20TH
OPEN ACCESS
ANNIVERSARY

OXFORD

# The differential effect of SARS-CoV-2 NSP1 on mRNA translation and stability reveals new insights linking ribosome recruitment, codon usage, and virus evolution

Juan José Berlanga[1,†], Tania Matamoros[1,†], Miguel Rodríguez Pulido[1], Margarita Sáiz[1],
Mercedes Núñez Bayón [1], René Toribio[2], Iván Ventoso [1,*]

[1]Centro de Biología Molecular Severo Ochoa (CSIC-UAM) and Departamento de Biología Molecular, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain
[2]Centro de Biotecnología y Genómica de Plantas (CBGP), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Campus de Montegancedo, Pozuelo de Alarcón, 28223 Madrid, Spain

[*]To whom correspondence should be addressed. Email: iventoso@cbm.csic.es
[†]The first two authors should be regarded as Joint First Authors.

## Abstract

The nonstructural protein 1 (NSP1) of SARS-CoV-2 blocks the messenger RNA (mRNA) entry channel of the 40S ribosomal subunit, causing inhibition of translation initiation and subsequent degradation of host mRNAs. However, target mRNA specificity and how viral mRNAs escape NSP1-mediated degradation have not been clarified to date. Here we found that NSP1 acts as a translational switch capable of blocking or enhancing translation depending on how preinitiation complex, 43S-PIC, is recruited to the mRNA, whereas NSP1-mediated mRNA degradation mostly depends on codon usage bias. Thus, fast-translating mRNAs with optimal codon usage for human cells that preferentially recruit 43S-PIC by threading showed a dramatic sensitivity to NSP1. Translation of SARS-CoV-2 mRNAs escapes NSP1-mediated inhibition by a proper combination of suboptimal codon usage and slotting-prone 5′ UTR. Thus, the prevalence of nonoptimal codons found in SARS-CoV-2 and other coronavirus genomes is favored by the distinctive effect that NSP1 plays on translation and mRNA stability.

## Graphical abstract
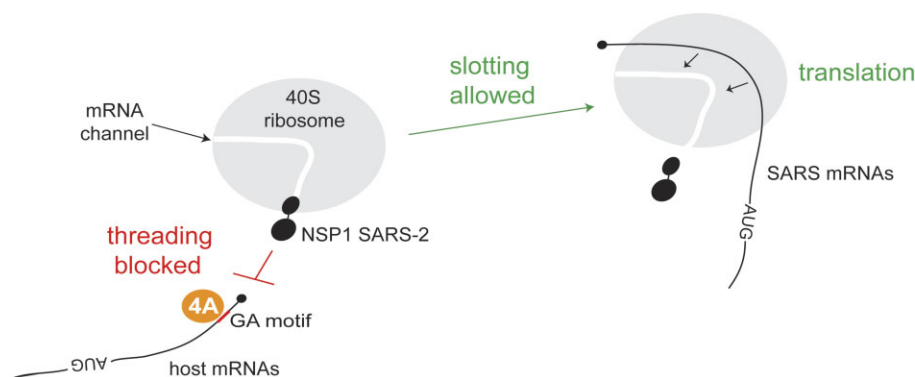


## Introduction

Translation in eukaryotes not only regulates protein outputs but also messenger RNA (mRNA) stability. Cap-dependent translation initiates with the recruitment of the 40S ribosomal subunit to the 5′ cap of mRNA, and the subsequent scanning of 43S-PIC through the 5′ UTR to locate the initiation codon [1–3]. Binding of the eIF4F complex (eIF4E + eIF4G + eIF4A) to the 5′ cap promotes local unwinding of the mRNA secondary structure that facilitates attachment of 43S-PIC(4). Features of the 5′ UTR including length, nucleotide composition and the presence of stable secondary structures can greatly influence translation initiation rates of particular mR-

NAs [3–6]. Thus, initiation on mRNAs with a complex 5′ UTR is more dependent on the helicase activities of eIF4A and DDX3, which are associated to 43S-PIC scanning, whereas translation of mRNAs with an unstructured 5′ UTR usually shows fewer requirements for initiation [7–11]. Although structural models of 43S-PIC bound to short mRNAs are available (48S-PIC), the way mRNA is inserted into the 40S subunit for codon inspection is still a matter of debate [1]. Threading into the mRNA entry channel or lateral slotting on the groove of the 40S subunit have both been proposed [12–15]. Threading and slotting would show different requirements for 5′ UTR length and topology, raising the possibility

that both mechanisms may operate depending on the 5′ UTR features found in a particular mRNA.

Translation elongation is increasingly recognized for its role in regulating mRNA stability and protein quality control, but it can also influence translation initiation by a yet unknown mechanism [16–19]. mRNAs with a synonymous codon composition adapted to aa-transfer RNA (tRNA) availability in a given cell type or organism tend to be more stable, and they are translated faster [19]. Since optimal codons usually end in C or G, mRNAs with optimal codon usage also show a higher G + C content. Although this may result in a more stable secondary structure that could hinder ribosome advance, mRNAs with optimal codons show an effective ribosome flow that minimizes ribosome stalling-mediated mRNA destabilization [17, 19].

In stark contrast, many viral genomes including SARS-CoV-2 show a low G + C content and a suboptimal codon usage for human cells, mainly shaped by a persistent C-to-U mutational bias [20–23]. Despite this, SARS-CoV-2 mRNAs are efficiently translated in a context of generalized translation shut-off and host mRNA degradation induced by the viral nonstructural protein 1 (NSP1) [24–28]. NSP1 of SARS-CoV-2 binds with high affinity to the 40S ribosomal subunit at the mRNA entry channel, thus preventing the penetration of mRNA through this region that results in a generalized translational blockade [29–31]. In addition, NSP1 binding to 40S subunit promotes host mRNA degradation that also requires the specific participation of the N-terminal domain of NSP1 [26–28, 32]. While the C-terminal domain of NSP1 is inserted into the mRNA entry channel, the remaining N-terminal domain stabilizes the NSP1-40S complex by interacting with surrounding ribosomal proteins (e.g. RPS3) and some initiation factors (e.g. eIF3G) [29, 33–37]. More recently, a dynamic binding of N-terminal domain of NSP1 to the decoding center of 40S subunit has been described [24]. Although most attempts failed to detect NSP1-associated nuclease activity, the N-terminal domain was shown to be essential for NSP1-mediated mRNA degradation, also requiring the presence of eIF3G and eIF4F in the PIC [34, 36]. Another unresolved question is how SARS-Cov-2 mRNAs escape the inhibitory activity of NSP1. The presence of a cap-proximal stem-loop (SL1) in the 5′ leader (5′ L) of genomic and subgenomic SARS-CoV-2 mRNAs has been reported as essential to escape the inhibitory effect of NSP1 [25, 34, 35, 38–41]. Although a proven mechanistic model for this resistance is still lacking, some authors have proposed that binding of SL1 to N-terminal domain could displace NSP1 from the mRNA entry channel, allowing the attachment of viral 5′ L to 43S-PIC [42]. However, a direct interaction of NSP1 with 5′ L was not observed in other reports [34].

Here, we have analyzed the contribution of the 5′ UTR and codon usage bias to the sensitivity of mRNAs to NSP1 in human cells. Our results reveal new insights and links between 43S-PIC recruitment to mRNA, codon usage bias and viral evolution.

## Materials and methods

### Plasmids and recombinant DNA

Plasmids encoding SARS-CoV-2 NSP1 variants were generated by exchanging the EGFP coding sequence of pEGFP-N1 plasmid (Clontech) for the corresponding nsp1 genes. For this, nsp1 coding sequences were amplified by polymerase chain reaction (PCR) from pDONR207 SARS-CoV-2 nsp1 (wild type (WT), Addgene #141255), pDONR207 SARS-CoV-2 nsp1 Delta RC (M1, K164A/H165A, Addgene #164522), and pDONR207 SARS-CoV-2 nsp1 R124A/K125A (M2, Addgene #164523) plasmids using the primers 5-EX-NSP1 and 3-NB-NSP1(Supplementary Table S3) and cloned into pEGFP-N1 digested with *EcoR*I and *Not*I enzymes. Some of the reporter plasmids used here have been described previously: pFFL (firefly luc), pRL (renilla luc), p242-RL, p427-RL, p657-RL, pG-less-RL, pG-low-FFL, pSt20-RL/FFL, and pSt30-RL/FFL [14]. pEGFP-N1 and pDsRed-N1 were purchased from Clontech and pbeta-Gal expresses the lacZ gene under the CMV-IE promoter. The new reporter plasmids generated in this study were:

-pNanoLuc. The coding sequence of NanoLuc was amplified by PCR using the primers 5-NanoL and 3-NanoL (Supplementary Table S3) from pNL1.1.CMV plasmid (Promega). The resulting fragment was digested with *BamH*I and *Not*I enzymes and cloned into pEGFP-N1 digested with the same enzymes, resulting in the replacement of EGFP by NanoLuc.

-pACTB. Human actin beta coding sequence (ACTB CDS) was amplified by reverse transcriptase-polymerase chain reaction (RT-PCR) using the primers 5-N-hACTB and 3-X-hACTB (Supplementary Table S3) and cloned into a plasmid bearing a myc epitope using *Not*I and *Xba*I enzymes. Then, the *BamH*I-*Not*I fragment was subcloned into pEGFP-N1 digested with the same enzymes. The resulting plasmid encodes a myc-tagged version of ACTB.

-pN SARS-2. The coding sequence of N gene with the natural codon usage of SARS-CoV-2 from pSARS-CoV-2 (N) plasmid (Addgene #153201) was digested with *BamH*I and *Not*I enzymes and cloned into pEGFP-N1 using the same enzymes.

-pG-low-RL/EGFP. The *Nde*I-*BamH*I fragment from pG-low-FFL plasmid was subcloned into pRL and pEGFP-N1 plasmids digested with the same enzymes.

All variants of pG-less-RL, p5′ L-N, p5′L-RL, p5′L(Gs)-RL, pTISU-EGFP, pSlot-RL, and the recoded versions of pEGFP (pEGFP-1, EGFP-2), pRL (pRL-1, pRL-2 and pRL-3) and N (p5′L-N1, pN1) were obtained from GenScript.

-pTISU-RL. A PCR fragment including TISU element from pTISU-EGFP plasmid was cloned into pRL using *Nde*I and *Sac*I enzymes.

All constructs generated in this work were verified by sequencing.

### Cell culture and transfection

HEK293T, HeLa, MEFs, Calu-3, and Vero E6 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and antibiotics. Cells were authenticated by microscopic examination and the absence of mycoplasma contamination was confirmed every 2 months of culture. For plasmid transfection, cells growing in a P24 plate were transfected with 1 µg of plasmid (or a combination of plasmids) and 2 µl of Turbofect$^{TM}$ (HEK293, MEFs and HeLa) or 2 µl of Lipofectamine$^{TM}$ 2000 transfection reagents (Calu-3 and VeroE6) according to the manufacturer's recommendation (Thermo Fisher and Invitrogen, respectively). For co-transfections, we combined up to 0.75 µg of pNSP1 plasmids with a mix of up to three reporter plasmids (0.15–0.2 µg each). We always included pEGFP-N1 plasmid as a readout

of transfection efficiency and sensitivity to NSP1. For titration with increasing amounts of pNSP1-WT or M2, the transfection mixtures were filled in with pNSP1-M1 (null) plasmid up to 0.5 μg. Transfected cells were incubated 24 h at 37°C and extracts were prepared for reporter activity measurement, western blot, or RNA extraction. Oligonucleotide transfection was carried out as described before [14]. For mRNA transfection, HEK293T cells growing in a P24 plate were transfected with 1 μg of capped *in vitro* synthesized mRNA (see below) and 3 μl of lipofectamine™ 2000 (Invitrogen) in Opti-MEM reduced serum medium, and luciferase activity was measured 6 h later.

## SARS-CoV-2 infection
Vero E6 cells growing in P24 wells were infected with SARS-CoV-2 (Wuhan-1) at a multiplicity of infection (moi) = 2 pfu/cell in 0.5 ml Opti-MEM (Gibco) under biosafety level 3 conditions. After 1 h adsortion, the inoculum was removed and cells were transfected with a combination of pEGFP, pRL-3- and pFFL plasmids (0.25 μg each) using Lipofectamine™ 2000 (Invitrogen). Cell extracts were prepared at 16- or 24-h post-infection (hpi) for luciferase activity measurement and western blot.

## Measurement of reporter gene expression
Enzymatic measurements were carried out in cell extracts prepared in passive lysis buffer (Biotium). Firefly, renilla, and NanoLuc luciferases were measured in a Berthold luminometer using D-luciferin (Promega), collantrazine (Biotium), and Nano-Glo (Promega), respectively. To avoid luminometer saturation, extracts were diluted to keep luciferase activities below $2 \times 10^6$ relative ligth units (RLUs). β-Galactosidase activity was measured using o-nitrophenyl-beta-D-galactoside (ONPG) as substrate (Promega). EGFP and dsRed accumulation was measured by western blot (see below) and by fluorescence emission of extracts using a GloMax microplate reader (Promega). In both cases, a standard curve was included for relative quantification in the linear range of measurement.

## Western blot
Western blots were carried out as described previously [14, 26] using the following primary antibodies: anti-EGFP (11 814 460 001, Roche), anti-dsRed (a gift from José María Requena, CBM, Madrid.), anti-FFLuc (a gift from Araceli del Arco, CBM, Madrid), anti-N SARS-CoV-2 (MA5-38034, Invitrogen), anti-NSP1 SARS-CoV-2 (STJ11103222, St. John's lab), anti-myc (05-724MG, Merck-Millipore), anti-eIF4A (STJ2724, St. John's lab), and anti-eIF3G (STJ23512, St. John's lab), anti-ABCE1 (STJ11102433, St. John's lab), and anti-eIF2α (sc-11386, Santa Cruz Biotechnology). Blots were developed with Luminata™/Immobilon® Crescendo Western HRP Substrate (Millipore) and bands were quantified by densitometry using the linear range of the signal.

## RNA extraction and quantitative RT-PCR
Total RNA was extracted from cells growing in P24 or P12 wells using the RNeasy plus kit (QIAGEN). The resulting RNA was digested with 2 U of DNAse I for 30 min at 37°C and further purified using RNA Clean & Concentrator 5 columns (Zymo Research). About 1.5 μg of total RNA were retrotranscribed with hexamers using Superscript VI retro-

transcriptase (Invitrogen). Real-time PCR (qPCR) amplification was carried out with GoTaq qPCR mix (Promega) in a Bio-Rad CFX-384 thermocycler using a 1/200 dilution of corresponding complementary DNAs. Relative quantification was performed using the Relative Expression Software Tool (REST) [43] and the following formula:

$$\text{ratio} = \frac{(Etarget)^{\Delta\text{CPtarget}(\text{control}-\text{sample})}}{(Eref)^{\Delta\text{CPPref}(\text{control}-\text{sample})}}$$

18S RNA was used as reference gene. For experiments analyzing the effect of NSP1 WT on reporter mRNA accumulation, samples expressing NSP1 M1 (null mutant) were used as control. Experimental details are included in MIQE cheklist (http://rdml.org/miqe.html).

## Measurement of translation efficiency of mRNAs
Translation efficiencies (TE) of mRNAs were calculated by polysome profiling of HEK293T extracts as described previously [14]. Briefly, one P100 plate of HEK293T cells were transfected with a combination of up to four different reporter-expressing plasmids. After 24 h, monolayers were lysed in polysome buffer (Tris–HCl 30 mM, pH 7.5, 100 mM KCl, 5 mM MgCl$_2$, 1 mM dithiothreitol (DTT), 1% Triton X-100, and 50 μg/ml cycloheximide) for 15 min on ice. Cell lysates were clarified by low-speed centrifugation and spun on a 10%–40% sucrose gradient at 35 000 rpm for 3 h at 4°C in an SW40 rotor. Nine to ten fractions were collected using an ISCO fractionator coupled to a ultraviolet recorder, followed by TRIzol™ reagent (Thermo Fisher) extraction and ethanol precipitation. The resulting RNA pellets were resuspended in water and digested with 2 U of DNAse I for 30 min at 37°C followed by purification using RNA Clean & Concentrator 5 columns (Zymo Research). mRNA levels in every fraction were quantified by RT-qPCR as described above. TE for every mRNA was calculated as the ratio between the relative amount found in polysomal region (fractions from 7 to 10) and monosomal (80S) and submonosomal regions (fractions from 2 to 5).

## Codon usage, tRNA adaptation index and codon stability coefficient calculations
Codon usage tables of human mRNAs were downloaded from HIVE-CUT repository [44]. Coronavirus coding sequences were extracted from reference genomes deposited in GenBank (SARS-CoV-2: NC_045512.2, SARS-CoV1: NC_004718, MERS: NC_019843.3, HCoV-OC43: NC_006213, HCoV-229E: NC_002645, HCoV-NL63: NC_005831, HCoV-HKU1: NC_006 577). Nucleotide composition, codon usage per thousand and codon adaptive index (CAI) of coronavirus genomes and reporter genes were calculated using the tools hosted in the CAIcal server [45]. Transcript-level average tRNA adaptation index (tAIg) and codon stability coefficient (CSCg) were calculated using the following formula:

$$\text{tAIg} = \Sigma \left( Codon.freqAAA * tAIAAA \right) + ...$$
$$+ \left( Codon.freqTTT * tAITTT \right)$$

$$\text{CSCg} = \Sigma \left( Codon.freqAAA * CSCAAA \right) + ...$$
$$+ \left( Codon.freqTTT * CSCTTT \right)$$

We used tAI and CSC values for all 61 nonstop codons previously calculated using experimental measurements of tRNA levels and mRNA half-lives in HEK293T cells, respectively [46, 47].

### Recoding of reporter genes

Codon usage optimization and deoptimization of EGFP, Renilla luciferase (RL), and N SARS-CoV-2 coding sequences were carried out using the harmonization tool described in http://www.codons.org/ [48]. We constructed a codon usage table of our HEK293T cells that included all mRNAs whose expression was detected by RNAseq. [14]. This table was used as reference for the recoding process.

### Purification of recombinant NSP1

NSP1 was cloned and expressed as fusion protein with LSL (mushroom lectin) in *Escherichia coli* rosetta as described previously [49]. Bacterial extract was incubated with Sepharose 4B (Merck) for 1 h followed by extensive washing in 25 mM Tris–HCl, pH 7.5, and 100 mM KCl for 3 h at 4°C. Then, NSP1 was released from the column by digestion with Tobacco Etch Virus (TEV) protease (NEB) overnight at 4°C. Finally, the His-tagged TEV protease was eliminated by passing the extract through a Ni Spin column (NEB), resulting in a near homogeneous preparation of NSP1 when analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and Coomassie blue staining.

### mRNA synthesis and *in vitro* translation

mRNAs were *in vitro* synthesized using HiScribe T7 Quick (NEB) and 5′ capped using the vaccinia capping system (NEB) as described previously [14]. When indicated, a poly(A) tail of 25 nt. was included in the reverse primer during the preparation of the DNA template for transcription. *In vitro* translation was programmed in 10 μl of Rabbit Reticulocyte Lysate (RRL, Promega) supplemented with 20% (v/v) of HEK293T S10 extract and 25 ng of the indicated RL mRNA. When indicated, lysates were pre-incubated with 0.5 μg of NSP1 M1 or WT for 5 min at 30°C before adding the mRNA. Luciferase activity was measured after 60 min of reaction.

### Small hairping RNA (shRNA)-mediated silencing

We silenced the expression of human ABCE1 gene by using lentivirus expressing two different shRNAs [TRCN0000158417 (shRNA_1) and TRCN0000331020 (shRNA_2), Sigma Mission] from pLKO.1 vector. HEK293T were transduced with the indicated lentivirus and selected with 4 μg/ml puromycin for three days. All experiments were performed in the next week after selection since ABCE1 knockdown cells stopped growing and died from 10 days after selection.

### Data processing and bioinformatic analysis

We used raw and processed datasets deposited in GEOs. Data normalization per kilobase per million reads (RPKM) was applied where necessary. Differential expression was calculated using DESeq2 and expressed as $\log_2 FC$ [50]. For further analysis, we selected only mRNA whose 5′ UTR sequence was experimentally confirmed using nanocage technology (5000–6000 mRNAs depending on the dataset) [51]. 5′ UTR features including length and nucleotide composition were calculated using excel formulas. A Python script was prepared to compute the number of G and G + A in 5′ UTR datasets using a 5 nt-sliding window. RNA secondary structure was predicted using the RNAfold program (ViennaRNA package). Human mRNAs with predicted fast ribosome transit were defined as those that showed a CDS length below the lower limit of the 95% confidence interval (CI) of the median and a tAIg above the upper limit of the 95% of the median. mRNAs with predicted slow ribosome transit were defined as those that showed a CDS length above the upper limit of the 95% CI of the median and a tAIg below the lower limit of the 95% CI of the median.

### Machine learning models

A machine learning pipeline was constructed in Python using scikit-learn [52]. As output variable, we selected data deposited in Gene Expression Omnibus (GSE200422) corresponding to the effect of SARS-CoV-2 WT versus ΔNSP1 infection on mRNA abundance in human Calu-3 cells at 8 hpi [27]. The dataset was filtered for a q-value < 0.01. We incorporated five features of 5′ UTR and 5 features of CDS of every mRNA: 5′ UTR length (nt), 5′ UTR G + C (%), predicted secondary structure of 5′ UTR ($\Delta G°$), number of Gs in 5–25 of 5′ UTR, number of G + A in 5–25 of 5′ UTR, CDS length (number of codons), CDS G + C (%), CDS GC3 (%), tAIg_HEK293, and CSCg_HEK293. For Random Forest classification, each feature was standardized so that the values were scaled to a mean of 0 and a standard deviation of 1. Continuous values ($\log_2 FC$) of output variable were transformed into two categorical classes: down (=0), showing $\log_2 FC < -0.5$ and up (=1) showing $\log_2 FC > 0.5$. The two classes were then balanced resulting in a dataset of 1652 genes. Then, the balanced dataset was split into training and testing sets using a ratio of 2.1 (two-thirds training and one-third testing). The hyperparameters for the Random Forest classifier were optimized using a grid search with 10-fold cross-validation. To determine feature importance and their contribution to the model, the mean decrease Gini index was used. CDS length was recognized as the most important feature, followed by 5′ UTR length and tAIg.

### Statistical analysis

Statistical analyses were carried out in GraphPad Prism 6. Two-tailed *t*-test and nonparametric tests were performed using at least three biological replicates per group and at least two technical replicates per sample. Significance was scored as follow: $P < .0001, ***P < .001, **P < .05, *P > .05$, ns = not significant. Pearson and Spearman's correlations were calculated using default parameters.

## Results

### Differential effect of NSP1 on reporter mRNA translation and stability

HEK293T cells were co-transfected with increasing amounts of NSP1-expressing plasmid in combination with several reporter genes available in our lab (Fig. 1A). Each reporter was expressed from the same plasmid, and the 5′ UTRs of the resulting mRNAs were nearly identical, consisting of 82–94 nt with a 60% G + C and moderate secondary struc-
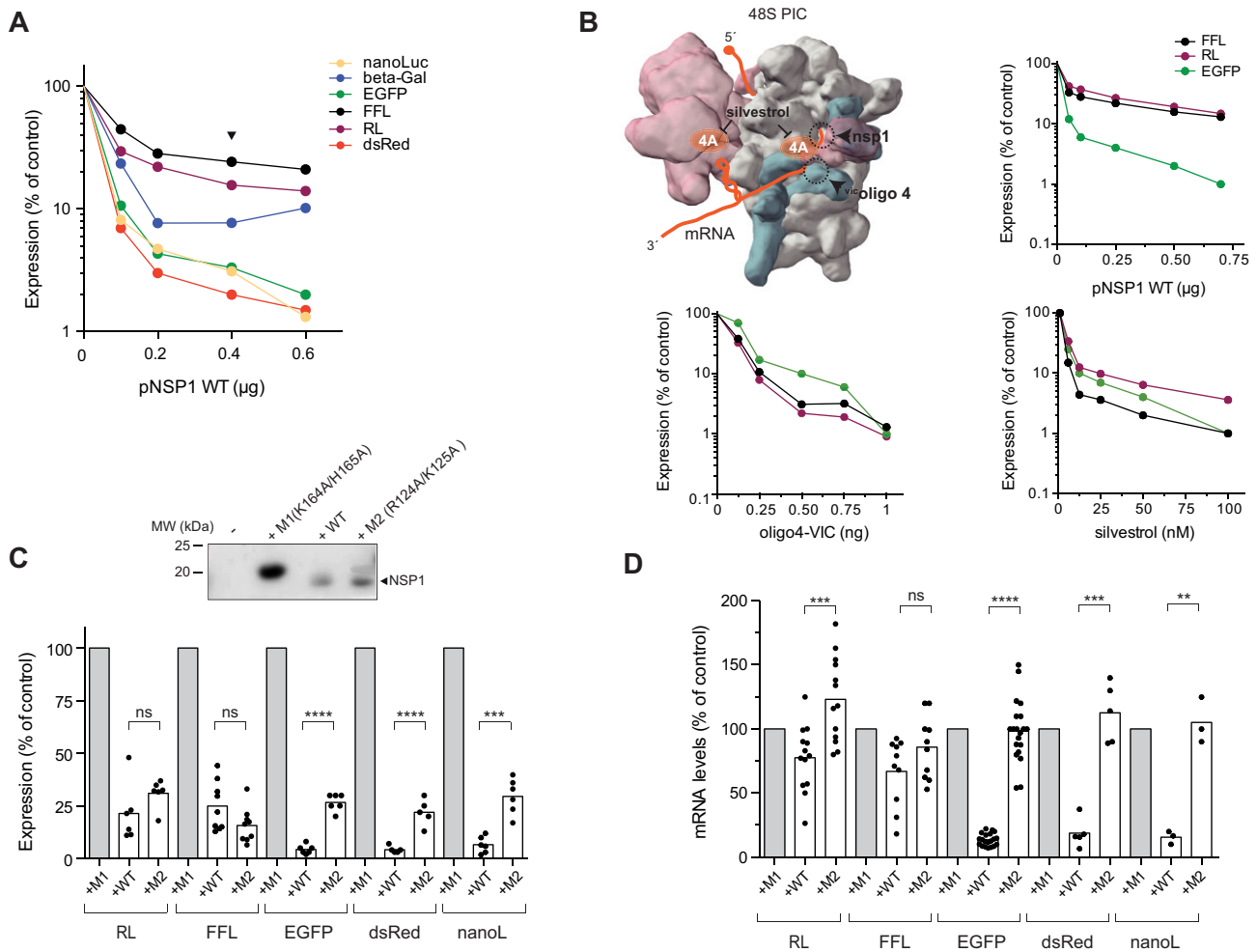
**Figure 1.** Differential sensitivity of reporter gene expression to NSP1. (**A**) Effect of increasing amount of NSP1 WT-expressing plasmid on expression of the indicated reporter genes in HEK293T cells. Data are the average of two independent experiments and are expressed as a percentage relative to the cells transfected with 0.6 μg of pNSP1 M1 (*null* mutant, control). Arrow indicates the amount of pNSP1 plasmid selected for further experiments. (**B**) Comparative analysis of the effect of NSP1 WT, oligo4^VIC, and silvestrol on expression of FFL, RL, and EGFP reporters. The binding sites of these inhibitors on 48S-PIC model are indicated (some eIFs were omitted for simplicity). Data are the average of two independent experiments. (**C**) Comparative analysis of the effect of NSP1 WT and M2 (R124A/K125A, mRNA degradation *null*) on the expression of the indicated reporter genes. A representative western blot monitoring the expression of different NSP1 variants is also shown. (**D**) Comparative analysis of the effect of NSP1 WT and M2 on accumulation of the indicated reporter mRNAs 24 h after transfection. mRNA levels were quantified by RT-qPCR as described in the 'Materials and methods' section. Significance was scored by two-tailed, unpaired *t*-test as described in the 'Materials and methods' section.

ture (5′ UTR 82/94) (Supplementary Fig. S1A). The transcription initiation site (TSS) for these constructs was confirmed by 5′ RACE (Supplementary Fig. S1B). We found that the level of NSP1 WT accumulated in transfected cells was comparable to those found in SARS-CoV-2-infected cells (Supplementary Fig. S1C). Furthermore, we detected a fraction of NSP1 in the 40S/PIC peak as described in some previous reports (Supplementary Fig. S1D) [29, 32]. Whereas the expression of some reporter genes such as dsRed, EGFP and NanoLuc was dramatically inhibited by NSP1 (up to 30-fold), the expression of Firefly luciferase (FFL) and RL was much less affected at any of the NSP1 doses tested (Fig. 1A). To determine whether this inhibitory effect was specific to NSP1, we analyzed the impact of other translation initiation inhibitors such as silvestrol and VIC-oligo 4, which block eIF4A helicase activity and the ES6S region of 40S, respectively [14, 53, 54]. Despite the proximity of the binding sites of all three inhibitors to the mRNA entry channel of the 40S

subunit, VIC-oligo 4 and silvestrol similarly inhibited the expression of all reporter genes tested (Fig. 1B), suggesting that the differential effect of NSP1 was not the simple result of translation inhibition. Since NSP1 binding to the ribosome not only blocks translation but also induces mRNA degradation, we compared the effect of wild-type (WT) NSP1 with previously described mutants deficient in ribosome binding (K164A/H165A, M1) (Supplementary Fig. S1E) or lacking the mRNA degradation activity (R124A/K125A, M2) [34, 36]. Interestingly, unlike NSP1 WT, M2 mutant inhibited the expression of all reporters at comparable extent (3–4-fold reduction), suggesting that the strongest effect of NSP1 WT on EGFP, dsRed, and nanoLuc expression could also involve mRNA degradation (Fig. 1C). To confirm this, we quantified the accumulation of reporter mRNA in cells cotransfected with NSP1 WT or M2. Indeed, NSP1 WT induced a strong reduction of EGFP, dsRed and NanoLuc mRNAs, but not of RL and FFL mRNAs (Fig. 1D).

## Codon usage of CDS mRNA influences sensitivity to NSP1

Since the reporter genes tested are transcribed from the cytomegalovirus (CMV) early promoter and contain almost identical 5′ and 3′ UTR regions flanking the coding sequences (CDS), we analyzed the correlation between sensitivity to NSP1 and a number of CDS parameters including codon composition (usage), G + C content, and predicted RNA secondary structure (Fig. 2A and Supplementary Fig. S2A). To strengthen the analysis, we included additional reporter genes of cellular (e.g. human actin, ACTB) and viral origin (e.g. nucleocapsid, N, from SARS-CoV-2) that showed different codon usage and nucleotide composition (Supplementary Table S1). As a measure of codon bias, we used codon adaptation index (CAI), tAIg, and CSCg estimated for HEK293T cells [46, 47]. Although based on different parameters, all three metrics are related and can be used as predictors of codon bias and expression level [55]. Interestingly, we found a positive and strong correlation between sensitivity to NSP1 WT and optimal codon usage of the reporter CDS for each of the three metrics used (Fig. 2A). Importantly, no correlation was found for NSP1 M2, suggesting that codon composition of reporter genes may also influence NSP1 WT-induced mRNA degradation (Fig. 2A). A similar positive correlation was found between sensitivity to NSP1 WT and G + C content, both analyzed globally and specifically at the third-codon position (GC3) (Supplementary Fig. S2A). This result was expected since both G + C and GC3 contents are linked to codon bias [19].

To test the influence of codon bias on sensitivity to NSP1 WT, we recoded the RL and EGFP CDS to generate variants with different codon usage. From the parental RL, which shows a suboptimal codon usage (tAIg = 0.2), we generated three variants (RL-1, RL-2, and RL-3) with increased CAI to HEK293T cells (Fig. 2B). The resulting codon-optimized RL-1 and RL-3 constructs showed an increased expression compared to the parental RL (Supplementary Fig. S2B), and a dramatic increase in sensitivity to NSP1 WT (up to 30-fold for RL-3 variant), but not to NSP1 M2, suggesting that the corresponding codon-optimized RL variants became more sensitive to NSP1-mediated mRNA degradation (Fig. 2B). To confirm this, we compared the amount of RL and RL-3 mRNA accumulating in cells expressing NSP1s, revealing a strong reduction of RL-3 mRNA in the presence WT NSP1 (Fig. 2B). The expression of RL-2 variant with a subtle increase in codon adaptation was inhibited by NSP1 to a similar extent than parental RL, showing that ultrasensitivity to NSP1 requires a substantial codon optimization (Fig. 2B). Conversely, we deoptimized the parental EGFP (tAIg = 0.3) to generate two variants (EGFP-1 and 2) with a predicted suboptimal codon usage for HEK293T cells (Supplementary Fig. S2C). Notably, codon deoptimization of EGFP drastically reduced the sensitivity to NSP1 WT from 30-fold to 2–3-fold inhibition, a rate that was similar to the parental RL and FFL reporters (Supplementary Fig. S2C). These results indicate that genes with optimal codon usage are more sensitive to NSP1-induced inhibition.

Since optimal codon usage accelerates the rate of translation elongation, we measured the TE of some reporter mRNAs by polysome profiling. As expected, reporter mRNAs with optimal tAIg (dsRed, EGFP, NanoLuc, RL-3) showed a higher TE, and conversely, reporter mRNAs with lower tAIg values (FFL and RL) accumulated preferentially in monosome and 40S fractions, which is indicative of a slower rate of translation initiation (Fig. 2C and Supplementary Fig. 2D). We found a strong direct correlation between TE and sensitivity to NSP1 when analyzed at both protein (reporter activity) and mRNA levels (Fig. 2C). To further confirm this, we used previously published data to calculate TE of human mRNAs in Calu3 cells and the impact of SARS-CoV-2 WT infection on mRNA levels. Similar to that found with reporter mRNAs, a significant correlation was observed between TE and NSP1-mediated degradation of human mRNAs; the mRNAs that were more sensitive to NSP1 in SARS-CoV-2-infected cells also showed higher translation rates in uninfected cells (Fig. 2C).

To experimentally confirm these findings in SARS-CoV-2-infected cells, we compared the expression of EGFP, RL-3 and FFL in VeroE6 cells, infected or not with SARS-CoV-2. Notably, expression of RL-3 and EGFP was more sensitive than FFL to SARS-CoV-2 infection, although the observed inhibition was less pronounced than that found in VeroE6 cells transfected with pNSP1 plasmid (which accumulated ∼2-fold more NSP1 than in infected cells at 24 hpi) (Fig. 2D). Therefore, the differential effect of NSP1 on reporter genes with different codon usages and TE was also seen in SARS-CoV-2-infected cells.

## Widely diverse influence of 5′ UTR on NSP1-mediated effect

Next, we systematically analyzed the influence of 5′ UTR on sensitivity to NSP1 using a collection of parental RL constructs bearing different 5′ UTRs (Fig. 3A and Supplementary Table S2). We analyzed the influence of 5′ UTR length (from 8 to 650 nt), the presence of stable secondary structures at different positions of the 5′ UTR and the influence of nucleotide composition and motifs in the 5′ UTR. We found great diversity in effects of 5′ UTRs on NSP1-mediated effect, ranging from those that conferred an extreme sensitivity (>40-fold inhibition) to those that allowed enhanced expression of the reporter gene in the presence of NSP1 WT. Regarding the influence of 5′ UTR length, a progressive increase in length from 82-RL up to 650 nt strongly reduced the expression of RL, but had a minor effect on sensitivity to NSP1 (Fig. 3A). However, shortening the 5′ UTR up to 5 nt in an RL construct bearing the TISU element (TISU-RL) dramatically increased sensitivity to NSP1 WT, and to a lesser extent to NSP1 M2 (Fig. 3A and B). TISU provides an optimal AUG context to promote substantial translation of mRNA bearing extremely short 5′ UTRs [56, 57]. We confirmed this effect by placing the TISU element upstream of EGFP CDS, resulting in an almost complete inhibition of EGFP expression by NSP1 WT (Fig. 3B). The levels of TISU-containing mRNAs were also further reduced by NSP1 WT compared to parental constructs, especially for TISU-EGFP mRNA that dropped 15-fold in the presence of NSP1 (Fig. 3B).

Next, we analyzed the contribution of stable secondary structures in the 5′ UTR to NSP1-mediated inhibition. It is well known that stable stem loops ($\Delta G° \leftarrow 20$ kcal.mol$^{-1}$) can delay or even prevent scanning of 43S-PIC on mRNA during translation initiation [5, 6]. We used previously described constructs generated from parental 82-FFL bearing a stem loop of $-20$ (SL20-FFL) or -50 kcal mol$^{-1}$ (SL50-FFL) immedi-
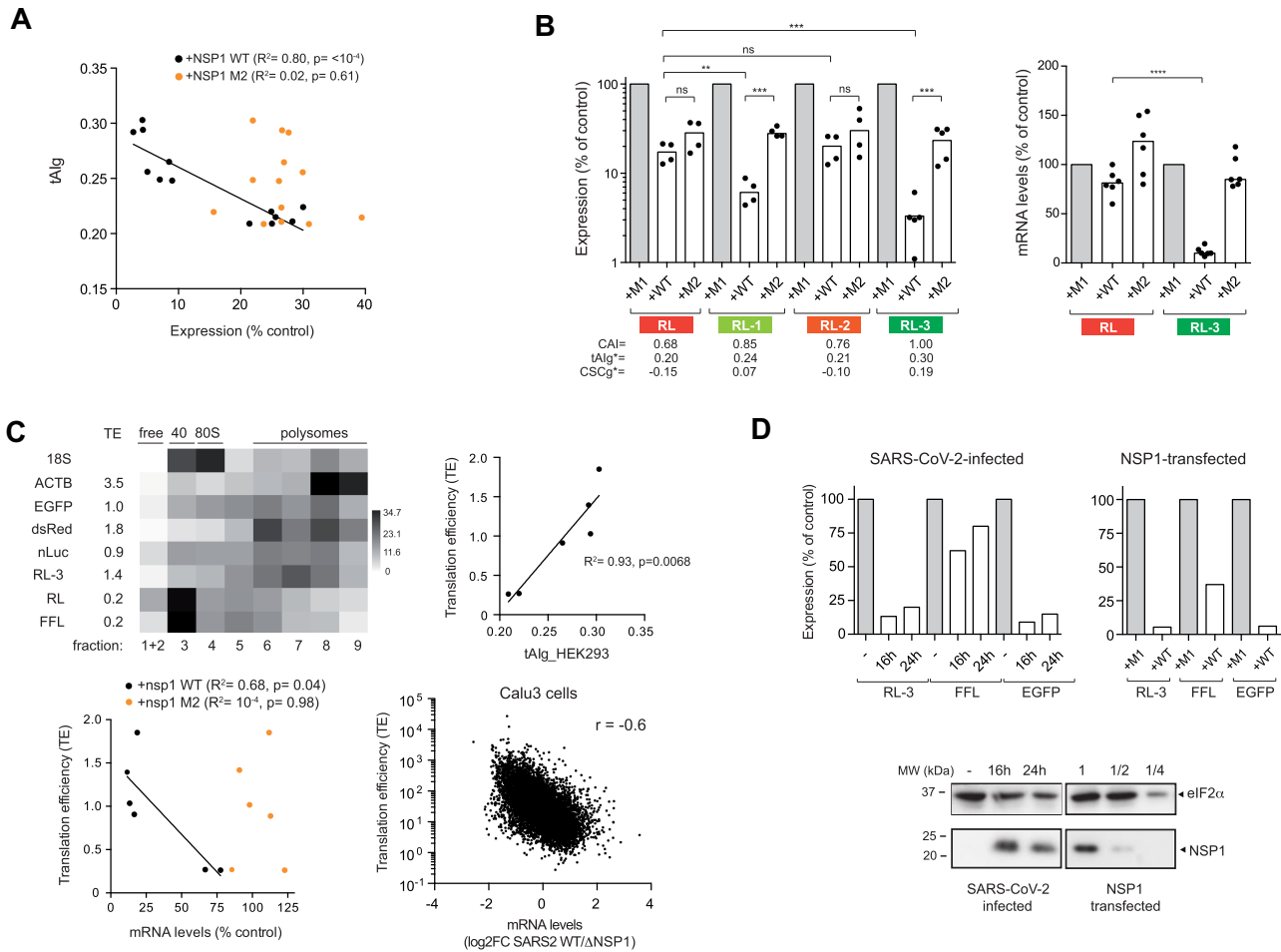
**Figure 2.** Influence of mRNA codon usage bias on sensitivity to NSP1. (**A**) Correlation between tAIg of reporter mRNAs and their corresponding sensitivities to NSP1 WT or M2. Data were scored as described in Fig. 1. The $R^2$ and *P*-values are indicated. (**B**) Effect of RL CDS recoding on sensitivity to NSP1 WT and M2. The CAI, tAIg, and CSCg values for parental (RL) and RL-1 to three variants are indicated. Right panel shows a relative quantification of RL and RL-3 mRNA levels in HEK293T cells expressing NSP1 WT or M2. (**C**) Correlation between TE, tAIg, and sensitivity to NSP1 WT and M2. Upper left panel shows TE calculated for the indicated mRNAs by polysome profiling followed of RT-qPCR. The resulting heat map of the relative quantification is shown. Correlation between TE and tAIg (upper right panel) or TE and sensitivity to mRNA-induced degradation by NSP1 WT or M2 (bottom panel). Data were scored as described above and the resulting $R^2$ is indicated. Bottom right panel shows the correlation between TE of human mRNAs and sensitivity to NSP1-induced mRNA degradation in Calu3 cells infected with SARS-CoV-2 WT compared with that infected by a mutant virus defective in NSP1 (ΔNSP1). Data were taken from GSE200422 [27] and analyzed as described in the 'Materials and methods' section; the resulting Spearman correlation coefficient (r) is indicated. (**D**) Effect of SARS-CoV-2 infection on expression of RL-3, FFL, and EGFP in VeroE6 cells. Cells were infected, transfected with a combination of the three plasmids and analyzed at 16 and 24 hpi for luciferase activity (upper panel) or by western blot against NSP1, RPS6, and EGFP. In parallel wells, cells were co-transfected with pNSP1 and a combination of RL-3, FFL, and EGFP plasmids and analyzed 24 h later (top panels). A comparative analysis of NSP1 levels accumulated in infected and transfected cells is shown (bottom panel). Blots were also probed with anti-eIF2α as loading control. The 1/2 and 1/4 dilutions of cell extracts are also indicated.

ately upstream of the initiator AUG (AUGi). The resulting constructs showed a reduced expression, especially for SL50-FFL that was 30- to 50-fold lower than parental 82-FFL (Fig. 3C). Moreover, we also observed reduced levels of SL50-FFL mRNA compared with 82-FFL, suggesting that SL50-induced translation blockade resulted in some mRNA destabilization (Fig. 3C). Surprisingly, the expression of SL50-FFL was stimulated ≈5-fold by NSP1 WT, but not by NSP1 M2 (Fig. 3C). We also found that NSP1 WT increased the levels of SL50-FFL mRNAs, suggesting that NSP1-mediated relief of SL50-FFL mRNA translation can result in mRNA stabilization (Fig. 3C). A similar result was observed when SL50 was placed in RL, although in this case the stimulation by NSP1 WT co-expression was only ≈2-fold (Supplementary Fig. S3A and B). We also confirmed by western blot that the size of the luciferase band accumulated in NSP1-expressing cells was sim-

ilar to that of control samples, showing that NSP1-mediated bypass of SL50 involves initiation at the natural AUGi of FFL mRNA (Fig. 3C).

## G-rich eIF4A binding motifs proximal to 5′ cap determines sensitivity to NSP1

We next analyzed the influence of nucleotide composition of the 5′ UTR on sensitivity to NSP1 (Fig. 4A). We used a previously described RL construct bearing a 5′ UTR lacking Gs (G-less-RL) except the one located in the second position from the TSS, which is necessary for transcription from the CMV promoter [14]. G-less 5′ UTR reduced the expression of parental RL, but it had little effect on the expression of codon-optimized RL-3 (Fig. 4B). Interestingly, the presence of NSP1 WT, but not M2, enhanced the expression of G-less-RL by
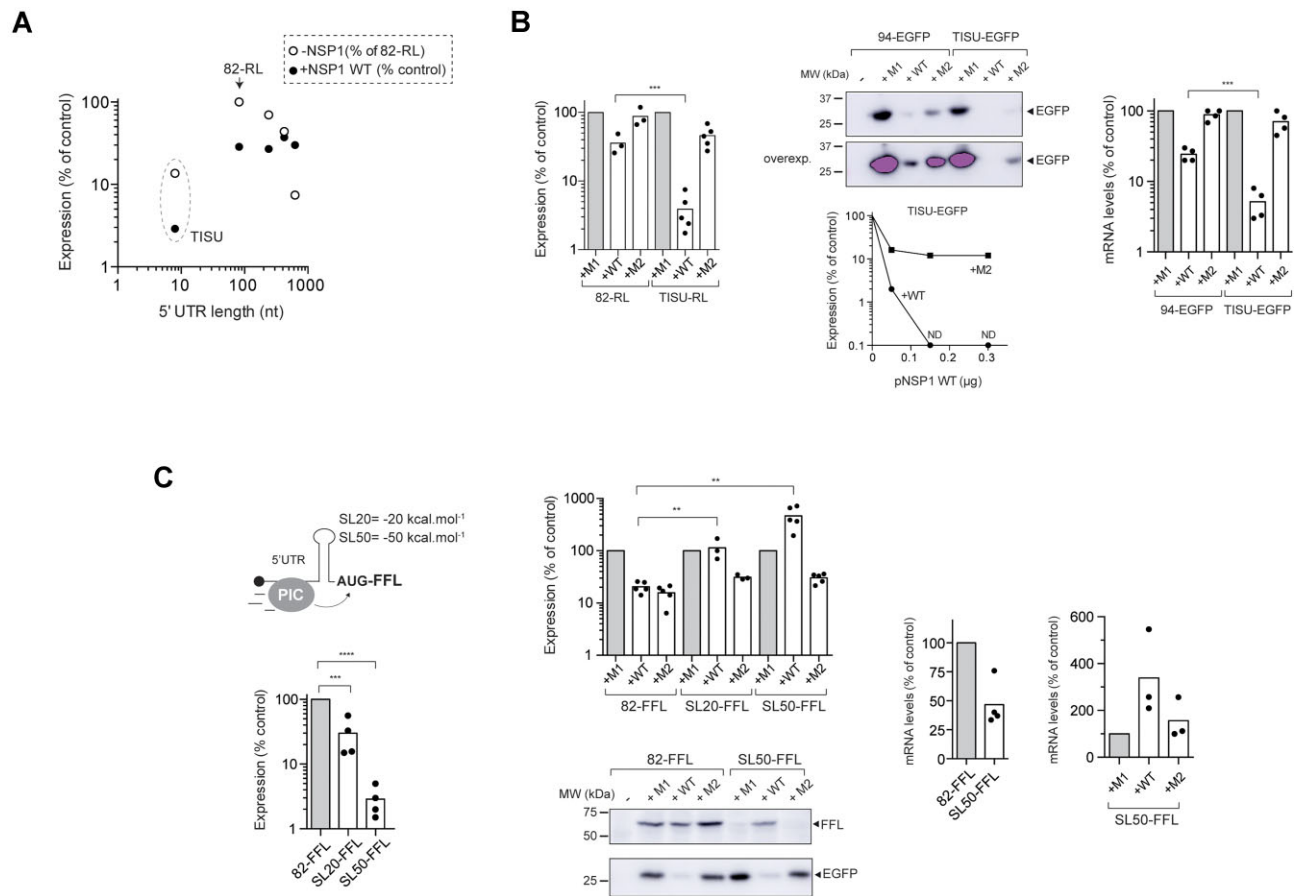
**Figure 3.** Influence of 5′ UTR length and secondary structure on sensitivity to NSP1. (**A**) RL plasmids with 5′ UTRs ranging from 8 to 650 nt were co-transfected with 82-FFL plasmid in HEK293T cells and the resulting normalized activities were expressed as a percentage relative to the 82-RL construct (-NSP1). The sensitivity to NSP1 WT expression (+NSP1) for every construct is expressed as described in previous figures. (**B**) Effect of NSP1 WT and M2 on 82-RL and TISU-RL expression (left panel) or 94-EGFP and TISU-EGFP expression assayed by western blot (middle panel). The effect of increasing amount of pNSP1 WT or M2 plasmids on TISU-EGFP expression is also shown. ND, not detected. Right panel shows mRNA level quantification of 94-EGFP and TISU-EGFP in cells expressing NSP1 WT or M2. Significance was scored as described before. (**C**) Influence of stable stem-loops (SLs) placed immediately before the AUG of FFL on the sensitivity to NSP1s. The placement and stability (ΔG°) of SLs are shown (left panel). Effect of SLs on FFL expression relative to 82-FFL (lower left panel) and sensitivity to NSP1 WT and M2 (top middle panel). SL50-FFL expression relief by NSP1 WT was also confirmed by western blot using anti-Luc antibodies (bottom middle panel). Transfections also included an EGFP expressing plasmid and the resulting western blot anti-EGFP is shown (bottom bottom panel). The right panels show relative quantification of 82-FFL and SL50-FFL mRNA levels in control and cells co-expressing NSP1 WT or M2.

up to 10-fold (Fig. 4B). We also observed a slight increase in mRNA levels of G-less-RL in the presence of NSP1 WT, suggesting that translational relief by NSP1 can result in some mRNA stabilization (Supplementary Fig. S4A). In addition, the presence of G-less 5′ UTR made RL-3 resistant to NSP1 WT (Fig. 4B). The effect of NSP1 on the expression of some of these constructs was also confirmed in other cell types including HeLa, MEF and Calu-3 cells (Supplementary Fig. S4B). Similar results were obtained when 5′ UTR-82 of EGFP and FFL was replaced by a 5′ UTR low in Gs (G-low), unrelated in sequence to G-less (Supplementary Fig. S4C and 4D). These results showed that NSP1 WT can act as a translation enhancer of mRNAs bearing G-poor 5′ UTR and suboptimal codon usage.

We next introduced Gs and/or secondary structure in G-less 5′ UTR distally or proximally to 5′ cap (constructs named as G-; Fig. 4A). Whereas introduction of Gs or a stem loop of moderate structure immediately before AUGi did not affect NSP1 WT-mediated stimulation, introduction of 9 Gs spaced by 1–5 nt near the 5′ cap (G-M1) drastically increased sensi-

tivity to NSP1. Thus, expression of G-M1-RL was inhibited by >10-fold, resulting in a net change of almost 100-fold when compared with G-less (Fig. 4C). Introduction of a stem loop of moderate stability also made G-less sensitive (G-SL1) to NSP1 WT-mediated inhibition, although the change was less dramatic (Fig. 4C). This result suggests that the presence of Gs near the 5′ cap had a greater impact on sensitivity to NSP1 than a tendency to form secondary structures. Since helicase eIF4A binds purine rich motifs in the 5′ UTR of many mRNAs [58, 59], we inserted a $(GA)_6$ motif near the 5′ cap (G-M4) to test the role of eIF4A in the susceptibility to NSP1. Similar to G-M1-RL, expression of G-M4-RL construct was strongly inhibited by NSP1, suggesting that eIF4A binding proximal to mRNA 5′ cap could determine the sensitivity to NSP1. To confirm this, we compared the sensitivity of G-less-RL constructs to silvestrol and RocA, two inhibitors that clamp eIF4A on target mRNAs [58]. Notably, we found that G-less-RL expression was much less sensitive to eIF4A inhibitors than reporter genes carrying the 5′ UTR-82, thus confirming our previous *in vitro* data [14]. Similar results were found when the sensitiv-
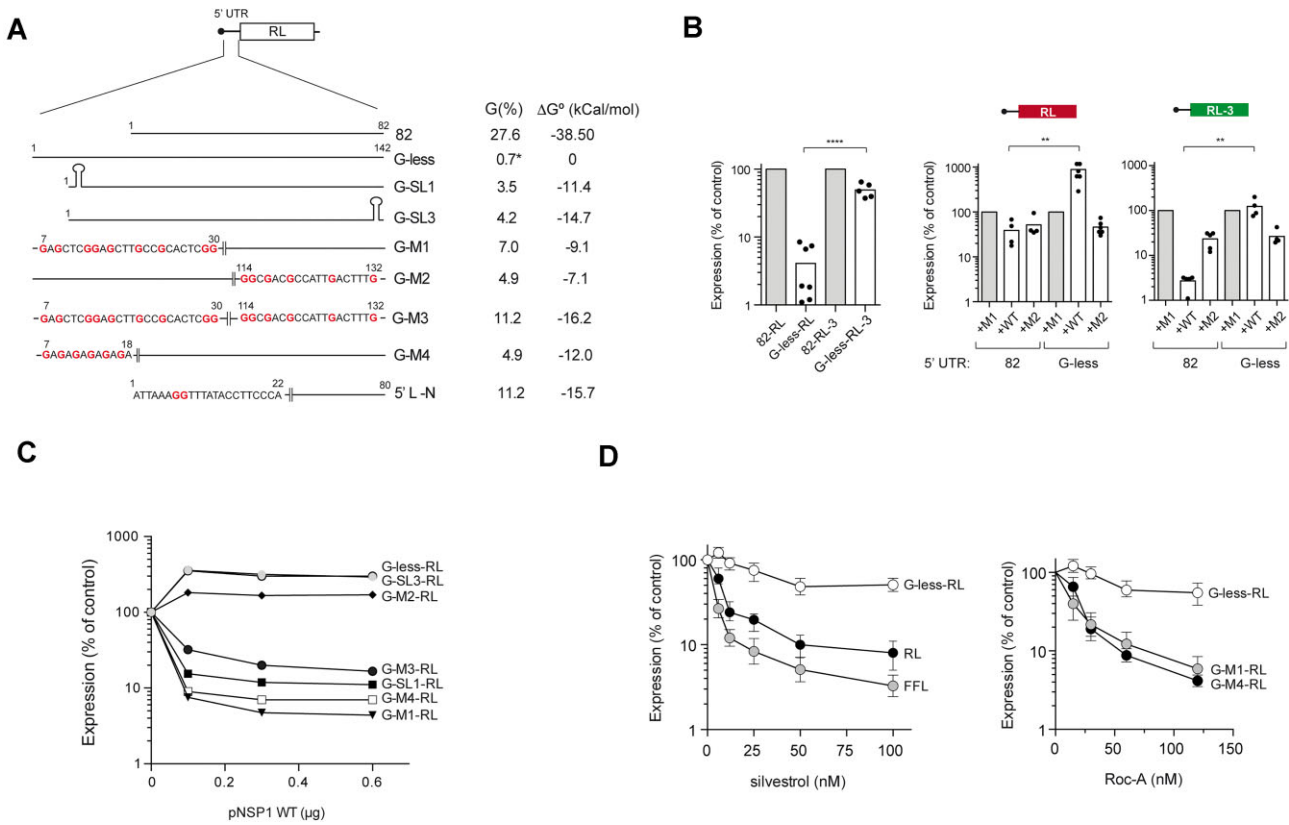
**Figure 4.** The presence of Gs near the 5′ cap critically determines the sensitivity to NSP1. (**A**) Schematic diagram of G-less 5′ UTR and its variants including G, SL, or GA motifs at the indicated positions. The composition (% G) and stability (ΔG° in kcal/mol) of the resulting sequences are indicated. The Gs introduced in the sequences are in bold. The first 22-nt sequence of 5′ L N of SARS-CoV-2 is also shown. (**B**) Left panel shows a comparative analysis of RL expression with G-less or 82 5′ UTRs in combination with optimal or suboptimal codon usage (RL versus RL-3) in HEK293T cells. Effect of NSP1 WT or M2 on the expression of G-less 5′ UTR in combination with optimal or suboptimal codon usage (middle and right panels). Significance was scored as described in previous figures. (**C**) Sensitivity of G-less and its variants to increasing amounts of NSP1 WT. Data are the average of two independent experiments. (**D**) Reduced sensitivity of G-less-RL expression to eIF4A inhibition by silvestrol (left panel) or RocA (right panel) compared to RL, FFL, or M1 and M4 variants of G-less-RL. Inhibitors were added 1 h after transfection and measurements were taken 16 h later. Data are the mean ± standard deviation from three independent experiments.

ity of G-low-FFL to RocA was compared to parental 82-FFL (Supplementary Fig. S4D). On the contrary, the expression of G-M4-RL and G-M1-RL showed an increased sensitivity to RocA similar or even superior to 82-RL (Fig. 4D).

## NSP1 blocks threading but allows slotting of mRNA into 43S-PIC

The differential effect observed on translation suggested that NSP1 could be blocking the threading of some mRNAs (e.g. TISU and G-M1/M4), while allowing the slotting of 43S-PIC on other mRNAs (e.g. G-less/G-low). To further substantiate this idea, we first progressively shortened parental 5′ UTR G-less from 142 to 8 nt, thereby reducing the chance of 43S-PIC slotting on the resulting 5′ UTR. Whereas NSP1 was still able to enhance the expression of 22 nt-long G-less 5′ UTR, and to a lesser extent of 15 nt-long G-less, a further reduction to 12 nt abruptly changed the effect of NSP1 on its expression and reached a 30-fold inhibition when the 5′ UTR was shortened to 8 nt (Fig. 5A). Since a full insertion of mRNA into the 40S channel of 43S-PIC requires a minimum of 24 nt, the 12 nt-long G-less 5′ UTR was probably too short to allow a productive slotting in the presence of NSP1 even if some back scanning is considered [12, 15, 54, 60]. These results agree well with a blind spot of 10–30 nt from the 5′ cap

detected when 48S-PIC was assembled *in vitro* on a 50 nt-long 5′ UTR mRNA, a result that has been taken as evidence of slotting [12]. Since this 5′ UTR is very similar in composition to our G-less and G-low 5′ UTRs (see Supplementary Table S2), we inserted it in our RL plasmid (Slot-RL) and the resulting sensitivity to NSP1 was tested. Slot-RL showed a similar expression to G-less-RL that was resistant to NSP1 WT (Fig. 5B).

Next, we inserted upstream AUGs out of frame at 8, 13, 18, 28, and 112 nt from the 5′ cap in G-less-RL (Fig. 5C). The resulting short upstream open reading frames (uORFs) overlap the first 25 nt of RL CDS to avoid the possibility of reinitiation, as occurs in ATF4 mRNA [61]. Notably, whereas insertion of uAUG at positions 8 and 13 had only a moderate effect on expression (3–4-fold inhibition), the placement of uAUG at position 28 and onwards strongly reduced the expression of G-less-RL (25–30-fold) (Fig. 5C), showing that a significant proportion of 43S-PIC can efficiently bypass any uAUG near the cap of G-less mRNA. Interestingly, the presence of NSP1 still enhanced the expression of G-uAUG_28, whereas the expression of G-uAUG_8 was unaffected by NSP1 (Fig. 5C). However, when (GA)₆ motif was placed in these constructs (derived from G-M4-RL; see Fig. 4A), a strong inhibition by NSP1 was observed (Fig. 5C). These results showed that NSP1 allowed the slotting of 43S-PIC on G-less-RL
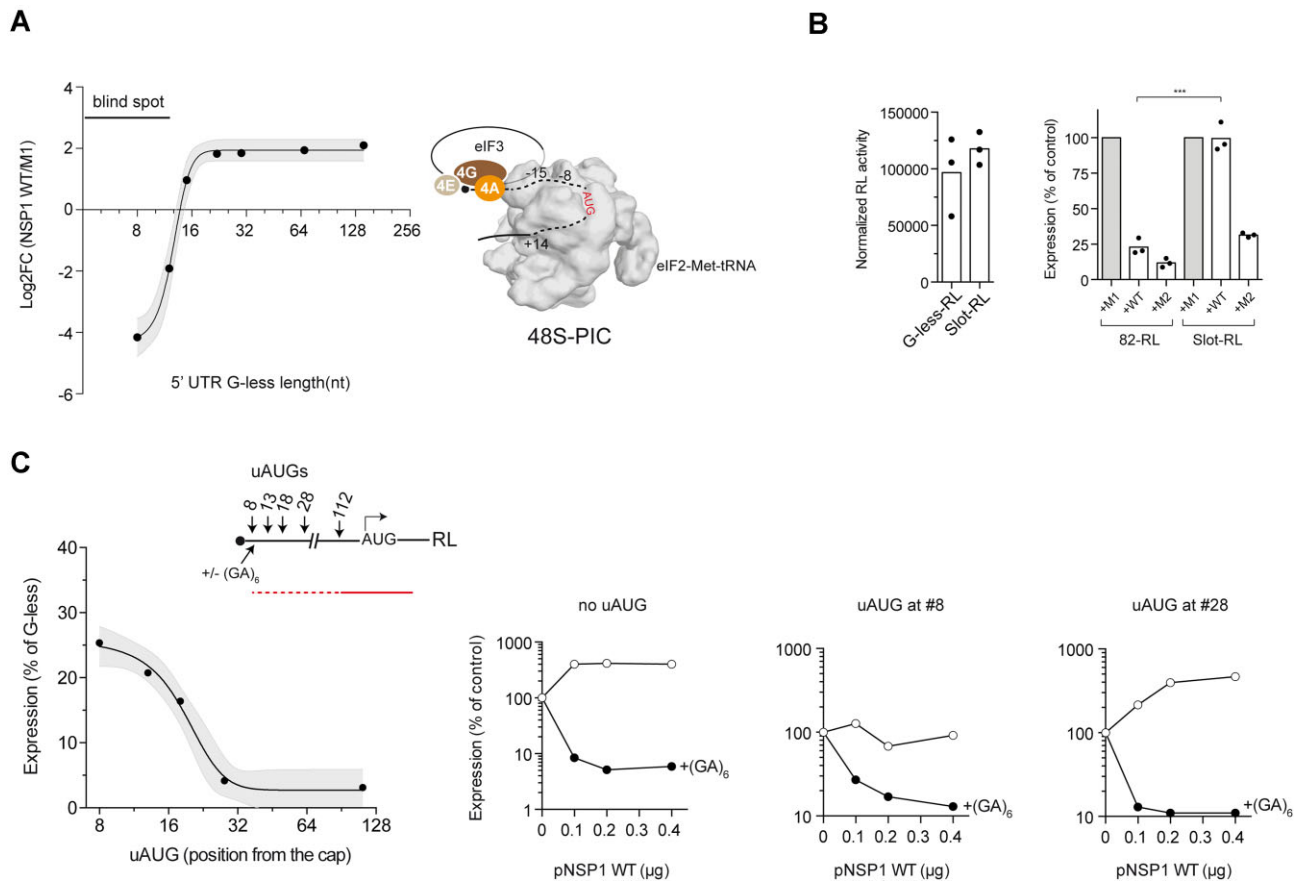
**Figure 5.** NSP1 promotes slotting of PIC to mRNA. (**A**) Effect of progressive shortening of 5′ UTR G-less on sensitivity to NSP1 WT in HEK293T cells. Co-transfections also included FFL plasmid to normalize for FFL activity. Data are expressed as log$_2$ fold-change expression. Data points correspond to the mean from at least three independent experiments that were adjusted to a sigmoid function (R$^2$ = 0.92); shading represents the 95% confidence interval. The resulting estimated blind spot of 12–15 nt is shown. A model of 48S-PIC showing the position of mRNA residues along the mRNA channel of 40S is also represented (right panel). (**B**) Comparative analysis of the expression of G-less-RL and Slot-RL constructs (left panel). Comparative analysis of the sensitivity of 82-RL and Slot-RL to NSP1s (right panel). Significance was scored as described in previous figures. (**C**) Effect of upstream out-of-frame AUGs (uAUG) on expression of G-less-RL and sensitivity to NSP1. The uAUG was inserted where indicated and the resulting uORFs partially overlap the 5′ of the main RL ORF (line). The presence of eIF4A-binding motif (GA)$_6$ is indicated. HEK293T cells were cotransfected with the indicated constructs together with pFFL and the resulting expression was normalized for FFL activity and expressed as % of G-less-RL. Data represent the mean of four independent experiments that were adjusted to a sigmoid function (R$^2$ = 0.89); shading represents the 95% confidence interval (left panel). Middle and right panels show the effect of increasing amounts of pNSP1 WT plasmid on the expression of the indicated construct derived from pG-less-RL (open circles) or pG-M4-RL (black circles).

mRNAs even beyond that expected according to the theoretical blind spot.

## Both 5′ UTR and codon usage of SARS-CoV-2 mRNAs promote resistance to NSP1

The next experiments were aimed at understanding how SARS-CoV-2 mRNAs can escape the inhibitory effect of NSP1. We first compared nucleotide composition and codon usage of full CDS of different coronavirus with those of human mRNAs (Fig. 6A and Supplementary Fig. S5A). As noticed before, a suppression in G + C and GC3 content is observed in coronavirus genomes that results in low CAI values [20, 23]. Notably, SARS-CoV-2 and the other human coronaviruses also showed low tAIg and CSCg values that clearly distance themselves from human mRNAs, especially from highly expressed mRNAs such as ACTB (Fig. 6A and Supplementary Fig. S5A). To analyze the effect of NSP1 on the expression of SARS-CoV-2 genes, we selected the mRNA encoding viral nucleocapsid (N) as representative since it is the most abundant viral mRNA

detected in infected cells [62]. The N mRNA includes the natural 80-nt 5′ leader (5′ L) that it is also present in the rest of viral mRNAs [62]. Moreover, N mRNA shows a tAIg representative of full CDS of SARS-CoV-2 ($\approx$0.21). As expected, the expression of 5′ L-N was resistant to NSP1 WT (Fig. 6B). Surprisingly, NSP1 M2 consistently reduced the expression of 5′ L-N to a similar extent to FFL, EGFP and RL reporters (Figs 1 and 6), confirming that expression of N mRNA is adapted to the presence of NSP1 WT. To test the contribution of 5′ UTR and codon usage on the resistance of N mRNA to NSP1, we replaced the 5′ L with 5′ UTR-82 alone or in combination with a variant of N that shows an improved codon usage for human cells (N1, tAIg = 0.248) (Fig. 6B). Expression of 5′ UTR-82-N was now sensitive to NSP1 WT and M2 to a similar extent to FFL and RL reporters (Figs 1 and 6). Notably, the combination of 5′ UTR-82 and N1 dramatically increased the sensitivity to NSP1 WT (15- to 20-fold inhibition) comparable to that seen for EGFP and ACTB mRNAs (Fig. 6B and Supplementary Fig. S5B). We also found reduced levels of 5′ UTR-82-N1 mRNAs in the presence of NSP1 WT,
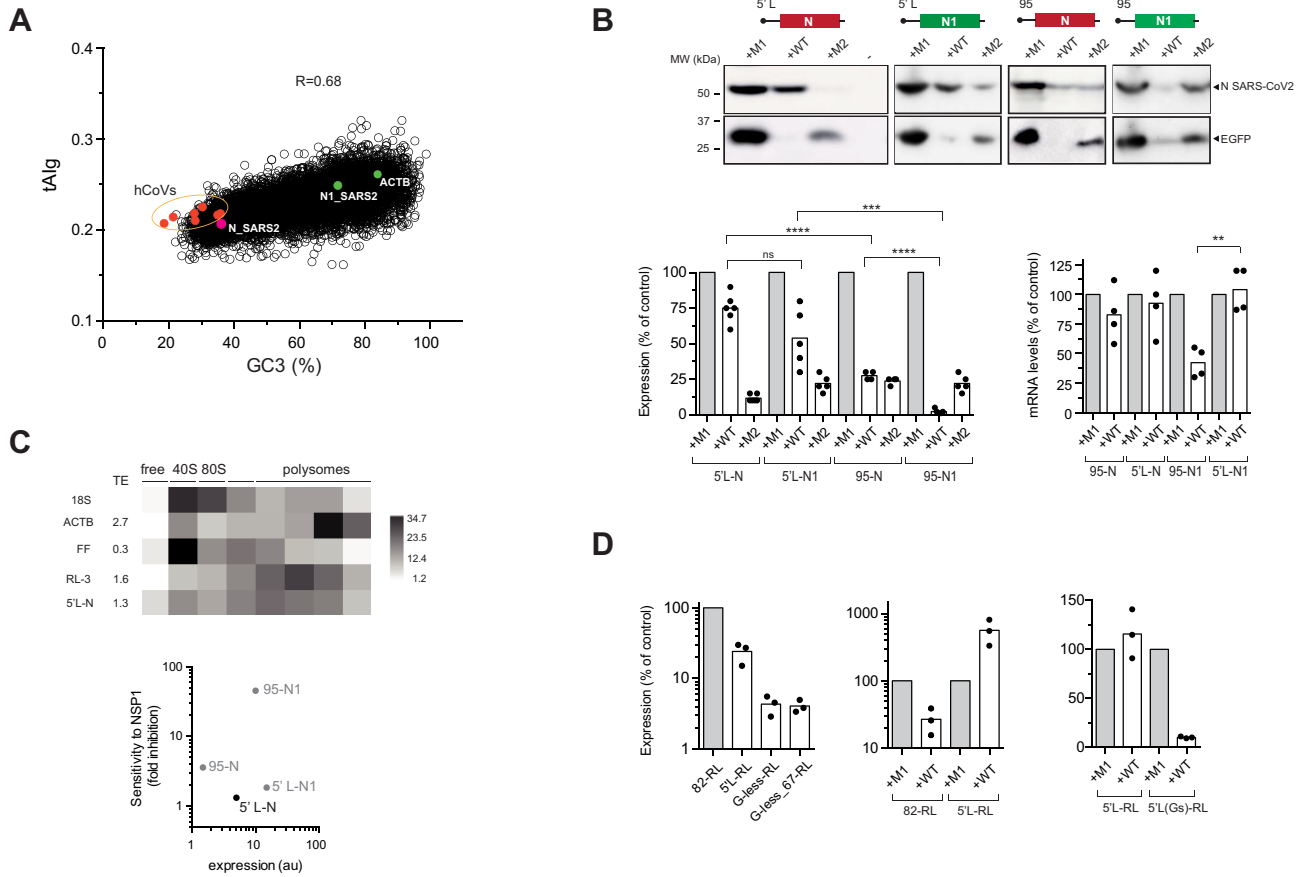
**Figure 6.** A combination of 5′ UTR and codon usage bias promotes resistance of N mRNA SARS-CoV-2 to NSP1. (**A**) Plot of tAIg and GC3 (%) distributions among human and coronavirus (hCoV) mRNAs. Values corresponding to human ACTB and SARS-CoV-2 mRNA N and its variant N1 are indicated, along with the Pearson correlation. (**B**) A representative western blot analysis of N and EGFP accumulated in HEK293T cells co-expressing NSP1 M1, WT or M2 (upper panel) and the corresponding quantification (bottom left panel). Relative quantification of mRNA levels of N variants in cells co-expressing NSP1 WT (bottom right panel). Significance was scored as described in previous figures. The combination of 5′ UTR (95 or 5′ L) with CDS N is indicated at the top of each panel. The tAIg values for N and N1 are 0.21 and 0.248, respectively. (**C**) TE of 5′ L-N mRNA in HEK293T cells compared to some other reporters included in the co-transfection experiment (upper panel). Comparative analysis of the expression level and sensitivity to NSP1 among N variants with different combinations of 5′ UTR and codon usage. (**D**) Analysis of the expression of 5′ L-RL, G-less-RL and G-less_67-RL compared to 82-RL (left panel). The sensitivities of 5′ L-RL expression (middle panel) and its variant (5′ L(Gs)-RL, right panel) to NSP1 WT are shown.

showing that unlike parental 5′ L-N mRNA, 5′ UTR-82-N1 mRNA became sensitive to NSP1-induced degradation (Fig. 6B). However, the replacement of 5′ UTR-82 with 5′ L significantly reduced the NSP1 sensitivity of the N1-expressing construct, showing that both natural 5′ L and codon usage of N mRNA contribute to resistance to NSP1.

Despite the suboptimal codon usage of natural N SARS-CoV-2 mRNA in human cells, polysome profiling experiments revealed a TE of N mRNA comparable to EGFP and much higher than FFL and RL mRNAs, which have similar tAIg (Fig. 6C and Supplementary Table S1). To understand this, we first compared the expression of the four variants of N in HEK293T cells. The presence of 5′ L significantly increased the expression of the N protein to a level comparable to that found in 5′ UTR-82-N1. Despite sharing a similar nucleotide composition, 5′ L was more efficient than G-less/G-less_67 in directing the expression of RL (Fig. 6D). Interestingly, NSP1 WT enhanced the expression of 5′ L-RL about six-fold, reaching an expression similar to 5′ UTR-82-RL. However, when we placed a $(GA)_6$ motif in the 5′ L-RL construct [5′L(Gs)-RL], a strong inhibition by NSP1 similar to the effect on G-M4-RL was observed. This result was expected since previ-

ous studies showed that insertion of Gs in the SL1 of 5′ L of SARS-CoV-2 increased its sensitivity to NSP1 [35, 41]. Moreover, we found an intermediate sensitivity of 5′ L to RocA compared to G-less and 5′ UTR 82 (Supplementary Fig. S5C). Taken together, these results suggest that SARS-CoV-2 mRNA has reached an optimal combination of 5′ UTR and codon usage bias to both support efficient translation and to resist the inhibitory effect of NSP1.

## A genome-scale model of NSP1 effect on translation and mRNA stability

To extend our analysis genome-wide, we used previously published datasets of mRNA abundance in cells expressing NSP1 alone [31] or in the context of infection [27] (Fig. 7A). For this analysis, we selected only mRNAs whose 5′ UTR sequence was experimentally confirmed using nanocage technology (6000 mRNAs) [51]. Notably, mRNAs with the greatest sensitivity to NSP1 (*down* group) showed shorter 5′ UTR and CDS than the median, and higher tAIg than the median in the two analyses (Supplementary Fig. S6A and B). Conversely, mRNAs whose levels increased upon SARS-CoV-2 WT infec-
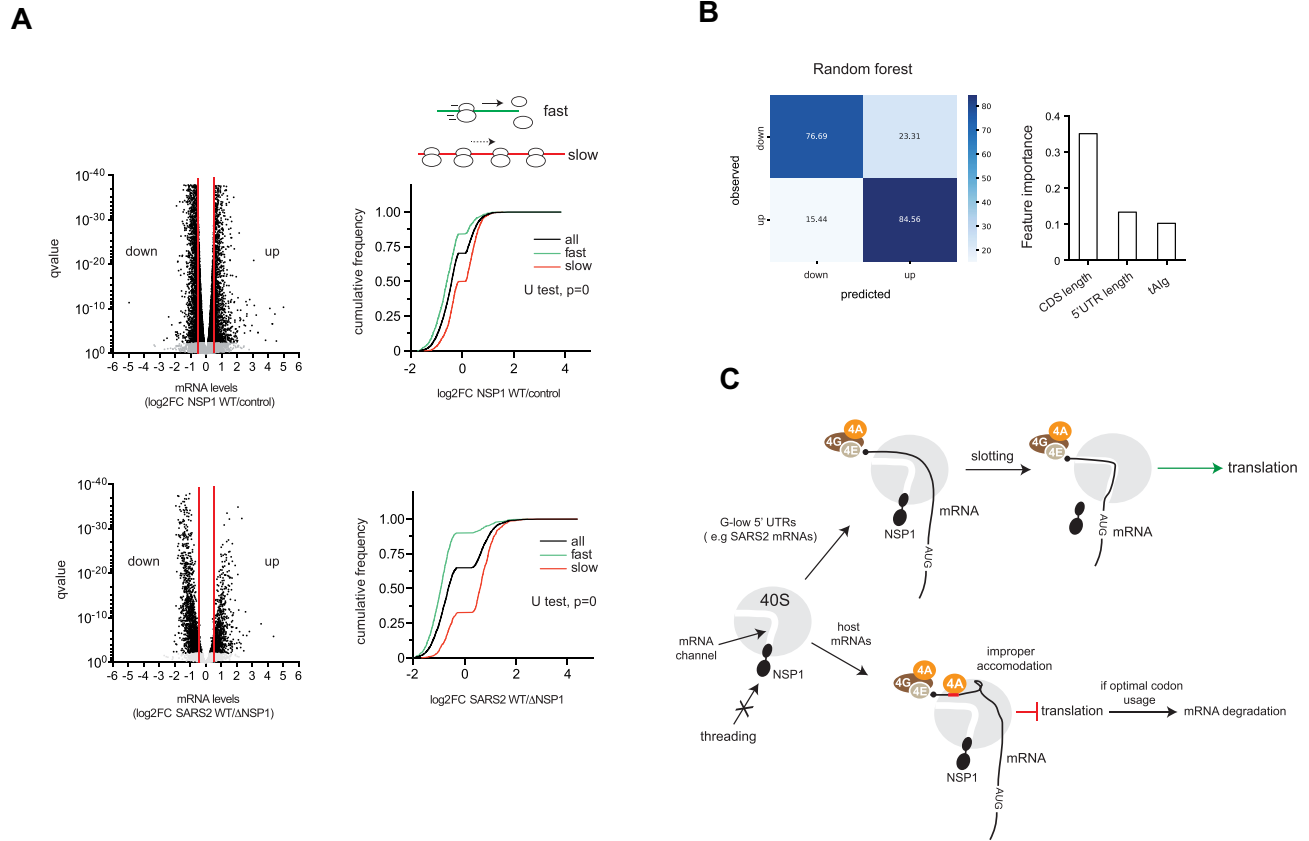
**A**



**B**



**C**



**Figure 7.** Genome-wide analysis of features in human mRNAs associated with sensitivity to NSP1. (**A**) Violin plots of the differential sensitivity of human mRNAs to NSP1 in transfected cells or in the context of infection. For the analysis, we used the datasets from Yuan *et al.* (2020; top) and from Fisher *et al.* (2022; GSE200422, bottom). mRNAs showing a $\log_2$FC ←0.5 or > 0.5 were classified as downregulated or upregulated by NSP1, respectively. The cumulative fraction of $\log_2$FC values for mRNAs showing fast or slow ribosome transit for every dataset is shown. Mann–Whitney test for fast and slow groups is shown (**B**) Prediction of *down* and *up* classes by Random Forest classifier. The mean accuracy obtained was 0.81. The resulting three most important features captured by the model are shown. (**C**) Model of the differential effect of NSP1 on mRNA translation and stability. NSP1 binding prevents the penetration of mRNA through the entry channel of 40S (threading). However, lateral attachment of mRNA by slotting is still possible in the presence of NSP1. The resulting complex can proceed efficiently for mRNAs with G-low 5′ UTR. However, for mRNAs with eIF4A binding motifs (Gs or GAs) near the cap, slotting can result in an improper accommodation of the mRNA in the 40S groove that blocks translation initiation. When combined with optimal codon usage of the CDS, NSP1-mediated translation blockade can also result in mRNA degradation.

tion or by NSP1 expression (*up* group) showed the opposite. When comparing *down* and *up* groups, these differences were exacerbated (Supplementary Fig. S6A and B). Since tAIg of mRNAs influences ribosome elongation rates, we compared the sensitivity to NSP1 among mRNAs with predicted fast (short CDS and high tAIg) and slow (long CDS and low tAIg) ribosome transits. A dramatic difference in sensitivity to NSP1 was found between these two groups, especially in SARS-CoV-2-infected cells (Fig. 7A), showing that mRNA with fast ribosome transits were particularly sensitive to NSP1-mediated inhibition. However, mRNAs with longer CDS and lower tAIg than the median showed resistance or even had increased levels in the presence of NSP1. The most abundant SARS-CoV-2 mRNAs fall into this group, showing a TE 2–3-fold higher in the presence of NSP1(27).

We also compared the nucleotide composition of 5′ UTR among *down* and *up* mRNA groups by computing the number of purines (G alone or G + A) throughout the 5′ UTR using a sliding 5-nt window to detect mRNAs enriched in eIF4A-binding motifs. Remarkably, the mRNA *down* group showed a significant enrichment in G content from + 11 to + 18 position with respect to the mRNA 5′ cap (Supplementary Fig. S6A and B). This finding is in good agreement with data shown

in Fig. 4 using G-less-RL variants with Gs and GA motifs near the 5′ cap.

Next, we used supervised machine learning models to predict the sensitivity of human mRNAs to NSP1 based on the discovered features. Among all models tested, Random Forest classification generated the best performance using the dataset of Fisher *et al.* (2022) [27]. The model predicted *down* and *up* classes with a high accuracy (81%) by selecting CDS length, tAIg, and 5′ UTR length as features with the greatest importance in the classification process (Fig. 7B). The accuracy increased to 83% when features were transformed using principal component analysis before training the classifier. Taken together, these analyses and predictions support the model depicted in Fig. 7C.

Finally, we analyzed the effect of ribosome recycling on sensitivity of mRNAs to NSP1. To do this, we first silenced the expression of the ribosome recycling factor ABCE1 [63] in HEK293T by means of lentivirus expressing shRNAs (see the 'Materials and methods' section). ABCE1 knockdown reduced translation in a dose-dependent manner as expected (Supplementary Fig. S7A). Interestingly, the sensitivity of RL-3 mRNA to NSP1, but not of G-low-FFL mRNA, was significantly reduced in those transduced cells with the lowest levels

of ABCE1 (Supplementary Fig. S7A), showing that ABCE1-mediated ribosome recycling influences sensitivity of mRNA to NSP1. We also analyzed the effect of poly(A) in facilitating ribosome recycling and reinitiation by bridging the 5′ and 3′ extremes of mRNA according to the closed loop model [64]. For this, we tested the effect of recombinant NSP1 protein on translation of RL-3 mRNA with or without a poly(A) tail of 25 nt in an *in vitro* cell-free translation system (RRL supplemented with a HeLa S10 extract). Under these conditions, the presence of the poly(A) tail increased translation of RL-3 mRNA about two-fold (Supplementary Fig. S7B). As found in transfected cells, NSP1 WT drastically blocked translation of RL-3 + poly(A) mRNA *in vitro* (Supplementary Fig. S7B). Notably, poly(A) tail elimination reduced the sensitivity of RL-3 mRNA to NSP1 about three-fold (Supplementary Fig. S7B). Taken together, these results confirm that the effective recycling of ribosomes for reinitiation influences sensitivity of mRNA to NSP1.

## Discussion

Among all known viral proteins that interfere with host translation, SARS-CoV-2 NSP1 is unique in the way it blocks cellular translation initiation, which may also result in mRNA degradation [65]. Taken together, our results show that the effect of NSP1 on translation initiation depends on the length and nucleotide composition of the 5′ UTR, reflecting a differential mode of 43S-PIC recruitment to mRNAs, whereas NSP1-mediated mRNA degradation is strongly influenced by codon usage bias. We propose the existence of both threading and slotting mechanisms of 43S-PIC recruitment to mRNA that are differentially affected by NSP1. According to our model, the presence of eIF4A-binding G/GA-motif near the 5′ cap promotes (or enforces) the threading of mRNA through the entry channel during 43S-PIC recruitment. When the mRNA entry channel of 43S-PIC is blocked by NSP1, eIF4A-bound mRNA could not be properly accommodated within the 40S groove, resulting in an aberrant 48S-PIC that may eventually undergo mRNA degradation (Fig. 7C); some previous data support this model. First, NSP1 binding to 40S did not prevent mRNA attachment so that the eventual mRNA degradation occurred on the 48S-PIC [36, 66, 67]. Second, eIF4A has been detected near the mRNA entry channel in 48S-PIC assembled with alphavirus or synthetic mRNAs [54, 68]. Third, the presence of eIF4A as part of the eIF4F complex increased the sensitivity to NSP1 *in vitro* even for those mRNAs that do not require eIF4F for translation [36]. Fourth, insertion of Gs in the SL1 of 5′ L of SARS-CoV-2 mRNA increased its sensitivity to NSP1, although this finding was not initially related to eIF4A binding [35, 40, 41].

When combined with a threading-prone 5′ UTR, optimal codon usage exacerbated NSP1-induced mRNA degradation. In our opinion, the simplest explanation for this is the positive influence that fast elongating CDS could have on translation initiation, feeding mRNA with new ribosomal subunits and eIFs for reinitiation that amplifies the inhibitory signal of NSP1 on target mRNAs. This idea is further supported by the findings that mRNAs with predicted fast ribosome transit showed a dramatic sensitivity to NSP1-induced mRNA degradation, and that maximum inhibition by NSP1 was achieved on polyadenylated mRNA that also required physiological levels of recycling factor ABCE1. However, it is important to note that optimal codon usage did not promote NSP1-mediated mRNA degradation when combined with a slotting-prone 5′ UTR. This shows that susceptibility of a given mRNA to NSP1-induced degradation is influenced by the way 43S-PIC is recruited to the mRNA. Although our data do not reveal the mechanistic details of NSP1-induced mRNA degradation, the fact that NSP1 can also promote stabilization of some mRNAs (e.g. SL50-FFL and G-less-RL) suggests that beyond an intrinsic activity of NSP1 bound to 40S, mRNA degradation could be the result of a safeguard activity associated with 48S-PIC to remove improperly accommodated mRNA within the 40S groove. This activity would require eIF4A (and perhaps DDX3) bound to mRNA and the participation of eIF3G as described recently [36].

According to the structural model of 40S-NSP1 complex, lateral accommodation of mRNA by slotting is sterically possible when mRNA entry channel is blocked by NSP1 [29, 30]. Our data strongly suggest that NSP1 allowed or even stimulated the slotting of 43S-PIC on G-less/low 5′ UTR mRNAs, including those of SARS-CoV-2 bearing the 5′ L. Despite they are unrelated in sequence, the striking compositional and functional similarities between G-less/low 5′ UTR and viral 5′ L are suggestive of an evolutionary adaptation of SARS-CoV-2 to exploit the ability of 43S-PIC to be slotted on certain mRNAs. Our results complement previous studies showing the participation of the 5′-proximal stem loop (SL1) in the translational resistance of 5′ leader mRNA to NSP1 [25, 42]. Perhaps both the structural composition and placement of SL1 are favoring the slotting of 43S-PIC on 5′ L during translation initiation of SARS-CoV-2 mRNAs. Thus, the combination of the slotting-prone activity of coronaviral 5′ L mRNA together with a suboptimal codon usage of CDS would provide an effective mechanism to escape the effect of NSP1. However, unlike reporter genes with G-less/low 5′ UTR that are translated less efficiently, 5′ L of SARS-CoV-2 directed efficient translation to reporter mRNA. We propose that the presence of G-poor secondary structure of moderate stability as SL1 in the 5′ L of SARS-CoV-2 mRNAs provides an optimal combination of TE and resistance to NSP1 as anticipated in previous studies [35, 40, 41]. Further studies on these adaptive trends could be useful to better understand long-term evolutionary trajectories of Coronavirus and other viruses infecting humans that show similar trends.

## Supplementary data

Supplementary data is available at NAR online.

## Conflict of interest

## Funding

## Data availability

The data underlying this article has been deposited in Zenodo under DOI https://doi.org/10.5281/zenodo.15045446.

## References

1. Brito Querido J, Díaz-López I, Ramakrishnan V. The molecular basis of translation initiation and its regulation in eukaryotes. *Nat Rev Mol Cell Biol* 2024;**25**:168–186. https://doi.org/10.1038/s41580-023-00624-9

2. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol Mol Biol Rev* 2011;**75**:434–67. https://doi.org/10.1128/MMBR.00008-11

3. Jackson RJ, Hellen CUT, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 2010;**11**:113–27. https://doi.org/10.1038/nrm2838

4. Pelletier J, Sonenberg N. The Organizing principles of eukaryotic ribosome recruitment. *Annu Rev Biochem* 2019;**88**:307–35. https://doi.org/10.1146/annurev-biochem-013118-111042

5. Kozak M. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol Cell Biol* 1989;**9**:5134–42.

6. Leppek K, Das R, Barna M. Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* 2018;**19**:158–74. https://doi.org/10.1038/nrm.2017.103

7. Pestova TV, Kolupaeva VG. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* 2002;**16**:2906–22.

8. Kumar P, Hellen CUT, Pestova TV. Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev* 2016;**30**:1573–88.

9. Soto-Rifo R, Rubilar PS, Limousin T *et al.* DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *EMBO J* 2012;**31**:3745–56. https://doi.org/10.1038/emboj.2012.220

10. Gupta N, Lorsch JR, Hinnebusch AG. Yeast Ded1 promotes 48S translation pre-initiation complex assembly in an mRNA-specific and eIF4F-dependent manner. *eLife* 2018;**7**:e38892. https://doi.org/10.7554/elife.38892

11. Svitkin YV, Pause A, Haghighat A *et al.* The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5′ secondary structure. *RNA* 2001;**7**:382–94. https://doi.org/10.1017/S135583820100108X

12. Brito Querido J, Sokabe M, Kraatz S *et al.* Structure of a human 48S translational initiation complex. *Science* 2020;**369**:1220–7. https://doi.org/10.1126/science.aba4904

13. Giess A, Torres Cleuren YN, Tjeldnes H *et al.* Profiling of small ribosomal subunits reveals modes and regulation of translation initiation. *Cell Rep* 2020;**31**:107534. https://doi.org/10.1016/j.celrep.2020.107534

14. Díaz-López I, Toribio R, Berlanga JJ *et al.* An mRNA-binding channel in the ES6S region of the translation 48S-PIC promotes RNA unwinding and scanning. *eLife* 2019;**8**:e48246. https://doi.org/10.7554/eLife.48246

15. Gu Y, Mao Y, Jia L *et al.* Bi-directional ribosome scanning controls the stringency of start codon selection. *Nat Commun* 2021;**12**:6603–4. https://doi.org/10.1038/s41467-021-26923-3

16. Presnyak V, Alhusaini N, Chen Y-H *et al.* Codon optimality is a major determinant of mRNA stability. *Cell* 2015;**160**:1111–24. https://doi.org/10.1016/j.cell.2015.02.029

17. Hanson G, Coller J. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol* 2018;**19**:20–30. https://doi.org/10.1038/nrm.2017.91

18. Barrington CL, Galindo G, Koch AL *et al.* Synonymous codon usage regulates translation initiation. *Cell Rep* 2023;**42**:113413. https://doi.org/10.1016/j.celrep.2023.113413

19. Wu Q, Bazzini AA. Translation and mRNA stability control. *Annu Rev Biochem* 2023;**92**:227–45.

20. Simmonds P. Rampant C→U hypermutation in the genomes of SARS-CoV-2 and other coronaviruses: causes and consequences for their short- and long-term evolutionary trajectories. *mSphere* 2020;**5**:e00408–20.

21. Markov PV, Ghafari M, Beer M *et al.* The evolution of SARS-CoV-2. *Nat Rev Mol Cell Biol* 2023;**21**:361–79.

22. Ratcliff J, Simmonds P. Potential APOBEC-mediated RNA editing of the genomes of SARS-CoV-2 and other coronaviruses and its impact on their longer term evolution. *Virology* 2021;**556**:62–72.

23. Mogro EG, Bottero D, Lozano MJ. Analysis of SARS-CoV-2 synonymous codon usage evolution throughout the COVID-19 pandemic. *Virology* 2022;**568**:56–71.

24. Schubert K, Karousis ED, Ban I *et al.* Universal features of Nsp1-mediated translational shutdown by coronaviruses. *Mol Cell* 2023;**83**:3546–57.

25. Karousis ED, Schubert K, Ban N. Coronavirus takeover of host cell translation and intracellular antiviral response: a molecular perspective. *EMBO J* 2024;**43**:151–67.

26. Finkel Y, Gluck A, Nachshon A *et al.* SARS-CoV-2 uses a multipronged strategy to impede host protein synthesis. *Nature* 2021;**594**:240–5.

27. Fisher T, Gluck A, Narayanan K *et al.* Parsing the role of NSP1 in SARS-CoV-2 infection. *Cell Rep* 2022;**39**:110954.

28. Burke JM, St Clair LA, Perera R *et al.* SARS-CoV-2 infection triggers widespread host mRNA decay leading to an mRNA export block. *RNA* 2021;**27**:1318–29.

29. Schubert K, Karousis ED, Jomaa A *et al.* SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat Struct Mol Biol* 2020;**27**:959–66.

30. Thoms M, Buschauer R, Ameismeier M *et al.* Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* 2020;**369**:1249–55.

31. Yuan S, Peng L, Park JJ *et al.* Nonstructural protein 1 of SARS-CoV-2 is a potent pathogenicity factor redirecting host protein synthesis machinery toward viral RNA. *Mol Cell* 2020;**80**:1055–66. https://doi.org/10.1016/j.molcel.2020.10.034

32. Kamitani W, Huang C, Narayanan K *et al.* A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat Struct Mol Biol* 2009;**16**:1134–40.

33. Slavin M, Zamel J, Zohar K *et al.* Targeted in situ cross-linking mass spectrometry and integrative modeling reveal the architectures of three proteins from SARS-CoV-2. *Proc Nat Acad Sci USA* 2021;**118**:e2103554118. https://doi.org/10.1073/pnas.2103554118

34. Mendez AS, Ly M, Gonzalez-Sanchez AM *et al.* The N-terminal domain of SARS-CoV-2 nsp1 plays key roles in suppression of cellular gene expression and preservation of viral gene expression. *Cell Rep* 2021;**37**:109841. https://doi.org/10.1016/j.celrep.2021.109841

35. Bujanic L, Shevchuk O, von Kugelgen N *et al.* The key features of SARS-CoV-2 leader and NSP1 required for viral escape of NSP1-mediated repression. *RNA* 2022;**28**:766–79. https://doi.org/10.1261/rna.079086.121

36. Abaeva IS, Arhab Y, Miscicka A *et al. In vitro* reconstitution of SARS-CoV-2 Nsp1-induced mRNA cleavage reveals the key roles of the N-terminal domain of Nsp1 and the RRM domain of eIF3g. *Genes Dev* 2023;**37**:844–60.

37. Havkin-Solomon T, Itzhaki E, Joffe N *et al.* Selective translational control of cellular and viral mRNAs by RPS3 mRNA binding. *Nucleic Acids Res* 2023;**51**.4208–22. https://doi.org/10.1093/nar/gkad269

38. Banerjee AK, Blanco MR, Bruce EA *et al.* SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. *Cell* 2020;**183**:1325–39. https://doi.org/10.1016/j.cell.2020.10.004

39. Vora SM, Fontana P, Mao T *et al.* Targeting stem-loop 1 of the SARS-CoV-2 5′ UTR to suppress viral translation and Nsp1 evasion. *Proc Nat Acad Sci USA* 2022;**119**:e2117198119. https://doi.org/10.1073/pnas.2117198119

40. Chen S-C, Xu C-T, Chang C-F *et al.* Optimization of 5′UTR to evade SARS-CoV-2 nonstructural protein 1-directed inhibition of protein synthesis in cells. *Appl Microbiol Biotechnol* 2023;**107**:2451–68. https://doi.org/10.1007/s00253-023-12442-2

41. Slobodin B, Sehrawat U, Lev A *et al.* Cap-independent translation and a precisely located RNA sequence enable SARS-CoV-2 to control host translation and escape anti-viral response. *Nucleic Acids Res* 2022;**50**.8080–92. https://doi.org/10.1093/nar/gkac615

42. Tidu A, Janvier A, Schaeffer L *et al.* The viral protein NSP1 acts as a ribosome gatekeeper for shutting down host translation and fostering SARS-CoV-2 translation. *RNA* 2020;**27**:253–64. https://doi.org/10.1261/rna.078121.120

43. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 2002;**30**:e36. https://doi.org/10.1093/nar/30.9.e36

44. Athey J, Alexaki A, Osipova E *et al.* A new and updated resource for codon usage tables. *BMC Bioinformatics* 2017;**18**:391–7. https://doi.org/10.1186/s12859-017-1793-7

45. Puigbo P, Bravo IG, Garcia-Vallve S. CAIcal: a combined set of tools to assess codon usage adaptation. *Biol Direct* 2008;**3**:38. https://doi.org/10.1186/1745-6150-3-38

46. Narula A, Ellis J, Taliaferro JM *et al.* Coding regions affect mRNA stability in human cells. *RNA* 2019;**25**:1751–64. https://doi.org/10.1261/rna.073239.119

47. Forrest ME, Pinkard O, Martin S *et al.* Codon and amino acid content are associated with mRNA stability in mammalian cells. *PLoS One* 2020;**15**:e0228730. https://doi.org/10.1371/journal.pone.0228730

48. Rodriguez A, Wright G, Emrich S *et al.* %MinMax: a versatile tool for calculating and comparing synonymous codon usage and its impact on protein folding. *Protein Sci* 2018;**27**:356–62. https://doi.org/10.1002/pro.3336

49. Angulo I, Acebrón I, De Las Rivas B *et al.* High-resolution structural insights on the sugar-recognition and fusion tag properties of a versatile-trefoil lectin domain from the mushroom laetiporus sulphureus. *Glycobiology* 2011;**21**.1349–61. https://doi.org/10.1093/glycob/cwr074

50. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;**11**:R106–r106. https://doi.org/10.1186/gb-2010-11-10-r106

51. Gandin V, Masvidal L, Hulea L *et al.* nanoCAGE reveals 5′ UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs. *Genome Res* 2016;**26**:636–48. https://doi.org/10.1101/gr.197566.115

52. Abraham A, Pedregosa F, Eickenberg M *et al.* Machine learning for neuroimaging with scikit-learn. *Front Neuroinform* 2014;**8**:14. https://doi.org/10.3389/fninf.2014.00014

53. Naineni SK, Itoua Maiga R, Cencic R *et al.* A comparative study of small molecules targeting eIF4A. *RNA* 2020;**26**:541–9. https://doi.org/10.1261/rna.072884.119

54. Brito Querido J, Sokabe M, Diaz-Lopez I *et al.* The structure of a human translation initiation complex reveals two independent roles for the helicase eIF4A. *Nat Struct Mol Biol* 2024;**31**:455–64.

55. Bahiri-Elitzur S, Tuller T. Codon-based indices for modeling gene expression and transcript evolution. *Comput Struct Biotechnol J* 2021;**19**:2646–63. https://doi.org/10.1016/j.csbj.2021.04.042

56. Elfakess R, Sinvani H, Haimov O *et al.* Unique translation initiation of mRNAs-containing TISU element. *Nucleic Acids Res* 2011;**39**:7598–609. https://doi.org/10.1093/nar/gkr484

57. Haimov O, Sinvani H, Martin F *et al.* Efficient and accurate translation initiation directed by TISU involves RPS3 and RPS10e binding and differential eukaryotic initiation factor 1A regulation. *Mol Cell Biol* 2017;**37**:e00150–17. https://doi.org/10.1128/MCB.00150-17

58. Iwasaki S, Floor SN, Ingolia NT. Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* 2016;**534**:558–61. https://doi.org/10.1038/nature17978

59. Schmidt T, Dabrowska A, Waldron JA *et al.* eIF4A1-dependent mRNAs employ purine-rich 5′UTR sequences to activate localised eIF4A1-unwinding through eIF4A1-multimerisation to facilitate translation. *Nucleic Acids Res* 2023;**51**:1859–79. https://doi.org/10.1093/nar/gkad030

60. Li K, Kong J, Zhang S *et al.* Distance-dependent inhibition of translation initiation by downstream out-of-frame AUGs is consistent with a Brownian ratchet process of ribosome scanning. *Genome Biol* 2022;**23**:251–4. https://doi.org/10.1186/s13059-022-02829-1

61. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Nat Acad Sci USA* 2004;**101**:11269–74. https://doi.org/10.1073/pnas.0400541101

62. Finkel Y, Mizrahi O, Nachshon A *et al.* The coding capacity of SARS-CoV-2. *Nature* 2021;**589**:125–30. https://doi.org/10.1038/s41586-020-2739-1

63. Hellen CUT. Translation termination and ribosome recycling in eukaryotes. *Cold Spring Harb Perspect Biol* 2018;**10**.a032656. https://doi.org/10.1101/cshperspect.a032656

64. Vicens Q, Kieft JS, Rissland OS. Revisiting the closed-loop model and the nature of mRNA 5′–3′ communication. *Mol Cell* 2018;**72**.805–12. https://doi.org/10.1016/j.molcel.2018.10.047

65. Rozman B, Fisher T, Stern-Ginossar N. Translation-A tug of war during viral infection. *Mol Cell* 2023;**83**:481–95. https://doi.org/10.1016/j.molcel.2022.10.012

66. Lapointe CP, Grosely R, Johnson AG *et al.* Dynamic competition between SARS-CoV-2 NSP1 and mRNA on the human ribosome inhibits translation initiation. *Proc Nat Acad Sci USA* 2021;**118**:e2017715118. https://doi.org/10.1073/pnas.2017715118

67. Tardivat Y, Sosnowski P, Tidu A *et al.* SARS-CoV-2 NSP1 induces mRNA cleavages on the ribosome. *Nucleic Acids Res* 2023;**51**:8677–90. https://doi.org/10.1093/nar/gkad627

68. Toribio R, Diaz-Lopez I, Boskovic J *et al.* Translation initiation of alphavirus mRNA reveals new insights into the topology of the 48S initiation complex. *Nucleic Acids Res* 2018;**46**:4176–87. https://doi.org/10.1093/nar/gky071