ABSTRACT

Title of Dissertation: KEEPING IT IN FRAME: MONITORING RIBOSOMAL FRAMESHIFTING DURING TRANSLATION

Jamie A. Kelly
Doctor of Philosophy, 2022

Dissertation directed by: Professor Jonathan D. Dinman, PhD.
Department of Cell Biology and Molecular Genetics

Programmed -1 ribosomal frameshifting (-1 PRF) is a molecular mechanism that redirects translating ribosomes into a new reading frame. It is widely used by RNA viruses to conserve genome space while expanding the viral proteome and it can help regulate gene expression in eukaryotic cells. Strict regulation of both programmed and non-programmed frameshift events are essential to translational fidelity. This dissertation explores the -1 PRF element of SARS coronavirus 2 (SARS-CoV-2) and the -1 PRF inhibitor, Shiftless. We comparatively analyzed the structural and functional conservation of -1 PRF elements in SARS-CoV and SARS-CoV-2. Both -1 PRF structure and frameshift efficiency were highly conserved between the two viruses and a small molecule effective against SARS-CoV -1 PRF significantly decreased frameshift efficiency in SARS-CoV-2. This suggests -1 PRF is an attractive antiviral target and could be a useful tool to combat the SARS-CoV-2 pandemic or future outbreaks of similar coronaviruses. The innate immune system targets viral frameshifting using an interferon-stimulated -1 PRF inhibitor called Shiftless (SFL) that binds, arrests, and terminates translation of -1 frameshifted ribosomes. We found that SFL is not only expressed in response to interferon but that it may have a role in general translational fidelity. SFL is constitutively expressed at low levels in human-derived cell lines and its effects are not limited to -1 PRF signals. Disruption of SFL homeostasis results in reciprocal 2-fold changes to recoding efficiencies in a panel of human and viral-derived translational recoding signals, decreases reporter gene expression, and decreases mRNA steady state abundances. Additionally, SFL over or under expression combined with knockdown of prominent ribosome-associated protein quality control (RQC) proteins reveals

that SFL is epistatic to RQC. These results suggest that SFL has a role in general translational fidelity monitoring for spontaneously frameshifted ribosomes in addition to its role as a member of the innate immune response.

KEEPING IT IN FRAME: MONITORING RIBOSOMAL FRAMESHIFTING DURING TRANSLATION

by

Jamie A. Kelly

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Jonathan D. Dinman, Chair
Professor Jeffrey J. DeStefano
Associate Professor Antony M. Jose
Assistant Professor Margaret A. Scull
Professor Theodore K. Dayie, Dean's Representative

Foreword

This dissertation uses manuscripts that were previously published in *Virology* and the *Journal of Biological Chemistry.* As first author of these manuscripts, I certify that I was a substantial contributor to the experimental design, data collection, data analysis and writing of each publication.

Dedication

This dissertation is dedicated to my family. You are my greatest source of strength and inspiration.

Acknowledgements

First and foremost, I would like to thank my advisor Dr. Jonathan Dinman. Your mentorship, guidance, and encouragement has helped me become the scientist I am today.

To my advisory committee: thank you for your insightful discussions, support, and guidance throughout my time at UMD.

I would also like to thank the past and present members of the Dinman laboratory, especially Dr. Carol Vieria and Alex Olson. You have been wonderful colleagues and friends, I cannot thank you enough for your support, discussions, troubleshooting tips, baseball games, and above all: helping me find where things are in the lab.

To my research collaborators, especially those from the laboratory of Dr. Michael Woodside at the University of Alberta, thank you for insightful discussions and lending your expertise to the manuscripts in this dissertation.

To the UMD Virology Program for their support, excellent seminars, and fun discussions about viruses.

To my friends who have stood by me during my time in graduate school: I cannot thank you enough for your love and unwavering support. A special thank you to Anders Johnson for our weekly chats, and to Kailey MacDonald and Alexis LaChapelle for curating and publishing a phenomenal collection of ‘Jamie stories’.

Most importantly, I would like to thank my parents and my sister, Shannon, for their endless love, support, and for fighting beside me since day one. I could not have endured the challenges that I have faced without you by my side every step of the way.

The works in this dissertation were supported in part by the Defense Threat Reduction Agency grant HDTRA 1-18-1-0045 to Dr. Jonathan D. Dinman, National Institutes of Health grant R01 GM117177 to Dr. Jonathan D. Dinman, University of Maryland Coronavirus Research Program Seed Grant to Dr. Jonathan D. Dinman, and by the National Institute of Health T32 training grant 2T32AI051967-06A1 to the University of Maryland Virology Program.

Table of Contents

List of Tables

List of Figures

Supplemental Figures

List of Abbreviations

µg	Microgram
µL	Microliter
µM	Micromolar
-1 PRF	Programmed -1 ribosomal frameshift
3' UTR	3' untranslated region
ABCE1	ATP Binding Cassette Subfamily E Member 1
AcGFP	Aequorea coerulescens green fluorescent protein
AH	Attenuator hairpin
ANOVA	Analysis of variance
ASCC2	Activating Signal Cointegrator 1 Complex Subunit 2
ASCC3	Activating Signal Cointegrator 1 Complex Subunit 3
BLCA	Bladder carcinoma
C19orf66	Chromosome 19 open reading frame 66
CCR5	Chemokine receptor 5
cDNA	Complementary DNA
Chr.	Chromosome
CMV	Cytomegalovirus
CO_2	Carbon dioxide
CoV	Coronavirus
COVID-19	Coronavirus disease 2019
CRISPR	Clustered regularly interspaced short palindromic repeats
DENV	Dengue virus
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
dsRNA	Double-stranded RNA
eIF2	Eukaryotic initiation factor 2
ER	Endoplasmic Reticulum
eRF1	Eukaryotic release factor 1
eRF3	Eukaryotic release factor 3
GFP	Green fluorescent protein
gRNA	Genomic RNA
GTP	Guanosine triphosphate
HCV	Hepatitis C virus
HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV-1	Human immunodeficiency virus 1
ISR	Integrated Stress Response
KCl	Potassium chloride
Kd	Dissociation constant
LARP1	La-related protein 1
LIX	Life Sciences X-ray Scattering Beamline
LUAD	Lung adenocarcinoma
mCherry	Monomeric cherry fluorescent protein
MERS	Middle Eastern Respiratory Syndrome
MESO	Mesothelioma
mL	Milliliter
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger RNA
MTDB	2-methylthiazol-4-ylmethyl)-[1,4]diazepane-1-carbonyl]amino}benzoic acid ethyl ester
Myc	Glu-Gln-Lys-Leu-Ile-Ser-Glu-Glu-Asp-Leu polypeptide tag
NaCl	Sodium chloride
ng	Nanogram

NGD	No-go decay
nM	Nanomolar
NMD	Nonsense mediated decay
NP -1FS	Non-programmed -1 frameshift
n.s	Not significant
NSD	Nonstop decay
OAZ1	Ornithine decarboxylase antizyme 1
ORF	Open reading frame
PABPC1	Poly(A) binding protein cytoplasmic 1
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PEG10	Paternally expressed gene 10
PIC	Preinitiation complex
Poly-A	Poly adenosine
pSGDluc	Stop-Go dual luciferase reporter
PTC	Premature Termination Codon
Puro	Puromycin
q	Scattering vector
qRT PCR	Quantitative Reverse Transcriptase Polymerase Chain Reaction
RCC	Renal Clear Cell Carcinoma
RdRP	RNA-dependent RNA Polymerase
RIBOTAC	Ribonuclease Targeting Chimera
RNA	Ribonucleic acid
Rq	Radius of gyration
RQC	Ribosome-associated protein Quality Control
RQT	Ribosome-associated protein Quality Control-trigger complex
RRL	Rabbit Reticulocyte Lysate
rRNA	Ribosomal RNA
RT PCR	Reverse Transcriptase PCR
SARS	Severe Acute Respiratory Syndrome
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SAXS	Small Angle X-ray Scattering
SEC	Size Exclusion Chromatography
SE	Standard error
SFL	Shiftless
SFLS	Shiftless Short splice isoform
sgRNA	Subgenomic RNA
shSFL	Shiftless short hairpin RNA
SINV	Sindbis virus
SKCM	Melanoma
ssM	Slippery site Mutant
St1	Stem 1
St2	Stem 2
St3	Stem 3
TGCA	The Cancer Genome Atlas
TMD	Transmembrane Domain
TRIP4	Thyroid Hormone Receptor Interactor 4
tRNA	Transfer RNA
UPF1	Up-Frameshift Suppressor 1
VEEV	Venezuelan Equine Encephalitis Virus
WT	Wildtype

Chapter 1: Introduction

Translation and the ribosome

Translation is the process of decoding the genetic information encoded in an mRNA into protein. This is executed by a molecular machine called the ribosome. In eukaryotes, ribosomes are composed of 40S and 60S subunits that combine to form the 80S ribosome (**Figure 1B**)[1,2]. A combination of four ribosomal RNAs (rRNA) and 76 proteins collectively make up these subunits[2,3]. During translation, the ribosome moves along a messenger RNA (mRNA), recruits transfer RNAs (tRNAs) to read mRNAs three nucleotides at a time and catalyzes the addition of amino acids to produce proteins. These decoding interactions between mRNA and tRNAs occur in the 40S or small ribosomal subunit while addition of amino acids to the nascent polypeptide chain occur in the 60S subunit. Translation occurs in three distinct steps: initiation, elongation, and termination (**Fig. 1A**).

Initiation

Initiation involves recruitment of translation factors to the beginning, i.e. the 5' ends, of mRNAs. Eukaryotic initiation factor 2 (eIF2) loads a Met-tRNA onto the small (40S) ribosomal subunit with the assistance of eIF1, eIF1A, eIF5, and eIF3, forming the preinitiation complex (PIC)[4]. The PIC attaches to the 5' end of an mRNA though interactions with the eIF4 complex and scans downstream in the 3' direction until it encounters an AUG start codon in a "favorable context". Once a start codon is recognized, eIF1 dissociates from the PIC and the 60S ribosomal subunit joins the complex. eIF5B and eIF2 dissociate when the 60S subunit joins and the 80S ribosome is ready for translation elongation[4].

Elongation

Each round of translation elongation involves three steps to add amino acids to a growing polypeptide chain: 1) tRNA decoding 2) peptide bond formation and 3) translocation of the mRNA-tRNA complex[1]. These steps are repeated for each codon in the mRNA until termination of the transcript. tRNA decoding, or selection, occurs when eEF1A delivers aminoacyl-tRNAs to the A-site of the ribosome. Once a matching codon-anticodon interaction is sensed, eEF1A hydrolyses a GTP on the tRNA, releasing it from

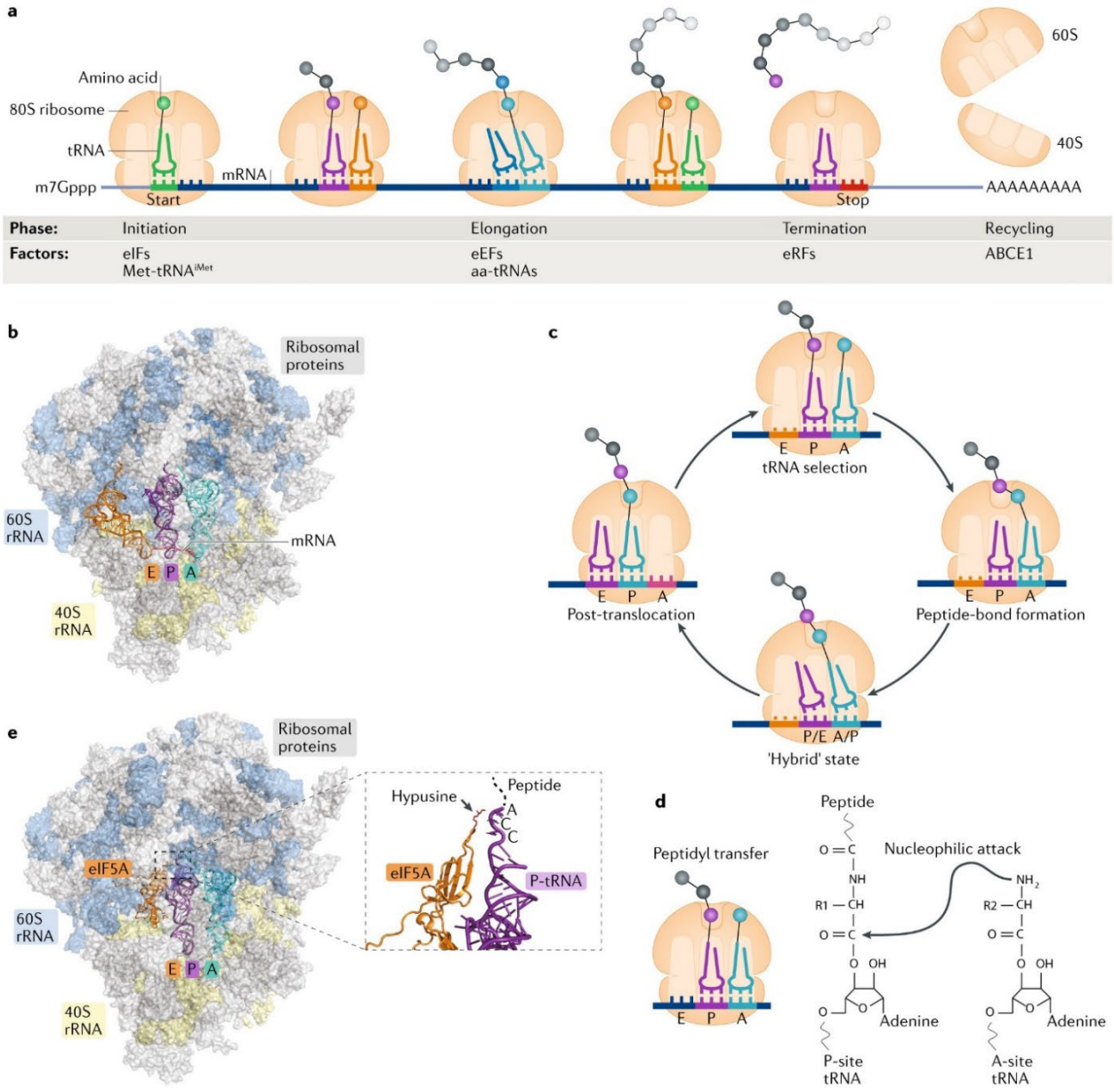
eEF1A and allowing it to fully accommodate into the A- site of the ribosome[4]. Next, the amino acid in the A- site reacts with the amino acid in the P- site and forms a peptide bond, extending the peptide chain by one residue (**Fig. 1D**)[4,5]. As the peptide bond forms between the two amino acids, the ribosomal subunits rotate and eEF2 translocates the mRNA-tRNA complex to move the tRNAs in the A- and P- sites into the P- and E- sites, respectively[1]. The tRNA in the E- site is released and the process repeats for the next codon in the mRNA.

Termination

Elongation continues until the ribosome encounters a stop codon in the ribosomal A- site. Termination codons are recognized by release factor 1 (eRF1) in eukaryotes. eRF3 delivers eRF1 to the stop codon where eRF1 binds and facilitates the release of the nascent peptide[1]. After the peptide is released, ABCE1 binds eRF1 and facilitates ribosome dissociation into the 60S and 40S ribosomal subunits[6]. The subunits are then bound by initiation factors and can be used to translate another mRNA.

Figure 1. Overview of eukaryotic translation.

(A) Overview of translation. (B). Structural model of the yeast 80S ribosome depicting 60S rRNA in light blue, 40S rRNA in yellow and ribosomal proteins in gray. The exit (E, orange), peptidyl (P, purple) and aminoacyl (A, cyan) sites for tRNA binding in the ribosome are indicated with mRNA-tRNA interactions taking place in the 40S subunit while tRNAs reside in the E, P, an A sites of the large subunit. (C) Cartoon depicting the elongation cycle. Clockwise from top: decoding and tRNA selection in the A-site, peptide bond formation between the amino acid in the A-side and the nascent peptide, 'Hybrid' state ribosome undergoing tRNA translocation, and post-translocation state with an empty A-site. (D) Peptidyl transfer reaction during elongation. (E) Structural model of the 80S ribosome bound to eIF5A. eIF5A stabilizes the peptidyl-tRNA for nucleophilic attack during peptide bond formation using a hypusine motif[1,7]. Figure 1 in Schuller and Green[1].


a
Amino acid
80S ribosome
tRNA
m7Gppp
mRNA
Start
Stop
AAAAAAAAA
60S
40S
Phase: Initiation Elongation Termination Recycling
Factors: eIFs Met-tRNA^iMet eEFs aa-tRNAs eRFs ABCE1
b
Ribosomal proteins
60S rRNA
mRNA
E P A
40S rRNA
c
E P A
tRNA selection
Peptide-bond formation
P/E A/P
'Hybrid' state
Post-translocation
d
Peptidyl transfer
Peptide
Nucleophilic attack
Adenine
P-site tRNA
A-site tRNA
e
Ribosomal proteins
eIF5A
60S rRNA
40S rRNA
Hypusine
Peptide
P-tRNA

Translational fidelity

Translational fidelity is a crucial aspect of gene expression. While DNA replication and RNA transcription utilize highly efficient co-transcriptional proofreading mechanisms, protein translation lacks such a mechanism, making it the rate-limiting step of gene expression. During translation, errors are the result of a missense peptide substitution or a failure to translate the full mRNA[8]. Missense errors in translation occur when one amino acid is substituted for another during protein synthesis and likely result from either a codon-anticodon mismatch in the ribosome or an erroneously charged tRNA[9]. These types of errors occur at a rate of $5x10^{-5}$ to $5x10^{-3}$ errors per codon, with an average of 10^{-4} errors per codon[8,9]. Processivity errors, or "translation accidents", are another major source of inaccuracy during translation. These types of errors occur when a ribosome fails to complete translation of a peptide due to events including ribosome drop off, early termination, or spontaneous frameshifting[8]. Errors in processivity such as spontaneous frameshifting occur at a rate of 10^{-4} to 10^{-5} frameshifts per codon[8–10].

Ribosome-associated protein quality control

Given that ribosome biogenesis and translation consume over 60% of the energy in the cell[11], translational errors pose a significant energy waste problem to cells. Accordingly, cells have evolved numerous mechanisms to prevent ribosome stalling and facilitating ribosome subunit recycling. These ribosome-associated quality control (RQC) mechanisms are used to identify and degrade defective mRNAs and to resolve other elongation distress causing events that inhibit productive translation. Accordingly, RQC is an important contributor to translational fidelity. It resides at the intersection of mRNA surveillance and protein degradation and is dictated by the state of the ribosome itself[12]. RQC is initiated by stalled ribosomes resulting from damaged or truncated mRNAs[13,14], lack of certain amino acids or tRNAs[15], or excessive secondary mRNA structures[12,16]. Stalling of a translating ribosome signals that there is an issue with the mRNA or nascent polypeptide, resulting in mRNA and peptide degradation[12].

Oftentimes these stalls result in ribosome collisions[17–19] with a trailing ribosome, leading to the formation of disomes. Disome formation initiates the RQC-trigger subcomplex (RQT)[17]. When collisions occur between ribosomes stalled on internal ORFs, the downstream or 5' ribosome enters a rotated conformation[17]. The collision interface between the two ribosomes is recognized by E3 ligase ZNF598 (Hel2 in yeast) which ubiquitinates ribosomal proteins eS10 and uS10 (uS3 and uS10 in yeast)[17,20,21]. The RQT

complex is composed of ASCC3, ASCC2 and TRIP4 in mammals (yeast homologs Slh1, Cue3, and Rqt4) and works to clear the lead ribosome involved in the collision.[17,18,22]. Mild collisions can be resolved by the RQT complex itself, resulting in continued translation and an intact mRNA[19,23]. More severe collisions require ribosomes to be 'rescued' by 1) the no go decay pathway (NGD) 2) the nonsense mediated decay pathway (NMD) or 3) the nonstop decay pathway (NSD) which results in termination of elongation and destruction of the mRNA[12,21,24].

Nonsense-mediated decay

NMD is triggered by a premature termination codon (PTC) (**Fig. 2A**). During canonical translation and termination, a stop codon in proximity to the poly(A) tail of an mRNA facilitates interactions between the ribosome, eRF3, and poly(A)- binding proteins that result in peptide release, ribosome recycling and end translation of that message. During NMD, however, the correct juxtaposition of the stop codon and appropriate release factors is not present, signaling the mRNA being translated is erroneous. For example, NMD can be activated by the position of splicing sites (exon-junction complex-stimulated decay) or the length of the 3' UTR (3' UTR- stimulated decay)[25–28]. Though components of NMD remain elusive, the surveillance mechanism is orchestrated by UPF proteins UPF1, UPF2, UPF3 and the SMG1 complex. UPF1 is an RNA helicase that associates with eRF1 and eRF3 and acts as the central player in NMD—determining which targets are substrates for decay[29]. In the context of a PTC, UPF1 is phosphorylated by SMG1[30]. Phosphorylated UPF1 recruits nucleases that degrade the targeted mRNA[29–31].

No-go decay

NGD occurs when a ribosome stalls during normal elongation, usually due to secondary RNA structures impeding the ribosome from continuing forward (**Fig. 2B**). Stalled ribosomes with an empty A-site are recognized by NGD factors Hbs1L and Pelota (Hbs1 and Dom34 in yeast)[16]. These factors are similar to ribosomal release factors eRF1 and eRF3, bind to the ribosome, destabilize the 80S subunit interface and promote dissociation by recruiting ABCE1 (Rli1 in yeast)[16,21,32,33].

Nonstop decay

NSD occurs when mutations in an mRNA remove the normal termination codon in the transcript and cause the ribosome to continue translation into the poly-A tail of the mRNA (**Fig 2C**). Translation of the poly-A tail ultimately causes ribosomes to stall, collide, and recruit the SKI complex for ribosome rescue and degradation of the problematic mRNA[21].

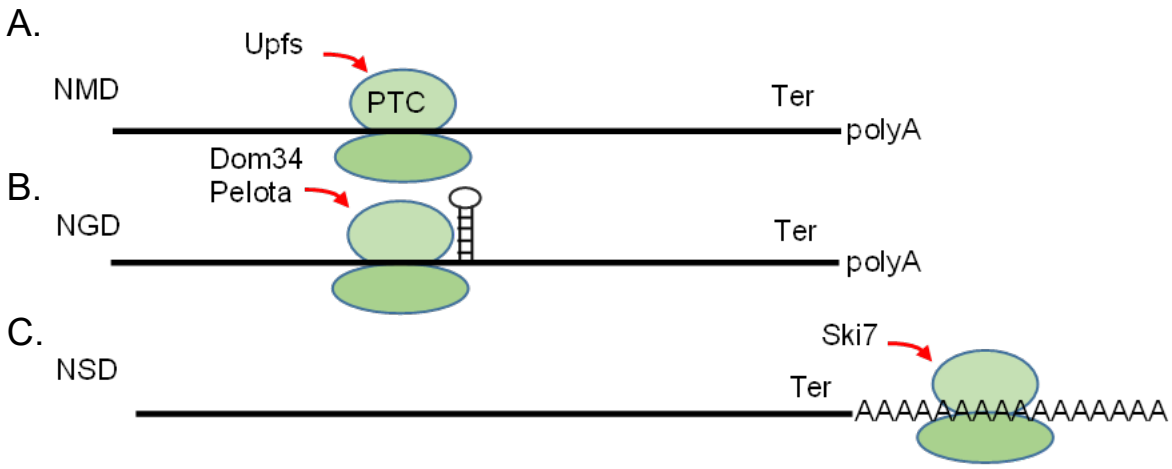


Figure 2. mRNA surveillance pathways

(A) Nonsense-mediated decay monitors for ribosomes paused at premature termination codons (PTC). (B) No go decay monitors for ribosomes stalled at secondary RNA structures. (C) Nonstop decay monitors for ribosomes that continue translation into the poly-A tail due to a missing termination codon.

Translational recoding

Although the translational apparatus goes to great lengths to maintain translational fidelity and ensure proteins are made as transcribed, translational recoding signals have evolved to reprogram the ribosome to deviate from what is written in the genetic code. Recoding can be used as a quality control check, to regulate levels of certain proteins or alter the rate of translation. Translational recoding mechanisms include termination suppression, ribosomal bypassing, or 'hopping', and programmed ribosomal frameshifting[34–37]. Termination suppression occurs when a ribosome does not stop at a termination codon at the end of an open reading frame. Though this occurs naturally at low rates, *cis*- and *trans*-acting factors can promote higher rates of termination codon readthrough[35,38,39]. Bypassing occurs

when the ribosome skips over a portion of the mRNA and produces one polypeptide from a discontinuous reading frame[34,36]. Programmed ribosomal frameshifting redirects the translating ribosome into a new reading frame. RNA structures and rare codons can cause ribosomes to shift one base in the 5' direction (-1 frameshift) or one base in the 3' direction (+1 frameshift)[34,37,40–42].

Programmed -1 ribosomal frameshifting

Programmed -1 ribosomal frameshifting (-1 PRF) is a translational recoding mechanism that redirects a translating ribosome to slip backwards (in the 5' direction) by one base, re-framing it into a new, -1 reading frame. It was first identified in Rous Sarcoma virus as a means to synthesize and regulate amounts of the viral polymerase[37]. Since then, it has been identified in a large number of RNA viruses and is a method to expand viral proteomes while conserving genome space[43–46]. In addition to viruses, some eukaryotic genes like CCR5 and PEG10 utilize -1 PRF to regulate gene expression and aid with protein folding[47–49].

Molecular mechanisms of -1 PRF

Programmed -1 ribosomal frameshifting functions by redirecting a translating ribosome into a new reading frame. Though frameshift signal sequence can vary greatly, their structure is composed of three parts: a heptameric slippery site, a 10-12 nt spacer and a stimulatory RNA element, generally a pseudoknot[49]. The slippery site has the sequence NNN WWW H in which N is a string of any 3 identical bases, W is A or U, and H is any nucleotide other than G[49] (**Fig. 3**). The "simultaneous slippage" model of frameshifting outlines the importance of the slippery site in -1 PRF[50]. In this model, an elongating ribosome encounters the -1 PRF stimulatory element and pauses with the slippery site positioned over the A- and P-sites in the 0-frame. During a -1 frameshift, the two adjacent tRNAs in the slippery site dissociate and the non-wobble bases in the tRNA re-anneal in the -1 frame after moving one base in the 5' direction[50].

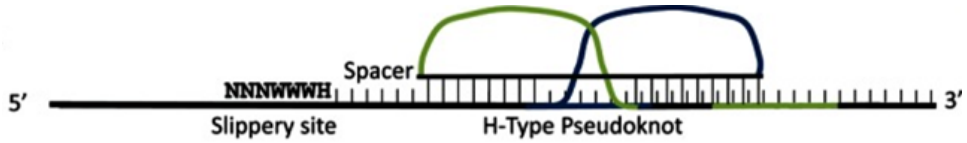


Figure 3. -1 PRF signal

Diagram of a typical -1 PRF signal comprising of a slippery site, spacer, and RNA pseudoknot.

In the "torsional restraint" model of -1 PRF, frameshifting is directed by torsional resistance in the stimulatory structure of the frameshift signal[51]. Many -1 PRF stimulatory elements are composed of an RNA pseudoknot structure in which a single stranded RNA folds to form two helical structures, or stems, connected by single-stranded loops[52]. When an elongating ribosome encounters a pseudoknot, it must unwind both stems of the RNA structure to continue translating. In the context of frameshifting, the ribosome will pause at the pseudoknot structure and position the ribosome A- and P- sites over the slippery sequence. The torsional restraint model suggests that while the ribosome is unwinding stem 1 of the RNA pseudoknot, stem 2 is forced to rotate and becomes supercoiled **(Fig. 4)**[51]. Eventually, the resistance caused by supercoiling of stem 2 will counteract the forward movement of the ribosome, positioning the A- and P- sites over the slippery sequence.

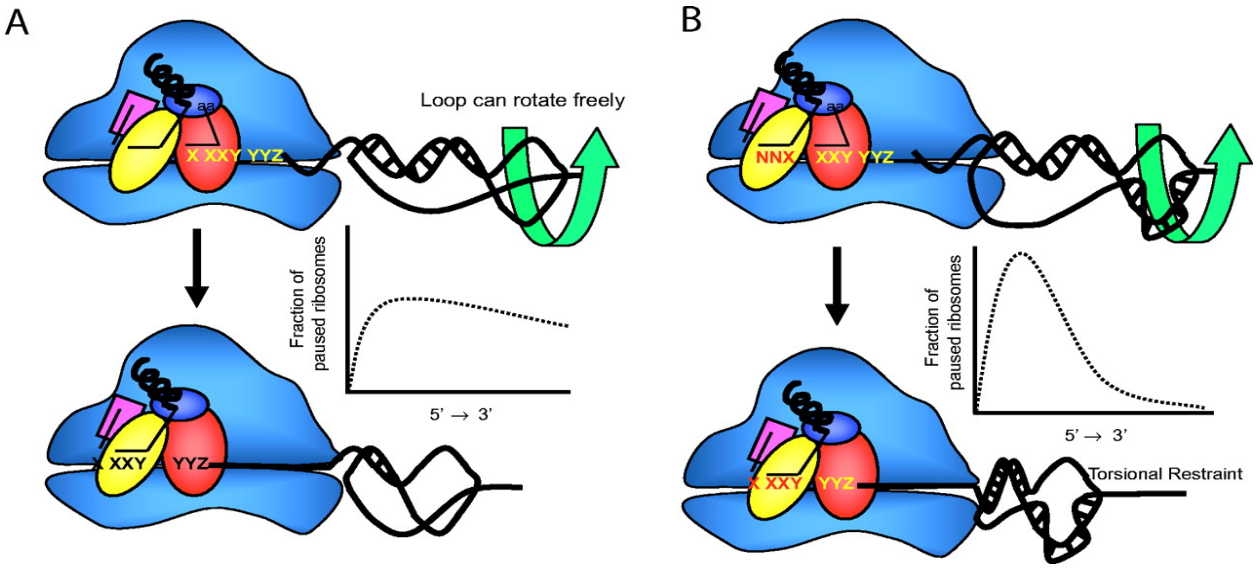


Figure 4. The Torsional Restraint model.

(A) Simple stem-loop structures in mRNA provide little resistance for an elongating ribosome as the loop can rotate freely while RNA helicase unwinds it. The energy required to denature the stem-loop is distributed along the length of the structure. (B) RNA pseudoknot structures provide more resistance for an elongating ribosome. The presence of stem 2 of the pseudoknot prohibits stem 1 from rotating freely. More energy is required to unwind the pseudoknot. In the context of a frameshift signal, torsional resistance in the RNA caused by unwinding stem 1 and super-coiling stem 2 forces the ribosome to pause with the slippery site positioned over the A- and P- sites of the ribosome. The theoretical 'graphs' are inserted to further illustrate this point, though they are not based on any actual measurements. Figure 1 from Plant and Dinman[51].

The 9Å solution was postulated from atomic resolution structural data of translating ribosomes[53,54]. During aminoacyl-tRNA accommodation, mRNA is pulled into the ribosome by 9Å[53]. The 9Å solution suggests that the downstream stimulatory structure of a -1 PRF element not only forces ribosomes to pause but also impedes the ribosome from pulling the mRNA during aa-tRNA accommodation. This stretches the mRNA segment between the slippery sequence and stimulatory structure[54]. In this situation, there are two options to resolve tension on the mRNA: 1) resolve the stimulatory structure or 2) shift one base in the 5' direction and enter the -1-reading frame.

-1 PRF in virology

-1 PRF is a crucial event during replication of many RNA viruses. It regulates protein ratios and directs the different stages of infection. Disrupting this process has been shown to have devastating consequences for viral propagation[55]. Because of its unique RNA structures and importance for regulating viral polymerase expression, -1 PRF is an attractive target for vaccine development and antiviral therapeutics[49,55–61]. For example, ablating -1 PRF in Venezuelan Equine Encephalitis virus (VEEV) by introducing silent mutations into the slippery site of the -1 PRF signal resulted in attenuated neuropathogenicity in mice[57]. Other work has shown that targeting different components of -1 PRF signals can be extremely effective as antivirals. Most recently, -1 PRF has emerged as a promising target to combat the COVID-19 pandemic caused by SARS Coronavirus 2 (SARS-CoV-2)[62–65].

Betacoronaviruses like SARS-CoV-2 have a ~30kb plus-sense RNA genome (**Figure 5**) [62]. While the coronavirus genome contains nine different ORFs, roughly two-thirds of the genome is comprised of ORF1a and ORF1b[66]. ORF1a encodes immediate early genes involved in dampening the host immune response upon entry into the cell, while ORF1b encodes the viral RNA-dependent RNA polymerase (RdRP). The remaining ORFs are transcribed from subgenomic RNAs (sgRNA) that encode proteins needed for virion assembly and egress from the host cell. One key component of this process is the juxtaposition of ORF1a and ORF1b. ORF1b is out of frame with respect to ORF1a and is expressed by directing ribosomes into the correct frame using a -1 PRF signal. In this case, -1 PRF functions as a switch between the initial viral takeover of the cell by ORF1a encoded proteins and the synthesis of viral RNAs by RdRP encoded in ORF1b.

Figure 5. Betacoronavirus genome organization.

The betacoronavirus genome contains nine ORFs that can be separated into three groups: immediate early genes in ORF1a (blue), early genes encoded in ORF1b (red), and late genes encoded by sgRNAs in ORF2-9 (purple). ORF1a and ORF1b are separated by a -1 PRF signal (yellow). Figure 1A from Kelly et al[67].

Coronaviruses like SARS-CoV contain a unique three-stemmed RNA pseudoknot that drives -1 PRF in addition to an upstream attenuator hairpin (AH) that acts as an additional regulator of frameshifting[58,59]. Targeting the AH or -1 PRF signal of coronaviruses like SARS-CoV and SARS-CoV-2 with small molecule compounds is a promising avenue to combat the SARS-CoV-2 pandemic in addition to future emergences of related coronaviruses[59,60,64,68]. Chapters 2 and 3 of this document describe our research characterizing the -1 PRF signal of SARS-CoV-2 and efforts to identify small molecules that target this element, and similar -1 PRF signals in other coronaviruses.

In addition to *cis*-acting factors or small molecules modifying viral -1 PRF, *trans*-acting factors like the -1 PRF inhibitor, Shiftless, can influence frameshift efficiency during infection[69,70]. Shiftless was originally identified as an effector in the type I interferon antiviral response[71]. Subsequent studies refined its role as an antiviral inhibitor involved with targeting viral replication by interacting with viral replication factories and translation of viral proteins[72,73] and eventually identified its role in targeting -1 PRF[69]. Chapter 4 of this document describes our research into the role of Shiftless with regard to both translational recoding and RQC.

Possible roles of -1 PRF in cellular gene expression

The role of -1 PRF in cellular gene expression is an emerging field and could have significant implications in human disease. It is estimated that roughly 10% of genes in eukaryotes contain a -1 PRF signal[74]. The first human -1 PRF signal was identified in chemokine receptor CCR5[47,75]. This frameshift signal has an efficiency of ~10-15% and redirects ribosomes into a reading frame containing a PTC,

triggering the NMD pathway[47]. Other human frameshift signals include ornithine decarboxylase antizyme 1 (OAZ1)[42,76], and parentally expressed gene 10 (PEG10)[48]. Though much still needs to be uncovered about the role of -1 PRF in eukaryotes, translational recoding events likely act as quality control measures and help ensure correct levels of these genes in the cell and regulate mRNA abundances[47].

Additionally, frameshifting may be linked to proper folding of the nascent peptide[77,78]. The forces generated by folding of a nascent peptide induce mechanical tension on the translating ribosome. In the context of a -1 PRF signal, this results in enhanced PRF as seen in Sindbis virus[78]. Though more investigation is needed, tensions on a -1 PRF signal could also be applied to human -1 PRF signals. For example, CCR5 contains five transmembrane domains[79], three of which are upstream of its -1 PRF signal. As the mRNA is translated, the nascent peptide is folded and TMDs are integrated into the membrane, providing additional tensions on the mRNA. If folded correctly, tensions on the ribosome and mRNA likely avoid a -1 frameshift. If improper folding occurs, this leads to a -1 frameshift and directs the ribosome into a PTC, subsequently tagging the transcript and peptide for degradation by NMD and RQC[47].

Research overview

This dissertation explores the importance of monitoring and regulating ribosomal frameshifting during translation and maintaining translational fidelity. First, it delves into using -1 PRF in coronaviruses as a target for antiviral therapeutics. Since -1 PRF is a critical and highly regulated component of the viral replication program, manipulating it could prove to be a useful tool in combatting the SARS-CoV-2 pandemic and future emergences of novel coronaviruses. Second, it characterizes the structural and functional conservation of -1 PRF in SARS-CoV and SARS-CoV-2. High conservation between the two viruses (as well as variants) recapitulates the importance of this mechanism in coronavirus replication and suggests the possibility of developing broad-spectrum antiviral therapeutics targeting frameshifting. Finally, this dissertation explores the -1 PRF inhibitor, Shiftless, and establishes it as a monitor of general translational fidelity and member of RQC in addition to its role in the innate immune response.

Chapter 2: Structural and functional conservation of the programmed −1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2)[80]

Jamie A. Kelly[1], Alexandra N. Olson[1], Krishna Neupane[2], Sneha Munshi[2], Josue San Emeterio[3], Lois Pollack[3],*, Michael T. Woodside[2],*, and Jonathan D. Dinman[1],*

1. Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland, USA, the

2. Department of Physics, University of Alberta, Edmonton, Alberta, Canada, and the

3. School of Applied and Engineering Physics,

Cornell University, Ithaca, New York, USA

*Corresponding authors Lois Pollack, lp26@cornell.edu; Michael T. Woodside, mwoodsid@ualberta.ca; Jonathan D. Dinman, dinman@umd.edu.

Work adapted from Journal of Biological Chemistry, July 2020. DOI: 10.1074/jbc.AC120.013449

Abstract

Approximately 17 years after the severe acute respiratory syndrome coronavirus (SARS-CoV) epidemic, the world is currently facing the COVID-19 pandemic caused by SARS corona virus 2 (SARS-CoV-2). According to the most optimistic projections, it will take more than a year to develop a vaccine, so the best short-term strategy may lie in identifying virus-specific targets for small molecule–based interventions. All coronaviruses utilize a molecular mechanism called programmed -1 ribosomal frameshift (-1 PRF) to control the relative expression of their proteins. Previous analyses of SARS-CoV have revealed that it employs a structurally unique three-stemmed mRNA pseudoknot that stimulates high -1 PRF rates and that it also harbors a -1 PRF attenuation element. Altering -1 PRF activity impairs virus replication, suggesting that this activity may be therapeutically targeted. Here, we comparatively analyzed the SARS-CoV and SARS-CoV-2 frameshift signals. Structural and functional analyses revealed that both elements promote similar -1 PRF rates and that silent coding mutations in the slippery sites and in all three stems of the pseudoknot strongly ablate -1 PRF activity. We noted that the upstream attenuator hairpin activity is also functionally retained in both viruses, despite differences in the primary sequence in this region. Small-angle X-ray scattering analyses indicated that the pseudoknots in SARS-CoV and SARS-CoV-2 have the same conformation. Finally, a small molecule previously shown to bind the SARS-CoV pseudoknot and inhibit -1 PRF was similarly effective against -1 PRF in SARS-CoV-2, suggesting that such frameshift inhibitors may be promising lead compounds to combat the current COVID-19 pandemic.

Introduction

SARS-CoV-2, the etiological agent of COVID-19, is a member of the coronavirus family[81]. Coronaviruses have single-strand RNA genomes that harbor two long ORFs that occupy approximately two-thirds of the 5' end of the genomic RNA[81] (ORF1 and ORF2), followed by several ORFs that are expressed late in the viral replication cycle from subgenomic RNAs (**Fig. 5A**)[82] . In general, the immediate early proteins encoded by ORF1a are involved in ablating the host cellular innate immune response, whereas the early proteins encoded in ORF1b are involved in genome replication and RNA synthesis. These functions include generating the minus-strand replicative intermediate, new plus-strand genomic RNAs, and subgenomic RNAs, which mostly encode structural, late proteins. ORF1b is out of frame with

respect to ORF1a, and all coronaviruses utilize a molecular mechanism called programmed -1 ribosomal frameshifting (-1 PRF) as a means to synthesize the ORF2-encoded proteins[83,84]. -1 PRF is a mechanism in which cis-acting elements in the mRNA direct elongating ribosomes to shift the reading frame by 1 base in the 5' direction. The use of a -1 PRF mechanism for expression of a viral gene was first identified in the Rous sarcoma virus[37]. A -1 PRF mechanism was shown to be required to translate ORF1ab in a coronavirus, avian infectious bronchitis virus, 2 years later[46]. In coronaviruses, -1 PRF functions as a developmental switch, and mutations and small molecules that alter this process have deleterious effects on virus replication[55,85].

The -1 PRF signal can be broken down into three discrete parts: the "slippery site," a linker region, and a downstream stimulatory region of mRNA secondary structure, typically an mRNA pseudoknot (reviewed in Ref. 3). The primary sequence of the slippery site and its placement in relation to the incoming translational reading frame is critical: it must be N NNW WWZ (codons are shown in the incoming or 0-frame), where NNN is a stretch of three identical nucleotides, WWW is either AAA or UUU, and Z ≠ G. The linker region is less well-defined, but typically is short (1–12 nt long) and is thought to be important for determining the extent of -1 PRF in a virus-specific manner. The function of the downstream secondary structure is to induce elongating ribosomes to pause, a critical step for efficient -1 PRF to occur (reviewed in Ref. [34]). The generally accepted mechanism of -1 PRF is that the mRNA secondary structure directs elongating ribosomes to pause with its A- and P-site bound aminoacyl- and peptidyl-tRNAs positioned over the slippery site. The sequence of the slippery site allows for re-pairing of the tRNAs to the -1 frame codons after they "simultaneously slip" by 1 base in the 5' direction along the mRNA. The subsequent resolution of the downstream mRNA secondary structure allows the ribosome to continue elongation of the nascent polypeptide in the new translational reading frame. The downstream stimulatory elements are most commonly H-type mRNA pseudoknots, so called because they are composed of two co-axially stacked stem loops where the second stem is formed by base pairing between sequence in the loop of the first-stem loop and additional downstream sequence[52]. The SARS-CoV pseudoknot is more complex because it contains a third, internal stem-loop element[86–88]. Mutations affecting this structure decreased the rates of -1 PRF and had deleterious effects on virus propagation, thus suggesting that it may present a target for small-molecule therapeutics[55,85]. In addition, the presence of a hairpin located immediately 5' of the slippery

site has been reported to regulate -1 PRF by attenuating its activity[89]. Here, we report on the -1 PRF signal from SARS-CoV-2. The core -1 PRF signal is nearly identical to that of SARS-CoV, containing only a single nucleotide difference, C to A. This change maps to a loop region in the molecule that is not predicted to affect the structure of the three-stemmed pseudoknot. The primary sequence of the attenuator hairpin is less well-conserved. However, genetic analyses reveal that both elements appear to have been functionally conserved. Conservation of RNA structure is further supported by the similarity of the small-angle X-ray scattering profiles for the two pseudoknots and by the similar anti-frameshifting activity of a small-molecule ligand against both frameshift signals.

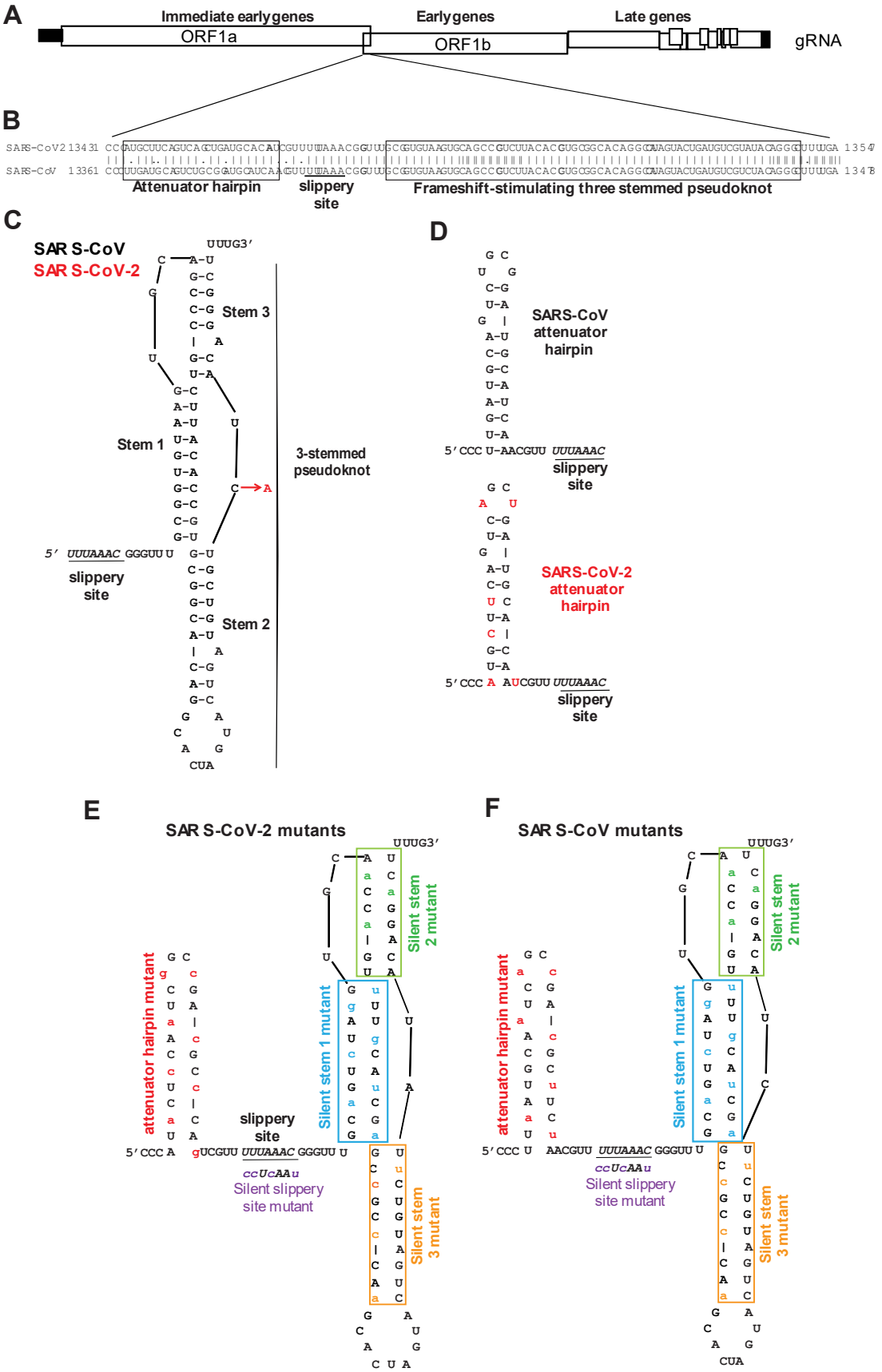
Results

Comparative structural analyses of the two -1 PRF signals

The core of the SARS-CoV -1 PRF signal begins with the U UUA AAC slippery site, followed by a 6-nt spacer region and then the three-stemmed mRNA pseudoknot that stimulates -1 PRF. A second regulatory element, called the attenuator hairpin, is located 5' of the slippery site. Pairwise analysis of the SARS-CoV and SARS-CoV-2 frameshift signals revealed that the sequence of the attenuator hairpin was less well-conserved than the frameshift-stimulating pseudoknot (**Fig. 6B**). The structure of the SARS-CoV -1 PRF signal was previously determined to include a three-stemmed pseudoknot[86]. Using this structure as a guide, the single C-to-A base difference between the core SARS-CoV and SARS-CoV-2 -1 PRF signals (**Fig. 6B**) that maps to a loop that is not predicted to alter the structure of the -1 PRF stimulating element[55] (**Fig. 6C**). In contrast, the attenuator hairpin contains six differences in the nucleotide sequence between the two viruses (**Fig. 6B**), and the SARS-CoV-2 element is predicted to be less stable than its SARS-CoV counterpart (**Fig. 6D**). To determine the importance of each of these elements, a series of silent coding mutants of both the SARS-CoV and SARS-CoV-2 sequences were constructed to disrupt the putative attenuators, slippery sites, and stems 1, 2, and 3 of the pseudoknots (**Fig. 6, E and F**).

Figure 6. Structural comparison of the SARS-CoV and SARS-CoV-2 -1 PRF signals.

(**A**) cartoon depicting SARS-CoV and SARS-CoV-2 genome organization including a -1 PRF between ORF1a and ORF1b. (**B**) pairwise analysis of the two -1 PRF signals. The attenuator elements and three-stemmed pseudoknot sequences are boxed as indicated. The U UUA AAC slippery site is underlined. (**C**) structure of the SARS-CoV -1 PRF signal[86] is composed of the 5' slippery site, a 6-nt spacer, and the three-stemmed pseudoknot stimulatory element. The single-base difference in SARS-CoV-2 (red) maps to the short loop linking stems 2 and 3. (**D**) comparison of the SARS-CoV and SARS-CoV-2 -1 PRF attenuator elements. SARS-CoV-2–specific bases are indicated in red. (**E** and **F**) silent coding mutations designed to disrupt the attenuators, slippery sites, and stems 1, 2, and 3 in the SARS-CoV-2 (**E**) and SARS-CoV (**F**) -1 PRF signals. gRNA, genomic RNA.


A
Immediate early genes
Early genes
Late genes
ORF1a
ORF1b
gRNA
B
Attenuator hairpin
slippery site
Frameshift-stimulating three stemmed pseudoknot
C
SAR S-CoV
SAR S-CoV-2
Stem 3
Stem 1
Stem 2
3-stemmed pseudoknot
slippery site
D
SARS-CoV attenuator hairpin
slippery site
SARS-CoV-2 attenuator hairpin
slippery site
E
SAR S-CoV-2 mutants
attenuator hairpin mutant
Silent stem 1 mutant
Silent stem 2 mutant
Silent stem 3 mutant
slippery site
Silent slippery site mutant
F
SAR S-CoV mutants
attenuator hairpin mutant
Silent stem 1 mutant
Silent stem 2 mutant
Silent stem 3 mutant
Silent slippery site mutant

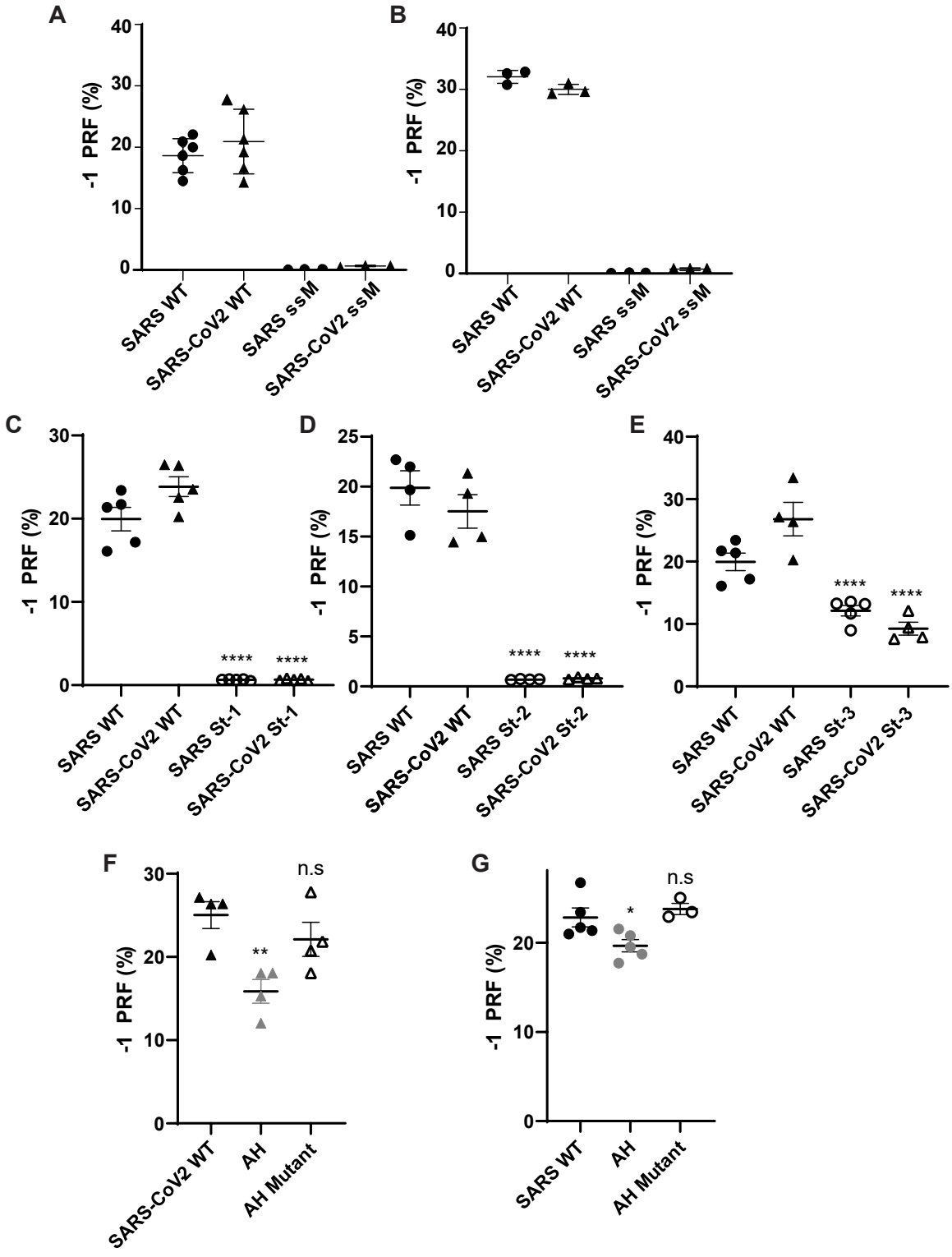
Comparative functional analyses of the two -1 PRF signals

Standard Dual-Luciferase assays were used to monitor -1 PRF activities of the two -1 PRF signals[90,91] in cultured human cell lines. For both of the elements, -1 PRF activity was ~20% in HEK (**Fig. 7A**) and ~30% in HeLa (**Fig. 7B**). Amino acid sequence silent coding mutation of the U UUA AAC slippery sites to C CUC AAC (the incoming 0-frame codons are indicated by spaces) ablated -1 PRF activity in both cases to less than 1% (**Fig. 7, A and B**), demonstrating the functional conservation of this central feature of the -1 PRF signal.

To test functional conservation of the three-stemmed pseudoknot, a series of silent 0-frame coding mutations were made to each of the stems in both the SARS-CoV and SARS-CoV-2 frameshift signals, and assays were performed in HEK cells. Disruption of stem 1 strongly suppressed the ability of both elements to promote -1 PRF, decreasing rates to 0.67$\pm$ 0.03 and 0.7 $\pm$ 0.1% for SARS-CoV and SARS-CoV-2, respectively, $p < 0.0001$ (**Fig. 7C**). Similarly, disruption of stem 2 had a strong negative impact on -1 PRF, decreasing rates to 0.68 $\pm$ 0.04% for SARS-CoV and 0.8 $\pm$ 0.1% for SARS-CoV-2; $p < 0.0001$ (**Fig. 7D**). In contrast, although disruption of stem 3 did decrease -1 PRF efficiencies, the effects were less severe, although the decreases were statistically significant (13.1 $\pm$ 0.9 and 8 $\pm$ 1% for SARS-CoV and SARS-CoV-2, respectively; $p < 0.0001$) (**Fig. 7E**). These findings support the hypothesis that the structure and function of the core -1 PRF signals have been conserved between the two viruses.

Figure 7. Functional characterization of the SARS-CoV and SARS-CoV-2 -1 PRF signals.

A and B, analyses of silent slippery site mutants. The efficiencies of -1 PRF promoted by the WT (U UUA AAC) and silent slippery site mutant (C CUC AAC) -1 PRF signals were assayed in HEK (A) and HeLa (B). ssM denotes silent slippery site mutant. C–E, analyses of the importance of the three stems in the -1 PRF stimulating RNA pseudoknot. Silent stem 1 (St-1, C), stem 2 (St-2, D), and stem3 (St-3, E) mutants were assayed in HEK cells. F and G, analyses of the attenuator hairpins. AH denotes constructs that included attenuator hairpin sequences. AH mutant denotes mutants harboring the silent coding attenuator hairpin sequences shown in Fig. 6 (E and F). Assays were performed using Dual-Luciferase assays as previously described[90,91] . Each data point represents a single biological replicate comprised of three technical replicates. Error bars denote standard error of the mean (S.E). n.s, not significant.


A
B
-1 PRF (%)
SARS WT
SARS-CoV2 WT
SARS ssM
SARS-CoV2 ssM
C
D
E
SARS St-1
SARS-CoV2 St-1
SARS St-2
SARS-CoV2 St-2
SARS St-3
SARS-CoV2 St-3
****
F
G
n.s
*
**
AH
AH Mutant

Conservation of the 5' attenuator function

Prior studies demonstrated the presence of an element located immediately 5' of the SARS-CoV slippery site that had the ability to decrease -1 PRF, called the attenuator hairpin[89]. Although less well-conserved at the primary sequence level (**Fig. 6, B and C**), addition of this sequence into the SARS-CoV-2 reporter also resulted in decreased -1 PRF efficiency:16 ± 3% compared with 25 ± 3% without the attenuator hairpin ($p < 0.01$), whereas disruption of the hairpin did not result in decreased efficiency (22 ± 4%, $p = 0.415$) (Fig. 2F). In the control experiment, the SARS-CoV attenuator also promoted decreased -1 PRF, albeit to a lesser extent (20 ± 2% compared with 23 ± 2% without the attenuator hairpin ($p = 0.04$) and 24 ± 1% with the disrupted hairpin ($p = 0.716$)) (**Fig. 7G**). Thus, the attenuation function has also been conserved between the two viruses despite the differences in primary nucleotide sequences.

Small-molecule frameshift inhibitor of SARS-CoV -1 PRF is also active against SARS-CoV-2

Based on the strong conservation of the frameshift signal between SARS-CoV and SARS-CoV-2, we tested whether a frameshift inhibitor active against the first also retained activity against the second. We focused on a small-molecule ligand previously shown to bind to the SARS-CoV pseudoknot and suppress -1 PRF, 2-{[4-(2-methyl-thiazol-4ylmethyl)-[1,4] diazepane-1-carbonyl]-amino}-benzoic acid ethyl ester, hereafter denoted as MTDB[60,61]. Comparing the -1 PRF activity from Dual-Luciferase measurements in rabbit reticulocyte lysates in the presence and absence of MTDB, we found that 5 μM MTDB reduced -1 PRF activity by almost 60%, from 36 ± 3 to 15 ± 1% (**Fig. 8**). This reduction was comparable with, but slightly smaller than, that seen previously for the SARS-CoV pseudoknot, where 0.8 μM MTDB reduced -1 PRF by roughly 60%[60].

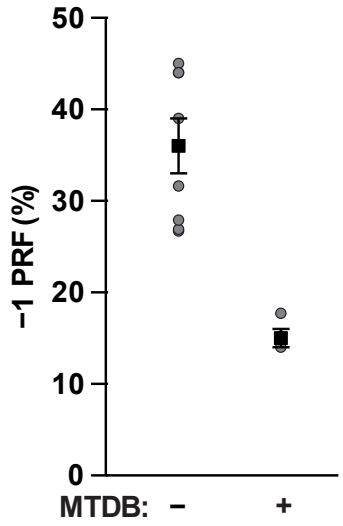


Figure 8. Small-molecule ligand MTDB inhibits -1 PRF stimulation by SARS-CoV-2 pseudoknot.

-1 PRF efficiency was reduced almost 60% in the presence of 5 μM MTDB (right), compared with -1 PRF efficiency in the absence of MTDB (left).

Solution scattering profiles of the SARS-CoV and SARS-CoV-2 pseudoknots are indistinguishable

Finally, we used small and wide-angle X-ray scattering (SAXS) to compare the solution scattering profiles of the two pseudoknots, which reflect their structure. The scattering profiles (intensity as a function of the scattering vector q) were indistinguishable for laboratory-purified samples of SARS-CoV (**Fig. 9A, blue**) and SARS-CoV-2 (**Fig. 9A, red**) pseudoknots. The difference between their scattering profiles is consistent with 0 at all q (**Fig. 9B**). The high-q portion of the profile is sensitive to the finer molecular details of the structure[92]; hence the similarity of the profiles for the two pseudoknots indicates that their structures are likely the same. Because SARS-CoV pseudoknots can dimerize[93], we also performed inline size exclusion chromatography (SEC) SAXS measurements, where the RNA was purified by SEC immediately before X-ray exposure to ensure only monomers were measured. From inline SEC-SAXS profiles (**Fig. 9A, inset**), we determined the monomer size, parameterized as the radius of gyration, Rg. We measured the same values for SARS-CoV and SARS-CoV-2 pseudoknots: Rg = 28.1 6 0.3 Å and 28.1 6 0.2 Å, respectively. The difference profile for this set is also consistent with 0 for all q (**Fig. 9C**).

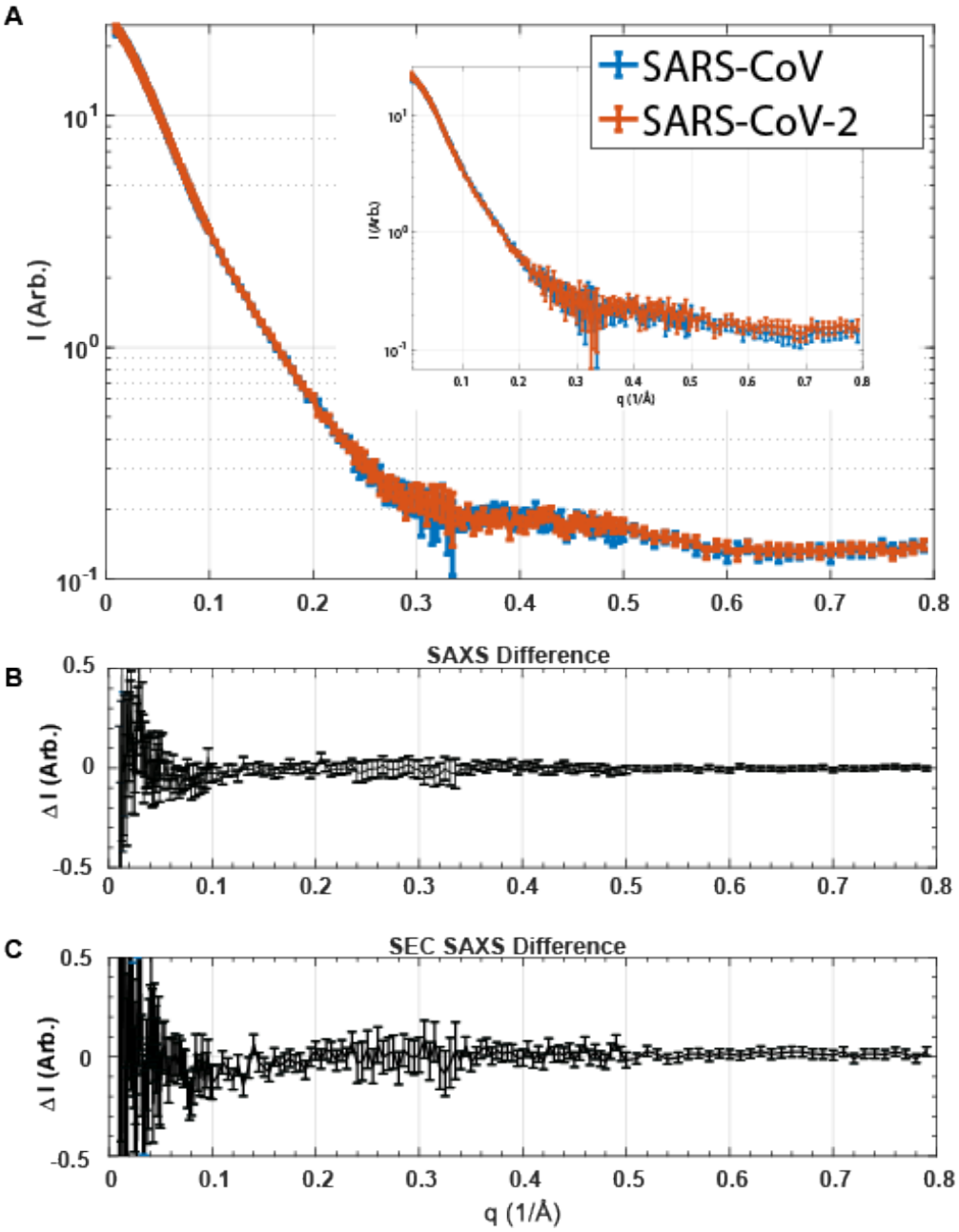


Figure 9. SAXS analyses.

A, scattering profiles from laboratory-purified SAXS samples containing pseudoknots from SARS-CoV (blue) and SARS-CoV-2 (red). Inset, scattering profiles from inline SEC-SAXS measurements, containing purely monomeric pseudoknots. B and C, difference between the scattering profiles for SARS-CoV and SARS-CoV-2 pseudoknots obtained from lab-purified SAXS (B) and inline SEC-SAXS (C) samples. Arb., arbitrary.

Discussion

These results verify that SARS-CoV-2 does indeed have a functional -1 PRF site. They also show that the properties of the frameshift signal in SARS-CoV-2 are very similar to those of the frameshift signal in SARS-CoV. Not only was the level of -1 PRF close to identical for both viruses, but disrupting stems 1 and 2 in the stimulatory pseudoknot abolished frameshifting in both cases, whereas disrupting stem 3 reduced -1 PRF but did not abolish it in each case. Furthermore, each frameshift signal featured an attenuator hairpin that promoted modestly decreased -1 PRF levels, and the global structures of two pseudoknots as reflected in the SAXS scattering profiles were virtually identical.

The very close correspondence in the properties of the frameshift-stimulatory pseudoknots from SARS-CoV and SARSCoV-2 suggests that other properties of the former that have been characterized in previous studies are highly likely to carry over to SARS-CoV-2. For example, deletion of stem 3 will likely lead to little or no change in -1 PRF, whereas mutation of the A bulge in stem 2 will likely abolish -1 PRF[55,85], the pseudoknot will likely dimerize via interactions between loop 2[93], and suppression of -1 PRF will most likely attenuate viral propagation[86]. This likely susceptibility of SARS-CoV-2 to attenuation by suppressing -1 PRF is of particular interest, because it suggests that targeting -1 PRF may provide a promising avenue for therapeutic intervention. Previous work on SARS-CoV found that antisense peptide nucleic acids could inhibit both -1 PRF and virus replication[94]. The fact that the compound MTDB, which was found in a computational search for -1 PRF inhibitors in SARS-CoV[60], is similarly active at suppressing -1 PRF in SARS-CoV-2 provides concrete evidence for small-molecule frameshifting inhibitors in SARS-CoV-2 and supports the hypothesis that the frameshift-stimulatory pseudoknot may be an attractive therapeutic target.

Experimental procedures

Identification of the SARS-CoV-2 -1 PRF signal and computational methods

The SARS-CoV-2 -1 PRF signal was identified from the complete genome sequence (NCBI sequence NC_045512.2). The EMBOSS Water pairwise alignment tool was used to identify sequences in the SARS-CoV-2 genome most similar to the SARS-CoV -1 PRF sequence. One hit was reported between bases 13,461 and 13,547 of SARS-CoV-2 that was 98.9% identical to the original SARS sequence. The SARS-CoV-2 sequence contains a single point mutation from C to A at base 13,533. EMBOSS Water was

used to generate pairwise alignments between sequences derived from SARS-CoV (GenBank entry NC_004718.3, begin nt 13361, end nt 13478) and SARS-CoV-2 (GenBank entry NC_045512.2, begin nt 13431, end nt 13547).

Preparation of plasmids and RNA transcription templates

Plasmids for cell-based Dual-Luciferase assays for SARSCoV-2 were generated by site-directed mutagenesis of the pJD2359 plasmid (SARS-CoV pSGDluc reporter plasmid)[85], introducing a single C-to-A point mutation at base 1873, corresponding to the point mutation in the SARS-CoV-2 genome (Q5 site-directed mutagenesis kit, NEB). Site-directed mutagenesis primers (Table 3) were synthesized and purified by IDT. Products were transformed into DH5a Escherichia coli cells (NEB) and spread onto LB agar plates containing 50mg/ml carbenicillin. Positive clones were verified by DNA sequencing (Genewiz). The frameshift reporter negative controls and reporter constructs containing silent mutations disrupting the -1 PRF slippery site (ssM), stem 1 (St1), stem 3 (St3), and attenuator hairpins were constructed similarly by site-directed mutagenesis. Reporters containing silent mutations to stem 2 were made by digesting pJD2257 with SalI and BamHI and ligating a DNA oligonucleotide insert (IDT) containing the silent mutations to stem 2 of SARS and SARS-CoV-2 (IDT) into the plasmid using T4 DNA ligase (NEB).

Plasmids for cell-free Dual-Luciferase assays were made as described previously[95]. Briefly, the reporter construct was made by cloning the sequence for Renilla luciferase and SARS- CoV-2 frameshift signal in the 0 frame upstream of the firefly luciferase sequence in the pISO plasmid (Addgene), with firefly luciferase in the -1 frame. A negative control was made by replacing part of the slippery sequence with a stop codon, and a positive control was made without a frameshift signal and the two luciferases in-frame. RNA transcription templates were amplified from these plasmids by PCR and transcribed in vitro by T7 RNA polymerase.

Plasmids for producing samples for SAXS were prepared by ligating an insert containing the sequences of the SARS-CoV and SARS-CoV-2 pseudoknots into the BamHI and SpeI sites of the pMLuc-1 plasmid as described previously[96]. RNA transcription templates were amplified from these plasmids by PCR, including three extra nucleotides upstream of the pseudoknot and four downstream (all Us). The forward PCR primer was extended on its 5' end to include the T7 polymerase promoter. The transcription

templates were then transcribed in vitro by T7 RNA polymerase. Plasmids used in this study are shown in Table 1.

Cell culture and plasmid transfection

Human embryonic kidney (HEK293T/17) (CRL-11268) and HeLa (CCL-2) cells were purchased from the American Type Culture Collection (Manassas, VA). HEK293T cells were maintained in Dulbecco's modified Eagle's medium (Fisher Scientific 10-013-CV) supplemented with 10% fetal bovine serum (Fisher Scientific 26140-079), 1% GlutaMAX (35050- 061), 1% nonessential amino acids (Fisher Scientific 11140-050), 1% HEPES buffer (Fisher Scientific 15630-030), and 1x penicillin/streptomycin (Fisher Scientific 15140-122) at 37°C in 5% CO_2. HeLa cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1% GlutaMAX, and 1x penicillin/streptomycin at 37°C in 5% CO_2. HEK293T and HeLa cells were seeded at 4 x 10^4 cells/well into 24-well plates. The cells were transfected 24 h after seeding with 500 ng of Dual-Luciferase reporter plasmid using Lipofectamine3000 (Invitrogen L3000015) per the manufacturer's protocol.

Dual-Luciferase assays of -1 PRF

The frameshifting efficiency of the reporter plasmids in cultured cells was assayed as described previously[90,91] using a Dual-Luciferase reporter assay system kit (Promega). 24 h after transfection, the cells were washed with 1x PBS and then lysed with 1x passive lysis buffer (E194A, Promega). Cell lysates were assayed in triplicate in a 96-well plate, and luciferase activity was quantified using a GloMax microplate luminometer (Promega). The percent frameshift was calculated by averaging the three firefly or Renilla luciferase technical replicate reads per sample and then forming a ratio of firefly to Renilla luminescence per sample. Each sample luminescence ratio was compared with a read-through control set to 100%. The ratio of ratios for each sample is the percent frameshift for the sample. A minimum of three biological replicates were assayed for each sample, each of which were assayed in triplicate (technical replicates). Mean technical replicate values of each biological replicate are depicted on graphs with standard error of the mean for biological replicates. Statistical analyses were conducted using Student's t test or one-way analysis of variance as appropriate using Prism 8 software (GraphPad).

To measure -1 PRF efficiency in cell-free assays, 2 µg of mRNA from each construct was heated to 65 °C, mixed with 35µL of nuclease-treated RRL (Promega) and 0.5µL of 1mM amino acid mixture lacking Leu and Met, and then incubated for 90 min at 30 °C. The firefly luminescence from each of the constructs was measured after incubating 20 µL of each reaction with 100 µL of Dual-Glo Luciferase reagent (Promega) for 10 min, and then Renilla luminescence was measured 10 min after adding 100 µL of Dual-Glo Stop and Glo reagent. The -1 PRF efficiency was calculated from the ratio of firefly and Renilla luminescence, subtracting the background measured from the negative control and normalizing by the positive control. Eight independent measurements were made without MTDB, and four were made with MTDB.

SAXS measurements

RNA samples for SAXS experiments were made by in vitro transcription of DNA templates followed by ethanol precipitation of the RNA. To avoid dimerization, RNA was resuspended in a low-salt solution (50 mM MOPS, 10 mM KCl, pH 7.5). The RNA was annealed by heating to 95C° for 5 min and then placed on ice. After concentration with a spin concentrator, a fraction of the RNA was set aside for inline SEC-SAXS, performed just prior to X-ray exposure, whereas the rest was purified by SEC in a column equilibrated with the SAXS buffer (50 mM MOPS, 130 mM KCl, pH 7.5). Selected peak fractions of these lab-purified samples were then concentrated to 17.3 µM for the SARS-CoV RNA and 19.2 µM for the SARS-CoV-2 RNA shown in the figure. All samples were sent to the National Synchrotron Light Source II for data acquisition.

The SAXS data were collected at the Life Sciences X-Ray Scattering Beamline (LIX) at Brookhaven National Laboratory using their standard solution scattering set-up, experimental procedures, and data-processing packages[97]. SEC-SAXS was performed on a Superdex 200 increase 5/150 GL column (GE) equilibrated in the SAXS buffer condition.

Acknowledgements

We thank Kevin Tu for help with reagent preparation and Shirish Chodankar for support with remote data acquisition at National Synchrotron Light Source II. J. D. D. thanks the doctors, nurses, and staff of the COVID-19 ward at Suburban Hospital for wonderful care.

Chapter 3: Programmed -1 Ribosomal Frameshifting in coronaviruses: A therapeutic target[67]

Jamie A. Kelly[a], Michael T. Woodside[b], Jonathan D. Dinman[a,*]

[a] Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, 20742, USA

[b] Department of Physics, University of Alberta, Edmonton, AB, T6G 2E1, Canada

*Corresponding author.

E-mail address: dinman@umd.edu (J.D. Dinman).

Work adapted from Virology, February 2021. DOI: 10.1016/j.virol.2020.12.010

Abstract

Human population growth, climate change, and globalization are accelerating the emergence of novel pathogenic viruses. In the past two decades alone, three such members of the coronavirus family have posed serious threats, spurring intense efforts to understand their biology as a way to identify targetable vulnerabilities. Coronaviruses use a programmed -1 ribosomal frameshift (-1 PRF) mechanism to direct synthesis of their replicase proteins. This is a critical switch in their replication program that can be therapeutically targeted. Here, we discuss how nearly half a century of research into -1 PRF have provided insight into the virological importance of -1 PRF, the molecular mechanisms that drive it, and approaches that can be used to manipulate it towards therapeutic outcomes with particular emphasis on SARS-CoV-2.

Main

Emerging viruses: the confluence of human population growth, globalization, and climate change.

Within the past generation, three interrelated but independent forces have driven the emergence of novel viral pathogens in the human population. Population growth has pushed humans into new ecosystems, increasing the number of contacts with zoonotic viruses, and hence the frequency of their adaptation to the new, human host[98]. Changes to ecosystems driven by climate change are altering the ranges of viral host species, enhancing the number of interspecies contacts[99]. Lastly, economic globalization during this time period has created the means for the rapid, worldwide dissemination of novel pathogens[100] . As a result, we are witnessing an alarming increase in the number and frequency of novel viral epidemics. These include: the introduction of West Nile into North America in 1999[101]; the Severe Acute Respiratory Syndrome (SARS) near-pandemic of 2002–2003, and the narrowly averted breakout of the closely related Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in 2012[63]; a similarly arrested Ebola threat in 2013–16[102]; the emergence of Zika and Chikungunya virus in the Americas in 2015–16[103]. Although a combination of sound policy and luck mitigated the impacts of these prior outbreaks, the ongoing COVID-19 crisis confirms the grim reality that emerging novel virial pathogens will continue to pose one of the defining challenges of the 21st century.

The coronavirus three step genetic program.

In humans, coronaviruses have traditionally been associated with causing approximately 10–30% of "common colds"[104]. The perception of coronaviruses as relatively benign was dramatically altered with the emergence of SARS-CoV, and was further reinforced by MERS-CoV[105]. As discussed elsewhere in this issue, these viruses are thought to be endemic in bats[106], and have moved to humans through intermediate species such as camels and exotic food animals in the context of crowded, unsanitary marketplaces[107].

Similar to other betacoronaviruses, SARS-CoV-2 has a plus-sense RNA genome that is roughly 30 kb in length[62]. The genome contains at least nine different open reading frames (ORFs), where ORF1a and ORF1b comprise of about two-thirds of the genome[62]. The organization of the betacoronavirus genome reveals the viral developmental program (**Fig. 10A**). Upon entering a cell, the viral genomic RNA (gRNA) is released into the cytoplasm, where it functions as an mRNA. Since ORF1a is located at the 5′ end of the gRNA, it is decoded first. ORF1a encodes proteins whose functions are to hijack the host cell by 1) securing the ribosomes, 2) dysregulating the host cellular innate immune response, and 3) cleaving polyproteins into individual proteins. For example, nsp 1 hijacks the ribosomes by binding to the small subunit and occluding the mRNA entrance tunnel[108]. How this interaction favors translation of the viral mRNA remains unanswered, but the 5′ untranslated region of the gRNA stimulates translation *in vitro*, suggesting that it contains a *cis*-acting element to bypass this block[108] , thus enabling the viral RNAs to "own" the ribosomes. Nsp 2 binds to prohibitin 1 and 2 and modulates host survival signaling in apoptosis[109]. In addition to its ability to counteract host innate immunity through its ability to de-ADP-ribosylate, de-ubiquitinate, and remove Interferon stimulated gene 15 (ISG15) modifications from cellular proteins (i.e. de-ISGylate), nsp 3 also has papain-like protease activity which the virus uses to cleave the ORF1a and ORF1b polyproteins[110]. Nsp 4 and nsp 6 function during the viral replication process to help the virus evade innate immune recognition[111]. ORF1b encodes proteins including the nsp 12 RNA-Dependent RNA polymerase (RDRP), the nsp 13 helicase, and the nsp 14/16 capping complex that are involved in the second stage of the viral replication program: RNA synthesis[112]. Specifically, these proteins direct the synthesis of the (-) strand antigenome, which serves as a template for production of new capped, (+) strand gRNAs and the subgenomic RNAs (sgRNAs), whose ORFs are located in the 3′ -most third of the viral genome. The sgRNAs encode mostly structural proteins, which defines the third step of the viral replication program,

synthesis of structural proteins and viral particle assembly. This process can be diagrammed as a software program flowchart (**Fig. 10B**).

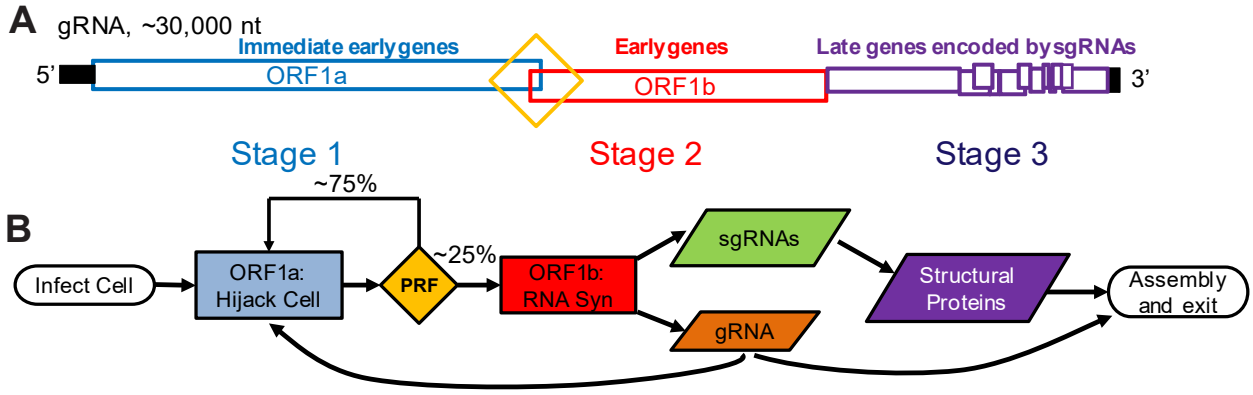


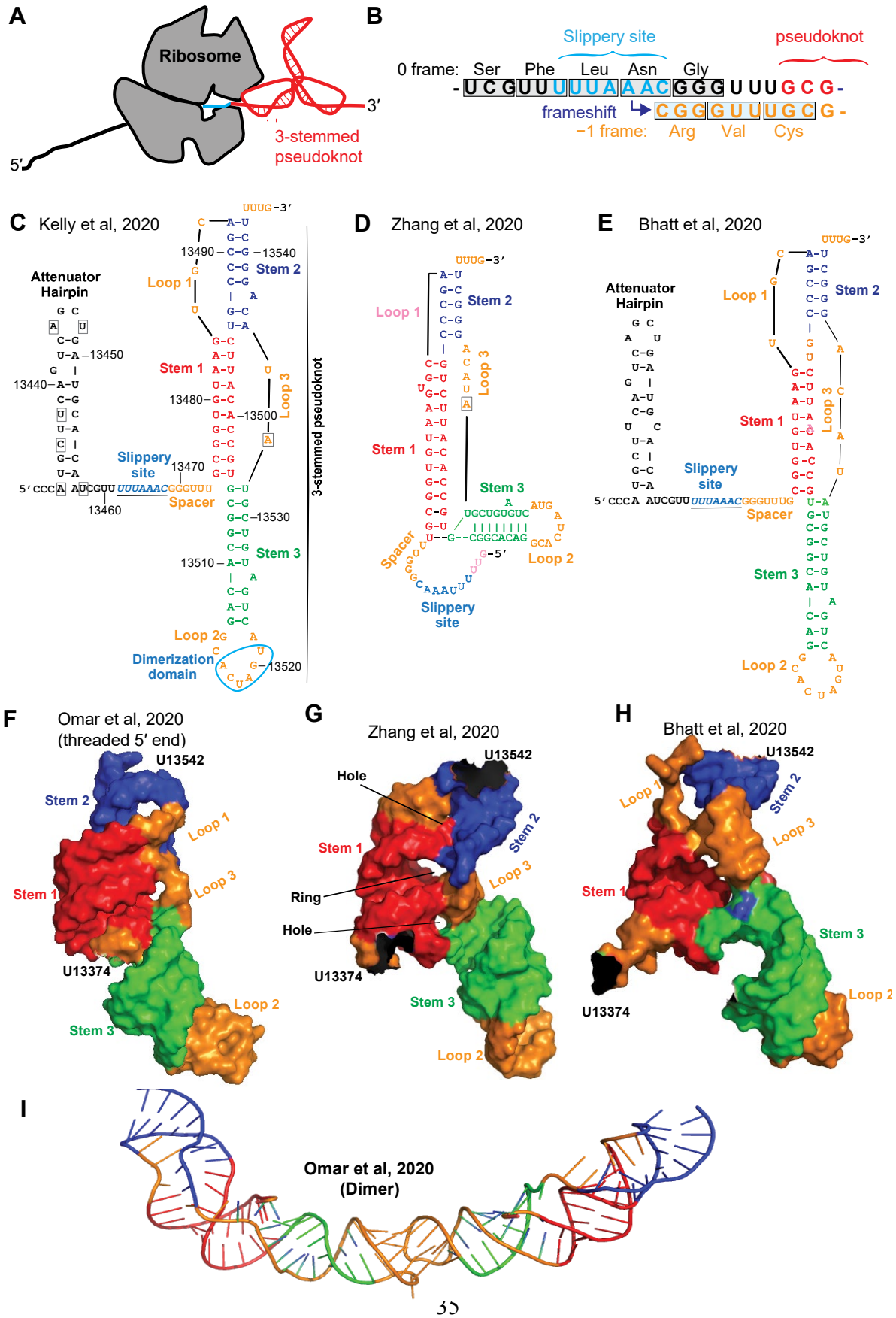
Figure 10. Betacoronavirus gene organization and expression flowchart

A. Map of the betacoronavirus genomic RNA (gRNA). Open reading frames are color-coded and the -1 PRF signal is indicated inside of the yellow diamond. B. Flowchart of the intracellular coronavirus (CoV) replication program. Upon infection and release of viral genomic RNA into the cytoplasm, ORF1a-encoded proteins are synthesized first, initiating Stage 1 of the program. Their function is to "hijack" the cell by securing the ribosomes and disrupting the host cellular innate immune response. Approximately one quarter of translating ribosomes are induced to shift reading frame at the -1 PRF signal. This -1 PRF signal represents a decision point: to continue with Stage 1 or to move into Stage 2, wherein proteins expressed from ORF1b are synthesized in order to transcribe new viral RNAs, including new genomic and subgenomic RNAs. New gRNAs also provide feedback to reinforce cellular takeover by the virus. The transition to Stage 2 may either be rapid, requiring the accumulation of a critical mass of ORF1b products to generate a rapid burst of RNA synthesis (e.g. the rate of viral factory assembly may be determined by -1 PRF rates), or it may be a gradual process instead. In Stage 3, structural proteins encoded in the subgenomic RNAs package the genomic RNAs to produce new viral particles, which exit to repeat the infectious cycle.

Expression of ORF1b requires a programmed -1 Ribosomal frameshift event. In coronavirus genomes, ORF1a and ORF1b partially overlap, where ORF1b is in the -1 reading frame relative to ORF1a (Fig. 1A). Embedded in this overlap region is a *cis*-acting RNA element that directs a fraction of elongating ribosomes to slip by one base in the 5′ (-1) direction in a process called Programmed -1 Ribosomal Frameshifting (-1 PRF). Upon a -1 frameshift, ribosomes are able to continue translating the ORF1b encoded proteins, enabling progression of the viral replication cycle from Stage 1 to Stage 2 as diagrammed in Fig. 1. A typical -1 PRF signal is composed of three elements. From 5′ → 3′, these are a heptameric slippery sequence at which the slippage occurs, a short spacer, and a proximal downstream stimulatory structure in the mRNA that directs the ribosome to pause over the slippery site (**Fig. 11A and B**). The slippery site most often has the sequence N NNW WWH (the incoming 0-frame is indicated by spaces), Where NNN = any three identical bases, WWW = three A's or three U's, and H ≠ G (**Fig. 11B**). With a few exceptions, the stimulatory structure is an RNA pseudoknot, in which the RNA strand folds back on itself one or more times to form a variety of complex but compact and stable structures (e.g., see **Fig. 11C–E**).

Figure 11. -1 PRF in SARS-CoV and SARS-CoV-2.

A-B. Cartoon showing elements involved in -1 PRF and example of shift in reading frame. Elongating ribosomes pause at the 3-stemmed pseudoknot with A- and P-site tRNAs base-paired respectively to AAG and UUA codons in slippery site; upon slippage, non-wobble bases of tRNAs can re-pair to -1 frame codons AAA and UUU. C-E. Comparison of the two-dimensional representations of the SARS-CoV-2 -1 PRF signals. Data from[80,113,114] are labeled and color-coded as indicated. The nucleotides that differ between SARS-CoV-2 and SARS-CoV are boxed in grey. The dimerization domain identified in[93] is circled in cyan. F–H. Space-filled models of the SARS-CoV-2 three stemmed pseudoknot. From left to right, an example of a 5′ -end threaded conformation generated by molecular dynamics simulations[115], the cryo-EM structure of an isolated pseudoknot[114], and the cryo-EM image of the pseudoknot in the context of a paused ribosome[113]. I. A model of the dimerized SARS-CoV-2 -1 PRF signal from molecular dynamics simulations[115].


A
Ribosome
5′
3′
3-stemmed pseudoknot
B
Slippery site
pseudoknot
0 frame: Ser Phe Leu Asn Gly
frameshift
−1 frame: Arg Val Cys
C Kelly et al, 2020
Attenuator Hairpin
Loop 1
Stem 2
Stem 1
Loop 3
Slippery site
Spacer
Stem 3
Loop 2
Dimerization domain
3-stemmed pseudoknot
D Zhang et al, 2020
Loop 1
Stem 2
Loop 3
Stem 1
Stem 3
Spacer
Loop 2
Slippery site
E Bhatt et al, 2020
Attenuator Hairpin
Loop 1
Stem 2
Stem 1
Loop 3
Slippery site
Spacer
Stem 3
Loop 2
F Omar et al, 2020 (threaded 5′ end)
U13542
Stem 2
Loop 1
Stem 1
Loop 3
U13374
Loop 2
Stem 3
G Zhang et al, 2020
U13542
Hole
Stem 2
Stem 1
Loop 3
Ring
Hole
U13374
Stem 3
Loop 2
H Bhatt et al, 2020
U13542
Loop 1
Stem 2
Loop 3
Stem 1
Stem 3
U13374
Loop 2
I
Omar et al, 2020 (Dimer)

The molecular mechanisms underlying -1 PRF have been deeply investigated. The “simultaneous slippage” model of -1 PRF[50] posits that the downstream stimulatory element makes elongating ribosomes pause with their A- and P-site tRNAs over the slippery site in the 0-frame. The nature of the tRNAs and slippery site are such that, upon a -1 slippage event, the tRNA non-wobble bases can re-pair to the -1 frame codons. The pseudoknot was first discovered in a coronavirus, Avian Infectious Bronchitis Virus[116], and this system was subsequently used to demonstrate pseudoknot-induced ribosomal pausing over the -1 PRF signal[117]. While the complete mechanism and structural biology underlying mRNA pseudoknot stimulation of recoding has not been fully elucidated, our current understanding is described as follows. The “torsional restraint” model[51] proposes that, as an elongating ribosome begins to unwind Stem 1 of the pseudoknot, supercoiling in Stem 2 impedes the ribosomes’ progress such that a point is reached where the forward motion of the ribosome is countered by the resistance of the pseudoknot to unwinding. This effect, in combination with a spacer of optimal length, serves to direct ribosomes to pause with their A- and P-sites at the slippery site. The “9 Å Solution” model of -1 PRF[54] was founded on atomic-resolution structural data indicating that the mRNA is pulled into the ribosome by one base during the process of aa-tRNA accommodation[53] . In this model, the downstream stimulatory structure impedes this movement, stretching the segment of mRNA located between the slippery site and the stimulatory structure. The resulting local tension in the mRNA can be resolved either by unwinding the stimulatory structure or by slippage into the -1 frame. A similar mechanism can also be applied to co-translocational -1 PRF events[10,118–120] . An important feature of this model is that the energy provided by the GTPase power stroke of either eEF1A or eEF2 is sufficient to drive tRNA unpairing from the 0-frame codons, a critical prerequisite for -1 PRF[10,54,118,121]. Additional structural and kinetic analyses using purified *E. coli* ribosomes and elongation factors have shown that the downstream pseudoknot in the mRNA can impede the closing movement of the large subunit head, arresting it in a hyper-rotated state, which delays dissociation of the translocase and the release of deacylated tRNA. Release of the tension on the mRNA by ribosomal slippage accelerates completion of translocation, providing a lower-energy path for the ribosome to continue translation[34,119,122]. More recently, an endogenous cellular protein called Shiftless was identified that binds

to frameshifted, hyper-rotated ribosomes, arresting their translation and promoting translation termination by recruiting the termination factors[69].

Whether downstream stimulatory structures play active roles in directing -1 PRF remains an open question. It has been known for some time that thermodynamic stability corresponds with -1 PRF efficiency, but only to a limited extent: pseudoknots that are too stable inhibit -1 PRF, presumably because they cannot be resolved by translating ribosomes[123]. Numerous studies suggest that dynamic mRNA structural remodeling is required for optimal -1 PRF efficiency[57,61,96,120,124–127]. Coordination of base triples in both major and minor grooves provides mechanical resistance to pseudoknot unwinding and stretches of adenosines confined along the minor groove of a helix also provide resistance. Together, these molecular features contribute to ribosome pausing at the slippery site to help stimulate -1 PRF[95,128]. Thus, while it was initially thought that downstream stimulatory structures were merely passive "roadblocks," the most recent research suggests that they are actively involved. What 'active' means in this context remains an evolving question. For example, it may involve the deformation of one structure that strongly impedes ribosome progress, followed by structural remodeling to another conformer that does not, or it may be that fluctuations in the mRNA tension induced by conformational switching play a role in inducing slippage. Additionally, biophysical and mutational analyses revealed that the terminal loop (i.e., loop 2) of the SARS-CoV -1 PRF stimulating pseudoknot mediates RNA dimerization and that this plays a role in determining -1 PRF efficiency[93]. Notably, this sequence is conserved in SARS-CoV-2 (**Fig. 11C and I**)[80]. The role of dimerization in -1 PRF remains unknown.

1. Structural biology: 3-stemmed RNA pseudoknots in coronavirus -1PRF signals

There is strong phylogenetic, genetic, and biophysical evidence showing that the alpha- and betacoronaviruses share a common 3-stem fold, rather than the 2-stem structure typical of most stimulatory pseudoknots[45,51,88,114,115,129] (**Fig. 11C, D, 11E**). This 3-stem architecture is highly conserved among betacoronaviruses[130] but seems to be unique to the coronavirus family[131]. The SARS-CoV-2 pseudoknot is nearly identical in sequence to the SARS-CoV pseudoknot, differing only in having A instead of C at position 13,533 in loop 3. The results from structural probes of the SARS-CoV pseudoknot[55,86,93] are thus likely to apply also to the SARS-CoV-2 pseudoknot. Indeed, small-angle X-ray scattering analyses of global

morphology confirms that the SARS-CoV and SARS-CoV-2 pseudoknots occupy effectively identical space-filling envelopes[80], and NMR spectra[132] revealed a similar secondary structure to that deduced previously for the SARS-CoV pseudoknot from nuclease-protection assays[86], although with the end of stem 2 unpaired to extend loop 3. Based on the secondary structure from nuclease-protection assays, atomistic molecular dynamics simulations found an ensemble of possible structures with networks of tertiary contacts consistent with the resistance of the SARS-CoV pseudoknot to mechanical unfolding[115]. Intriguingly, this ensemble encompasses different fold topologies, including one with the 5′ end threaded through the junction between stems 1 and 3 (**Fig. 11F**). The 5′ -end threading creates an unusual "ring-knot" structure previously only seen in viral exoribonuclease-resistant RNAs[133].

Cryo-EM imaging of the SARS-CoV-2 pseudoknot confirms that it forms an ensemble of structures, with both rod-like and bent conformers observed, including some with a ring-like feature[114]. This 88-nt RNA is also notable for being the smallest biomolecule resolved by single-particle cryo-EM to date. A preliminary 5.9 Å resolution structure of the ring-like subpopulation reveals the presence of two "holes" in addition to the open ring (**Fig. 11G**), which may present small-molecule binding/docking sites. The ring provides space for threading of the 5′ end, similar to what was seen in molecular dynamics modeling[115] (**Fig. 11F**), confirming the presence of this unusual topology. However, the base-pairing is different in this structure from the pairing deduced from nuclease-protection results[86] and NMR[132], featuring an extended stem 1, loss of loop 1, shortened stem 2, and lengthened loop 3. As a result, none of the stems are stacked, unlike what is often seen in stimulatory pseudoknots. These differences may arise from the high Mg^{2+} concentration used for the imaging. This cryo-EM reconstruction also includes a mini-helix formed within the slippery site and spacer region upstream of the pseudoknot. However, RNA chemical modification studies of the SARS-CoV-2 genome reveal that the slippery site is single-stranded[114,134–137], suggesting that the mini-helix is an artifact of the construct used for cryo-EM imaging.

A second cryo-EM study has imaged the SARS-CoV-2 -1 PRF signal not in isolation, but on an arrested mammalian ribosome that is primed for frameshifting[113] (**Fig. 11H**). In this ~2.3 – 7 Å reconstruction, the RNA is positioned with its slippery-site codons in the 0-frame, a peptidyl-phenylalanyl tRNA base-paired to the ribosomal P-site, and the spacer region pulled into the mRNA entrance tunnel. It should be noted that in order to stall ribosomes at the -1 PRF signal, the 0-frame A-site codon was changed from AAC to UAA

and *in vitro* translation reactions were supplemented with an excess of a mutant eRF1 (AAQ) in order to trap ribosomes in the act of decoding the A-site. These non-rotated ribosomes thus represent a pre-frameshift complex, which necessarily limits what can be learned about the -1 PRF process itself. Nonetheless, this structure presents a wealth of novel information. Consistent with the unwinding of the pseudoknot by the intrinsic ribosomal helicase[138,139], the spacer and stem 1 of the pseudoknot interact with basic residues in the C-terminal domain of ribosomal proteins uS5 and eS30. A direct interaction between helix 16 of the 18S rRNA and minor groove of stem 1 was also noted: this interaction may restrict the relative rotation between the head and body of the small subunit during translocation, which has been shown to be important for the -1 PRF process[118]. Similar to the structures described above, the 5′ end is threaded through a ring formed by the junction between the 3 stems. Whereas the quasi-coaxial stacking of stems 1 and 2 resembles what is seen in the threaded structures described by Omar et al[115]. the interaction of the -1 PRF signal with the ribosome appears to have caused significant restructuring of this element (compare **Fig. 11F and G** with **Fig. 11H**). In particular, stem 1 is distorted towards the 5′ end and shortened by 1 base-pair, loop 1 is less compact and breaks triples with stem 2 in favor of interactions with the ribosome, loop 3 is extended by shortening stem 2 but also loses tertiary contacts with stem 1, while stem 3 is lengthened by 1 base-pair and most notably is rotated nearly perpendicular to the axis formed by stems 1 and 2. These results suggest that the interaction of the pseudoknot with the ribosome results in significant restructuring of the frameshift-stimulating element, consistent with SAXS analyses[80] and molecular dynamics simulations indicating the presence of a complex structural ensemble of conformers[115]. These findings support an emerging theme wherein “shapeshifting” RNAs are important for regulating gene expression[140,141]. Beyond well-documented examples in mRNA splicing, others include the ability of different conformers of the nc886 RNA to control activation of RNase L and its ability to activate the immune response[142], and the interactions of mRNAs with Argonaute[143]. Currently, small- and wide-angle x-ray scattering (SAXS and WAXS)[80], single-molecule force spectroscopy[95], time-resolved cryo-EM[144], new biophysical assays of -1 PRF including ribosome profiling[47] and nanopore-based applications[145], and computational advances are being exploited to visualize and model the process of -1 PRF.

2. Functional analyses of the SARS-CoV and SARS-CoV-2 -1PRF signals

Historically, a series of molecular genetics and biochemical analyses of the Avian Infectious Bronchitis Virus -1 PRF signal established the foundation for much of our understanding of this phenomenon[146,147]. Analyses of SARS-CoV-2 sequence variations reveal the highly conserved nature of the -1 PRF signal; the vast majority of variants are very infrequently represented in the population, supporting the importance of this element for viral fitness[148]. Functional studies of single-nucleotide polymorphisms seen in different regions of the pseudoknot found that most of them had little effect on -1 PRF efficiency[58], with only a few leading to significant decreases, including a ~2-fold decrease from C13476U and C13501U in stem 1[65] and a roughly 3-fold decrease from U13494C in stem 2[58]; notably, each of these mutations involved converting G:C pairs to G:U (or vice versa), and hence would be expected to leave the secondary structure unchanged. Stems 1 and 2 of the SARS-CoV and SARS-CoV-2 pseudoknots are absolutely required to promote -1 PRF, but stem 3 is not; rather it appears to function to further stimulate this activity[55,80,86,88]. In loop 1, changing G13486 to A reduces -1 PRF to roughly one-third of wild-type levels, while changing it to C reduced it even further[113]; similarly, the U13485C mutation reduces -1 PRF more than two-fold[65]. Mutations to A13535 and A13537, located in loop 3 and/or stem 2 (depending on the structural model), also abrogated efficient -1 PRF[86,113]. These observations also support the idea that structural plasticity plays an important role in the -1 PRF mechanism. Additionally, the placement of the 0-frame stop codon appears to play an important role in determining -1 PRF efficiency, and a model has been proposed in which the process of termination by a leading ribosome provides the pseudoknot time to re-fold before a trailing ribosome encounters the -1 PRF stimulating sequence[113].

The SARS-CoV and SARS-CoV-2 -1 PRF signals also harbor a novel “attenuator hairpin” element located immediately upstream of the slippery site[89,149] (**Fig. 11C and E**). The attenuation model posits that the hairpin is initially unwound by an elongating ribosome as it approaches the frameshift signal. As it enters the slippery site, the ribosome clears the attenuator sequence, enabling the stem-loop to re-form. Its formation enables it to block the backwards slippage of the ribosome. While the primary sequence of the SARS-CoV-2 attenuator element is not as well conserved with its SARS-CoV counterpart as compared to their core -1 PRF signals, both have been shown to have -1 PRF-tempering activities[80]. Additionally, *in*

silico analysis of the SARS-CoV-2 "structurome" suggests that the -1 PRF signal is nested inside of a larger, double-stranded RNA superstructural domain[150].

3. -1PRF as a critical developmental switch

As noted above, expression of the ORF1b proteins require a -1 PRF event. From the programmatic point of view shown in Fig. 4B, -1 PRF represents a decision nexus: either remain in Stage 1 of the infectious program or progress to Stage 2. Notably, -1 PRF does not happen with 100% efficiency; rather, -1 PRF directed by the SARS-CoV and SARS-CoV-2 elements occur at an efficiency of ~15–30%, depending on the assay system[80,86,113]. In viruses such as HIV-1, -1 PRF rates determine the ratio of structural (e. g. the Gag polyprotein) to enzymatic proteins (the Gag-pol polyprotein), and the prevailing model is that the rate of -1 PRF ensures the production of the correct ratios of structural to non-structural proteins[151]. However, this situation does not apply to coronaviruses because the ORF1a proteins do not encode structural proteins. Instead, we suggest that -1 PRF in these viruses may have a timing function. We propose that by delaying the accumulation of the RNA replication machinery until some critical concentration is reached—which could be important for building a viral factory[152], for example, a process that may involve a concentration-dependent phase transition of the viral replication complex[153,154]—the virus may buy time for the ORF1a-encoded non-structural proteins to amass to high enough concentrations that they can incapacitate the host cell's innate immune response. This time delay may be important because of the transient production of dsRNAs during the RNA replicative phase, which may activate various arms of the innate immune response[155]. From a biochemical/biophysical vantage point, slowing the buildup of viral replicase may maximize the timing at which a critical concentration of this enzyme is achieved, enabling a burst of RNA synthesis at the right time during the viral replication cycle.

4. -1PRF is a novel target for antiviral therapeutics

An early study of -1 PRF in a totivirus demonstrated that the native rate of frameshifting produced the correct stoichiometric ratios of structural to enzymatic viral proteins, and that either increasing or decreasing -1 PRF efficiencies inhibited viral replication[56]. Consistent with this model, overexpression of retroviral Gag-pol protein inhibited viral replication[156]. Later studies revealed that -1 PRF rates can also be altered by small molecules to interfere with viral replication, thereby identifying -1 PRF as a potential

therapeutic target[157,158]. These findings were later extended to the SARS-CoV -1 PRF signal, showing that mutants[55,85], antisense peptide nucleic acids[94], and a small-molecule inhibitor of -1 PRF, 2-methylthiazol-4-ylmethyl)-[1,4]diazepane-1-carbonyl]amino}benzoic acid ethyl ester (MTDB)[60], all negatively impacted virus replication. The -1 PRF signals of the SARS-CoV family may be particularly good drug targets because a) there is no known case of -1 PRF promoted by a three-stemmed pseudoknot structure in host cellular mRNAs; b) the -1 PRF signal is highly conserved because it has to maintain structure while coding for two overlapping genes, and thus it is not likely to mutate to evade drug interactions; and c) the structure of the pseudoknot is sufficiently complex to contain well-defined binding pockets, with the 5′ -end threading in particular generating a unique pocket geometry. This notion has elicited a burst of recent research aimed at identifying small molecules that target the SARS-CoV-2 -1 PRF signal[137]. For example, MDTB was also shown to inhibit SARS-CoV-2 -1 PRF[80] and viral replication[113]. Similarly, this agent appears to be resistant to natural variants of the SARS-CoV-2 -1 PRF stimulating pseudoknot[58]. A recent screen of a bank of approved drugs identified numerous small molecules that either stimulated or inhibited SARS-CoV-2 mediated -1 PRF[159]. Independently, another group identified merafloxacin, a fluoroquinolone antibacterial, as a potent inhibitor of SARS-CoV-2 -1 PRF and viral replication in Vero-E6 cells, which also showed resistance to natural mutations and activity against other human betacoronaviruses[65]. We have also identified numerous small-molecule inhibitors (Dinman, unpublished). Although there does not appear to be overlap among all of the screens reported to date, the compounds identified thus far are rich in hydrophobic cyclic structures (**Fig. 12A**), suggesting that they may bind to the “ring” and “holes” identified by molecular dynamics simulations[115] and cryo-EM[114]. Indeed, computational modeling of the binding of MTDB to the SARS-CoV[60] and SARS-CoV-2 (Woodside, unpublished) pseudoknots shows that it binds to a cleft formed by the threading of the 5′ end (**Fig. 12B**). Intriguingly, interactions with the pseudoknot alone are insufficient to explain the inhibitory effect of MTDB, since its *K*d of ~200 μM for pseudoknot binding[61] is many times higher than IC50 for suppressing -1 PRF[60,80] or viral replication[113], suggesting that its binding is enhanced by the presence of ribosomes, for example owing to direct contacts with the ribosome or effects from ribosome-induced remodeling of the pseudoknot. It remains unclear to what extent similar considerations may apply to other small-molecule inhibitors. In parallel to exploration of small-molecule inhibitors, antisense targeting of the -1 PRF signal is also being explored as a therapeutic approach[85,114].

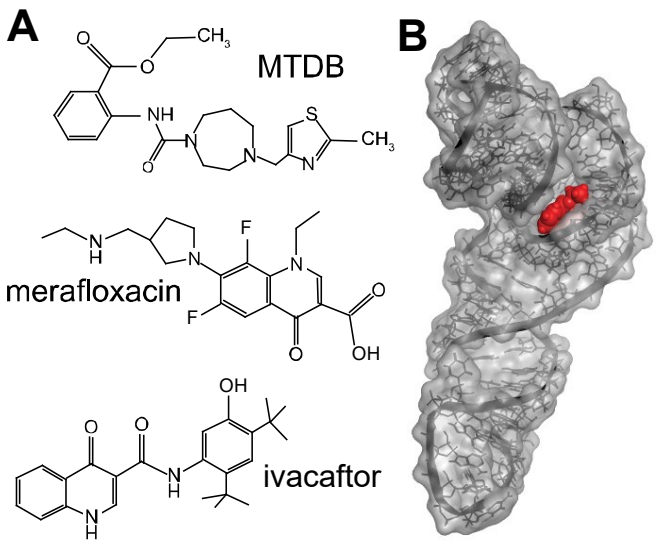


Figure 12. Small-molecule inhibitors of -1 PRF in SARS-CoV-2

A. Three examples of small molecules that have been found to inhibit -1 PRF: MTDB[80], merafloxacin[65], and ivacaftor[159]. B. Binding site of MTDB on SARS-CoV-2 pseudoknot. Model of the binding site of MTDB (red) found from docking calculations and molecular dynamics simulations of the bound complex shows the ligand binds in a cleft formed because of 5′ -end threading in the pseudoknot.

The -1 PRF attenuator hairpin also presents a target for antiviral intervention. For example, annealing of an antisense RNA or DNA oligonucleotide to upstream of the MERS-CoV -1 PRF signal strongly inhibited frameshifting[160]. Similarly, a drug-like small molecule has been identified that binds with high affinity to the SARS-CoV-2 frameshift-attenuator hairpin, stabilizing it in its folded state and attenuating -1 PRF in a cell-based assay[64]. Additionally, when ligated to RIBOTAC, a ribonuclease targeting chimera, it can recruit a cellular protease to degrade the viral RNA.

An alternative approach to small-molecule or anti-sense inhibitors may be to develop attenuated viral vaccine strains that incorporate mutated -1 PRF signals. These RNA elements may be particularly amenable to such an approach because multiple silent coding mutations can be incorporated into the slippery-site and pseudoknot-forming regions, thus decreasing the chances of mutational reversion. For example, mutations of the slippery site of Venezuelan Equine Encephalitis Virus that promoted decreased rates of -1 PRF only mildly delayed the kinetics of VEEV accumulation in cultured cells, but strongly inhibited its pathogenesis in an aerosol infection mouse model, including decreasing viral titers in the brain[57]. Preliminary data indicate that mice infected with this mutant are protected from subsequent challenge with a highly pathogenic version of the virus (Dinman and Kehn-Hall, unpublished). These findings suggest a novel approach to the development of safe and effective live attenuated vaccines directed against -1 PRF-utilizing viruses, including members of the SARS-like coronaviruses. As a final thought, it may be possible to exploit the -1 PRF inhibitor Shiftless as a means to control viral infection[70].

Acknowledgements

We wish to thank members of the Dinman lab, both past and present for their efforts investigating the SARS-CoV and SARS-CoV-2 -1 PRF signals, with special thanks to Dr. Ewan Plant. We also thank Drs. Lois Pollack, Kylene Kehn-Hall, and Nenad Ban for their strong collaborative spirit and sharing of preliminary information. JDD extends his heartfelt thanks to the doctors, nurses and staff of the Covid-19 ward at Suburban Hospital in Bethesda MD for their wonderful care.

Chapter 4: Shiftless is a regulator of translational fidelity

Abstract

The innate immune system targets viral frameshifting using an interferon-stimulated -1 PRF inhibitor called Shiftless (SFL) that binds, arrests, and terminates translation of -1 frameshifted ribosomes. We found that SFL is not only expressed in response to viral infection, but that it may have a role in monitoring general translational fidelity. SFL is constitutively expressed at low levels in human-derived cell lines and its effects are not limited to -1 PRF signals. Disruption of SFL homeostasis results in reciprocal 2-fold changes to recoding efficiencies in a panel of human and viral-derived translational recoding signals, global decreases in reporter protein synthesis and global decreases in reporter and cellular mRNA steady state abundances. Additionally, SFL over or under expression combined with knockdown of prominent ribosome-associated protein quality control (RQC) proteins reveals that SFL is epistatic to RQC. These results suggest that SFL has a role in general translational fidelity monitoring for spontaneously frameshifted ribosomes in addition to its role as a member of the innate immune response.

Introduction

Programmed -1 ribosomal frameshifting (-1 PRF) is a translational recoding mechanism that directs ribosomes to slip backwards one base in the 5' direction and change the reading frame. It is widely used by RNA viruses to regulate the synthesis of viral RNA-dependent RNA polymerase and is used in some eukaryotic genes to regulate gene expression[37,44,47,48,75,87,88]. In viruses, -1 PRF is highly regulated, and disruption of frameshift efficiency can have devastating consequences for the virus[55,64,80,150,156]. Shiftless is an interferon inducible -1 PRF inhibitor[69]. Originally identified as an antiviral protein involved with targeting viral replication via interactions with viral replication factories, its role was later refined to targeting -1 PRF[69,72,161]. Shiftless binds, arrests and liberates frameshifted ribosomes and can target both human and viral derived -1 PRF signals[69]. Recent studies have reported that Shiftless binding is not specific to -1 PRF and can bind with similar affinity to other translational recoding signals[162].

With the exception of programmed frameshifting and other translational recoding mechanisms, translational fidelity is highly regulated in the cell. Ribosome-associated protein quality control (RQC) mechanisms identify defective mRNAs and resolve elongation distress that inhibits productive

translation[12,17,21,163]. RQC can be induced by secondary RNA structures impeding ribosome progression, premature termination codons, or lack of a normal termination codon causing the ribosome to continue into the poly-A tail[16,21,27].

Here, we report on the role of Shiftless in translational fidelity and establish it as a member of ribosome-associated protein quality control (RQC). Shiftless is constitutively expressed at low levels in human-derived cell lines. Contrary to previous studies, it is not -1 PRF specific and can also influence efficiency of other translational recoding signals as well as "normal" or non-recoding messages. Disruption of Shiftless homeostasis significantly decreases mRNA steady-state abundances. Genetic analyses suggest that Shiftless is epistatic to RQC and that it may assist with triggering mRNA surveillance mechanisms by binding and arresting frameshifted ribosomes. Lastly, Shiftless expression is reduced in multiple cancers, further supporting its importance in translational fidelity.

Results

Shiftless is constitutively expressed in human-derived cells

Previous studies demonstrated that Shiftless is constitutively expressed at varying levels in human-derived cell lines and can be upregulated by the type I interferon response[69,72]. To confirm this, we used RT-PCR to probe for SFL expression in HeLa, HEK293T, and U87 MG cells using *E. coli* as a negative control. Our results show that both SFL and the shorter splice isoform, SFLS[69], are expressed in the human-derived cell lines but not in bacteria (**Fig. S1A**).

Shiftless overexpression or knockdown alters -1 PRF

A bifluorescent reporter assay was used to monitor -1 frameshift efficiency in HEK293T cells that over- or under-express SFL. Disruption of SFL homeostasis resulted in reciprocal 2-fold changes to -1 PRF efficiencies in a panel of human and viral-derived -1 PRF signals **(Fig. 13**). Consistent with previous studies, SFL overexpression resulted in ~2-fold decreases in -1 PRF efficiency promoted by the -1 PRF signals derived from SARS-CoV, HIV-1 and CCR5 (**Fig. 13**). Specifically, SARS-CoV -1 PRF decreased from ~35% $\pm$ 5.5% in WT HEKs to ~18.9% $\pm$ 2.2% in SFL overexpressed cells (SFL OE) (p = 0.0057) (**Fig. 13A**). HIV-1 frameshift efficiency decreased from 4.5% $\pm$ 0.45% in WT to 2.89% $\pm$ 0.36% in SFL OE (p = 0.157)

(Fig. 13B). Frameshifting promoted by the CCR5 -1 PRF signal decreased from 1.76% ± 0.13% in WT to 0.95% in SFL OE (p = 0.0298) (**Fig. 13C)**. While significant decreases in -1 PRF were seen using the bifluorescent reporter system, SFL OE did not yield significant decreases in -1 PRF using dual luciferase reporters (**Fig. S2).** In each individual experiment, we did observe that SFL overexpression did decrease -1 PRF as monitored by the dual-luciferase reporters, but that the variability in % -1 PRF between experimental replicates masked this effect. Plotting the data as fold-change compared to wild-type revealed decreases in -1 PRF efficiencies between 10 to 30% **(Fig. S3**).

Conversely, knockdown of expression using an shRNA targeting SFL resulted in a roughly two-fold increase in apparent frameshifting. SARS-CoV promoted -1 PRF increased from ~35% in WT HEKs to 51.2% ± 6.08% in SFL shRNA knockdown (shSFL) (p = 0.0081). HIV-1 stimulated -1 PRF increased from 4.5% in WT to 8.24% ± 1.65% in shSFL (p = 0.003), and CCR5 frameshifting increased from 1.76% in WT to 2.96% ± 0.51% (p = 0.0036). Similar results were observed using dual luciferase reporters (**Fig. S2 and S3**).

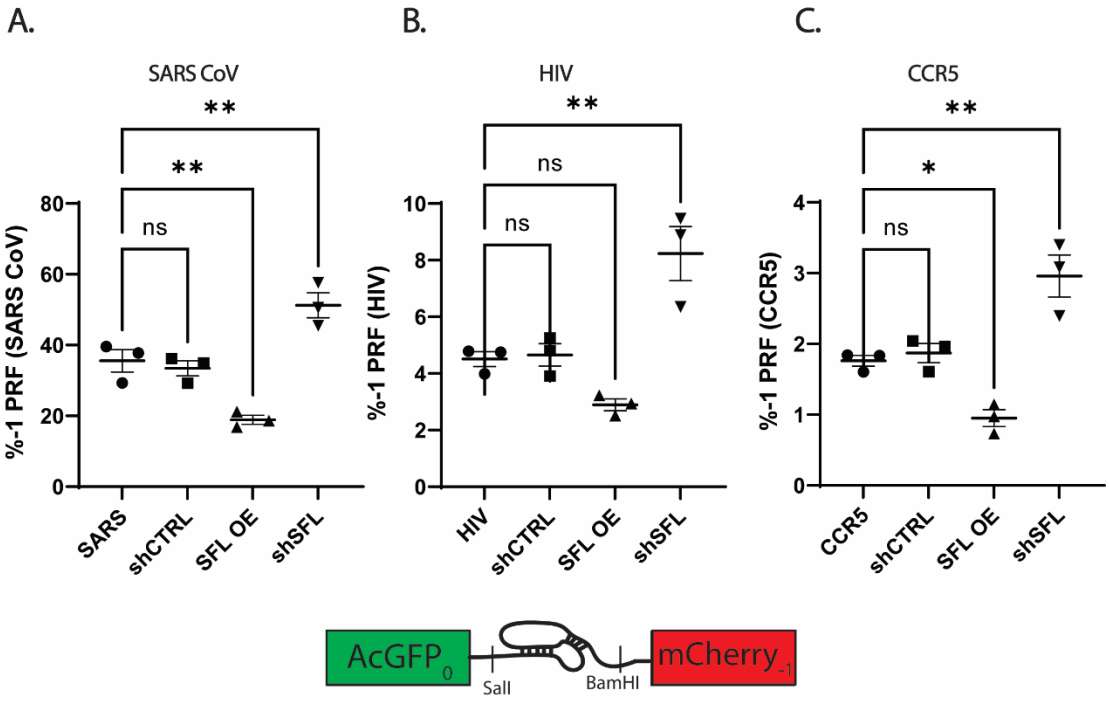


Figure 13. Shiftless overexpression or knockdown alters -1 PRF.

Frameshift efficiency of three translational recoding elements measured using dual fluorescence reporters in HEK293T cells over or under-expressing SFL. (A) SARS-CoV -1 PRF signal, (B) HIV-1 -1 PRF signal, (C) CCR5 -1 PRF signal. Each data point represents a single biological replicate. Error bars denote S.E. ns, not significant.

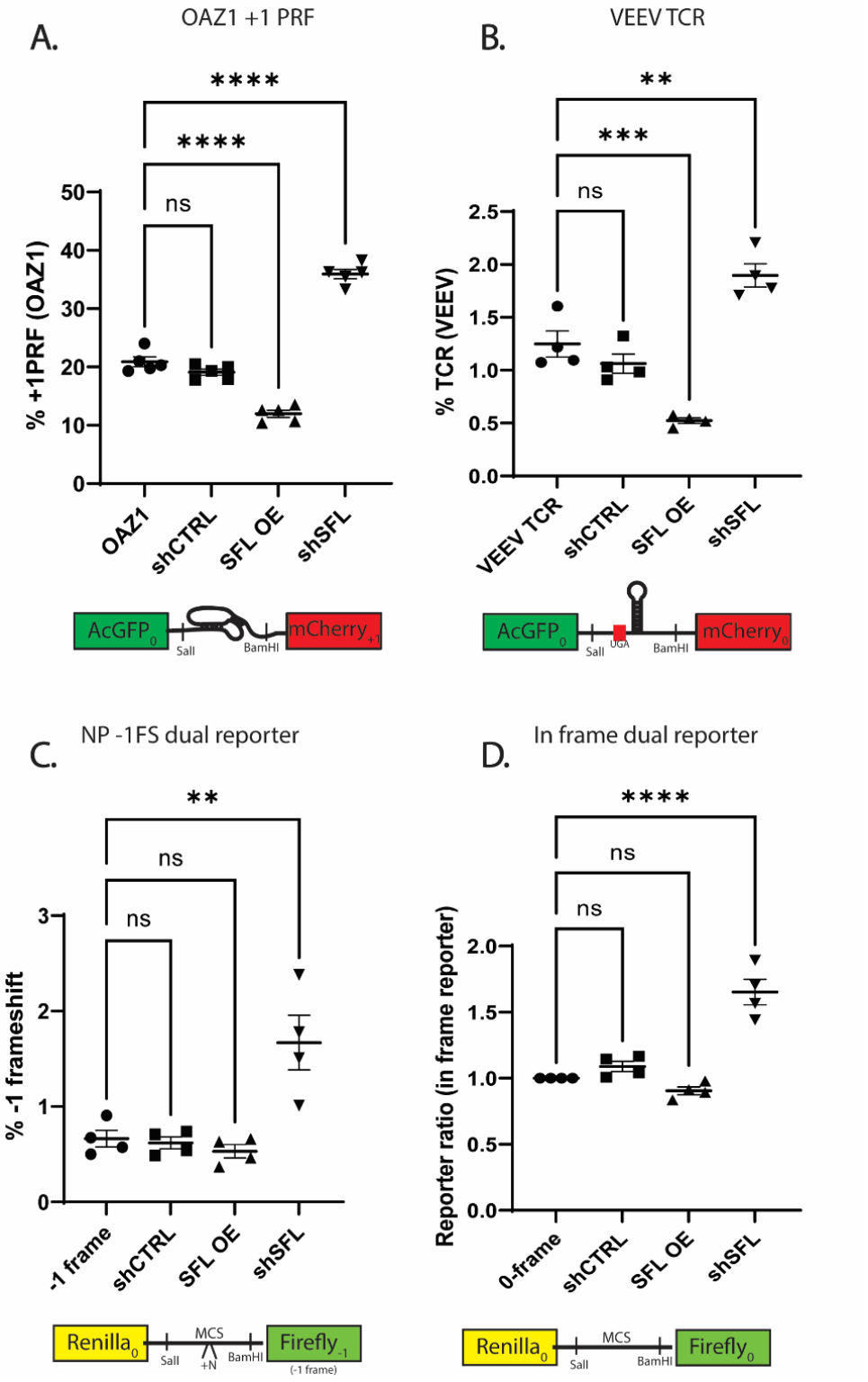
Shiftless is not limited to -1 PRF signals

Although Shiftless was originally identified as a -1 PRF specific frameshift inhibitor[8], a recent study demonstrated SFL can also modulate programmed termination codon readthrough and that it has no specificity for certain mRNA structures, e.g. pseudoknots over others, e.g. stem-stem loops[162]. To further probe the range of possible SFL substrates, SFL was over or under expressed in the context of cells expressing either a +1 PRF signal from ornithine decarboxylase antizyme 1 (OAZ1)[42,76] or a termination codon readthrough (TCR) signal from Venezuelan equine encephalitis virus (VEEV)[35,164]. Similar to the findings with -1 PRF signals, reciprocal 2-fold changes were observed with these recoding signals. Specifically, SFL overexpression in OAZ1 decreased +1 PRF from 20.9% $\pm$ 1.85% to 12% $\pm$ 1.35% (p < 0.0001) and shSFL increased apparent frameshifting to 35.9% $\pm$ 1.78% (p < 0.0001), respectively (**Fig. 14A**). SFL OE in VEEV decreased TCR from 1.25% $\pm$ 0.24% to 0.52% $\pm$ 0.05% (p = 0.0005) and shSFL increased TCR to 1.89% $\pm$ 0.22% (p = 0.0011) (**Fig. 14B**).

To determine whether these effects are limited to *bona fide* recoding signals, the effects of SFL over- or under-expression were assayed using a bicistronic reporter in which the downstream (firefly luciferase) reporter is in the -1 frame with regard to the upstream reporter (non-programmed -1 frameshift, i.e. NP-1FS). SFL OE using NP-1FS did not yield a significant change in non-programmed frameshifting, however SFL knockdown resulted in a 2.5-fold increase, from 0.66% $\pm$ 0.18% to 1.67% $\pm$ 0.57% (p = 0.0018) (**Fig. 14C**). Following this observation, we examined the effects of SFL over- and under-expression on the 0-frame reporter, i.e. where both reporters are in-frame with one another. Surprisingly, SFL overexpression decreased the 3' to 5' reporter ratio from 1.0 to 0.905, while knocking down SFL expression resulted in a 1.65-fold increase in reporter ratio (p <0.0001) **(Fig. 14D**).

Figure 14. Shiftless is not limited to -1 PRF signals.

Translational recoding efficiency and reporter ratio measured using dual fluorescence reporters in HEK293T cells over or under-expressing SFL. (A) OAZ1 +1 ribosomal frameshift signal, (B) VEEV termination codon readthrough signal, (C) NP -1FS dual luciferase reporter, (D) 0-frame dual luciferase reporter. Each dual luciferase data point represents a single biological replicate comprised of three technical replicates. Each dual fluorescent data point represents a single biological replicate. Error bars denote S.E. n.s, not significant.


A.
OAZ1 +1 PRF
****
****
ns
% +1PRF (OAZ1)
50
40
30
20
10
0
OAZ1
shCTRL
SFL OE
shSFL
AcGFP0
SalI
BamHI
mCherry+1
B.
VEEV TCR
**
***
ns
% TCR (VEEV)
2.5
2.0
1.5
1.0
0.5
0.0
VEEV TCR
shCTRL
SFL OE
shSFL
AcGFP0
SalI
UGA
BamHI
mCherry0
C.
NP -1FS dual reporter
**
ns
ns
% -1 frameshift
3
2
1
0
-1 frame
shCTRL
SFL OE
shSFL
Renilla0
SalI
MCS
+N
BamHI
Firefly-1
(-1 frame)
D.
In frame dual reporter
****
ns
ns
Reporter ratio (in frame reporter)
2.0
1.5
1.0
0.5
0.0
0-frame
shCTRL
SFL OE
shSFL
Renilla0
SalI
MCS
BamHI
Firefly0

Disruption of Shiftless homeostasis inhibits reporter activity

The experiments described above examined ratios of ratios of two reporter proteins expressed in cells either over- or under-expressing SFL as compared to mock transfected cells. Since calculation of recoding efficiency can reduce information content by averaging ratios of ratios, an issue that particularly impacts dual luciferase-based assays due to their greater variability, measuring actual reporter gene activity presents a more accurate view of the effects of SFL on gene expression. Because SFL over- or under-expression influenced these ratios independent of the type, or even the presence of a recoding element, we examined actual reporter gene activities to determine the extent to which SFL directly influences reporter output. In WT cells, expression of both GFP and mCherry from the bicistronic dual-fluorescence reporter is ~14000 fluorescence units per 150uL cell lysate. Over-expression of SFL promoted a ~2-fold decrease in both GFP and mCherry activities using this reporter (**Fig. 15A**). Knockout of SFL resulted in a 1.6-fold decrease in GFP expression with no significant change in mCherry expression. Similar inhibitory trends of SFL over- or under-expression of gene expression were observed with 0-frame dual-luciferase reporters (**Fig. 15B**) and the non-programmed -1 frameshift (NP -1FS) dual-luciferase reporter **Fig. 15C**). Over-expression of SFL promoted >10-fold decreases in both Renilla and firefly luciferase activities using the 0-frame reporter. $SFL^{-/-}$ reduced reporter activities to a lesser extent, i.e. a ~6.8-fold in Renilla and ~5-fold in firefly (**Fig. 15B**). We note that the baseline firefly activity is almost three orders of magnitude less than Renilla activity in the NP-1FS reporter construct, reflecting the baseline rate of non-programmed -1 frameshifting (**Fig. 15C**). These observations suggested that SFL may normally monitor in-frame messages for spontaneous frameshifting activity. Consistent with this, the activity of a monocistronic GFP reporter was similarly impacted by SFL overexpression (> 10-fold decrease) or in an isogenic $SFL^{-/-}$ HEK cell line (~30% decrease) (**Fig. 15D**).

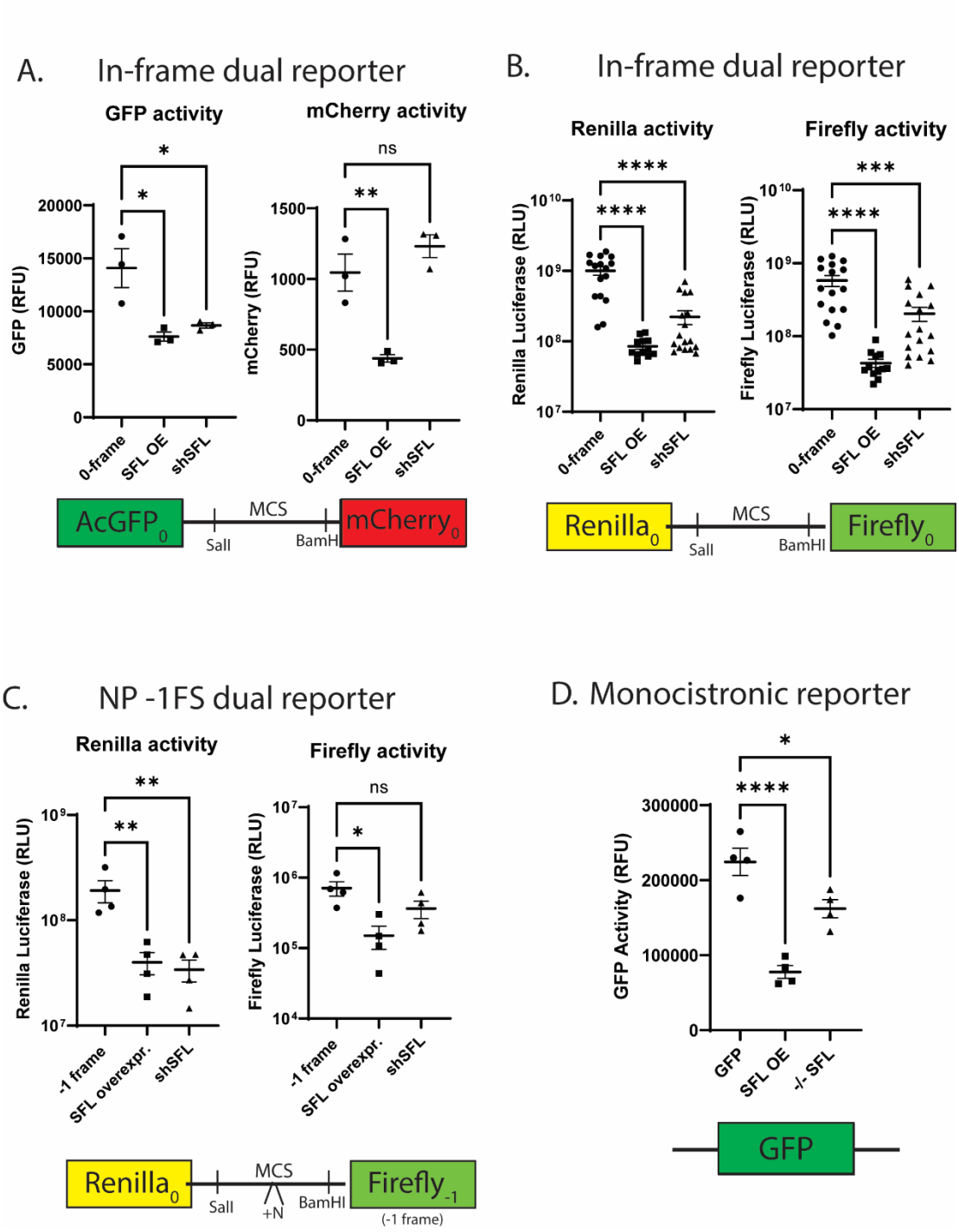


Figure 15. Disruption of SFL homeostasis inhibits reporter activity.

Activity of bicistronic and monocistronic reporters in HEK293T cells over or under-expressing SFL (A) 0-frame dual fluorescence reporter, (B) 0-frame dual luciferase reporter, (C) NP -1FS reporter, and (D) monocistronic GFP reporter.

Disruption of Shiftless homeostasis decreases mRNA steady-state abundances

Given the negative impacts of increased or decreased SFL expression on reporter gene activities, we examined the impacts on mRNA steady state abundances in isogenic WT HEK cells or HEK cells overexpressing SFL by qRT-PCR using 100ng of total cDNA derived from cellular RNAs. Shiftless overexpression decreased the steady state abundances of mRNAs expressed from the GFP monocistronic reporter by 77% ± 5.2% *(p* = 0.0002), Renilla luciferase in the 0-frame dual luciferase reporter by 82% + 13.9% *(p* = 0.0003), and Renilla luciferase in the SARS-CoV-2 by 75.6% $\pm$ 26.1% (*p* = 0.0001) compared to WT HEKs. Importantly, SFL overexpression also decreased the abundance of the endogenous GAPDH mRNA by 44% + 17.9% (*p* = 0.0021) (**Fig. 16**).

In parallel, comparison of mRNA steady state abundances in isogenic WT HEK cells versus $SFL^{-/-}$ HEK cells revealed that the lack of SFL also reduced the abundance of all four mRNAs, although to lesser extents. Specifically, in $SFL^{-/-}$ cells, GFP mRNA was reduced by 25.6% $\pm$ 17.8% *(p* = 0.046), 0-frame Renilla mRNA levels by 53.7% + 15% *(p* = 0.0001), SARS-CoV-2 Renilla mRNA levels by 72.4% $\pm$ 6% (*p* = 0.0002) and GAPDH levels were decreased by 40.9% + 14.3% (*p* = 0.0034) (**Fig. 16**).

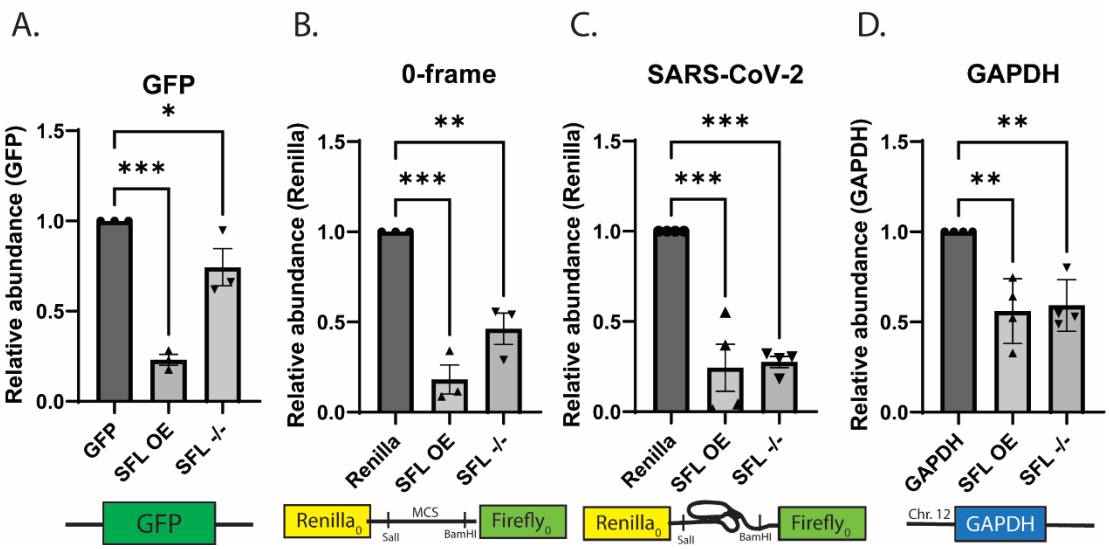


Figure 16. Disruption of SFL homeostasis decreases mRNA steady-state abundances.

Relative mRNA abundance in $SFL^{+/+}$ HEK293T or $SFL^{-/-}$ HEK293T cells. Relative abundance of (A) GFP mRNA from a monocistronic reporter, Renilla luciferase mRNA from the (A) 0-frame control or (B) SARS-CoV-2 dual luciferase reporter, (D) GAPDH mRNA.

Genetic analysis of Shiftless in the context of ribosome quality control (RQC).

RQC is initiated by disome formation when a trailing ribosome encounters a paused downstream ribosome[18,19]. SFL is known to bind tightly to hyper-rotated, frameshifted ribosomes, where it presumably causes them to pause[69]. This suggests that defects in SFL expression may affect the kinetics of pausing of frameshifted ribosomes, thus affecting RQC activation. Paused ribosomes direct recruitment of numerous factors including the GIGYF2-4EBP complex, ZNF598, and EDF1[165]. GIGYF2-4EBP inhibits initiation by new ribosomes. Easily resolved pauses result in recruitment of the tRNA-eEF1A-GTP ternary complex, resulting in resumption of translation elongation and depression of initiation. In contrast, persistent collisions stimulate eS10 and uS10 ubiquitination by ZNF598, which in turn stimulates ASCC3 recruitment to the disome[17,20,21]. This complex then recruits the no-go mRNA machinery which includes Pelota and Hbs1L.

The effects of SFL homeostasis disruption on RQC was tested by knocking down the expression of ZNF598, ASCC3, HBS1L or Pelota in HEK cells either overexpressing SFL, or in the $SFL^{-/-}$ background using the in-frame dual luciferase reporter. shRNA knockdown of any of the RQC factors inhibited expression of the upstream Firefly reporter to approximately the same extent as SFL overexpression, but not as much as in $SFL^{-/-}$ cells (**Fig. 17A**). The combination of ZNF598 knockdown plus SFL overexpression was similar to either of these conditions on their own. In contrast, ZNF598 knockdown in $SFL^{-/-}$ cells appeared to have a synergistic effect, reducing Renilla expression to an even greater extent than either of the two conditions alone. Similar patterns were observed upon shRNA knockdown of ASCC3 (**Fig. 17B**), and HBS1L (**Fig. 17C**). In contrast, synergy between Pelota and SFL overexpression was not observed (**Fig. 17D**). Rather, all conditions inhibited Renilla expression to similar extents. To probe potential positional effects, the effects of these conditions on expression of the in-frame downstream firefly luciferase reporter were also evaluated. These analyses revealed similar results, i.e. the absence of SFL had synergistic effects in combination with shRNA knockdown of ZNF598, ASCC3 and HBSl1, while SFL overexpression was dominant to Pelota knockdown (**Fig. 17**).

Nonsense-mediated mRNA decay (NMD) represents a second arm of the RQC apparatus. Importantly, ribosomes that encounter a termination codon in the wrong context have been observed to slip back and forth on mRNAs[166], suggesting that they too may recruit SFL. Knockdown of the NMD factor

SMG1 also inhibited expression of the upstream Renilla reporter, there was no synergy with either SFL overexpression or lack of expression (**Fig. 17E**). in contrast, expression of the downstream firefly reporter was more strongly affected by SFL overexpression, and lack of SMG1 in combination with SFL overexpression more closely resembled SMG1 knockout than SFL overexpression alone, suggesting that SMG1 may be epistatic to SFL.

The effects of knocking down RQC factors in the context of the HIV-1 -1 PRF were also examined. Lack of ZNF598 enhanced expression of the downstream firefly reporter relative to upstream Renilla luciferase reporter, resulting in a net increase in -1 PRF efficiency (**Fig. S4A**). SFL overexpression resulted in a small but statistically insignificant decrease in the downstream reporter relative to the upstream one, resulting in a net decrease in -1 PRF efficiency (again, not statistically significant). The same patterns were observed upon shRNA knockdown of ASCC3, HBSL1, Pelota or SMG1 (**Fig. S4B-D**).

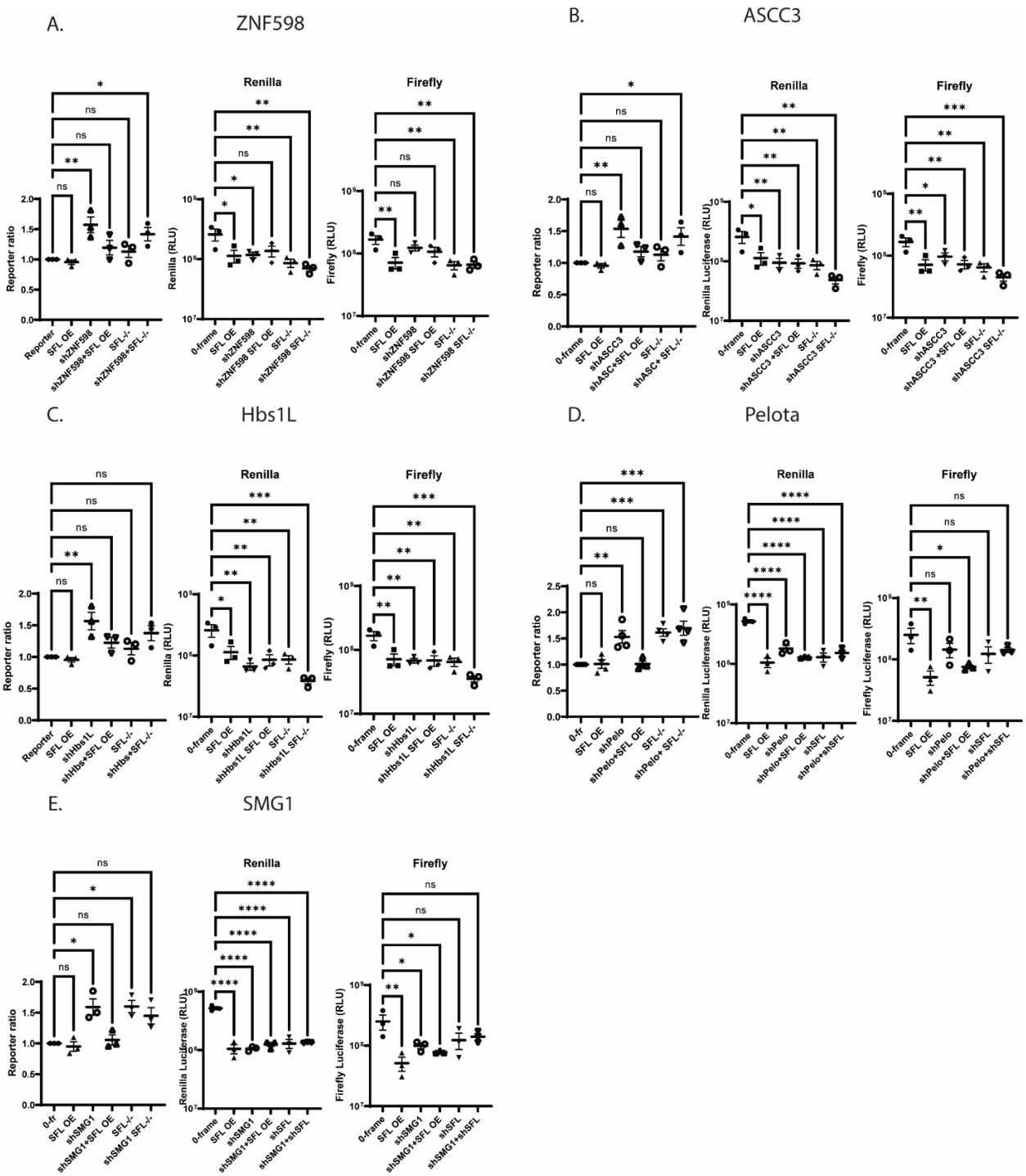


Figure 17. Shiftless and RQC.

In-frame control reporter protein ratios and luciferase activity with shRNA knockdown of RQC proteins in WT HEK293T or SFL$^{-/-}$ HEK293T. (A) ZNF598, (B) ASCC3, (C) Hbs1L, (D) Pelota, or (E) SMG1 shRNA knockdown.

Shiftless expression is significantly reduced in many common cancers and correlates with worse clinical outcomes.

The Cancer Genome Atlas (https://www.cancer.gov/tcga) and Genome Browser[167] databases were used to examine potential changes in SFL expression across a panel of the most common cancer types[168] and their corresponding normal tissues. SFL expression is decreased in all cancers examined by more than 30% except in cutaneous melanoma (SKCM) (**Fig. 18A, Table S1**). In addition, ZNF598 expression was also decreased in most cancer types except mesothelioma (MESO) (**Fig. 18B**). Expression of other members of RQC including Pelota (**Fig. 18C**), SMG1 (**Fig. 18D**) did not correlate across cancer types, nor were patterns observed in the expression of reference genes identified to have stable expression patterns in human cancer and matching normal cell types[169,170]. Strikingly, lower levels of SFL expression correlated with reduced survival in lung mesothelioma, bladder, and skin cancers (**Fig. S5**).

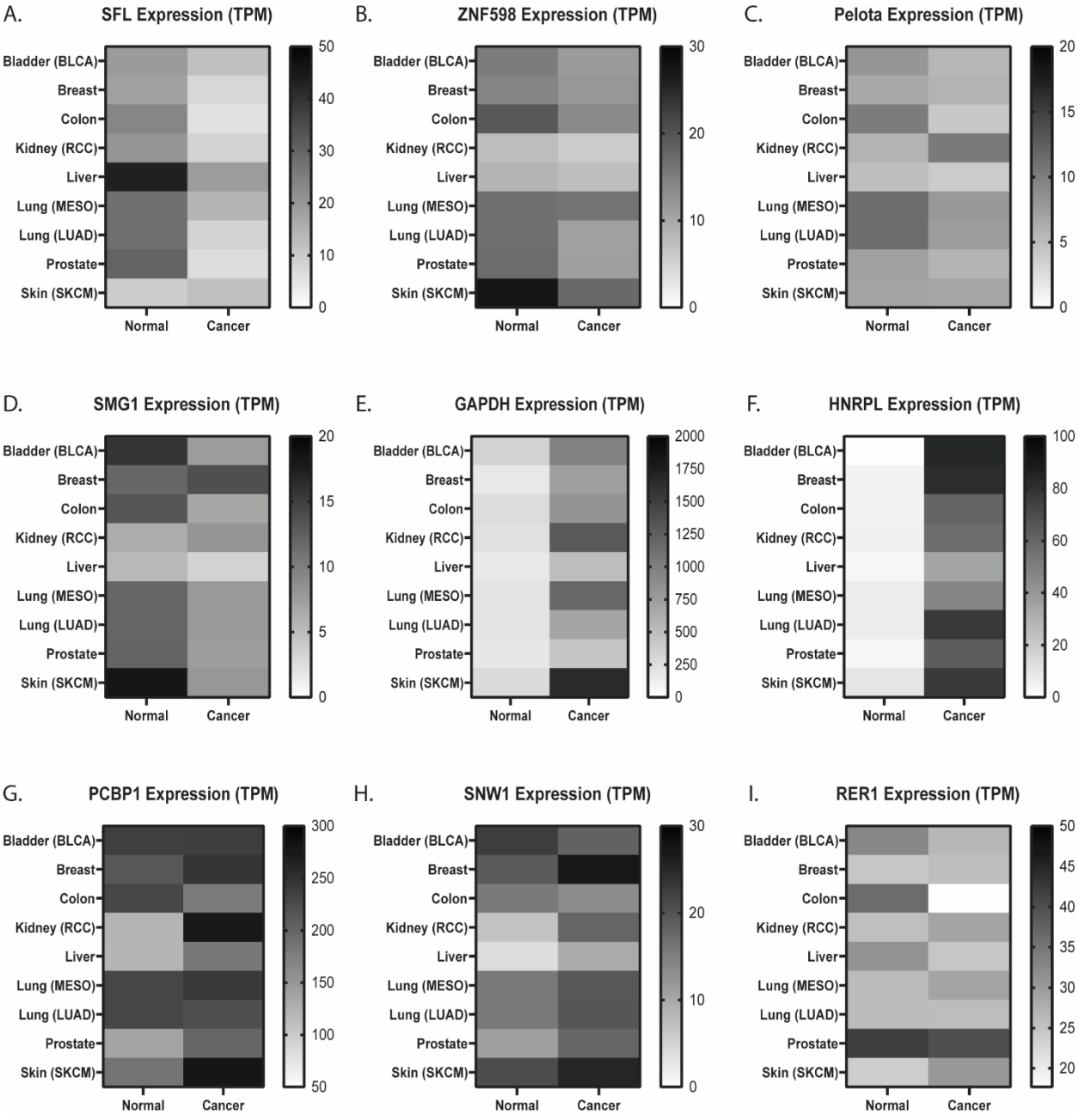


Figure 18. Gene expression pattern changes in Cancer.

Average TPM of RQC components or control genes in common cancers and corresponding normal tissue. (A) Shiftless, (B) ZNF598, (C) Pelota, (D) SMG1, (E) GAPDH, (F) HNRPL, (G) PCBP1, (H) SNW1, and (I) RER1.

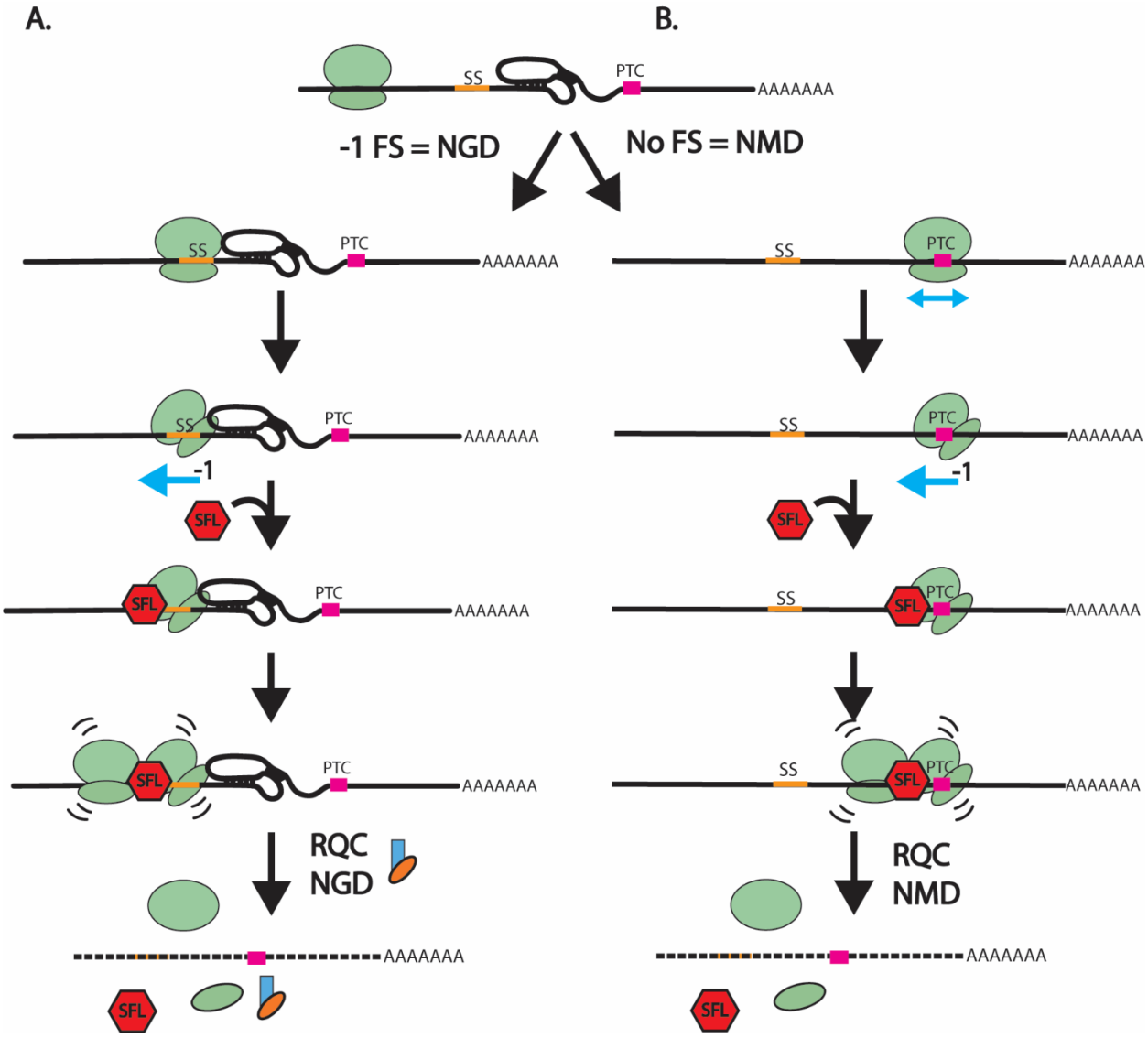
Model

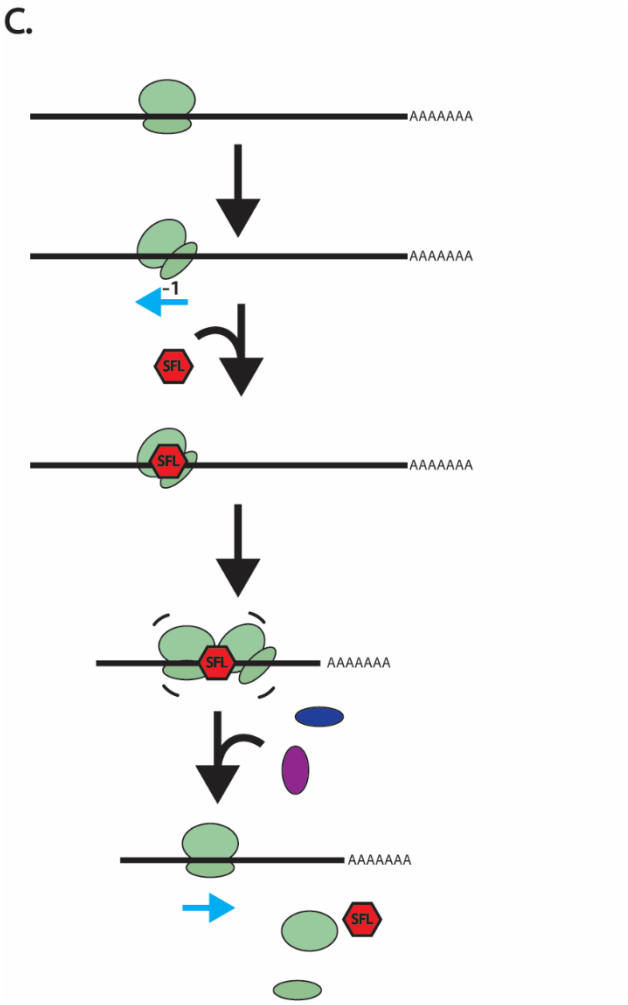
Prolonged ribosomal pausing and collisions are emerging as common denominator between the RQC surveillance pathways[18]. Here, we propose that Shiftless is a member of RQT and acts in concert with NGD and NMD (**Fig. 19**). Although the role of Shiftless in the context of a viral infection has been well described[69], -1 PRF signals can act as substrates for both NGD and NMD. In the NGD arm, shiftless binds to the -1 frameshifted ribosome and arrests translation, resulting in a ribosome collision that triggers the NGD pathway (**Fig. 19A**). In the NMD arm, the ribosome encounters a PTC adjacent to the -1 PRF signal and pauses (**Fig. 19B**). Ribosomes paused at PTCs in the middle of ORFs have been shown to slide back and forth on the mRNA while "waiting" for recruitment of release factors and or a ribosome collision to instigate NMD[166]. In cells expressing SFL, this pausing/oscillation may result in a spontaneous -1 frameshift, thus recruiting shiftless to the ribosome, followed by eRF1, eRF3, and NMD mediators[28,29,69].

During canonical translation, SFL likely monitors spontaneous -1 frameshifting (**Fig. 19C**). Here, a ribosome translating a "flat" message undergoes a spontaneous frameshift. Shiftless binds to the frameshifted ribosome, recruits NGD factors Hbs1L and Pelota and ribosome rescue ensues via NGD.

Figure 19. Model of Shiftless as a member of RQC.

Proposed mechanism of Shiftless in mRNA surveillance during infection (**A-B**) and canonical translation (**C**). (**A**) Shiftless in the No go decay pathway. A translating ribosome encounters a viral -1 PRF signal and undergoes a -1 frameshift. SFL recognizes, binds, and arrests translation of the shifted ribosome. A trailing ribosome will collide with the shifted ribosome, inducing RQC and the no go decay pathway. (**B**) Shiftless in the nonsense mediated decay pathway. If a ribosome does not shift at a viral frameshift signal and encounters a PTC. The ribosome, stalls, slides back and forth on the message and undergoes a spontaneous -1 frameshift. SFL is recruited to the shifted ribosome, arrests translation, induces a collision with the trailing ribosome and triggers nonsense mediated decay. (**C**) Shiftless as a monitor of spontaneous -1 frameshifting triggers the no go decay pathway. A ribosome undergoes a spontaneous -1 frameshift during translation. SFL binds, arrests translation and collides with a trailing ribosome. The frameshifted, lead ribosome is "rescued" by eRF1 and eRF3 while the trailing ribosome resumes translation of the mRNA.


A.
B.
SS
PTC
AAAAAAA
-1 FS = NGD
No FS = NMD
-1
SFL
RQC
NGD
RQC
NMD


C.
AAAAAAA
AAAAAAA
-1
SFL
SFL
AAAAAAA
SFL
AAAAAAA
AAAAAAA
SFL

Discussion

The current study sought to expand our understanding of Shiftless with respect to its effects on translational recoding, gene expression in general, and its relationship with RQC. Though SFL is significantly upregulated during viral infection, it is expressed constitutively in a variety of human-derived cell lines (**Figure S1** and ref. 7)[72]. Furthermore, there are conflicting studies on the mechanism of Shiftless[69,72,161]. In Wang et al, Shiftless was identified as a -1 PRF-specific inhibitor, active against a panel of human and viral-derived -1 PRF signals, but with no influence on +1 frameshifting or termination codon readthrough[69]. Other studies agree that this protein has a potent impact on viral replication yet is active against viruses that do not harbor known -1 PRF signals like Dengue virus and HCV[72,161,171–173]. A more recent study demonstrated SFL is not -1 PRF specific and can bind to other translational recoding signals with similar affinity[162], but it still leaves an ambiguous role for the true mechanism of Shiftless. The results presented here are in accordance with the latter study, indicating that SFL can also modulate +1 PRF and termination codon readthrough.

The current study and previous work found constitutive SFL expression in the cell (**Fig. S1A** and Balinsky et al[72]). Disruption of SFL homeostasis not only results in altered -1 PRF efficiency, but also influences reporters that do not contain a programmed recoding signal (**Fig. 13 and 14**). Most intriguingly, reduction of SFL led to a significant increase in a reporter monitoring non-programmed ribosomal frameshifting (**Fig. 14C**). Spontaneous, non-programmed ribosomal frameshifting occurs at a rate of 10^{-4} to 10^{-5} per codon and monitoring or resolution of this issue has been an outstanding question for decades[8–10]. Though NP -1FS is not likely to influence translation of small to medium-sized mRNAs as the average eukaryotic mRNA is ~1.2kb[174], it is a major issue for larger proteins like titin and BRCA2 that are encoded by mRNAs >10kb. Since this type of recoding error will not immediately invoke established mRNA surveillance mechanisms like NMD, these results suggest it is highly likely that SFL acts as a monitor of NP -1FS to detect and rescue spontaneously frameshifted ribosomes.

Recoding data is generated by calculating the ratios of ratios of reporter protein activities. This approach tends to reduce information content by masking general effects on translation. Deconvoluting the reporter data revealed significant effects on reporter gene expression consequent to SFL over- and lack of expression. This and further analyses revealed that SFL is a general monitor of translational fidelity.

Specifically, disruption of Shiftless homeostasis generally decreases gene expression, both at the levels of mRNA abundance and translational output.

The importance of Shiftless in cellular gene expression is further revealed when looking at cancer. Though cancer causes major changes to the translational landscape of a cell, expression patterns of cellular genes can vary widely depending on the type of tumor. Oftentimes genes involved with growth checkpoints and translation regulation are down regulated while genes promoting cell growth and metabolism are increased. Expression of SFL in cancerous tissue is consistently and significantly decreased among common cancer types compared to normal tissues (**Fig. 18** and **Table 6**). In contrast, other members of RQC are not uniformly dysregulated (**Fig. 18**). This supports the hypothesis that Shiftless works to regulate cellular gene expression and serves an important role maintaining cellular homeostasis.

Experimental Procedures

Cell culture and plasmid transfections

Human embryonic kidney (HEK293T/17) (CRL-11268) and HeLa (CCL-2) cells were purchased from the American Type Culture Collection (Manassas, VA). HEK 293T cells were maintained in Dulbecco's modified Eagle's medium (Corning Life Sciences 10-013-CV Durham, NC) supplemented with 10% fetal bovine serum (Life Technologies 26140079 Carlsbad, CA) 1% GlutaMAX (35050061), 1% nonessential amino acids (Gibco 11140050), 1% HEPES buffer (Cytiva Life Sciences SH30237.01) and 1x Penicillin/Streptomycin (Gibco 15140122). HeLa cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1% GlutaMAX, and 1x Penicillin/Streptomycin. Both cell lines were incubated at 37°C in 5% CO_2. The cells were transfected with a total of 500ng (dual luciferase assay) or 1ug (bi-fluorescence and qRT PCR assays) plasmid DNA 24 hours after seeding using Lipofectamine 3000 (Invitrogen L3000015) per the manufacturer's protocol.

shRNA knockdowns

A panel of short hairpin shRNA plasmids targeting SFL and commercially validated shRNAs targeting Pelota, SMG1, ASCC3, Hbs1L, and ZNF598 were purchased from Millipore Sigma. shRNA targets are listed in **Table 5.** Five different SFL shRNAs were assayed and knockdowns were conducted by

transfecting 250ng (dual luciferase assay) or 500ng (bi-fluorescence assay) shRNA plasmid DNA into HEK293T cells 24 hours after seeding using Lipofectamine 3000 per the manufacturer's protocol. RT-PCR analysis revealed that shRNA 1 and shRNA 5 promoted the strongest knockdown of SFL (**Fig S2)** SFL knockdown was also validated using RT PCR and qRT PCR (**Supplemental figure S1C).**

Generation of CRISPR knockout cell lines

A SFL$^{-/-}$ HEK293T cell line was generated using CRSIPR-Cas9 as described in Ran et al[175]. HEK923T cells were transfected with spCas9(BB)-2A-Puro plasmid containing a gRNA specific to C19orf66 (5' CGTGTATCCAACACGGATCC 3') designed to result in a 1 base deletion in the ORF using Lipofectamine3000 per manufacturer's protocol. Cells were selected for the presence of Cas9 by incubating with 1.0ug/mL puromycin. Clonal cell lines were obtained by seeding cells at low density in a 10cm tissue culture dish and then selecting well-isolated colonies for expansion. Clonal lines were screened for mutations by PCR-amplifying C19orf66 from genomic DNA and assessing for mutations using the Surveyor mutation detection kit (Integrated DNA Technologies 706020) per manufacturer's protocol. Positive hits from the Surveyor screen were sequence-verified (**Figure S1D)** and knockout of SFL was validated via qRT PCR **(Figure S1C).**

Growth curve of SFL$^{-/-}$ HEK293T cells

HEK WT and SFL$^{-/-}$ cells were seeded at a density of 10^4 cells per well of a 12-well plate. Cells were trypsinized and counted at 24-hour intervals for a total of 168 hours. Cell doubling times were determined using the equation:

$$DT = \frac{T*\ln(2)}{\ln(\frac{X_e}{X_b})}$$

Where T = time in hours, X_e = endpoint cell count, X_b = beginning cell count.

These data are shown in **Figure S2.**

Preparation of reporter plasmids

Dual luciferase reporters containing -1 PRF signals of SARS-CoV, SARS-CoV-2, and human CCR5 listed in **Table 2** were made by digesting pJD2257 with *Sal* I and *Bam* HI, gel-purifying digest products, and ligating a DNA oligonucleotide insert (IDT) containing the -1 PRF signal of interest into the plasmid using T4 DNA ligase (NEB). pJD2257 is derived from pSGDLuc[176] into which Sal I and Bam HI sites were inserted so as to preclude possible distortions of luciferase reporter readouts due to insertion of sequences of viral or cellular origin. The non-programmed -1 frameshift (NP-1FS) reporter plasmid was made using site-directed mutagenesis of pJD2257. Site directed mutagenesis primers (**Table 4**) were synthesized and purified by IDT.

Bi-fluorescent reporters containing a panel of human and viral translational control element inserts listed in **Table 2** were made by digesting pJD2261 with *Sal* I and *Bam* HI, gel-purifying digest products, and ligating a DNA oligonucleotide insert (IDT) containing sequences of interest into the plasmid using T4 DNA ligase (NEB).

Products were transformed into DH5α Escherichia coli cells (NEB) and spread onto LB agar plates containing 50mg/ml carbenicillin. Positive clones were verified by DNA sequencing (Genewiz).

Dual luciferase assays of -1 PRF

The frameshifting efficiency of luciferase reporter plasmids in cultured cells was assayed as previously described using a dual luciferase reporter assay system kit (Promega)[90,91]. 24 hours after transfection, cells were washed with 1x PBS then lysed with 1x passive lysis buffer (E194A, Promega). Reporter activity was calculated by measuring the luminescence of firefly or Renilla luciferase in 50µL of cell lysate. Assays were conducted in triplicate in 96-well plates and quantified using a GloMax microplate luminometer (Promega). Percent frameshift was calculated by averaging the three Firefly or Renilla luciferase technical replicate reads per sample then forming a ratio of firefly to Renilla luminescence per sample. Each sample ratio was compared to a 0-frame control set to 100%. At least three biological replicates with three technical replicates each were assayed for each sample. Statistical analyses were conducted using one-way analysis of variance using Prism 9 software (GraphPad).

Bi-fluorescence assays of -1 PRF

Frameshift efficiency of bi-fluorescent reporters in cultured cells were assayed as described previously[59]. HEK293T or HeLa cells were seeded at density of 10^5 cells per well of a 12-well plate in appropriate growth media. After a 24-hour incubation, cells were transfected with 500ng bi-fluorescence reporter plasmid and 500ng of either a Shiftless overexpression plasmid, shRNA plasmid targeting Shiftless, or shRNA scramble control. Additional media was added to cells 24 hours post transfection and incubated for an additional 24 hours. Cells were collected by scraping into Dulbecco's phosphate-buffered saline (Corning 21-031-CV), pelleted by centrifugation then lysed in 150uL Triton lysis buffer (1% Triton X-100, 150mM NaCl, 50mM Tris pH8, 1x Halt protease inhibitor cocktail (Thermo Scientific)). Cell lysates were clarified by centrifugation and assayed in a clear-bottom black-walled 96 well plate (Grenier Bio-One). Fluorescence was quantified using a GloMax microplate luminometer (Promega) with the "green" optical kit (Excitation 525nm, Emission 580-640nm) for mCherry and the "blue" optical kit (excitation 490nm, emission 510-570nm) for AcGFP. Reporter activity and -1 PRF efficiency was corrected for AcGFP bleed over into the mCherry channel by subtracting background fluorescence from mock transfected cells. Reporter activity was calculated by measuring the fluorescence of 150µL clarified cell lysate and subtracting background fluorescence from mock transfected cells. -1 PRF efficiency was calculated as previously described[59]. Statistical analyses were conducted using one-way analysis of variance using Prism 9 software (GraphPad).

qRT PCR and RT PCR methods

To quantitatively measure mRNA abundances (qRT-PCR), total RNA was extracted from cells grown to 70-80% confluency using the Total RNA miniprep kit (NEB T2010S) according to the manufacturer's protocol. cDNA synthesis was performed using the iScript gDNA Clear cDNA synthesis kit (BioRad 1725034) according to the manufacturer's protocol using 500ng total RNA. qPCR was performed with 100ng of total cDNA, 250nM primers in a final volume of 10µL using BioRad CFX96 and SSOAdvanced SYBR green master mix (BioRad 1725270). mRNA abundances for SFL OE and $SFL^{-/-}$ conditions were normalized to WT HEK293T conditions. Primers for this assay are listed in **Table 4**. qPCR conditions were:

4 minutes at 95°C followed by 40 cycles at 95°C for 5 seconds and 60°C for 30 seconds followed by a melt curve from 65°C to 95°C. Experiments were repeated with at least three independent biological replicates.

To monitor the presence of specific mRNAs (RT-PCR), RNA was extracted from HEK293T, HeLa, and U87MG cells grown to 70-80% confluency and an overnight *E. coli* culture using the Total RNA miniprep kit (NEB T2010S) according to the manufacturer's protocol. cDNA synthesis was performed using the iScript gDNA Clear cDNA synthesis kit (BioRad 1725034) according to the manufacturer's protocol. RT PCR was performed using 1µL of cDNA, and 500nM primers in a final volume of 35µL using Q5 Hotstart Mastermix (NEB M0494S). Primers for this assay are listed in **Table 4**. PCR conditions were: 98°C for 30 seconds followed by 40 cycles at 98°C for 10 seconds, 71°C for 15 seconds and 72°C for 20 seconds, then 72°C final extension for 2 minutes.

Bioinformatics analysis of Shiftless expression in cancer cells

The expression of SFL was analyzed in common cancer tissues and normal tissues deposited in the TGCA database using Genome Browser[167] (**http://genome.ucsc.edu**). Average transcripts per million (TPM) for each normal and cancerous cell type were displayed on a heatmap using GraphPad Prism 9 software. Survival corresponding to SFL expression levels among patients with various cancers was analyzed using the UALCAN platform[177].

Chapter 5: Conclusions and future directions

Regulation of programmed and non-programmed ribosomal frameshifting is an essential aspect of translational fidelity. Programmed frameshifting mechanisms and their regulation have been studied in great detail and are emerging as a target for the next generation of antiviral therapeutics. Regulation of non-programmed, or spontaneous frameshifting, however, remains somewhat elusive. Though quality control mechanisms have been established for ribosomes that encounter a premature termination codon, secondary RNA structures that impede ribosome processivity, and collisions between ribosomes, to date there is no established surveillance pathway that monitors for spontaneously frameshifted ribosomes. This thesis explored the programmed -1 frameshift signal of SARS-CoV-2 and its potential role as an antiviral target in chapters 2 and 3. Non-programmed frameshifting was investigated in chapter 4 and establishes the -1 PRF inhibitor Shiftless as a monitor of both spontaneous and programmed frameshifting in translational fidelity.

Chapters 2 and 3 surveyed the importance of -1 PRF in coronaviruses and established the frameshift element of SARS-CoV-2 as an attractive target for antiviral therapeutics. In chapter 2, we found high structural and functional similarity between the frameshift elements of SARS-CoV and SARS-CoV-2. Next, we demonstrated that a small molecule compound, MTDB, that reduces frameshifting in SARS-CoV is similarly effective in reducing frameshifting in SARS-CoV-2. Subsequent research has focused on identifying FDA-approved compounds that interfere with viral frameshifting in SARS-CoV-2 in addition to other crucial RNA elements in the viral genome. In chapter 3, we explored decades worth of research into the importance of -1 PRF and applied it to current measures for combatting the COVID-19 pandemic and preparation for the next emergence of new RNA viruses into the human population.

To date, several small molecule frameshift inhibitors have been identified to alter -1 PRF efficiency in SARS-CoV-2 and other related coronaviruses[58,64,65,80,159,178–180]. Though there are numerous potential targets for antiviral compounds, the -1 PRF signal is attractive because of its importance for viral replication, high sequence conservation among other SARS-CoV-2 variants as well as high structural conservation in other coronaviruses. RNA viruses often have high mutation rates. Though most mutations are harmful for viral replication, a small but significant number drive viral evolution to stay ahead of their hosts immune system. Mutation rates are highest in areas where the host antibodies can recognize and bind, such as the

viral spike protein, while other areas essential for viral propagation tend to be more conserved[148]. Because of its importance to viral replication, the -1 PRF signal is highly conserved in SARS-CoV-2 variants (**Fig. 20**) while general structure is well-conserved among different types of coronaviruses[82]. A recent study characterized frameshifting in bat-derived coronaviruses which are highly likely to cause future zoonoses in the human population[59]. The authors found several compounds that inhibited frameshifting significantly in a panel of bat and human-derived coronavirus -1 PRF signals. This suggests coronaviruses are less likely to gain resistance to -1 PRF targeting small molecule compounds over time and that these types of molecules could be used as effective broad-spectrum antivirals.

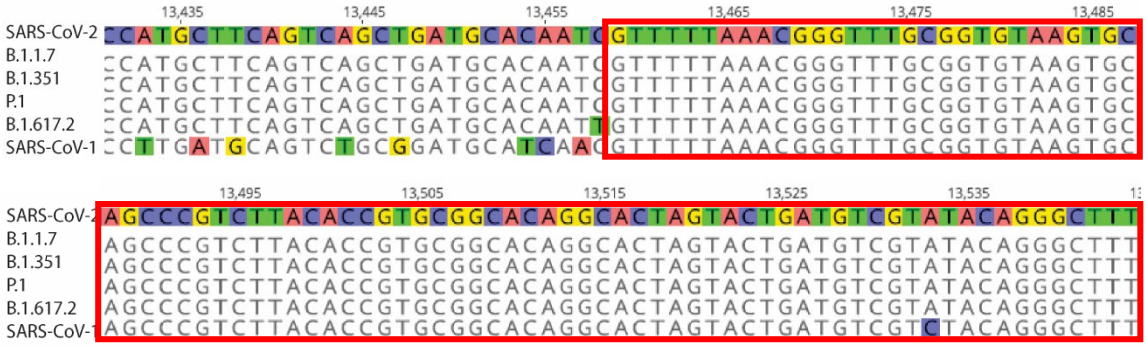


Figure 20. Sequence alignment of -1 PRF element from SARS-CoV-2 reference genome, variants B.1.1.17, B.1.351, P.1, B.1.617.2, and SARS-CoV.

Variant sequences from genome position 13,430-13,545 obtained from the GISAID database[181,182]. -1 PRF sequence boxed in red.

Future studies of -1 PRF in coronaviruses will likely be focused on identifying potent, frameshift-specific small molecule compounds. Though multiple compounds have been identified to alter coronavirus -1 PRF *in vitro*, including the work presented in this thesis, additional studies are needed to 1) identify or modify compounds to bind the CoV -1 PRF element with greater affinity and specificity 2) test the efficacy of these compounds using *in vivo* models and 3) test their efficacy in a clinical setting. Additionally, targeting -1 PRF in other types of viruses will likely be a valuable therapeutic approach. For example, encephalitic

alphaviruses like VEEV and Sindbis virus (SINV) use -1 PRF to generate a protein involved in neuropathogenesis[57,183,184]. Mutating the -1 PRF signal in SINV and VEEV resulted in attenuated neuropathogenicity in mice without compromising the ability of the virus to replicate[57,183]. Identification of small molecule compounds targeting -1 PRF in these types of alphaviruses could prove extremely beneficially to combatting devastating neurological sequelae associated with infection.

In Chapter 4, we explored the -1 PRF inhibitor Shiftless and established a secondary role as a monitor of translational fidelity. Though RQC mechanisms like NGD, NMD, and NSD effectively identify and resolve issues during translation, to date there is no established mechanism that directly monitors spontaneous ribosomal frameshifting in the cell. Since reading frame maintenance is essential for accurate translation of mRNAs and spontaneous frameshifts occur at moderate rates[8–10], it is no surprise that the cellular translational fidelity surveillance machinery includes a factor that identifies and resolves the problem of frameshifted ribosomes. Numerous proteins associated with translation and mRNA binding co-precipitate with Shiftless including poly-A binding proteins PABPC1 and LARP1[161,185], as well as RNA helicases MOV10 and UPF1[72,161]. Though Shiftless may not work with these proteins directly, their proximity suggest that it is indeed involved in translational surveillance. Paired with the knowledge that there is no frameshift-specific factor involved with ribosome quality control, Shiftless likely has other roles besides its function as an antiviral.

Knockdown or knockout of Shiftless expression in cells resulted in a maximum 2-fold increase in apparent frameshifting in both -1 PRF reporters as well as other translational recoding and control reporters. A similar limit was seen in a study of ribosomal pausing kinetics during -1 PRF[186]. This study in *Saccharomyces cerevisiae* hypothesized that increasing the space between ribosomes on an mRNA would allow more time for the -1 PRF element of ScV-L-A virus to reform after denaturation by a “leading” ribosome and increase the likelihood the trailing ribosome on the mRNA would encounter a completely refolded -1 PRF element in contrast to “normal” ribosome spacing where only a portion of ribosomes encounter an intact -1 PRF element. While increased space between ribosomes did increase frameshift efficiency of ScV-L-A virus, the maximum increase in frameshifting was limited to a 2-fold increase[186]. These results suggest that even through the -1 PRF element has more time to recover from one translating ribosome, and more

ribosomes encounter an intact -1 PRF signal, frameshift efficiency is not exclusively dependent on ribosomes encountering an intact frameshift element.

A similar pattern was seen with Shiftless knockdown. We observed a maximum 2-fold increase in frameshifting in the absence of Shiftless using a panel of -1 PRF signals. If Shiftless is a monitor of -1 PRF and captures frameshifted ribosomes, then why does frameshifting efficiency not reach 100% when its expression is eliminated? -1 PRF signals are substrates for both the NGD and NMD pathways. In the context of NGD, ribosomes pause at the RNA stimulatory structure of a -1 PRF signal during a frameshift and can become "stuck" at this location until the NGD surveillance pathway is triggered. Similarly, many -1 PRF signals will result in a PTC if a frameshift does not occur. This will activate the NMD pathway to rescue the ribosome. Both NMD and NGD result in ribosomal pausing resulting in ribosome collisions if the issue is not resolved quickly[18,19,23,187]. While the ribosome spacing studies in *S. cerevisiae* and the Shiftless knockout studies presented here both yielded 2-fold increases in frameshifting, ribosomes in both circumstances will inevitably become substrates for NGD or NMD.

Translation and ribosome biogenesis consume over 60% of the energy in the cell[11]. Because of this, many mechanisms are in place to a) ensure translation is accurate, b) limit the amounts of aberrant mRNAs or peptides and c) ensure that ribosomes are actively translating and promptly recycled after completing peptide synthesis. Established RQC pathways monitor for PTCs, stalled ribosomes, or aberrant mRNAs, yet to date there is no established monitor of frameshifting. Spontaneous frameshifting is an intrinsic property of the ribosome[34]. The frequency of spontaneous frameshifting is relatively high (~10^{-4}/codon). The average length of a eukaryotic protein is 400 amino acids/1200 nt[174], meaning that ~1 in 10^4 ribosomes will spontaneously frameshift while translating such an mRNA. Additionally, the coding regions of many essential proteins, e.g. alpha 5 laminin, BRCA2 and titin are >10,000 nt long, effectively rendering them theoretically impossible to translate without at least one spontaneous frameshift event occurring during translation. Thus, the evolution of a frameshift error correction mechanism must have been essential. Here, we observed disruption of SFL homeostasis not only influences programmed frameshifting, but also non-programmed frameshifting **(Figure 14)**.

In the context of RQC, Shiftless is likely one of the first players recruited to the scene. Similar to SFL knockdown, shRNA knockdowns of various RQC factors resulted in 2-fold increases in reporter ratios

or frameshifting (**Figure 17**). When SFL is overexpressed, ratios return to normal levels. Since SFL overexpression in combination with RQC component knockdown yields no net change in reporter ratio, this suggests SFL likely acts upstream of these proteins. This is further supported by the fact that SFL has been demonstrated to interact with RNA binding proteins and translation regulators including: Poly(A) binding protein cytoplasmic 1 (PABPC1) , La-motif related protein 1 (LARP1), MOV10, and UPF1[72,161].

PABPC1 and LARP1 both interact with the poly(A) tail of mRNA. PABPC1 is an RNA binding protein involved in translation and mRNA stability[188]. During translation, PABPC1 associates with both the poly(A) tail of mRNAs and eIF4G in the PIC, forming a closed-loop structure[189,190]. PABPC1 also plays a role during viral infection. HIV-1 protease targets and cleaves PABPC1 during infection to shut down host mRNA translation while viruses like influenza and dengue virus (DENV) interact with it to promote more efficient translation of viral RNAs[191–193]. LARP1 binds to the 3' terminus of the poly(A) tail, helps regulate gene expression in eukaryotes, and is required for efficient replication of DENV[161,185,194]. UPF1 has multiple functions during translation, is a key component of the nonsense-mediated decay pathway and its interactions with SFL could be the physical link of SFL to NMD and RQC[26,30,72]. MOV10 is an RNA helicase involved with translation processing bodies, translation suppression as an interacting partner of RISC, and degradation of viral RNA during infection[195]. Furthermore, several studies have demonstrated MOV10 interacts with polysomes[196,197]. Though the role of SFL in ribosome collisions needs to be evaluated, its interactions with MOV10 support the hypothesis that SFL may cause ribosome collisions and the subsequent formation of disomes or polysomes.

Cancerous cells harbor mutations that result in over-proliferation and a lapse in checkpoints ensuring proper replication timing and often translational control. We found that Shiftless is significantly under expressed in a panel of common human cancers while other proteins involved in translational fidelity like ZNF598, Pelota, or SMG1 did not have consistent expression pattern changes in cancerous cell types (**Figure 18**). Furthermore, tumors expressing lower levels of Shiftless may be linked to decreased survival in bladder cancer, mesothelioma, or cutaneous melanoma (**Figure S5**). Not only does this suggest Shiftless is an important monitor of translation regulation but also may be a survival marker for some types of cancer.

Many questions remain about the mechanism of SFL in the context of RQC. In this work, we established an additional role of SFL as a monitor of spontaneous ribosomal frameshifting and added it to

the roster of RQC players. Not only did this solve a longstanding problem regarding how ribosomes maintain the translational reading frame for large mRNAs but also brought new insight into how frameshift events trigger RQC. The biggest remaining question with this work is the involvement of ribosome collisions. Recent studies have brought ribosomal stalling and ribosome collisions to the forefront of RQC, demonstrating the context of these collisions is crucial for how the stall will be resolved. In the case of minor or transient collisions, ribosomes are free to continue translating while more severe collisions will invoke RQC or the integrated stress response (ISR)[18,19,23,198]. Here, we hypothesize that SFL will stall the frameshifted ribosome on its mRNA and invoke a ribosome collision with the trailing ribosome. Ongoing experiments are determining how disruption of SFL homeostasis influences disome and polysome formation.

The ISR is a pathway involved in cell stress response and is triggered by a variety of conditions like hypoxia, nutrient deprivation, viral infection, oncogene activation, and ER stress[199]. Triggered by the phosphorylation of eIF2α, activation of the ISR results in a global decrease in translation while select genes work to aid cell recovery[200,201]. Because of its role as an antiviral and monitor of translational fidelity, SFL may also be involved with the ISR.

Finally, the structure of SFL remains elusive. Though SFL binds RNA and has established interactions with ribosomal and translation-associated proteins, structural data could provide important clues to the mechanism and interactions of Shiftless[69,72,161,162].

Appendices

Appendix 1: Plasmids

Table 1: Plasmids used in Chapter 2

Plasmid	Description
pJD2257	Modified pSGDluc (dual luciferase with inteins) Readthrough control
pJD2359	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV -1 PRF insert
pJD2514	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 -1 PRF insert
pJD2515	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV silent slippery site mutant insert
pJD2516	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 silent slippery site mutant insert
pJD2517	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV attenuator hairpin and -1 PRF insert
pJD2518	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV disrupted attenuator hairpin and -1 PRF insert
pJD2519	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV2 attenuator hairpin and -1 PRF insert
pJD2520	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 disrupted attenuator hairpin and -1 PRF insert
pJD2521	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV -1 PRF insert containing silent mutations in Stem 1
pJD2522	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV -1 PRF insert containing silent mutations in Stem 2
pJD2523	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV -1 PRF insert containing silent mutations in Stem 3
pJD2524	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 -1PRF insert containing silent mutations in Stem 1
pJD2525	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 -1 PRF insert containing silent mutations in Stem 2
pJD2526	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 -1PRF insert containing silent mutations in Stem 3

Table 2: Plasmids used in Chapter 4

Plasmid name	Description
pJD2257	Modified pSGDluc (dual luciferase with inteins) with 0-frame control sequence insert
pJD2256	Modified pSGDluc (dual luciferase with inteins) with HIV-1 -1 PRF sequence insert
pJD2258	Modified pSGDluc (dual luciferase with inteins) with CCR5 -1 PRF sequence insert
pJD2359	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV -1 PRF sequence insert
pJD2514	Modified pSGDluc (dual luciferase with inteins) with SARS-CoV-2 -1 PRF sequence insert
pJD2450	Modified pSGDluc (dual luciferase with inteins) with non-programmed -1 FS sequence insert
pJD2262	Dual fluorescent construct with 0-frame control sequence insert
pJD2261	Dual fluorescent construct with HIV-1 -1 PRF sequence insert
pJD2281	Dual fluorescent construct with CCR5 -1 PRF sequence insert
pJD2529	Dual fluorescent construct with SARS-CoV-2 sequence insert
pJD2350	Dual fluorescent construct with OAZ1 +1 PRF sequence insert
pJD2455	Dual fluorescent construct with VEEV stop codon readthrough sequence insert
pJD2260	pMAX GFP monocistronic reporter
pJD2612	C19orf66 ORF in pCMV-Myc

Appendix 2: Oligonucleotides

Table 3: Oligonucleotides in Chapter 2

Oligonucleotide Description	Sequence (5'→3')
SARS-CoV2 -1 PRF WT Fwd	CTGATGTCGTaTACAGGGCTTTTG
SARS-CoV2 -1 PRF WT Rev	TACTAGTGCCTGTGCCGC
SARS-CoV -1 PRF Attenuator hairpin Fwd	gtctgcggatgcatcaacGTTTTTAAACGGGTTTGCGGTGTAA GTGCAG
SARS-CoV -1 PRF Attenuator hairpin Rev	tgcatcaagggttcgcggGTCGACGAGGGCCCGGGG
SARS-CoV2 -1 PRF Attenuator hairpin Fwd	gtcagctgatgcacaatcGTTTTTAAACGGGTTTGCGGTGTAA GTGCAG
SARS-CoV2 -1 PRF Attenuator hairpin Rev	tgaagcatgggttcgcggGTCGACGAGGGCCCGGGG
SARS-CoV -1 PRF Attenuator mutant Fwd	ccgacgcttctaCGTTTTTAAACGGGTTTGCGGTGTAAG
SARS-CoV -1 PRF Attenuator mutant Rev	ctgattgcattaaGGGTTCGCGGGTCGACGA
SARS-CoV2 -1 PRF Attenuator mutant Fwd	ccgacgcccagtCGTTTTTAAACGGGTTTGCGGTGTAAG
SARS-Cov2 -1 PRF Attenuator mutant Rev	ccgattggagtatGGGTTCGCGGGTCGACGA
SARS-CoV -1 PRF Stem 1 mutant Fwd	cccgttttgcatcgaGCGGCACAGGCACTAGTA
SARS-CoV -1 PRF Stem 1 mutant Rev	ctgcacctagactgcAAACCCGTTTAAAAACGTTGATGC
SARS-CoV2 -1 PRF Stem 1 mutant Fwd	cccgttttgcatcgaGCGGCACAGGCACTAGTA
SARS-CoV2 -1 PRF Stem 1 mutant Rev	ctgcacctagactgcAAACCCGTTTAAAAACGATTGTGC
SARS-CoV -1 PRF Stem 3 mutant Fwd	agtactgatgtcttCTACAGGGCTTTTGATGGATC
SARS-CoV -1 PRF Stem 3 mutant Rev	agtgcttgggcggcACGGTGTAAGACGGGCTG
SARS-CoV2 -1 PRF Stem 3 mutant Fwd	agtactgatgtcttATACAGGGCTTTTGATGGATCC
SARS-CoV2 -1 PRF Stem 3 mutant Rev	agtgcttgggcggcACGGTGTAAGACGGGCTG
SARS-CoV -1 PRF Stem 2 mutant Top	tcgacgTTTTTAAACGGGTTTGCGGTGTAAGTGCAaCCaG TCTTACACCGTGCGGCACAGGCACTAGTACTGATGTC GTCTACAGGaCTTTTGATg
SARS-CoV -1 PRF Stem 2 mutant Bottom	gatccATCAAAAGtCCTGTAGACGACATCAGTACTAGTGC CTGTGCCGCACGGTGTAAGACtGGtTGCACTTACACCG CAAACCCGTTTAAAAAcg
SARS-CoV2 -1 PRF Stem 2 mutant Top	tcgacgTTTTTAAACGGGTTTGCGGTGTAAGTGCAaCCaG TCTTACACCGTGCGGCACAGGCACTAGTACTGATGTC GTATACAGGaCTTTTGATg
SARS-CoV2 -1 PRF Stem 2 mutant Bottom	gatccATCAAAAGtCCTGTATACGACATCAGTACTAGTGC CTGTGCCGCACGGTGTAAGACtGGtTGCACTTACACCG CAAACCCGTTTAAAAAcg

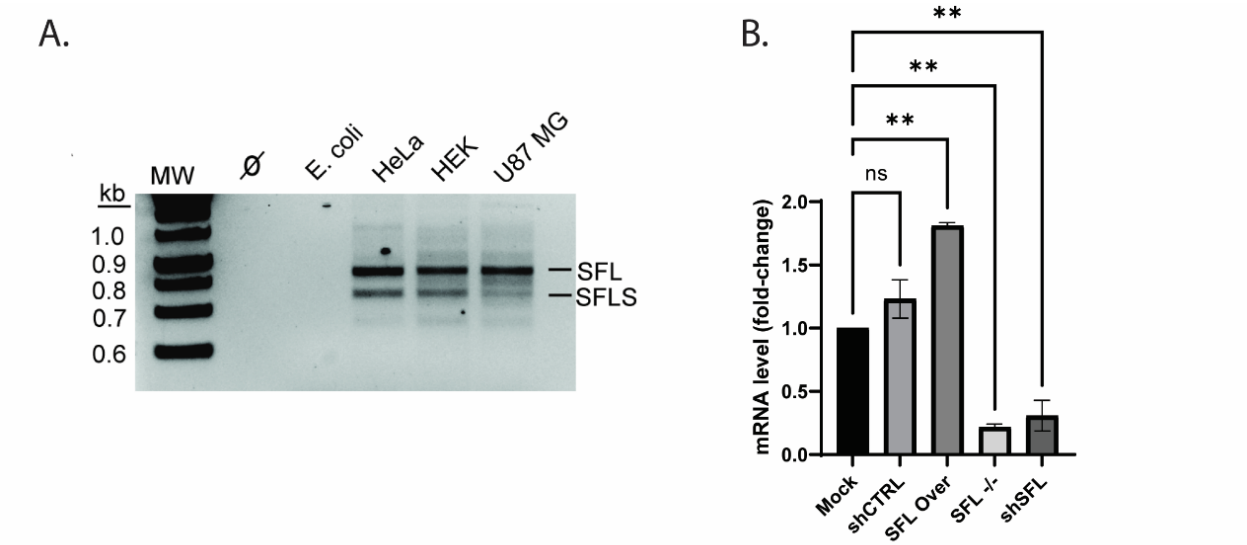
Table 4: Oligonucleotides used in Chapter 4

Oligo name	Sequence (5'→3')
H_GAPDH_F	GGATGATGTTCTGGAGAGCC
H_GAPDH_R	CATCACCATCTTCCAGGAGC
GFP_qPCR_F	GGCTACGGCTTCTACCACTT
GFP_qPCR_R	CTCGTACTTCTCGATGCGGG
Firefly Luciferase qPCR Fwd	TCGCCTCTCTGATTAACGCC
Firefly Luciferase qPCR Rev	ATTACACCCGAGGGGGATGA
C19orf66_qPCR	Biorad PrimePCR qHsaCED0003572
C19orf66_RT_PCR_F	CTCAGGAAGGTGTGGAGCTG
C19orff_RT_PCR_R	GCCACTGCTAATGTGAGGGT
Pelo_qPCR_F	GACCGACAACAAACTGCTCCTG
Pelo_qPCR_R	AGCCACAGTAGGGTCACAAAGG
SMG1_qPCR_F	ATGCTGGTGAGCTTCGGCAGTA
SMG1_qPCR_R	CGCACATACACTTCAGGGTGGT
CRISPR validation primer Fwd	AAATCTGGCTTCTGAACCTCCT
CRISPR validation primer Rev	GTGGGAGACAAAGTGGACTGAG
SFL gRNA Top	CACCG GGATCCGTGTTGGATACACG
SFL gRNA Bottom	AAAC CGTGTATCCAACACGGATCC C
CRISPR validation primer Fwd	AAATCTGGCTTCTGAACCTCCT
CRISPR validation primer Rev	GTGGGAGACAAAGTGGACTGAG
NP-1FS_SDM_F	AAAGAGGCTGCGGCAAAAGC
NP-1FS_SDM_R	GCGGCTGCTTCGGTCGAC

Table 5: shRNA target sequences used in Chapter 4

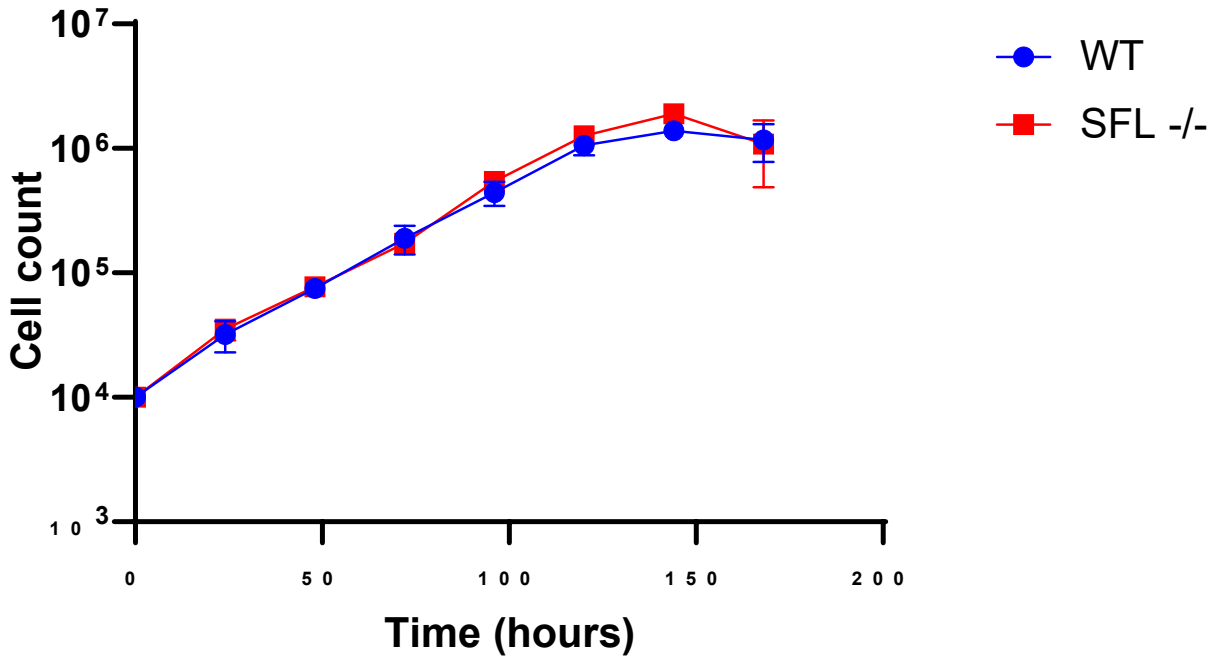
shRNA Target	Target sequence (5'→3')	Cat#
SFL 1	GTGTATCCAACACGGATCCTC	TRCN0000434142
SFL 2	AGCAACCCTCACATTAGCAGT	TRCN0000420530
SFL 3	GAAGTTTCATGGGAAGGTATC	TRCN0000164654
SFL 4	GAAGTTCTGTGGGACACATTG	TRCN0000418723
SFL 5	CCAAGAACTAAGTAACGATCT	TRCN0000161786
ASCC3	TGAGGAGCGAACTGGATATTT	TRCN0000296023
Hbs1L	GCGATCTATTGACAAACCTTT	TRCN0000353597
PELO	GCAGTGAAGACCGACAACAAA	TRCN0000163394
SMG1	GCACTGTAACTACGGCTACAA	TRCN0000037413
ZNF598	CCAACCCTCTAAAGTTGGGAA	TRCN0000222610

Appendix 3: Supplementary figures for Chapter 4



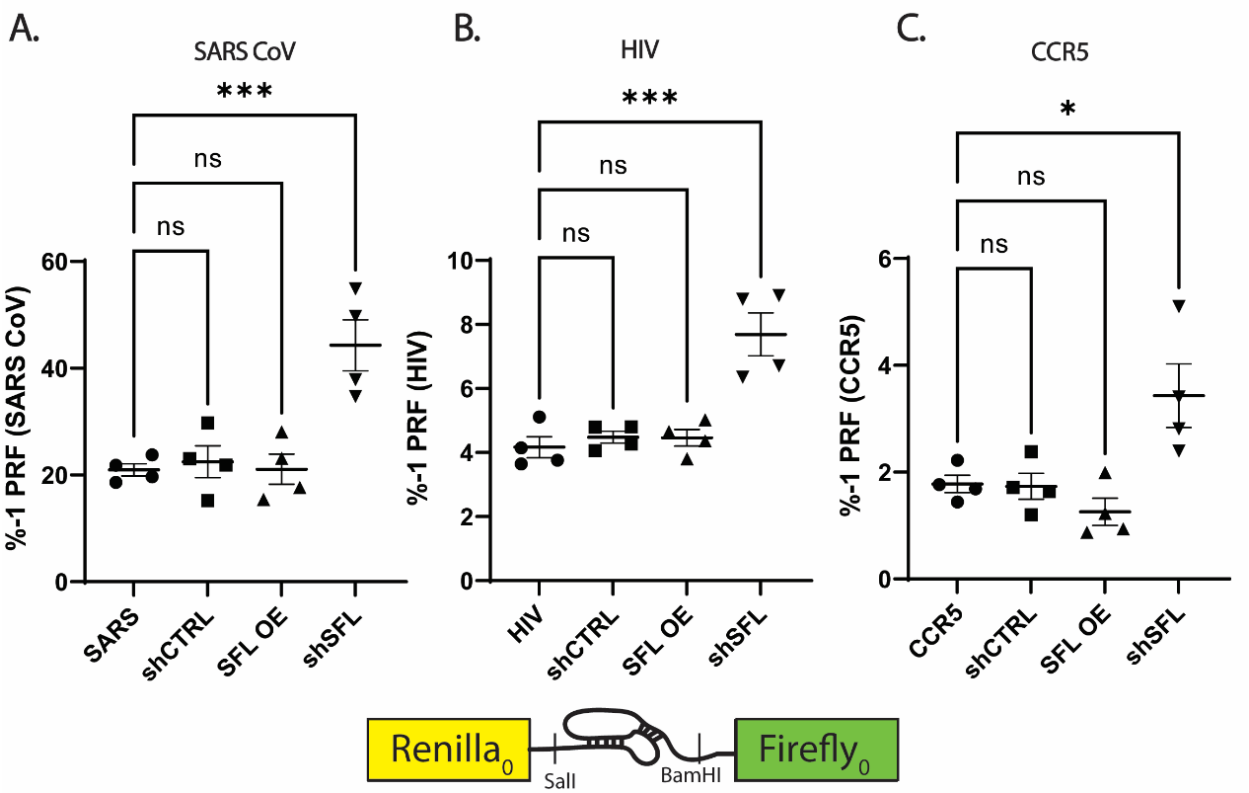
Supplemental figure S 1. Modifying SFL expression in the cell.

(A) RT-PCR analysis of SFL in a blank sample (Ø), and RNA extracted from E. coli, HeLa, HEK293T and U87 MG cells. SFL denotes the full-length Shiftless mRNA while SFLS denotes a shorter Shiftless splice isoform. (B) qRT-PCR analysis of SFL expression in HEK293T cells. (C) Sequence validation of SFL$^{-/-}$ HEK293T cells.



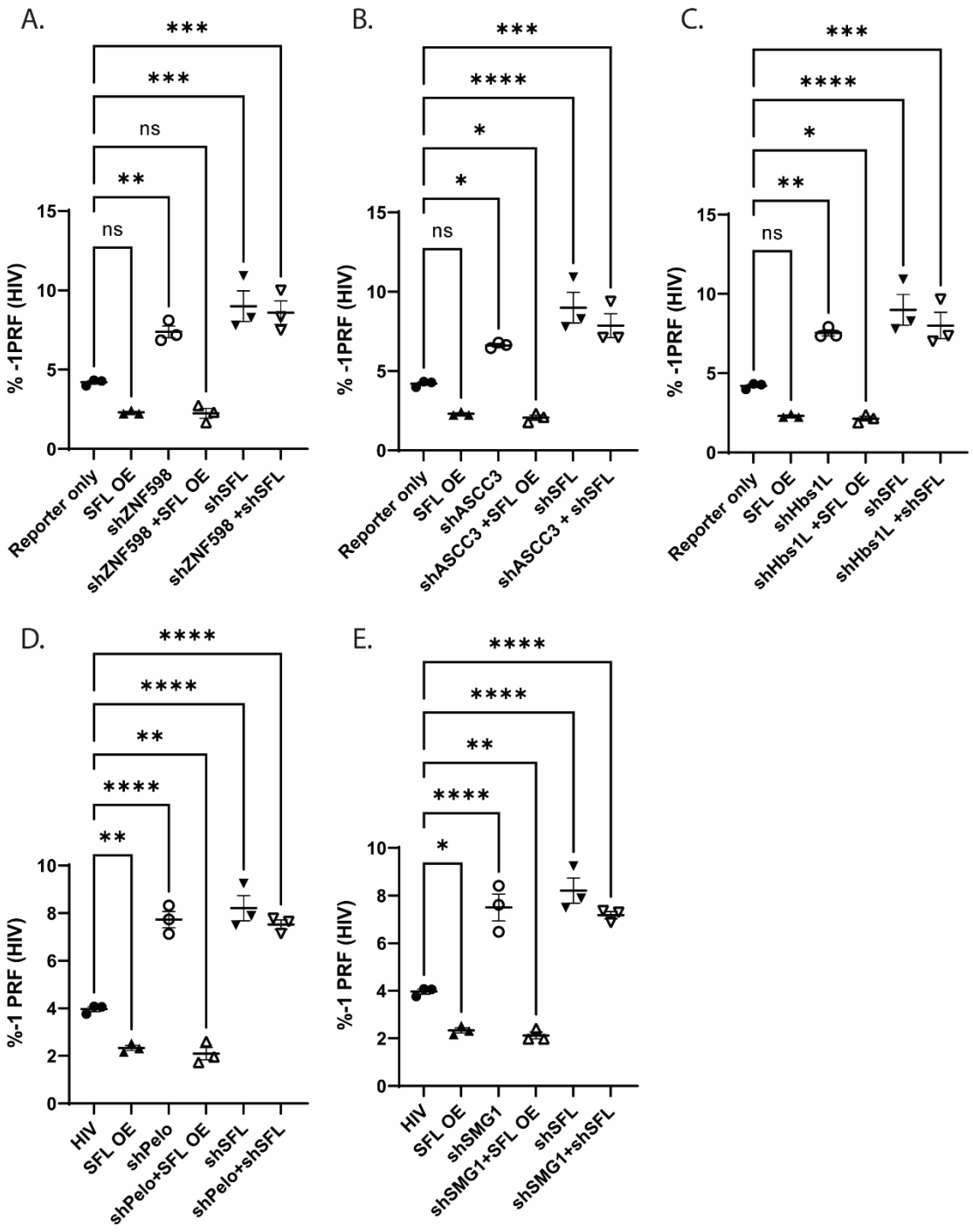
Supplemental figure S 2. SFL knockout does not alter cell growth

Growth curve of WT HEK293T cells (blue) and SFL$^{-/-}$ HEK293T cells (red). Error bars denote S.E. Each data point represents the mean of three biological replicates.



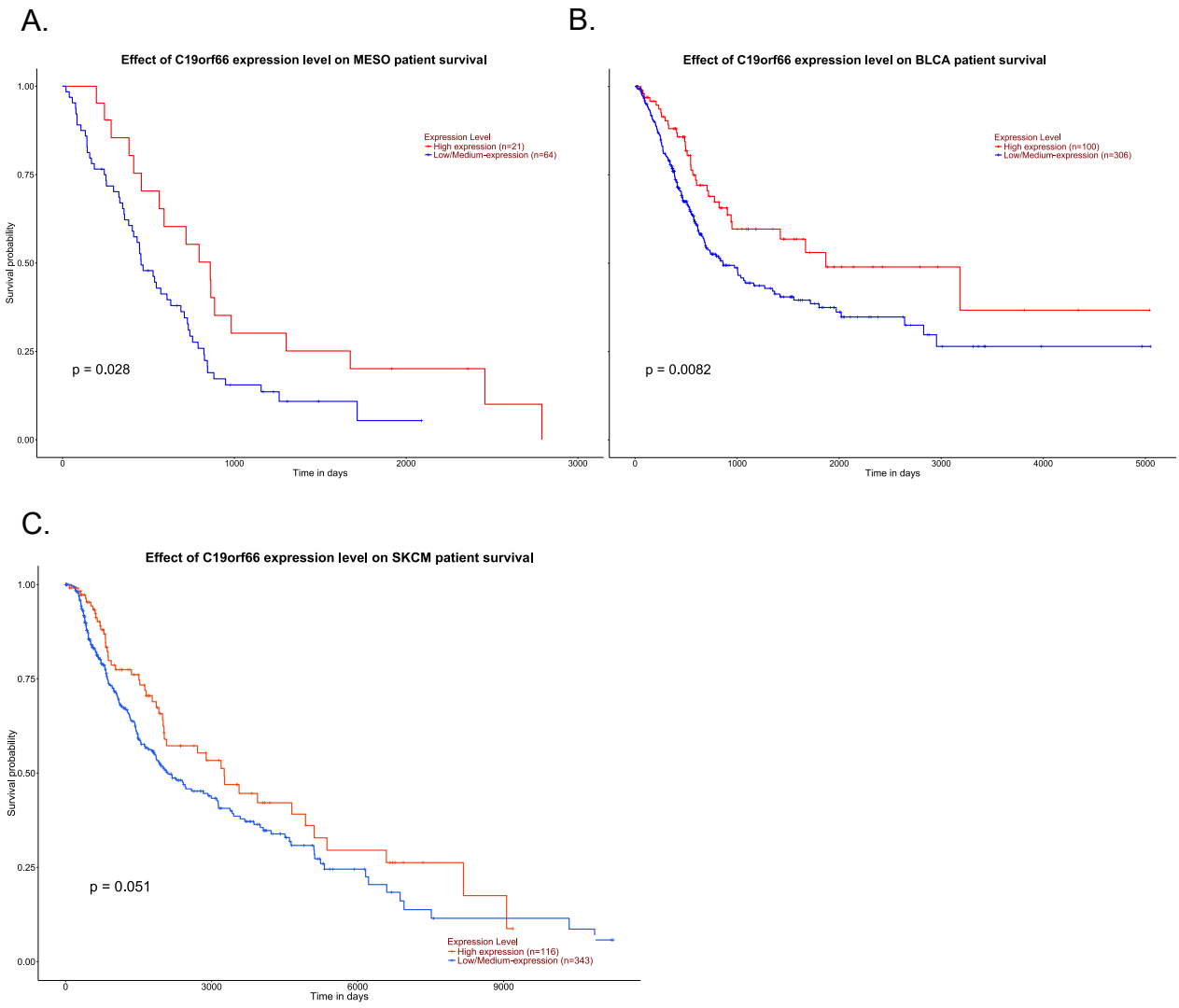
Supplemental figure S 3. SFL overexpression or knockdown alters -1 PRF.

Frameshift efficiency of three translational recoding elements measured using dual luciferase reporters in HEK293T cells over or under-expressing SFL. (A) SARS-CoV -1 PRF signal, (B) HIV-1 -1 PRF signal, (C) CCR5 -1 PRF signal. Error bars indicate S.E. Each data point represents one biological replicate.



Supplemental figure S 4. Influence SFL and RQC on the HIV-1 -1 PRF signal

Frameshift efficiency of the HIV-1 -1 PRF signal in HEK293T cells over or under-expressing SFL in combination with shRNA knockdown of (A) ZNF598, (B) ASCC3, (C) Hbs1L, (D) Pelota, (E) SMG1.



Supplemental figure S 5. Lower SFL expression decreases cancer patient survival.

Kaplan-Meyer survival plots comparing low and high-level Shiftless expression in (A) Lung mesothelioma, (B) Bladder urothelial carcinoma, and (C) Cutaneous melanoma. Graphs generated using the ULACAN platform.

Appendix 4: Supplementary tables for Chapter 4

	Normal			Cancer			Expression Change	
	Mean TPM	SD	N	Mean TPM	SD	N	Fold mean	% Change
Bladder (BLCA)	19.348	3.791429	9	12.09574	7.454498	432	0.625167	-37.4833
Breast	18.40608	6.584216	181	7.722593	5.202915	1253	0.419568	-58.0432
Colon	23.36609	9.238391	141	5.811777	4.246281	543	0.248727	-75.1273
Kidney (RCC)	20.41811	15.46794	28	8.765442	5.709652	615	0.429297	-57.0703
Liver	68.51388	24.48586	110	19.12129	11.57723	422	0.279086	-72.0914
Lung (MESO)	28.0186	10.74209	295	14.38403	6.311042	87	0.513374	-48.6626
Lung (LUAD)	28.0186	10.74209	295	8.378763	4.990853	598	0.299043	-70.0957
Prostate	30.51811	10.45481	100	6.631655	3.429536	556	0.217302	-78.2698
Skin (SKCM)	9.825529	3.824696	325	12.23792	7.540857	470	1.245522	24.55225
Average all							0.475232	-52.4768

Table 6. Shiftless expression in common cancers.

Average transcripts per million reads (TPM) of Shiftless in Normal or cancerous tissues. Data mined from Genome Browser[167].

Bibliography

1. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nature Reviews Molecular Cell Biology* vol. 19 (2018).
2. Ben-Shem, A. *et al.* The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science (1979)* **334**, (2011).
3. Lecompte, O., Ripp, R., Thierry, J. C., Moras, D. & Poch, O. Comparative analysis of ribosomal proteins in complete genomes: An example of reductive evolution at the domain scale. *Nucleic Acids Research* vol. 30 (2002).
4. Hinnebusch, A. G. The scanning mechanism of eukaryotic translation initiation. *Annual Review of Biochemistry* vol. 83 (2014).
5. Beringer, M. & Rodnina, M. v. The Ribosomal Peptidyl Transferase. *Molecular Cell* vol. 26 (2007).
6. Pisarev, A. v. *et al.* The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling. *Molecular Cell* **37**, (2010).
7. Schmidt, C. *et al.* Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. *Nucleic Acids Research* **44**, (2015).
8. Kurland, C. G. Translational accuracy and the fitness of bacteria. *Annual Review of Genetics* vol. 26 (1992).
9. Parker, J. Errors and alternatives in reading the universal genetic code. *Microbiological reviews* vol. 53 (1989).
10. Bock, L. v. *et al.* Thermodynamic control of −1 programmed ribosomal frameshifting. *Nature Communications* **10**, (2019).
11. Zhou, X., Liao, W. J., Liao, J. M., Liao, P. & Lu, H. Ribosomal proteins: Functions beyond the ribosome. *Journal of Molecular Cell Biology* **7**, (2015).
12. Brandman, O. & Hegde, R. S. Ribosome-associated protein quality control. *Nature Structural & Molecular Biology* **23**, 7–15 (2016).
13. Meaux, S. & van Hoof, A. Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly(A) tail in translation and mRNA decay. *RNA* **12**, (2006).
14. Simms, C. L., Hudson, B. H., Mosior, J. W., Rangwala, A. S. & Zaher, H. S. An Active Role for the Ribosome in Determining the Fate of Oxidized mRNA. *Cell Reports* **9**, (2014).
15. Letzring, D. P., Dean, K. M. & Grayhack, E. J. Control of translation efficiency in yeast by codon-anticodon interactions. *RNA* **16**, (2010).
16. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, (2006).
17. Matsuo, Y. *et al.* Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nature Communications* **8**, (2017).
18. Wu, C. C. C., Peterson, A., Zinshteyn, B., Regot, S. & Green, R. Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell* **182**, (2020).
19. Meydan, S. & Guydosh, N. R. Disome and Trisome Profiling Reveal Genome-wide Targets of Ribosome Quality Control. *Molecular Cell* **79**, (2020).
20. Garzia, A. *et al.* The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nature Communications* **8**, (2017).
21. D'Orazio, K. N. & Green, R. Ribosome states signal RNA quality control. *Molecular Cell* vol. 81 (2021).
22. Juszkiewicz, S., Speldewinde, S. H., Wan, L., Svejstrup, J. Q. & Hegde, R. S. The ASC-1 Complex Disassembles Collided Ribosomes. *Molecular Cell* **79**, (2020).
23. Olson, A. N. & Dinman, J. D. Two Ribosomes Are Better Than One. Sometimes. *Molecular Cell* **79**, (2020).
24. Joazeiro, C. A. P. Mechanisms and functions of ribosome-associated protein quality control. *Nature Reviews Molecular Cell Biology* vol. 20 (2019).
25. Goetz, A. E. & Wilkinson, M. Stress and the nonsense-mediated RNA decay pathway. *Cellular and Molecular Life Sciences* vol. 74 (2017).

26. Chamieh, H., Ballut, L., Bonneau, F. & le Hir, H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nature Structural and Molecular Biology* **15**, (2008).
27. Kim, V. N., Kataoka, N. & Dreyfuss, G. Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. *Science (1979)* **293**, (2001).
28. Kishor, A., Fritz, S. E. & Hogg, J. R. Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. *Wiley Interdisciplinary Reviews: RNA* vol. 10 (2019).
29. Karousis, E. D. & Mühlemann, O. Nonsense-mediated mRNA decay begins where translation ends. *Cold Spring Harbor Perspectives in Biology* **11**, (2019).
30. Kashima, I. *et al.* Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes and Development* **20**, (2006).
31. Lejeune, F., Li, X. & Maquat, L. E. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Molecular Cell* **12**, (2003).
32. Shoemaker, C. J., Eyler, D. E. & Green, R. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science (1979)* **330**, (2010).
33. Pisareva, V. P., Skabkin, M. A., Hellen, C. U. T., Pestova, T. v. & Pisarev, A. v. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO Journal* **30**, (2011).
34. Rodnina, M. v. *et al.* Survey and summary: Translational recoding: Canonical translation mechanisms reinterpreted. *Nucleic Acids Research* **48**, 1056–1067 (2020).
35. Firth, A. E., Wills, N. M., Gesteland, R. F. & Atkins, J. F. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic Acids Res* **39**, 6679–6691 (2011).
36. Huang, W. M. *et al.* A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. *Science* vol. 239 (1988).
37. Jacks, T. & Varmus, H. E. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science (1979)* **230**, (1985).
38. Pelham, H. R. B. Leaky UAG termination codon in tobacco mosaic virus RNA. *Nature* **272**, (1978).
39. Blanchet, S., Cornu, D., Argentini, M. & Namy, O. New insights into the incorporation of natural suppressor tRNAs at stop codons in Saccharomyces cerevisiae. *Nucleic Acids Research* **42**, (2014).
40. Lainé, S., Thouard, A., Komar, A. A. & Rossignol, J. M. Ribosome can resume the translation in both + 1 or - 1 frames after encountering an AGA cluster in Escherichia coli. *Gene* **412**, (2008).
41. Gurvich, O. L., Baranov, P. v., Gesteland, R. F. & Atkins, J. F. Expression levels influence ribosomal frameshifting at the tandem rare arginine codons AGG_AGG and AGA_AGA in Escherichia coli. *Journal of Bacteriology* **187**, (2005).
42. Rom, E. & Kahana, C. Polyamines regulate the expression of ornithine decarboxylase antizyme in vitro by inducing ribosomal frame-shifting. *Proc Natl Acad Sci U S A* **91**, 3959–3963 (1994).
43. Jacks, T., Townsley, K., Varmus, H. E. & Majors, J. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. *Proc Natl Acad Sci U S A* **84**, (1987).
44. Jacks, T. *et al.* Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**, 280–283 (1988).
45. dos Ramos, F., Carrasco, M., Doyle, T. & Brierley, I. Programmed -1 ribosomal frameshifting in the SARS coronavirus. in *Biochemical Society Transactions* vol. 32 (2004).
46. Brierley, I. *et al.* An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J* **6**, (1987).
47. Belew, A. T. *et al.* Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature* **512**, 265–269 (2014).
48. Clark, M. B. *et al.* Mammalian gene PEG10 expresses two reading frames by high efficiency -1 frameshifting in embryonic-associated tissues. *J Biol Chem* **282**, 37359–69 (2007).
49. Belew, A. T. & Dinman, J. D. Cell cycle control (and more) by programmed -1 ribosomal frameshifting: implications for disease and therapeutics. *Cell Cycle* **14**, 172–178 (2015).

50. Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. Signals for ribosomal frameshifting in the rous sarcoma virus gag-pol region. *Cell* **55**, (1988).
51. Plant, E. P. & Dinman, J. D. Torsional restraint: A new twist on frameshifting pseudoknots. *Nucleic Acids Research* **33**, (2005).
52. Puglisi, J. D., Wyatt, J. R. & Tinoco, I. A pseudoknotted RNA oligonucleotide. *Nature* **331**, (1988).
53. Noller, H. F., Yusupov, M. M., Yusupova, G. Z., Baucom, A. & Cate, J. H. D. Translocation of tRNA during protein synthesis. in *FEBS Letters* vol. 514 (2002).
54. Plant, E. P. *et al.* The 9-Å solution: How mRNA pseudoknots promote efficient programmed -1 ribosomal frameshifting. *RNA* vol. 9 (2003).
55. Plant, E. P., Rakauskaite, R., Taylor, D. R. & Dinman, J. D. Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. *J Virol* **84**, 4330–4340 (2010).
56. Dinman, J. D. & Wickner, R. B. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J Virol* **66**, 3669–3676 (1992).
57. Kendra, J. A. *et al.* Ablation of Programmed -1 Ribosomal Frameshifting in Venezuelan Equine Encephalitis Virus Results in Attenuated Neuropathogenicity. *J Virol* **91**, e01766-16 (2017).
58. Neupane, K. *et al.* Anti-Frameshifting Ligand Active against SARS Coronavirus-2 Is Resistant to Natural Mutations of the Frameshift-Stimulatory Pseudoknot. *Journal of Molecular Biology* **432**, (2020).
59. Munshi, S. *et al.* Identifying Inhibitors of −1 Programmed Ribosomal Frameshifting in a Broad Spectrum of Coronaviruses. *Viruses* **14**, (2022).
60. Park, S.-J., Kim, Y.-G. & Park, H.-J. Identification of RNA Pseudoknot-Binding Ligand That Inhibits the −1 Ribosomal Frameshifting of SARS-Coronavirus by Structure-Based Virtual Screening. *J Am Chem Soc* **133**, 10094–10100 (2011).
61. Ritchie, D. B., Soong, J., Sikkema, W. K. A. & Woodside, M. T. Anti-frameshifting Ligand Reduces the Conformational Plasticity of the SARS Virus Pseudoknot. *J Am Chem Soc* **136**, 2196–2199 (2014).
62. Wu, A. *et al.* Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* **27**, 325–328 (2020).
63. Xie, M. & Chen, Q. Insight into 2019 novel coronavirus — An updated interim review and lessons from SARS-CoV and MERS-CoV. *International Journal of Infectious Diseases* vol. 94 (2020).
64. Haniff, H. S. *et al.* Targeting the SARS-COV-2 RNA genome with small molecule binders and ribonuclease targeting chimera (RiboTAC) degraders. *ACS Central Science* **6**, 1713–1721 (2020).
65. Sun, Y. *et al.* Restriction of SARS-CoV-2 replication by targeting programmed -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **118**, (2021).
66. Fehr, A. R. & Perlman, S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* **1282**, 1–23 (2015).
67. Kelly, J. A., Woodside, M. T. & Dinman, J. D. Programmed −1 Ribosomal Frameshifting in coronaviruses: A therapeutic target. *Virology* **554**, 75–82 (2021).
68. Ritchie, D. B., Soong, J., Sikkema, W. K. A. & Woodside, M. T. Anti-frameshifting Ligand Reduces the Conformational Plasticity of the SARS Virus Pseudoknot. *J Am Chem Soc* **136**, 2196–2199 (2014).
69. Wang, X. *et al.* Regulation of HIV-1 Gag-Pol Expression by Shiftless, an Inhibitor of Programmed -1 Ribosomal Frameshifting. *Cell* **176**, 625-635.e14 (2019).
70. Dinman, J. D. Scaring Ribosomes Shiftless. *Biochemistry* **58**, 1831–1832 (2019).
71. Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type i interferon antiviral response. *Nature* **472**, (2011).
72. Balinsky, C. A. *et al.* IRAV (FLJ11286), an Interferon-Stimulated Gene with Antiviral Activity against Dengue Virus, Interacts with MOV10. *J Virol* **91**, e01606-16 (2017).
73. Kinast, V. *et al.* C19orf66 is an interferon-induced inhibitor of HCV replication that restricts formation of the viral replication organelle. *Journal of Hepatology* (2020) doi:https://doi.org/10.1016/j.jhep.2020.03.047.
74. Belew, A. T., Hepler, N. L., Jacobs, J. L. & Dinman, J. D. PRFdb: a database of computationally predicted eukaryotic programmed -1 ribosomal frameshift signals. *BMC Genomics* **9**, (2008).

75. de Messieres, M. *et al.* Single-molecule measurements of the CCR5 mRNA unfolding pathways. *Biophys J* **106**, 244–252 (2014).
76. Matsufuji, S. *et al.* Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell* **80**, 51–60 (1995).
77. Penn, W. D., Harrington, H. R., Schlebach, J. P. & Mukhopadhyay, S. Regulators of Viral Frameshifting: More Than RNA Influences Translation Events. *Annual Review of Virology* vol. 7 (2020).
78. Harrington, H. R. *et al.* Cotranslational folding stimulates programmed ribosomal frameshifting in the alphavirus structural polyprotein. *Journal of Biological Chemistry* vol. 295 (2020).
79. Ling, K. *et al.* Five-transmembrane domains appear sufficient for a G protein-coupled receptor: Functional five-transmembrane domain chemokine receptors. *Proc Natl Acad Sci U S A* **96**, (1999).
80. Kelly, J. A. *et al.* Structural and functional conservation of the programmed −1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2). *Journal of Biological Chemistry* (2020) doi:10.1074/jbc.AC120.013449.
81. Gorbalenya, A. E. *et al.* The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology* **5**, 536–544 (2020).
82. Brian, D. A. & Baric, R. S. Coronavirus genome structure and replication. *Current Topics in Microbiology and Immunology* vol. 287 (2005).
83. Dinman, J. D. Mechanisms and implications of programmed translational frameshifting. *WIREs RNA* **3**, 661–673 (2012).
84. Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E. & Baranov, P. v. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Research* **44**, (2016).
85. Plant, E. P., Sims, A. C., Baric, R. S., Dinman, J. D. & Taylor, D. R. Altering SARS coronavirus frameshift efficiency affects genomic and subgenomic RNA production. *Viruses* **5**, 279–294 (2013).
86. Plant, E. P. *et al.* A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol* **3**, e172–e172 (2005).
87. Brierley, I. & dos Ramos, F. J. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* **119**, 29–42 (2006).
88. Baranov, P. v. *et al.* Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, (2005).
89. Cho, C.-P., Lin, S.-C., Chou, M.-Y., Hsu, H.-T. & Chang, K.-Y. Regulation of Programmed Ribosomal Frameshifting by Co-Translational Refolding RNA Hairpins. *PLOS ONE* **8**, e62283- (2013).
90. Harger, J. W. & Dinman, J. D. An in vivo dual-luciferase assay system for studying translational recoding in the yeast Saccharomyces cerevisiae. *RNA* **9**, (2003).
91. Jacobs, J. L. & Dinman, J. D. Systematic analysis of bicistronic reporter assay data. *Nucleic Acids Res* **32**, e160–e160 (2004).
92. Chen, Y. L. & Pollack, L. Salt Dependence of A-Form RNA Duplexes: Structures and Implications. *Journal of Physical Chemistry B* **123**, (2019).
93. Ishimaru, D. *et al.* RNA dimerization plays a role in ribosomal frameshifting of the SARS coronavirus. *Nucleic Acids Res* **41**, 2594–2608 (2013).
94. Ahn, D. G. *et al.* Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses SARS coronavirus replication. *Antiviral Research* **91**, (2011).
95. Halma, M. T. J., Ritchie, D. B., Cappellano, T. R., Neupane, K. & Woodside, M. T. Complex dynamics under tension in a high-efficiency frameshift stimulatory structure. *Proc Natl Acad Sci U S A* **116**, 19500–19505 (2019).
96. Ritchie, D. B., Foster, D. A. N. & Woodside, M. T. Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci U S A* **109**, (2012).
97. Yang, L. *et al.* Solution scattering at the Life Science X-ray Scattering (LiX) beamline. *Journal of Synchrotron Radiation* **27**, (2020).

98. Karesh, W. B. *et al.* Ecology of zoonoses: Natural and unnatural histories. *The Lancet* vol. 380 (2012).
99. Mills, J. N., Gage, K. L. & Khan, A. S. Potential influence of climate change on vector-borne and zoonotic diseases: A review and proposed research plan. *Environmental Health Perspectives* **118**, (2010).
100. Saker, L., Lee, K., Cannito, B., Gilmore, A. & Campbell-Lendrum, D. *Globalization and Infectious Diseases: A review of linkages (WHO)*. *Social, Economic and Behavioural (SEB) Research* (2002).
101. Hadfield, J. *et al.* Twenty years of West Nile virus spread and evolution in the Americas visualized by Nextstrain. *PLoS Pathogens* vol. 15 (2019).
102. Holmes, E. C., Dudas, G., Rambaut, A. & Andersen, K. G. The evolution of Ebola virus: Insights from the 2013-2016 epidemic. *Nature* vol. 538 (2016).
103. Weaver, S. C., Charlier, C., Vasilakis, N. & Lecuit, M. Zika, Chikungunya, and Other Emerging Vector-Borne Viral Diseases. *Annual Review of Medicine* **69**, (2018).
104. Paules, C. I., Marston, H. D. & Fauci, A. S. Coronavirus Infections-More Than Just the Common Cold. *JAMA - Journal of the American Medical Association* vol. 323 (2020).
105. Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology* vol. 17 (2019).
106. Wong, A. C. P., Li, X., Lau, S. K. P. & Woo, P. C. Y. Global epidemiology of bat coronaviruses. *Viruses* vol. 11 (2019).
107. Volpato, G., Fontefrancesco, M. F., Gruppuso, P., Zocchi, D. M. & Pieroni, A. Baby pangolins on my plate: Possible lessons to learn from the COVID-19 pandemic. *Journal of Ethnobiology and Ethnomedicine* vol. 16 (2020).
108. Schubert, K. *et al.* Author Correction: SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation (Nature Structural & Molecular Biology, (2020), 27, 10, (959-966), 10.1038/s41594-020-0511-8). *Nature Structural and Molecular Biology* vol. 27 (2020).
109. Cornillez-Ty, C. T., Liao, L., Yates, J. R., Kuhn, P. & Buchmeier, M. J. Severe Acute Respiratory Syndrome Coronavirus Nonstructural Protein 2 Interacts with a Host Protein Complex Involved in Mitochondrial Biogenesis and Intracellular Signaling. *Journal of Virology* **83**, (2009).
110. Lei, J., Kusov, Y. & Hilgenfeld, R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Research* vol. 149 (2018).
111. Angelini, M. M., Akhlaghpour, M., Neuman, B. W. & Buchmeier, M. J. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *mBio* **4**, (2013).
112. Amor, S., Fernández Blanco, L. & Baker, D. Innate immunity during SARS-CoV-2: evasion strategies and activation trigger hypoxia and vascular damage. *Clinical and Experimental Immunology* **202**, (2020).
113. Bhatt, P. R. *et al.* Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science (1979)* **372**, 1306–1313 (2021).
114. Zhang, K. *et al.* Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. *Nature Structural and Molecular Biology* **28**, 747–754 (2021).
115. Omar, S. I. *et al.* Modeling the structure of the frameshift-stimulatory pseudoknot in SARS-CoV-2 reveals multiple possible conformers. *PLoS Computational Biology* **17**, (2021).
116. Brierley, I., Digard, P. & Inglis, S. C. Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell* **57**, (1989).
117. Somogyi, P., Jenner, A. J., Brierley, I. & Inglis, S. C. Ribosomal pausing during translation of an RNA pseudoknot. *Molecular and Cellular Biology* **13**, (1993).
118. Caliskan, N., Katunin, V. I., Belardinelli, R., Peske, F. & Rodnina, M. v. Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell* **157**, (2014).
119. Chen, J. *et al.* Dynamic pathways of -1 translational frameshifting. *Nature* **512**, 328–332 (2014).
120. Kim, H.-K. *et al.* A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. *Proc Natl Acad Sci U S A* **111**, 5538–5543 (2014).
121. Rodnina, M. v., Fricke, R., Kuhn, L. & Wintermeyer, W. Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome. *EMBO Journal* **14**, (1995).

122. Yan, S., Wen, J. der, Bustamante, C. & Tinoco, I. Ribosome Excursions during mRNA Translocation Mediate Broad Branching of Frameshift Pathways. *Cell* **160**, (2015).
123. Marczinke, B., Fisher, R., Vidakovic, M., Bloys, A. J. & Brierley, I. Secondary structure and mutational analysis of the ribosomal frameshift signal of Rous sarcoma virus. *Journal of Molecular Biology* **284**, (1998).
124. Chen, J., Petrov, A., Tsai, A., O'Leary, S. E. & Puglisi, J. D. Coordinated conformational and compositional dynamics drive ribosome translocation. *Nature Structural and Molecular Biology* **20**, (2013).
125. Moomau, C., Musalgaonkar, S., Khan, Y. A., Jones, J. E. & Dinman, J. D. Structural and functional characterization of programmed ribosomal frameshift signals in West Nile virus strains reveals high structural plasticity among cis-acting RNA elements. *Journal of Biological Chemistry* **291**, (2016).
126. Ritchie, D. B. *et al.* Conformational dynamics of the frameshift stimulatory structure in HIV-1. *RNA* **23**, (2017).
127. Yang, L. *et al.* Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A·C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. *J Am Chem Soc* **140**, (2018).
128. Chen, Y. T. *et al.* Coordination among tertiary base pairs results in an efficient frameshift-stimulating RNA pseudoknot. *Nucleic Acids Research* **45**, (2017).
129. Harcourt, J. *et al.* Severe acute respiratory syndrome coronavirus 2 from patient with coronavirus disease, United States. *Emerging Infectious Diseases* **26**, (2020).
130. Rangan, R. *et al.* RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: A first look. *RNA* **26**, (2020).
131. Plant, E. P. & Dinman, J. D. The role of programmed-1 ribosomal frameshifting in coronavirus propagation. *Frontiers in Bioscience* vol. 13 (2008).
132. Wacker, A. *et al.* Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. *Nucleic Acids Research* **48**, 12415–12435 (2020).
133. Akiyama, B. M. *et al.* Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science (1979)* **354**, (2016).
134. Huston, N. C. *et al.* Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Molecular Cell* **81**, (2021).
135. Iserman, C. *et al.* Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid. *Molecular Cell* **80**, (2020).
136. Lan, T. C. T. *et al.* Structure of the full SARS-CoV-2 RNA genome in infected cells. *bioRxiv* (2020) doi:10.1101/2020.06.29.178343.
137. Manfredonia, I. *et al.* Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Research* **48**, 12436–12452 (2020).
138. Rabl, J., Leibundgut, M., Ataide, S. F., Haag, A. & Ban, N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science (1979)* **331**, (2011).
139. Takyar, S., Hickerson, R. P. & Noller, H. F. mRNA helicase activity of the ribosome. *Cell* **120**, (2005).
140. Dinman, J. D. Shapeshifting RNAs guide innate immunity. *Journal of Biological Chemistry* **293**, (2018).
141. Dinman, J. D. Slippery ribosomes prefer shapeshifting mRNAs. *Proceedings of the National Academy of Sciences of the United States of America* vol. 116 (2019).
142. Calderon, B. M. & Conn, G. L. Human noncoding RNA 886 (nc886) adopts two structurally distinct conformers that are functionally opposing regulators of PKR. *RNA* **23**, (2017).
143. Ruijtenberg, S. *et al.* mRNA structural dynamics shape Argonaute-target interactions. *Nature Structural and Molecular Biology* **27**, (2020).
144. Frank, J. Time-resolved cryo-electron microscopy: Recent progress. *Journal of Structural Biology* **200**, (2017).
145. Zhang, X. *et al.* Mimicking Ribosomal Unfolding of RNA Pseudoknot in a Protein Channel. *J Am Chem Soc* **137**, (2015).
146. Brierley, I., Pennell, S. & Gilbert, R. J. C. Viral RNA pseudoknots: Versatile motifs in gene expression and replication. *Nature Reviews Microbiology* vol. 5 (2007).

147. Brierley, I. & Pennell, S. Structure and function of the stimulatory RNAs involved in programmed eukaryotic -1 ribosomal frameshifting. in *Cold Spring Harbor Symposia on Quantitative Biology* vol. 66 (2001).
148. Ryder, S. P., Morgan, B. R., Coskun, P., Antkowiak, K. & Massi, F. Analysis of Emerging Variants in Structured Regions of the SARS-CoV-2 Genome. *Evolutionary Bioinformatics* **17**, (2021).
149. Su, M. C., Chang, C. te, Chu, C. H., Tsai, C. H. & Chang, K. Y. An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Research* **33**, (2005).
150. Andrews, R. J. *et al.* A map of the SARS-CoV-2 RNA structurome. *NAR Genomics and Bioinformatics* **3**, (2021).
151. Dever, T. E., Dinman, J. D. & Green, R. Translation elongation and recoding in eukaryotes. *Cold Spring Harbor Perspectives in Biology* **10**, (2018).
152. Neuman, B. W., Angelini, M. M. & Buchmeier, M. J. Does form meet function in the coronavirus replicative organelle? *Trends in Microbiology* vol. 22 (2014).
153. Galloux, M. *et al.* Minimal elements required for the formation of respiratory syncytial virus cytoplasmic inclusion bodies in vivo and in vitro. *mBio* **11**, (2020).
154. Zhou, Y., Su, J. M., Samuel, C. E. & Ma, D. Measles Virus Forms Inclusion Bodies with Properties of Liquid Organelles. *Journal of Virology* **93**, (2019).
155. Maillard, P. v, Veen, A. G., Poirier, E. Z. & Reis e Sousa, C. Slicing and dicing viruses: antiviral RNA interference in mammals . *The EMBO Journal* **38**, (2019).
156. Karacostas, V., Wolffe, E. J., Nagashima, K., Gonda, M. A. & Moss, B. Overexpression of the HIV-1 Gag-Pol Polyprotein Results in Intracellular Activation of HIV-1 Protease and Inhibition of Assembly and Budding of Virus-like Particles. *Virology* **193**, 661–671 (1993).
157. Dinman, J. D., Ruiz-Echevarria, M. J., Czaplinski, K. & Peltz, S. W. Peptidyl-transferase inhibitors have antiviral properties by altering programmed -1 ribosomal frameshifting efficiencies: Development of model systems. *Proc Natl Acad Sci U S A* **94**, (1997).
158. Goss Kinzy, T. *et al.* New targets for antivirals: The ribosomal A-site and the factors that interact with it. *Virology* **300**, (2002).
159. Chen, Y. *et al.* A drug screening toolkit based on the –1 ribosomal frameshifting of SARS-CoV-2. *Heliyon* **6**, (2020).
160. Hu, H. T., Cho, C. P., Lin, Y. H. & Chang, K. Y. A general strategy to inhibiting viral -1 frameshifting based on upstream attenuation duplex formation. *Nucleic Acids Research* **44**, (2016).
161. Suzuki, Y. *et al.* Characterization of RyDEN (C19orf66) as an Interferon-Stimulated Cellular Inhibitor against Dengue Virus Replication. *PLoS Pathog* **12**, e1005357–e1005357 (2016).
162. Napthine, S., Hill, C. H., Nugent, H. C. M. & Brierley, I. Modulation of viral programmed ribosomal frameshifting and stop codon readthrough by the host restriction factor shiftless. *Viruses* **13**, (2021).
163. Ikeuchi, K., Izawa, T. & Inada, T. Recent progress on the molecular mechanism of quality controls induced by ribosome stalling. *Frontiers in Genetics* vol. 10 (2019).
164. Strauss, E. G., Rice, C. M. & Strauss, J. H. Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. *Proc Natl Acad Sci U S A* **80**, (1983).
165. Hickey, K. L. *et al.* GIGYF2 and 4EHP Inhibit Translation Initiation of Defective Messenger RNAs to Assist Ribosome-Associated Quality Control. *Molecular Cell* **79**, (2020).
166. Amrani, N. *et al.* A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432**, (2004).
167. Lee, B. T. *et al.* The UCSC Genome Browser database: 2022 update. *Nucleic Acids Res* **50**, (2022).
168. Howlader N *et al.* SEER Cancer Statistics Review, 1975-2017, National Cancer Institute. Bethesda, MD. *Https://Seer.Cancer.Gov/Csr/1975_2017/* (2020).
169. Rácz, G. A., Nagy, N., Tóvári, J., Apáti, Á. & Vértessy, B. G. Identification of new reference genes with stable expression patterns for gene expression studies using human cancer and normal cell lines. *Scientific Reports* **11**, (2021).
170. Jo, J. *et al.* Conventionally used reference genes are not outstanding for normalization of gene expression in human cancer research. *BMC Bioinformatics* **20**, (2019).

171. Park, S. B., Seronello, S., Mayer, W. & Ojcius, D. M. Hepatitis C virus frameshift/alternate reading frame protein suppresses interferon responses mediated by pattern recognition receptor retinoic-acid-inducible gene-I. *PLoS ONE* **11**, (2016).
172. Firth, A. E., Blitvich, B. J., Wills, N. M., Miller, C. L. & Atkins, J. F. Evidence for ribosomal frameshifting and a novel overlapping gene in the genomes of insect-specific flaviviruses. *Virology* **399**, (2010).
173. Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *Journal of General Virology* vol. 93 (2012).
174. Brocchieri, L. & Karlin, S. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Research* **33**, (2005).
175. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281–2308 (2013).
176. Loughran, G., Howard, M. T., Firth, A. E. & Atkins, J. F. Avoidance of reporter assay distortions from fused dual reporters. *RNA* **23**, (2017).
177. Chandrashekar, D. S. *et al.* UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia (United States)* **19**, (2017).
178. Hoffmann, M. *et al.* Nafamostat Mesylate Blocks Activation of SARS-CoV-2: New Treatment Option for COVID-19. *Antimicrobial Agents and Chemotherapy* **64**, e00754-20 (2020).
179. Jeon, S. *et al.* Identification of Antiviral Drug Candidates against SARS-CoV-2 from FDA-Approved Drugs. *Antimicrobial Agents and Chemotherapy* **64**, e00819-20 (2020).
180. Munshi, S. *et al.* Small-molecule ligands can inhibit −1 programmed ribosomal frameshifting in a broad spectrum of coronaviruses. *bioRxiv* (2021).
181. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* vol. 22 (2017).
182. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall* **1**, 33–46 (2017).
183. Snyder, J. E. *et al.* Functional Characterization of the Alphavirus TF Protein. *Journal of Virology* **87**, (2013).
184. Chung, B. Y. W., Firth, A. E. & Atkins, J. F. Frameshifting in Alphaviruses: A Diversity of 3′ Stimulatory Structures. *Journal of Molecular Biology* **397**, (2010).
185. Aoki, K. *et al.* LARP1 specifically recognizes the 3′ terminus of poly(A) mRNA. *FEBS Letters* **587**, (2013).
186. Lopinski, J. D., Dinman, J. D. & Bruenn, J. A. Kinetics of Ribosomal Pausing during Programmed −1 Translational Frameshifting. *Molecular and Cellular Biology* **20**, (2000).
187. Meydan, S. & Guydosh, N. R. A cellular handbook for collided ribosomes: surveillance pathways and collision types. *Current Genetics* vol. 67 (2021).
188. Gorgoni, B. & Gray, N. K. The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: A developmental perspective. *Briefings in Functional Genomics and Proteomics* vol. 3 (2004).
189. Tarun, S. Z. & Sachs, A. B. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO Journal* **15**, (1996).
190. Vicens, Q., Kieft, J. S. & Rissland, O. S. Revisiting the Closed-Loop Model and the Nature of mRNA 5′–3′ Communication. *Molecular Cell* vol. 72 (2018).
191. Álvarez, E., Castelló, A., Menéndez-Arias, L. & Carrasco, L. HIV protease cleaves poly(A)-binding protein. *Biochemical Journal* **396**, (2006).
192. Burgui, I., Aragón, T., Ortín, J. & Nieto, A. PABP1 and eIF4GI associate with influenza virus NS1 protein in viral mRNA translation initiation complexes. *Journal of General Virology* **84**, (2003).
193. Polacek, C., Friebe, P. & Harris, E. Poly(A)-binding protein binds to the non-polyadenylated 3′ untranslated region of dengue virus and modulates translation efficiency. *Journal of General Virology* **90**, (2009).
194. Bayfield, M. A., Yang, R. & Maraia, R. J. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* vol. 1799 (2010).
195. Nawaz, A., Shilikbay, T., Skariah, G. & Ceman, S. Unwinding the roles of RNA helicase MOV10. *Wiley Interdisciplinary Reviews: RNA* vol. 13 (2022).

196. Kenny, P. J. *et al.* MOV10 and FMRP Regulate AGO2 Association with MicroRNA Recognition Elements. *Cell Reports* **9**, (2014).
197. Kute, P. M., Ramakrishna, S., Neelagandan, N., Chattarji, S. & Muddashetty, R. S. NMDAR mediated translation at the synapse is regulated by MOV10 and FMRP. *Molecular Brain* **12**, (2019).
198. Sinha, N. K. *et al.* EDF1 coordinates cellular responses to ribosome collisions. *Elife* **9**, (2020).
199. Pakos-Zebrucka, K. *et al.* The integrated stress response. *EMBO Rep* **17**, (2016).
200. Lu, P. D., Harding, H. P. & Ron, D. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *Journal of Cell Biology* **167**, (2004).
201. Ron, D. Translational control in the endoplasmic reticulum stress response. *Journal of Clinical Investigation* **110**, (2002).