An RNA Element That Facilitates Programmed Ribosomal Readthrough in Turnip Crinkle Virus Adopts Multiple Conformations

Micki M. Kuhlmann, Maitreyi Chattopadhyay, Vera A. Stupina, Feng Gao, Anne E. Simon
Department of Cell Biology and Molecular Genetics, University of Maryland—College Park, College Park, Maryland, USA

ABSTRACT
Ribosome recoding is used by RNA viruses for translational readthrough or frameshifting past termination codons for the synthesis of extension products. Recoding sites, along with downstream recoding stimulatory elements (RSEs), have long been studied in reporter constructs, because these fragments alone mediate customary levels of recoding and are thus assumed to contain complete instructions for establishment of the proper ratio of termination to recoding. RSEs from the Tombusviridae and Luteoviridae are thought to be exceptions, since they contain a long-distance RNA-RNA connection with the 3′ end. This interaction has been suggested to substitute for pseudoknots, thought to be missing in tombusvirid RSEs. We provide evidence that the phylogenetically conserved RSE of the carmovirus Turnip crinkle virus (TCV) adopts an alternative, smaller structure that extends an upstream conserved hairpin and that this alternative structure is the predominant form of the RSE within nascent viral RNA in plant cells and when RNA is synthesized in vitro. The TCV RSE also contains an internal pseudoknot along with the long-distance interaction, and the pseudoknot is not compatible with the phylogenetically conserved structure. Conserved residues just past the recoding site are important for recoding, and these residues are also conserved in the RSEs of gammaretroviruses. Our data demonstrate the dynamic nature of the TCV RSE and suggest that studies using reporter constructs may not be effectively recapitulating RSE-mediated recoding within viral genomes.

IMPORTANCE
Ribosome recoding is used by RNA viruses to enable ribosomes to extend translation past termination codons for the synthesis of longer products. Recoding sites and a downstream recoding stimulatory element (RSE) mediate expected levels of recoding when excised and placed in reporter constructs and thus are assumed to contain complete instructions for the establishment of the proper ratio of termination to recoding. We provide evidence that most of the TCV RSE adopts an alternative structure that extends an upstream conserved hairpin and that this alternative structure, and not the phylogenetically conserved structure, is the predominant form of the RSE in RNA synthesized in vitro and in plant cells. The TCV RSE also contains an internal pseudo-knot that is not compatible with the phylogenetically conserved structure and an RNA bridge to the 3′ end. These data suggest that the TCV RSE is structurally dynamic and that multiple conformations are likely required to regulate ribosomal readthrough.

Positive-sense RNA viruses employ a variety of gene expression strategies for the diversification of their proteomes (1, 2). Two noncanonical mechanisms, −1 programmed ribosomal frameshifting (−1PRF) and programmed ribosomal readthrough (PRT), circumvent stop codons for the expression of carboxy-terminal extension products, effectively economizing on the limited genome size of most RNA viruses (3, 4). In −1PRF, the elongating ribosome shifts back 1 residue at a 7-residue slippery sequence upstream of the stop codon and then continues translating in the new reading frame, generating a fusion polypeptide. In PRT, a tRNA decodes the stop codon, and translation continues in the same reading frame. These processes are used by viruses mainly to express their RNA-dependent RNA polymerase (RdRp) or a coat protein (CP) extension product that is necessary for vector transmission and/or virion assembly (5). Ribosome recoding leads to the tightly controlled synthesis of 5% to 10% fusion products, and the maintenance of this level is critical for efficient virus amplification (6–10).

Early studies used deletions to define minimal sequences capable of stimulating −1PRF from recoding regions inserted into heterologous RNA. These studies identified a stem-loop structure in Rous sarcoma virus (11) and a pseudoknotted structure in Avian infectious bronchitis virus (IBV) (12) that enhanced recoding independently of additional viral sequences. Since then, nearly all recoding studies have used similar approaches, which generally involve inserting the recoding site and downstream sequences between two luciferase reporter open reading frames (ORFs) and assaying for recoding-required luciferase activity either in cell extracts (in vitro) or in cell culture (in vivo) (13–17). Using this approach, studies have confirmed that the majority of −1PRF and PRT events require the presence of a stable hairpin or pseudoknotted recoding stimulatory element (RSE) just downstream of the recoding site (18, 19).

Despite 25 years of study, the specific mechanism(s) that mediates the critical ratio of recoding to termination products re-
mains unresolved. For −1PRF, pseudoknotted structures are proposed to compromise the translocation process by forming a physical barrier that blocks the mRNA entrance channel, impeding the mechanical unwinding of the mRNA by the translating ribosome (20). The stability of the stem of simple RSE hairpins has also been suggested to serve a similar function in promoting PRT and −1PRF (17, 19, 21). However, no consistent correlation has been found between frameshifting efficiency in reporter constructs and the mechanical force required to unfold the RSE using optical tweezers (22, 23). Instead, the suggestion has been made that frameshifting efficiency correlates with conformational plasticity and the tendency of the pseudoknot to form alternative local folds. The structural plasticity of the West Nile virus RSE was recently demonstrated, with both tandem stem-loop and pseudoknotted structures fitting the structural data, suggesting that this RSE may comprise a molecular switch necessary for the control of levels of frameshifting (24). In addition, protonation of the H-type pseudoknot RSE of Murine leukemia virus (MuLV) was suggested to be important for the formation of the readthrough-permissive conformation (25); however, whether this represents a general mechanism remains an open question.

All of these studies make the assumption that fragments active for recoding in vitro contain all instructions necessary for obtaining the proper ratio of termination to recoding products during the translation of the virus. However, long-distance RNA-RNA interactions are required for efficient −1PRF in Pea enation mosaic virus RNA 2 (PEMV2), Barley yellow dwarf virus (BYDV), and Red clover necrotic mosaic virus (21, 26–29) and for PRT in Carnation Italian ringspot virus and Tobacco necrosis virus D (30, 31). Since internal H-type pseudoknots have not been reported in RSEs from any of these plant viruses, it has been proposed that these long-distance RNA bridges form an atypical pseudoknot that functionally replaces the internal pseudoknot that is common in animal virus RSEs (2, 28, 32). In addition, hairpins or artificial duplexes just upstream of the recoding site repress frameshifting in coronaviruses and PEMV2 (21, 33–35). Strikingly, in the absence of the upstream hairpin, the PEMV2 RSE is no longer required to stimulate frameshifting from the slippery site, suggesting that upstream hairpins may play significant roles in suppressing −1PRF.

To circumvent methods that exclude proximal or distal viral sequences that might also be important for readthrough in vivo, we assessed PRT within full-length viral genomic RNA (gRNA) and chose Turnip crinkle virus (TCV) because of its small size, its well-characterized genome, and the simplicity of studying virus accumulation in single cells. TCV is a member of the Carmovirus genus within the Tombusviridae family and has a 4,053-nucleotide (nt) positive-sense genome containing five overlapping ORFs (Fig. 1A). The 5’-proximal ORF (p28, encoding a replicase-associated protein) and the p88 RdRp readthrough product are translated from the gRNA, which lacks a 5’ cap and 3’ poly(A) tail (36, 37). The TCV RSE, which is located just downstream of the p28 amber termination codon, contains two asymmetric internal loops, one of which pairs with the apical loop of the terminal 3’-end Pr hairpin (30).

Here we report that in addition to the long-distance interaction between the RSE and 3’ sequences, the TCV RSE contains an internal H-type pseudoknot, and both tertiary interactions are important for competent ribosome recoding. We also demonstrate that a short conserved sequence just downstream from the recoding site is required for efficient PRT and that this sequence is similar to the sequence previously found at the gag-pol junction in gammaretroviruses (38, 39). Most significantly, we provide structural evidence that the lower stem of the TCV RSE adopts an alternative conformation that extends an upstream, phylogenetically conserved hairpin and that this alternative structure is the predominant form of the RSE in vitro and in plant cells. These data reveal that at least three RSE conformations exist for TCV, showcasing the dynamic nature of RSEs and adding to the growing evidence that the current approaches—examining RSEs in isolation—may not effectively recapitulate the recoding process as it occurs within full-length viral gRNA.

**MATERIALS AND METHODS**

**Generation of plasmid constructs.** Mutant clones were generated using PCR-based site-directed mutagenesis from plasmid pTCV66, which contains the full-length TCV genome downstream of a T7 RNA polymerase promoter. Overlapping oligonucleotides (Integrated DNA Technologies) were used to introduce the desired mutations by use of Phusion high-fidelity polymerase (New England BioLabs) according to a one-step mutagenesis protocol (40). PCR products were subjected to DpnI digestion for 2 h at 37°C prior to transformation into DH5α competent cells. Mutations were confirmed by sequencing in the region (Eurofins Genomics).

**In vitro translation.** Smal-digested TCV plasmids were subjected to in vitro transcription for 2 h at 37°C using T7 RNA polymerase. Uncapped RNA transcripts (1 pmol) were translated in 10 μl of wheat germ extract (WGE; Promega), supplemented with 100 mM potassium acetate and [35S]methionine, according to the manufacturer’s instructions. Translated products were treated either with [35S]methionine, according to the manufacturer’s instructions. The reaction mixture was incubated at 25°C for 2 h and was resolved on 10% SDS-PAGE gel. The dried gel was exposed to a phosphorimager screen and was scanned by a FLA-5100 fluorescent image analyzer (Fujiﬁlm). Band intensity was quantiﬁed using Quantity One software (Bio-Rad). Experiments were performed in triplicate using independently transcribed RNAs.

**Protoplast transfection and RNA gel blots.** Protoplasts generated from seed-derived Arabidopsis thaliana callus tissue (ecotype Col-0) were transfected with in vitro-transcribed full-length TCV gRNA as described previously (41). Briefly, 20 μg of uncapped gRNA was transfected into 5 × 106 cells and was incubated in the dark for 40 h. Total RNA was extracted, subjected to electrophoresis, transferred to a nitrocellulose membrane, and probed using [γ-32P]ATP-labeled oligonucleotides complementary to positions 3931 to 3953, 3869 to 3883, and 4035 to 4054 in the 3’ untranslated region of the gag-pol extension (3’ UTR). The membrane was exposed to a phosphorimagga screen as described above. Experiments were performed in triplicate using independently transcribed RNAs.

**SHAPE structure probing.** Structure probing was performed using SHAPE (selective 2’-hydroxyl acylation analyzed by primer extension) as described previously (42). Briefly, for in vitro SHAPE, 6 pmol of in vitro-transcribed, full-length TCV gRNA was denatured for 5 min at 65°C, snap-cooled on ice, and then incubated in folding buffer at 37°C for 20 min. The folded RNA was treated with N-methylisothiocyanide (NMI; final concentration, 15 mM) for base modification or with an equal volume of dimethyl sulfoxide (DMSO) as a negative control. The reaction mixtures were incubated at 37°C for 35 min, ethanol precipitated, and resuspended in 8 μl of Tris-EDTA (TE) buffer. Primer extension was carried out using [γ-32P]ATP-labeled oligonucleotides and SuperScript III reverse transcriptase (Invitrogen) as described previously (43). Primers complementary to TCV positions 940 to 964, 875 to 894, 859 to 862, and 995 to 1018 were used for probing of the RSE, stem-loop A (SLA), and upstream and downstream region structures, respectively. Reaction products and dideoxy sequencing ladders (Roche) were resolved on an 8% denaturing polyacrylamide gel, exposed overnight to a phosphorimager screen, and imaged using the FLA-5100 fluorescent analyzer (Fujiﬁlm).
thaliana seed callus cultures (5 × 10^6 cells) were transfected with 20 μg of uncapped full-length wild-type (WT) TCV gRNA or an equal amount of a nonreplicating (GDD → GAA) polymerase active-site mutant. After 40 h of incubation at room temperature in the dark, either 1-methyl-7-nitroisatoic anhydride (1M7; final concentration, 5 mM) was added to the protoplasts for base modification or an equal volume of DMSO was added as a negative control, and cultures were gently shaken for 5 min at room temperature. IM7 was used in place of NMIA because NMIA has been reported to be inefficient at modifying RNA in a cellular environment (44, 45). Unmodified WT TCV was used to generate dideoxy sequencing ladders. Cells were collected by centrifugation; total RNA was extracted; and the RNA was resuspended in H2O. Primer extension, structure probing, resolution, and visualization were performed as described above. All SHAPE structure mapping was conducted in duplicate, independent experiments.

**RESULTS**

The TCV RSE contains an internal pseudoknot. Computational and phylogenetic approaches indicate that members of the Tombusviridae contain two hairpins in the vicinity of the recoding site irrespective of whether they use frameshifting or readthrough to synthesize their RdRp (Fig. 2). The 5′ hairpin, known as stem-loop A (SLA), is within 10 nt of the amber termination codon. The 3′ hairpin, previously identified as an RSE (30), is an unbranched structure with two large, asymmetric internal loops (L2 and L3), of which L2 is engaged in the long-distance interaction with sequences at the 3′ end. The RSE is located just downstream of the UAG, with the guanylate nearly always participating in the terminal base pair at the base of the G·C-rich lower stem (S1).

Although RSE pseudoknots have not been reported in tombusvirids, visual inspection of TCV and other carmovirus RSEs revealed the potential for an internal pseudoknot between guanylate-rich residues in upper internal loop L3 and cytidylates that are also predicted to be base paired in the highly conserved S1 stem (Fig. 1B and 2, blue). Internal loop mutation m1, which disrupts the putative pseudoknot, reduced the synthesis of p88 (and thus readthrough) by 97% in WGE, whereas stem S1 mutation m2, which would convert G·C to G:U pairs in S1 and the putative pseudoknot, did not affect readthrough (Fig. 1C). When the gRNA contained both m1 and m2, which should maintain the pseudoknot and S1, the level of p88 synthesis increased to 127% of the WT level. Although the m1 mutation was not a silent mutation (a proline was converted to a serine in the RdRp), mutant and WT gRNAs were evaluated for accumulation in Arabidopsis protoplasts. m1 reduced gRNA accumulation by 97%, and m2 reduced gRNA accumulation by 34%. The combination of the m1 and m2 mutations (m1/m2) was compensatory, with accumulation increasing (over m1 levels) to 71% of WT accumu-

**FIG 1** TCV RSE contains an internal pseudoknot. (A) Genome organization of TCV. Genomic and subgenomic RNAs are shown. The long-distance RNA-RNA interaction connecting the RSE with the 3′-terminal Pr hairpin is indicated by a dashed red line. (B) Phylogenetically conserved structure of the TCV RSE and the upstream hairpin SLA. Designations of stems and loops in the RSE are shown. Hairpins in the location of the SLA relative to the RSE are conserved in the Tombusviridae (see Fig. 2). Sequences in L2 and the Pr loop that engage in the long-distance interaction are shown in red. Putative pseudoknot residues are shown in blue. The locations of m1 and m2 mutations, generated to test for the pseudoknot, are shown. (C) In vitro translation of full-length wild-type (WT) and mutant gRNAs in WGE. The positions of p28 and the readthrough product, p88, are shown. p38 is the capsid protein and is translated by internal initiation. The product below p38 is an internal initiation product that is coterminal with the RdRp (data not shown). p38 was used as a loading control throughout this study. Average percentages relative to the WT value and standard deviations for three independent experiments are shown below panels throughout the figures. (D) RNA gel blot of WT and mutant gRNAs accumulating at 40 hpi in Arabidopsis protoplasts. The locations of the gRNA and the 26S rRNA are shown. Data are from three independent experiments.
FIG 2. Phylogenetic conservation of RSE and SLA structures. (A) Carmoviruses: Cardamine chlorotic fleck virus (CCFV), Saguaro cactus virus (SCV), Melon necrotic spot virus (MNSV), Pelargonium flower break virus (PFBV), Calibrachoa mottle virus (CbMV). Sequences in blue are predicted to form an internal pseudoknot in an alternative structure. Sequences in red are predicted to pair with complementary sequences at the 3’ end of the genome. (B) Carnation Italian ringspot virus (CIRV) is a member of the genus Tombusvirus. Panicum mosaic virus (PMV) is a member of the genus Panicovirus. Umbravirus PEMV2 is unclassified.
lation. These results strongly suggest that in addition to the phylogenetically conserved RSE structure that contains stem S1, the RSE must form an alternative structure without S1 and with an internal pseudoknot for efficient readthrough in vitro.

The structure of the TCV RSE is important for function. To further analyze the importance of RSE sequence and structure, mutations in loops and mutations designed to be compensatory in stems were generated throughout the hairpin (Fig. 3A). If stem alterations could be designed that were both compensatory in the RSE and silent in the RdRp, levels of mutant gRNAs were also assessed in protoplasts. Mutating the terminal loop (m3) or the base pair just below internal loop L3 (m6, m7) did not affect read-through. FIG 3 Analysis of the TCV RSE. (A) Mutations generated in the RSE. Asterisks indicate residues conserved at the readthrough site throughout the Tombusviridae (see panel D). When possible, base alterations were chosen for being silent in the encoded RdRp and for maintaining codon usage similar to that of the WT residue, which allowed for the gRNA to be assayed in vivo (see panel C). The designations of the mutations are given in brackets. (B) In vitro translation of full-length WT and mutant gRNAs in WGE. (C) Accumulation of selected gRNA mutants in protoplasts, assayed using RNA gel blots. The line within the upper right blot indicates that a lane was discarded. (D) Sequence conservation adjacent to the amber codon in carmoviruses and all other members of the Tombus-viridae that were examined. (E) In vitro translation of full-length WT and mutant gRNAs with alterations in the conserved residues. The line indicates that a lane from the blot was discarded.
through or gRNA accumulation \textit{in vivo} (Fig. 3B and C). Mutating the three base pairs just below L3 (m8, m9) reduced PRT to background levels, whereas the compensatory mutation (m10) restored p88 synthesis to WT levels. Altering the base pair in S4 that is adjacent to L3 (m4, m5) decreased gRNA accumulation \textit{in vivo} but did not affect readthrough \textit{in vitro}. The combination of the two alterations was compensatory for gRNA accumulation \textit{in vivo}, indicating that S4 disruption might have consequences that were not discernible in the WGE assay. Altering the guanylate just below L2 to an adenylate (m11) disrupted pairing at this location and reduced PRT to 8% of the WT level and gRNA accumulation to 30% of WT accumulation. Maintaining pairing in this location (m12) had no effect on readthrough and reduced gRNA accumulation by only 26%. gRNA containing both mutations (m13) gave WT levels of readthrough and accumulation \textit{in vivo}, suggesting that a base pair at this position is important. Altering two base pairs simultaneously in stem S2 (m15, m16) reduced readthrough and gRNA accumulation by 20-fold, whereas the compensatory mutation (m17) restored readthrough to 75% of WT levels and gRNA accumulation to WT levels. Stabilizing the S2 stem by converting the adjacent G:U pair and U-U mismatch to G·C and A·U (m14) enhanced readthrough to 275% of the WT level. The latter result correlates with published reports indicating that strengthening the stability of RSE lower stems positively impacts PRT and \textit{−}1PRF (17, 21). These results establish the importance of RSE stems S2 and S3 for stimulating PRT.

Sequence adjacent to the readthrough site is conserved in tombusvirids and is important for readthrough. The readthrough sites in all 15 carmoviruses and in all other tombusvirids examined (Fig. 3D) contain a conserved sequence [UAG GGR (G/U)G YNU] (the termination codon is underlined)], Conservation of the GGR, but not the (G/U)G, 3' of the UAG has been reported previously for viruses in the \textit{Tombusviridae} undergoing readthrough (19), and a similar 5-base purine-rich sequence follows the stop codon between the \textit{gag} and \textit{pol} genes in \textit{Marinella leukemia virus} and is important for readthrough (38, 39). Either the G821C (m18) or the C986G (m19) alteration reduced readthrough to 13% or 16% of WT readthrough, respectively (Fig. 3A and E). The two alterations together (m20), which should re-form the S2 stem, improved readthrough to only 33% of the WT level. This is in contrast with the two nearby base pair alterations (m15 and m16) that decreased readthrough to 5% of the WT level, with the compensatory change (m17) improving readthrough to 75% of the WT level. These results suggest that both base pairing and the identity of G821 and/or C986 are important for readthrough.

To determine the importance of the guanylates adjacent to the termination codon, the three guanylates in positions +1, +2, and +3 were converted to cytidylates (m21), and the opposing cytidylates were converted to guanylates (m22). Because the m22 mutations would disrupt the required pseudoknot, they were combined with corresponding changes in internal loop L3 (m24) to maintain the tertiary interaction. m21 reduced readthrough by nearly 50%, whereas m22 (+m24) had no detrimental effect (Fig. 3E). These results indicate that the S1 stem can be disrupted without negatively impacting readthrough as long as mutations are on the 3' side and the required pseudoknot is not disrupted. The combination of m21 with m22 (compensatory mutation m23), which, along with m24, should restore the S1 stem and maintain the pseudoknot, reduced readthrough from that with m22 + m24. The noncompensatory nature of the residues in stem S1 suggests that one or more of the conserved guanylates just downstream of the amber codon are required for efficient readthrough.

SHAPE structure probing of the RSE region \textit{in vitro} and \textit{in vivo}. SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension) was used to structurally map the SLA/RSE region in full-length TCV gRNA prepared \textit{in vitro} (Fig. 4A to C). SHAPE detects flexible residues whose 2' OH can form covalent bonds with NMIA, which is an impediment for reverse transcriptase-mediated primer extension. Residues with moderately high to high reactivity with NMIA are shown in red, and residues with lower reactivity with NMIA are shown in green. In Fig. 4A, SHAPE data supported the predicted structure of the SLA but was consistent only with the upper portion of the RSE (Fig. 4A to C). In the lower portion of the RSE, three residues on the 3' side of stem S2 that were predicted to be paired were strongly reactive with NMIA, whereas their partner residues were not. This suggests that despite the importance of the S2 stem for PRT \textit{in vitro}, it was not present at discernible levels within the gRNA population that was used for both SHAPE and WGE. The strong reactivity of L3 pseudoknot guanylates, but not their paired cytidylates, also suggests that the PRT-required pseudoknot was not present in the \textit{in vitro}-synthesized gRNA. Internal loop L2, which engages in the long-distance interaction with the Pr loop, contained a mixture of flexible and nonflexible residues. These residues were similarly flexible when gRNA that contained mutations in the Pr loop designed to disrupt the long-distance interaction was used (see Fig. 7E). Thus, this necessary tertiary interaction also was not detected in gRNA prepared \textit{in vitro}.

Since RNA folds cotranscriptionally, the absence of both the phylogenetically conserved stem S2 and the two required tertiary interactions could be due to misfolding of the \textit{in vitro}-synthesized, denatured, and then renatured gRNA. Therefore, SHAPE was repeated using viral gRNA accumulating within infected protoplasts (\textit{in vivo SHAPE}). gRNA that was capable (WT) or incapable (GDD) of amplification (due to mutations in the RdRp active site) was subjected to SHAPE at 40 h postinoculation (40 hpi), since plant viruses are restricted to initially inoculated protoplasts, and little virus accumulates after this point. Cells were treated with 1M7 (a membrane-permeant electrophile that generated an \textit{in vitro} SHAPE profile identical to that with NMIA [data not shown]), followed by cell lysis, total-RNA extraction, and primer extension. As shown in Fig. 4E and F, and in the secondary structure in Fig. 4D, the \textit{in vivo} SHAPE profile for the RSE region at this time point was very similar to the \textit{in vitro} SHAPE profile (excluding consideration of the guanylates in L3 that could not be assessed), suggesting that the structure of the RSE region in the \textit{in vitro}-synthesized gRNA does not represent a kinetically trapped artificial structure. The \textit{in vivo} SHAPE profile was not due to residual input RNA, since no SHAPE data were obtained using the GDD mutant (with the exception of the strong signal in all lanes at the position of the pseudoknot guanylates, which was likely associated with the radiolabeled oligonucleotide). The one major difference between the \textit{in vivo} and \textit{in vitro} SHAPE profiles was additional reactive residues on the 3' side of stem S2, strongly suggesting that at least some portion of the phylogenetically conserved S2 stem is not present in the majority of gRNAs sampled \textit{in vivo} and \textit{in vitro}.

To investigate whether the S1 stem is present in \textit{in vitro}-synthesized gRNA, all four S1 3'-side cytidylates were changed to guanylates (m32), and the mutant gRNA was subjected to \textit{in vitro}}
SHAPE (Fig. 5). The flexibilities of the altered residues on the 3′ side of S1 (and adjacent sequences) were enhanced; however, no change in the flexibility of the 5′-side guanylates was detected (Fig. 5B). In addition, the flexibility of L3 internal-loop guanylates that form a pseudoknot with these cytidylates was also unchanged. When the L3 guanylates were also converted to cytidylates (m32/53) with the hope of driving the formation of the pseudoknot, the mutated L3 residues were no longer flexible. However, the S1 3′ guanylates in the m32/53 mutant showed little change in flexibility from that with m32 alone, suggesting that the pseudoknot still was not forming in the majority of gRNAs in the population and implying that the new cytidylate residues in L3 were likely pairing elsewhere, m32 reduced readthrough by 20-fold in WGE, and m32+53 restored WT readthrough levels, indicating that the pseudoknot is still required and still capable of formation during the translation of the gRNA (Fig. 5C). These results strongly suggest that the S1 stem and RSE pseudoknot are not present in the majority of the gRNAs synthesized in vitro.

Sequences upstream of SLA pair with RSE 5′-side S1 and S2 residues. To determine if the middle portion of the S2 stem exists in gRNA synthesized in vitro, two additional mutant gRNAs were generated (m25, m26), each with seven altered residues in S2 that,
together, are complementary (Fig. 6A). In vitro translation of m25, m26, and m25+m26 gRNAs generated only low (6 to 8% of WT) levels of p88 (Fig. 6B). Probing of the structure of m25 gRNA, which contained mutations on the 5’ side of S2, indicated that the mutated residues, as well as a few adjacent residues, became flexible (a total of 12 nt). On the 3’ side of the RSE, there were only minor changes in the S2 SHAPE profile, while a guanylate (G783) in loop L2 became strongly reactive (Fig. 6C). m26 gRNA, with mutations on the 3’ side of S2, contained residues in the altered region that both gained and lost flexibility, with flexibility changes extending into the L2 internal loop. No flexibility changes were observed on the opposing 5’ side of S2. In m25+m26 gRNA, the mutated residues on the 3’ side were no longer flexible, suggesting that the majority of the S2 stem was now forming. In addition, the upper portion of S2 and the lower portion of L2 contained large numbers of residues with enhanced flexibility, including G783. Taken together, these results strongly suggest that the entire S2 stem, in addition to the S1 stem, does not form in the gRNA synthesized in vitro.

In m25+m26 gRNA, the structure of a portion of L2, which engages in the long-distance interaction with 3’ sequences, was significantly altered when either the S2 stem formed or the alternative pairing of S2 residues was disrupted. The SHAPE profile of m14 gRNA, which contains a strengthened S2 stem (Fig. 3), displayed structural changes in this region similar to those of m25+m26 gRNA (Fig. 7B and D). Since readthrough was enhanced in m14 gRNA (Fig. 3) and depressed in m25+m26 gRNA (Fig. 6B), the altered L2 structure was not by itself responsible for weak readthrough in m25+m26 gRNA. To determine if the S2/L2 structural changes affected the long-distance interaction between L2 and the 3’ end, SHAPE was conducted on m25+m26 gRNA that also contained a 2-base alteration in the 3’-end interacting sequence (mutation m34). No difference in L2 flexibility from that in WT gRNA was found in m34 or m25+m26+m34 gRNA, suggesting that the flexibility changes in L2 did not lead to detection of the long-distance interaction (Fig. 7E). In addition, m34 reduced WT gRNA readthrough by 80% but had no comparable effect on m25+m26 gRNA (Fig. 7F). These results support our conclusion that the S2 stem forms in m25+m26 gRNA but not in WT gRNA synthesized in vitro and that either the presence of the S2 stem or the disruption of alternative pairing of S2 residues affects residue flexibility in and near the L2 internal loop.

A region upstream of SLA is implicated in alternative pairing with RSE S2 residues. In the SHAPE profiles for m25 and m25+m26 gRNAs, a small region located just upstream of SLA (residues 720 to 735) showed enhanced flexibility (Fig. 6C, blue bracket). No discernible differences were detected either upstream or downstream of this region, and no flexibility changes were discernible in this region in m26 gRNA. When residues 720 to 735 were altered (m27) and the SHAPE profile of m27 gRNA was compared with that of WT gRNA, the mutated region in m27 gRNA became more flexible (Fig. 8). In addition, 13 of 26 residues in S2 showed altered flexibility, as did residues in the lower portion of L2. Since most S2 residues had increased flexibility, mutation of residues 720 to 735 did not by itself result in the formation of the phylogenetically conserved S2 stem. The presence of very similar flexibility changes in upper 3’ S2 and lower L2 residues suggests that this flexibility signature (also found for m14 and m25+m26 gRNAs [Fig. 7D and E]) is not directly associated with the formation of the S2 stem but rather reflects a loss of pairing between S2 residues and residues in positions 720 to 735.

Alternative structure of the TCV RSE. When the sequence between residues 720 and 735 was examined, clear base-pairing possibilities with residues on the 5’ side of S1 and S2 were discernible, which extended the length of SLA (Fig. 9). This alternative conformation, labeled the “basal” structure, agrees well with the SHAPE structure profile determined both in vitro (Fig. 9A) and in vivo (Fig. 9B). To determine the importance of SLA and the basal structure, three sets of silent mutations were generated on the 5’ side of the presumptive elongated SLA: mutations of the lower stems S1 and S2 (S1+S2) (m28), the middle stem S5 (m29), and the upper stem S6 (m30) (Fig. 10A). m28 gRNA caused several residues to gain flexibility on the opposing (3’) side of the stem (Fig. 10B). These changes were accompanied by the signature structural changes in the upper 3’ side and lower L2 regions of the RSE (Fig. 10C), as was found when RSE stem S2 was strengthened (m14, m25+m26 [Fig. 7D and E]) and/or when alternative pairing for S2 residues was disrupted (m27 [Fig. 8]). In addition, m28 mutations increased the flexibility of residues in the S5 stem of the SLA, suggesting the presence of higher-order structure within the extended hairpin (Fig. 10B). SHAPE structure probing of gRNA with m29 mutations in SLA S5 showed significant disruptions on both sides of the S5 stem (Fig. 10B). In addition, structural changes very similar to those found for m28 were evident in the S1+S2 region of the extended SLA, accompanied by signature changes in the RSE (Fig. 10B and C). These results support a higher-order structure encompassing residues in the lower half of the basal SLA structure and the L2/upper S2 3’ region of the RSE. In contrast, m30 mutations in the upper stem S6 of SLA enhanced the flexibility of several S6 stem residues without discernibly affecting the flexibility of the remainder of the SLA.

m28 and m29 caused minor (<20%) increases in p88 translation in WGE, whereas m30 did not affect readthrough (Fig. 10D). This suggests that weakening the alternative pairing of RSE S1+S2 may only slightly promote the formation of the active form of the RSE. m29 and m30 reduced gRNA levels in protoplasts to 48% and 28% of WT levels, respectively (Fig. 10E). In contrast, m28 mutations did not affect gRNA levels. Although these results support the importance of SLA for gRNA accumulation in vivo, freeing the 5’ side of the RSE lower stem is not sufficient to cause a discernible effect on virus accumulation.

DISCUSSION
Conserved features of tombusvirid RSEs. We initiated this study with the hypothesis that few overarching mechanisms exist for ribosome recoding because most currently used approaches might be excluding sequences necessary for proper regulation of the event. To provide evidence for or against this hypothesis, we examined RSE regions in tombusvirids involved in either −1PRF or PRT. With few exceptions, these RSEs are ca. 90 nt long, with a G-C-rich lower stem that incorporates the guanylate of the amber stop codon at the base. Tombusvirid RSEs contain at least two large, asymmetric internal loops, with the lower loop engaged in an important long-distance interaction with the 3’ end (28, 30, 31). This RNA bridge had been proposed to form an atypical pseudoknot in the RSE that functionally replaces the pseudoknot commonly found in animal virus RSEs (30). However, our compensatory mutagenesis revealed that a phylogenetically conserved pseudoknot is also a critical feature of the TCV RSE (Fig. 1), sug-
FIG 6 Alternative pairing partners for RSE 5’-side stem S2 residues. (A) (Upper left) WT SHAPE profile. The locations of mutations (m25 and m26) generated in full-length TCV gRNA are shown. Residues are color coded as described in the legend to Fig. 4A. (Top right) Effects of m25 mutations on NMIA-reactive residues in the RSE region. Dark blue and light blue circles designate residues that were more or less reactive, respectively, than WT residues. (Bottom left) Effects of m26 mutations on reactive residues in the RSE region. (Bottom right) Effects of a combination of m25 and m26 mutations that should be compensatory. (B) In vitro translation of full-length WT and mutant gRNAs in WGE. (C) SHAPE mapping of the RSE and SLA regions in WT and mutant gRNAs prepared in vitro. The first three lanes are nucleotide ladder lanes. N, NMIA treated; D, DMSO treated. The blue bracket delineates a region upstream of SLA that becomes flexible when 5’ S2 residues are altered (m25). Note that the mutated residues on the 5’ side of S2 are flexible in m25, whereas the partner residues on the 3’ side of S2 retain WT-like flexibility.
FIG 7 m14 and m25+m26 share similar structural alterations in the upper 3′ S2 and lower L2 regions. (A) Locations of mutations in the RSE that stabilize the stem of S2. (B) Effects of m14 mutations on the SHAPE profile of the RSE. Dark blue and light blue circles mark residues that were more or less reactive, respectively, than WT residues. (C) SHAPE profile of m25+m26 (compensatory mutations in the lower S2 stem) gRNA from Fig. 6A, provided here for purposes of comparison. The hairpin at the right is the 3′-terminal Pr hairpin, whose apical loop is the pairing partner of L2. Residues complementary with the L2 sequence are shown in red. The m34 mutations that disrupt the long-distance interaction are shown. (D) SHAPE mapping of the RSE region in WT and m14 gRNAs prepared in vitro. The first three lanes are nucleotide ladder lanes. N, NMIA treated; D, DMSO treated. Note that the structural changes in upper 3′ S2 and lower L2 are very similar to those found in m25+m26 gRNA (see panel E). (E) SHAPE mapping of the L2 region when the RSE contains m25+m26, and the long-distance interaction is either permitted (m25+m26) or not permitted (m25+m26+m34). This gel also shows that there is no difference in the WT L2 sequence when the long-distance RNA-RNA interaction is disrupted. (F) In vitro translation of WT and m25+m26 gRNAs that can or cannot (+m34) engage in the long-distance RNA-RNA interaction.
gesting that the function of the long-distance interaction is not a simple substitution for a local H-type pseudoknot.

Conserved residues [UAG GGR (G/U)G YNU (the termination codon is underlined)] were also identified just downstream of the amber termination codon in tombusvirids that use PRT to synthesize their RdRp. The conserved guanylates adjacent to the UAG contributed to PRT efficiency, but not as simple participants in the S1 stem. In addition, the G at position G110015(G110015G) was also important for readthrough (Fig. 3). Gammaretroviruses share a similar conserved sequence downstream of their amber termination codon at the gag-pol junction (UAG GGN NGN) (Fig. 11A). Alterations of G110015G, G110012G, and G110015G (but not the intervening G110013A and G110014G) have been reported to be highly detrimental for readthrough in Murine leukemia virus (MuLV) by use of dual-luciferase reporter constructs (8, 39). +5G is located within the ribosome when the amber codon is in the A site, suggesting that specific contacts may exist between this guanylate and ribosome components that support readthrough in these viruses.

The TCV RSE adopts several alternative conformations. Tombusvirid RSEs are structurally conserved, particularly in the lower stem regions (S1 and S2) (Fig. 2). However, SHAPE conducted both in vivo and in vitro did not support the presence of these stems in the TCV RSE. Rather, SHAPE profiles supported the predominance of an alternative structure, where residues on the 5′ side of RSE S1 and S2 pair with sequences located 100 nt upstream, extending the base of the upstream SLA hairpin. Hairpins in the same location as SLA are predicted to form upstream of RSEs in all tombusvirids examined (Fig. 2). A hairpin in this loca-
tion was also previously reported in BYDV (26), as well as in umbraviruses (21) and coronaviruses (33), where they serve as repressors of recoding. The PEMV2 SLA strongly repressed frameshifting from the slippery sequence when the RSE structure was absent, possibly by blocking the backward movement of the ribosome (21). In TCV, mutations disrupting the middle and upper portions of the extended TCV SLA reduced gRNA accumulation in vivo (Fig. 10E). Although SLA in tombusvirids differ in sequence and structure, strong conservation of a hairpin just upstream of the recoding site suggests a role for this hairpin in both readthrough and frameshifting.

While SHAPE only provides a structural snapshot of the majority of RNAs sampled, our results demonstrate that the structure with expanded SLA and truncated RSE (designated the basal structure) is the predominant form in vitro and at 40 hpi in vivo, when little if any translation would be occurring (plant viruses are restricted to initially infected cells in culture). We propose that the basal structure is the readthrough-inactive form of the RSE, which, when present, results in the ribosome terminating translation at the p28 amber codon. A major issue that remains unresolved is the pairing location(s) for residues on the 3' side of the RSE in the basal structure. Although a number of possible pairing sites both upstream and downstream of the RSE were examined, none of these locations were structurally altered by mutations in the 3' side of the RSE (and elsewhere) that would have been disruptive (data not shown). m27 mutations, which extensively disrupted the pairing partner for the 5' side of RSE S1/H11001 S2, did not lead to widespread formation of the phylogenetically conserved RSE S1/H11001 S2 stem (Fig. 8B). This suggests that it may also be necessary to free the 3' side of the RSE S1+S2 stem from its paired location before the S1+S2 stem can form. Silent m28 mutations that disrupted four of the pairing partners for the 5' side of RSE S1+S2 were not detrimental in vivo (Fig. 10E), suggesting that the basal structure locations for both sides of the RSE may need to be modified before an effect on virus accumulation is discernible.

Although no structural evidence for the phylogenetically conserved active form of the TCV RSE currently exists, genetic evidence clearly demonstrated that an RSE containing the S2 stem is critical for recoding in vitro (Fig. 3). The phylogenetically conserved structure that includes stem S1 is also not compatible with
Extended SLA contains higher-order structure and is important for gRNA accumulation in vivo. (A) Locations of the m28, m29, and m30 mutations in the basal and active structures of the RSE. Each set of mutations has a different background color for its residues: blue for m28, green for m29, and pink for m30. (B) Effects of SLA mutations on the SHAPE profiles of SLA and the RSE. Dark blue and light blue circles designate residues that were more or less reactive, respectively, than WT residues. Note that m28 and m29 cause nearly reciprocal structural changes in the adjacent stems. (C) Effects of SLA mutations on the structure of the RSE. Note that m28 and m29 cause alterations in residue flexibility in lower L2 and adjacent upper S2 (3') similar to those caused by m14, m25+m26, and m27 (Fig. 6 and 7). (D) In vitro translation of full-length WT and mutant gRNAs in WGE. (E) RNA gel blot of WT and mutant gRNAs accumulating at 40 hpi. The line indicates that a lane was removed from the blot.
the required pseudoknot between S1 and L3 residues. This suggests that a series of conformational changes are likely needed for the recoding event. One possible scenario would have ribosomes unwinding both sides of the lower stem of the extended SLA basal structure during translation, which causes the signature structural changes in RSE L2 that might lead to the establishment of the long-distance interaction with the 3′ end, possibly helping to stabilize the RSE active structure, leading to ribosome pausing and insertion of a tRNA that decodes the amber stop codon.

Evidence for additional RSE conformations in unrelated viruses. Since viral RNAs are known to adopt alternative conformations of overlapping sequences to facilitate transitions between reversible processes (46), conformational switches modulating ribosome recoding might be widespread. In addition to West Nile virus (24), the presence of alternative structures has been suggested by SHAPE structure modeling of the RSE region in full-
length HIV in vitro, which indicated that the previously proposed two-helix model (9, 47, 48) was missing upstream and downstream sequences that together generate a more complex structure (49). Ribosomes translating through this larger frameshifting domain were proposed to remodel portions of the extended structure into the conventionally accepted two-helix active structure. The RSE of MuLV, in addition to having the same conserved guanylates proximal to the termination codon as the TCV RSE (38, 39) (Fig. 1B), can also adopt an alternative hairpin conformation (50, 51). Mutations specifically affecting this alternative hairpin structure (but not the pseudoknotted structure) were highly detrimental to virus viability (51). These data are consistent with MuLV also having basal and active conformations of its RSE.

Since the severe acute respiratory syndrome coronavirus (SARS-CoV) SLA hairpin was also repressive for recoding (33), we examined the sequence of the SARS-CoV SLA/RSE region to determine whether an alternative conformation that extends the length of the SLA, as in TCV, might be possible. As shown in Fig. 11C, an alternative structure that extends the length of SLA and contains a 10-bp stem is predicted. This SLA extension is formed by pairing sequence upstream of SLA with sequence from the lower stem by base-pairing residues upstream of SLA with RSE residues, thereby disrupting the structure of the active RSE.

In conclusion, current models proposed to explain how RSEs mediate either −1PFR or PFT at the precise ratios found in infected cells require the assumption that all sequences necessary for regulating recoding are encompassed within the recoding site and downstream RSE (13, 20, 25). However, the original strategies used to identify active RSEs would have eliminated repressive/regulatory upstream hairpins that might be key components in the formation of alternative basal forms of RSEs. Our results support the necessity of examining recoding within full-length viral RNA in order to account for the full spectrum of RNA elements necessary to precisely regulate this critical event in the life cycle of RNA viruses.

ACKNOWLEDGMENTS

We thank all members of the Simon lab for helpful discussions and Pamela Katali for technical assistance. We also thank J. Dinman for providing the 1M7. This work was supported by the National Science Foundation (MCB-1411836) and National Institutes of Health (R21AI177882-01) (to A.E.S.). M.M.K. and M.C. were partly supported by NIH Institutional Training Grant 2T32AI051967-06A1.

FUNDING INFORMATION

This work, including the efforts of Anne E. Simon, was funded by HHS | National Institutes of Health (R21AI177882-01). This work, including the efforts of Micki M. Kuhlmann and Maistry Chattopadhyay, was funded by HHS | National Institutes of Health (NIH) (2T32AI051967-06A1). This work, including the efforts of Anne E. Simon, was funded by National Science Foundation (NSF) (MCB-1411836).

REFERENCES


