Structural and functional conservation of the programmed –1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2)

Received for publication, March 13, 2020, and in revised form, June 19, 2020. Published, Papers in Press, June 22, 2020, DOI 10.1074/jbc.AC120.013449

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

From the 1Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland, USA, the 2Department of Physics, University of Alberta, Edmonton, Alberta, Canada, and the 3School of Applied and Engineering Physics, Cornell University, Ithaca, New York, USA

Edited by Craig E. Cameron

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 coronavirus 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.

SARS-CoV-2, the etiological agent of COVID-19, is a member of the coronavirus family (1). Coronaviruses have (+) single-strand RNA genomes that harbor two long ORFs that occupy approximately two-thirds of the 5’ end of the genomic RNA (ORF1 and ORF2), followed by several ORFs that are expressed late in the viral replication cycle from subgenomic RNAs (Fig. 1A) (2). 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 (3, 4). –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 (5). A –1 PRF mechanism was shown to be required to translate ORF1ab in a coronavirus, avian infectious bronchitis virus, 2 years later (6). In coronaviruses, –1 PRF functions as a developmental switch, and mutations and small molecules that alter this process have deleterious effects on virus replication (7, 8).

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, WWZ is either AAA or UUU, and Z ≠ G. The linker region is 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. 9). 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...
ACCELERATED COMMUNICATIONS: Frameshifting in SARS-CoV-2

A

Immediate early genes

Early genes

Late genes

gRNA

B

SARS-CoV-2

SARS-CoV

Attenuator hairpin

slippery site

Frameshift-stimulating three stemmed pseudoknot

C

SARS-CoV

SARS-CoV-2

UUG3'

Stem 1

Stem 2

Stem 3

Slippery site

3-stemmed pseudoknot

D

SARS-CoV attenuator hairpin

Slippery site

SARS-CoV-2 attenuator hairpin

Slippery site

E

SARS-CoV-2 mutants

F

SARS-CoV mutants

Slippery site

Attenuator hairpin

Silent stem 1 mutant

Silent stem 2 mutant

Silent stem 3 mutant

ccUcAAu
Silent slippery site mutant

ccGcAAu
Silent slippery site mutant
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 (10). The SARS-CoV pseudoknot is more complex because it contains a third, internal stem-loop element (11–13). 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 (7, 8). 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 (14). 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, a 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 antishuffling activity of a small-molecule ligand against both frameshift signals.

Results

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 (15, 16) in cultured human cell lines. For both of the elements, −1 PRF activity was −20% in HEK (Fig. 2A) and −30% in Hela (Fig. 2B). 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. 2, 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 ± 0.03 and 0.7 ± 0.1% for SARS-CoV and SARS-CoV-2, respectively, p < 0.0001 (Fig. 2C). Similarly, disruption of stem 2 had a strong negative impact on −1 PRF, decreasing rates to 0.68 ± 0.04% for SARS-CoV and 0.8 ± 0.1% for SARS-CoV-2; p < 0.0001 (Fig. 2D). 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 ± 0.9 and 8 ± 1% for SARS-CoV and SARS-CoV-2, respectively; p < 0.0001) (Fig. 2E). These findings support the hypothesis that the structure and function of the core −1 PRF signals have been conserved between the two viruses.

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 (14). Although less well-conserved at the primary sequence level (Fig. 1, 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. 2G). 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

Figure 1. 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 (11) 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.
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 \(-1\) frameshifting, 2-\{4-(2-methyl-thiazol-4-ylmethyl)-[1,4]diazepane-1-carbonyl]-amino\}-benzoic acid ethyl ester, hereafter denoted as MTDB (17, 18). Comparing the \(-1\) frameshifting activity from Dual-Luciferase measurements in rabbit reticulocyte lysates in the presence and absence of MTDB, we found that 5 \(\mu M\) MTDB reduced \(-1\) frameshifting by almost 60%, from 36 ± 3 to 15 ± 1\% (Fig. 3). This reduction was comparable with, but slightly smaller than, that seen previously for the SARS-CoV pseudoknot, where 0.8 \(\mu M\) MTDB reduced \(-1\) frameshifting by roughly 60\% (17).

**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-2.
The difference between their scattering profiles is consistent with 0 at all q (Fig. 4B). The high-q portion of the profile is sensitive to the finer molecular details of the structure (19); hence the similarity of the profiles for the two pseudoknots indicates that their structures are likely the same. Because SARS-CoV pseudoknots can dimerize (20), 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. 4A, inset), we determined the monomer size, parameterized as the radius of gyration, $R_g$. We measured the same values for SARS-CoV and SARS-CoV-2 pseudoknots: $R_g = 28.1 \pm 0.3$ Å and $28.1 \pm 0.2$ Å, respectively. The difference profile for this set is also consistent with 0 for all q (Fig. 4C).

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 SARS-CoV-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 (7, 8), the pseudoknot will likely dimerize via interactions between loop 2 (20), and suppression of $-1$ PRF will most likely attenuate viral propagation (11). 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 (21). The fact that the compound MTDB, which was found in a computational search for $-1$ PRF inhibitors in SARS-CoV (17), 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.
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 (23). 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 S2.

**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-
structs was measured after incubating 20 μl of dual luciferase mixture lacking Leu and Met, and then incubated for 90 min, and then luminescence was measured 10 min after addition of 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 95°C 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 (24). SEC-SAXS was performed on a Superdex 200 increase 5/150 GL column (GE) equilibrated in the SAXS buffer condition.

**Data availability**

Full data sets of −1 PRF assays are available upon request. Please contact Dr. Jonathan D. Dinman, University of Maryland, dinman@umd.edu.

**Acknowledgments**—We thank Kevin Tu for help with reagent preparation and Shrish 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.


**Funding and additional information**—This work was supported by Defense Threat Reduction Agency Grant HDTRA1-13-1-0005; NIGMS, National Institutes of Health Grant R01 GM117177; a University of Maryland Coronavirus Research Program Seed Grant (to J. D. D.); Canadian Institutes of Health Research Grant O3-170709 (to M. T. W.); and National Institutes of Health Institutional Training Grant 2T32AI051967-06A1 (to J. A. K.). The solution scattering work was supported by National Science Foundation Award DBI-1930046 (to L. P.). Support for work performed at Beamline LIX (16ID) of the Center for BioMolecular Structure at National Synchrotron Light Source II was provided by National Institutes of Health Grants P30 GM133893, S10 OD012331, and BER-BO 070. National Synchrotron Light Source II is supported by Department of Energy Grant BES-FWP-PS001. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: SARS-CoV, severe acute respiratory syndrome coronavirus; SARS-CoV-2, SARS...
coronavirus 2; −1 PRF, programmed −1 ribosomal frameshift; nt, nucleotide(s); MTDB, 2′-[(4-(2-methyl-thiazol-4-ylmethyl)-[1,4]diazepane-1-carbonyl]-amino]-benzoic acid ethyl ester; SAXS, small and wide-angle X-ray scattering; SEC, size-exclusion chromatography.

References