J. Mol. Biol. (1995) **254**, 6–14

JMB

Requirement of a 3'-Terminal Stem-loop in *in Vitro* Transcription by an RNA-dependent RNA Polymerase

Chuanzheng Song and Anne E. Simon*

Department of Biochemistry and Molecular Biology and Program in Molecular and Cellular Biology, University of Massachusetts at Amherst Amherst, MA 01003, USA

**Corresponding author*

Partially purified RNA-dependent RNA polymerase (RdRp) isolated from plants infected with turnip crinkle virus (TCV) is capable of template-dependent synthesis of TCV-associated RNAs. To determine the *cis*-sequences required for the synthesis of TCV satellite (sat-) RNA C (–) strands *in vitro*, templates containing interior deletions were subjected to transcription using RdRp-active fractions. Results indicated that the promoter for (–)-strand synthesis was contained within the 3'-terminal 29 bases of the (+)-strand. Structural probing by enzymatic digestion and chemical modification revealed the presence of a hairpin structure within this terminal region. Compensatory exchanges of four bases in the lower stem or alterations in the sequence and size of the loop region did not affect *in vitro* transcription, implying that the primary sequence in the loop and lower part of the stem is not important for interaction with the viral RdRp. However, single mutations in the base of the stem or double mutations in the upper stem strongly reduced template activity *in vitro*, suggesting that the stability of the hairpin is an important functional consideration. Relocation of the 3'-terminal 37 bases containing this stem-loop to an inactive template RNA rendered the resultant hybrid RNA competent for *in vitro* transcription by RdRp activity, suggesting that the promoter for (–)-strand synthesis *in vitro* is completely contained within the 3'-terminal region.

Keywords: RNA virus; RNA-dependent RNA polymerase; RNA replication; satellite RNA; turnip crinkle virus

Introduction

Replication of single-stranded, positive sense RNA viruses involves specific interactions between viral replicases, composed of virus-encoded RNA-dependent RNA polymerases (RdRp) and possibly host factors, and their cognate RNA templates. The first step in the replication of (+)-strand RNA viruses most likely involves the recognition by the replicase of specific promoter sequences and/or structures within 19 to 200 bases of the 3' ends of the (+)-strand template (Levis *et al*., 1986; French & Ahlquist, 1987; Pacha *et al*., 1990; Ball & Li, 1993; Boccard & Baulcombe, 1993; Kim *et al*., 1993; Duggal *et al*., 1994; Lin *et al*., 1994), followed by polymerization of complementary (–)-strands. The 3' ends of many (+)-strand viruses contain defined sequence or structural elements, such as tRNA-like structures, pseudoknots or poly(A) tracts that have been identified by deletion and base alteration analysis as *cis*-acting factors required for replication and/or replicase binding (Jupin *et al*., 1990; Jacobson *et al*., 1993; Duggal *et al*., 1994; Cui & Porter, 1995). In addition, internal regions necessary for *in vivo* accumulation have been identified for brome mosaic virus (French & Ahlquist, 1987), flock house virus (Li & Ball, 1993) and mouse hepatitis virus (Lin & Lai, 1993).

Since *cis*-sequences important for virus accumulation *in vivo* can affect (–)- or (+)-strand synthesis as well as translation (Gallie & Walbot, 1990; Leathers *et al*., 1993), packaging (Levis *et al*., 1986), or movement (Petty *et al*., 1990), identifying elements specifically involved in (–)-strand synthesis has involved quantification of (–)-strands *in vivo* by RNase protection (Lin & Lai, 1993) or transcription *in vitro* using partially purified replicase (Miller *et al*., 1986; Dreher & Hall., 1988;

Abbreviations used: RdRp, RNA-dependent RNA polymerases; DI, defective interfering; sat, satellite; TCV, turnip crinkle virus; MDV, midivariant; MCMV, maize chlorotic mottle virus; CarMV, carnation mottle virus; CMCT, carbodiimide; DEPC, diethylpyrocarbonate; DMS, dimethylsulfate; PCR, polymerase chain reaction; CCFV, cardamine chlorotic fleck virus.

0022–2836/95/460006–09 $12.00/0

van der Kuyl *et al*., 1990). Subviral RNAs amplified by association with a specific helper virus, such as defective interfering (DI) RNAs or satellite (sat-) RNAs, have also been useful in determining sequence requirements for replication (Levis *et al*., 1986; Li & Simon, 1991; Lin & Lai, 1993). These subviral RNAs, many of which are considerably smaller than viral genomic RNAs and encode no translation products, must still contain functional *cis*-acting sequences and structures that can be recognized by the helper virus replicase.

Turnip crinkle virus (TCV), a member of the carmovirus group, is a positive-strand virus with a single RNA genome of 4054 bases (Heaton *et al*., 1989; Hacker *et al*., 1992). TCV is naturally associated with sat-RNAs and DI RNAs that range in size from 194 to 356 bases, do not contain functional open reading frames, and require co-inoculation of TCV genomic RNA for accumulation in plants (Simon & Howell, 1986; Li *et al*., 1989). The 3' untranslated region of TCV genomic RNA and the 3'-ends of the subviral RNAs do not form a tRNA-like structure or contain a poly(A) tail. The best characterized TCV sat-RNA is sat-RNA C, a 356 base RNA composed of the nearly complete sequence of a second TCV sat-RNA (sat-RNA D) at the 5'-end joined to two regions of TCV genomic RNA at the 3'-end (Simon & Howell, 1986; Figure 1A). Sat-RNA C shares 90% identity with the corresponding 3'-terminal sequence of TCV genomic RNA, and accumulates to very high levels (equivalent to 5 S ribosomal RNA) in infected cells. Therefore, sat-RNA C should be a useful template in the analysis of *cis*-sequences important for the synthesis of (–)-strands by the TCV replicase.

An *in vitro* RNA transcription system has been established that uses partially purified RdRp activity from TCV-infected turnip plants to synthesize complementary strands of either (+)- or (–)-strand sat-RNA C template (Song & Simon, 1994). Plus-strand sat-RNA C containing a deletion of the 3'-terminal 100 bases was not an active template for synthesis of (–)-strands, suggesting that this region contained a sequence indispensable for its transcription *in vitro*. We now report that the 3'-terminal 29 bases of sat-RNA C, containing a small hairpin and conserved six base single-stranded region, can function as an independent promoter for (–)-strand synthesis *in vitro*.

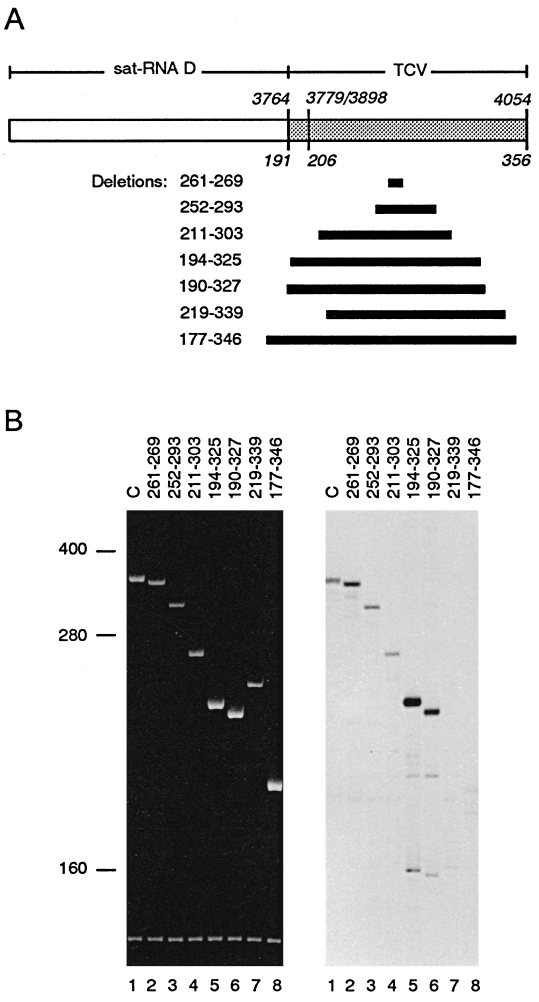

Figure 1. Effect of internal deletions in (+)-strand sat-RNA C on transcription *in vitro* by peak II RdRp activity. A, Schematic representation of wild-type (+)-strand sat-RNA C and location of the deletions. Portons of sat-RNA C (rectangle) sharing homology with sat-RNA D and two regions of TCV are indicated. Upper numbers reflect positions in the TCV genomic RNA of the related sequence. Lower numbers denote positions in sat-RNA C. Numbers to the left of the deletions denote their positional location. B, Denaturing gel analysis of ^{32}P-labeled (–)-strand products synthesized using peak II RdRp activity and the templates shown in A. Numbers above the lanes refer to the location of the deletions. The ethidium bromide-stained gel showing the migration positions and relative levels of (+)-strand templates is shown to the left of the autoradiogram. Positions of RNA markers are indicated (in bases). C, Wild-type sat-RNA C template.

Results

Localization of the promoter for (–)-strand synthesis on (+)-strand sat-RNA C by deletion analysis

We previously reported that detergent-solubilized membrane fractions from TCV-infected turnip contained three RdRp activities separable by size-exclusion chromatography (Song & Simon, 1994). Peak I, which contained endogenous TCV genomic RNA, and peak II were able to accept exogenous TCV-associated RNAs as templates, but could not synthesize full-length complementary strands of non-specific RNAs or sat-RNA C (+)-strand templates lacking the 3'-terminal 100 bases. Peak III contained non-specific RdRp activity that accepted all RNA templates assayed. The products produced by the peak II RdRp activity using (+)-strand templates were double-stranded and no evidence of cyclic replication was observed.

To better localize the sequences on the (+)-strands of sat-RNA C required for synthesis of complementary strands, deletions were generated in both directions from the *Spe*I site located 90 bases upstream from the 3'-end in a full-length cDNA of sat-RNA C (Figure 1A). Transcription of sat-RNA C (+)-strands from an upstream bacteriophage T7 DNA-dependent RNA polymerase promoter produced transcripts with the wild-type 5'- and 3'-ends (Song & Simon, 1994). The activities of the deletion-containing transcripts in generating full-length (–)-strand products were tested by combining the templates with peak II RdRp extracts and [α-^{32}P]UTP, followed by analysis of the products by denaturing polyacrylamide gel electrophoresis and autoradiography (Figure 1B). RNAs with deletions encompassing positions 190 to 327 were active templates for the RdRp activity, generating template-length radiolabeled products (Figure 1B, lanes 2 to 6). However, deletion of positions 219 to 339 or 177 to 346 abolished detectable full-length complementary strand synthesis (Figure 1B, lanes 7 and 8). Since the inactive sat-RNA C template with a deletion of positions 219 to 339 has a shorter 5'-directional deletion than the functional template containing a deletion of positions 190 to 327 (Figure 1B, lanes 6 and 7), the loss of template efficiency is likely to be due to the removal of nucleotides between positions 328 and 339. These results suggested that sequences within the 3'-terminal 29 bases (positions 328 to 356) of (+)-strand sat-RNA C are important for template activity.

Probing the secondary structure of the 3'-end of (+)-strand sat-RNA C

To determine the secondary structure of the 3'-end of (+)-strand sat-RNA C, chemical and enzymatic modifications were performed, followed by reverse transcriptase-mediated extension from oligonucleotide primers. The modification pattern of the 3'-end of sat-RNA C was obtained using (+)-strand sat-RNA C that contained 220 bases of vector-derived sequence at the 3'-end (C + 220); extension on a primer hybridizing in the vector sequence 39 bases downstream from the sat-RNA C 3'-end could determine the modification pattern of bases at the sat-RNA 3'-end. Since *in vitro* transcription of C + 220 (and other sat-RNA C templates containing various lengths of vector 3'-end sequences) by peak II RdRp activity gave rise to a major product the size of wild-type sat-RNA C (Song & Simon, 1994), the RdRp activity was able to recognize the correct 3'-end within an internal location. The assumption was therefore made that the wild-type 3'-end of sat-RNA C located internally within C + 220 should have the same structural features for the initiation of (–)-strand synthesis as wild-type sat-RNA C. To determine if sat-RNA C and C + 220 templates had similar modification patterns, the extension products following chemical and enzymatic modifications

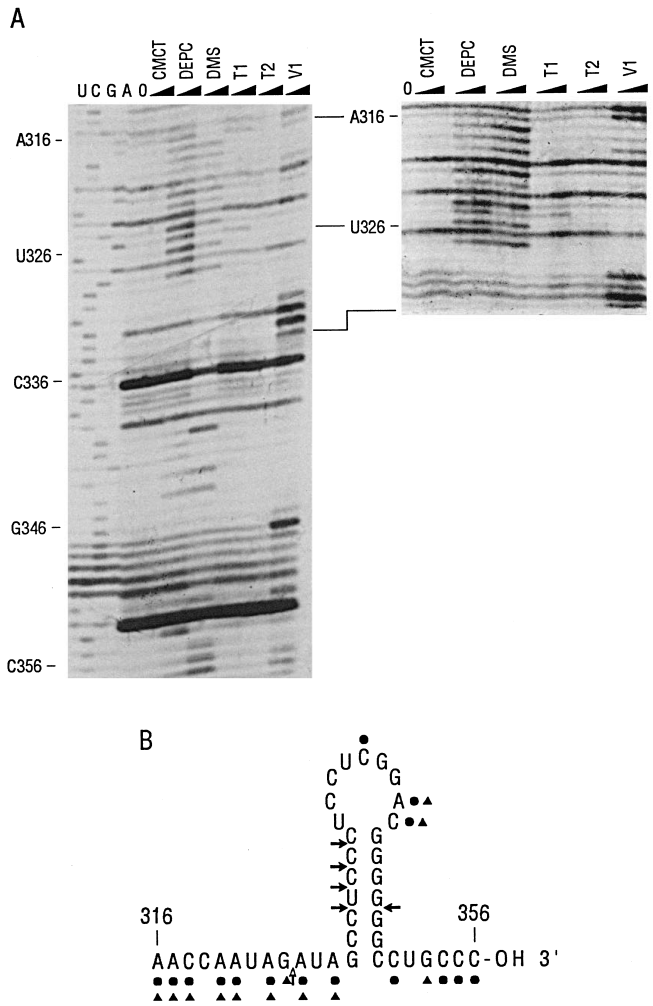

Figure 2. Biochemical analysis of the structure at the 3'-end of (+)-strand sat-RNA C. A, Comparison of modification patterns of the 3'-ends of (+)-strand sat-RNA C (right) and C + 220 (left). Purified wild-type sat-RNA C or C + 220 were treated with chemicals (CMCT, DEPC, and DMS), enzymes (RNase T_1, T1; RNase T_2, T2; and RNase V_1, V1) or untreated (0) followed by primer extension using oligonucleotides specific for the 3'-terminal 19 bases of sat-RNA C (right) or the M13/pUC (–40) sequencing primer, which is complementary to sequence 39 bases downstream from the 3'-end of sat-RNA C (left). Triangles denote increasing time of the modification reactions. The RNA sequencing ladder is shown to the left of the C + 220 modification reactions. The positions of nucleotides in sat-RNA C are indicated. B, *In vitro* secondary structure model of (+)-strand sat-RNA C. Modifications by DEPC, DMS and CMCT are represented by filled circles, triangles and squares, respectively. Cleavages by RNases T_1 and V_1 are denoted by the open arrow and filled arrows, respectively.

were compared between positions 316 and 334, the region nearest to the 3'-end that was accessible by both sat-RNA C– and C + 220-specific primers (Figure 2A). Nucleotides in this region reacted in an identical fashion with chemicals and enzymes specific for single-stranded and double-stranded bases, suggesting that the addition of 220 bases to

the 3'-end of sat-RNA C did not affect the sat-RNA C secondary structure near the junction region.

Based on the modification data, a secondary structure model for the 3'-end was generated using the M-FOLD program in the University of Wisconsin GCG package (Figure 2B). The region was predicted to contain a stable hairpin surrounded by single-stranded sequences. As described above, templates containing deletions towards the 3'-end that terminated at or before position 327 were active for *in vitro* transcription. Position 327 is located precisely at the base of the hairpin. Deletions that extended 3' of nucleotide 327 eliminated template activity, suggesting that the 3'-terminal hairpin is at least one important component in promoting (–)-strand synthesis.

Effects of mutations in the 3' hairpin on (–)-strand synthesis *in vitro*

To test if the 3' hairpin is involved in initiating (–)-strand synthesis, (+)-strand sat-RNA C templates with mutations in the loop region (UCCUCGGAC), the upper stem (CCC/GGG), lower stem (GCC/CGG), and the midstem (U/G) were assayed for the efficiency of (–)-strand synthesis *in vitro* (Figure 3). Peak II RdRp activity was able to synthesize at least wild-type levels of full-length product, using templates where the nine-base loop was changed to random five-base (Cm13 and Cm16) or six-base (Cm14 and Cm15) sequences (Figure 3B, lanes 15 to 18). Single-base transversions at position 334 would disrupt the base-pair adjacent to the loop, possibly expanding the loop by two bases (Cm6, C to A; Cm7, C to G). Templates containing these alterations still resulted in the production of high levels of full-length complementary RNA. In addition, Cm6 produced increased levels of an unidentified species that migrated to a position about 10 to 15 bases shorter than full-length (Figure 3B, lane 7). Templates containing alterations of two neighboring bases in the upper stem adjacent to the loop (Cm8; CC to GG, positions 333 and 334) resulted in reduced levels of full-length product compared with Cm7, which contained one of the Cm8 mutations (Figure 3B, compare lanes 8 and 9). Templates containing alterations of two non-adjacent bases in the upper stem (Cm9; CCC to GCA, positions 332 and 334) or three consecutive bases in the upper stem (Cm10; CCC to GGG, positions 332 to 334) did not generate detectable full-length product (Figure 3B, lanes 10 and 11).

A single transversion in the midstem (Cm12; U to G, position 331) did not affect the activity of the template, suggesting that a symmetrical bulge in the stem does not eliminate activity. However, alteration or deletion of the G residue at the base of the lower stem (position 328; Cm1 and Cm3) resulted in a decrease in full-length product synthesized and an increase in synthesis of the faster-migrating species (Figure 3B, lanes 2 and 4). Changing the lower 3' side of the stem from CGGG to GCCU (positions 347 to 350, Cm4) also eliminated the synthesis of detectable levels of full-length product (Figure 3B, lane 5).

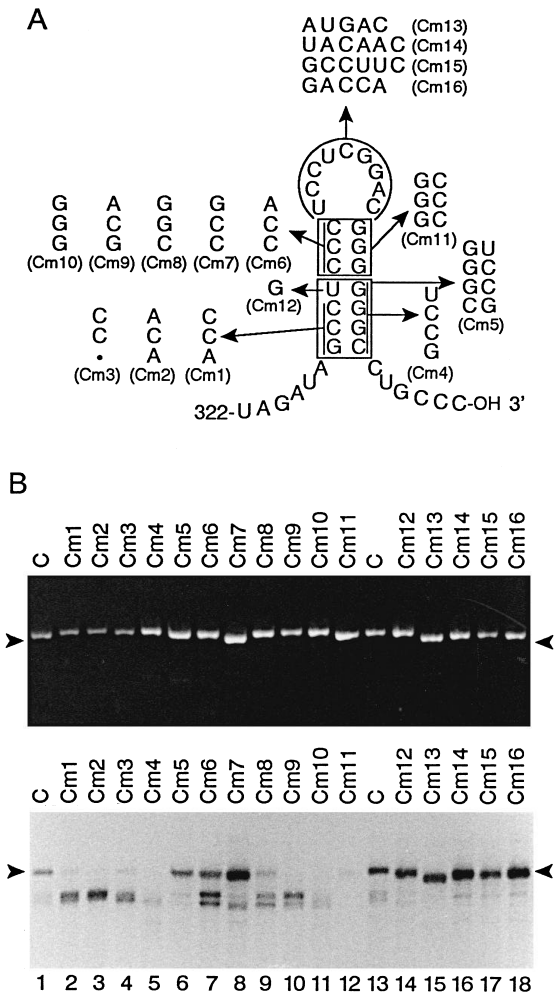

Figure 3. Efficiency of (+)-strand sat-RNA C templates containing mutations in the 3'-end stem-loop on transcription *in vitro* by peak II RdRp activity. A, The 3'-end stem-loop structure and location of the mutations. The names of the templates containing the mutations are in parentheses. B, Denaturing gel analysis of ^{32}P-labeled (–)-strand products synthesized using peak II RdRp activity and the templates shown in A. The ethidium bromide-stained gel showing the migration positions and relative levels of (+)-strand templates is shown above the autoradiogram. The templates used are indicated above each lane. The wild-type sat-RNA C template shown in lane 1 was assayed at the same time as the mutant templates in lanes 2 to 12. The wild-type sat-RNA C template (C) in lane 13 was assayed with the mutant templates in lanes 14 to 18. Arrowheads denote the position of full-length sat-RNA C. The identity of the faster migrating species is not known. This unknown species can also be found migrating slightly below the full-length product in Figure 1B, lane 1.

To test whether the primary sequence and/or secondary structure of the hairpin is important for template activity, two sets of compensatory mutations were constructed. Exchanging the CGGG and GCCU in the lower stem (Cm5) gave wild-type

levels of full-length product, compared with the substantially reduced levels produced by the four base alteration on the 3' side of the stem (Figure 3B, compare lanes 1, 5 and 6). The second compensatory exchange involved the three C·G base-pairs in the upper stem (positions 332 to 334 and 344 to 346; Cm11). Unlike the compensatory mutations in the lower stem, this exchange did not result in the synthesis of wild-type levels of full-length product (Figure 3B, lane 12). However, the level of product produced using Cm11 was greater than that detected using template containing only the three base alteration on one side of the stem (Cm10). Lack of high level product synthesis due to this compensatory exchange in the upper stem could be due to a requirement for specific residues on one face of the stem helix, or it is possible that the compensatory exchanges did not restore the hairpin. While the 3' hairpin in the wild-type was present in five of 12 structures predicted using the M-FOLD secondary structure analysis program, none of the seven structures produced using Cm11 contained this hairpin structure. In addition, chemical and enzymatic modification analysis using C + 220 containing the Cm11 compensatory mutations indicated very different susceptibility to single-stranded and double-stranded specific reagents compared with the wild-type C + 220 template (data not shown). Taken together, these results suggest that the compensatory alterations in Cm11 may not have restored the stem-loop structure, and therefore it is not clear whether the change in primary sequence or altered secondary structure is responsible for the poor activity of the Cm11 template.

Relocation of the 3'-terminal promoter region of (+)-strand sat-RNA C

A series of chimeric RNA templates were prepared to test if the 3'-terminal 37 bases of (+)-strand sat-RNA C comprising the stem-loop, the single-stranded six bases at the 3' terminus and eight bases 5' of the hairpin was sufficient to function independently as a promoter. These RNAs contained the MDV (midivariant) RNA (221 bases) associated with the Qβ bacteriophage at the 5' end joined to the 3' terminal 37 bases of either wild-type (+)-strand sat-RNA C or mutant Cm10 (containing three C to G transversions at positions 332 to 334 that greatly reduced the level of (–)-strand sat-RNA C produced; Figure 3B, lane 11) in forward and reverse orientations (Figure 4A). MDV RNA alone is not an active template for peak II RdRp activity (Song & Simon, 1994; Figure 4B, lane 2). However, MDV ligated to either one or two copies of the sat-RNA C wild-type 3'-terminal 37 bases in the forward (correct) orientation were efficient templates for full-length (–)-strand synthesis (Figure 4B, lanes 3 and 7). MDV ligated to the Cm10 3'-terminal 37 bases was a less efficient template for RdRp activity (Figure 4B, lane 4). Chimeric templates containing either the wild-type or mutant sat-RNA C 3'-terminal sequence in the reverse orientation did not result in detectable full-length (–)-strand

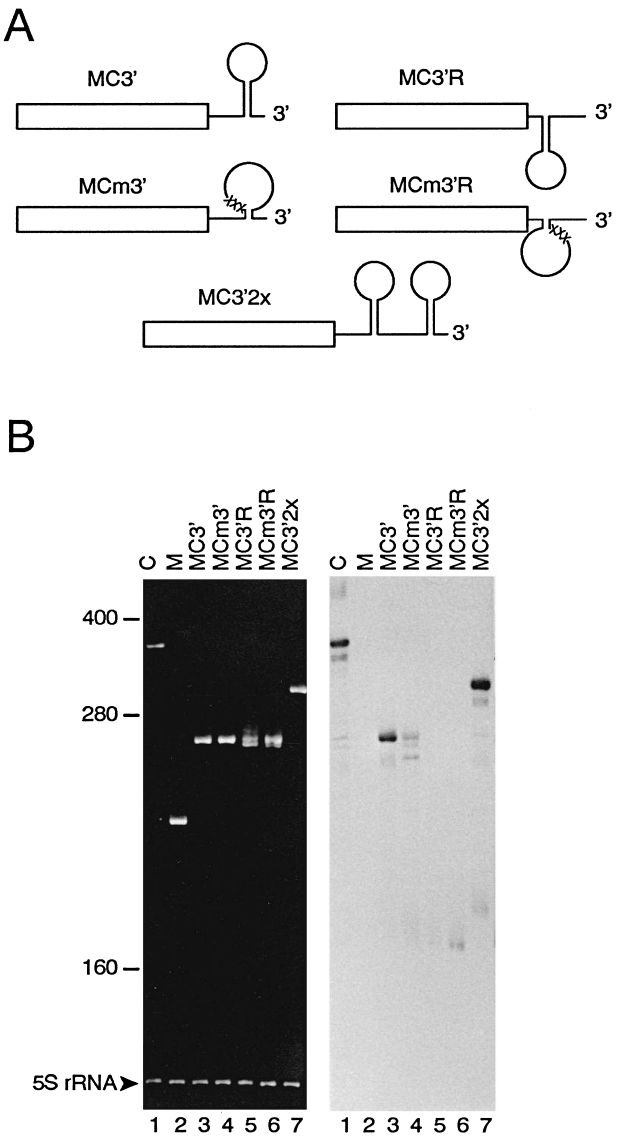

Figure 4. Template efficiencies of hybrid RNAs derived from MDV RNA and the 3'-end of sat-RNA C. A, Schematic representation of the hybrid molecules. Open bars denote the complete MDV RNA sequence at the 5' portion of each molecule. Hairpins with long stems are from the 3'-end of wild-type sat-RNA C and those with short stems are from the Cm10 mutant (see Figure 3A; X represents the three mutated bases in Cm10 RNA). Upright stem-loops indicate 3'-end sequences that are in the wild-type (forward) orientation, while the downward stem-loops denote 3'-end sequences which are in the reverse orientation. B, Denaturing gel analysis of ^{32}P-labeled (–)-strand products synthesized *in vitro* using peak II RdRp activity. The ethidium bromide-stained gel showing the migration positions and relative levels of (+)-strand templates is shown to the left of the autoradiogram. Positions of RNA markers (in bases) and yeast 5 S rRNA (added as a carrier) are indicated. Names above the lanes refer to the template names shown in A. C, Wild-type RNA C template; M, wild-type MDV RNA template.

synthesis (Figure 4B, lanes 5 and 6). These results suggest that the 3' 37 bases of (+)-strand sat-RNA C can function as a promoter for transcription *in vitro*, independent of the remaining sat-RNA C sequence. In addition, since chimeric templates containing two head-to-tail copies of the wild-type promoter produced a predominantly full-length RNA product, the internally located promoter was suppressed by a functional 3'-terminal promoter *in vitro*.

Discussion

We have constructed a secondary structural model for the 3'-end of (+)-strand sat-RNA C and provided evidence for the involvement of the 3'-terminal stem-loop in the initiation of (–)-strand synthesis. Involvement of the 3' stem-loop in (–)-strand synthesis was suggested by deletion analysis (Figure 1) and then confirmed by site-directed mutagenesis. Analyses of RNA templates with altered loop sequences (Figure 3B, lanes 15 to 18) strongly suggest that the RdRp does not recognize the size or sequence of the loop. Although single-base mutations that disrupted the base-pair adjacent to the loop (Cm6 and Cm7) still produced high levels of full-length (–)-strand transcripts (Figure 3B, lanes 7 and 8), mutations with two or three bases changed in the upper stem region decreased (Cm8) or eliminated (Cm9 and Cm10) the synthesis of full-length transcripts (Figure 3B, lanes 10 to 12). Exchanging three bases in the upper stem region (Cm11) did not restore template activity, possibly because of a disruption in the local secondary structure. Therefore, it is still unresolved whether the interaction between the RdRp and the upper stem has specific sequence requirements.

Deletion or mutation of the G residue at the base of the stem (position 328; Cm1 and Cm3), disrupted a G·C base-pair and substantially decreased template efficiency. Similar effects were also observed for multiple mutations in the left or right sides of the lower stem (Cm2 and Cm4). However, the compensatory mutation that switched four base-pairs in the lower stem completely restored template activity. These results indicate that a stable lower stem is important for (–)-strand initiation, but the sequence may not be specifically recognized by the RdRp. However, the U to G mutation in the midstem (position 331; Cm12; Figure 3B, lane 14), which results in a G·G mismatch, did not affect template efficiency, suggesting that a midstem symmetrical bulge is tolerated.

The identity of the products that migrated slightly faster than the full-length products is unknown (see Figure 3B, lanes 2 to 4, 7 and 9 to 10). Although transcription of wild-type sat-RNA C resulted in the synthesis of low levels of these products, substantial increase in their synthesis was observed for several templates with mutations in the 5' side of the stem (positions 328 to 334). One possible origin of these products is from the RdRp initiating synthesis at an alternative initiation site 10 to 15 bases from the 3' end, which is within a sequence normally base-paired in the 3'-proximal hairpin. Initiation from this alternative site could be activated in some of the mutants (e.g. Cm1 to 3) when base-pairing in the region is disrupted by the alteration of residues on the 5' side of the hairpin. Alternatively, these products may be due to the RdRp terminating transcription prematurely on the (+)-strand template.

In addition to the 3' stem-loop, the complete promoter may also include the single-stranded 3'-terminal sequence CUGCCC-OH that is found at the 3' termini of all RNAs replicated by the TCV RdRp, including the TCV genomic RNA, sat-RNAs C, D and F (Simon & Howell, 1986) and DI RNA G (Li *et al*., 1989). Recently, studies have been performed that indicate the importance of the CUGCCC-OH sequence. The sat-RNA D molecules with 3' deletions of up to 13 bases were repaired *in planta* by two weeks postinoculation. These newly repaired molecules contained heterogeneous sequences downstream from the sat-RNA D sequence nearly always joined to $CUG(C)_{1\text{-}4}$ at the 3'-termini (Carpenter & Simon, 1995). These results imply that the 3' CUGCCC sequence is likely to be important for RNA accumulation *in vivo*.

The 3'-terminal 37 bases of sat-RNA C, when present at the 3'-end of MDV RNA, directs the initiation of complementary strand RNA synthesis *in vitro*, indicating that this sequence can independently promote (–)-strand synthesis. Since deletions up to the terminal 29 nucleotides did not disrupt sat-RNA C (–)-strand synthesis, it is likely that only the terminal 29 nucleotides are necessary to function as a promoter *in vitro*. While *in vitro* mapping results do not rule out additional sequence requirements for high level replication *in vivo*, to our knowledge, this is the smallest sequence shown to function as an independent promoter for full-length complementary strand synthesis by an RNA-dependent RNA polymerase.

Our previous results demonstrated that the major product produced from sat-RNA C (+)-strands with 7 to 220 bases of plasmid sequence at the 3'-ends was the size of full-length sat-RNA C (–)-strands, indicating that transcription can initiate opposite the natural sat-RNA C 3'-end sequence even when located internally in the template (Song & Simon, 1994). It is therefore curious that full-sized RNA was the major product produced from a chimeric MDV template containing two head-to-tail copies of the sat-RNA C 3'-terminal 37 base sequence. This result suggests that a functional promoter located at the 3'-end of the template may suppress the function of an internally located promoter. Alternatively, the presence of a hairpin directly downstream from the internal initiation site may impede access of the RdRp to the internal promoter.

Although sat-RNA C and TCV share 90% sequence similarity in their common regions, the

sequences differ by a four base insertion and two base alterations within the sat-RNA C 3'-terminal hairpin. Interestingly, a similar stem-loop structure is still predicted by computer analysis to form in the 3'-end region of TCV (Figure 5). The TCV 3' hairpin is predicted to have a longer stem than sat-RNA C (three additional C·G pairs) and a shorter loop. The lack of conservation in the loop sequences between TCV and sat-RNA C and the variable stem lengths support our functional predictions based on the mutagenesis analysis of the sat-RNA C hairpin. Similar structures can also be formed from the 3'-end sequences of the related viruses cardamine chlorotic fleck (Skotnicki *et al*., 1993), maize chlorotic mottle virus (MCMV; Nutter *et al*., 1989) and carnation mottle (CarMV; Guilley *et al*., 1985). For MCMV and CarMV, the hairpin is present even though the 3'-end sequences share little similarity to TCV (Figure 5). Conservation of a 3'-terminal hairpin among carmoviruses suggests that this structure may be of functional importance in the replication of genomic and subviral RNAs.

Materials and Methods

Probing the structure of (+)-strand sat-RNA C using chemical and enzymatic modifications

A single modification buffer was used for all the chemical and enzymatic modifications to avoid possible conformational change of the RNA at different salt concentrations. Modification reactions were performed at 20°C to conform to the plant growth temperature and the temperature used for *in vitro* transcription by peak II RdRp fractions.

Wild-type (+)-strand sat-RNA C or (+)-strand sat-RNA C containing 220 bases derived from the vector at the 3'-end (C + 220; used to obtain the modification pattern of the extreme 3'-end of (+)-strand sat-RNA C) were synthesized by T7 RNA polymerase *in vitro* (Song & Simon, 1994) and purified by agarose gel electrophoresis. Eleven micrograms of purified (+)-strand sat-RNA C transcripts plus 11 µg of crude yeast tRNA/rRNA were added to 700 µl of modification buffer (70 mM Hepes (pH 7.5), 10 mM $MgCl_2$, 0.1 mM EDTA, 100 mM KCl) heated to 60°C, slowly cooled to 35°C, incubated at 20°C for three minutes and then placed on ice. Fifty microliter samples of the RNA were added to equal volumes of modification buffer containing: no additions (control); 42 mg/ml carbodiimide (CMCT; Aldrich); 10% (v/v) diethylpyrocarbonate (DEPC; Sigma); 1% (v/v) dimethylsulfate (DMS; Sigma); 0.05 units of RNase T_1; 0.03 units of RNase T_2; or 0.03 units of RNase V_1 (Sigma). CMCT, DEPC and DMS reactions were incubated at 20°C for ten and 20 minutes. RNase T_1, T_2 and V_1 reactions were incubated for five and ten minutes. The control tube was incubated at 20°C for ten minutes. The modified RNA samples were extracted with phenol/chloroform, precipitated with NH_4OH and isopropanol and then dissolved in 10 µl of distilled water.

For primer extension, 1 µl of primer (6 µM) was mixed with 2 µl of the modified RNA, heated briefly at 90°C, and then slow cooled to 40°C. Reverse transcription was performed in a 40 µl volume by the addition of 0.1 mM each of dGTP, dCTP and dTTP, 40 µCi of [α-^{35}S]dATP (New England Nuclear, 1398 Ci/mM), 60 units of AMV reverse transcriptase (Amersham) and buffer recommended by the manufacturer. After incubation at 42°C for five minutes, 1 µl containing 0.6 mM each of dATP, dGTP, dCTP and dTTP and 50 mM NaCl was added. After an additional 15 minutes incubation, reactions were terminated by the addition of 3 µl of 95% (v/v) formamide, 20 mM EDTA, 0.05% (w/v) bromophenol blue, and 0.05% (w/v) xylene cyanol. The RNA sequence ladder was generated using an RNA sequencing kit (Amersham). Samples were subjected to electrophoresis on 6% (w/v) polyacrylamide/50% (w/v) urea gels and exposed to X-ray film overnight. An oligonucleotide complementary to sat-RNA C positions 339 to 356 was used as a primer to determine the modified bases from positions 316 to 334. The M13/pUC (–40) sequencing primer (New England Biolabs) was used to determine the modified bases from positions 316 to 356 using C + 220 transcripts.

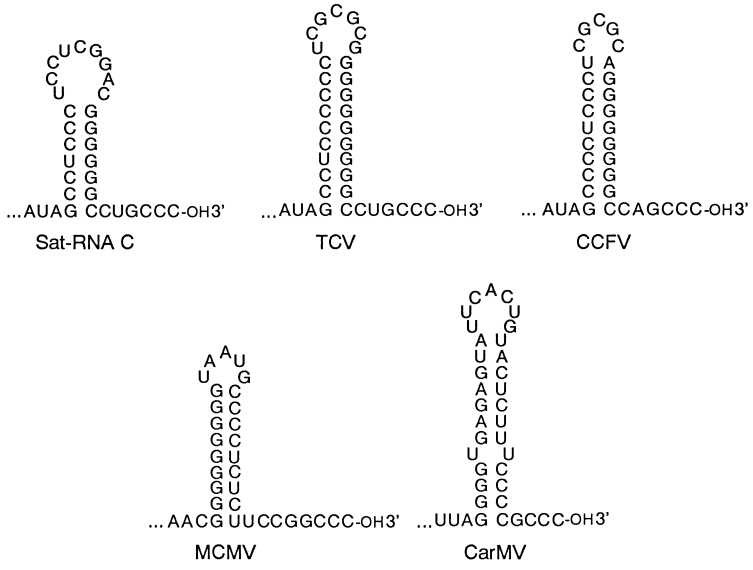

Figure 5. Comparison between the hairpin at the 3'-end of sat-RNA C and potential hairpin structures at the 3'-ends of TCV genomic RNA and related carmoviruses. CCFV, cardamine chlorotic fleck (Skotnicki *et al*., 1993); MCMV, maize chlorotic mottle virus (Nutter *et al*., 1989); CarMV, carnation mottle virus (Guilley *et al*., 1985).

Computer prediction of the RNA secondary structure

The M-FOLD program in the GCG package (Genetics Computer Group, Inc., University of Wisconsin) was used for secondary structure predictions.

Construction of deletion mutants

Deletions were constructed in plasmid pUC19T7C(+), containing a full-length sat-RNA C cDNA downstream of a T7 RNA polymerase promoter (Song & Simon, 1994). pUC19T7C(+) was digested with *Spe*I (New England Biolabs), treated with the slow form of *Bal*31 (Kodak) in buffer recommended by the manufacturer, and the ends religated.

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

Mutations were introduced into the 3'-end region of (+)-strand sat-RNA C as described by Kunkel (1985). The oligonucleotides used were: 5'-CAATAGATAA[A/C][A/C]TCCTCCTC-3', 5'-ATAGCCTCCC[A/C/G][A/G/T][A/G/T][A/C/T][G/C/T][A/G/T]GGGGGGCCTG-3', 5'-AGATAGCCTGGGTCCTCGGACCCCGGGCCTGCCC-3', 5'-AGATAGCC[A/C/G]CCCTCCTCGG-3', 5'-CAATAGATACGGGCCCTCCTCGGACGGGTCCGCTGCCCGGG-3' and 5'-TAGATAGCCT[G/C][G/C]GTCCTCGGAC-3', where the mutated bases are underlined.

Generation of MDV containing the 3' 37 bases of (+)-strand sat-RNA C

The 3'-terminal 37 bases of wild-type (+)-strand sat-RNA C or the mutant Cm10 (Figure 3A) were amplified by PCR using oligonucleotides complementary to the 3' 20 bases of (+)-strand sat-RNA C and 5'-ACAGCTGAATAGATAGCCT-3' (the underlined bases represent a *Pvu*II site introduced for cloning purposes). The PCR products were gel-purified and cloned into the *Sma*I site of plasmid pT7MDV (Axelrod *et al*., 1991), a kind gift from Dr Axelrod (American Cyanamid). Plasmids were recovered containing wild-type and Cm10 sat-RNA C sequences in the forward and reverse orientations. In addition, one plasmid contained two head-to-tail copies of the sat-RNA C 3'-end sequence.

In vitro transcription using RdRp from TCV-infected leaves and product analysis

Templates were synthesized using T7 RNA polymerase from plasmids digested with *Sma*I with the following exceptions. (1) Plasmids containing MDV and the 3'-end of sat-RNA C or Cm10 in the reverse orientation were digested with *Pvu*II. (2) To prepare full-length (+)-strand sat-RNA C from mutant Cm11, two oligonucleotides (M13/pUC [–40] sequencing primer and 5'-GGGCAGGCCCGGGGTCCG-3') were used to amplify the cDNA insert by polymerase chain reaction (PCR); the modification in Cm11 created a new internal *Sma*I site). The PCR products were gel-purified and used for *in vitro* transcription by T7 RNA polymerase. All RNA templates were extracted and precipitated as described by Song & Simon (1994).

RNA-directed RNA synthesis was performed using peak II RdRp fractions partially purified from TCV-infected turnip plants as described by Song & Simon (1994), except that actinomycin D was omitted from the RdRp reactions.

Acknowledgements

We thank Dr Clifford Carpenter for critical reading of the manuscript. This work was supported by National Science Foundation Grants MCB-9315948 and MCB-9419303 to A.E.S. and a University of Massachusetts graduate fellowship to C.S.

References

Axelrod, V. D., Brown, E., Priano, C. & Mills, D. R. (1991). Coliphage Qβ RNA replication: RNA catalytic for single-stranded release. *Virology*, **184**, 595–608.

Ball, L. A. & Li, Y. (1993). *cis*-Acting requirement for the replication of flock house virus RNA 2. *J. Virol.* **67**, 3544–3551.

Boccard, F. & Baulcombe, D. (1993). Mutational analysis of *cis*-acting sequences and gene function in RNA3 of cucumber mosaic virus. *Virology*, **193**, 563–578.

Carpenter, C. D. & Simon, A. E. (1995). *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.* In the press.

Cui, T. & Porter, A. G. (1995). Localization of binding site for encephalomyocarditis virus RNA polymerase in the 3'-noncoding region of the viral RNA. *Nucl. Acids Res.* **23**, 377–382.

Dreher, T. W. & 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. & 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. & 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.

Gallie, D. R. & Walbot, V. (1990). RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. *Genes Dev.* **4**, 1149–1157.

Guilley, H., Carrington, J. C., Balàzs, E., Jonard, G., Richards, K. & Morris, T. J. (1985). Nucleotide sequence and genome organization of carnation mottle virus RNA. *Nucl. Acids Res.* **13**, 6663–6677.

Hacker, D. L., Petty, I. T. D., Wei, N. & Morris, T. J. (1992). Turnip crinkle virus genes required for RNA replication and virus movement. *Virology*, **186**, 1–8.

Heaton, L. A., Carrington, J. C. & Morris, T. J. (1989). Turnip crinkle virus infection from RNA synthesized *in vitro*. *Virology*, **170**, 214–218.

Jacobson, S. J., Konings, D. A. M. & 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. & 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. & Makino, S. (1993). Analysis of *cis*-acting sequences essential for coronavirus defective interfering RNA replication. *Virology*, **197**, 53–63.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488–492.

Leathers, V., Tanguay, R., Kobayashi, M. & Gallie, D. R. (1993). A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation. *Mol. Cell. Biol.* **13**, 5331–5347.

Levis, R., Weiss, B. G., Tsiang, M., Huang, H. & 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. & Simon, A. E. (1991). *In vivo* accumulation of a turnip crinkle virus defective interfering RNA is affected by alterations in size and sequence. *J. Virol.* **65**, 4582–4590.

Li, X. H., Heaton, L. A., Morris, T. J. & 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, Y. & Ball, L. A. (1993). Nonhomologous RNA recombination during negative-strand synthesis of flock house virus RNA. *J. Virol.* **67**, 3854–3860.

Lin, Y.-J. & Lai, M. M. C. (1993). Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontiguous sequence for replication. *J. Virol.* **67**, 6110–6118.

Lin, Y.-J., Liao C.-L. & 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.

Miller, W. A., Bujarski, J. J., Dreher, T. W. & Hall, T. C. (1986). Minus-strand initiation by brome mosaic virus replicase within the 3' tRNA-like structure of native and modified RNA templates. *J. Mol. Biol.* **187**, 537–546.

Nutter, R. C., Scheets, K., Panganiban, L. C. & Lommel, S. A. (1989). The complete nucleotide sequence of maize chlorotic mottle virus genome. *Nucl. Acids Res.* **17**, 3163–3177.

Pacha, R. F., Allison, R. F. & Ahlquist, P. (1990). *cis*-Acting sequences required for *in vivo* amplification of genomic RNA3 are organized differently in related bromoviruses. *Virology*, **174**, 436–443.

Petty, I. T. D., Edwards, M. C. & Jackson, A. O. (1990). Systemic movement of an RNA plant virus determined by a point substitution in a 5' leader sequence. *Proc. Natl Acad. Sci. USA*, **87**, 8894–8897.

Simon, A. E. & 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.

Skotnicki, M. L., Mackenzie, A. M., Torronen, M. & Gibbs, A. J. (1993). The genomic sequence of cardamine chlorotic fleck carmovirus. *J. Gen. Virol.* **74**, 1933–1937.

Song, C. & 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.

van der Kuyl, A. C., Langereis, K., Houwing, C. J., Jaspars, E. M. J. & Bol, J. (1990). *cis*-Acting elements involved in replication of alfalfa mosaic virus RNAs *in vitro*. *Virology*, **176**, 346–354.

Edited by D. E. Draper

(Received 16 June 1995; accepted 30 August 1995)