Biologically-supported structural model for a viral satellite RNA

Peter Ashton[1], Baodong Wu[1], Jessica D'Angelo[1], Jörg Grigull[2] and K. Andrew White[1,*]

[1]Department of Biology, York University, Toronto, Ontario, M3J 1P3 Canada and [2]Department of Mathematics and Statistics, York University, Toronto, Ontario, M3J 1P3 Canada

ABSTRACT

Satellite RNAs (satRNAs) are a class of small parasitic RNA replicon that associate with different viruses, including plus-strand RNA viruses. Because satRNAs do not encode a polymerase or capsid subunit, they rely on a companion virus to provide these proteins for their RNA replication and packaging. SatRNAs recruit these and other required factors via their RNA sequences and structures. Here, through a combination of chemical probing analysis of RNA structure, phylogenetic structural comparisons, and viability assays of satRNA mutants in infected cells, the biological importance of a deduced higher-order structure for a 619 nt long tombusvirus satRNA was assessed. Functionally-relevant secondary and tertiary RNA structures were identified throughout the length of the satRNA. Notably, a 3′-terminal segment was found to adopt two mutually-exclusive RNA secondary structures, both of which were required for efficient satRNA accumulation. Accordingly, these alternative conformations likely function as a type of RNA switch. The RNA switch was also found to engage in a required long-range kissing-loop interaction with an upstream sequence. Collectively, these results establish a high level of conformational complexity within this small parasitic RNA and provide a valuable structural framework for detailed mechanistic studies.

INTRODUCTION

Plus-strand RNA viruses represent the largest group of plant-infecting viruses. Some of these viruses are associated with small subviral RNAs, termed satellite (sat) RNAs, which depend on a companion or 'helper virus' for their reproduction (1,2). Typically, these parasitic RNAs share little sequence identity with their helper virus genome and are often non-coding. The latter property dictates that they utilize the RNA-dependent RNA polymerase (RdRp) en-

coded by their helper virus for replication, which, as for their helper genome, occurs via synthesis of a minus-strand RNA intermediate (3,4). SatRNAs must also pilfer the capsid protein (CP) produced by the helper virus for the encapsidation of their RNA (1,2). In contrast, satRNA variants, termed satRNA viruses, encode their own distinct CP, but are still dependent on their helper genome for RNA replication (2). In both cases, these subviral RNAs must maintain RNA elements that identify them as templates for RNA replication and packaging by the viral RdRp and their CP, respectively (1–4).

SatRNAs of plus-strand RNA plant viruses can be either linear or circular single-stranded RNA (1,2). For circular satRNAs, the structures and catalytic activities of their self-cleaving ribozymes have been extensively characterized (5) and global RNA structure-function analyses of the satRNA of Cereal yellow dwarf virus allowed for the identification of domains corresponding to promoters, ribozymes, and a packaging signal (6). In linear satRNAs and satRNA viruses, promoters for synthesis of minus- and plus-strands are located at terminal positions (4), while the packaging signals for satRNA viruses can exist as multiple dispersed secondary structures (7–11). Determinants of encapsidation in linear satRNAs have not yet been reported.

Detailed studies on specific regions in linear satRNAs, most often terminal regions, have led to some important advances in the understanding of their mode of reproduction (4). Currently, the best structurally characterized example is satC of Turnip crinkle virus (TCV), which is a chimeric molecule that consists of a smaller satRNA at its 5′ end, satD, which is connected to a 3′-terminal portion of the TCV genome (12). The virally-derived portion of satC has been extensively studied and includes many functionally important secondary and tertiary RNA structures, including a 3′-terminal conformational switch, which are important for satC replication (13–20). The activity of the 5′ half of satC corresponding to satD is less well characterized, but terminal promoters for plus-strand synthesis have been defined (21–23) and a large internal hairpin is known to contribute to its accumulation (24). In the satRNA of Bamboo mosaic virus (satBaMV), the 3′-terminal sequence is structurally similar to that of its helper genome, however

*To whom correspondence should be addressed. Tel: +1 416 736 5243; Fax: +1 416 736 5698; Email: kawhite@yorku.ca

direct structural equivalence at the functional level is not observed, indicating the presence of satRNA-specific features (25). Replication of satBaMV is also dependent on a large 5′-proximal stem loop structure that, additionally, is a major determinant of interference with BaMV replication (26–31).

Defining global and higher-order RNA structures in subviral RNAs and investigating their functionality is an important step toward determining how different RNA-directed processes occur. Notably, studying the entire molecule, as opposed to fragments, is important because structural components may fold and/or function differently when positioned in a complete context. The global structure of the RNA of satellite Tobacco mosaic virus (STMV) has been studied by both atomic force microscopy (9,32) and cryo-electron microscopy (33), and the results are consistent with biochemical data suggesting a branched configuration (9,10). Secondary structure models based on chemical probing data of full-length RNAs have also been generated for STMV (9–11), however, biological activity has been reported for only a few substructures (8). Similarly, complete structure models based on solution probing have been reported for satC of TCV and satRNA of Cucumber mosaic virus (CMV satRNA), but not all of the predicted substructures have been assessed for biological relevance (24,34).

Tombusviruses are among the best experimentally-explored plus-strand RNA plant viruses, both at the protein and RNA level (35–40). Two members of this genus, Cymbidium ringspot virus (CymRSV) and Tomato bushy stunt virus (TBSV), have cognate satRNAs (41–43). The satRNA associated with CymRSV (sat-Cym) has been most extensively characterized, and several noteworthy features have been uncovered (44–48). First, the 5′-proximal one-third of sat-Cym adopts similar secondary and tertiary structures to those present in the 5′ UTRs of helper tombusvirus genomes, thus some structural correspondence exists between this satRNA and its helper (48–50). Second, replication of sat-Cym requires that its helper genome is also replicable (51). That is, sat-Cym replication in cells does not occur when it is provided with the viral replication proteins *in trans* from a non-viral source (51), suggesting a mechanistic link between viral genome replication and satRNA replication. Third, and paradoxically, the infectious clone of CymRSV is a poor helper for sat-Cym replication, which was recently determined to be due to a mutation gained during cloning that altered one of its replication proteins (47). Fortunately, the infectious clone of TBSV (T100) is also a very efficient helper for sat-Cym (Figure 1) and has been used successfully in previous analyses of this satRNA (48).

In terms of higher-order structural features, only the 5′-proximal region of sat-Cym has been functionally characterized (48). Here we report a complete structure-function model for this satRNA. Through chemical probing analysis, phylogenetic structural comparisons and mutational analyses, we deduced and assessed the biological importance of different higher-order structures and identified functionally-relevant secondary and tertiary structures that contribute to sat-Cym viability.
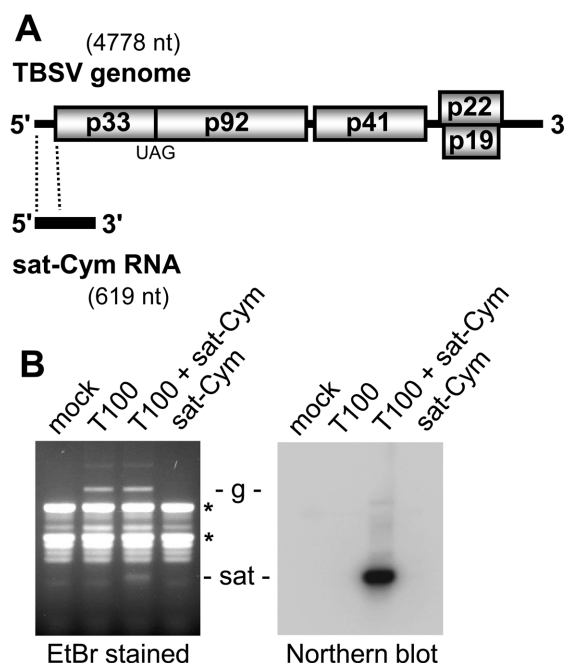

Figure 1. TBSV RNA genome and sat-Cym RNA. (**A**) Linear representation of the TBSV genome showing encoded open reading frames. P33 is an accessory replication protein, while its readthrough product, p92, is the viral RdRp. P41, p22 and p19 are the capsid, movement and suppressor of gene silencing proteins, respectively. The sat-Cym RNA is shown below with the 5′-proximal portion that is structurally equivalent to the TBSV 5′ UTR delineated by dotted lines. (**B**) Northern blot analysis of TBSV helper virus (T100) and/or sat-Cym infections of protoplasts. In the left panel, total nucleic acids isolated from infections were separated in a 1.5% agarose gel and then stained with ethidium bromide. The RNA transcripts used to transfect the cells are shown above each lane and the positions of the TBSV genome and sat-Cym RNA in the gel are indicated. Major rRNAs are indicated by asterisks. The right panel represents a northern blot of the gel on the left that was hybridized with sat-Cym-specific ^{32}P-labeled oligonucleotide probes.

MATERIALS AND METHODS

Construction of satRNA mutants

An infectious clone of the wild-type (wt) sat-Cym (sat181), which has been described previously (48), was used to generate all satRNA mutants using standard PCR-based mutagenesis. The regions corresponding to the PCR-based inserts in mutants were completely sequenced to ensure that only the desired changes were present. The specific changes introduced into each mutant are depicted in the accompanying figures.

In vitro transcription

In vitro transcription, using T7 RNA polymerase, was carried out on SmaI-linearized cloned DNA templates of the TBSV genome (T100) or sat-Cym RNA, as described previously (48). The concentration and integrity of the RNA transcripts were determined by spectrophotometry and gel analysis, respectively.

RNA structure analysis

Four pmol of in vitro transcribed RNA in 24 μl of 0.5×
TE buffer (5 mM Tris, pH 8, 0.5 mM EDTA) was heated
for 5 min at 95°C and then transferred to ice for 2 min. Af-
ter adding 12 μl of 3.3× folding buffer (333 mM HEPES
pH 8.0, 16.5 mM MgCl₂, 333 mM NaCl), RNA was re-
natured by incubating at 37°C for 30 min. Eighteen mi-
croliters of the renatured RNA was treated with 2 μl of
50 mM 1-methyl-7-nitroisatoic anhydride (1M7) in DMSO
and another 18 μl aliquot of the renatured RNA was treated
with 2 μl neat DMSO. Both were subsequently incubated at
37°C for 4 min. 1M7-modified RNA (+), and DMSO con-
trol RNA (−) were recovered by adding 20 μl of a solution
containing 400 mM NaCl, 4 mM EDTA and 40 μg glyco-
gen, followed by ethanol precipitation. Both RNAs were
resuspended in 20 μl of 0.5× TE and analyzed by primer
extension using fluorescently-labeled primers, as described
previously (52). Selective 2′-hydroxyl acylation analyzed
by primer extension (SHAPE) electropherogram intensities
were quantified using *SHAPEfinder* (53) and the average re-
activity at each position from two separate SHAPE analy-
ses of the sat-Cym RNA were incorporated into the *RNAs-
tructure* software (54,55), as described previously (52); ex-
cept that updated pseudo-free energy parameters of $m =$
1.9 and $b = $ –0.7 were used (56). Formation of the previ-
ously confirmed RNA structures in the 5′-region, the T-
shaped domain (TSD), stem 4 (S4) and the down-stream
domain (DSD) (48) (Figure 2), was added as a constraint
in the input file for *RNAstructure*, and no limit was im-
posed on the maximal base pairing range. The 3′ region, for
which SHAPE data were incomplete, was modelled using
a combination of the available partial SHAPE data for the
5′-proximal portion of this region, thermodynamics-based
structure prediction and comparative structural analysis, as
described in the Results section. The two structures that
were consistent with the three types of analyses used are pre-
sented in Figure 2. Sat-Cym RNA structures were generated
using *RnaVis2* software (57).

Protoplast infections

Cucumber cotyledon protoplasts were prepared from 6 to
7 day-old plants, as described previously (48). Coinfections
were initiated by transfecting ∼3 × 10⁵ protoplasts with 3
μg of TBSV helper genome transcript (T100) and 1 μg of
sat-Cym transcript. Transfected protoplasts were incubated
at 22°C under constant fluorescent light for 24 h.

SatRNA accumulation was determined by northern blot
analysis using sat-specific probes, as previously described
(48). Quantification of satRNA was carried out by radio-
analytical scanning of blots, and values with standard error
from three independent experiments are presented.

RESULTS

Secondary structure model for sat-Cym RNA

To gain insight into higher-order structural features present
in the sat-Cym RNA we employed SHAPE methodology,
which assesses the flexibility of individual nucleotides in
an RNA molecule (54). These chemical probing results
were then integrated into the thermodynamics-based RNA
secondary structure predicting program *RNAstructure* and
used to generate a structural model for sat-Cym. In the 5′-
terminal region, which corresponds to nucleotides 1–250,
the previously functionally-validated substructures TSD, S4
and DSD were constrained to form in the structural predic-
tion (Figure 2A, aqua shading) (48). Additional sequence
in this region (between the two halves of S4) was not con-
strained, and adopted four stem loop (SL) structures (SL5
through SL8) that were accompanied by two 6 to 7 nt long
single-stranded spacer segments. The central region of sat-
Cym, nucleotides 250–450, contained two small (SL10 and
SL12) and two large (SL11 and SL13) SL structures (Figure
2A). The helix in the smallest of these, SL12, was uninter-
rupted, whereas those in the other three included internal
loops and/or bulges. Only partial SHAPE data was avail-
able for the 3′-terminal region encompassing nucleotides
450–619. It is possible to obtain SHAPE data on 3′-terminal
regions by adding extra sequence to the 3′ end of the ex-
perimental RNA, which provides a downstream priming
site (58). However, we found that addition of the primer
sequence used by Merino et al. (58) to the 3′-end of the
sat-Cym RNA severely inhibited its accumulation in cells,
suggesting that its normal RNA structure was disturbed
by the extra segment. Consequently, this approach was not
used and the structure of the 3′-terminal region was in-
stead modelled using a combination of the partial SHAPE
data available for the 5′-proximal portion of this region (i.e.
nucleotides 450–515; Figure 2A), thermodynamics-based
structure prediction, and comparative structural analysis.
Interestingly, the results supported the existence of two dif-
ferent conformations for the 3′-terminal region (Figure 2A
and B), which will be described in later sections.

A novel tertiary interaction in sat-Cym RNA

One limitation of the structure modelling approach used
here is that it does not predict tertiary interactions.
Nonetheless, candidate tertiary interactions were sought
by looking for inconsistencies between predicted single-
stranded regions and their corresponding reactivity data.
For instance, the terminal loop of SL11 in the central region
contained unpaired nucleotides that showed lower levels of
reactivity, suggesting that these residues could have pair-
ing partners elsewhere in the RNA molecule (Figure 2A).
Analysis of the sat-Cym sequence for potential partner se-
quences located a 6 nt long complementary sequence in the
3′ region within the loop of SL16 (Figure 2A). To address a
possible role for this interaction, compensatory mutational
analysis of the partner sequences was performed (Figure
3). Sat-Cym mutants containing substitutions targeting two
base pairs in the partner sequences were tested via co-
transfection with helper TBSV genome (T100) into plant
protoplasts (Figure 3B, upper panel). Assessment of sat-
Cym accumulation by northern blotting revealed that dis-
ruptive mutants (T-1A and T-1B) abolished sat-Cym accu-
mulation, while the restoration of pairing with transposed
bases (T-1C) partially recovered accumulation to ∼30% that
of wt. When additional compensatory mutant series that in-
dividually targeted these base pairs were tested, one yielded
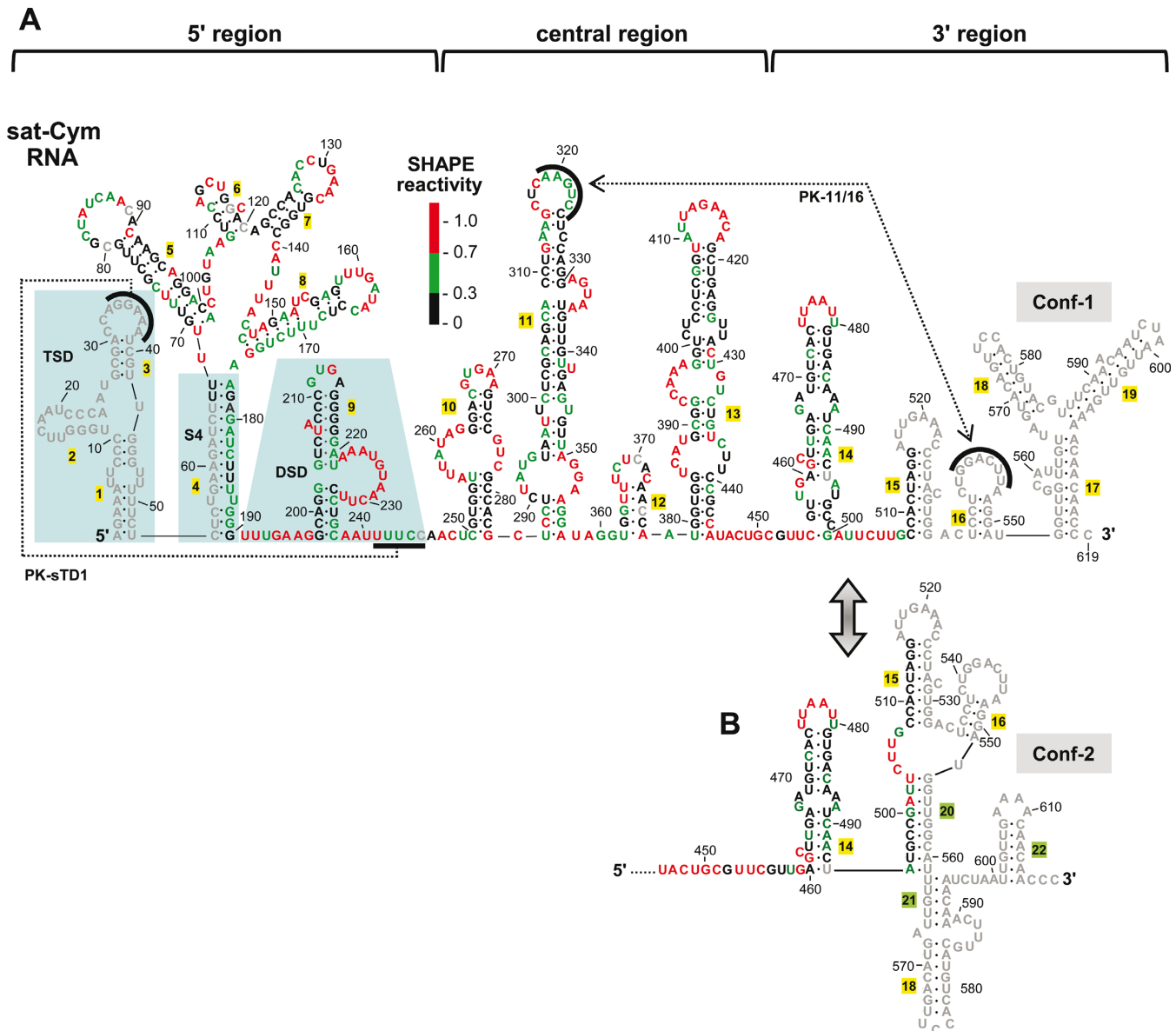results similar to that observed for the double base pair mu-

Figure 2. Structural model for the sat-Cym RNA. (**A**) SHAPE-guided prediction of sat-Cym RNA secondary structure. The brackets at the top divide the RNA into three regions: 5′, central, and 3′. Nucleotides are colour-coded according to their relative SHAPE reactivities, as indicated in the colour key. Nucleotides in grey correspond to regions for which no SHAPE data were available. Residue coordinates are provided and stems are labeled left to right with ascending numbers that are highlighted in yellow. Note that S4 in this figure corresponds to previously defined S5 in reference 48; the S4 designation will be used henceforth. The 5′-proximal structural features that are equivalent to those found in tombusvirus 5′ UTRs are shaded in aqua. Sequences that participate in long-range tertiary RNA-RNA interactions are indicated by black bars that are joined by dotted lines. (**B**) Conformation 2 (conf-2) represents an alternative structure for the 3′ region of conf-1 in (A). Alternative stems in conf-2 are numbered and highlighted in green.

tant (Figure 3B, compare bottom and top panels), whereas the other mediated very efficient recovery (Figure 3B, middle panel). Together, these results provide solid support for the functional importance of the long-range interaction (termed PK-11/16) between the terminal loops of SL11 and SL16 (Figure 3A).

Functional analysis of the central region of sat-Cym

Predicted helical regions and terminal loops in the central region were targeted with substitutions, and the resulting sat-Cym mutants were tested in protoplast infections.

The roles of helical regions in SL10 through SL13 were assessed by compensatory mutational analysis, while their terminal loops were evaluated by replacing them with a GAAA tetraloop sequence (Figure 4). Modifications in SL10 were generally well tolerated, except for mutant series PA-28, where stability of the upper helix correlated moderately with function (Figure 4). For SL11, mismatches in the lower helical region were largely neutral (mutant series JD-2 and PA-26), whereas alterations to the upper stem region and terminal loop were highly detrimental (mutant series PA-27 and mutant PA-34, respectively) (Figure 4). This latter result is consistent with the involvement of the terminal loop
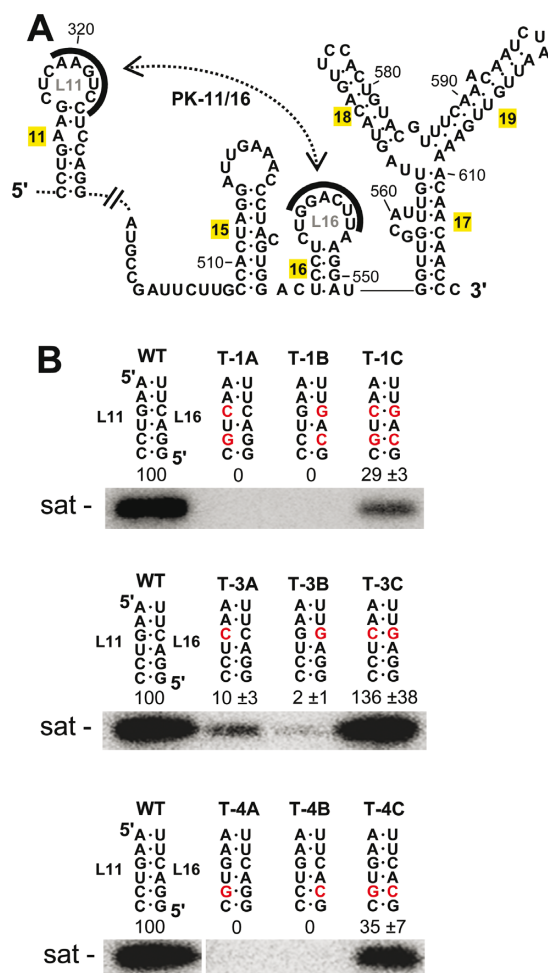
Figure 3. Mutational analysis of a tertiary interaction in sat-Cym RNA. (**A**) The predicted long-range interaction between L11 and L16 that forms PK-11/16 is shown. (**B**) Compensatory mutational analysis of the PK-11/16 interaction showing the substitutions in each mutant and their relative accumulation levels determined by northern blot analysis. Total nucleic acids were isolated from protoplasts co-transfected with helper TBSV genome and wt or mutant sat-Cym RNA. Substituted nucleotides are shown in red for each mutant. The values represent the mean percentages from three independent infections with standard errors.

of SL11 in the PK-11/16 long-range interaction (Figure 3) and implicates the upper stem as being important for presenting the loop sequence. In the 3′-adjacent small hairpin, SL12, stem stability was of intermediate importance, while the loop sequence was dispensable (mutant series JD-3 and mutant JD-4, respectively) (Figure 4). Finally, with the exception of one internal stem (mutant series JD-5), both helical regions and the terminal loop of SL13 were found to be functionally relevant, with the latter being of greatest significance (mutant PA-33) (Figure 4). Accordingly, several substructures within the central domain of sat-Cym contribute to its viability in cells.

Structural analysis of the 3′ region of sat-Cym

As limited SHAPE data were available for the 3′ region (i.e. nts 450–619), a combination of partial SHAPE data, thermodynamics-based structure prediction, and comparative structural analysis, was used to deduce probable structures. By surveying different potential secondary structures, two alternative conformations, termed conf-1 and conf-2, were selected as candidate structures for the 3′ region (Figure 2A, B). The rationale for selecting these two structures was derived from comparative structural analysis which revealed that they could potentially form in both sat-Cym RNA and its most closely-related satRNA counterpart, sat-L RNA (Figure 5) (43). In sat-Cym, the 3′-terminal sequence in conf-1 formed a 3-helix branched structure (S17, SL18 and SL19) with an unpaired terminal cytidylate (Figure 5A, upper panel). Refolding of this sequence into conf-2 would involve unpairing of residues in S17 and S19 and their repairing to alternative partner sequences to form three new helices S20, S21 and SL22, terminated by three unpaired 3′-terminal cytidylates (Figure 5A, lower panel). In such a transition, the 5′-half of S20 in conf-2 would be provided primarily by the lower 3′ portion of SL14 in conf-1. Consequently, the upper part of SL14 would remain intact in both conformations, as would SL15, SL16 and SL18 (Figure 5A). The proposed conformations were also consistent with the available SHAPE data for sat-Cym in terms of the formation of SL14 and SL15 in conf-1 (Figure 2A) and formation of SL15, S20 and the smaller SL14 in conf-2 (Figure 2B).

Notably, comparable secondary structure-based conformations were also predicted to form in sat-L RNA, but with significantly different primary sequences, thus supporting the importance of the proposed higher-order structures (Figure 5B). With the exception of SL15, the correspondence of structural elements in the two sets of conformations is apparent, and include: (i) a 3-helix branched structure in conf-1, (ii) a single unpaired 3′-terminal cytidylate in conf-1, (iii) three consecutive similarly-spaced helices in conf-2, and (iv) three unpaired 3′-terminal cytidylates in conf-2 (Figure 5A and B). This high level of conservation of these distinct structural features suggested that the proposed mutually-exclusive RNA conformations are functionally relevant.

Functional analysis of the 3′ region of sat-Cym

Mutational analysis was performed to assess the possible relevance of the proposed alternative conformations in the 3′ region (Figure 6). Initially, substructures that were predicted to be maintained in both conformations (i.e., SL15, SL16, SL18 and the upper third of SL14) were mutationally-targeted and assessed for functional relevance via protoplast infections, as described earlier. The results indicated that the upper helical region of SL14 contributed to sat-Cym viability (mutant series PA-23 and PA-22), whereas the terminal loop was not important (mutant PA-32) (Figure 6A). For SL15, the lower portion of its stem was most relevant (mutant series PA-20 and PA-21), while its terminal loop was of moderate importance (mutants JD-7 and PA-31). Consistent with its demonstrated involvement with a long-range interaction (i.e. PK-11/16), both the stem and loop of SL16 proved to be highly pertinent (mutant series JD-8 and mutant JD-9). The stem of SL18 was also important (mutant series PA-4), but its terminal loop sequence was expendable (mutant PA-30) (Figure 6A). The re-
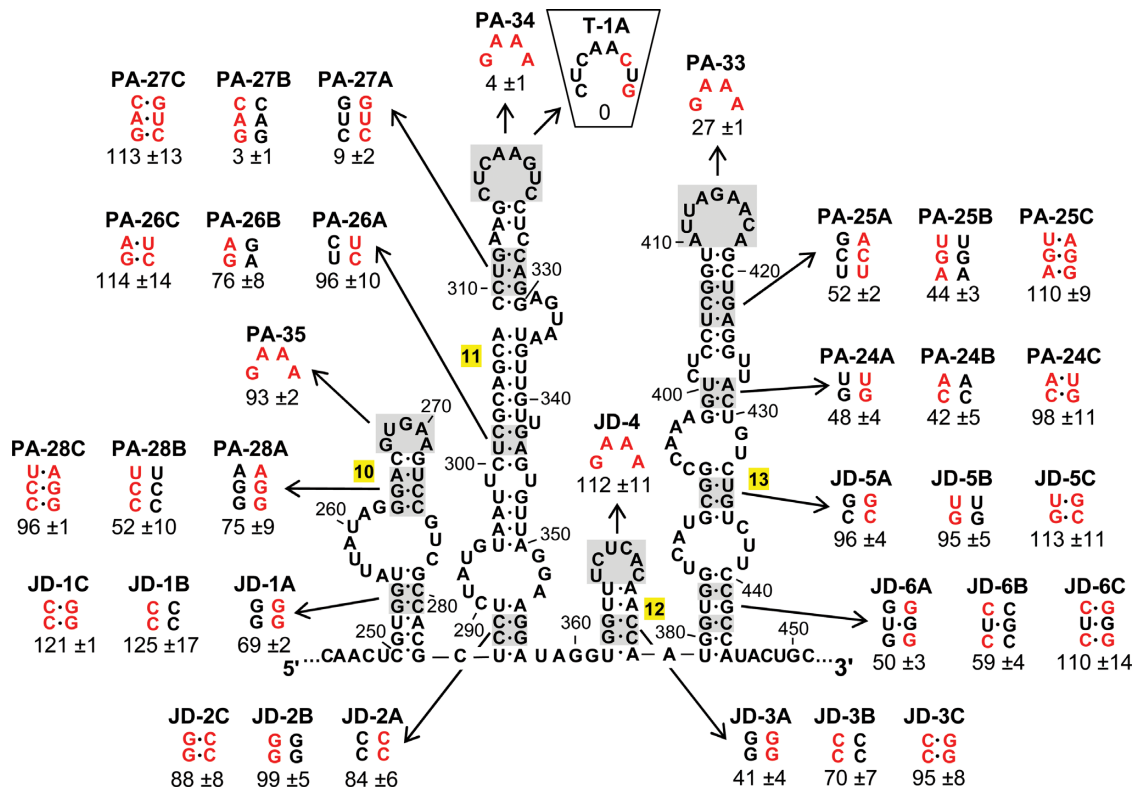
Figure 4. Mutational analysis of the central region of sat-Cym RNA. Regions in the structure that were targeted are shaded and the nucleotides that were substituted in corresponding mutants are shown in red. For reference, mutant T-1A (in the trapezoid) from Figure 3 is also shown. The values represent the mean percentages from three independent infections with standard errors. Analysis was performed as described in the legend to Figure 3.

sults of these compensatory mutational analyses were therefore consistent with the proposed invariant nature of SL15, SL16, SL18 and the upper portion of SL14 in the two proposed conformations. That is, the high level of recovery observed for mutants with transposed base pairs makes it very unlikely that the segments forming these helices have functionally-relevant alternative partner sequences.

Next, we sought to address the relevance of the proposed dynamic portion of the 3′ region. Compensatory mutations introduced into the small 3′-terminal hairpin SL22 in conf-2 showed no recovery when base pairing was restored in mutant PA-5C (Figure 6B). When the positions of the substituted residues were mapped onto conf-1, their locations were predicted to destabilize S17 and S19 (Figure 6A, blue boxes), which would explain the lack of recovery of activity in mutant PA-5C. Consequently, when mutants to assess the importance of S20 in conf-2 were designed, the two base-paired positions targeted were selected to minimize destabilization of corresponding paired regions in conf-1 (Figure 6A, green boxes). In the mutant series T-2, disruption and restoration of S20 pairing correlated with low and high levels of sat-Cym, supporting a functional role for this helix (Figure 6B). Additional sets of compensatory mutations were also introduced into S17 and S19 in conf-1 to test their significance (mutant series PA-1, PA2 and PA-3) (Figure 6A). In all three cases, no recovery of activity was observed when pairing was restored, and, in all instances, the substitutions were predicted to disrupt S20, S21 and/or S22 in conf-2 (Figure 6B). In contrast, substitu-

tion of the terminal loop sequence of SL19 in conf-1 (mutant PA-29) had a relatively minor effect (Figure 6A) and these changes mapped to a predicted single-stranded region in conf-2 (Figure 6B, black box). Taken together, these results are in agreement with the requirement for both conformations, because mutations that were predicted to allow for the formation of only conf-1 (mutants PA-1C, PA-2C and PA-3C) or only conf-2 formation (PA-5C) were detrimental, while those expected to support formation of both conformations (T-2C and PA-29) were permissive.

To obtain further support for the two-conformation model, additional mutants were constructed. In mutant PA-7, a transposed base pair in the 3′-terminal hairpin, S22, notably impaired sat-Cym (Figure 7B). However, combining these two substitutions with an additional substitution in S21 in conf-2 (mutant PA-7+U) (Figure 7B), which regenerated pairing potential in S19 in conf-1 (Figure 7A, upper blue box), restored activity to wt levels (Figure 7B). Inhibition of sat-Cym levels was also observed for a second defective mutant with a different transposed base pair in S22 (mutant PA-9), and the defect was partially restored by adding a third substitution that allowed for S19 pairing (mutant PA-9+U) (Figure 7B). In both sets of mutants above, a UU mismatch would be present in S17 in conf-1 (Figure 7A, lower blue and green boxes); however, these mismatches were tolerated. Accordingly, the results support a strong requirement for both S19 in conf-1 and S22 in conf-2, and suggest a degree of flexibility for pairing of S21 in conf-2 and S17 in conf-1.
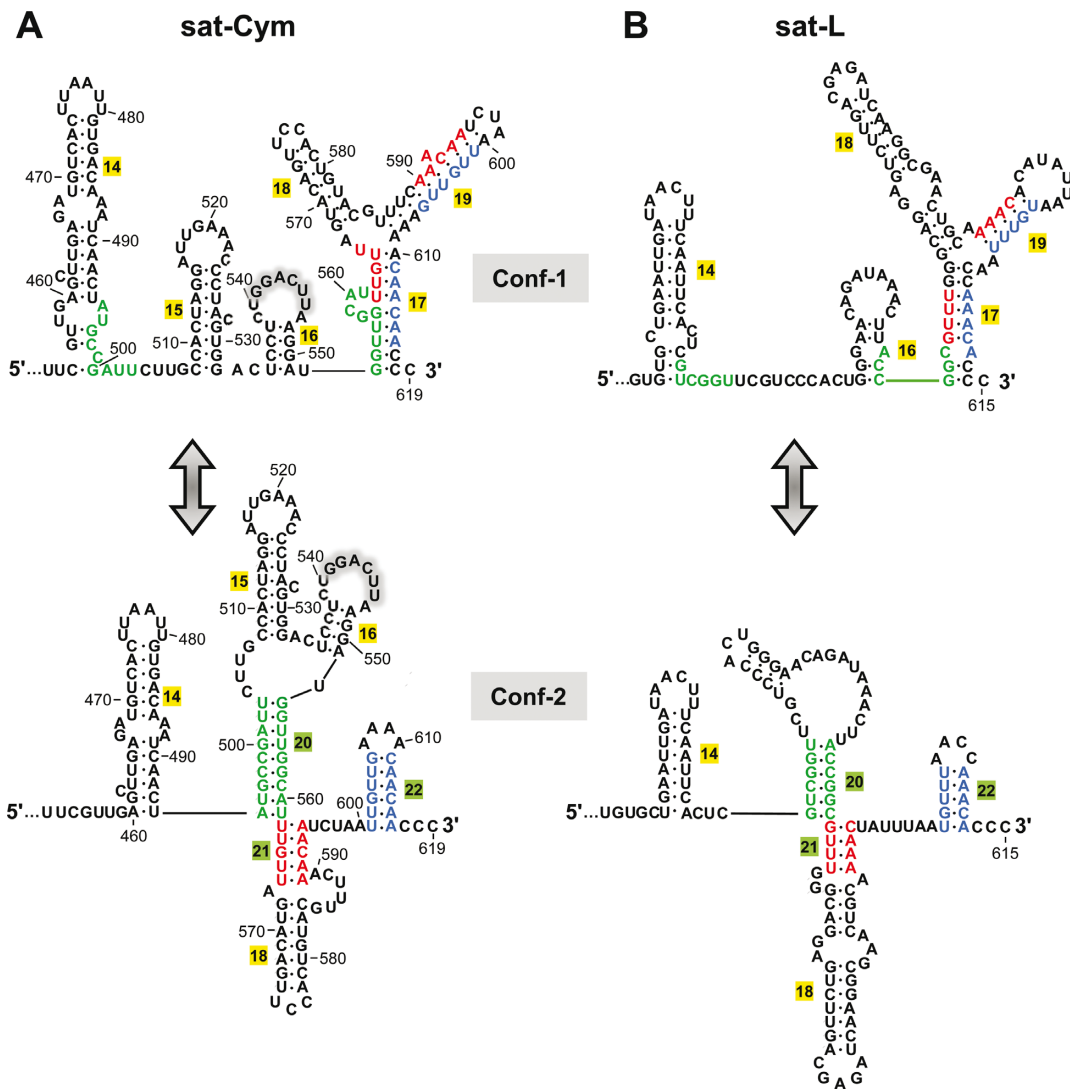
Figure 5. Structural equivalence of alternative 3′-terminal conformations in sat-Cym and sat-L. (**A**) Conf-1 and conf-2 for sat-Cym. In the upper panel, the sequences in conf-1 that form alternative stems in conf-2 are colour coded. The grey shaded nucleotides in the loop of SL16 participate in the long-range interaction PK-11/16. In the lower panel, alternative stems in conf-2 are numbered and highlighted in green. (**B**) Corresponding predicted conformations in sat-L.

Based on these findings, we anticipated that mutant PA-13-1, which maintained pairing potential with both S19 in conf-1 and S22 in conf-2 with a UU mismatch in S17 in conf-1, would also show moderate to high levels of activity, but only minimal sat-Cym levels were observed (Figure 7A). However, when these changes were combined with a fourth substitution that restored S17 pairing (mutant PA-13-1+2), near wt levels of sat-Cym were detected; whereas, on its own, the latter substitution was highly inhibitory (mutant PA-13-2) (Figure 7A). A comparable mutational strategy at different positions in mutants PA-12-1, PA-12–1+2 and PA-12-2 yielded consistent results (Figure 7A). Thus, in addition to S19 in conf-1 and S22 in conf-2, the formation of S17 in conf-1 is also functionally relevant. Collectively, these data support an important role for both of the proposed alternative conformations in the accumulation of sat-Cym RNA.

DISCUSSION

General structural features of sat-Cym RNA

Combined biochemical and biological analyses of the sat-Cym RNA allowed for the generation of a functionally-relevant higher-order RNA structural model for this molecule. At the secondary structure level, the RNA is composed of a series of consecutive helical structures, some of which contain limited (e.g. S1, S17 and S20) or more extensive (e.g. S4) branched components (Figure 2). Many of the sequential stem loops are located immediately adjacent to each other (e.g. S1/S4, S14/S20 and S20/S21) or are separated by a single nucleotide (e.g. S6/S7, S10/S11, S12/S13 and S16/S17), thus some of these helical ends could coaxially stack to further stabilize and/or condense the structure.

Overall, 58% of the sat-Cym RNA is predicted to be involved in base pairing interactions forming secondary structures. Similarly, the secondary structure model for the
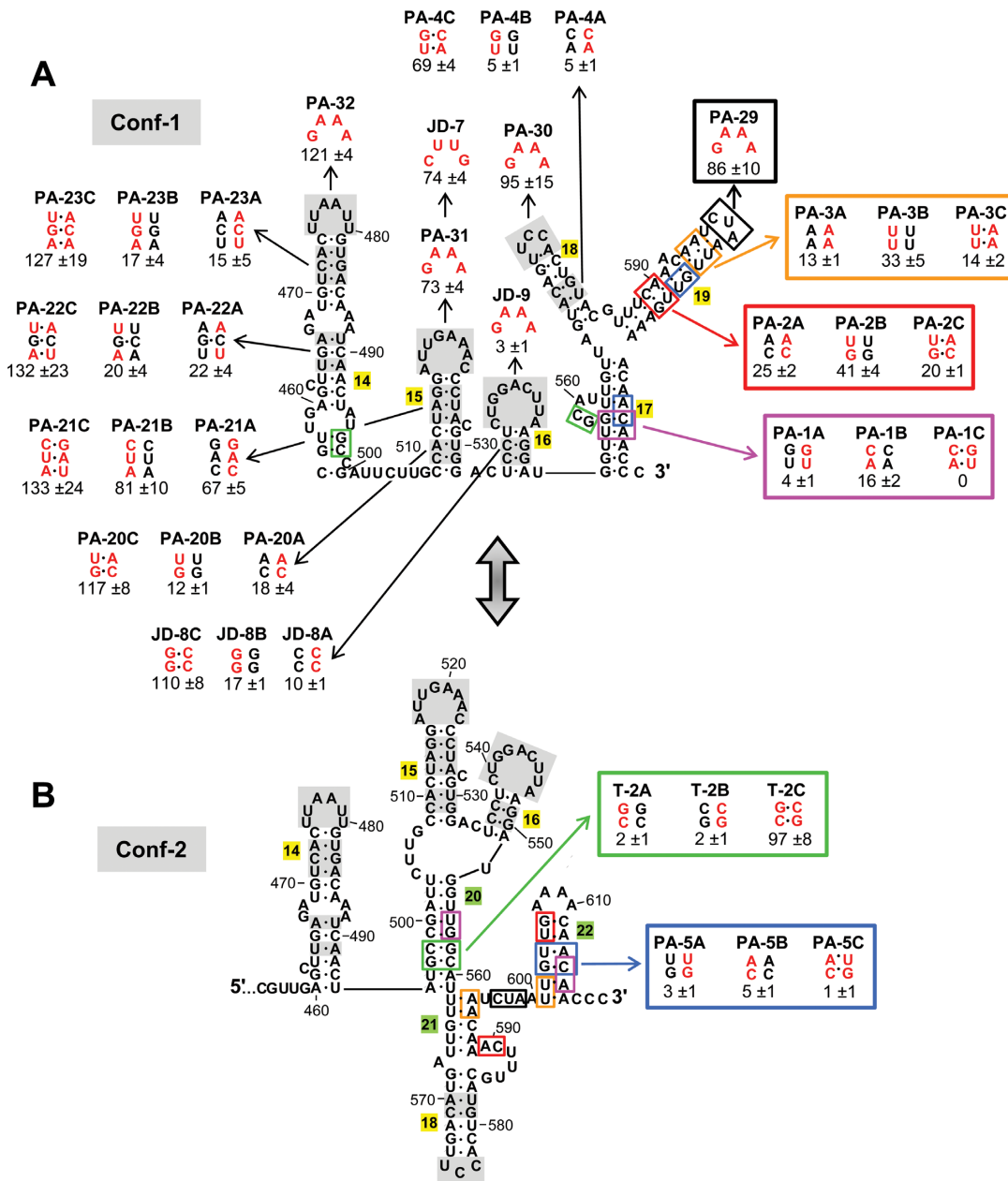
Figure 6. (**A**) Mutational analysis of structures in the 3′ region of sat-Cym. Grey shading shows regions in the structure common to both conformations that were targeted with mutations. Nucleotides in the dynamic parts of the structure that were modified by substitution are identified by black, orange, red and purple boxes. Substitutions in each mutant are indicated and depicted as red nucleotides. The values represent the mean percentages from three independent protoplast infections with standard errors. Green and blue boxes denote nucleotides that were modified in mutants described in (B). (**B**) Nucleotides in the structure that were modified by substitution are delineated by green and blue boxes. The values represent the mean percentages from three independent infections with standard errors. Black, orange, red and purple boxes denote nucleotides that were modified in mutants described in (A).

helper TBSV genome predicts 60% of this RNA to be base paired (52), and comparable values were observed for CMV-associated satRNA (51%), TCV-associated satC (47%) and STMV (62%) (9,24,34). Interestingly, the structural models for all three of the aforementioned satellite species contain a large central or 5′-proximal domain comprising ∼47–60% of their sequence. In contrast, the largest substructure predicted in sat-Cym (i.e. S4 and its intervening sequence; Figure 2A) corresponds to only ∼22% of this RNA, with the remainder of the molecule occupied primarily by closely-

spaced smaller-sized stem loop structures, akin to beads on a string. Thus, in comparison with available linear satellite structural models, the global secondary structure organization of sat-Cym RNA appears to be somewhat distinct.

Intriguingly, disruption of several of the helical segments in sat-Cym RNA did not notably alter its ability to accumulate. Such substructures could be non-functional or play nonspecific and/or passive roles, like contributing to RNA length requirements for efficient RNA replication or particle assembly. Since our protoplast assay monitors primarily
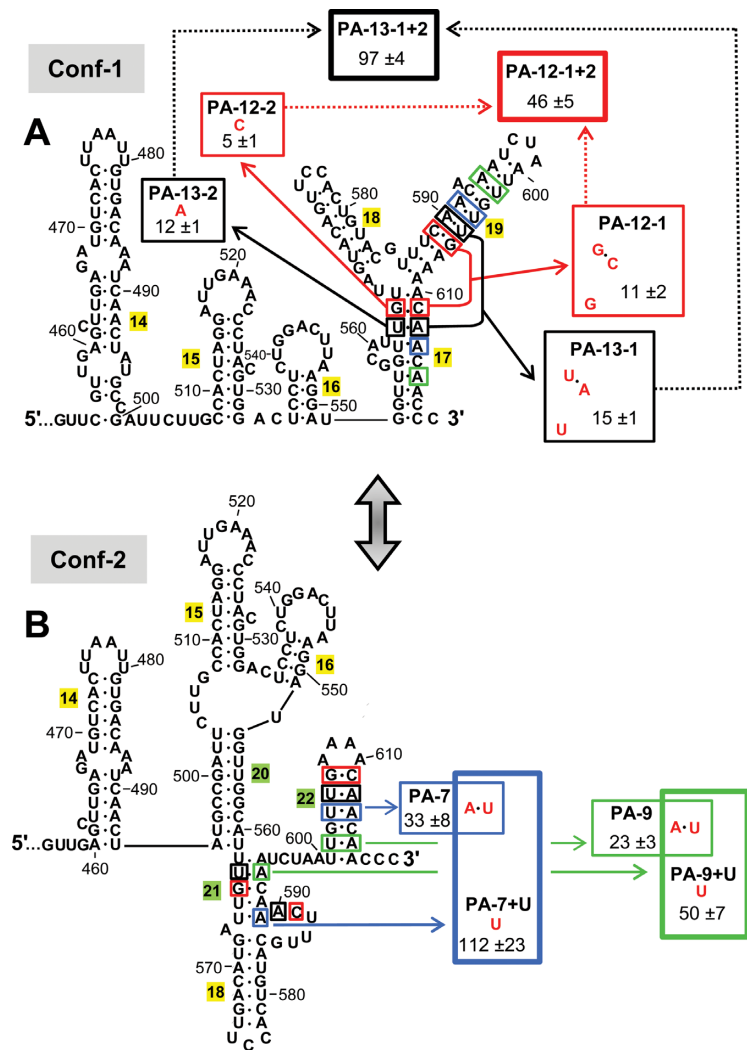
Figure 7. Mutational analysis of conf-1 and conf-2 structures in sat-Cym RNA. (**A**) Nucleotides in conf-1 structure modified by substitution are identified by black and red boxes. The dotted arrows point to mutants in which the substitutions in two different mutants, those from which the arrows originate, were combined. Substitutions in mutants are depicted as red nucleotides. The values represent the mean percentages from three independent infections with standard errors. Blue and green boxes correspond to the nucleotide positions that were modified in mutants described in (**B**). (**B**) Nucleotides in conf-2 structure modified by substitution are identified by blue and green boxes. PA-7 contains only the UA to AU substitution, whereas PA-7+U contains both the former substitution and an additional A to U substitution. The same applies, respectively, for PA-9 and PA-9+U. Black and red boxes correspond to the nucleotide positions that were modified in mutants described in (A).

RNA replication and/or *in vivo* stability, the absence of a discernible phenotype in some mutants could also be due to a structure contributing to a process not monitored by this assay, such as particle assembly or movement within infected plants. Additionally, some undetected defects may be demonstrable only if assessed under competitive conditions, such as head-to-head competition with a wt satRNA. Accordingly, future comprehensive functional analyses will be required to precisely assign roles to the identified RNA secondary structures.

Much less information is known about tertiary interactions in satRNAs. One exception is satC of TCV, where three local pseudoknots in its 3′-proximal region facilitate its replication (59). In sat-Cym, only one functional tertiary interaction, PK-sTD1, which also occurs in tombusvirus helper genomes, was known prior to this study (48). Herein,

we discovered a novel satRNA-specific tertiary interaction, PK11/16, that spans a distance of 216 nt. Long-range interactions are prevalent in helper tombusvirus genomes, where they participate in a variety of viral processes such as translation, genome replication and transcription, and operate via different mechanisms, including generating bipartite RNA structures, repositioning bound proteins, and acting as regulatory RNA switches (52). Similar RNA-based mechanisms can be envisaged for the far-reaching interactions in sat-Cym RNA, which would also act to compress and/or rearrange the global architecture of this satRNA. Below, the biological significance of the structural features in the 5′, central, and 3′ regions of this molecule are discussed.

5′ region of sat-Cym RNA

The substructures in the 5′ region of sat-Cym that emulate those in tombusvirus 5′ UTRs (i.e. TSD, S4 and DSD) have been analyzed previously (48) and are required for RNA replication (49,50). Similarly, the 109 nt long segment intervening S4 in sat-Cym has also been investigated, and substitutions and deletions introduced into this region were found to reduce satRNA levels ranging from 2 to 38% that of wt (48). In the helper virus, the corresponding intervening sequence is only 3–5 nts long and is part of the 5′ UTR of the viral genome (50). Consequently, the insertion of additional sequences in this context could interfere with the 5′ UTR's role in cap-independent translation (60,61). In contrast, the satRNA is non-coding and not subject to such constraints, thus it would be able to maintain and utilize the additional RNA appendage.

Central region of sat-Cym RNA

Of the four stem loop structures present in this region, SL10 was of least functional significance (Figure 4). A higher level of relevance was observed for the small SL12 located between the larger SL11 and SL13. Since formation of the stem in SL12 was moderately important, but its loop sequence identity was not, this hairpin could potentially contribute to optimal relative positioning of the two flanking SLs and/or the stabilization of SL13 via coaxial stacking. The lower stem region of SL11 was less relevant, whereas the upper region was predictably significant, because its terminal loop participates in the formation of PK-11/16 with the 3′ region. SL13 was interesting, as its helical parts were moderately important, while its terminal loop proved to contribute significantly to sat-Cym viability; *i.e.* loop replacement reduced accumulation to 27% that of wt (Figure 4). A search for potential complementary partners for the SL13 loop sequence (410-AUUAGAACA-418) uncovered one in the terminal loop of SL14 (475-UUAAUU-480) and one in the terminal loop of SL19 (595-AUCUAAU-601) (Figure 2A). However, substitution of either of the latter two loop sequences were well tolerated by sat-Cym (i.e. 121% and 86%, respectively; Figure 6A), suggesting that the SL13 loop sequence may have alternative unidentified partner sequences or interacts with a protein factor.

3′ region of sat-Cym RNA

Our structure-function analyses suggest that the 3′ region represents a central hub for RNA-mediated regulation, because it contains a localized RNA switch and an associated long-range interaction. The long-range interaction involves SL11 in the central region pairing with SL16 within the RNA switch. SL16 is predicted to form in both conf-1 and conf-2 (Figure 5A); however, it is possible that these two contexts differentially affect its ability to interact with SL11. Alternatively, if interaction does occur in either conformation, the effect may be different. Similar mechanistic scenarios have been proposed for long-range RNA interactions in Hepatitis C virus that control different steps in virus reproduction (62–65).

The sat-Cym RNA switch comprises a local conformational rearrangement of the 3′-terminal segment. This region is relevant to satRNA replication because the viral polymerase initiates minus-strand synthesis at this location. In tombusvirus genomes, two alternative conformations, closed and open, are required for viral genome replication (Figure 8A). The closed conformation is formed by the 3′-terminal sequence pairing to an internal loop, which renders the 3′ end unavailable as a template for the polymerase (66). Consequently, this structure down-regulates genome replication and concurrently mediates the assembly of the viral RNA replicase complex while protecting the end from 3′-to-5′ exoribonucleases (66–68). In the open conformation, the aforementioned interaction does not occur, and this makes available a small RNA hairpin (SL1) and three unpaired 3′-terminal cytidylates (Figure 8A). The hairpin and short tail constitute the core promoter for minus-strand synthesis that is used by the viral polymerase (69); accordingly, this conformation mediates the initiation of viral RNA replication (66).

The 3′-portion of the viral genome also harbours a small RNA switch comprised of two small, mutually-exclusive RNA hairpins, SL2 and SL-T, that coordinates translational readthrough and genome replication (70) (Figure 8A, right panel). This RNA switch is not present in non-coding tombusvirus satRNAs, however, one of these satRNAs, sat-B10, does contain a genome-like 3′-terminal region capable of forming structurally-comparable open and closed conformations (Figure 8B) (71). Thus, for this satRNA, the maintenance of required functions is accomplished through direct higher-order structural mimicry. A similar approach is also used by SatC, which has a 3′ terminus that is structurally equivalent to that of its helper virus TCV and can fold into comparable alternative conformations (72).

In contrast, direct structural mimicry is not utilized by sat-Cym (Figure 8C). Nonetheless, its conf-1 and conf-2, respectively, do show structural equivalence to some of the key features of the closed and open conformations in tombusvirus genomes. For instance, in conf-1, two of the three 3′-terminal cytidylates are paired, and a possible non-canonical pairing of the ultimate C residue with a proximal U could allow for coaxial stacking of S17 with SL16 (Figure 8C, left panel, dotted arrow). This configuration would embed and protect the 3′-terminus in a quasi-continuous helix and inhibit polymerase access, comparable to the closed conformation in the viral genome (Figure 8A, left panel). Conversely, conf-2 in sat-Cym possesses a 3′-terminal hairpin and tail that are similar to those in the core promoter of the viral open conformation, thus resembling a replication-competent configuration (Figure 8C and A, compare right panels).

Notably, while the 3′-terminal hairpins in the open conformation of both TBSV and sat-B10 contain ultra-stable loops (i.e. aGAAAu and cUUCGg, respectively) (73), the one in sat-Cym is not predicted to possess enhanced stability (i.e. gAAAAc). These distinct structural features are in agreement with the invariant nature of the upper portion of these hairpins in the open and closed conformations in TBSV and sat-B10 and the conformational flexibility required by SL22 in sat-Cym to transition from conf-2 into conf-1 (Figure 8). Indeed, for the latter event, binding of the loop sequence in SL22 to the complementary bulge sequence between S21 and SL18 represents an attractive
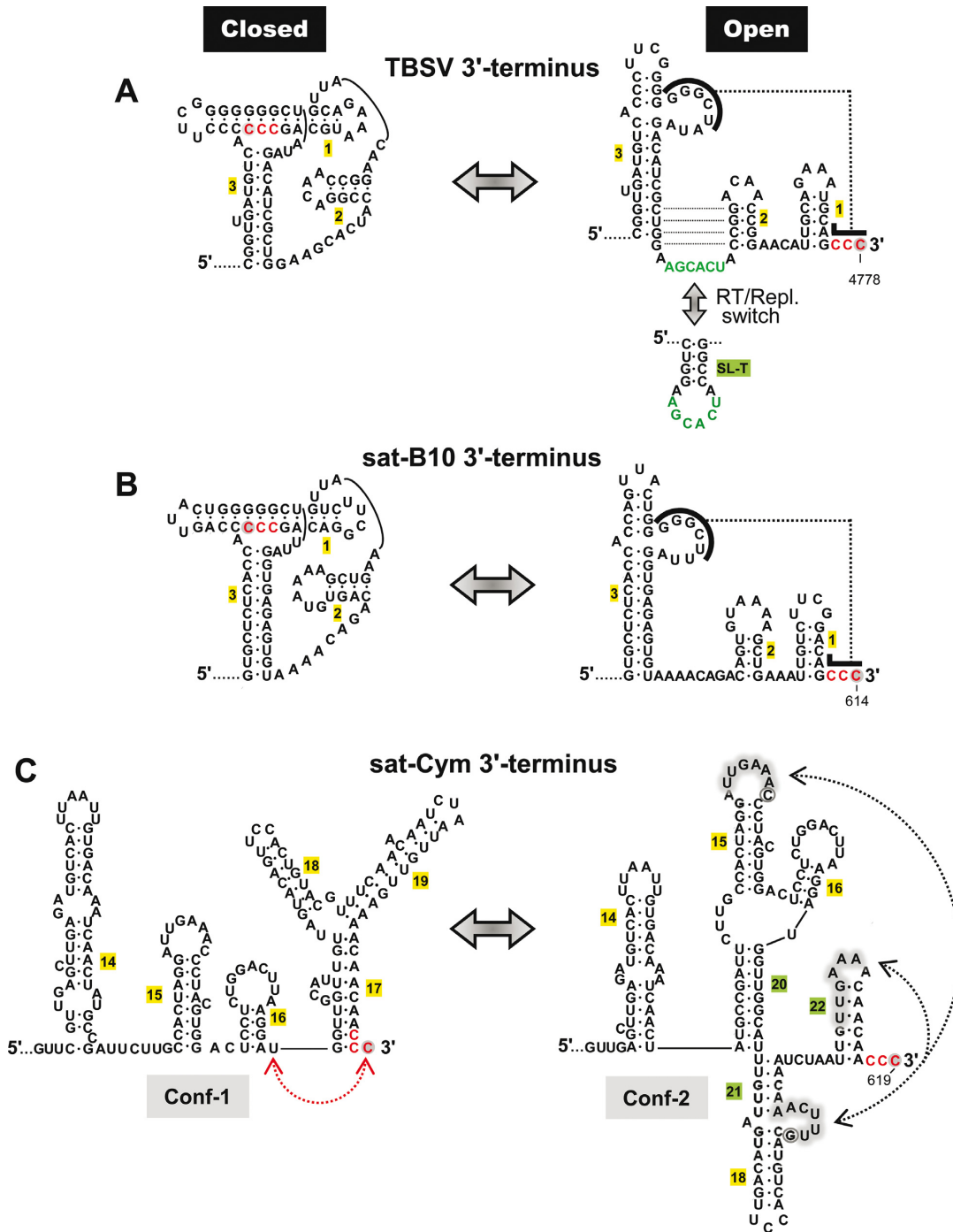
Figure 8. Structures of 3′-terminal regions of the TBSV genome and tombusvirus satRNAs. (**A**) Structure of the 3′-terminal region of the TBSV genome that forms closed and open conformations. The sequences that form a tertiary interaction in the closed conformation are identified in the open conformation by black bars connected by a dotted line. The three 3′-terminal cytidylates are depicted in red and the ultimate residue is shaded grey. A small RNA switch, SL2↔SL-T, that activates RNA replication or translational readthough of p92, is shown. Nucleotides in green participate in a long-range interaction with sequences proximal to the readthough site. (**B**) Structure of the 3′-terminal region of sat-B10 that forms closed and open conformations similar to the TBSV genome. (**C**) Structure of the 3′-terminal region of sat-Cym in which conf-1 and conf-2, respectively, mimic features of closed and open conformations of the TBSV genome. The red dotted line depicts a possible UC base pair, the formation of which could allow for coaxial stacking of stems 16 and 17. The black dotted lines connect grey shaded residues that could interact and regulate the transition of conf-2 to conf-1. The circled nucleotides could form an additional base pair between SL15 and the bulge sequence between S21 and SL18.

prospect for the nucleation step of the transition (Figure 8C, right panel). Moreover, this initial step could potentially be modulated by competition from the loop residues in SL15 (Figure 8C, right panel), the substitution of which reduced sat-Cym accumulation to ∼74% that of wt (Figure 6A, mutants JD-7 and PA-31). The relevance of these and other possible determinants of conformational dynamics will be sought in future detailed studies.

It is clear that sat-Cym has evolved a very different structural solution to control accessibility of its 3′-terminal sequence. This alternate approach, however, when compared with that in the genome, comes with a comparatively large energy barrier between the conformations. That is, conversion for the viral conformations would involve un/pairing of only a few residues (Figure 8A), whereas either transition in the satRNA would require unpairing of significantly longer segments (Figure 8C). Thus, as an RNA switch, the question remains as to how this higher transitional energy barrier is traversed in sat-Cym. Likely possibilities include viral or host proteins, with tombusvirus auxiliary replication protein p33 representing a plausible candidate, since it has been shown to possess RNA chaperone activity (74).

In conclusion, our extensive structural examination of sat-Cym RNA has revealed functionally-relevant secondary and tertiary structures, as well as an RNA switch. This comprehensive catalogue of biologically-important higher-order structures provides a foundation for future investigations into the mechanistic details of RNA-based activities.

REFERENCES

1. Simon,A.E., Roossinck,M.J. and Havelda,Z. (2004) Plant virus satellite and defective interfering RNAs: new paradigms for a new century. *Annu. Rev. Phytopathol.*, **42**, 415–437.
2. Hu,C.C., Hsu,Y.H. and Lin,N.S. (2009) Satellite RNAs and Satellite Viruses of Plants. *Viruses*, **1**, 1325–1350.
3. Rao,A.L. and Kalantidis,K. (2015) Virus-associated small satellite RNAs and viroids display similarities in their replication strategies. *Virology*, **479–480C**, 627–636.
4. Huang,Y.W., Hu,C.C., Lin,N.S. and Hsu,Y.H. (2010) Mimicry of molecular pretenders: the terminal structures of satellites associated with plant RNA viruses. *RNA Biol.*, **7**, 162–171.
5. Ferré-D'Amaré,A.R. and Scott,W.G. (2010) Small self-cleaving ribozymes. *Cold Spring Harb. Perspect. Biol.,* **2**, a003574.
6. Song,S.I. and Miller,W.A. (2004) Cis and trans requirements for rolling circle replication of a satellite RNA. *J Virol.*, **78**, 3072–3082.
7. Rao,A.L. (2006) Genome packaging by spherical plant RNA viruses. *Annu. Rev. Phytopathol.*, **44**, 61–87.
8. Bunka,D.H., Lane,S.W., Lane,C.L., Dykeman,E.C., Ford,R.J., Barker,A.M., Twarock,R., Phillips,S.E. and Stockley,P.G. (2011) Degenerate RNA packaging signals in the genome of Satellite Tobacco Necrosis Virus: implications for the assembly of a T = 1 capsid. *J. Mol. Biol.*, **413**, 51–65.
9. Archer,E.J., Simpson,M.A., Watts,N.J., O'Kane,R., Wang,B., Erie,D.A., McPherson,A. and Weeks,K.M. (2013) Long-range architecture in a viral RNA genome. *Biochemistry*, **52**, 3182–3190.
10. Athavale,S.S., Gossett,J.J., Bowman,J.C., Hud,N.V., Williams,L.D. and Harvey,S.C. (2013) In vitro secondary structure of the genomic RNA of satellite tobacco mosaic virus. *PLoS One*, **8**, e54384.
11. Schroeder,S.J., Stone,J.W., Bleckley,S., Gibbons,T. and Mathews,D.M. (2011) Ensemble of secondary structures for encapsidated satellite tobacco mosaic virus RNA consistent with chemical probing and crystallography constraints. *Biophys. J.*, **101**, 167–175.
12. Simon,A.E. and Howell,S.H. (1986) The virulent satellite RNA of turnip crinkle virus has a major domain homologous to the 3′ end of the helper virus genome. *EMBO J.*, **5**, 3423–3428.
13. Song,C. and Simon,A.E. (1995) Requirement of a 3′-terminal stem-loop in in vitro transcription by an RNA-dependent RNA polymerase. *J. Mol. Biol.*, **254**, 6–14.
14. Stupina,V. and Simon,A.E. (1997) Analysis in vivo of turnip crinkle virus satellite RNA C variants with mutations in the 3′-terminal minus-strand promoter. *Virology*, **238**, 470–477.
15. Carpenter,C.D. and Simon,A.E. (1998) Analysis of sequences and predicted structures required for viral satellite RNA accumulation by in vivo genetic selection. *Nucleic Acids Res.*, **26**, 2426–2432.
16. Nagy,P.D., Pogany,J. and Simon,A.E. (2001) In vivo and in vitro characterization of an RNA replication enhancer in a satellite RNA associated with turnip crinkle virus. *Virology*, **288**, 315–324.
17. Zhang,G. and Simon,A.E. (2003) A multifunctional turnip crinkle virus replication enhancer revealed by in vivo functional SELEX. *J. Mol. Biol.*, **326**, 35–48.
18. Zhang,J., Stuntz,R.M. and Simon,A.E. (2004) Analysis of a viral replication repressor: sequence requirements for a large symmetrical internal loop. *Virology*, **326**, 90–102.
19. Zhang,G., Zhang,J., George,A.T., Baumstark,T. and Simon,A.E. (2006) Conformational changes involved in initiation of minus-strand synthesis of a virus-associated RNA. *RNA*, **12**, 147–162.
20. Zhang,J., Zhang,G., Guo,R., Shapiro,B.A. and Simon,A.E. (2006) A pseudoknot in a preactive form of a viral RNA is part of a structural switch activating minus-strand synthesis. *J. Virol.*, **80**, 9181–9191.
21. Guan,H., Carpenter,C.D. and Simon,A.E. (2000) Analysis of cis-acting sequences involved in plus-strand synthesis of a turnip crinkle virus-associated satellite RNA identifies a new carmovirus replication element. *Virology*, **268**, 345–354.
22. Guan,H., Carpenter,C.D. and Simon,A.E. (2000) Requirement of a 5′-proximal linear sequence on minus strands for plus-strand synthesis of a satellite RNA associated with turnip crinkle virus. *Virology*, **268**, 355–363.
23. Guan,H., Song,C. and Simon,A.E. (1997) RNA promoters located on (-)-strands of a subviral RNA associated with turnip crinkle virus. *RNA*, **3**, 1401–1412.
24. Murawski,A.M., Nieves,J.L., Chattopadhyay,M., Young,M.Y., Szarko,C., Tajalli,H.F., Azad,T., Jean-Jacques,N.B., Simon,A.E. and Kushner,D.B. (2015) Rapid evolution of in vivo-selected sequences and structures replacing 20% of a subviral RNA. *Virology*, **483**, 149–162.
25. Huang,Y.W., Hu,C.C., Lin,C.A., Liu,Y.P., Tsai,C.H., Lin,N.S. and Hsu,Y.H. (2009) Structural and functional analyses of the 3′ untranslated region of Bamboo mosaic virus satellite RNA. *Virology*, **386**, 139–153.
26. Chen,H.C., Kong,L.R., Yeh,T.Y., Cheng,C.P., Hsu,Y.H. and Lin,N.S. (2012) The conserved 5′ apical hairpin stem loops of bamboo mosaic virus and its satellite RNA contribute to replication competence. *Nucleic Acids Res.*, **40**, 4641–4652.
27. Chen,S.C., Desprez,A. and Olsthoorn,R.C. (2010) Structural homology between bamboo mosaic virus and its satellite RNAs in the 5′ untranslated region. *J. Gen. Virol.*, **91**, 782–787.
28. Chen,H.C., Hsu,Y.H. and Lin,N.S. (2007) Downregulation of Bamboo mosaic virus replication requires the 5′ apical hairpin stem loop structure and sequence of satellite RNA. *Virology*, **365**, 271–284.
29. Hsu,Y.H., Chen,H.C., Cheng,J., Annamalai,P., Lin,B.Y., Wu,C.T., Yeh,W.B. and Lin,N.S. (2006) Crucial role of the 5′ conserved structure of bamboo mosaic virus satellite RNA in downregulation of helper viral RNA replication. *J. Virol.*, **80**, 2566–2574.
30. Yeh,W.B., Hsu,Y.H., Chen,H.C. and Lin,N.S. (2004) A conserved secondary structure in the hypervariable region at the 5′ end of

Bamboo mosaic virus satellite RNA is functionally interchangeable. *Virology*, **330**, 105–115.

31. Annamalai,P., Hsu,Y.H., Liu,Y.P., Tsai,C.H. and Lin,N.S. (2003) Structural and mutational analyses of cis-acting sequences in the 5′-untranslated region of satellite RNA of bamboo mosaic potexvirus. *Virology*, **311**, 229–239.

32. Kuznetsov,Y.G., Dowell,J.J., Gavira,J.A., Ng,J.D. and McPherson,A. (2010) Biophysical and atomic force microscopy characterization of the RNA from satellite tobacco mosaic virus. *Nucleic Acids Res.*, **38**, 8284–8294.

33. Garmann,R.F., Gopal,A., Athavale,S.S., Knobler,C.M., Gelbart,W.M. and Harvey,S.C. (2015) Visualizing the global secondary structure of a viral RNA genome with cryo-electron microscopy. *RNA*, **21**, 877–886.

34. Bernal,J.J. and García-Arenal,F. (1997) Analysis of the in vitro secondary structure of cucumber mosaic virus satellite RNA. *RNA*, **3**, 1052–1067.

35. Nicholson,B.L. and White,K.A. (2014) Functional long-range RNA-RNA interactions in positive-strand RNA viruses. *Nat. Rev. Microbiol.*, **12**, 493–504.

36. Nicholson,B.L. and White,K.A. (2015) Exploring the architecture of viral RNA genomes. *Curr. Opin. Virol.*, **12**, 66–74.

37. Newburn,L.R. and White,K.A. (2015) Cis-acting RNA elements in positive-strand RNA plant virus genomes. *Virology*, **479–480C**, 434–443.

38. Nagy,P.D., Barajas,D. and Pogany,J. (2012) Host factors with regulatory roles in tombusvirus replication. *Curr. Opin. Virol.*, **2**, 691–698.

39. Nagy,P.D. and Pogany,J. (2008) Multiple roles of viral replication proteins in plant RNA virus replication. *Methods Mol. Biol.*, **451**, 55–68.

40. White,K.A. and Nagy,P.D. (2004) Advances in the molecular biology of tombusviruses: gene expression, genome replication, and recombination. *Prog. Nucleic Acids Res. Mol. Biol.*, **78**, 187–226.

41. Rubino,L., Burgyan,J., Grieco,F. and Russo,M. (1990) Sequence analysis of cymbidium ringspot virus satellite and defective interfering RNAs. *J. Gen. Virol.*, **71**, 1655–1660.

42. Celix,A., Rodriguez-Cerezo,E. and Garcia-Arenal,F. (1997) New satellite RNAs, but no DI RNAs, are found in natural populations of tomato bushy stunt tombusvirus. *Virology*, **239**, 277–284.

43. Rubino,L. and Russo,M. (2010) Properties of a novel satellite RNA associated with tomato bushy stunt virus infections. *J. Gen. Virol.*, **91**, 2393–2401.

44. Dalmay,T. and Rubino,L. (1995) Replication of cymbidium ringspot virus satellite RNA mutants. *Virology*, **206**, 1092–1098.

45. Burgyán,J., Dalmay,T., Rubino,L. and Russo,M. (1992) The replication of cymbidium ringspot tombusvirus defective interfering-satellite RNA hybrid molecules. *Virology*, **190**, 579–586.

46. Pantaleo,V. and Burgyán,J. (2008) Cymbidium ringspot virus harnesses RNA silencing to control the accumulation of virus parasite satellite RNA. *J. Virol.*, **82**, 11851–11858.

47. Rubino,L. and Russo,M. (2012) A single amino acid substitution in the ORF1 of cymbidium ringspot virus determines the accumulation of two satellite RNAs. *Virus Res.*, **168**, 84–87.

48. Chernysheva,O.A. and White,K.A. (2005) Modular arrangement of viral cis-acting RNA domains in a tombusvirus satellite RNA. *Virology*, **332**, 640–649.

49. Wu,B., Vanti,W.B. and White,K.A. (2001) An RNA domain within the 5′ untranslated region of the tomato bushy stunt virus genome modulates viral RNA replication. *J. Mol. Biol.*, **305**, 741–756.

50. Ray,D., Wu,B. and White,K.A. (2003) A second functional RNA domain in the 5′ UTR of the Tomato bushy stunt virus genome: intra- and interdomain interactions mediate viral RNA replication. *RNA*, **9**, 1232–1245.

51. Rubino,L., Pantaleo,V., Navarro,B. and Russo,M. (2004) Expression of tombusvirus open reading frames 1 and 2 is sufficient for the replication of defective interfering, but not satellite, RNA. *J. Gen. Virol.*, **85**, 3115–3122.

52. Wu,B., Grigull,J., Ore,M.O., Morin,S. and White,K.A. (2013) Global organization of a positive-strand RNA virus genome. *PLoS Pathog.*, **9**, e1003363.

53. Vasa,S.M., Guex,N., Wilkinson,K.A., Weeks,K.M. and Giddings,M.C. (2008) ShapeFinder: a software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA*, **14**, 1979–1990.

54. Low,J.T. and Weeks,K.M. (2010) SHAPE-directed RNA secondary structure prediction. *Methods*, **52**, 150–158.

55. Mathews,D.H., Disney,M.D., Childs,J.L., Schroeder,S.J., Zuker,M. and Turner,D.H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 7287–7292.

56. Hajdin,C.E., Bellaousov,S., Huggins,W., Leonard,C.W., Mathews,D.H. and Weeks,K.M. (2013) Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 5498–5503.

57. De Rijk,P., Wuyts,J. and De Wachter,R. (2003) RnaViz2: an improved representation of RNA secondary structure. *Bioinformatics*, **19**, 299–300.

58. Merino,E.J., Wilkinson,K.A., Coughlan,J.L. and Weeks,K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.*, **127**, 4223–4231.

59. Simon,A.E. (2015) 3′ UTRs of carmoviruses. *Virus Res.*, **206**, 27–36.

60. Fabian,M.R. and White,K.A. (2006) Analysis of a 3′-translation enhancer in a tombusvirus: a dynamic model for RNA-RNA interactions of mRNA termini. *RNA*, **12**, 1304–1314.

61. Nicholson,B.L., Wu,B., Chevtchenko,I. and White,K.A. (2010) Tombusvirus recruitment of host translational machinery via the 3′ UTR. *RNA*, **16**, 1402–1419.

62. Tuplin,A., Struthers,M., Simmonds,P. and Evans,D.J. (2012) A twist in the tail: SHAPE mapping of long-range interactions and structural rearrangements of RNA elements involved in HCV replication. *Nucleic Acids Res.*, **40**, 6908–6921.

63. Shetty,S., Stefanovic,S. and Mihailescu,M.R. (2013) Hepatitis C virus RNA: molecular switches mediated by long-range RNA-RNA interactions? *Nucleic Acids Res.*, **41**, 2526–2540.

64. Romero-López,C., Barroso-Deljesus,A., García-Sacristán,A., Briones,C. and Berzal-Herranz,A. (2014) End-to-end crosstalk within the hepatitis C virus genome mediates the conformational switch of the 3′ X-tail region. *Nucleic Acids Res.*, **42**, 567–582.

65. Fricke,M., Dünnes,N., Zayas,M., Bartenschlager,R., Niepmann,M. and Marz,M. (2015) Conserved RNA secondary structures and long-range interactions in hepatitis C viruses. *RNA*, **21**, 1219–1232.

66. Pogany,J., Fabian,M.R., White,K.A. and Nagy,P.D. (2003) A replication silencer element in a plus-strand RNA virus. *EMBO J.*, **22**, 5602–5611.

67. Wu,B., Pogany,J., Na,H., Nicholson,B.L., Nagy,P.D. and White,K.A. (2009) A discontinuous RNA platform mediates RNA virus replication: building an integrated model for RNA-based regulation of viral processes. *PLoS Pathog.*, **5**, e1000323.

68. Na,H., Fabian,M.R. and White,K.A. (2006) Conformational organization of the 3′ untranslated region in the tomato bushy stunt virus genome. *RNA*, **12**, 2199–2210.

69. Panavas,T., Pogany,J. and Nagy,P.D. (2002) Analysis of minimal promoter sequences for plus-strand synthesis by the Cucumber necrosis virus RNA-dependent RNA polymerase. *Virology*, **296**, 263–274.

70. Cimino,P.A., Nicholson,B.L., Wu,B., Xu,W. and White,K.A. (2011) Multifaceted regulation of translational readthrough by RNA replication elements in a tombusvirus. *PLoS Pathog.*, **7**, e1002423.

71. Fabian,M.R., Na,H., Ray,D. and White,K.A. (2003) 3′-Terminal RNA secondary structures are important for accumulation of tomato bushy stunt virus DI RNAs. *Virology*, **313**, 567–580.

72. Zhang,G., Zhang,J. and Simon,A.E. (2004) Repression and derepression of minus-strand synthesis in a plus-strand RNA virus replicon. *J. Virol.*, **78**, 7619–7633.

73. Hall,K.B. (2015) Mighty tiny. *RNA*, **21**, 630–631.

74. Stork,J., Kovalev,N., Sasvari,Z. and Nagy,P.D. (2011) RNA chaperone activity of the tombusviral p33 replication protein facilitates initiation of RNA synthesis by the viral RdRp in vitro. *Virology*, **409**, 338–347.