489903
  104. Young SK, Willy JA, Wu C, Sachs MS, Wek RC. Ribosome Reinitiation Directs Gene-specific Translation and Regulates the Integrated Stress Response. *J Biol Chem*. 2015 Nov 20; 290(47):28257–71. <https://doi.org/10.1074/jbc.M115.693184> PMID: 26446796
  105. Berthelot K, Muldoon M, Rajkowitsch L, Hughes J, McCarthy JEG. Dynamics and processivity of 40S ribosome scanning on mRNA in yeast. *Mol Microbiol*. 2004 Feb; 51(4):987–1001. <https://doi.org/10.1046/j.1365-2958.2003.03898.x> PMID: 14763975
  106. Mohammad MP, Munzarová Pondělíčková V, Zeman J, Gunišová S, Valášek LS. In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation. *Nucleic Acids Research* [Internet]. 2017 Mar 17 [cited 2021 Dec 31]; 45(5):2658–74. Available from: <https://doi.org/10.1093/nar/gkx049> PMID: 28119417
  107. Sanchez M, Lin Y, Yang CC, McQuary P, Rosa Campos A, Aza Blanc P, et al. Cross Talk between eIF2 $\alpha$  and eEF2 Phosphorylation Pathways Optimizes Translational Arrest in Response to Oxidative Stress. *iScience* [Internet]. 2019 Oct [cited 2021 Sep 7]; 20:466–80. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2589004219303694>
  108. Shu XE, Mao Y, Jia L, Qian SB. Dynamic eIF3a O-GlcNAcylation controls translation reinitiation during nutrient stress. *Nat Chem Biol* [Internet]. 2021 Dec 9 [cited 2021 Dec 14]; Available from: <https://www>.

[nature.com/articles/s41589-021-00913-4](https://doi.org/10.1038/s41589-021-00913-4) <https://doi.org/10.1038/s41589-021-00913-4> PMID: 34887587

109. Yan LL, Zaher HS. Ribosome quality control antagonizes the activation of the integrated stress response on colliding ribosomes. *Molecular Cell* [Internet]. 2020 Dec 17 [cited 2020 Dec 18]; Available from: <http://www.sciencedirect.com/science/article/pii/S1097276520308339> <https://doi.org/10.1016/j.molcel.2020.11.033> PMID: 33338396
110. Wu CCC, Peterson A, Zinshteyn B, Regot S, Green R. Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell* [Internet]. 2020 Jun 30 [cited 2020 Jun 30]; Available from: <http://www.sciencedirect.com/science/article/pii/S0092867420306930> <https://doi.org/10.1016/j.cell.2020.06.006> PMID: 32610081
111. Ferrin MA, Subramaniam AR. Kinetic modeling predicts a stimulatory role for ribosome collisions at elongation stall sites in bacteria. *eLife*. 2017 May 12;e23629. <https://doi.org/10.7554/eLife.23629> PMID: 28498106
112. Wallace EWJ, Maufrais C, Sales-Lee J, Tuck LR, de Oliveira L, Feuerbach F, et al. Quantitative global studies reveal differential translational control by start codon context across the fungal kingdom. *Nucleic Acids Research* [Internet]. 2020 Mar 18 [cited 2021 May 17]; 48(5):2312–31. Available from: <https://doi.org/10.1093/nar/gkaa060> PMID: 32020195
113. Aliouat A, Hatin I, Bertin P, François P, Stierlé V, Namy O, et al. Divergent effects of translation termination factor eRF3A and nonsense-mediated mRNA decay factor UPF1 on the expression of uORF carrying mRNAs and ribosome protein genes. *RNA Biol* [Internet]. 2019 Oct 17 [cited 2020 Nov 20]; 17(2):227–39. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6973328/> <https://doi.org/10.1080/15476286.2019.1674595> PMID: 31619139
114. Jia L, Mao Y, Ji Q, Dersh D, Yewdell JW, Qian SB. Decoding mRNA translatability and stability from the 5' UTR. *Nat Struct Mol Biol* [Internet]. 2020 Sep [cited 2021 Aug 17]; 27(9):814–21. Available from: <http://www.nature.com/articles/s41594-020-0465-x> <https://doi.org/10.1038/s41594-020-0465-x> PMID: 32719458
115. Wang H, Sun L, Gaba A, Qu X. An in vitro single-molecule assay for eukaryotic cap-dependent translation initiation kinetics. *Nucleic Acids Res*. 2020 Jan 10; 48(1):e6. <https://doi.org/10.1093/nar/gkz1066> PMID: 31722415
116. Yu J, Chen M, Huang H, Zhu J, Song H, Zhu J, et al. Dynamic m6A modification regulates local translation of mRNA in axons. *Nucleic Acids Research* [Internet]. 2018 Feb 16 [cited 2022 Feb 23]; 46(3):1412–23. Available from: <https://doi.org/10.1093/nar/gkx1182> PMID: 29186567
117. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*. 2009 May; 6(5):343–5. <https://doi.org/10.1038/nmeth.1318> PMID: 19363495
118. Sambrook J. *Molecular Cloning: A Laboratory Manual, Third Edition*. 3rd ed. Cold Spring Harbor Laboratory Press; 2001. 2344 p.
119. Pichon X, Bastide A, Safieddine A, Chouaib R, Samacoits A, Basyuk E, et al. Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *Journal of Cell Biology* [Internet]. 2016 Sep 5 [cited 2021 Jan 21]; 214(6):769–81. Available from: <https://doi.org/10.1083/jcb.201605024> PMID: 27597760
120. Michel AM, Fox G, Kiran M. A, De Bo C, O'Connor PBF, Heaphy SM, et al. GWIPS-viz: development of a ribo-seq genome browser. *Nucleic Acids Res* [Internet]. 2014 Jan 1 [cited 2022 Aug 12]; 42(Database issue):D859–64. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3965066/> <https://doi.org/10.1093/nar/gkt1035> PMID: 24185699