RNA dimerization plays a role in ribosomal frameshifting of the SARS coronavirus

Daniella Ishimaru[1], Ewan P. Plant[2], Amy C. Sims[3], Boyd L. Yount Jr[3], Braden M. Roth[1], Nadukkudy V. Eldho[1], Gabriela C. Pérez-Alvarado[4], David W. Armbruster[1], Ralph S. Baric[3], Jonathan D. Dinman[5], Deborah R. Taylor[2] and Mirko Hennig[1,*]

[1]Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC 29425, USA, [2]Laboratory of Hepatitis and Related Emerging Agents, Division of Emerging and Transfusion-Transmitted Diseases, Food and Drug Administration, Bethesda, MD 20892, USA, [3]Departments of Epidemiology and Microbiology and Immunology, University of North Carolina, Chapel Hill, NC 27599, [4]Department of Chemistry and Biochemistry, Southern Illinois University, Carbondale, IL 62901, USA and [5]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA

ABSTRACT

Messenger RNA encoded signals that are involved in programmed -1 ribosomal frameshifting (-1 PRF) are typically two-stemmed hairpin (H)-type pseudoknots (pks). We previously described an unusual three-stemmed pseudoknot from the severe acute respiratory syndrome (SARS) coronavirus (CoV) that stimulated -1 PRF. The conserved existence of a third stem–loop suggested an important hitherto unknown function. Here we present new information describing structure and function of the third stem of the SARS pseudoknot. We uncovered RNA dimerization through a palindromic sequence embedded in the SARS-CoV Stem 3. Further *in vitro* analysis revealed that SARS-CoV RNA dimers assemble through 'kissing' loop–loop interactions. We also show that loop–loop kissing complex formation becomes more efficient at physiological temperature and in the presence of magnesium. When the palindromic sequence was mutated, *in vitro* RNA dimerization was abolished, and frameshifting was reduced from 15 to 5.7%. Furthermore, the inability to dimerize caused by the silent codon change in Stem 3 of SARS-CoV changed the viral growth kinetics and affected the levels of genomic and subgenomic RNA in infected cells. These results suggest that the homodimeric RNA complex formed by the SARS pseudoknot occurs in the cellular environment and that loop–loop kissing interactions involving Stem 3 modulate -1 PRF and play a role in subgenomic and full-length RNA synthesis.

INTRODUCTION

A novel coronavirus was responsible for the sudden epidemic, severe acute respiratory syndrome (SARS) outbreak, in 2003. Coronaviruses are positive-strand RNA viruses with large genomes [∼30 000 nucleotides (nt)] that serve as templates for translation of viral proteins and for replication. The production of proteins from these viral RNAs does not follow the usual rules governing translation. The first polyprotein encoded by open reading frame (ORF)1a, which encodes non-structural proteins, is defined by initiation and termination codons and is translated normally. Signals embedded within the RNA just before the termination codon of ORF1a redirect a fraction of translating ribosomes to bypass the stop codon and continue translation in the -1 reading frame, thus creating the larger ORF1ab polyprotein (1–3). These programmed -1 ribosomal frameshift (-1 PRF) stimulating signals are typically composed of a heptameric slippery site, on which the ribosome can change register by 1 nt in the 5′ direction, followed by a pseudoknot. Slippery

site sequence requirements have been characterized for several cell types (4) but the range and diversity of frameshift-stimulating pseudoknots continues to grow (5).

Most frameshift-stimulating pseudoknots are two-stemmed H-type structures. However, we and others have shown that the SARS coronavirus (SARS-CoV) -1 PRF signal is composed of three stems [Figure 1A; (1,2,6)]. Secondary structure predictions indicate that the potential to form the third stem is conserved among Group II coronavirus even though the RNA sequences themselves are not well conserved (1). Interestingly, removal of the third stem from the coronavirus frameshift signal still allows for frameshifting (1,2,6,7). Thus, it is not clear what the molecular role of the additional stem–loop (Stem 3–Loop 2 [S3L2]) is, and this requires further study.

Here we scrutinize features of the third stem of the SARS-CoV frameshift-stimulating pseudoknot that are important for RNA structure and frameshifting efficiency. We demonstrate the importance of the capping loop sequence in promoting Stem 3 stability and maintaining near–wild-type levels of frameshifting. Specifically, a hexanucleotide, self-complementary sequence in the loop capping Stem 3 raises the possibility that dimerization of the pseudoknot may play a role in viral lifecycle. While the palindromic sequence embedded in the SARS-CoV Stem 3 is not strictly conserved among severe acute respiratory syndrome–related coronaviruses (SARSr-CoV), genomic sequences from *Rhinolophus* Chinese horseshoe bats (SARSr-Rh-BatCoV) (8) that were previously identified as a natural reservoir of SARS-CoV–related CoV (9) also encode the hexanucleotide palindrome.

'Loop–loop kissing' interactions involving Watson–Crick base pairing between complementary RNA loops are common RNA-tertiary structural motifs and are used e.g. in retroviral replication (10–14). Here, the retroviral nucleocapsid proteins can chaperone conversion of the non-covalently linked kissing dimer to a more thermodynamically stable extended duplex-mediated dimer linkage in a structural rearrangement suggested to be associated with viral particle maturation (15). Pioneering work on Infectious Bronchitis Virus suggested that the genomic RNA (gRNA) of CoVs is not packaged as a dimer (16). However, CoV replication is mediated by the synthesis of a negative-strand RNA and also includes a discontinuous step involving synthesis of five to eight nested subgenomic RNA (sgRNA) intermediates (17).
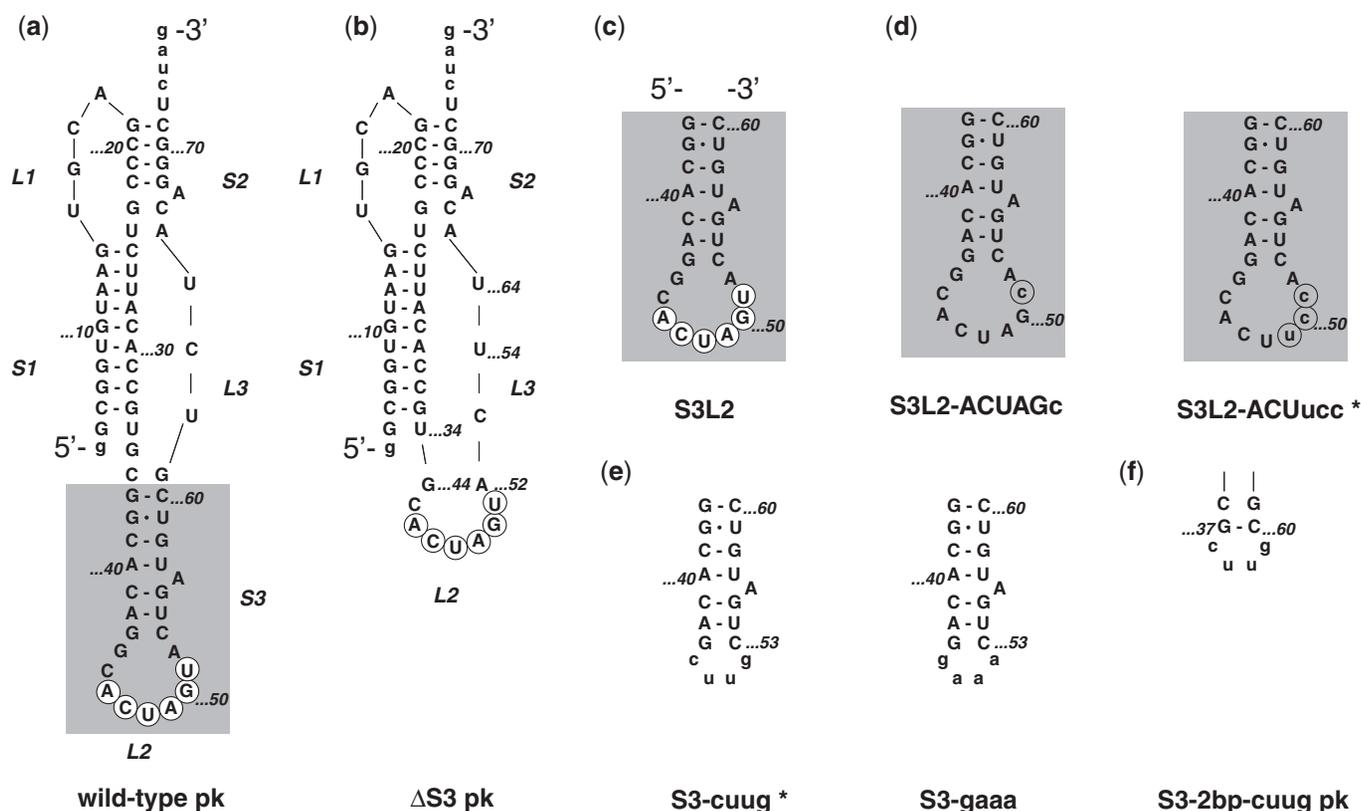

Figure 1. SARS constructs. (**a**) The three-stemmed wild-type SARS pseudoknot. Stems are labeled S1, S2 and S3 in the order that they occur 5′ to 3′ along the RNA. Accordingly, loops are labeled L1, L2 and L3. Note that L1 and L3 join adjacent stems, while L2 closes S3 (highlighted using gray box). Only the last two digits of the wild-type sequence numbering are used for clarity. The palindromic sequence 5′-ACUAGU-3′ embedded into L2 is indicated using white circles. Dashes represent Watson–Crick and the dot G•U Wobble base-pairing as confirmed by NMR spectroscopy. (**b**) Stem 3 deletion mutant ΔS3 pk. (**c**) S3L2 hairpin construct S3L2 spanning nucleotides G_{37} to C_{60}. (**d**) S3L2 hairpin constructs S3L2-ACUucc and S3L2-ACUAGc with L2 mutations that render the palindromic sequence asymmetrical (highlighted using gray circles) while conserving a Serine codon. (**e**) S3L2 hairpin constructs S3-cuug and S3-gaaa where the 9 nt L2 is replaced with the smaller tetraloops 5′-cuug-3′ and 5′-gaaa-3′, respectively. (**f**) SARS pseudoknot variants S3-2 bp-cuug pk with a shortened Stem 3 is capped with a 5′-cuug-3′ tetraloop. S3-cuug and S3L2-ACUucc variant constructs highlighted with an asterisk (*) were also generated in the context of full-length pk.

In the current work, we demonstrate that a previously overlooked loop–loop kissing interaction involving the conserved Stem–loop 3 embedded within the SARS-CoV pseudoknot occurs under physiological conditions *in vitro*. We further show that kissing dimer formation plays a role in frameshift-stimulation and modulates the relative abundance of full-length and subgenomic viral RNAs.

MATERIALS AND METHODS

Preparation of RNA samples

Plasmids containing wild-type pseudoknot as well as the ΔS3 pk mutant were described in Plant *et al* (1). The wild-type pk plasmid was used as a template for the generation of S3L2 mutants. S3L2-only transcripts incorporated a *cis*-acting, 3′-hammerhead ribozyme. This plasmid was the template for site-directed mutagenesis replacing Loop 2 5′-GCACTAGTA-3′ with GCACTAG cA, GCACTtccA, cttg or gaaa.

In vitro transcriptions were optimized and performed as described (1,18,19) using unlabeled nucleotide triphosphates (MP Biomedicals). RNA transcripts were purified by fast-performance liquid chromatography with HiTrap Q column (GE Healthcare) (18,19). S3L2-only transcripts and variants were purified with a HiTrap Q (GE Healthcare), followed by a DNAPac PA200 columns (Dionex) (18). Purified RNA was equilibrated with nuclear magnetic resonance (NMR) buffer [25 mM KCl unless otherwise noted, 10 mM sodium phosphate (pH 6.5), 500 μM EDTA (Ethylenediaminetetraacetic acid), 50 μM sodium azide, 9:1 $H_2O:D_2O$].

NMR spectroscopy

All NMR spectra were recorded at 288 K, 295 K or 298 K on Bruker Avance 900, 800, 700 or 600 MHz spectrometer equipped with either a triple resonance inverse detection cryoprobe (800 and 700) or standard triple resonance inverse detection-probeheads. NMR experiments were performed on samples of 500 μl volume containing 0.2–1.2 mM SARS-CoV pk and stem–loop S3L2 variant RNA. Data were processed using NMRPipe (20) and analysed using Sparky (21). One-dimensional imino proton spectra were acquired using a jump-return echo sequence (22). Imino resonances were assigned sequence specifically from water flip-back, WATERGATE 2D nuclear Overhauser effect spectroscopy (NOESY) spectra (τ_{mix} = 150–200 ms) (23). Typically, for the $^1H,^1H$-NOESY spectra, 384 complex points were recorded with an acquisition time of 24 ms for ^1H (ω_1), and 2048 complex points with an acquisition time of 127 ms for ^1H (ω_2). Repetition delays ranging from 1.2 (pk variants) to 1.7 s (S3L2 variants) were used between transients, with 128 scans per increment (total measuring times 36–49 h, respectively).

Dimerization assay and native gel electophoresis

Unlabeled (150 μM) and ^{32}P-labeled (1 pM) SARS-CoV pk and S3L2 RNA transcripts were annealed in NMR buffer unless stated otherwise. When 5 mM $MgCl_2$ was added, samples were incubated for 6 h at 37°C. Temperature, time of incubation and RNA concentration varied, as specified in the text. RNA samples were separated on 10% native polyacrylamide gel in Tris Borate buffer pH 8.3 at 4°C when $MgCl_2$ was added to the dimerization reaction. Otherwise, Tris Borate EDTA pH 8.3 was used and the gel analysis performed at 4°C or 25°C, as specified in the text. Gels were dried and analysed by phosphorimaging. In Figures 3A and 3B, gels were analysed by ethidium bromide staining.

Determination of crosslinking sites

RNA samples were incubated overnight at 37°C to promote dimerization. Samples were crosslinked in a ultraviolet (UV) crosslinker (Spectroline) for 30 min, 254 nm, 1.5 mW/cm^2 on ice (24). Crosslinked RNA was separated on denaturing PAGE (polyacrylamide gel electrophoresis) and eluted with 0.3 M sodium acetate, 2% sodium dodecyl sulfate (SDS), pH 5.5 at 42°C for 18 h. Precipitated RNA was subjected to partial alkaline hydrolysis. Control, non-crosslinked, ^{32}P-labeled RNA was subjected to RNase T1 digestion or partial alkaline hydrolysis (Applied Biosystems). All samples were separated on denaturing 15% PAGE. Gels were dried and densitometry of bands was determined using ImageQuant TL (GE Healthcare) (25).

Dual luciferase assays

Transfected VeroE6 cells were grown overnight in Dulbecco's Modified Eagle Medium supplemented with 10% Fetal Bovine Serum at 37°C. Cells were disrupted using the passive lysis buffer (Promega). Luminescence reactions were initiated by addition of 10–20 μl of cell lysates to 100 μl of the Promega LAR II buffer and completed by addition of 100 μl Stop-n-Glo reagent. Luminescence was measured using a Turner Design TD20/20. At least three replicates were performed within each assay, and all assays were repeated at least three times until the data were normally distributed (26). The frequency of frameshifting is expressed as a ratio of firefly to *Renilla* luciferase from a test plasmid divided by the analogous ratio from the read-through control plasmid multiplied by 100%. Fold change, standard error and estimates of the *P*-values for ratiometric analyses were performed as previously described (26).

Infection and quantitation of viral titer

VeroE6 cells inoculated with SARS-CoV or S3L2-ACU ucc pk (multiplicity of infection of 1) or were mock-infected and incubated at 37°C. Media were harvested at 0, 5, 8 or 12 h post-infection (pi) and titers assessed by plaque assay as previously published (27). Viral detection limit was 50 pfu/ml. Error bars are the standard deviation of three measurements.

Northern blot analysis

Total RNA from SARS-CoV and S3L2-ACUucc pk was isolated from infected cell monolayers (Trizol Invitrogen) and purified using Oligotex mRNA spin column reagents

(QIAGEN). RNA was separated on an agarose gel using Northern-Max-Gly (Ambion), transfered to a BrightStar-Plus membrane (Ambion) and cross-linked to the membrane with UV light. The blot was pre-hybridized and probed with a SARS-CoV nucleocapsid-specific oligodeoxynucleotide probe (5′-CTTGACTGCCGCCTC TGCT^{b}T^{b}CCCT^{b}CT^{b}GC^b-3′), where biotinylated nucleotides are designated with a superscript b. Blots were hybridized overnight and washed with low- and high-stringency buffers. Filters were incubated with streptavidin-AP, washed and then developed using the chemiluminescent substrate CDP-STAR (New England Biolabs).

Western blot analysis

At 5, 8 and 12 h pi, SARS-CoV and S3L2-ACUucc pk infected or mock-infected cells were washed and lysed in buffer containing 20 mM Tris–HCl (pH 7.6), 150 mM NaCl, 0.5% deoxycholine, 1% Nonidet P-40, 0.1% SDS and post-nuclear supernatants added to an equal volume of 5 mM EDTA and 0.9% SDS, resulting in a final SDS concentration of 0.5%. Samples were heat inactivated twice before usage. On 4–20% Criterion gradient gels (Bio-Rad), 10 μg of protein was loaded and transferred to a polyvinylidene difluoride membrane. Blots were probed with polyclonal rabbit antisera directed against nsp1 (diluted 1:500) or nsp16 (diluted 1:200) (28) and developed using enhanced chemiluminescence reagents (Amersham Biosciences).

Sequence comparisons and prediction of Stem 3/Loop 2 secondary structures

All CoV genome sequences were obtained from Gen Bank (http://www.ncbi.nlm.nih.gov/genbank/). Accession numbers for sequences discussed are summarized in Supplementary Table S1. Multiple sequence alignments were performed using ClustalW2 (version 2.0.12) (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (29). Free energies of the proposed S3L2 structures at 37°C in 1 M NaCl were calculated with Mfold (30).

RESULTS

Stem 3 of SARS pseudoknot forms an autonomous substructure

Our previous NMR analysis of exchangeable imino protons of the SARS-CoV pseudoknot (Figure 1A, wild-type pk) provided unequivocal evidence for the existence of Stem 3 (1). In the present study, secondary structure analysis by NMR provided further insight into the complex global architecture of the wild-type pk. To establish experimentally whether Stem 3 interacts with the two-stemmed H-type structure, we prepared a transcript lacking the base-paired region of Stem 3 while retaining the hexanucleotide 5′-ACU AGU-3′ palindromic sequence (Figure 1B, ΔS3 pk) and compared this construct with wild-type pk. The imino NOESY spectrum of the ΔS3 pk transcript is virtually identical to the wild-type pk native construct with a few marked exceptions. The missing sequential imino assignment path

for the mutant ΔS3 pk is indicated by dashed lines in the superposition of wild-type pk and ΔS3 pk NOESY spectra shown in Figure 2A. For clarity, these nuclear Overhauser effect (NOE) connectivities including the characteristic G38-U59 wobble pair are highlighted in the schematic secondary structure of the three-stemmed pseudoknot structure (dashed box) and correspond to Stem 3. No significant chemical shift or linewidth changes can be observed for imino protons located outside of Stem 3. The only notable exception is the severely broadened crosspeak for the G14-C25 base-pair in the ΔS3 pk construct located in the vicinity of the Stem 1–Stem2 junction (Figure 2A, red box). A possible explanation for this observation is that a longer unpaired Loop 3 consisting of 12 (ΔS3 pk) rather than 3 nt (wild-type pk) may affect the degree of overrotation at the S1–S2 junction. Addressing this question in more detail would require a complete structure determination. At this stage, the large intrinsic linewidth of a 67-nt wild-type pk RNA in combination with severe line-broadening observed for NMR samples concentrated to >200 μM made this procedure unrealistic (data not shown). However, the overall comparison of exchangeable imino proton spectra of the wild-type pk with ΔS3 pk suggests that Stem 3 does not noticeably engage in stable tertiary interactions involving the two-stemmed H-type structure and likely constitutes an autonomous substructure within the frameshift signal.

SARS pseudoknot dimerizes via palindromic sequence in Stem 3/Loop 2

The NMR line-broadening observed motivated a closer examination of the SARS pseudoknot and revealed a palindromic sequence in the loop capping Stem 3, designated Loop 2 (L2) (Figure 1A). To determine whether this palindromic sequence could mediate dimerization, Stem 3 transcripts (Figure 1C, S3L2) were incubated at 37°C for 30 min in the presence of KCl and subjected to native PAGE. Lane 1 of Figure 3A shows two bands and demonstrates that S3L2 transcripts form homodimers when analysed at 4°C. To evaluate the role of the palindromic sequence for dimer formation, a series of mutations were generated (Figure 1D–F) with the intention of altering the palindromic sequence (Figure 1D), replacing the entire Loop 2 (Figure 1E) or significantly reducing the size of both Stem 3 and Loop 2 (Figure 1F).

When the palindromic sequence was mutated from 5′-A CUAGU-3′ to ACUAGc or ACUucc, the resulting S3L2-A CUAGc and -ACUucc transcripts migrated as single species, indicating that dimer formation was abolished (Figure 3A, lanes 2 and 3). Unexpectedly, replacement of the entire Loop 2 (5′-GCACUAGUA-3′) with the stable tetraloop 5′-cuug-3′ generated a transcript migrating as a dimeric species (Figure 3A, lane 4). This could be explained by the formation of an extended duplex featuring two central U-U wobble mismatches and was further investigated using NMR methods. In contrast, a 5′-gaaa-3′ tetraloop containing transcript, S3-gaaa, efficiently prevented dimer formation *in vitro* (Figure 3A, lane 5).

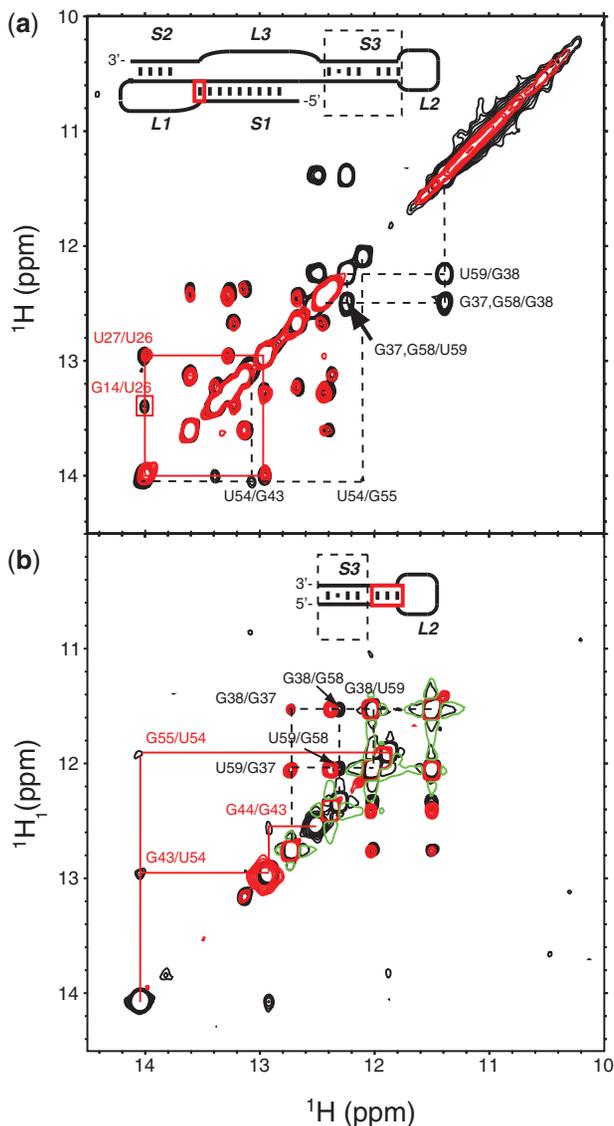To verify whether dimers could be observed in the context of the full-length pseudoknot, SARS wild-type

Figure 2. NMR secondary structure comparison of wild-type SARS-CoV pk, ΔS3 pk, S3 and S3L2-ACUucc mutants. (a) Imino regions of 2D ^{1}H,^{1}H-NOESY experiments collected on wild-type SARS-CoV pk (black contours) and ΔS3 pk mutant (red contours), respectively. Dashed black lines show the imino proton walk in the S3 stem. The base-paired region of S3 is deleted in the ΔS3 pk mutant; however, L2 is left intact. Solid red lines show the sequential NOE correlations involving the imino proton U26 located in S1, and the red box highlights the cross peak connecting imino protons U26 and G14 adjacent to the S1-S2 junction, which is absent in the ΔS3 pk mutant. Only the last two digits of the wild-type sequence numbering are used for clarity. The schematic SARS-CoV pk inset highlights the corresponding S3 stem (dashed box) as well as the G14-C25 basepair location in S1 (solid red box). (b) Imino regions of 2D ^{1}H,^{1}H-NOESY experiments collected on wild-type SARS S3 (black contours) and the S3L2-ACUucc mutant (red contours), respectively. Dashed black lines show the imino proton walk in the lower portion of the S3 stem. Solid red lines highlight the sequential cross peaks in the upper portion of S3 correlating imino protons G55, U54 and G43 adjacent to L2, which are broadened beyond detection in the S3L2-ACUucc mutant. The schematic SARS S3L2 inset highlights the corresponding lower S3 stem (dashed box) as well as the base-paired region in the upper S3 stem (solid red box).

pk transcripts were incubated in the same conditions described above. Full-length pk homodimers were observed (Figure 3C, lane 2) as demonstrated for S3L2 constructs (Figure 3C, lane 1).

To confirm that pk dimerization occurs via Loop 2, full-length pk, ΔS3 pk and S3-2bp-cuug pk were incubated with S3L2 transcripts (Figure 3C). Detection of heterodimers was performed by incubation of unlabeled pk with ^{32}P-labeled S3L2 (S3L2*) transcripts (Figure 3C, lane 5), where the radioactive-labeled transcripts showed mobility shifts compatible with both homo- and heterodimers. Similarly, ΔS3 pk, a construct that lacks Stem 3 but retains Loop 2, was found to self-associate (Figure 3C, lane 3) and to form heterodimers when incubated with S3L2 (Figure 3C, lane 6), albeit weakly. S3L2 was unable to form heterodimers with S3-2bp-cuug pk (Figure 3C, lane 7), a variant lacking sequence complementarity in Loop 2.

These results collectively indicate that the SARS pseudoknot can homodimerize *in vitro* via the palindromic sequence located in Loop 2. Because ΔS3 pk is able to self-associate *weakly*, we conclude that stable S3 formation facilitates dimerization but is not a requirement.

S3L2 self-association is modulated by temperature, time and Mg²⁺

The stability of loop–loop kissing interactions in mitochondrial transfer RNA (31) and viral RNA (32,33) are highly dependent on cation concentration. To test whether SARS S3L2 dimer formation is favored in the presence of Mg²⁺-ions (34), [^{32}P] 5′-end-labeled transcripts were incubated in the absence or presence of MgCl₂ at 37°C and separated by native PAGE. Comparison of Figures 4A and B shows that S3L2 self-association is a Mg²⁺-dependent event. In the absence of Mg²⁺, no appreciable change in dimer population was observed between 0 and 48 h at 37°C (Figure 4A). Next, we investigated S3L2 self-association as a function of increasing monovalent potassium cation concentration (Figure 4E). We found that S3L2 self-association was predominantly driven by the presence of Mg²⁺ because significant but slow dimerization resulted from the addition of 6–250 mM KCl in the absence of Mg²⁺ (Figure 4E, open circles), while responses to varying KCl concentrations were negligible in the presence of 5 mM Mg²⁺ (Figure 4E, closed circles).

To examine the influence of temperature on dimer formation, S3L2 RNAs were incubated in the presence of MgCl₂ at 25°C and 37°C (Figures 4C and 4D). Aliquots at various time-points were collected and stored at −20°C until all samples were collectively analysed by native PAGE. As shown in Figure 4D, S3L2 dimers readily formed at physiological temperatures (T ½ = 48 ± 12 min), while dimer formation was considerably slower at room temperature (T ½ = 19.8 ± 6.5 h; Figure 4C). We also quantified the concentration dependence of S3L2 dimerization, in the presence of MgCl₂, and determined the dissociation constant for the dimer to be 2.6 ± 0.15 µM at 37°C (Figure 4F). Taken together these results suggest that Stem 3/Loop 2 readily dimerizes under physiological conditions, i.e. at 37°C in the presence of
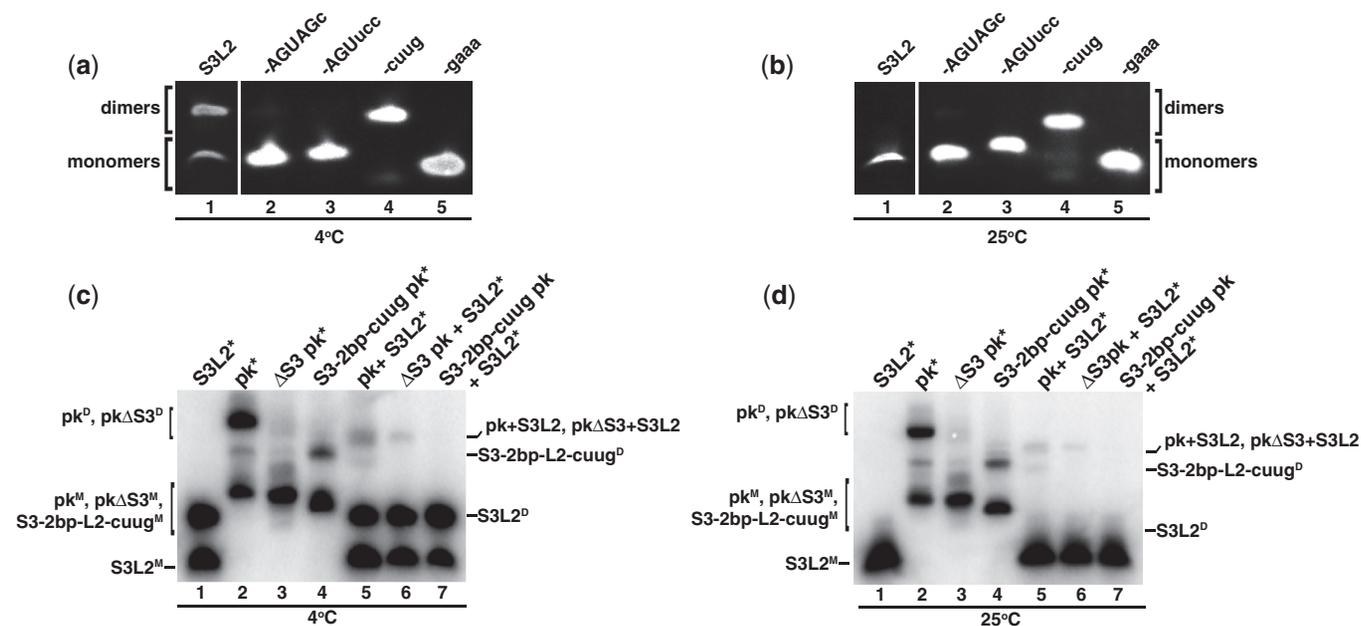
Figure 3. Detection of SARS RNA dimers. RNA transcripts (150 μM/each) were incubated at 37°C for 30 min in the presence of 200 mM KCl without MgCl$_2$. Samples were separated on a 10% native PAGE. **(a)** Native gel analysis at 4°C of S3L2-only variants: S3L2, S3L2-ACUAGc, S3L2-ACUucc, S3-cuug and S3-gaaa. Gels were stained with ethidium bromide and visualized by UV. S3L2 transcript was independently analysed. **(b)** Same as in (a) but at 25°C. **(c)** Native gel analysis at 4°C of the following transcripts: wild-type S3L2, wild-type pk, Stem 3 deletion mutant ΔS3 pk and S3-2 bp-cuug. RNA variants were incubated with ^{32}P-labeled transcripts in trace [1 pM] indicated by *. **(d)** Same as in (c), but at 25°C. For (c) and (d) samples were separated on a 10% native PAGE, which was dried and analysed by phosphorimaging. Uppercase 'M' denotes monomer and uppercase 'D', dimer.

5 mM MgCl$_2$, and that homodimers tolerate a broad range of ionic strengths.

S3L2 dimerization occurs via a loop–loop kissing interaction

Intermolecular loop–loop kissing of retroviral gRNA are initially metastable and subsequently converted to more stable mature duplexes catalysed by nucleocapsid proteins. Such processes involving palindromic 6-nt sequences have been extensively studied and described for Moloney murine leukemia virus (35), Hepatitis C virus (36) and Human immunodeficiency virus (HIV) dimers (15). We thus asked the question whether S3L2 loop–loop kissing complexes (Figure 5B) can potentially form extended duplexes (Figure 5C). Seminal work by Laughrea and Jetté have established that loose (loop–loop kissing) dimers are unstable when subjected to native electrophoresis at 25°C, while tight (extended) duplexes resist these conditions (37,38). As shown in lane 1 of Figures 3A and B, S3L2 dimers, while detectable at 4°C, are not favored at 25°C, suggesting the formation of loop–loop kissing complexes.

To further demonstrate that S3L2 dimerizes via L2-mediated loop–loop kissing interactions, UV-cross-linked dimers were subjected to partial hydrolysis. Normalized densitometry of bands in each lane revealed a significant reduction in ^{32}P-signal beginning at nucleotide 45 (Figure 6B), indicating that this nucleotide is in close proximity to the dimerization interface. The nucleotide at position 45 represents the adenosine 5′-neighbor of the 6-nt palindrome, consistent with SARS S3L2 loop–loop kissing

formation (Figure 5B) and cannot be explained on the basis of extended duplex formation (Figure 5C).

Detection of loop–loop kissing dimers for S3L2 transcripts prompted us to investigate whether pk, ΔS3 pk and S3-2bp-cuug pk homodimers formed loose or tight complexes. As shown in Figure 3D, all three homodimers remained stable when subjected to native gel at room temperature, an indication of tight dimer formation (37,38). However, as observed previously, only faint bands were detected for ΔS3 pk homodimers (Figures 3C and 3D, lane 3).

Altogether, these results suggest that S3L2 initially forms loop–loop kissing dimers and, when embedded into larger SARS pk constructs, can mature to form tight dimers. Extended duplex formation can also be observed in case of constructs capped by stable 5′-cuug-3′ tetraloops.

Dimerization monitored by NMR

To obtain direct structural information about the observed homodimers, a series of RNA transcripts were further investigated by NMR. Watson–Crick base pairing for the isolated Stem 3 RNA was confirmed by two-dimensional nuclear NOESY spectra and a sequential imino walk is indicated in Figure 2B. As compared with the native S3L2 construct, a number of exchangeable imino resonances disappeared at 283 K in the NOESY spectrum of the asymmetric S3L2-ACUucc Loop 2 mutant. The missing sequential imino assignment path for the mutant S3L2-ACUucc corresponds to the upper portion of Stem 3 and is indicated by solid red lines in the superposition of S3L2 and S3L2-ACUucc imino NOESY
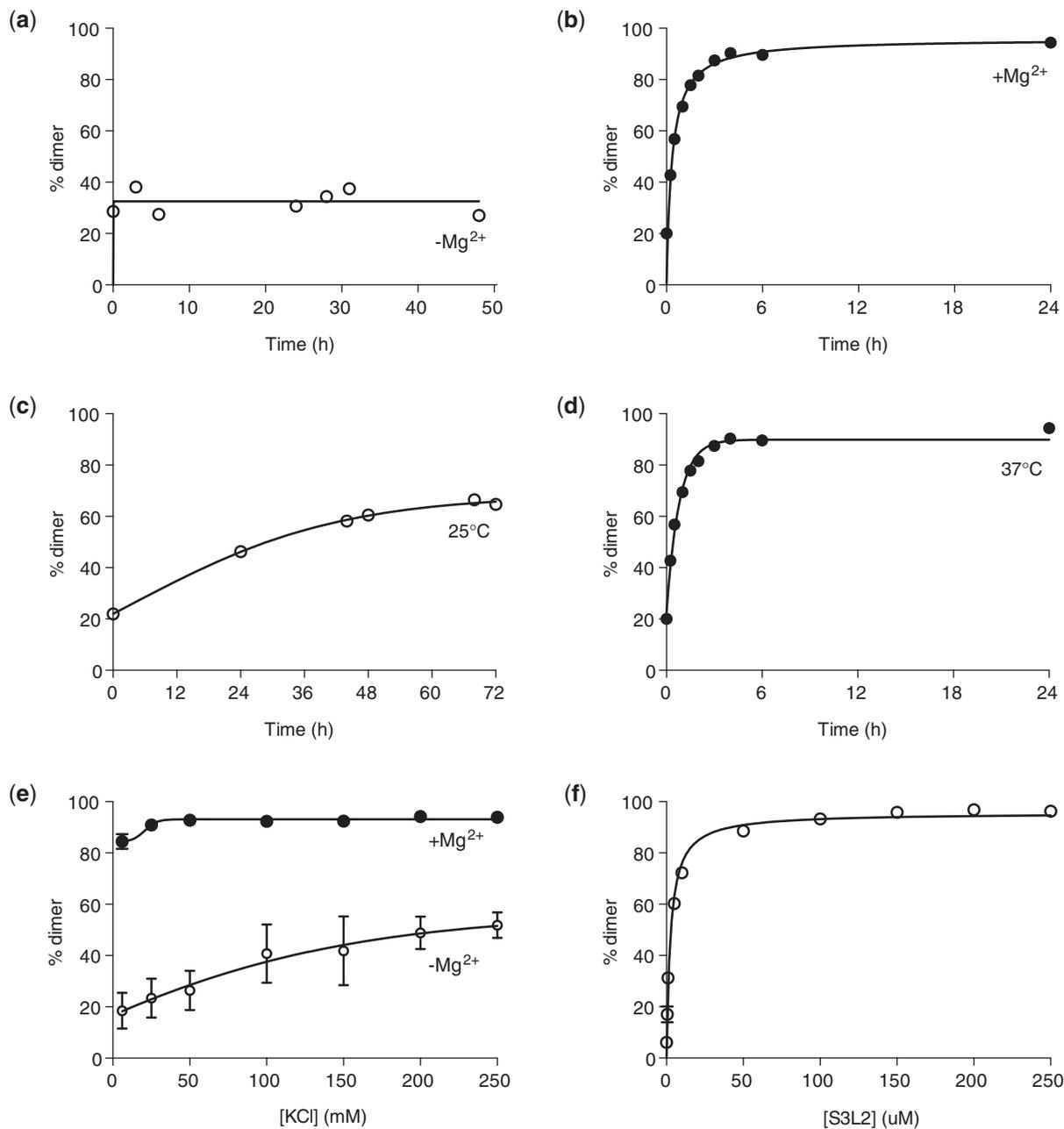
Figure 4. Evaluation of dimer-promoting conditions for SARS S3L2. Wild-type SARS S3L2 transcripts (150 µM) were incubated in 10 mM Na_2HPO_4, pH 6.5, 0.5 mM EDTA, 100 mM KCl and 5 mM $MgCl_2$ for 6 h at 37°C, unless otherwise stated. Samples were separated on a native PAGE in Tris Borate buffer. Gels were dried and analysed by phosphorimaging. (a) RNAs were incubated in the absence of $MgCl_2$. Aliquots were removed at the following time-points: 0 h, 3 h, 6 h, 24 h, 28 h, 31 h and 48 h. (b) Same as in (a) but in the presence of 5 mM $MgCl_2$. Aliquots were removed at the following time-points: 0 min, 15 min, 30 min, 1 h, 1:30 h, 2 h, 3 h, 4 h, 6 h and 24 h. (c) RNAs were incubated at 25°C in the presence of 5 mM $MgCl_2$. Aliquots were removed at the following time-points: 0 h, 24 h, 44 h, 48 h, 68 h and 72 h. (d) Same as in (c) but at 37°C. Aliquots were removed at the following time-points: 0′, 15′, 30′, 1 h, 1:30 h, 2 h, 3 h, 4 h, 6 h and 24 h. (e) RNAs were incubated in the absence (open symbols) or in the presence of 5 mM $MgCl_2$ (closed symbols) and varying concentrations of KCl (6–250 mM). (f) RNAs (0.1–250 µM) were incubated at 37°C for 6 h in the presence of $MgCl_2$.

spectra shown in Figure 2B. No significant chemical shift or linewidth changes can be observed for imino protons located in the lower part of Stem 3. For clarity, NOE connectivities including the characteristic G38-U59 wobble pair are highlighted in the schematic secondary structure of the isolated Stem 3 structure (dashed box). Due to the observed differences in the upper Stem 3

portion, it is concluded that while the S3L2-ACUucc adopts a hairpin structure, it features an open Loop 2 that destabilizes base-pairing interactions in the upper part of Stem 3.

Because kinetics associated with loop–loop kissing formation in the presence of Mg^{2+}-cations at 25°C are too fast at RNA concentrations required for NMR
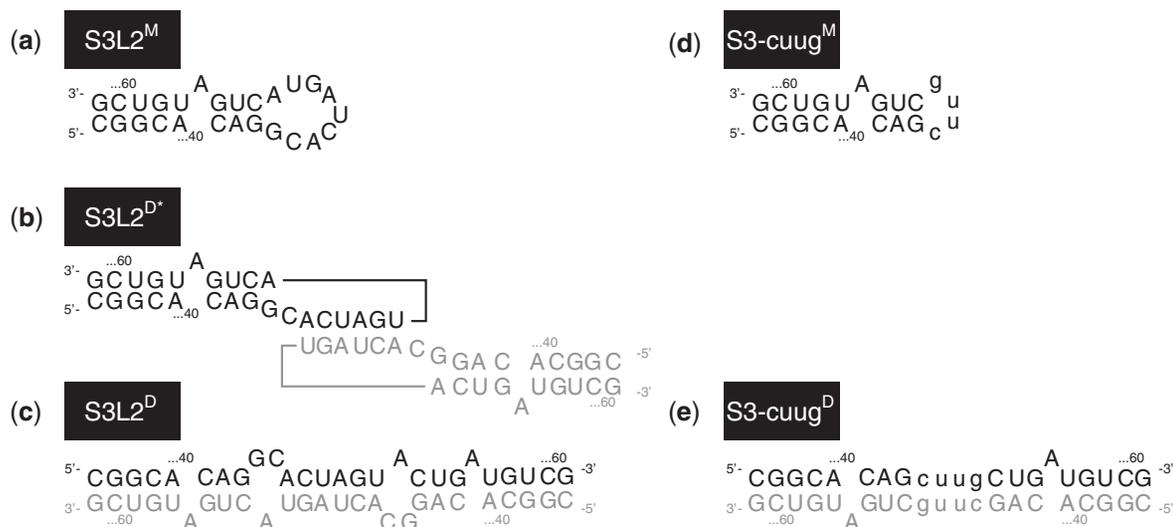
Figure 5. Schematic structures of SARS RNA dimers. **(a)** Stem–loop sequence and secondary structure for Stem 3 from SARS-CoV (S3L2^M). **(b)** Schematic representation of the Stem 3 loop–loop kissing homodimer, with individual stem–loops shown using black and gray letters, respectively (S3L2^D*). **(c)** Schematic of an extended S3 duplex, with individual stem–loops shown using black and gray letters, respectively (S3L2^D). **(d)** Stem–loop sequence and secondary structure for the 5′-cuug-3′-capped Stem 3 variant (S3-cuug^M). **(e)** Schematic of an extended S3-cuug duplex, with individual stem–loops shown using black and gray letters, respectively (S3-cuug^D).

investigation, we attempted to monitor the conversion by NMR with nucleotide-specific resolution in the presence of only potassium chloride. Therefore, we recorded a series of consecutive one-dimensional jump-return echo NMR spectra (Figure 7) to monitor S3L2 imino proton resonances in real-time. The overlapping G43 and U48 as well as G44 and G58 imino proton resonance assignments were confirmed using 2D heteronuclear ^{1}H,^{15}N-HMQC correlations (Supplementary Figure S1). Dimer formation at 25°C was induced by addition of 125 mM potassium chloride to an NMR sample containing 250 µM S3L2 RNA and (interrupted) 1D spectra recorded over a period of >20 days. Figure 7 shows that half of the kissing complex is formed after ~9 h and conversion progresses for >20 days before reaching a plateau characterized by a S3L2:(S3L2)$_2$ ratio of 2:8. No evidence for further progression to a potential extended duplex structure could be detected.

Curiously, the replacement of the 9-nt loop L2 with the stable 5′-cuug-3′ tetraloop (S3-cuug, Figure 1E) designed to stabilize hairpin formation generated a transcript that efficiently dimerized (Figure 3A, lane 2). The imino assignments through the stem could be easily followed based on the resonance assignments and NOE patterns observed in S3L2 (Supplementary Figure S2). Close inspection of the sequential assignments and NOE patterns revealed the existence of tandem U-U wobble pairs involving the two uridine nucleotides of the tetraloop. In addition, substantial line broadening was observed consistent with extended duplex formation of S3-cuug transcripts (Supplementary Figure S2). Subsequently, we examined the NOESY spectra of the larger, three-stemmed 5′-cuug-3′-tetraloop containing SARS mutants S3-2bp-cuug (Figure 1F) and S3-cuug pk that were subjected to functional frameshifting analysis. We confirmed the formation of the tandem U-U wobble pairs by 2D NOESY. The NOE patterns for the

U-U wobbles are an almost perfect subspectrum of the same region of the NOESY for the isolated S3-cuug transcripts (Supplementary Figure S3). Other exchangeable protons in the Stem 1 and Stem 2 regions comprising the H-type pseudoknot overlay with the corresponding imino protons observed in the context of wild-type pk. The excellent agreement in the overlay of the imino connectivities traced by the NOESY experiments indicates that the stable 5′-cuug-3′ tetraloop capping an otherwise wild-type Stem 3 facilitates the formation of extended duplex structures in the context of larger S3-2bp-cuug and S3-cuug pk mutants and locally retains the same structure as S3-cuug in isolation.

In contrast to the SARS S3L2 variants capped with 5′-cuug-3′ tetraloops, S3-gaaa (Figure 1E) NMR samples showed pronounced differences and narrower imino proton linewidths. Based on NOESY spectra and the observation of an additional upfield-shifted guanosine resonance located in the sheared G-A loop basepair, it was concluded that a construct containing a 5′-gaaa-3′ tetraloop predominantly adopts a hairpin structure (Supplementary Figure S4), which is consistent with the native gel mobility, indicating a monomeric species (lane 3, Figure 3A).

Mutations in the Loop 2 palindrome reduces frameshifting efficiency

To determine the importance of L2-mediated self-association for SARS-CoV -1 PRF, several pseudoknot variants were subjected to a dual luciferase reporter-based frameshifting analysis as previously described (Figure 8A) (1,39). A Loop 2 ACUucc silent (serine codon) mutation was made to investigate the role of the palindrome (S3L2-ACUucc; Figure 1D) in promoting -1 PRF. Phylogenetic analysis indicated that the sequence UCC is present at a
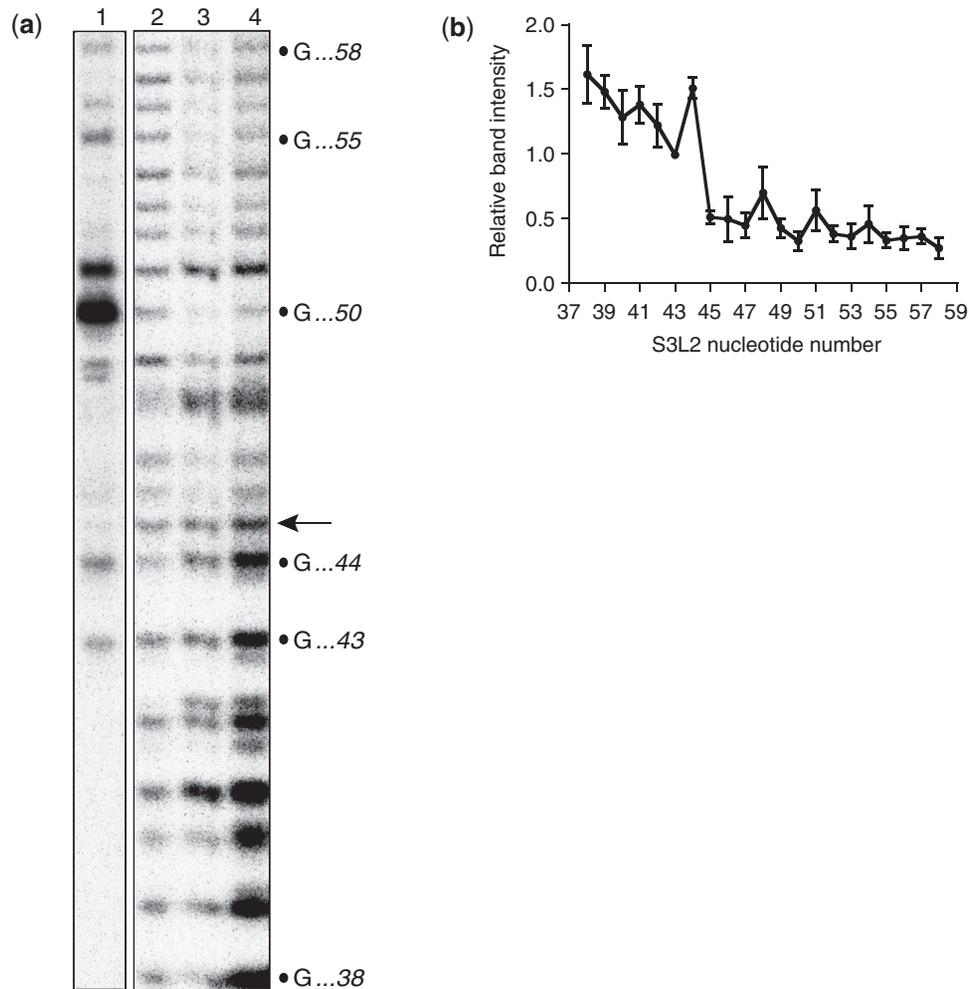
Figure 6. Identification of the crosslinking site of the S3L2 dimer. **(a)** Partial alkaline hydrolysis of control S3L2 and crosslinked S3L2 [^{32}P]ATP-RNAs. Lane 1, RNase T1 digestion of control S3L2 RNA; locations of Gs are shown on the right side. Arrow, position of C...*45*, the first nucleotide with reduced intensity in the partial hydrolysis. Lane 2, alkaline hydrolysis ladder. Lanes 3 and 4, partial alkaline hydrolysis of crosslinked and non-crosslinked samples, respectively. **(b)** Ratio of intensities from the partial hydrolysis of lanes 3 and 4 in Panel (a). Values for each lane were normalized and the intensities of each band of crosslinked over control RNAs were plotted. Error bars indicate the values obtained in four independent experiments.

similar position in Loop 2 of the transmissible gastroenteritis CoV (1). Dual luciferase measurements of -1 PRF from the S3L2-ACUucc pk construct showed a 63% reduction in frameshift stimulation compared with wild-type pk (5.7% versus 15.0%, Figure 8A). Next, we tested two S3 mutants for frameshift stimulation in the luciferase assay. S3-cuug (Figure 1E), which we showed to readily self-associate (Figure 3A, lane 2), demonstrated a nominal increase in frameshift frequency (17.5%), while a S3 truncation of the same construct (S3-2bp-cuug, Figure 1F) exhibited near wild-type efficiency (15.2%). It is important to note that the previously characterized S3-deletion ΔS3 pk construct maintained dimerization capability albeit with reduced efficiency (Figures 3B and C, lane 3). Our results are consistent with previous analyses by three different groups that have shown mutations to the third stem of a three-stemmed pseudoknot to have less impact on frameshifting efficiency than mutations to Stem 2 (1,2,6,7). However, our data indicate that silent codon

changes to Loop 2 that disrupt the palindrome play a role in regulating the frequency of frameshifting.

Effects of silent codon changes on viral viability

Having shown that the SARS-CoV pseudoknot dimerizes *in vitro* and alters frameshifting in Vero cells, we asked whether the ability to self-associate is important for viral propagation. Site-directed mutagenesis was used to introduce the Loop 2 ACUucc silent mutation into SARS subclone D (40), and the recombinant virus mutants were recovered from full-length virus transcripts. The observed effects should therefore reflect changes related to RNA structure within the frameshifting signal, rather than modifying the encoded proteins. To determine the effects on viral growth kinetics, stocks of rescued wild-type SARS-CoV and S3L2-ACUucc pk (multiplicity of infection (MOI) of 1) were used to infect VeroE6 cells, media was harvested at 0, 5 and 8 h pi and tested by plaque assay. Viral titers for SARS-CoV increased by
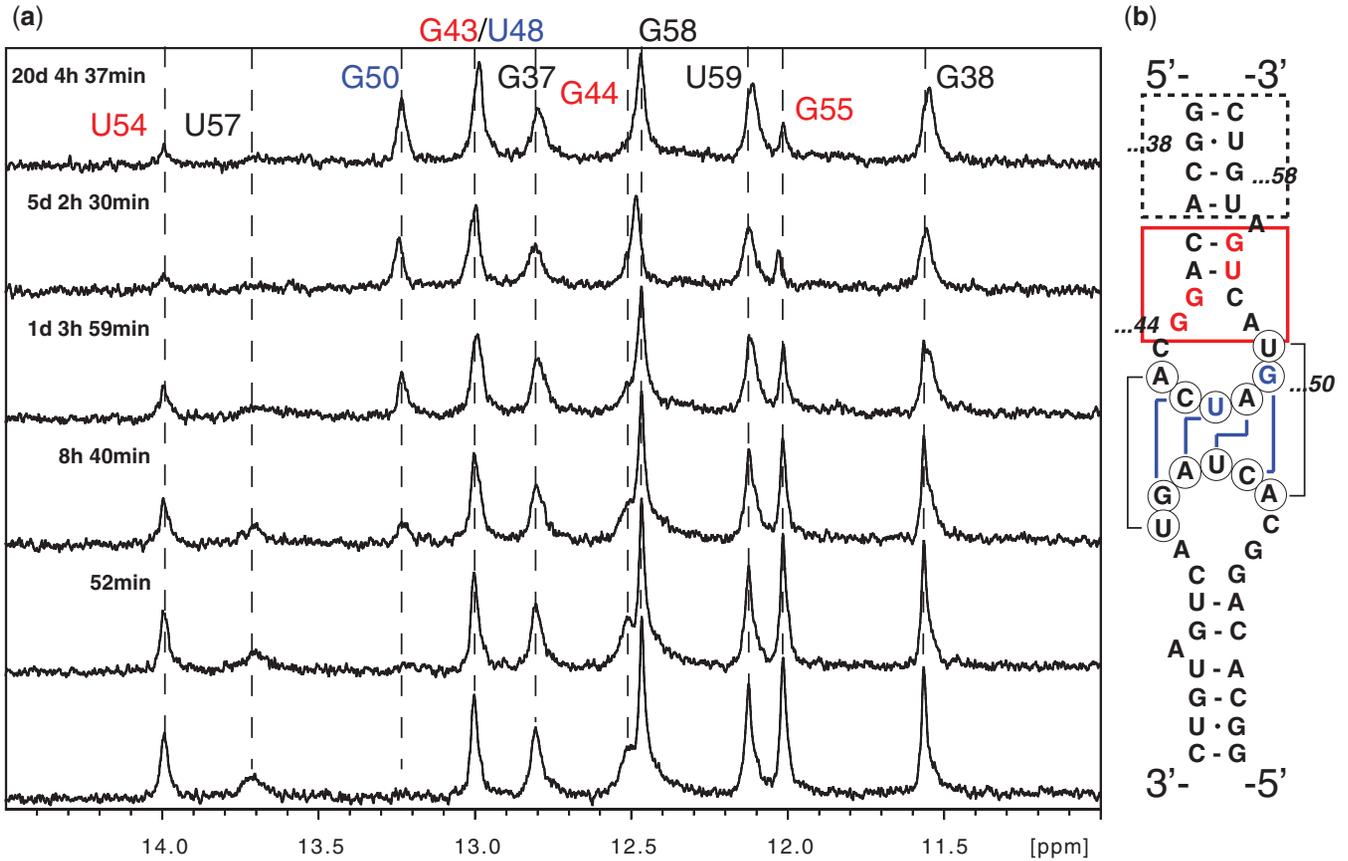
Figure 7. NMR real-time monitoring of salt-dependent loop–loop kissing formation of wild-type SARS-CoV S3L2 stem–loop. **(a)** Time course of imino regions of 1D ^{1}H-jump-return echo experiments. Data were collected on a SARS S3L2 sample containing 0.25 mM RNA in 500 μl volume of NMR buffer. Spectra were recorded at 295 K on a Bruker Avance III 600 (S3) MHz spectrometer. Loop–loop kissing formation was induced by spiking the NMR buffer [10 mM sodium phosphate (pH 6.5), 500 μM EDTA, 50 μM sodium azide, 9:1 H$_2$O:D$_2$O] with 125 mM potassium chloride. Dashed black lines follow assigned imino proton resonances at various time points. Only the last two digits of the wild-type sequence numbering are used for clarity. **(b)** Schematic S3L2 inset showing the loop–loop kissing geometry highlighting the lower S3 stem (black dashed box), the base-paired region in the upper S3 stem (solid red box) as well as intermolecular loop–loop kissing interactions (solid blue lines).
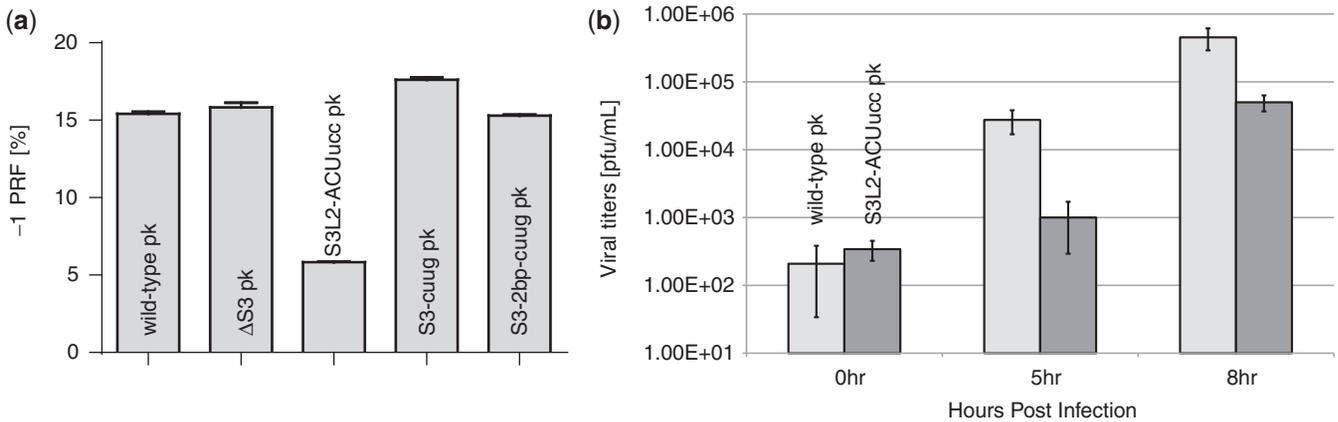

Figure 8. Functional analysis of wild-type SARS-CoV and S3L2-ACUucc pk mutants. **(a)** Frameshifting Frequencies of SARS S3L2 Pseudoknot Mutants. -1PRF (%) efficiency was determined as described in the methods. -1 PRF efficiency for ΔS3 pk is included for comparison and was reported previously (1). **(b)** Cultures of VeroE6 cells were inoculated with wild-type or S3L2-ACUucc pk mutant SARS-CoV at an MOI 1. Samples were taken at the indicated times and viral titers determined by plaque assay. Titers are indicated as plaque forming units/ml (ml). Error bars represent the standard deviations of three measurements. Student's t-test: for 0 h, $P = 0.6184$; for 5 h, $P = 0.0310$; for 8 h, $P = 0.0015$.

2.5 logs at 5 h pi and by 3.5 logs at 8 h pi, while the growth kinetics for S3L2-ACUucc, a mutant that does not form dimers, lagged behind at ~1 log increase at 5 h pi and ~2.5 logs by 8 h pi (Figure 8B). Our analysis demonstrates that disruption of the palindrome via silent codon changes does impede growth kinetics for S3L2-ACUucc pk at early times pi; however, by 12 h pi, similar titers were detected (data not shown).

To determine if the differences in viral growth kinetics were due to reduced levels of RNA replication and/or synthesis, we harvested total RNA at 0, 5, 8 and 12 h pi and performed northern blot analysis. Total RNA was probed using a nucleocapsid specific biotinylated probe (Figure 9A) to examine levels of gRNA. S3L2-ACUucc gRNA levels were reduced ~2-fold when compared with SARS-CoV at 5 and 8 h pi, suggesting that RNA replication contributed to the reduction in viral titers (Figure 9B). To determine the quantities of sgRNA species, the total RNA was enriched for poly A containing mRNA species (Figure 9C) before separation on agarose gel. All sgRNA species were readily detectible in both SARS-CoV and S3L2-ACUucc samples; however, ~1.5-fold reduction in the quantity of sgRNA species was detected at 5 and 8 h pi (Figure 9D), suggesting that reduced levels of RNA transcription also contributed to the titer loss. Finally we compared the levels of ORF1a and ORF1b replicase proteins to determine if these protein levels were altered as well. Equivalent protein concentrations were probed with polyclonal rabbit antisera directed against nsp1 or nsp16 as indicated (Figure 9E and F). Reduced ORF1a and ORF1b replicase protein levels were seen at 5 and 8 h pi but similar levels were detected by 12 h pi (Supplementary Figure S5). Together these data indicate that dimerization is not required for viral viability but it is important for maintaining efficient rates of virus growth and genomic and subgenomic mRNA synthesis.

While Stem 3 is conserved in group 2 coronaviruses, the palindrome in loop 2 is not

The common ancestor of the civet and human strains seems to have been a bat virus (9,41). Therefore, we attempted to obtain information about the combined conservation of sequential and structural features of Stem 3 from a compilation of 22 sequences from related human [GenBank accession codes AY304495 (42) and NC_004718 (43,44)], masked palm civet [GenBank accession code AY304488 (42)] and bat coronaviruses. For this purpose, Mfold (30) was used to predict lowest energy structures of aligned CoV S3 sequences. While 22 out of 24 nt of the S3L2 region are absolutely invariant between the 22 viral species (Supplementary Figure S6), the variation that does occur most frequently disrupts the hexanucleotide palindromic sequence at its 3'-end exhibiting a U51C substitution (Supplementary Figure S6). The U51C substitution stabilizes the hairpin and lowers the calculated ΔG from -4.8 kcal/mol to -7.5 kcal/mol (Supplementary Table S1). We previously showed that this 'ancestor' S3L2-AGUAGc$_{51}$ variant was dimerization incompetent (lanes 2, Figure 3A, D).

On the basis of phylogeny clustering, CoVs are classified into three Groups (Group 1–3). The SARS–CoV lineage has been proposed to cluster with Group 2 (44,45). To further analyse conservation patterns within the S3 element in a broader context, we aligned 28 complete genomic CoV sequences (summarized in Supplementary Table S1). As previously shown, pseudoknots containing a structured Stem 3 can be predicted for all Group 2 coronaviruses (1). The calculated average free energy for predicted Stem 3s of Group 2 CoV is $\Delta G = -6.8 \pm 1.2$ kcal/mol. Except for Feline CoV ($\Delta G = -5.6$ kcal/mol), all Group 1 sequences are not predicted to form stable Stem 3 substructures and are characterized by average free energies around zero ($\Delta G = -1.4 \pm 1.7$ kcal/mol). Similarly, in Group 3 coronaviruses the S3L2 region is most likely a single-stranded loop as indicated by a predicted ΔG of -1.8 kcal/mol (1). Remarkably, among the predicted S3L2 structures of all 17 Group 2 coronaviruses investigated, SARS-CoV codes for the second most labile Stem 3 with a ΔG of -4.8 kcal/mol; only the Human CoV HKU1 Stem 3 is predicted to be less stable as indicated by its ΔG of -4.7 kcal/mol (Supplementary Figure S7). Thus, SARS-CoV may form homodimers using its unique hexanucleotide palindromic sequence to compensate for a lack of thermodynamic stability.

DISCUSSION

Although a coronavirus frameshift signal was among the first pseudoknots to be identified and has been extensively studied using mutagenesis (7), the size of the CoV pseudoknots have limited structural analyses at the atomic level. NMR and X-ray crystallography have been used to describe several smaller pseudoknots but the differences between the pseudoknots makes it impossible to extrapolate features from the smaller pseudoknots to the larger ones (5). From studies of HIV and other retroviruses, it is clear that modulation of frameshift efficiency can have a dramatic effect on virus viability, and the same is true for positive-strand RNA viruses (46,47). Recently, it was demonstrated that the SARS-CoV pseudoknot could potentially be a target for antiviral agents making it imperative to understand its structure and function in greater detail (48).

Our studies of the third stem of the SARS coronavirus pseudoknot illuminate features of Stem 3 that affect RNA structure, frameshifting frequency and viral replication. Typically, G-C-rich Watson–Crick complementarity is required for stable loop–loop base pairing interactions between two RNA hairpins. Among all natural HIV isolates, two palindromic hexanucleotide sequences are most commonly found: GCGCGC and GUGCAC (49,50). However, mutagenesis work investigating HIV RNA dimerization also demonstrated that the GUUAAC palindrome found in SIV$_{mnd}$ (Simian Immunodeficiency Virus) could yield up to 20% dimers in the presence of 5 mM MgCl$_2$ (51). In good agreement with those studies, we confirm that weaker palindromic sequences such as AC UAGU can readily facilitate intermolecular loop–loop kissing RNA–RNA interactions.
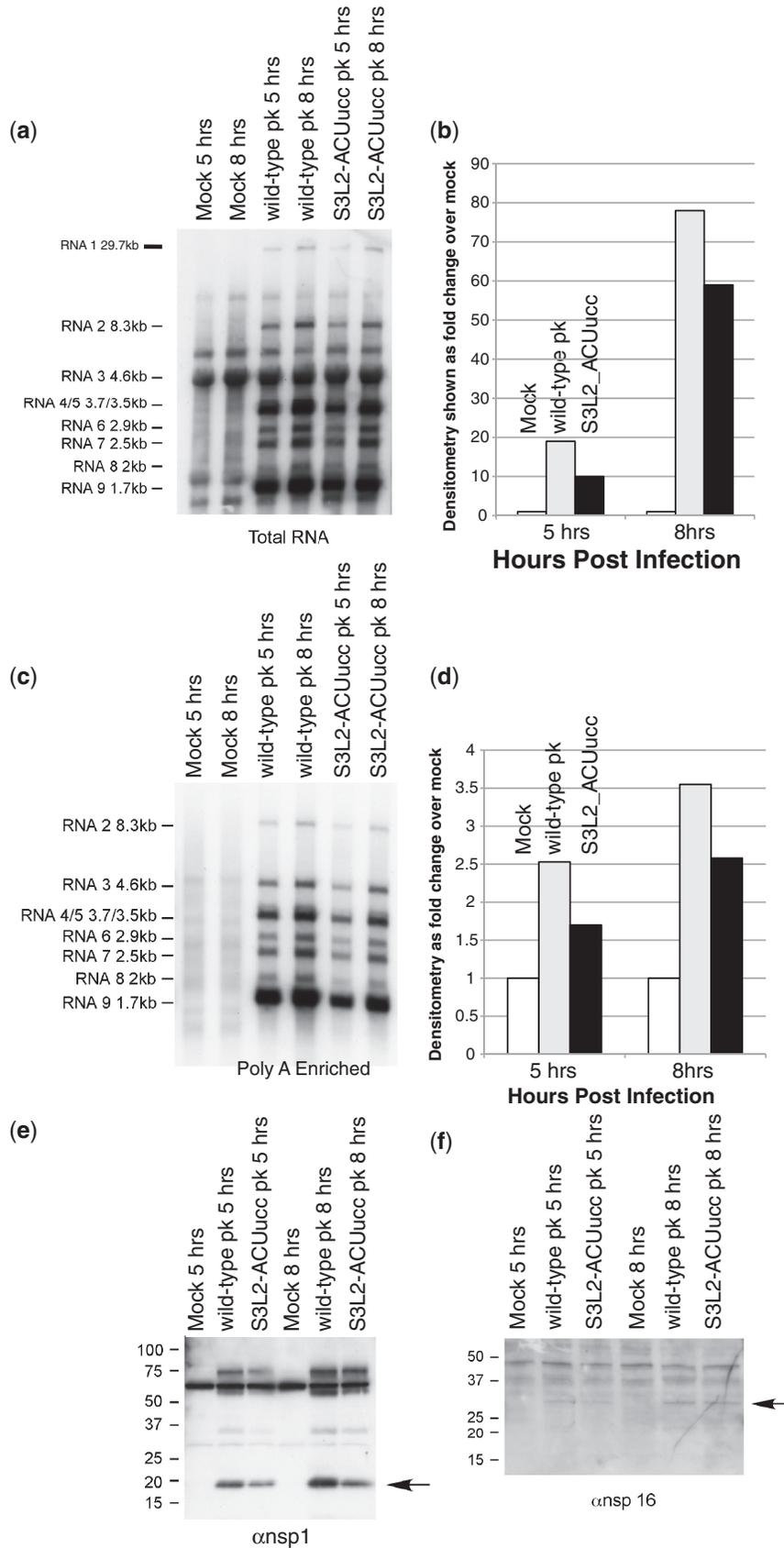
Figure 9. S3L2-ACUucc SARS-CoV mutant RNA synthesis and protein expression kinetics. Cultures of VeroE6 cells were infected with wild-type or S3L2-ACUucc mutant SARS-CoV at an MOI of 1 or mock-infected. Representative northern blots with sgRNA and gRNA species and molecular weights indicated by arrows. (a) probed total RNA to visualize genomic and subgenomic viral RNA and (c) probed polyA enriched RNA to visualize

(continued)

RNA pseudoknots implicated in -1 PRF stimulation are not static and are in fact dismantled during translation (52,53). If a frameshift event does not occur during the first round of translation then the pseudoknot will have to refold for frameshifting to occur in subsequent rounds. Long unstructured loops between the 3′ portion of Stem 1 and 5′ portion of Stem 2 would be expected to reduce the chances of Stem 2 forming and reduce frameshifting efficiency. If this sequence can form a stable restricting S3L2 substructure, as is the case for the SARS-CoV pseudoknot steadied by an intermolecular kissing complex, then the rapid re-formation of Stem 2, which is required for efficient frameshifting, will be more probable.

In previous work, the functional importance of the 31 nt comprising Stem 3 and Loop 2 in the SARS-CoV pseudoknot (C36 through A67) were investigated and no obvious determinants for efficient frameshift stimulation for this specific element could be identified. Deletion analysis characterizing variants with 5(6) and 11 nt(2) crossing the minor groove side of Stem 1 yielded 35 and 52% of wild-type frameshifting levels. However, because of the common lack of precise structural information for pseudoknot varaints, determining the exact contribution of specific nucleotides of the pseudoknot to frameshifting is often difficult. There is a potential pitfall in the mutational methods because longer-range distortions that are difficult to predict using e.g. Mfold especially in complicated three-stemmed architectures of Group II CoV could lead to false hypotheses regarding the physical mechanism of frameshifting.

Secondary structures and basepairing patterns for all SARS-CoV variants investigated in this study were characterized using NMR methods. Because the 5′-ACU AGU-3′ palindromic sequence was left intact, even the S3-deletion ΔS3 pk mutant could dimerize and thus efficiently stimulated -1 PRF. On the other hand, the S3L2-ACUucc did not engage in loop–loop kissing interactions, was characterized by an open Loop 2 structure, which destabilized Stem 3, and its ability to promote -1 PRF was reduced almost 3-fold.

When we capped Stem 3 with a stable CUUG tetraloop that normally facilitates formation of stable hairpin structures (54) and is highly conserved (55), the frequency of frameshifting almost returned to wild-type levels (Figure 8). Surprisingly, in the context of the SARS-CoV Stem 3 sequence, 5′-cuug-3′ tetraloop-capped mutants readily formed extended duplex structures as revealed by native gel and NMR analysis. Thus, the promotion of wild-type levels of frameshifting exhibited by S3-cuug and S3-2bp-cuug SARS pseudoknot variants is consistent with a mechanism involving dimer formation. The metastable SARS-CoV Stem 3 sequence, which features a G38·U59 wobble pair and a bulged A56, apparently evolved to facilitate a certain degree of

dimerization if seeded through loop–loop kissing interactions that can even tolerate tandem U-U mismatches.

The S3L2-ACUucc pk mutant, in which frameshifting was reduced nearly 3-fold, was viable. However, while the mutant replicated to levels similar to the wild-type virus at later times pi, the inability to dimerize in infected Vero cells significantly changed the viral growth kinetics, RNA species and replicase protein levels. The fact that a silent codon change in the Loop 2 of SARS-CoV reproducibly affected the levels of gRNA and sgRNA strongly suggests that dimer formation occurs in the cellular environment and that loop–loop kissing interactions involving Stem 3 of the pseudoknot are important for accumulation of sgRNA. The *in vitro* findings of reduced total RNA levels and ORF1a/b protein following infection leading to a lag in replication are consistent with reduced levels of ORF1b translation products as a result of the 3-fold reduction in -1 PRF. The downregulation of the ORF1a gene product nsp1 is likely a result of the decrease in the amount of gRNA in the infected cells.

Several studies have demonstrated that facile formation of the H-type skeleton involving Stem 1 and 2 of a frameshift-stimulating pseudoknot structure is critical for efficient frameshifting (56,57). Intriguingly, several long-distance intramolecular loop–loop kissing has also been implicated in -1 PRF stimulation. Specifically, the Group 1 Human CoV 229 E has been shown to use a kissing stem loop to promote frameshifting rather than a compact H-type pseudoknot (58). Based on sequence alignments, such 'elaborated' pseudoknots could be used by Human CoV NL63, porcine epidemic diarrhea virus and transmissible gastroenteritis virus (6). Similarly, the P1–P2 pseudoknot from the luteovirus barley yellow dwarf virus can be considered a variation on the 'elaborated' CoV pseudoknot structure. It contains a loop L2 of nearly 4 kb, with S2 formed by long-range kissing interactions with nucleotides near the 3′ end of the genome (59). Such a long-distance RNA–RNA kissing interaction is also implicated in -1 PRF of red clover necrotic mosaic virus (60). Thus, an intermolecular dimerization in SARS-CoV may simply represent a variation of the 'elaborated' pseudoknot.

In conclusion, we demonstrate that the structured third stem of SARS-CoV represents an integral part of the -1 PRF stimulating pseudoknot topology. While the exact functional interplay in the viral lifecycle and timing of Stem 3 folding–unfolding interconversions remain to be determined, our results provide important first insight into a palindromic sequence element embedded into Stem 3 that permit loop–loop kissing formation for efficient -1 PRF stimulation, efficient translation of ORF1a/b encoded proteins and relative amounts of sgRNA to remain close to wild-type levels.

Figure 9. Continued
the sgRNA species. Delays in RNA synthesis kinetics in S3L2-ACUucc frameshift mutant. Bars represent the densitometry measurements of fold change over mock for mock-infected, wild-type and S3L2-ACUucc SARS-CoV mutant RNA at the indicated times pi. (**b**) Densitometry of gRNA. (**d**) Densitometry of sgRNA. ORF1a versus ORF1b protein expression kinetics in S3L2-ACUucc frameshift mutant. Proteins were separated on polyacrylamide gels and probed with rabbit sera directed against (**e**) the replicase proteins nsp1 or (**f**) nsp16 as indicated. Size markers are indicated to the left of each blot.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–7.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Plant,E.P., Perez-Alvarado,G.C., Jacobs,J.L., Mukhopadhyay,B., Hennig,M. and Dinman,J.D. (2005) A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol.*, **3**, e172.
2. Su,M.C., Chang,C.T., Chu,C.H., Tsai,C.H. and Chang,K.Y. (2005) An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Res.*, **33**, 4265–4275.
3. Thiel,V., Ivanov,K.A., Putics,A., Hertzig,T., Schelle,B., Bayer,S., Weissbrich,B., Snijder,E.J., Rabenau,H., Doerr,H.W. *et al.* (2003) Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol.*, **84**, 2305–2315.
4. Plant,E.P. and Dinman,J.D. (2006) Comparative study of the effects of heptameric slippery site composition on -1 frameshifting among different eukaryotic systems. *RNA*, **12**, 666–673.
5. Giedroc,D.P. and Cornish,P.V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.*, **139**, 193–208.
6. Baranov,P.V., Henderson,C.M., Anderson,C.B., Gesteland,R.F., Atkins,J.F. and Howard,M.T. (2005) Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology*, **332**, 498–510.
7. Brierley,I., Rolley,N.J., Jenner,A.J. and Inglis,S.C. (1991) Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.*, **220**, 889–902.
8. Lau,S.K., Li,K.S., Huang,Y., Shek,C.T., Tse,H., Wang,M., Choi,G.K., Xu,H., Lam,C.S., Guo,R. *et al.* (2010) Ecoepidemiology and complete genome comparison of different strains of severe acute respiratory syndrome-related Rhinolophus bat coronavirus in China reveal bats as a reservoir for acute, self-limiting infection that allows recombination events. *J. Virol.*, **84**, 2808–2819.
9. Shi,Z. and Hu,Z. (2008) A review of studies on animal reservoirs of the SARS coronavirus. *Virus Res.*, **133**, 74–87.
10. Clever,J.L., Wong,M.L. and Parslow,T.G. (1996) Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. *J. Virol.*, **70**, 5902–5908.
11. Haddrick,M., Lear,A.L., Cann,A.J. and Heaphy,S. (1996) Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. *J. Mol. Biol.*, **259**, 58–68.
12. Laughrea,M. and Jette,L. (1994) A 19-nucleotide sequence upstream of the 5′ major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. *Biochemistry*, **33**, 13464–13474.
13. Skripkin,E., Paillart,J.C., Marquet,R., Ehresmann,B. and Ehresmann,C. (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc. Natl Acad. Sci. USA*, **91**, 4945–4949.
14. Shen,N., Jette,L., Liang,C., Wainberg,M.A. and Laughrea,M. (2000) Impact of human immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B on RNA dimerization and reverse transcription and dissociation of dimerization from packaging. *J. Virol.*, **74**, 5729–5735.
15. Paillart,J.C., Shehu-Xhilaga,M., Marquet,R. and Mak,J. (2004) Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev. Microbiol.*, **2**, 461–472.
16. Lomniczi,B. and Kennedy,I. (1977) Genome of infectious bronchitis virus. *J. Virol.*, **24**, 99–107.
17. Enjuanes,L., Almazan,F., Sola,I. and Zuniga,S. (2006) Biochemical aspects of coronavirus replication and virus-host interaction. *Annu. Rev. Microbiol.*, **60**, 211–230.
18. Roth,B.M. and Hennig,M. (2011) In: Dingley,A. and Pascal,S.M. (eds), *Biomolecular NMR Spectroscopy*, Vol. 3. IOS Press, Amsterdam, pp. 205–228.
19. Scott,L.G. and Hennig,M. (2008) RNA structure determination by NMR. *Methods Mol. Biol.*, **452**, 29–61.
20. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
21. Goddard,T.D. and Kneller,D.G. (2004). University of California, San Francisco, San Francisco.
22. Sklenar,V., Brooks,B.R., Zon,G. and Bax,A. (1987) Absorption mode two-dimensional NOE spectroscopy of exchangeable protons in oligonucleotides. *FEBS Lett.*, **216**, 249–252.
23. Lippens,G., Dhalluin,C. and Wieruszeski,J.M. (1995) Use of a water flip-back pulse in the homonuclear noesy experiment. *J. Biomol. NMR*, **5**, 327–331.
24. Harris,M.E. and Christian,E.L. (2009) RNA crosslinking methods. *Methods Enzymol.*, **468**, 127–146.
25. Nanda,K. and Wollenzien,P. (2004) Pattern of 4-thiouridine-induced cross-linking in 16S ribosomal RNA in the Escherichia coli 30S subunit. *Biochemistry*, **43**, 8923–8934.
26. Jacobs,J.L. and Dinman,J.D. (2004) Systematic analysis of bicistronic reporter assay data. *Nucleic Acids Res.*, **32**, e160.
27. Sims,A.C., Baric,R.S., Yount,B., Burkett,S.E., Collins,P.L. and Pickles,R.J. (2005) Severe acute respiratory syndrome coronavirus infection of human ciliated airway epithelia: role of ciliated cells in viral spread in the conducting airways of the lungs. *J. Virol.*, **79**, 15511–15524.
28. Prentice,E., McAuliffe,J., Lu,X., Subbarao,K. and Denison,M.R. (2004) Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. *J. Virol.*, **78**, 9977–9986.
29. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
30. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
31. Wittenhagen,L.M. and Kelley,S.O. (2002) Dimerization of a pathogenic human mitochondrial tRNA. *Nat. Struct. Biol.*, **9**, 586–590.
32. Jossinet,F., Paillart,J.C., Westhof,E., Hermann,T., Skripkin,E., Lodmell,J.S., Ehresmann,C., Ehresmann,B. and Marquet,R. (1999) Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding. *RNA*, **5**, 1222–1234.
33. Reblova,K., Spackova,N., Sponer,J.E., Koca,J. and Sponer,J. (2003) Molecular dynamics simulations of RNA kissing-loop motifs reveal structural dynamics and formation of cation-binding pockets. *Nucleic Acids Res.*, **31**, 6942–6952.
34. Horiya,S., Li,X., Kawai,G., Saito,R., Katoh,A., Kobayashi,K. and Harada,K. (2003) RNA LEGO: magnesium-dependent formation of specific RNA assemblies through kissing interactions. *Chem. Biol.*, **10**, 645–654.
35. Ly,H. and Parslow,T.G. (2002) Bipartite signal for genomic RNA dimerization in Moloney murine leukemia virus. *J. Virol.*, **76**, 3135–3144.

36. Shetty,S., Kim,S., Shimakami,T., Lemon,S.M. and Mihailescu,M.R. (2010) Hepatitis C virus genomic RNA dimerization is mediated via a kissing complex intermediate. *RNA*, **16**, 913–925.

37. Laughrea,M. and Jette,L. (1996) Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248-271 are dispensable for dimer formation. *Biochemistry*, **35**, 1589–1598.

38. Laughrea,M. and Jette,L. (1997) HIV-1 genome dimerization: kissing-loop hairpin dictates whether nucleotides downstream of the 5′ splice junction contribute to loose and tight dimerization of human immunodeficiency virus RNA. *Biochemistry*, **36**, 9501–9508.

39. Plant,E.P., Rakauskaite,R., Taylor,D.R. and Dinman,J.D. (2010) Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. *J. Virol.*, **84**, 4330–4340.

40. Yount,B., Curtis,K.M., Fritz,E.A., Hensley,L.E., Jahrling,P.B., Prentice,E., Denison,M.R., Geisbert,T.W. and Baric,R.S. (2003) Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc. Natl Acad. Sci. USA*, **100**, 12995–13000.

41. Graham,R.L. and Baric,R.S. (2010) Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.*, **84**, 3134–3146.

42. Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L., Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J. *et al.* (2003) Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science*, **302**, 276–278.

43. Marra,M.A., Jones,S.J., Astell,C.R., Holt,R.A., Brooks-Wilson,A., Butterfield,Y.S., Khattra,J., Asano,J.K., Barber,S.A., Chan,S.Y. *et al.* (2003) The genome sequence of the SARS-associated coronavirus. *Science*, **300**, 1399–1404.

44. Snijder,E.J., Bredenbeek,P.J., Dobbe,J.C., Thiel,V., Ziebuhr,J., Poon,L.L., Guan,Y., Rozanov,M., Spaan,W.J. and Gorbalenya,A.E. (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.*, **331**, 991–1004.

45. Eickmann,M., Becker,S., Klenk,H.D., Doerr,H.W., Stadler,K., Censini,S., Guidotti,S., Masignani,V., Scarselli,M., Mora,M. *et al.* (2003) Phylogeny of the SARS coronavirus. *Science*, **302**, 1504–1505.

46. Shehu-Xhilaga,M., Crowe,S.M. and Mak,J. (2001) Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol.*, **75**, 1834–1841.

47. Telenti,A., Martinez,R., Munoz,M., Bleiber,G., Greub,G., Sanglard,D. and Peters,S. (2002) Analysis of natural variants of the human immunodeficiency virus type 1 gag-pol frameshift stem-loop structure. *J. Virol.*, **76**, 7868–7873.

48. Ahn,D.G., Lee,W., Choi,J.K., Kim,S.J., Plant,E.P., Almazan,F., Taylor,D.R., Enjuanes,L. and Oh,J.W. (2011) Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses SARS coronavirus replication. *Antiviral Res*, **91**, 1–10.

49. Berkhout,B. (1996) Structure and function of the human immunodeficiency virus leader RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **54**, 1–34.

50. Paillart,J.C., Marquet,R., Skripkin,E., Ehresmann,C. and Ehresmann,B. (1996) Dimerization of retroviral genomic RNAs: structural and functional implications. *Biochimie*, **78**, 639–653.

51. Laughrea,M., Shen,N., Jette,L. and Wainberg,M.A. (1999) Variant effects of non-native kissing-loop hairpin palindromes on HIV replication and HIV RNA dimerization: role of stem-loop B in HIV replication and HIV RNA dimerization. *Biochemistry*, **38**, 226–234.

52. Hansen,T.M., Reihani,S.N., Oddershede,L.B. and Sorensen,M.A. (2007) Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc. Natl Acad. Sci. USA*, **104**, 5830–5835.

53. Chen,G., Chang,K.Y., Chou,M.Y., Bustamante,C. and Tinoco,I. Jr. (2009) Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc. Natl. Acad. Sci. USA*, **106**, 12706–12711.

54. Jucker,F.M. and Pardi,A. (1995) Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. *Biochemistry*, **34**, 14416–14427.

55. Woese,C.R., Winker,S. and Gutell,R.R. (1990) Architecture of ribosomal RNA: constraints on the sequence of ''tetra-loops''. *Proc. Natl. Acad. Sci. USA*, **87**, 8467–8471.

56. Farabaugh,P.J. (1996) Programmed translational frameshifting. *Annu. Rev. Genet.*, **30**, 507–528.

57. Giedroc,D.P., Theimer,C.A. and Nixon,P.L. (2000) Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J. Mol. Biol.*, **298**, 167–185.

58. Herold,J. and Siddell,S.G. (1993) An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res.*, **21**, 5838–5842.

59. Paul,C.P., Barry,J.K., Dinesh-Kumar,S.P., Brault,V. and Miller,W.A. (2001) A sequence required for -1 ribosomal frameshifting located four kilobases downstream of the frameshift site. *J. Mol. Biol.*, **310**, 987–999.

60. Tajima,Y., Iwakawa,H.O., Kaido,M., Mise,K. and Okuno,T. (2011) A long-distance RNA-RNA interaction plays an important role in programmed -1 ribosomal frameshifting in the translation of p88 replicase protein of Red clover necrotic mosaic virus. *Virology*, **417**, 169–178.