Evolution of virus-derived sequences for high-level replication of a subviral RNA

Jiuchun Zhang, Guohua Zhang, John C. McCormack, Anne E. Simon *

Department of Cell Biology and Molecular Genetics, University of Maryland, 1109 Microbiology Building, College Park, MD 20742, USA

Received 23 February 2006; returned to author for revision 7 March 2006; accepted 9 March 2006

Available online 6 May 2006

Abstract

Turnip crinkle virus (TCV) and its 356-nt satellite RNA satC share 151 nt of 3′-terminal sequence, which contain 8 positional differences and are predicted to fold into virtually identical structures, including a series of four phylogenetically inferred hairpins. SatC and TCV containing reciprocal exchanges of this region accumulate to only 15% or 1% of wild-type levels, respectively. Step-wise conversion of satC and TCV 3′-terminal sequences into the counterpart’s sequence revealed the importance of having the cognate core promoter (Pr), which is composed of a single hairpin that differs in both sequence and stability, and an adjacent short 3′-terminal segment. The negative impact of the more stable TCV Pr on satC could not be attributed to lack of formation of a known tertiary interaction involving the 3′-terminal bases, nor an effect of coat protein, which binds specifically to TCV-like Pr and not the satC Pr. The satC Pr was a substantially better promoter than the TCV Pr when assayed in vitro using purified recombinant TCV RdRp, either in the context of satC or when assayed downstream of non-TCV-related sequence. Poor activity of the TCV Pr in vitro occurred despite solution structure probing indicating that its conformation in the context of satC is similar to the active form of the satC Pr, which is thought to form following a required conformational switch. These results suggest that evolution of satC following its initial formation generated a Pr that can function more efficiently in the absence of additional TCV sequence that may be required for full functionality of the TCV Pr.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Turnip crinkle virus; RNA conformational switch; RNA virus evolution; Satellite RNA; RNA replication promoters

Introduction

Replication of genomic and subviral RNAs requires specific interactions between the replicase complex and RNA cis-acting elements. While core promoters located near the 3′ ends of plus (+)- and minus (−)-strands can independently recruit replication complexes resulting in low levels of de novo synthesized complementary strands (Dreher, 1999), additional elements throughout viral genomes aid in enzyme complex assembly or enhance or repress transcription (e.g., French and Ahlquist, 1987; Frolov et al., 2001; Herold and Andino, 2001; Khromykh et al., 2001; Klovins and van Duin, 1999; McCormack and Simon, 2004; Monkewich et al., 2005; Nagashima et al., 2005; Nagy et al., 1999; Panavas and Nagy, 2003; Panaviene et al., 2005; Ray and White, 2003; Vlot and Bol, 2003; Zhang and Simon, 2003). In addition, recent evidence suggests that RNA conformational rearrangements play key roles in coordinating translation and replication, regulating subgenomic RNA synthesis or producing asymmetric levels of (+)- and (−)-strands by masking or exposing elements required for a particular process (Barry and Miller, 2002; Isken et al., 2004; Khromykh et al., 2001; Olsthoorn et al., 1999; Koev et al., 2002; Na and White, 2006; Pogany et al., 2003; van den Born et al., 2005; Zhang et al., 2004a, 2006). For example, conformational changes at the 3′ ends of Barley yellow dwarf virus (Koev et al., 2002) and Tomato bushy stunt virus (TBSV; Na and White, 2006; Pogany et al., 2003) may control accessibility of the RNA-dependent RNA polymerase (RdRp) to the initiation site for (−)-strand synthesis.

The complexities inherent in RNA virus replication has led to the use of untranslated subviral RNAs such as defective interfering (DI) RNAs or satellite (sat) RNAs, as models for...
their larger, multifaceted helper viral genomic RNAs. Many (+)-strand RNA viruses are naturally associated with subviral RNAs, which depend on their helper virus for replication and host trafficking components (David et al., 1992; White and Morris, 1999; Simons et al., 2004). While DI RNAs are mainly derived from 5’ and 3’ portions of viral genomic RNAs, most satRNAs share little consecutive sequence similarity with their helper virus genomes and may have arisen from an unrelated RNA or a series of recombination events joining short segments of viral and non-viral RNAs that further evolved into a functional molecule (Carpenter and Simon, 1996). An unusual satRNA, satC (356 bases), is associated with Turnip crinkle virus (TCV; single (+)-strand RNA of 4054 bases), a member of the family Tombusviridae, genus Carmovirus (Simon, 2001).

SatC has features of both DI and satRNAs with its 5’ 190 bases originating from nearly full-length TCV satRNA satD (194 bases) and its 3’ 166 bases derived from two regions at the 3’ end of TCV genomic RNA (Fig. 1A; Simon and Howell, 1986). TCV is also naturally associated with DI RNAs, such as diG, whose sequence is mainly derived from 5’ and 3’ regions of the genomic RNA (Fig. 1A; Li et al., 1989).

The 3’-terminal 100 bases shared by TCV and satC differ at only eight positions (“positions” refers to particular locations where one or more consecutive bases may differ) and are predicted to be structurally similar by mFold (Fig. 1B; Zhang et al., 2004b; Zuker, 2003). This observation suggested that satC would be a good model for determining the function of TCV cis-acting sequences within this region in the replication process. A combination of in vivo studies using Arabidopsis thaliana protoplasts, in vitro assays for transcription initiation using purified recombinant TCV RdRp, and in vitro RNA solution structure probing revealed that this region in satC assumes two very different RNA conformations: an unresolved preactive conformation stabilized by extensive tertiary structure that...
includes pseudoknot 2 (Ψ2) (Zhang et al., 2006, submitted for publication); and an active conformation that includes pseudoknot 1 (Ψ1) and four hairpins found in similar locations in TCV and the related carmoviruses, Cardamine chlorotic fleck virus (CCFV) and Japanese iris necrosis virus (JINRV) (Fig. 1B; Zhang et al., 2006). These four hairpins have been designated as (from 3′ to 5′) (1) Pr, the core promoter of satC (Song and Simon, 1995) and TCV (Sun and Simon, manuscript in revision) for synthesis of (−)-strands; (2) H5, which contains a large symmetrical internal loop (LSL) that interacts with 3′-terminal bases to form Ψ1 in satC (Zhang et al., 2004a) and is proposed to help organize the replication complex in TCV (McCormack and Simon, 2004); (3) H4b, which contains terminal loop sequences that forms Ψ2 with sequence flanking the 3′ base of H5 (Zhang et al., submitted for publication); and (4) H4a, which forms a single unit with H4b in satC and is flanked by a short GC-rich element (DR) that is proposed to help mediate the conformational switch (Zhang et al., 2004a, submitted for publication). All four hairpins are required for satC replication in vivo (Zhang et al., 2004b, submitted for publication).

Despite sequence and structural similarities between satC and TCV, satC with the 3′-terminal 100 bases of TCV (renamed C3′100T) accumulated very poorly in plants and protoplasts (Wang and Simon, 2000). This result suggested that one or more of the 3′ cis-acting elements that are necessary for efficient amplification of TCV (J. C. McCormack and A. E. Simon, unpublished) function poorly when associated with satC sequence or when separated from the remainder of the viral genome. To gain an understanding of how the 3′ region of TCV has evolved in the context of satC to allow for high-level satRNA accumulation, satC and TCV 3′-terminal sequences were converted in a step-wise fashion into the counterpart’s sequence, which revealed the importance of having the cognate Pr. Our results also indicate that the TCV Pr is a much weaker core promoter than the satC Pr in vitro, even though structural analyses suggested that the TCV Pr assumes a form similar to the active form of the satC promoter. These results suggest that the TCV Pr requires additional elements upstream of the region shared between satC and TCV to function optimally in vivo and vitro.

Results

TCV H5 participates in a tertiary interaction that is similar to Ψ1 of satC

We previously demonstrated that the four 3′-terminal bases of satC (353GCCC–OH) likely interact with complementary residues in the LSL of the upstream hairpin H5 (297GGGC) forming Ψ1 (Zhang et al., 2004a). When tested in an in vitro assay programmed with purified recombinant p88 RdRp, compensatory alterations between the satC G:C pair in positions 353 and 300 restored wt levels of transcription to full-length RNA transcripts compared with transcripts containing the individual mutations (Zhang et al., 2004a). Interestingly, both single and double mutants accumulated below the level of detection in protoplasts (Zhang et al., 2004a). A similar interaction between the 3′ end and the 3′ proximal hairpin, SL3, has also been demonstrated experimentally for the Tombusvirus TBSV (Pogany et al., 2003).

While this interaction is phylogenetically conserved in carmoviruses, it had not been experimentally confirmed for carmoviral genomic RNAs. Therefore, to address the possibility that the 3′ 100 bases of TCV cannot efficiently substitute for the analogous region in satC because it does not form Ψ1, a cytidylate to guanylate transversion was engineered at position 3994 in the TCV H5 LSL, generating C3994G, and a guanylate to cytidylate transversion was constructed at position 4051 generating G4051C. In addition, the mutations were combined, generating C3994/G4051, which reestablishes the putative pseudoknot (Fig. 2A). Arabidopsis protoplasts were inoculated with T7 polymerase-synthesized transcripts of the TCV single and double mutants, and levels of (+)-strands were examined at 40 h postinoculation (hpi). C3994G was more severely impacted than G4051C, accumulating to only 7% of wt levels

Fig. 2. Confirmation of Ψ1 in TCV. (A) Putative paired bases are connected by a dotted line. Location of point mutations generated in TCV are shown. The large symmetrical loop (LSL) of H5 is also indicated. (B) RNA gel blot of TCV (+)-strand accumulation in protoplasts at 40 h postinoculation (hpi). The blot was stripped and reprobed with a fragment complementary to rRNAs. Values are averages of three independent assays.
while G4051 accumulated to 29% of wt (Fig. 2B). C3994/G4051 restored TCV accumulation to 50% of wt (Fig. 2B). These results support the existence of Ψ1 in TCV. In addition, the detectable accumulation of the three TCV mutants in protoplasts compared with undetectable accumulation for the analogous satC mutants suggests reduced importance for Ψ1 formation in TCV compared with satC in vivo, and/or that the sequences involved in satC Ψ1 participate in additional activities crucial for satC accumulation in vivo.

**Two positions in the satC/TCV Pr are critical for satC accumulation**

To determine which base differences in the 3′ 100 bases of C3′ 100T were most responsible for reduced satC accumulation, satC constructs were generated with single and multiple positional changes to incorporate TCV-specific bases and transcripts tested for accumulation in protoplasts (Fig. 3). The 3′ 100 bases of TCV and satC differ at 8 positions, which are labeled 1 through 8 in Fig. 3A. Most variations are in the Pr, where two bases differences at position 6 and the single variance at position 8 reduce the 11-bp stem of the TCV Pr to 7 bp for satC. The satC and TCV Pr loops also vary substantially (compare the satC Pr to the TCV Pr in Fig. 1B). The two bases that vary between TCV and satC in H5 (positions 1 and 2) also reduce the stability of the satC H5 stem compared with H5 of TCV. Two base differences (positions 3 and 4) are also located in the linker region (LinkH5-Pr) between H5 and Pr and a single base variation (position 7) is in the 3 nt linker (LinkH4b-H5) between H4b and H5.

SatC construct C12T, with H5 of TCV (i.e., positions 1 and 2 of TCV; most mutants are labeled with the backbone construct satC [C] or TCV [T] followed by the numerical positions that

Fig. 3. Effect on satC accumulation in protoplasts after sequence conversion to residues found in TCV and diG. (A) Left, nomenclature of positional differences between satC and TCV in their 3′ 100 bases. The sequence of satC is given and positional variances in TCV are underlined. Locations of differences are identified by circled numbers. TCV has a 4 nt insert at position 5. Right, the Pr of diG. Nucleotides in triangles denote extra bases in diG compared with satC. Base alterations compared with satC are boxed. (B) Arabidopsis protoplasts were inoculated with TCV genomic RNA transcripts and transcripts of constructs shown above each lane. Total RNA was extracted at 40 hpi and probed with an oligonucleotide complementary to TCV, satC and diG. Ethidium bromide staining of the gel before blotting shows ribosomal RNA (rRNA) loading control. Numbers reflect average accumulation levels for two repeats with the exception of the first four lanes, which reflect a single assay. None, no added satC. wtC, wt satC; wtG, wt diG; C3′100T, satC with the 3′ 104 bases of TCV; G-PrC, diG with the Pr of satC; G-PrT, diG with the Pr of TCV. Numbers in the names of all other constructs reflect which positions were changed to those of TCV in the background of satC.
have been changed to the alternate sequence identified by a subscript T or C), accumulated slightly better than wt satC (115%), as previously reported (Zhang and Simon, 2005). The base change in Link{sub}4b-115 (position 7; C7T) also did not negatively affect accumulation of the satRNA, indicating that positions 1, 2 and 7 do not independently reduce C3′100T accumulation. Constructs containing both the H5 and link{sub}4b-115 variances (C127T), however, accumulated to 80% of wt, suggesting that the slight beneficial effect of incorporating the TCV H5 may be eliminated when transcripts contain all three differences. Alteration of positions 3 and 4 in link{sub}4b-Pr (C34T) reduced satRNA accumulation to 75% of wt, which was further reduced by inclusion of position 7 (C347T; 58% of wt). In contrast, satC containing TCV-specific bases in link{sub}4b-Pr and H5 (C1234T) accumulated to 105% of wt.

We previously determined that satC with TCV Pr positions 5 or 6 accumulated to wt levels in protoplasts (Wang and Simon, 2000). In addition, satC with position 6 of TCV and a single CG dinucleotide at position 5 (this generates the Pr of diG and the construct has been renamed C-PrG) accumulated to 64% of wt levels (Wang and Simon, 2000; see Fig. 4B). To examine the effect of additional combinations of TCV-specific bases in the Pr, satC was altered at position 8 (C8T), positions 5 and 6 (C56T) and at all three locations (C568T). C8T accumulated to 54% of wt, while C56T and C568T accumulated to 31 and 34% of wt, respectively. Accumulation was not enhanced when alterations at positions 1 and 2 in H5 or 3 and 4 in link{sub}4b-Pr were included with positional changes at 5 and 6 (C1256T and C3456T, 34%). These results, combined with our earlier results, suggest (1) satC accumulation is most negatively affected when containing both positions 5 and 6 of TCV; and (2) no further reduction occurs when satC also contains TCV-specific H5 or the TCV link{sub}4b-Pr region.

We previously determined that diG accumulation increases when its Pr included positions 5 or 6 of satC, or both positions of satC, the latter of which generates diG with the Pr of satC (renamed G-PrC; Wang and Simon, 2000). To determine if diG accumulation is affected when containing position 5 of TCV (i.e., an insertion of CG at this location producing diG with the Pr of TCV [G-PrT]), diG, G-PrT and G-PrC were assayed for accumulation in protoplasts. As previously found, diG was a poor template compared with satC, accumulating to only 11% of wt satC (Fig. 3B). Also similar to prior results, accumulation was enhanced nearly 7-fold when diG contained the Pr of satC (G-PrC). In contrast, G-PrT accumulated as poorly as diG. Altogether, these results suggest that (1) the TCV Pr is a limited promoter in the context of satC or diG and (2) base changes in the TCV Pr that occurred after the recombination events that produced satC or diG allow the Pr to function efficiently in the context of satC.

**Coat protein ability to bind TCV-like Pr is not a primary factor in reduced accumulation of satC Pr variants**

We previously reported that TCV CP binds specifically to the TCV-like Pr of diG but not to the Pr of satC (Wang and Simon, 2000). While efficiency of CP-binding to the TCV Pr was not determined, the possibility existed that CP interaction with the TCV-like Pr was responsible for the reduced accumulation of C56T and related constructs. To examine whether CP differentially interferes with the accumulation of satC with either the Pr of TCV or diG, wt satC, C3′100T, C56T and C-PrG, were coinoculated with either wt TCV or CPmT, a TCV variant with mutations at the primary and secondary CP translation initiation codons that eliminate detectable CP (Wang and Simon, 1999, Fig. 4A). As previously shown, CPmT accumulated to lower levels than wt TCV in protoplasts when compared with the levels of control ribosomal RNA (rRNA) (Fig. 4B). Whether this reflects a negative effect of the mutations on replication/translation of the genomic RNA or a requirement for CP to achieve high-level TCV accumulation is not known. Despite the reduction in TCV levels, both wt satC and satC variants accumulated to proportionately higher levels compared with rRNA when inoculated with CPmT (Fig. 4B). Since this effect was not restricted to constructs with Pr of TCV or diG, CP binding to TCV-like Pr is apparently not responsible for the reduced accumulation of satC with Pr elements from TCV.

**Efficient TCV accumulation depends on its cognate Pr**

To determine if TCV accumulation is affected when containing the satC 3′-terminal region, the 3′-terminal 104 bases of TCV were replaced with the analogous region of satC, generating T3′100C. Transcripts of T3′100C accumulated to barely detectable levels (1% of wt TCV), indicating that one or more of the eight positional changes was strongly detrimental to TCV accumulation (Fig. 5). TCV with the satC H5 (T12C) or

---

**Fig. 4. Effect of CP on TCV subviral RNA replication in protoplasts.** (A) Sequence in the vicinity of the CP translation initiation codon in TCV and its CP-minus derivative CPmT. The normal AUG initiation codon and the alternative initiation codon that is used when the normal codon is mutated are underlined with arrows reflecting positions of translation initiation. The altered nucleotides in CPmT are indicated by lowercase letters. (B) RNA gel blot of mutant satC and TCV genomic RNA (+)-strands. Total RNA was extracted at 40 hpi from Arabidopsis protoplasts. Ethidium bromide staining of the gel before blotting shows ribosomal RNA loading control (below the blot). None, no added satC; wtC, wt satC.

---
satC-specific residues in positions 3, 4, 7 and 8 (T3478C) accumulated to 84 or 81% of wt TCV, respectively, suggesting these positions are only marginally involved in reduced accumulation of T3′100C. Additional combinations that retained TCV Pr positions 5 and 6 also had less than a 2-fold reduction in TCV accumulation (T1278C, 66%; T12347C, 56%; T123478C, 61%). In contrast, TCV with positions 5 and 6 of satC, either alone or in combination with other satC-specific bases, accumulated between <1 and 5% of wt TCV (Fig. 5). We also determined the effect of deleting one of the two CG dinucleotide repeats at position 5 in the TCV Pr loop, which generates TCV with the diG Pr (T-PrG). T-PrG accumulated to only 2% of wt levels, indicating that altering position 5 strongly impacts on the efficiency of the TCV promoter.

CCFV and JINRV Pr are similar to the TCV Pr in having extensively paired stems (10 and 12 bp, respectively; Fig. 6A). The CCFV Pr contains a similar loop as the TCV Pr and eight consecutive guanylates compared with 10 for TCV. We had previously determined that satC with the Pr of CCFV accumulates to 35% of wt satC (Zhang et al., submitted for publication), similar to satC with the TCV Pr (C568T, 37%; Fig. 3B). To determine if the Pr of other carmoviruses can function in the context of TCV, TCV was constructed to include the Pr of CCFV (Ts-PrCCFV) or JINRV (Ts-PrJINRV). To facilitate cloning, a new restriction site was created in the TCV H5/Pr linker by insertion of 3 nt. The resultant TCV, designated as TCVs, accumulated to identical levels as wt TCV in protoplasts (data not shown). Ts-PrCCFV accumulated to only 11% of TCVs, while Ts-PrJINRV did not reach detectable levels in protoplasts (Fig. 6B). These results support a relationship between carmoviral genomic RNA promoters and cognate upstream sequences.

The TCV Pr is a less efficient promoter than the satC Pr when assayed in vitro

To determine the in vitro transcriptional activity of satC with various positions converted to TCV, wt satC, C3′100T, C56T and C1256T were subjected to transcription by purified recombinant TCV p88 RdRp (Rajendran et al., 2002) and radiolabeled products examined following gel electrophoresis. As shown in Fig. 7, transcription of C3′100T was only 13% of wt satC, accounting for the decrease in accumulation obtained when this construct was assayed in protoplasts. C1256T transcription was reduced to 22% of wt satC. Unexpectedly C56T was transcribed nearly twice as efficiently as C1256T in vitro (42% of wt), suggesting the TCV Pr, in combination with TCV H5 is responsible for much of the reduced in vitro transcription of C3′100T. Since transcription in vitro mainly assays initiation (products are not templates for further synthesis), these results indicate that the TCV Pr is not as efficient a promoter as the satC Pr in the context of satC. Both C56T and C1245T produced products that migrated as doublets.
in these gels. We currently do not know if this is due to internal initiation occurring for a portion of the templates or if a portion of the products are uncharacteristically migrating to two different positions in the partially denaturing gels.

As a further test of the comparative promoter efficiencies of the TCV and satC Pr, both sequences were removed from their natural location and placed downstream of an unrelated sequence (Qβ bacteriophage-associated midivariant [MDV] RNA [220 nt]) that is not a template for TCV RdRp. We previously determined that the 3′-terminal 37 bases of (+)-strand satC, comprising the Pr hairpin, six flanking 3-terminal bases and eight flanking 5′ bases, can efficiently direct complementary strand synthesis of MDV using TCV RdRp partially purified from infected turnip plants (Song and Simon, 1995). To directly compare Pr activities in vitro, analogous Pr-containing segments from satC, TCV and diG were joined to MDV and template activity measured in vitro using p88 RdRp. As shown in Fig. 8, the satC Pr was able to direct 6-fold or 50-fold more complementary strand synthesis than the TCV Pr or the diG Pr, respectively. These results suggest that modifications in the TCV Pr that generated the Pr of satC resulted in a promoter that can function more efficiently in the absence of additional viral sequences in vitro.

Comparisons of the structures of satC and C56T reveals differences in the Pr, DR and H4a regions

SatC transcripts synthesized by T7 RNA polymerase assume an initial preactive structure that does not contain the phylogenetically conserved hairpins shown in Fig. 1B (Zhang et al., 2006). Template activity of these transcripts for RdRp-directed complementary strand synthesis is substantially enhanced when specific mutations within or outside the Pr region disrupt the initial Pr structure (Pr-1), causing the Pr to assume a structure that resembles its phylogenetically inferred form (active structure) known as Pr-2 or Pr-2*. Pr-2 and Pr-2* structures are very similar except that the 3′-terminal three cytidylates in Pr-2* remain in the Pr-1 configuration (double-stranded or stacked), whereas these residues in Pr-2 are single-stranded (Zhang et al., 2006). Although transcription is substantially enhanced for Pr-2*-containing mutant satC transcripts in vitro, none of these transcripts accumulated to detectable levels when inoculated into protoplasts. These results were interpreted to reflect a satC requirement for both preactive and active conformations in vivo, whereas in vitro, the initial presence of the active Pr-2* form enhances transcription independent of a (poorly executed) conformational switch.

Our initial plans were to compare the structures of satC, C56T and C3′100T with the expectation that the increased stability of the TCV-like and TCV Pr in C56T and C3′100T, respectively, would result in a Pr that resembled the phylogenetically inferred Pr-2* form. Transcripts were radioactively end-labeled, purified from acrylamide gels and partially digested with RNase T1 (specific for single-stranded guanylates), RNase A (specific for single-stranded pyrimidines) or RNase V1 (specific for double-stranded or stacked nucleotides). Following RNase digestion, RNA fragments were subjected to electrophoresis through 20% or 10% polyacrylamide gels to examine the Pr region or upstream regions, respectively. A ladder of RNase T1-digested RNA that had been denatured prior to enzyme treatment was also included. For ease in describing specific cleavages, residues are referred to their positions in active structure elements and, when presented, are mapped on these active structure elements since the preactive structure remains unknown.

![Fig. 7. Transcription in vitro using purified recombinant TCV RdRp. Transcripts of constructs indicated above each lane were synthesized by T7 RNA polymerase and subjected to complementary strand synthesis by TCV RdRp in the presence of [α-32P]-UTP. Ethidium stained gel of template RNAs is presented on the left. Nomenclature is as previously described in legend to Fig. 3. MDV, non-specific 220 nt RNA associated with Qβ bacteriophage used as a control for template specificity of the RdRp. Numbers below each lane are average values for three independent experiments.](image1)

![Fig. 8. Transcription efficiencies of TCV, satC and diG core promoters. (A) Schematic representation of the chimeric RNAs containing Pr hairpin and short flanking sequences that were joined to MDV RNA. Actual core promoter sequences are illustrated in Figs. 1B and 3A. (B) In vitro RdRp assay. Transcripts of constructs indicated above each lane were synthesized by T7 RNA polymerase and subjected to complementary strand synthesis by TCV RdRp in the presence of [α-32P]-UTP. Identities of the attached Pr are shown above each lane. The gel was stained with ethidium–bromide to indicate template levels. Values below lanes are averages for two independent experiments. C-Pr, Pr of satC; G-Pr, Pr of diG; T-Pr, Pr of TCV; MDV, MDV RNA without any Pr sequence.](image2)
A comparison of the Pr regions for wt satC, C56\textsubscript{T} and the satC mutant H5RL/G218C that folds into the Pr-2* form is shown in Fig. 9. H5RL/G218C contains three base changes in the H5 LSL, which result in the Pr-2* configuration and an additional variance in the DR region (Zhang and Simon, unpublished). As previously reported (Zhang et al., 2006), the Pr-1 structure is distinguished by strong RNase A cleavages at positions U331/C336, non-specific cleavages between C334 and G315 and reproducible compression between positions C334 and G324. In contrast, Pr-2* has weak or undetectable RNase A cleavages at U335/C336, very few non-specific RNase A cleavages at U335/C336, (1) few non-specific cleavages compared to wt satC, although the non-specific RNase T\textsubscript{1} cleavage at position C341 between two highly reactive guanylates in the Pr loop was an exception; (2) no strong RNase A cleavages at U335/C336; and (3) no compression within the Pr region (Fig. 9). However, unlike Pr-2*, C330 and U331, located in the lower half of the hairpin stem, were susceptible to RNase A cleavage. Interestingly, despite the very stable C56\textsubscript{T} hairpin predicted by mFold
\( \Delta G = 20.4 \text{ kcal/mol} \), the lower three presumptive paired bases of the hairpin were not susceptible to RNase V1, similar to satC Pr-1 and Pr-2* structures. This result suggested possible fold-back pairing of the terminal cytidylates with the guanylates at the base of the stem (Zhang et al., 2006). Unfortunately, despite numerous attempts over a 4-month period, we were unable to obtain a decipherable pattern for C3′100, including the control ladder sample, although the patterns obtained in repetitive assays with different transcript preparations were identical (data not shown). We currently have no explanation for this previously unencountered problem.

We have reported that all transcripts containing the 3′ end Pr-2/Pr-2* configuration also contain distinctive structural differences in the DR/H4a region compared with Pr-1-containing transcripts (Zhang et al., 2006). To determine whether the Pr-2*-like configuration of C56\( _T \) was similarly associated with these upstream structural variances, the RNase cleavage pattern of this region was compared for wt satC, C56\( _T \) and satC with a deletion of the 5′ two guanylates (CΔ2G), which folds into the Pr-2* configuration at its 3′ end (Zhang and Simon, unpublished). As previously shown, Pr-1-containing constructs contain three weak RNase A or RNase T1 cleavages in the DR, five weak cleavages in the H4a stem and a strong RNase A cleavage at U233 (Fig. 10). In contrast, Pr-2/Pr-2* containing constructs have no cleavages in the DR, only two weak RNase A or RNase T1 cleavages in the H4a stem and maintain the strong RNase A cleavage at U233. Pr-2/Pr-2* also contains a new RNase T1 cleavage at G230. Examination of this region in C56\( _T \) indicates the presence of the identical pattern found in all Pr-2/Pr-2* containing constructs (Fig. 10). These results suggest that satC containing positions 5 and 6 of TCV has a terminal hairpin that adopts a Pr-2*-like configuration. However, unlike satC variants with promoters in the form of Pr-2/Pr-2*, transcripts of C56\( _T \) are not transcribed at high efficiency in vitro.

**Discussion**

Efficient accumulation of (+)-strand RNA viruses requires direct or indirect long distance interactions between their 3′ and 5′ ends (Barton et al., 2001; Frolov et al., 2001; Herold and Andino, 2001; Khromykh et al., 2001; You et al., 2001; Vlot and Bol, 2003). Such long distance interactions may be necessary to coordinate translation and replication, present the initiation site to the replicase and/or stabilize the viral RNA. The initial event that created satC, however, resulted in a molecule that retained only the 3′ region of TCV joined to satD, a molecule with limited similarity to the TCV genomic RNA. While modern satC accumulates to levels similar to 5S rRNA, making it one of the most prevalent RNAs in an infected cell, satC with the 3′ 100 bases of TCV (C3′100\( _T \)), which should more closely resemble the progenitor RNA, is a poor template that accumulates to only 15% of modern (wt) satC levels.

Mutational analysis revealed that satC constructs were inefficient templates when they contained TCV positions 5 and 6 located in the core Pr promoter. Little difference in accumulation was found when C56\( _T \) also contained additional TCV-specific bases in position 8 in the Pr, positions 1 and 2 in the upstream hairpin H5 or positions 3 and 4 in the linkH5-Pr region (34 to 37% of wt satC for all constructs). The two-fold difference in accumulation between these constructs and C3′...
showing that TCV full-length genomic RNA, but not C3 upstream sequences is also supported by our previous results. The TCV RdRp can similarly recognize the Pr of TCV and CCFV in the context of satC. The increased efficiency of the satC Pr in directing complementary strand synthesis in vitro suggests that evolutionary adaptation has allowed the Pr to function independent of its upstream interacting element. diG, whose 3′-end region closely resembles that of TCV (Li et al., 1989), accumulates poorly in vivo (11% of wt satC; Fig. 3B), suggesting that the diG Pr is less adapted for high-level accumulation than the Pr of satC. However, the 2-nt deletion at position 5 in the diG Pr, while not obviously affecting accumulation of diG (compare wtG with G-PrT, Fig. 3B), results in a nearly 2-fold enhancement in accumulation of satC compared to satC with the TCV Pr (Fig. 5). This difference suggests that when associated with a possibly “more evolved” molecule like satC, the diG Pr is able to function with greater independence than the TCV Pr. The importance of these two nucleotides at position 5 in the TCV Pr loop for TCV accumulation (deletion resulted in a 50-fold decrease [Fig. 5]) suggests that the Pr loop region may be the 3′ element that interacts with upstream sequences missing in diG and satC.

The recently discovered satC conformational switch involving the 3′-terminal 140 nt must also be considered as resulting from the base differences between TCV and satC (Zhang et al., 2006, submitted for publication). In TCV genomic RNA, as with other viral genomic RNAs, a conformational switch is probably needed to convert the template from one that is translationally competent to one that is active for replication (van Dijk et al., 2004). Since satC is not translated, its switch may be important to reduce or eliminate the ability of newly synthesized (+)-strands to serve as templates for further (−)-strand synthesis (Chao et al., 2002; Brown et al., 2004). Base differences at positions 1, 2, 5 and 6 all reduce the stability of the satC active structure, which, along with other base changes, may have helped in evolving an alternative structure comprising only available local sequences.

Materials and methods

Construction of satC, diG and TCV mutants

To construct plasmid C7T, oligonucleotides C268U and oligo7 were used as primers (all primers are listed in Table 1), for PCR with template pT7C+, a plasmid containing wt satC cDNA downstream from a T7 RNA polymerase promoter (Song and Simon, 1994). PCR products were treated with T4 DNA polymerase, digested with appropriate restriction enzymes and cloned into the analogous location in pT7C+, which had been treated with the same restriction enzymes. Most other constructs were generated in a similar fashion unless noted, and all mutants created for this study were confirmed by sequencing. C8T was constructed using primers T7C5′ and U335C while C34T, C347T and C1278C, T12568C, T123478C and T12347C were generated in a similar fashion as C56T except that oligonucleotides C568C* and C34T, C347T and C1278C, T12568C, T123478C and T12347C were used as templates, respectively. C568T was generated in a similar fashion as C56T except that oligonucleotide C56C* was replaced by C568C*.

T12C was generated using primers T7C5′ and C568C* and template C347T. PCR products were digested with appropriate enzymes and cloned into the analogous location in pT7C+. T12568C, T125678C, T1278C, T125678C, T123478C and T12347C were generated by digesting plasmids C12T, C1234T, C125678C, C1278C, T125678C, T123478C and T12347C with SpeI and SmaI. The
Table 1  
Summary of the oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Application/Construct</th>
<th>Name</th>
<th>Position&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sequence&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Polarity&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutagenesis in satC</td>
<td>T7C'</td>
<td>1–19</td>
<td>5'-GTAATACGACTCACTATAGGGAUAACUAAGGGTTTCA</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Oligo 7</td>
<td>338–356</td>
<td>5'-GGGCAGGCCCTCTCCGGA</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>C268U</td>
<td>253–279</td>
<td>5'-GAAAATAGGCTCTCTGGGGTAAACAC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>U335C</td>
<td>319–356</td>
<td>5'-GGCGAGGCCCTCTCGAGGGAGGCTATCTATTG</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>C34C*</td>
<td>307–356</td>
<td>5'-GGCGAGGCCCTCTCCGAGGGAGGCTATCTTTTAGTCCGAGGAGTC</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>C56C*</td>
<td>324–356</td>
<td>5'-GGCGAGGCCCTCTCCGAGGGAGGCTATTGAC</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>C568C*</td>
<td>318–356</td>
<td>5'-GGGCAGGCCCTCTCCGAGGGAGGCTATCTTTAGTCCGAGGAGTC</td>
<td>−</td>
</tr>
<tr>
<td>Mutagenesis in TCV</td>
<td>Oligo 3495</td>
<td>3495–4012</td>
<td>5'-GGGCAGACTGGCAAGGGTTTGA</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>T5G</td>
<td>4023–4054</td>
<td>5'-GGGCAGGCCTCTCCGGAAGGAGGG</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Oligo 3164</td>
<td>3164–3181</td>
<td>5'-ATGAGCCCTTCAACGCCAC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>C3994S</td>
<td>3983–4054</td>
<td>5'-aggatccccGGGAGGGCCCCCCTCCTCGAGGGTACCTTTTTAGTCCGAGGAGTCACCAACACCTTTTC</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>G4051S</td>
<td>3948–4030</td>
<td>5'-AAAAGCTAGTCCCTTTGGTAAACACTAAAATCCCGAAAGGGTGTCGGTTGTCGACCCTCCGAAC(TACTC)AAATAGCCCTCCAC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>SNAB1</td>
<td>3936–4054</td>
<td>5'-GGGCAGGCCCTCTCCGGAAGGAGGG</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>KK57</td>
<td>4036–4054</td>
<td>5'-GGGCAGGCCCTCTCCGGAAGGAGGG</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>JINRVPr+</td>
<td>4015–4045</td>
<td>5'-GTAAAGATACCCCTCTCCGAGGGGAGGGCTGAGCC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>JINRVPr−</td>
<td>4015–4054</td>
<td>5'-GGGCAGGCCCTCTCCGAGGGGAGG</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>CCFVPr+</td>
<td>4015–4054</td>
<td>5'-GGGCAGGCCCTCTCCGAGGGGAGGCTATCTTTAC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>CCFVPr−</td>
<td>4015–4054</td>
<td>5'-GGGCAGGCCCTCTCCGAGGGGAGGCTATCTTTAC</td>
<td>−</td>
</tr>
<tr>
<td>Mutagenesis in diG</td>
<td>GST</td>
<td>324–353</td>
<td>5'-GGGCAGGCCCTCTCCGAGGGGAGG</td>
<td>−</td>
</tr>
<tr>
<td>Construction of</td>
<td>GP5</td>
<td>315–322</td>
<td>5'-AAAAGAAGACACCTCCCCTCC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>MDV/Pr</td>
<td>320–356</td>
<td>5'-ATTAAGATACCTCCCTCCTCGGGAGGGGCGCTGCCC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>chimera</td>
<td>301–356</td>
<td>5'-AAAAGAAGACACCTCCCTCCTCGGGAGGGGCGCTGCCC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>TPr</td>
<td>4014–4054</td>
<td>5'-AAAAGAAGACACCTCCCTCCTCGGGAGGGGCGCTGCCC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>NDE</td>
<td>176–195</td>
<td>5'-AGTGGACCATATGCGGTGTG</td>
<td>−</td>
</tr>
<tr>
<td>RNA gel blots</td>
<td>Oligo 13</td>
<td>249–269</td>
<td>5'-AGAGAGACAGTTCTCCAGG</td>
<td>−</td>
</tr>
</tbody>
</table>

<sup>a</sup> Coordinates correspond to those of the TCV genome, diG, satC or pUC19 (Oligo NDE) as indicated.

<sup>b</sup> Bases in italics indicate T7 RNA polymerase promoter sequence. Bases in lowercase were added to achieve efficient digestion. Mutant bases are underlined. Bold bases denote nucleotides inserted in TCV promoter compared to satC or diG promoter. Bold and underlined bases indicate nucleotides inserted in TCV to generate a SnaBI site as shown in parentheses.

<sup>c</sup> ‘+’ and ‘−’ polarities refer to homology and complementarity with sat C plus strands, respectively.

<sup>d</sup> Oligo 13 is also complementary to positions 3950 to 3970 of TCV genomic RNA.
Hybridized oligonucleotides were treated with T4 polynucleotide kinase and inserted into pUC19T7MDV (Guan, 2000) that had been treated with SmaI with inserted SnaB1 site in the linker between H5 and Pr, was generated with oligonucleotides SNA1B and KK57 and template pTCV66. Transcripts of pTSN-L5 (TCVs) and TCV66 (wtTCV) accumulated to similar levels in protoplasts (data not shown). Ts-PrJINRV and Ts-PrCCFV were made by annealing either oligonucleotides JINRVPr+ and JINRVPr− or oligonucleotides CCFVPPr+ and CCFFPPr−. Hybridized oligonucleotides were treated with T4 polynucleotide kinase, digested with SmaI and cloned into the analogous location in pT7diG that had been treated with SpeI and SmaI.

To construct plasmid G5Pr, primers G5′ and GST were used with template pT7diG (Li and Simon, 1991). PCR products were subsequently treated with T4 DNA polymerase and SpeI and cloned into the analogous location in pT7diG digested with SnaB1 and SmaI.

**Generation of MDV/Pr constructs**

To construct plasmid MDV-GPr, oligonucleotides GPr5 and NDE and NDE were used as primers with template C3456T. Following treatment with T4 DNA polymerase and NdeI, the fragment was inserted into pUC19T7MDV (Guan, 2000) that had been treated with SmaI and NdeI. Plasmid MDV-CPr was generated in similar fashion using primers CPr and NDE and template MDV-GPr. For plasmid MDV-TPr, oligonucleotide CPr was replaced with CPr in a similar fashion using primers CPr and NDE and template MDV-GPr. The linker between H5 and Pr, was generated with oligonucleotides SpeI and cloned into pTCV66, which had been treated with the same restriction enzymes.

**Transcript preparation, inoculation of Arabidopsis protoplasts and RNA gel blots**

TCV genomic RNA, satC and diG transcripts were synthesized by T7 RNA polymerase using plasmids linearized with SmaI, which generates transcripts with precise 5′ and 3′ ends. Protoplasts (5 × 10⁶) prepared from Arabidopsis (Arabidopsis thaliana) were inoculated with 2 μg of TCV genomic RNA transcripts with or without 2 μg of satC RNA transcripts using PEG-CaCl₂, as previously described (Kong et al., 1997). Total RNAs were isolated from protoplasts at 40 hpi and subjected to RNA gel blot analysis. Plus-strand RNA was probed with a [γ-³²P]-ATP-labeled oligonucleotide that is complementary to both satC and TCV sequences (Zhang and Simon, 2003).

**RNA solution structure probing of 3′ end-labeled transcripts**

RNA structure probing of satC and its mutants was performed using a protocol and reagents obtained from Ambion. Briefly, gel-purified transcripts were labeled at the 3′ end in a final volume of 20 μl containing 6 μg of transcripts, 50 μCi of [³²P]–pCp (Amersham), 20 U of T4 RNA ligase, 50 mM Tris–HCl (pH 7.8), 10 mM MgCl₂, 10 mM dithiothreitol and 1 mM ATP. After overnight incubation at 4 °C, reactions were terminated by phenol–chloroform extraction. Labeled transcripts were separated in a 6% sequencing gel and eluted by soaking overnight with constant shaking in a buffer containing 25 mM Tris–HCl (pH 7.5), 400 mM NaCl and 0.1% SDS. Labeled transcripts were added to a mixture containing RNA structure buffer (10 mM Tris–HCl pH 7.0, 100 mM KCl, 10 mM MgCl₂), yeast tRNA and either water or RNase T₁ (0.01 or 0.001 U/μl), RNase A (0.01 or 0.001 μg/ml), or RNase V₁ (0.001 or 0.0001 U/μl) and incubated at 22 °C for 15 min. Samples were precipitated, resuspended in 10 μl of loading buffer and subjected to electrophoresis through 10% or 20% sequencing gels. Alkaline hydrolysis ladders were obtained by treatment of 3′ end-labeled transcripts with alkaline hydrolysis buffer (Ambion) at 95 °C for 5 min. To obtain RNase T₁ ladders, 3′ end-labeled transcripts were heated at 95 °C for 5 min in buffer supplied by the manufacturer (Ambion), then cooled on ice, followed by RNase T₁ (0.1 or 0.01 U/μl) digestion at 22 °C for 15 min.

**In vitro RdRp assay**

In vitro RdRp assays were carried out using recombinant TCV p88. The p88/maltose binding protein-expressing plasmid, a kind gift from P.D. Nagy (University of Kentucky), was transformed into E. coli Rosetta (DE3) pLacI competent cells (Novagen), and expression and purification of the recombinant protein carried out as described (Rajendran et al., 2002). In vitro RdRp assays were performed in the presence of 10 mM MgCl₂, the same concentration as used for RNA structure probing. Briefly, 1 μg of purified RNA template was added to a 25 μl reaction mixture containing 50 mM Tris–HCl (pH 8.2), 100 mM potassium glutamate, 10 mM MgCl₂, 10 mM dithiothreitol, 1 mM each of ATP, CTP, GTP, 0.01 mM UTP, 10 μCi of [α-³²P]-UTP (Amersham) and 2 μg of recombinant p88. After a 90-min incubation at 20 °C, 1 μg of RNA was added, and the mixture was subjected to phenol–chloroform extraction and ammonium acetate–isopropanol precipitation. Radiolabeled products were analyzed by denaturing 8 M urea–5% polyacrylamide gel electrophoresis followed by autoradiography.

**Acknowledgments**

Funding was provided by grants from the U.S. Public Health Service (GM61515-01) and the National Science Foundation (MCB-0086952) to AES. JCM was supported by NIH predoctoral training grant T32-AI51967-01.

**References**


487


