

## Analysis *in Vivo* of Turnip Crinkle Virus Satellite RNA C Variants with Mutations in the 3'-Terminal Minus-Strand Promoter

Vera Stupina<sup>1</sup> and Anne E. Simon<sup>2</sup>

*Department of Biochemistry and Molecular Biology and Program in Molecular and Cellular Biology,  
University of Massachusetts, Amherst, Massachusetts 01003*

*Received August 7, 1997; returned to author for revision August 25, 1997; accepted September 17, 1997*

Turnip crinkle virus and its associated RNA, sat-RNA C, share similar, but not identical hairpins near their 3' ends and terminate with CCUGCCC-OH, which forms a single-stranded tail. With an *in vitro* transcription system containing partially purified TCV RdRp, the 3'-terminal 29 bases making up the hairpin and single-stranded tail were previously demonstrated to be required for transcription, and alterations in the stem, but not the loop, could affect template activity (C. Song and A. E. Simon, 1995, *J. Mol. Biol.* 254, 6–14). We have now analyzed sat-RNA C mutants in the 3' hairpin for ability to accumulate *in vivo*. While active templates *in vitro* were able to accumulate *in vivo*, some very weak templates *in vitro* were also able to accumulate *in vivo* without reversion or second-site alterations. Computer models of hairpin structure indicated that biologically active promoters could have hairpins less stable than wild type, with loops of variable length and sequence, and without a need for a 6-base single-stranded tail. In addition, transcripts containing compensatory exchanges in the upper stem region that had limited activity *in vitro* were biologically active *in vivo*, indicating that positioning of specific bases in the stem is not required to produce an active minus-strand promoter. © 1997 Academic Press

### INTRODUCTION

Replication of positive-sense, single-stranded RNA viruses begins with an interaction between the viral RNA-dependent RNA polymerase (RdRp) and sequences or structural elements that make up the promoter for minus-strand synthesis and are usually contained within the 3'-terminal 200 bases of the viral RNA (French and Ahlquist, 1987; Dreher and Hall, 1988; Takamatsu *et al.*, 1990; Bocard and Baulcombe, 1993; Ball and Li, 1993; Rohll *et al.*, 1993; Duggal *et al.*, 1994; Lin *et al.*, 1994; Havelda *et al.*, 1995). Structural elements contained in the 3'-terminal region of many viruses include tRNA-like structures, pseudoknots, and small hairpins, some of which have been shown to be important in virus replication (Jupin *et al.*, 1990; Pilipenko *et al.*, 1992; Tsai and Dreher, 1992; Jacobson *et al.*, 1993; Rohll *et al.*, 1993, 1995; Duggal *et al.*, 1994; Havelda and Burgyn, 1995).

Many viruses are associated with small, dispensable subviral RNAs that require the replication machinery of a specific helper virus for amplification. Defective interfering (DI) RNAs are deletion derivatives of the helper virus (Roux *et al.*, 1991), while satellite (sat-) RNAs share no or partial sequence similarity with their helper virus (Simon and Howell, 1986; Roossinck *et al.*, 1992). Since DI and sat-RNAs must contain specific *cis*-sequences and/or struc-

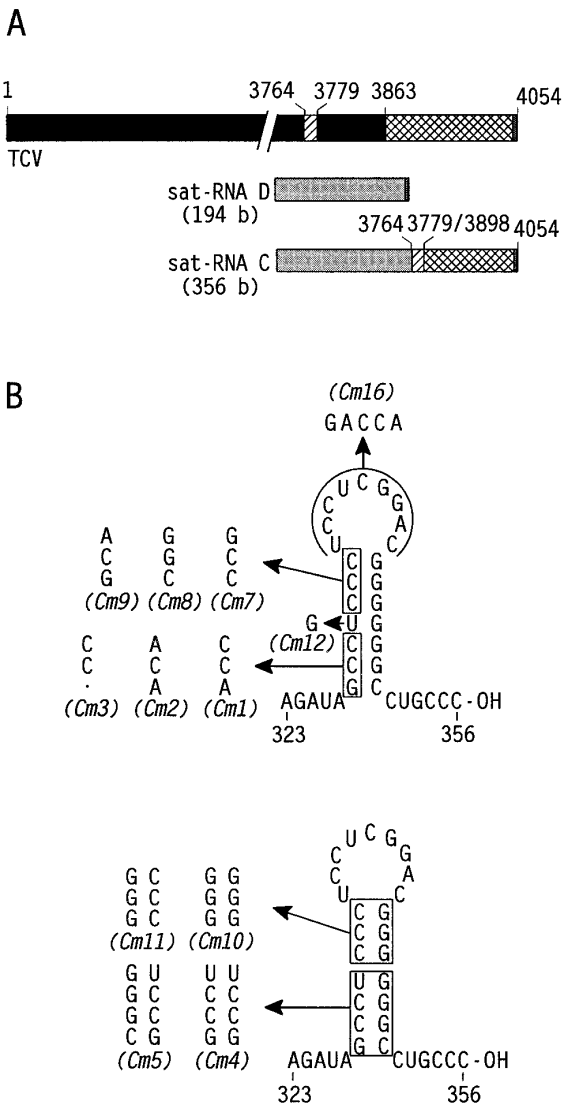
tures required for replication by their helper virus RdRp and are usually much smaller than their corresponding helper virus, they have proven to be valuable tools for elucidating RdRp promoter sequences (Levis *et al.*, 1986; Tsiang *et al.*, 1988; Li *et al.*, 1991; Kim *et al.*, 1993; Lin and Lai, 1993; Thomson and Dimmock, 1994; Chang *et al.*, 1995; Havelda *et al.*, 1995; Song and Simon, 1995).

Turnip crinkle virus (TCV) is a 4054-base single-stranded, positive-sense RNA virus that belongs to the carmovirus group (Heaton *et al.*, 1989; Oh *et al.*, 1995). TCV is naturally associated with several sat- and DI RNAs that do not possess open reading frames and require a co-inoculated viral genomic RNA for replication (Simon and Howell, 1986; Li *et al.*, 1989). Sat-RNA C is a chimeric RNA of 356 bases that consists of a nearly full-length smaller sat-RNA (sat-RNA D) at its 5' end linked to two regions from the 3' end of TCV (Simon and Howell, 1986, Fig. 1A). TCV and sat-RNA C share similar, but not identical, hairpins near their 3' ends. In addition, all TCV-associated RNAs terminate with 5' CCUGCCC-OH at their 3' ends, the last 6 bases of which form a single-stranded tail and are required for replication (Carpenter *et al.*, 1995; Carpenter and Simon, 1996a,b; Nagy *et al.*, 1997). The genomic RNAs of other carmoviruses (Guilley *et al.*, 1985; Nutter *et al.*, 1989; Skotnicki *et al.*, 1993) also contain 3'-terminal sequences predicted to fold into a stable hairpin and similar 6-base single-stranded tails (Song and Simon, 1995).

An *in vitro* transcription system containing partially purified TCV RdRp is able to synthesize full-length comple-

<sup>1</sup> Current address: Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138.

<sup>2</sup> To whom correspondence and reprint requests should be addressed. Fax: (413) 545-4529. E-mail: simon@biochem.umass.edu.



**FIG. 1.** TCV and its associated sat-RNA C. (A) Origin of sequences that constitute sat-RNA C. Sat-RNA C is a hybrid RNA formed from recombination between sat-RNA D and TCV. Similar regions are shaded alike. Numbering above the RNA diagrams corresponds to positions in TCV isolate TCV-M (Oh *et al.*, 1995). (B) Structure of the sat-RNA C 3'-terminal promoter required for minus-strand synthesis. The structure of the wild-type hairpin was determined by structure-probing experiments (Song and Simon, 1995). The 3'-terminal 29 bases (from position 328 at the 5' base of the hairpin) were required for template activity *in vitro* in a deletion analysis (Song and Simon, 1995). Sequences of mutants derived from sat-RNA C are shown, with names of mutants in italics below the altered sequences.

mentary strands of plus- or minus-strand sat-RNA C template (Song and Simon, 1994). In this *in vitro* system, sat-RNA C containing deletions extending into the 3'-terminal 29 bases were no longer templates for transcription (Song and Simon, 1995). In addition, the 37-base 3'-terminal sequence was able to promote transcription of an inactive, non-TCV-associated RNA. The 3'-terminal 29 bases comprise the hairpin and single-stranded tail described above (Fig. 1B). Testing of sat-RNA C mutant

templates containing mutations in the lower stem, upper stem, and loop regions of the hairpin in *in vitro* transcription reactions indicated: (1) the size and sequence of the loop are not important, (2) mutations that disrupt the lower stem substantially reduce template activity that is restored by compensatory mutations, and (3) mutation of multiple bases in the upper stem substantially reduces template activity that is not substantially restored by compensatory mutations (Song and Simon, 1995).

For the current study, we analyzed a number of sat-RNA C mutants for ability to accumulate *in vivo* and compared the results with the RNA's *in vitro* transcription activity. While templates that were highly active *in vitro* were all able to accumulate *in vivo* without reversion or second-site alterations, some very weak templates *in vitro* were also able to accumulate *in vivo*. Analysis of computer-predicted structures for sat-RNA C mutants that accumulated in plants confirmed and extended the prior *in vitro* analysis, indicating that biologically active promoters could have hairpins both more and less stable than wild type, with loops of variable length and sequence, and without a need for a 6-base single-stranded tail. In addition, transcripts containing compensatory exchanges in the upper stem region were biologically active *in vivo*, indicating that positioning of specific bases in the stem is not required to produce an active minus-strand promoter.

## MATERIALS AND METHODS

### Preparation of infectious transcripts, plant inoculations, and detection of sat-RNA C

Construction of plasmids containing mutations in the 3'-terminal hairpin have been previously described (Song and Simon, 1995). Prior to transcript synthesis, plasmids were linearized with *Sma*I, except for mutant Cm11, which was linearized with *Bam*HI since it contains an internal *Sma*I site. Linearized plasmids were transcribed *in vitro* using T7 RNA polymerase as described (Carpenter *et al.*, 1995). The RNA transcripts were full-length sat-RNA C containing the natural 5' end and either the natural 3' end (*Sma*I-digested plasmids) or 5 additional 3'-end bases (*Bam*HI-digested plasmid). Two-week-old turnip seedlings were inoculated with sat-RNA C mutant or wild-type transcripts and TCV transcripts prepared from pT7TCVms (Carpenter *et al.*, 1995) in inoculation buffer as previously described (Li *et al.*, 1989). Two weeks postinoculation, total RNA was extracted from uninoculated leaves and sat-RNA C was purified following electrophoresis on 5% denaturing acrylamide gels. Purified sat-RNA C was polyadenylated using poly(A) polymerase and reverse transcribed using MmLV reverse transcriptase to obtain sat-RNA C cDNA, as previously described (Carpenter and Simon, 1996b). Oligonucleotide T<sub>17</sub> adapter (Table 1) was used to synthesize cDNA, and then oligonucleotides DM4 and Adapter (Table 1)

TABLE 1  
List of Oligonucleotides

Application	Name	Position <sup>a</sup>	Sequence <sup>b</sup>	Polarity <sup>c</sup>
cDNA amplification for cloning and sequencing	DM4	206–223	5'-GGGACCAAAAACGGG	+
	Oligo 7	338–356	5'-GGGCAGGCCCCCGTCCGA	–
	T <sub>17</sub> adapter		5'-CCACTCGAGTCGACATCGA(T) <sub>17</sub>	
	Adapter		5'-CCACTCGAGTCGACATCGA	
Site-directed mutagenesis	SL22	57–78	5'-AAACTTGACTGATGACCCCTAC	+
	Oligo V	342–356	5'-GGGCAGGCCCGCCCTCCGA	–
	Oligo V2	338–356	5'-GGGCAGGCCCGCCGTCCTCCGA	–
	Oligo V3	327–356	5'-GGGCAGGCCCCCGTCCGAGGACTCAGGCT	–

<sup>a</sup> Positions in sat-RNA C.

<sup>b</sup> Bases in italics are primary mutation sites within the oligonucleotide. Bases in bold italics are second-site mutations.

<sup>c</sup> "–" and "+" polarities are complementary or identical, respectively, to plus-strand sat-RNA C.

were used to amplify the 3' region of sat-RNA C cDNA by polymerase chain reaction (PCR) using conditions previously described (Carpenter and Simon, 1996b). Following treatment with DNA polymerase large fragment, sat-RNA C cDNA was ligated to *Sma*I-digested pUC19T7 (Song and Simon, 1995). Plasmids purified from colonies were sequenced using Sequenase Version I (Amersham) and primer DM4.

#### Construction of marker mutations

Several sat-RNA C mutants gave rise to progeny *in vivo* with wild-type sequence. Since discrimination was required between reversion of the mutant transcripts to wild-type sequence and possible contamination of wild-type sat-RNA C, a 4-base marker mutation was introduced into the 5' region of these sat-RNA C mutants (Cm1 and Cm2) by digesting the plasmids with *Nco*I, which cleaves 105 bp from the 5' end of the sat-RNA C cDNA, and treating the ends with *Escherichia coli* DNA polymerase I (large subunit) followed by intramolecular ligation. Progeny of these transcripts accumulating in plants were cloned as described above except that oligonucleotide SL22 (Table 1) was used in place of DM4. Sat-RNA cDNA in the plasmids was sequenced with oligonucleotide SL22 to confirm disruption of the *Nco*I site.

#### RNA secondary structure analysis

The M-fold program from the GCG package (Genetics Computer Group, Inc., University of Wisconsin) was used to predict the most stable secondary structure for the 3' ends of sat-RNA C mutants.

#### Site-directed mutagenesis of sat-RNA C within the 3'-end hairpin region

To prepare transcripts from mutants containing second-site alterations for *in vitro* analysis of template activity, PCR was used to introduce site-specific mutations into full-length sat-RNA C cDNA using oligonucleotides DM4 and V, V2, and V3 (Table 1) for second-site mutants Cm10b,

Cm10a, and Cm10c, respectively. Oligonucleotides Oligo 7 and DM4 were used to amplify DNA directly from cDNA clones of second-site mutants Cm2a, Cm7a, and Cm9a. The amplified regions were digested with *Bst*EII and ligated into *Sma*I/*Bst*EII-digested pUC19T7C(+). Sequences were confirmed by DNA sequencing using primer DM4.

#### *In vitro* transcription of sat-RNA C using partially purified TCV RdRp

pUC19T7C(+)-derived plasmids containing full-length wild-type or mutant sat-RNA C cDNA were linearized with *Sma*I and transcripts synthesized using T7 RNA polymerase. Two micrograms of RNA transcripts were subjected to RNA-directed RNA synthesis in the presence of [ $\alpha$ -<sup>32</sup>P]UTP and partially purified RdRp extracts (Song and Simon, 1994). Products were analyzed by electrophoresis on denaturing 5% polyacrylamide sequencing-length gels. Following electrophoresis, gels were incubated in ethidium bromide for 20 min for template detection, dried, and exposed to X-ray film. Autoradiograms from three independent experiments were quantified using an imaging densitometer and software provided by the manufacturer (Bio-Rad). Product levels were normalized to the level of template, which was quantified by densitometry following photography of the stained gels.

## RESULTS AND DISCUSSION

#### Effects of mutations in the 3'-end hairpin on accumulation of sat-RNA C *in vivo*

To analyze the relationship between template activity *in vitro* and ability to accumulate *in vivo*, sat-RNA C mutants containing modifications in the 3'-end hairpin were inoculated onto turnip plants along with TCV helper virus. All mutations altered the sequence and stability of the hairpin, the sequence of the hairpin while maintaining hairpin stability, or the sequence of the loop region (Fig. 1B). At 2 weeks postinoculation, total RNA was isolated from plants and analyzed for the presence of sat-RNA

TABLE 2

Effect of Mutations in the 3' Hairpin on Accumulation and Composition of Sat-RNA C *in Vivo*

Group	Mutant	<i>In vitro</i> activity <sup>a</sup>	Sat-RNA C present <sup>b</sup>	Analyzed <sup>c</sup>	No. of clones with original mutations	No. of clones reverted to wild type	Second-site alterations <sup>d</sup>
I	Cm4	—	0/6	—	—	—	—
II	Cm1	+	12/12	7/3	0	7	0
II	Cm2	—	2/6	8/2	0	5	Cm2a: from U <sub>331</sub> to G (3)
II	Cm10	+	8/12	13/3	0	0	Cm10a: from G <sub>346</sub> to C (1) Cm10b: from C <sub>352</sub> UGCCC to UGGCUGCCC (9) Cm10c: from G <sub>333</sub> to A (2) Cm10d: from G <sub>345</sub> to U (1)
II/III	Cm9	—	6/6	10/2	4	0	Cm9a: from A <sub>334</sub> to G (6)
III	Cm3	+	12/12	9/4	9	0	0
III	Cm5	++	6/6	3/1	3	0	0
III	Cm7	++++	6/6	8/1	7	0	Cm7a: from G <sub>334</sub> to U (1)
III	Cm8	++	6/6	4/1	4	0	0
III	Cm11	+	6/6	4/2	4	0	0
III	Cm12	++	12/12	6/2	5	0	Cm12a: from U <sub>331</sub> to A (1)
III	Cm16	+++	6/6	5/1	5	0	0

<sup>a</sup> Activity levels are from Song and Simon (1995). Wild-type sat-RNA C level was ++.

<sup>b</sup> Number of plants containing sat-RNA C species visible in ethidium bromide-stained gel/total number of plants inoculated. The amount of sat-RNA that accumulated in positive plants was approximately equivalent for all mutants and wild type.

<sup>c</sup> Number of clones sequenced/number of plants from which the clones were derived.

<sup>d</sup> Numbers in parentheses indicate the number of clones recovered with the sequence shown.

C-sized species, which are normally clearly visible in ethidium bromide-stained gels. Sat-RNA C species accumulating in plants were then cloned and their sequences determined. The mutants chosen for the *in vivo* analysis were Cm1, Cm2, Cm4, Cm9, Cm10, and Cm11 (low *in vitro* transcription activity using TCV RdRp extracts; Song and Simon, 1995) and Cm5, Cm7, Cm8, Cm12, and Cm16 (nearly as active or more active in *in vitro* transcription than wild type).

Based on the results of the *in vivo* accumulation and sequence analysis described below and summarized in Table 2, the mutants were subdivided into four general groups based on whether or not inoculation of transcripts led to accumulation in plants and whether the original mutations were maintained in the accumulating sat-RNA C. Group I was designated for mutants that did not produce any detectable sat-RNA C in plants. The only mutant in this class was Cm4, which contained a 4-base alteration in the lower stem (G<sub>347</sub>GGC to UCCG; number refers to position in sat-RNA C) that was predicted to substantially alter the structure of the promoter region (Fig. 2). The nonviability of Cm4 transcripts in plants was in correspondence with prior *in vitro* transcription results using partially purified TCV RdRp where no detectable full-length minus strands were observed (Song and Simon, 1995).

Group II mutants were designated as those that were able to accumulate in plants, but all cDNAs derived from the accumulating sat-RNA C had reversions of the mutations or second-site alterations. Sat-RNA C in plants inoculated with Group II mutants (and Groups II/III and III

described below) accumulated to the same level as wild type in all plants in which there was detectable sat-RNA (data not shown). Cm1, containing a single base alteration in the lower stem region (G<sub>328</sub> to A), accumulated sat-RNA C; however, all recovered sat-RNA C species sequenced had reverted to wild type (Table 2). Lack of biological activity of Cm1 was in correspondence with the low level of activity found *in vitro* (Song and Simon, 1995). The predicted structure of the promoter region of Cm1 differed somewhat from the wild-type structure (Fig. 2), with a 32% loss in stability of the hairpin and an additional residue in the single-stranded tail. Cm2, also categorized as a group II mutant, contained a 2-base alteration in the lower stem region (G<sub>328</sub> to A and C<sub>330</sub> to A) that was predicted to substantially alter the structure of the lower stem (Fig. 2). Only two of six plants inoculated with Cm2 contained sat-RNA C. Of the eight Cm2-derived species cloned from the two plants, five had reverted to wild type and three contained a second site alteration (U<sub>331</sub> to G; Cm2a). The predicted 3'-end structure for Cm2a was a hairpin that was 30% more stable than Cm2. The lack of biological activity *in vivo* of Cm2 was in correspondence with low levels of sat-RNA C synthesized during *in vitro* transcription using RdRp extracts (Song and Simon, 1995).

A third Group II mutant, Cm10, contained three consecutive base alterations in the upper stem (C<sub>332</sub>CC to GGG). Cm10 was predicted to have a longer 3' hairpin than wild type, with a shorter loop sequence and only a 3-base single-stranded tail (Fig. 2). Eight of 12 plants inoculated

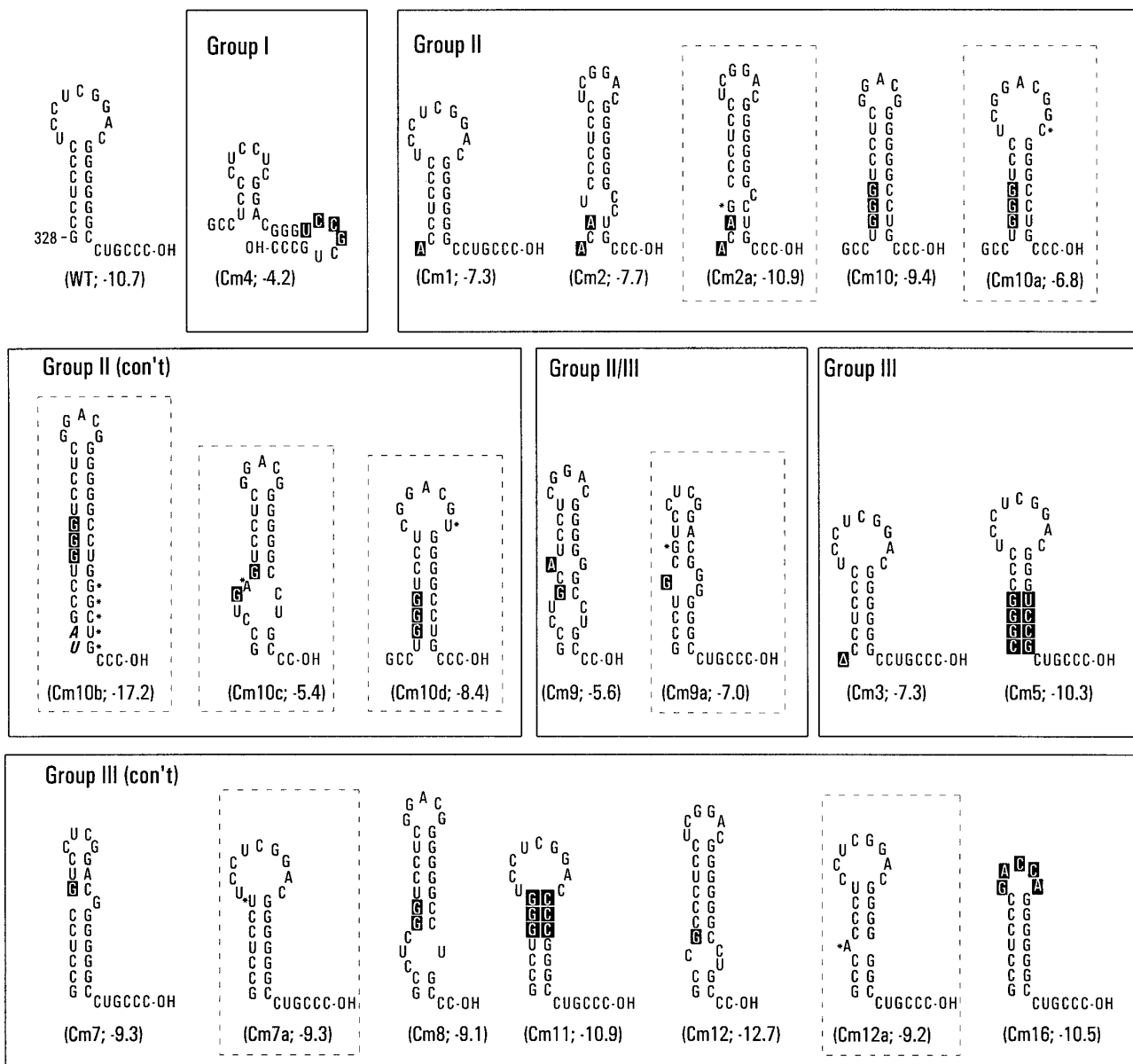


FIG. 2. Secondary structures of original sat-RNA C mutants and their derivatives. The names of the mutants and the stability (in kcal/mol) of the hairpins (as determined by the FOLD RNA program) are given below the structures. Equal lengths of RNA were used for the analyses with the exception of Cm10a. Original mutations are shown as white letters in black boxes. Sat-RNA C with second-site alterations accumulating in plants derived from the original mutants are boxed with a hatched line. Asterisks denote the positions of second-site mutations. Nucleotides in bold italics in Cm10b are upstream of the sat-RNA C wild-type hairpin and are predicted to participate in the mutant hairpin. Group I mutant produced no detectable accumulation of sat-RNA C species in plants. Group II mutants only accumulated following reversion to wild type or following reversions and/or second-site alterations. Group III mutants were able to accumulate in plants without secondary alterations. Group II/III mutant was intermediate between Group II and Group III (see text). WT, wild-type sat-RNA C.

with Cm10 accumulated sat-RNA C, although all recovered species contained various second-site alterations (Table 2). Nine of 13 clones derived from 2 plants contained a 5-base insertion (GGCUC) at position 353 that was predicted to further enlarge the hairpin of Cm10 and increase its stability by 1.7-fold while maintaining the 3-base single-stranded tail (Fig. 2). Since insertion of the same 5-base sequence was found in clones from more than a single plant, the insertion probably did not occur

by simple nucleotide misincorporation mediated by the RdRp, but may involve primer extension of fragments generated by polymerase abortive cycling, as found for repair of deletions at the 3' ends of TCV sat-RNAs (Carpenter and Simon, 1996a,b; Nagy *et al.*, 1997). The remaining minor species recovered from Cm10-inoculated plants had either a single second-site alteration or secondary alterations at the original mutation sites. The inability of Cm10 to accumulate *in vivo* in the absence of additional

mutations reflects prior *in vitro* analysis in which only very low levels of transcription products were observed using TCV RdRp extracts (Song and Simon, 1995).

Group III mutants (Cm3, 5, 7, 8, 11, 12, and 16) were able to accumulate in all inoculated plants and the original sequence was maintained in all or nearly all clones derived from the plants (Table 2). Cm5 (compensatory exchange of 4 bases in the lower stem at positions 328–331 and 347–350), Cm7 (G<sub>334</sub> to C), Cm8 (C<sub>333</sub>C to GG), Cm12 (U<sub>331</sub> to G), and Cm16 (U<sub>335</sub>CCUCGGAC to GACCA) were also good templates for *in vitro* transcription (Song and Simon, 1995). On the other hand, Cm3 (deletion of G<sub>328</sub>) and Cm11 (compensatory exchange of 3 bases in the upper stem at positions 332–334 and 344–346) were weak templates *in vitro* (Song and Simon, 1995), yet were able to accumulate without reversions or second-site mutations *in vivo*.

The ability of Cm3 to accumulate in all inoculated plants was surprising, as this mutant was similar to Group II mutant Cm1 in sequence and structure. In Cm1, G<sub>328</sub> was changed to an A residue, while in Cm3, deletion of G<sub>328</sub> results in A<sub>327</sub> being adjacent to the identical hairpin as predicted for Cm1 (Fig. 2). One explanation for why such similar sequences/structures at the 3' end (differing only in the presence of an additional A residue in Cm1 just 5' of the hairpin) led to such differences in accumulation *in vivo* is as follows: since sat-RNA C wild type and mutants can accumulate to similar levels in plants, a saturating level of sat-RNA may be reached even by an attenuating mutant. Both Cm1 and Cm3 may be able to accumulate *in vivo* and frequent nucleotide misincorporation by the error-prone RdRp is able to repair Cm1 (requiring a transition of A<sub>328</sub> to G), leading to the much more efficient wild-type template, which becomes the only species that was cloned from noninoculated leaves in Cm1-infected plants. In support of this possibility was the finding that all 12 plants inoculated with Cm1 accumulated sat-RNA C, suggesting that there was sufficient replication of Cm1 in all inoculated plants to allow for reversions to take place. This is in contrast with the other Group II mutants Cm2 and Cm10, of which sat-RNA C accumulated in only one-third or two-thirds of the inoculated plants, respectively. For Cm3-inoculated plants, the RdRp may be unable to "repair" the hairpin of Cm3 (requiring an insertion of a G residue at position 328 or a transition of A<sub>327</sub> to G) because of a low probability for insertion events or a requirement for maintenance of A<sub>327</sub> for accumulation *in vivo*. Therefore, with a lack of competition from any wild-type revertants, Cm3 would be the only sat-RNA accumulating in Cm3-infected plants. Alternatively, there may be structural differences between Cm1 and Cm3 due to the additional A residue in Cm1 that affect other functions required for viability leading to the accumulation of Cm3 but not Cm1.

Cm9, a weak template *in vitro* (Song and Simon, 1995), contained two nonadjacent base alterations (C<sub>332</sub> to G

and C<sub>334</sub> to A). While Cm9 progeny were able to accumulate in all plants, the majority species cloned from both plants was Cm9a, which contained a mutation of A<sub>334</sub> back to C (Table 2, Fig. 2). Since Cm9 was able to accumulate sufficiently in plants to be readily cloned, but not as the majority species cloned from the plants, Cm9 was placed in an intermediate group (Group II/III) between Group II and Group III. Cm9a was predicted to contain a hairpin that was 1.3-fold more stable than Cm9, with a reformed 6-base single-stranded tail.

The only mutant with a promoter structure that differed significantly from wild-type, Cm4, did not produce detectable sat-RNA C in plants, possibly because of the inability of the TCV RdRp to recognize the terminal structure as a promoter element. All other mutants, including those of Group II, contained at least partially active promoters, since transcription by the RdRp is required for generation of reversions or second-site alterations. The predicted structures for Group II mutants were not noticeably distinguished from the predicted structures for Group III mutants, Cm9, and derivatives of Group II mutants that were the majority cloned species in plants (Fig. 2). This lack of defined features for less biologically active promoters makes determination of what is required to produce an efficient promoter difficult at present. However, a direct correlation was found between higher stability of the hairpin and the majority species cloned from plants inoculated with mutants Cm1, Cm9, Cm10, and Cm12. For the two remaining mutants, the majority species cloned from Cm2 was the wild-type revertant, which has a hairpin slightly less stable than the second-site mutant Cm2a. Cm7a, constituting one of eight clones from Cm7-inoculated plants, contained a hairpin of stability equal to that of Cm7. These results suggest that stability of the hairpin likely plays a role, but not the only role, in establishing what constitutes an efficient *in vivo* promoter. The current results also suggest that sat-RNA C containing less than a 6-base single-stranded tail and a wide variety of hairpin structures are recognized by the TCV RdRp *in vivo*, including hairpins only 50% as stable as wild type. Our results confirm previous *in vitro* findings that the sequence of the loop region is unimportant for promoter recognition. However, the conclusion from the *in vitro* study that the sequence of the upper stem was important, based on the low activity of Cm11 (compensatory exchanges of 3 bp in the upper stem), must be reassessed, since Cm11 was a Group III mutant for *in vivo* accumulation. The *in vivo* results for Cm11 suggest that the positioning of specific bases in the upper stem is not required to produce an active minus-strand promoter.

#### Analysis of effects of secondary mutations on efficiency of *in vitro* transcription using partially purified TCV RdRp

As described above, all mutants that were active templates for previous transcription studies *in vitro* (Cm5,

Cm7, Cm8, Cm12, and Cm16) were able to accumulate without second-site alterations *in vivo*. The converse, however, was not true, as several mutants that were poor templates *in vitro* (Cm3, Cm11, and Cm9) were also able to accumulate *in vivo*. To determine if species containing secondary alterations derived from Group II mutants had improved template activity for minus-strand synthesis *in vitro*, Cm2a, Cm10a, Cm10b, and Cm10c derived from Cm2 and Cm10 were tested for activity in an *in vitro* RdRp assay using partially purified TCV RdRp (Song and Simon, 1994). Cm7a and Cm9a, derived from Cm7 and Cm9, were also tested.

As shown in Fig. 3A and as described previously (Song and Simon, 1995), transcription of plus strands of sat-RNA C resulted in full-length minus strands (indicated by an arrow, Fig. 3A) and an unidentified doublet product that migrated slightly faster than full length and was present at a ratio with the full-length product that was dependent on the template (data not shown). Since mutations within the 3'-terminal hairpin led to different levels of the doublet product, it is more likely that these products are the result of internal initiation within the 3'-end hairpin and not premature termination. For this reason, only the full-length products from three independent experiments were quantified in the histogram shown in Fig. 3B. Cm2a, containing a second-site alteration of U<sub>331</sub> to G, produced more than 2-fold more full-length minus strands compared with Cm2 (Fig. 3), while the major species recovered from plants inoculated with Cm10 (Cm10b) produced almost 8-fold more minus strands compared with Cm10, and was a much more active template than the minor species derived from Cm10 (Cm10a and Cm10c). Cm10b, which contained the most stable hairpin of all sat-RNA C accumulating *in vivo*, was also 4-fold more active than wild-type as a template for *in vitro* transcription. Cm9a, the major species cloned from plants also accumulating Cm9, contained a more stable hairpin than Cm9 and was about 200-fold more active in synthesis of the full-length product when compared with Cm9. The minor species cloned from Cm7-inoculated plants, Cm7a, was a poor template *in vitro*, although the stability of the hairpin was not predicted to change.

Therefore, with the exception of Cm2 and Cm7, the major species cloned from plants containing more than one related species had the most stable hairpin and the highest activity for *in vitro* transcription. However, results of these studies combined with previous *in vitro* transcription studies (Song and Simon, 1995) suggest that caution should be used when determining promoter requirements based solely on *in vitro* analyses. While all templates that were active in the *in vitro* reactions were also active templates *in vivo*, several templates with weak activity *in vitro* were nevertheless able to accumulate *in vivo* in the absence of secondary alterations. One possibility for differences in *in vitro* and *in vivo* results is if some templates generated for *in vitro* RdRp transcription did not fold into

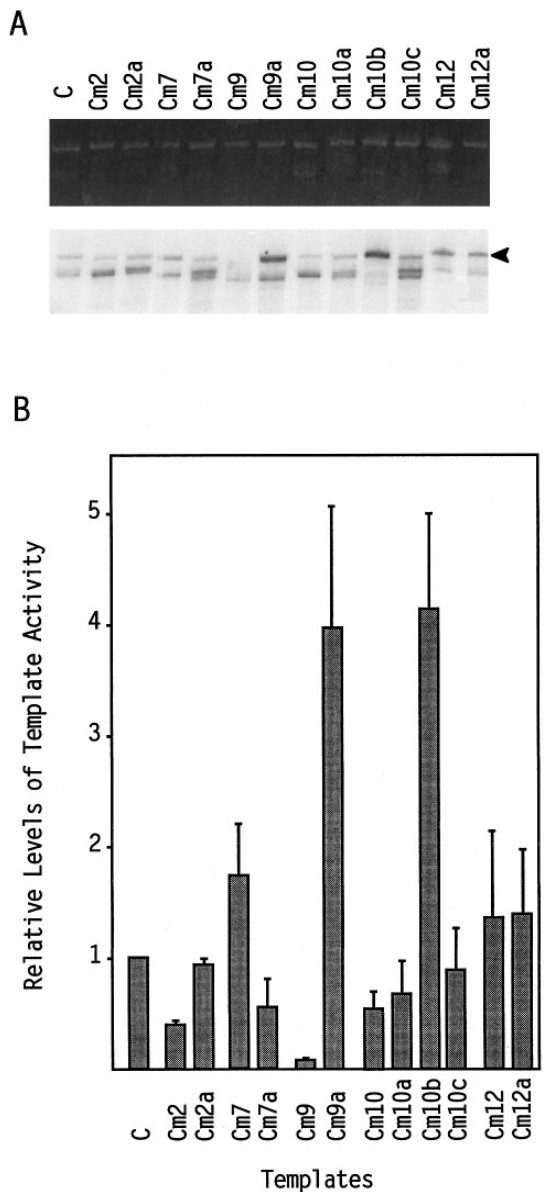


FIG. 3. Activity of plus-strand sat-RNA C templates in *in vitro* transcription using partially purified TCV RdRp extracts. (A) Example of denaturing gel analysis of <sup>32</sup>P-labeled minus-strand products synthesized using partially purified RdRp. The ethidium bromide-stained gel showing the relative levels of plus-strand templates is shown above the autoradiogram of the products. Arrow denotes the product migrating to the position of full-length templates. The identity of the faster migrating doublet products is not known but the ratio between the full-length and the doublet products was similar for each experiment. The amounts of truncated templates visible in the ethidium-stained gels varied between experiments and the amounts did not correlate with the levels of the doublet products. (B) Quantitative analysis of template activity. Autoradiograms and photographs of ethidium bromide-stained gels from three independent experiments (including template synthesis) were scanned with an imaging densitometer. Sat-RNA levels were normalized to the amounts of template RNA. Error bars denote standard error.

a structure required for activity *in vitro* but did assume the proper conformation *in vivo*. A second possibility is that our partially purified RdRp extract does not completely reflect the enzyme produced *in vivo*.

## ACKNOWLEDGMENTS

We thank Drs. Peter D. Nagy and Clifford D. Carpenter for helpful discussions and critical review of the manuscript. The work was supported by National Science Foundation Grants MCB-9419303 and MCB-9630191 to A.E.S.

## REFERENCES

- Ball, L. A., and Li, Y. (1993). Cis-acting requirement for the replication of flock house virus RNA 2. *J. Virol.* **67**, 3544–3551.
- Boccard, F., and Baulcombe, D. (1993). Mutational analysis of cis-acting sequences and gene function in RNA 3 of cucumber mosaic virus. *Virology* **193**, 563–578.
- Carpenter, C. D., Oh, J.-W., Zhang, C., and Simon, A. E. (1995). Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. *J. Mol. Biol.* **245**, 608–622.
- Carpenter, C. D., and Simon, A. E. (1996a). *In vivo* repair of deletions at the 3' end of a TCV satellite RNA in vivo may involve two abortive synthesis and priming events. *Virology* **226**, 153–160.
- Carpenter, C. D., and Simon, A. E. (1996b). *In vivo* restoration of biologically active 3' ends of virus-associated RNAs by non-homologous RNA recombination and replacement of a terminal motif. *J. Virol.* **70**, 478–486.
- Chang, Y. C., Borja, M., Scholthof, H. B., Jackson, A. O., and Morris, T. J. (1995). Host effects and sequences essential for accumulation of defective interfering RNAs of cucumber necrosis and tomato bushy stunt tobusviruses. *Virology* **210**, 41–53.
- Dreher, T. W., and Hall, T. C. (1988). Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for minus strand promoter activity. *J. Mol. Biol.* **201**, 31–40.
- Duggal, R., Lahser, F. C., and Hall, T. C. (1994). Cis-acting sequences in the replication of plant viruses with plus-sense RNA genomes. *Annu. Rev. Phytopathol.* **32**, 287–309.
- French, R., and Ahlquist, P. (1987). Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA 3. *J. Virol.* **61**, 1457–1465.
- Guilley, H., Carrington, J. C., Balazs, E., Jonard, G., Richards, K., and Morris, T. J. (1985). Nucleotide sequence and genome organization of carnation mottle virus RNA. *Nucleic Acids Res.* **13**, 6663–6677.
- Havelda, Z., and Burgyn, J. (1995). 3' terminal putative stem-loop structure required for the accumulation of cymbidium ringspot viral RNA. *Virology* **214**, 269–272.
- Havelda, Z., Dalmay, T., and Burgyn, J. (1995). Localization of cis-acting sequences essential for cymbidium ringspot tobusvirus defective interfering RNA replication. *J. Gen. Virol.* **76**, 2311–2316.
- Heaton, L. A., Carrington, J. C., and Morris, T. J. (1989). Turnip crinkle virus infection from RNA synthesized *in vitro*. *Virology* **170**, 214–218.
- Jacobson, S. J., Konnings, D. A. M., and Sarnow, P. (1993). Biochemical and genetic evidence for a pseudoknot structure at the 3' terminus of the poliovirus RNA genome and its role in viral RNA amplification. *J. Virol.* **67**, 2961–2971.
- Jupin, I., Richards, K., Jonard, G., Guilley, H., and Pleij, C. W. A. (1990). Mapping sequences required for productive replication of beet necrotic yellow vein virus RNA 3. *Virology* **178**, 273–280.
- Kim, Y. N., Jeong, Y. S., and Makino, S. (1993). Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication. *Virology* **197**, 53–63.
- Levis, R., Weiss, B. G., Tsiang, M., Huang, H., and Schlesinger, S. (1986). Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequence essential for replication and packaging. *Cell* **44**, 137–145.
- Li, X. H., Heaton, L. A., Morris, T. J., and Simon, A. E. (1989). Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated *de novo*. *Proc. Natl. Acad. Sci. USA* **86**, 9173–9177.
- Li, X. H., and Simon, A. E. (1991). *In vivo* accumulation of a turnip crinkle virus DI RNA is affected by alterations in size and sequence. *J. Virol.* **65**, 4582–4590.
- Lin, Y.-J., and Lai, M. M. C. (1993). Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J. Virol.* **67**, 6110–6118.
- Lin, Y.-J., Liao, C.-L., and Lai, M. M. C. (1994). Identification of the cis-acting signal for minus-strand RNA synthesis of a murine coronavirus: Implication for the role of minus-strand RNA in RNA replication and transcription. *J. Virol.* **68**, 8131–8140.
- Nagy, P. D., Carpenter, C. D., and Simon, A. E. (1997). A novel 3' end repair mechanism in an RNA virus. *Proc. Natl. Acad. Sci. USA* **94**, 1113–1118.
- Nutter, R. C., Scheets, K., Panganiban, L. C., and Lommel, S. A. (1989). The complete nucleotide sequence of maize chlorotic mottle virus genome. *Nucleic Acids Res.* **17**, 3163–3177.
- Oh, J.-W., Kong, Q., Song, C., Carpenter, C. D., and Simon, A. E. (1995). Open reading frames of turnip crinkle virus involved in satellite symptom expression and incompatibility with *Arabidopsis thaliana* ecotype Dijon. *Mol. Plant-Microbe Interact.* **8**, 979–987.
- Pilipenko, E. V., Maslova, S., Sinyakov, A. N., and Agol, V. I. (1992). Towards identification of cis-acting elements involved in the replication of enterovirus and rhinovirus RNAs: A proposal for the existence of tRNA-like terminal structures. *Nucleic Acids Res.* **20**, 1739–1745.
- Rohli, J. B., Holness, C. L., Lomonosoff, G. P., and Maule, A. J. (1993). 3'-terminal nucleotide sequences important for the accumulation of cowpea mosaic virus M-RNA. *Virology* **193**, 672–679.
- Rohli, J. B., Moon, D. H., Evans, D. J., and Almond, J. W. (1995). The 3' untranslated region of picornavirus RNA: Features required for efficient genome replication. *J. Virol.* **69**, 7835–7844.
- Roossinck, M. J., Sleat, D. E., and Palukaitis, P. (1992). Satellite RNAs of plant viruses: Structures and biological effects. *Microbiol. Rev.* **56**, 265–279.
- Roux, L., Simon, A. E., and Holland, J. J. (1991). Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. *Adv. Virus Res.* **40**, 181–211.
- Skotnicki, M. L., Mackenzie, A. M., Torronen, M., and Gibbs, A. J. (1993). The genomic sequence of cardamine chlorotic fleck carmovirus. *J. Gen. Virol.* **74**, 1933–1937.
- 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**, 3423–3438.
- Song, C., and Simon, A. E. (1994). RNA-dependent RNA polymerase from plants infected with turnip crinkle virus can transcribe (+)- and (-)-strands of virus-associated RNAs. *Proc. Natl. Acad. Sci. USA* **91**, 8792–8796.
- Song, C., and Simon, A. E. (1995). Requirement of a 3'-terminal stem-loop *in vitro* transcription by an RNA-dependent RNA polymerase. *J. Mol. Biol.* **254**, 6–14.
- Takamatsu, N., Watanabe, Y., Meshi, T., and Okada, Y. (1990). Mutational analysis of the pseudoknot region in the 3' noncoding region of tobacco mosaic virus RNA. *J. Virol.* **64**, 3686–3693.
- Thomson, M., and Dimmock, N. J. (1994). Common sequence elements in structurally unrelated genomes of defective interfering Semliki Forest virus. *Virology* **199**, 354–365.
- Tsai, C.-H., and Dreher, T. W. (1992). Second-site suppressor mutations assist in studying the function of the 3' noncoding region of turnip yellow mosaic virus RNA. *J. Virol.* **66**, 5190–5199.
- Tsiang, M., Weiss, B. G., and Schlesinger, S. (1988). Effects of 5'-terminal modifications on the biological activity of defective interfering RNAs of Sindbis virus. *J. Virol.* **62**, 47–53.