Virology **275,** 177–192 (2000)
doi:10.1006/viro.2000.0511, available online at http://www.idealibrary.com on IDEAL®

Mapping of the *Tobacco Mosaic Virus* Movement Protein and Coat Protein Subgenomic RNA Promoters *in Vivo*

Valery Z. Grdzelishvili, Sean N. Chapman,[1] William O. Dawson, and Dennis J. Lewandowski[2]

Department of Plant Pathology, University of Florida, CREC, Lake Alfred, Florida 33850

Received May 17, 2000; returned to author for revision June 20, 2000; accepted July 6, 2000

The *Tobacco mosaic virus* movement protein (MP) and coat protein (CP) are expressed from 3′-coterminal subgenomic RNAs (sgRNAs). The transcription start site of the MP sgRNA, previously mapped to positions 4838 (Y. Watanabe, T. Meshi, and Y. Okada (1984), *FEBS Lett.* 173, 247–250) and 4828 (K. Lehto, G. L. Grantham, and W. O. Dawson (1990), *Virology* 174, 145–157) for the TMV OM and U1 strains, respectively, has been reexamined and mapped to position 4838 for strain U1. Sequences of the MP and CP sgRNA promoters were delineated by deletion analysis. The boundaries for minimal and full MP sgRNA promoter activity were localized between −35 and +10 and −95 and +40, respectively, relative to the transcription start site. The minimal CP sgRNA promoter was mapped between −69 and +12, whereas the boundaries of the fully active promoter were between −157 and +54. Computer analysis predicted two stem–loop structures (SL1 and SL2) upstream of the MP sgRNA transcription start site. Deletion analysis and site-directed mutagenesis suggested that SL1 secondary structure, but not its sequence, was required for MP sgRNA promoter activity, whereas a 39-nt deletion removing most of the SL2 region increased MP sgRNA accumulation fourfold. Computer-predicted folding of the fully active CP sgRNA promoter revealed one long stem–loop structure. Deletion analysis suggested that the upper part of this stem–loop, located upstream of the transcription start site, was essential for transcription and that the lower part of the stem had an enhancing role.

INTRODUCTION

The production of subgenomic mRNAs (sgRNAs) is one of the strategies that eukaryotic RNA viruses employ for expression and regulation of genes from a multicistronic RNA. Although several mechanisms of sgRNA synthesis have been proposed for various viruses (Spaan *et al.,* 1983; Miller *et al.,* 1985; Sit *et al.,* 1998; Sawicki and Sawicki, 1998), only internal initiation on the negative-strand template has been experimentally demonstrated for plant RNA viruses: *in vitro* for *Brome mosaic virus* (BMV) (Miller *et al.,* 1985) and *in vivo* for *Turnip yellow mosaic virus* (Gargouri *et al.,* 1989).

The boundaries of sgRNA promoters have been delineated for several plant viruses. As a rule, sgRNA promoters are located primarily upstream of the transcription start site (Marsh *et al.,* 1988; French and Ahlquist, 1988; van der Kuyl *et al.,* 1990, 1991; Boccard and Baulcombe, 1993; Johnston and Rochon, 1995; van der Vossen *et al.,* 1995; Wang and Simon, 1997; Koev *et al.,* 1999). The sgRNA promoter of *Beet necrotic yellow vein virus* and *Barley yellow dwarf virus* (BYDV) promoters for sgRNAs 2 and 3 are the exceptions, in that most of the promoter is located downstream of the transcription start site (Balmori *et al.,* 1993; Koev and Miller, 2000).

The sgRNA promoters of *Sindbis virus,* BMV and related plant tricornaviruses, and some other members of the alphavirus-like supergroup (Koonin and Dolja, 1993) contain a number of conserved sequence motifs that suggest possible parallels in sgRNA transcription among members of this supergroup (Marsh *et al.,* 1988; French and Ahlquist, 1988). In addition to the core (minimal) promoter sequences, the complete sgRNA promoter of BMV contains several elements required for full activity (Marsh *et al.,* 1988; Smirnyagina *et al.,* 1994). Using "proscripts" comprising both the BMV core sgRNA promoter and template sequences, Kao and co-workers demonstrated *in vitro* that nucleotides (nts) at positions −17, −14, −13, and −11, relative to the transcription start site, were essential for sgRNA synthesis and probably interact directly with the polymerase (Siegel *et al.,* 1997, 1998; Adkins and Kao, 1998). Moreover, they found that these nts were conserved not only in the *Bromoviridae,* but also among animal-infecting members of the *Alphavirus* genus.

Recent studies have revealed that not only primary sequence, but also secondary structure contributes to promoter activity *in vivo:* a stable stem–loop structure is predicted within the *Red clover necrotic mosaic* virus sgRNA promoter (Zavriev *et al.,* 1996); a 21-nt hairpin and

[1] Present address: Department of Biochemistry and Cell Biology, Scottish Crop Research Institute, Invergowrie, Dundee, UK DD2 5DA.

[2] To whom correspondence and reprint requests should be addressed at University of Florida, Citrus Research and Education Center, 700 Experiment Station Road, Lake Alfred, FL 33850. Fax: (863) 956-4631. E-mail: djlew@lal.ufl.edu.

0042-6822/00 $35.00

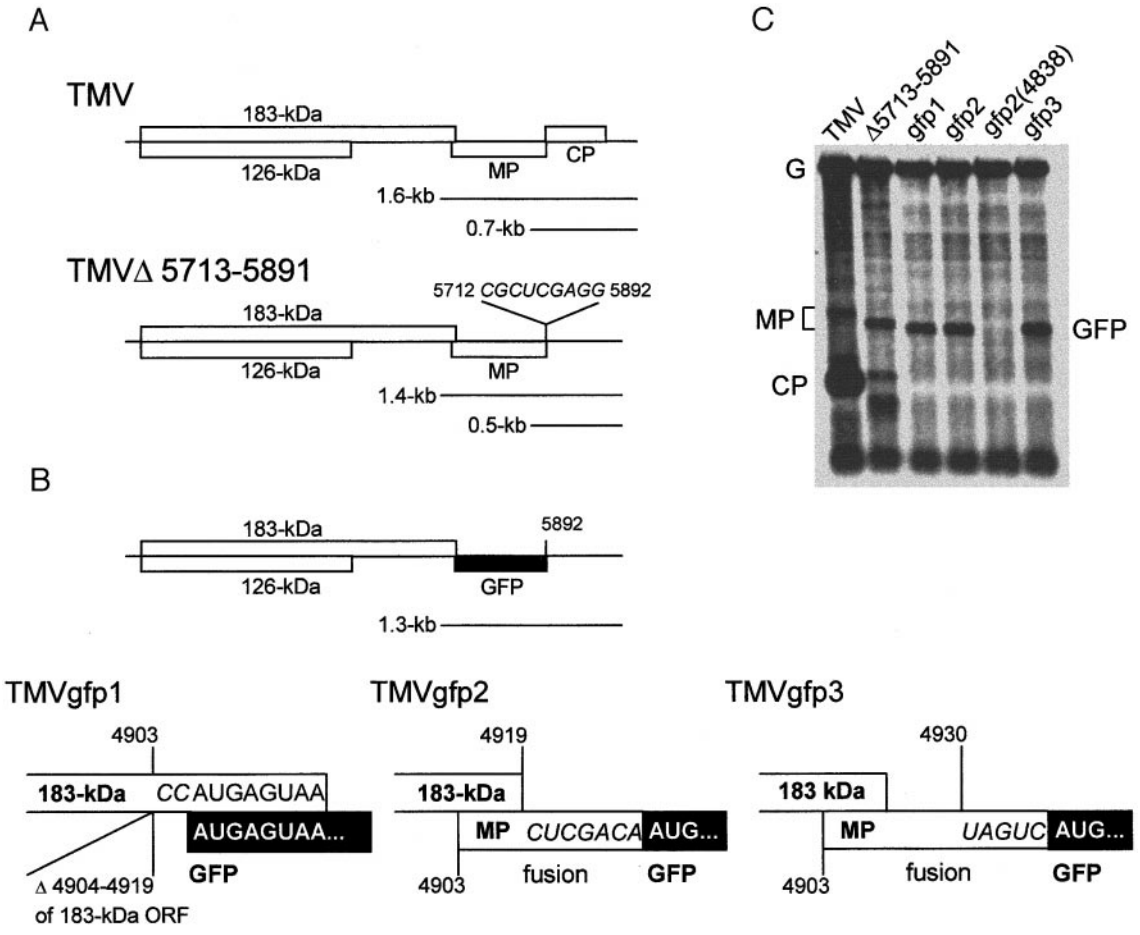

FIG. 1. Expression of GFP from the TMV MP sgRNA promoter. (A) Schematic diagram of TMV and TMVΔ5713–5891. MP, movement protein; CP, coat protein. Lines below genome diagrams represent sgRNAs and their approximate sizes. Numbering corresponds to TMV nts. Non-TMV nts (italics) between artificial junctions are indicated. (B) Genome organization of TMV hybrids with the MP ORF replaced with the green fluorescent protein (GFP) ORF. Detail of junctions of nonviral (italics) and GFP ORF nts are indicated below the generic hybrid. (C) Northern blot of total RNA extracted from tobacco protoplasts inoculated with *in vitro* transcripts of the viruses shown in (A) and (B). gfp2(4838) is a derivative of gfp2 containing a +1 C to G substitution. Each lane contains RNA extracted from $\sim 5 \times 10^3$ tobacco suspension cell protoplasts at 22 h postinoculation. Blots were probed with a digoxigenin-labeled riboprobe complementary to TMV nts 6202–6395 and analyzed as described (Lewandowski and Dawson, 1998). Positions of genomic RNA (G) and sgRNAs are indicated.

a 9-nt flanking single-stranded region are contained within the core promoter of *Turnip crinkle virus* (TCV; Wang *et al.,* 1999); stem–loops are predicted in the sgRNA promoters of BYDV (Koev *et al.,* 1999; Koev and Miller, 2000).

Tobacco mosaic virus (TMV), the type member of the *Tobamovirus* genus, has a positive-stranded RNA genome of 6395 nts that encodes at least four proteins (Goelet *et al.,* 1982). Only the 126- and 183-kDa proteins are translated directly from the genomic RNA, whereas the 30-kDa movement protein (MP) and the 17.5-kDa coat protein (CP) are translated from individual 3′-coterminal sgRNAs (Jackson *et al.,* 1972; Siegel *et al.,* 1976; Beachy and Zaitlin, 1977) (Fig. 1). A third sgRNA containing an ORF for a 54-kDa protein, corresponding to the C-terminal portion of the 183-kDa protein, has been isolated, although the protein has not been detected (Sulzinski *et al.,* 1985).

Although the MP and CP are expressed through sgRNAs, their expression patterns are markedly different. MP is produced maximally early in the infection (Watanabe *et al.,* 1984; Lehto *et al.,* 1990) and in low amounts (Ooshika *et al.,* 1984), whereas CP is expressed late, reaching an extremely high level (Siegel *et al.,* 1978). The time course of MP and CP expression is correlated with synthesis of the corresponding sgRNAs, suggesting transcriptional regulation (Ogawa and Sakai, 1984; Watanabe *et al.,* 1984). Translational regulation might also contribute to the differential synthesis of these proteins, as the sgRNAs have distinctly different leaders. The CP sgRNA has a 5′ 7-methyl guanosine cap and a short leader (Guilley *et al.,* 1979), while the leader of the MP sgRNA is believed to be uncapped (Hunter *et al.,* 1983; Joshi *et al.,* 1983) and is much longer (Watanabe *et al.,* 1984; Lehto *et al.,* 1990). Furthermore, a recent study suggests that MP might be translated from its sgRNA via

an internal ribosome initiation pathway (Skulachev *et al.,* 1999).

The TMV MP and CP sgRNA promoters have not been mapped, although the 253 nts upstream of the CP start codon directed sgRNA synthesis (Lehto *et al.,* 1990) and the addition of the 5′-proximal 45 nts of the CP ORF further enhanced activity of the CP sgRNA promoter in TMV-based vectors (Shivprasad *et al.,* 1999). Also, sequences upstream of TMV nt 4924, located within the MP ORF, are sufficient to direct transcription (Deom *et al.,* 1994; Giesman-Cookmeyer *et al.,* 1995).

As a first step to elucidate the mechanisms by which TMV is able to independently regulate expression of its genes from sgRNAs, we have mapped the MP and CP sgRNA promoters and dissected these regions into functional components. We have delineated the borders of the core and fully active MP and CP sgRNA promoters. We reexamined the transcription start site for the MP sgRNA and, using site-directed mutagenesis, investigated the importance of individual nts flanking the MP sgRNA transcription start site for promoter activity. In addition, we used deletion analysis and site-directed mutagenesis to reveal sequences and putative secondary structures important for sgRNA promoter activity.

RESULTS

Sequences within the MP ORF are not required for MP sgRNA synthesis

Sequences upstream of TMV nt 4924, including 21 nts of the MP ORF, are sufficient to direct transcription (Deom *et al.,* 1994; Giesman-Cookmeyer *et al.,* 1995). To determine whether the MP sgRNA promoter extends into the MP ORF, the jellyfish green fluorescent protein (GFP) ORF was inserted at three positions (Fig. 1B). In TMVgfp1, the GFP ORF replaced the complete MP ORF, which deleted 16 nts from the 3′-end of 183-kDa protein ORF. TMVgfp2 contains the complete 183-kDa protein ORF and a translational fusion between the first 17 nts of the MP ORF and the GFP ORF. TMVgfp3 contains a translational fusion between the first 28 nts of the MP ORF and the GFP ORF.

Tobacco suspension cell protoplasts were transfected with *in vitro* transcripts of each of the GFP constructs. Total RNA was extracted and analyzed by Northern blot hybridization. Each lane contained RNA extracted from $\sim 5 \times 10^3$ protoplasts and GFP sgRNA levels were directly compared between lanes. All three constructs containing GFP behind the putative MP sgRNA promoter replicated to similar levels and produced similar levels of GFP sgRNA in tobacco protoplasts (Fig. 1C), indicating that neither the sequences within the MP ORF nor the last 16 nts of the 183-kDa ORF were essential for TMV replication and MP sgRNA synthesis.

Mapping of the transcription start site for the MP sgRNA

The 5′-end of the MP sgRNA was mapped by Watanabe *et al.* (1984) to the G residue at position 4838 for the Japanese common strain (OM) of TMV and by Lehto *et al.* (1990) to a C residue at position 4828 for the U1 strain of TMV. These results predict MP sgRNA leaders of 65 and 75 nts for the OM and U1 strains, respectively.

Because transcription start sites have been mapped for a number of viruses from alphavirus-like supergroup and none of the sgRNAs have been found to initiate with a C residue (Adkins *et al.,* 1998), we decided to reexamine the transcription start site for the TMV U1 MP sgRNA using primer extension analysis. We designed primers that bound closer to the transcription start site than those used previously (Lehto *et al.,* 1990). For additional confirmation, we also mapped the transcription start site for GFP sgRNAs driven by the TMV MP sgRNA promoter using a primer that bound within the GFP ORF. Included in our experiments was a series of mutants, a resource not available when the transcription start site was initially mapped by Lehto *et al.* (1990).

Initial attempts to map the 5′-end of the MP sgRNA using total RNA extracted from tobacco protoplasts infected with wild-type TMV at 20 h postinoculation were inconclusive (data not shown). This may have been due to the high ratio of genomic RNA to MP sgRNA late in infection (Fig. 2A). However, the constructs cpΔ−238/−27 and cpΔ−238/+1, which contain deletions in the CP sgRNA promoter and are described later, accumulated increased levels of MP sgRNA compared to TMV (Fig. 2A). Primer extension with the MP leader-specific primer (complement of TMV nts 4867–4891) on total RNA extracted from protoplasts transfected with cpΔ−238/−27 or cpΔ−238/+1 yielded bands 53 and 54 bases in length, which correspond to nts 4838 and 4837 in the TMV genome (Fig. 2B). The smaller product corresponds to a 65-nt untranslated leader for the MP sgRNA. The larger band suggested that the MP sgRNA is capped, as doublets are typical of capped RNAs (Ahlquist and Janda, 1984), including TMV genomic RNA and CP sgRNA (Zimmern, 1975; Guilley *et al.,* 1979; Lewandowski and Dawson, 2000).

To determine the transcription start site for the GFP sgRNA produced by TMVgfp2, two different primers were used for primer extension. The MP leader-specific primer produced bands at positions 4838 and 4837 (data not shown), consistent with the results for wild-type MP sgRNA. The GFP-specific primer, which was complementary to GFP nts 16–38, also produced a doublet corresponding to nts 4838 and 4837 (Fig. 2C). As a control, a primer extension reaction was run on total RNA extracted from protoplasts inoculated with a TMV-based vector (TMV413) containing the GFP ORF behind the TMV CP sgRNA promoter. Primer extension reactions on total

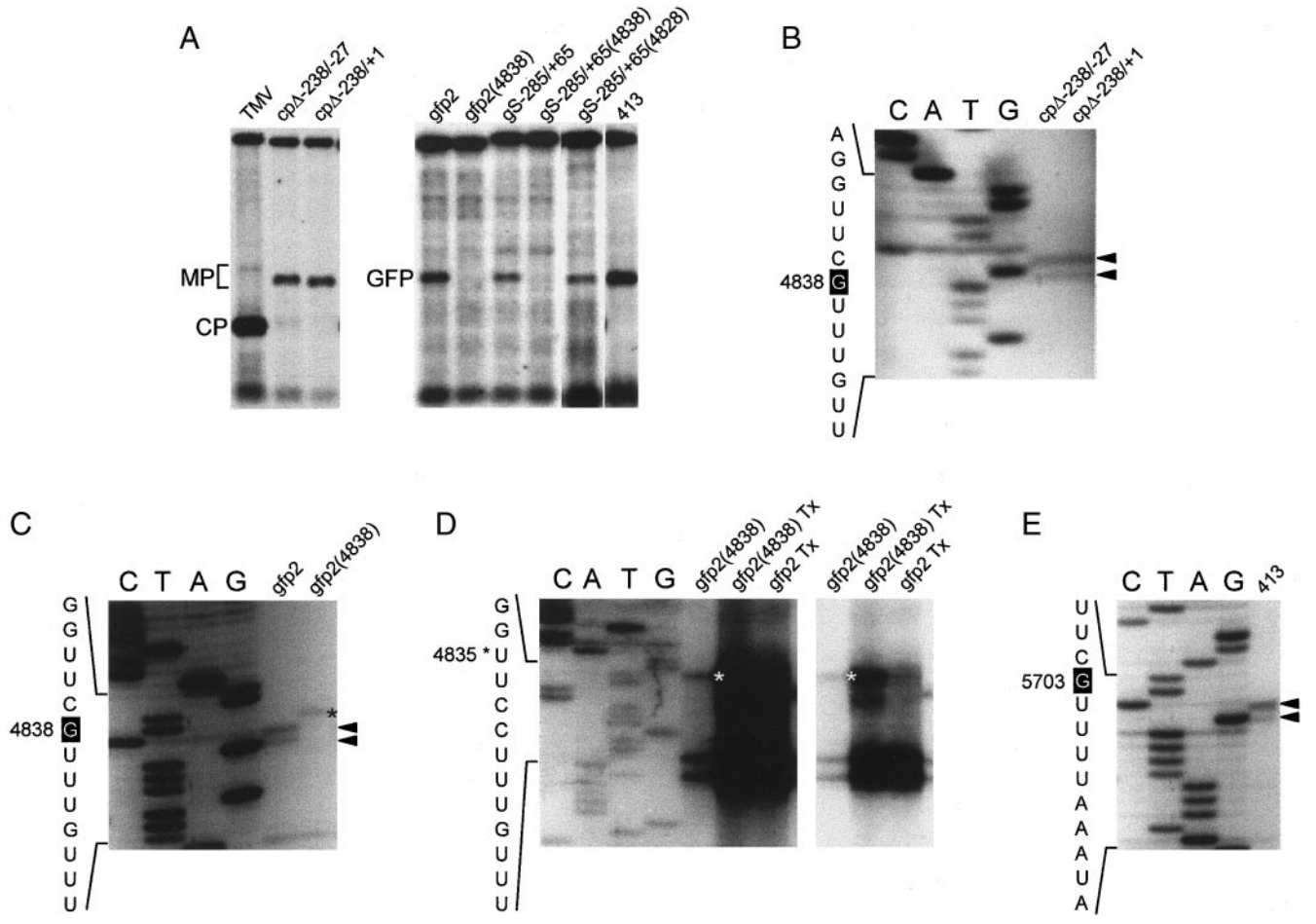

FIG. 2. Analysis of 5′ termini of TMV sgRNAs. (A) Replication and accumulation of sgRNAs for viruses with wild-type or mutant sgRNA promoters. Northern blots were probed as described in Fig. 1. Positions of sgRNAs are indicated. (B) Primer extension analyses of TMV mutants producing elevated levels of MP sgRNA. Total RNA extracted from infected protoplasts was subjected to primer extension analysis using a ^{32}P-end-labeled primer complementary to TMV nts 4867–4891. Products were resolved on a sequencing gel next to the DNA sequencing ladder of pTMV004. (C, D) Primer extension analyses of total RNA extracted from protoplasts infected with hybrids containing the GFP ORF behind the wild-type (gfp2) or +1 C to G mutant [gfp2(C4838)] MP sgRNA promoter. Primer extension reaction products using a primer complementary to GFP ORF nts 16–38 were run opposite the DNA sequence of pTMVgfp2 (C) or pTMVgfp2(4838) (D). *Band specific to +1 C to G mutant. (D) Control primer extension reactions of wild-type (gfp2) and +1 C to G [gfp2(4838)] *in vitro* RNA transcripts (Tx). (Right) A lighter exposure of the primer extension products. *Band derived from gfp2(4838)-infected protoplasts and *in vitro* transcripts. (E) Primer extension analysis of the GFP sgRNA expressed from the CP sgRNA promoter of TMV413 using the GFP-specific primer described in (C). Products were run next to the DNA sequence of pTMV413. Bands derived from sgRNAs are marked with arrowheads. The viral sequence derived from the sequencing ladders is indicated to the left of each figure. The 5′-proximal nt of the sgRNA (white letter on black background) is numbered according to the corresponding TMV nt.

RNA extracted from TMV413-infected protoplasts produced bands that corresponded to TMV nts 5703 and 5702, suggesting that the sgRNA was correctly initiated and was also capped (Fig. 2E).

To prove that nt 4838 was the transcription start site, a G to C substitution was introduced at position 4838 in TMVgfp2, which maintained the amino acid sequence of the 183-kDa replicase protein [TMVgfp2(4838)]. Although the level of replication was unaffected, this single nt substitution abolished GFP sgRNA synthesis (Figs. 1C and 2A). Primer extension of total RNA extracted from TMVgfp2(4838)-infected protoplasts with either primer did not produce bands at positions 4838 and 4837, but instead yielded a single band that corresponded to nt 4835 (Fig. 2C and data not shown). To determine whether the primer extension product that corresponded to nt 4835 was a transcription start site, control primer extension reactions with TMVgfp2 and TMVgfp2(4838) *in vitro* transcripts were run. Neither *in vitro* transcript yielded bands at positions 4838 or 4837 (Fig. 2D). However, primer extension of TMVgfp2(4838) *in vitro* transcripts produced a band at 4835, indicating that this band was an artifact from full-length genomic RNA (Fig. 2D).

To confirm the transcription start site, we made substitutions in the functional duplicated sgRNA promoter of TMVgS−285/+65 (described below), which directed synthesis of the artificial GFP sgRNA (Fig.2A). Similar to TMVgfp2(4838), loss of transcriptional activity occurred

with a G to C substitution at +1 in the duplicated MP sgRNA promoter [TMVgS−285/+65(4838)] (Fig. 2A). However, a C to G substitution at TMV nt 4828 [considered in Lehto *et al.* (1990) as a transcription start site] in the tandem MP sgRNA promoter [TMVgS−285/+65(4828)] did not affect transcription from the duplicated sgRNA promoter (Fig. 2A).

In the light of the large body of evidence, it is clearly demonstrated that the transcription start site for the MP sgRNA of TMV strain U1 is position 4838, corresponding to a length of 65 nts for the leader sequence. The mass of data presented, which is in accordance with the results obtained for the OM strain of TMV (Watanabe *et al.,* 1984), overturns the previous results for the U1 strain of TMV (Lehto *et al.,* 1990) that indicated a transcription start site at position 4828. Further, the doublet observed at positions 4838 and 4837 for all the transcriptionally active constructs tested suggests that the MP sgRNA is capped.

Mapping the boundaries of the MP sgRNA promoter

The sgRNA promoter for the MP gene is located within the ORF of the 183-kDa protein that is required for TMV replication (Ishikawa *et al.,* 1988; Lewandowski and Dawson, 2000). Thus, deletion mapping and site-directed mutagenesis were done on a duplicated fragment containing the putative MP sgRNA promoter. A fragment containing TMV nts 4503–4919 followed by the GFP ORF was inserted downstream of the 183-kDa ORF at position 5023 (Fig. 3A). The resulting construct TMVgS−335/+65 produced a 1.3-kb sgRNA (GFP sgRNA) and GFP (Fig. 3B), indicating that the fragment −335 to +65, relative to the +1 transcription start site, contained the MP sgRNA promoter.

Using appropriate primers, progressively smaller fragments ranging in size from −335/+65 to −25/+65 were amplified by PCR and introduced into pTMVgS (Fig. 3B). *In vitro* RNA transcripts from these constructs were used to inoculate tobacco protoplasts. Total RNA and proteins were analyzed by Northern and Western blot techniques, respectively (Fig. 3B).

The 3′ boundary of the fully active MP sgRNA promoter (in the negative strand) was mapped to −95 (complement of nt 4743) relative to the transcription start site (Fig. 3B). Larger portions of the 183-kDa ORF did not increase levels of GFP sgRNA or GFP.

Constructs with promoters −75/+65 (gS−75/+65) and larger produced a clearly distinguishable GFP sgRNA (Fig. 3A). However, a band with a slightly slower migration than the GFP sgRNA accumulated in gS−45/+65-, gS−35/+65-, and gS−25/+65-infected protoplasts (Fig. 3A), which made resolution of the extremely low levels of GFP sgRNA difficult. However, the smallest 3′ portion of the sgRNA promoter sequence that directed production of detectable amounts of GFP was 35 nts, indicating that −35 was the 3′ boundary of the core MP sgRNA promoter.

To minimize possible influences of the adjacent MP ORF sequences on activity of the second MP sgRNA promoter, some of the duplicated promoter/GFP cassettes were also inserted further downstream within the MP ORF at position 5463 [pTMVg(N)S; Fig. 3A]. We found that in both contexts, TMVgS and TMVg(N)S, the addition of only 5 nts turned a transcriptionally weak −90/+65 promoter into the fully active −95/+65 promoter (Fig. 3B and data not shown).

To determine whether sequences 5′ of the transcription start site (in the negative strand) are part of the MP sgRNA promoter, we created a series of constructs (in the TMVgS context) with progressively shorter leaders (Fig. 3C). GFP sgRNA levels from the −285/+40 promoter were similar to the −285/+65 promoter (Fig. 3C). However, larger deletions reduced GFP sgRNA accumulation to 15% or less of wild-type levels (Fig. 3C). The largest deletion able to produce detectable levels of sgRNA and GFP contained the first 10 nts of the leader. Thus, the 5′ boundary of the fully active MP sgRNA promoter was defined as +40, whereas the 5′ border of the core promoter was +10.

To determine whether the 135-nt region (−95/+40; complement of nts 4743–4877) had full MP sgRNA promoter activity, we created pTMVgS−95/+40. TMVgS−95/+40 produced levels of GFP sgRNA and GFP similar to TMVgS−95/+65 (data not shown). Thus, we delineated the boundaries of the fully active MP sgRNA promoter between −95 and +40, whereas the core promoter was localized between −35 and +10. In addition, we showed that a 135-nt region of the TMV genome comprising the MP sgRNA promoter (complement of nts 4743–4877) can be duplicated and retain full activity.

Mutations near the MP sgRNA transcription start site

A 6-nt region surrounding the transcription start site (−3 to +3) is identical for the TMV MP and CP sgRNA promoters (Watanabe *et al.,* 1984; Fig. 4B). A region −5 to +5 relative to the transcription start site of the MP sgRNA was found to be highly conserved among eight tobamoviruses (Fig. 4A). To investigate the importance of these 10 nts for MP sgRNA promoter activity, we introduced single nt substitutions between −5 and +5. All mutants were derivatives of pTMVgS−285/+65 (Fig. 5A).

The +1 C to G substitution (all changes shown in the negative strand) resulted in no detectable sgRNA accumulation (Fig. 5B). Substitutions at positions −5 (C to G), −4 (C to G), and −1 (G to C) reduced sgRNA accumulation to ~5% of wild-type levels (Fig. 5B). Surprisingly, individual transversions at positions −3, −2, +3, and +4 led to sgRNA levels about two times higher than with the wild-type motif of TMVgS−285/+65 (Fig. 5B). Transver-

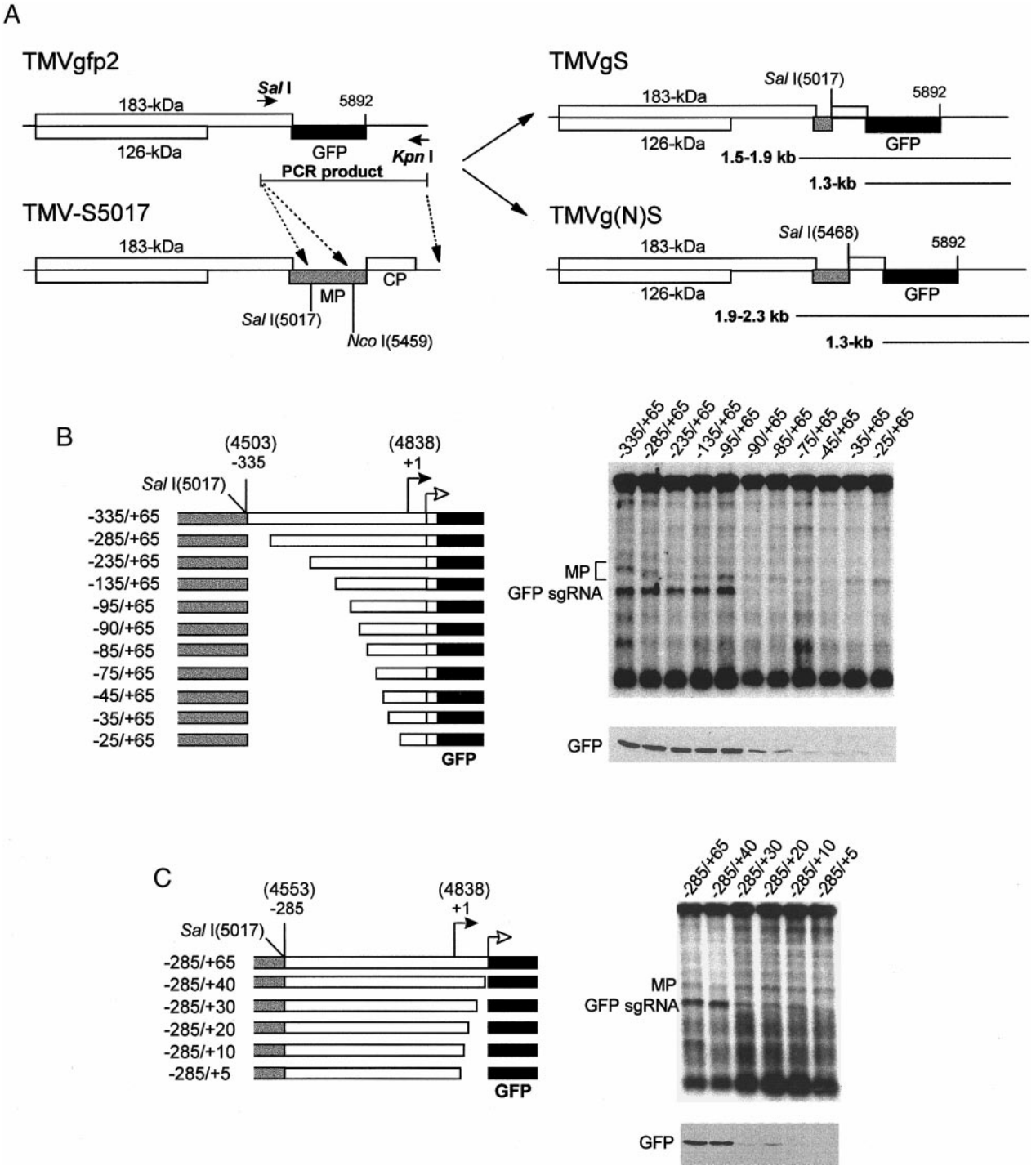

FIG. 3. Delineation of the boundaries of the TMV MP sgRNA promoter. (A) Schematic diagram of TMVgS and TMVg(N)S derivatives and strategies to amplify portions of the MP sgRNA promoter. Expected sgRNAs are indicated below the genome diagrams. Schematic diagram of strategy to map the 3′ (B) and 5′ (C) borders of the MP sgRNA promoter (in the negative strand) and representative Northern and Western blot data. The duplicated region is indicated to the left of each construct. Gray boxes, MP ORF; black boxes, GFP ORF; gaps, deletions; black arrowheads, transcription start site; empty arrowheads, translation start site. Northern blots (upper panels) were analyzed as described in Fig. 1. Western blots (lower panels) represent total soluble protein from $\sim 10^5$ protoplasts harvested at 22 h postinoculation that was resolved on 12% SDS–PAGE, electroblotted, and probed with polyclonal antiserum against GFP (Clontech). Positions of sgRNAs and GFP are indicated.

sions at positions +2 and +5 had no measurable effect on sgRNA accumulation (Fig. 5B).

Thus, the C to G substitution in the duplicated MP sgRNA promoter at the complement of TMV nt 4838 was the only change that abolished detectable transcription, which is in agreement with the primer extension data

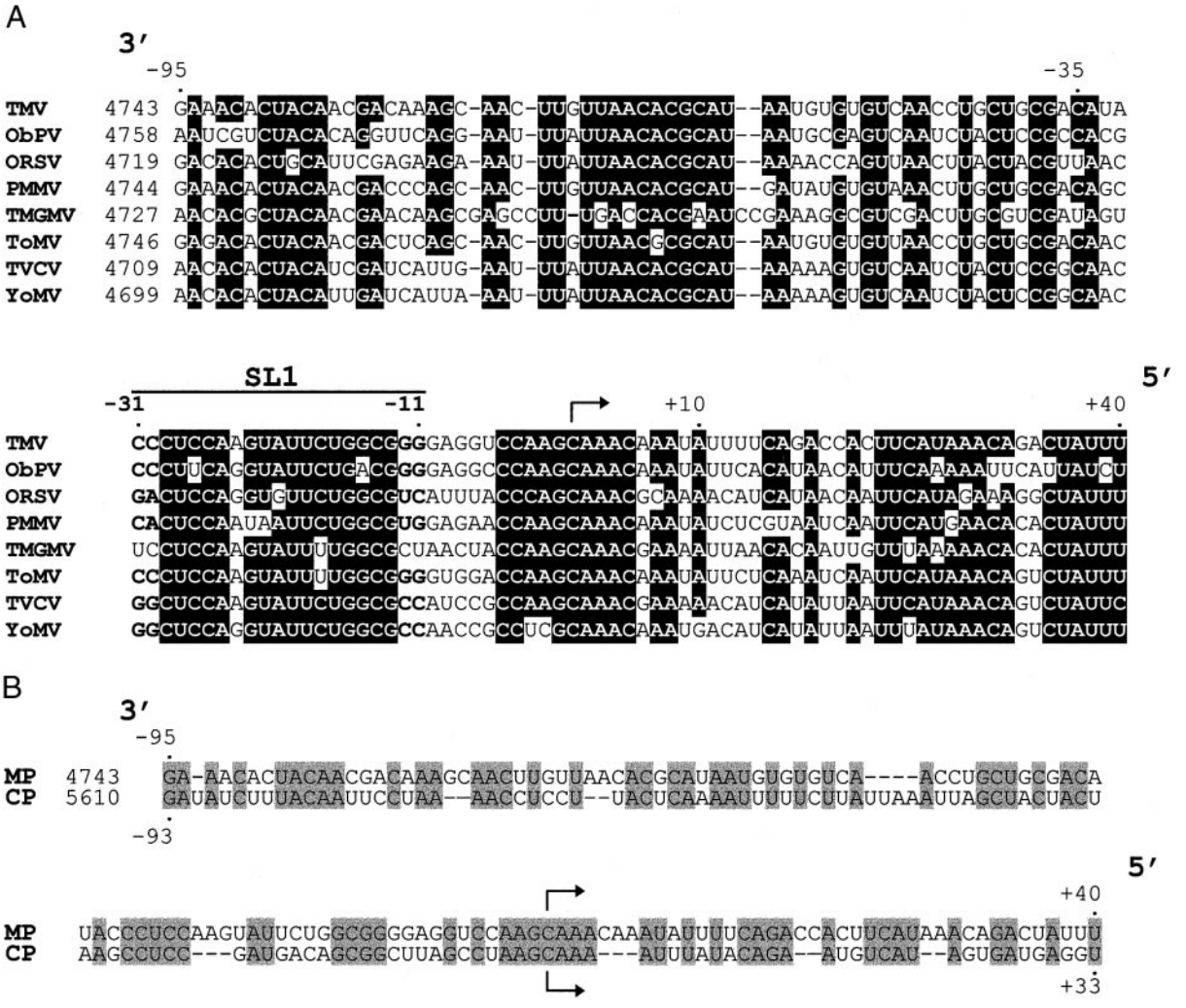

FIG. 4. (A) Alignment of the TMV MP sgRNA promoter with the putative MP sgRNA promoters of seven tobamoviruses. Sequences −95 to +40 relative to the transcription start site of TMV MP sgRNA were aligned with the region −95 to +40 relative to the putative MP sgRNA transcription start site for *Obuda pepper virus* (ObPV; Ikeda *et al.,* 1993), *Odontoglossum ringspot virus* (ORSV; Ryu and Park, 1995), *Pepper mild mottle virus* (PMMV; Alonso *et al.,* 1991), *Tomato mosaic virus* (ToMV; Ohno *et al.,* 1984), *Turnip vein-clearing virus* (TVCV; Lartey *et al.,* 1995), *Youcai mosaic virus* (YoMV; Aguilar *et al.,* 1996), and the region −98 to +40 for *Tobacco mild green mosaic virus* (TMGMV; Solis and Garcia-Arenal, 1990) using ClustalW (Thompson *et al.,* 1994). Black boxes indicate nts conserved in at least six tobamoviruses. SL1, stem–loop structure upstream of MP sgRNA transcription start site. Arrow, transcription start site. Bold letters at the borders of SL1 indicate compensatory mutations predicted to base-pair in SL1. (B) Alignment of the TMV MP and CP sgRNA promoters. MP, MP sgRNA promoter. CP, CP sgRNA promoter. Identical nts are shaded. Arrow, transcription start site. The sequences are negative sense, numbering is positive sense.

identifying this position as the transcription start site. However, positions −5, −4, and −1 were also found to be important for MP sgRNA synthesis.

Secondary structure prediction for the MP sgRNA promoter

To investigate the possible involvement of RNA secondary structure in MP sgRNA synthesis, we analyzed the 135-nt (−95/+40) promoter sequence using MFOLD (Zuker, 1989). Because sgRNA synthesis has been demonstrated to occur by internal initiation of transcription on the negative-stranded template (Miller *et al.,* 1985), we used the complement of nts 4743–4877 for the secondary structure predictions. Most of the MFOLD computational results contained two stem–loop (SL) structures (SL1 and SL2; Fig. 6A). SL1 is the complement of TMV nts 4807–4827 and is located between −31 and −11 relative to the transcription start site. SL2 is located between −83 and −37 and is the complement of nts 4755–4801.

Secondary structure of SL1 is required for transcription of MP sgRNA

Alignment of the TMV MP sgRNA promoter (−95 to +40) with the putative MP sgRNA promoter sequences of seven tobamoviruses revealed that the sequence of the SL1 region and location (−31 to −11) are highly

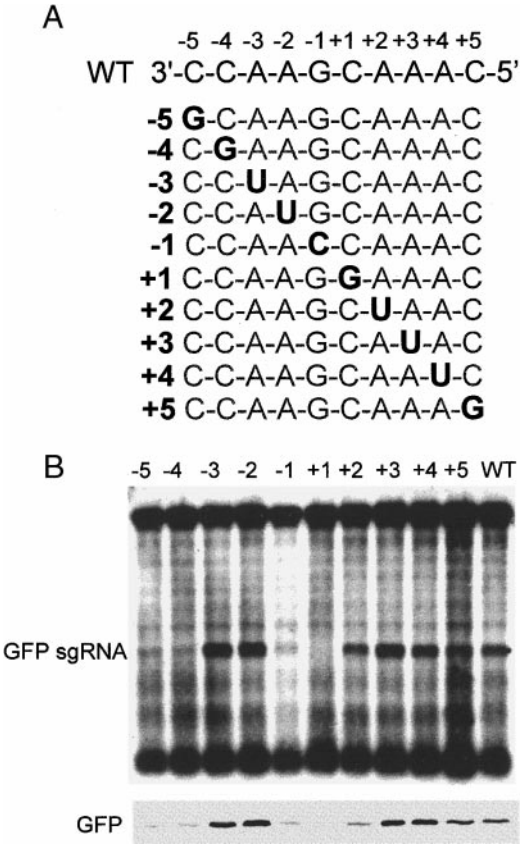

FIG. 5. Effect of point mutations near the MP sgRNA transcription start site. (A) Sequence of nts −5 to +5 (negative strand) flanking the MP sgRNA transcription start site for the wild-type (WT) MP sgRNA promoter and point mutations (bold) introduced into the duplicated MP sgRNA promoter of pTMVgS−285/+65. (B) Northern and Western blots were analyzed as described in Figs. 1 and 3, respectively. Positions of GFP sgRNA and GFP are indicated.

conserved (Fig. 4A). Precise deletion of the putative SL1 element (ΔSL1) prevented MP sgRNA synthesis (Fig. 6B). To test whether the SL1 secondary structure prediction has biological significance, we attempted to disrupt the putative base pairing in the stem of the SL1 in the duplicated MP sgRNA promoter of TMVgS−285/+65. Mutants SL1B and SL1C had alterations in opposite halves of the stem (Fig. 6A). Both sets of mutations abolished GFP sgRNA accumulation (Fig. 6B). However, the double mutant SL1BC containing compensatory changes predicted to restore base pairing in SL1 produced wild-type levels of sgRNA and GFP (Fig. 6B). To determine whether the specific sequence in the loop of SL1 is important for promoter activity, we changed the sequence (in the negative strand) 3′UAUU to 3′AUAA. This mutation (SL1D) had no effect on the sgRNA synthesis (Fig. 6B). Thus, our results suggest that the predicted secondary structure but not sequence of SL1 is important for MP sgRNA promoter activity.

Modifications upstream of the SL1 region

To determine whether the predicted SL2 structure is necessary for transcription, we created mutant ΔSL2 with a deletion of −76 to −38 in the duplicated MP sgRNA promoter of TMVgS−285/+65. Surprisingly, deletion of these 39 nts resulted in a fourfold increase in GFP sgRNA levels (Fig. 6B). Mutations designed to disrupt potential base pairing at the upper part of the putative SL2 structure (SL2B) did not affect sgRNA synthesis (Fig. 6B). However, modifications near the base of the putative SL2 structure had a negative effect on transcription. An 8-nt substitution (SL2C) and a 7-nt deletion (mpΔ−83/−77) reduced sgRNA levels to ~15% of wild-type levels (Fig. 6B). Together, these data suggest that SL2 is not necessary by itself for promoter activity, but might play a role in bringing discontinuous elements into closer proximity.

The region −95 to −84 of the MP sgRNA promoter was predicted by MFOLD as a single-stranded region (Fig. 6A). As shown in Fig. 3B, the presence of five additional TMV nts −95 to −91 turned a minimally active promoter into a fully active one, which indicated the importance of this region for promoter activity. To test whether sequence specificity or the absence of secondary structure in this region is important for promoter activity, nts −94 to −84 were replaced with A_{11} (in the negative strand; Fig. 6A). This 4-nt substitution decreased sgRNA levels to ~10% of wild-type levels (Fig. 6B). As only one of the four substitutions was within −95 to −91, it suggests that the sequence of the entire region (−95 to −84) is important for MP sgRNA promoter activity.

Mapping of the CP sgRNA promoter

The region between the *Nco*I site (nt 5459) and the CP translation start (nt 5712) contains sequences that are able to direct transcription (Lehto *et al.,* 1990). Insertion of heterologous ORFs in place of the CP ORF markedly reduced transcriptional activity relative to wild-type levels (Dawson *et al.,* 1989; Donson *et al.,* 1991; Kumagai *et al.,* 1993), suggesting that RNA sequences within the CP ORF may form part of the sgRNA promoter. In fact, when TMV-based transient expression vectors included 5′ proximal CP ORF sequences upstream of the foreign gene, sgRNA levels increased (Shivprasad *et al.,* 1999).

Because the CP sgRNA promoter overlaps the MP ORF, which is not essential for TMV replication, mapping was done with the CP sgRNA promoter in its natural context. To map the 3′ boundary of the CP sgRNA promoter, we created a series of constructs with progressively longer deletions from −238 toward the transcription start site (Fig. 7A). Deletion of nts 5465–5545 (cpΔ−238/−158) had no effect on the level of CP sgRNA (Fig. 7A). However, deletion of an additional 89 nts (cpΔ−238/−70) reduced the level of CP sgRNA to approximately one-third of wild-type levels (Fig. 7A). Larger

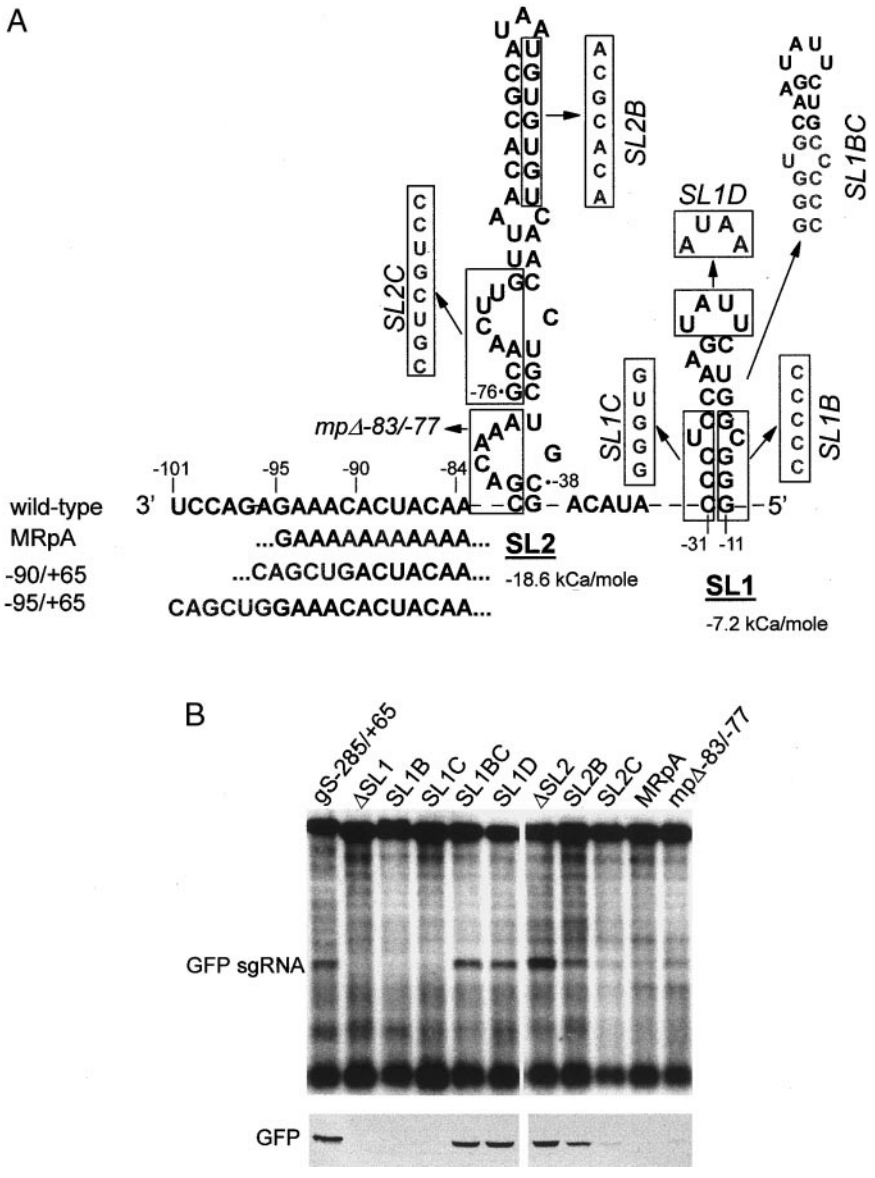

FIG. 6. RNA sequence and secondary structure analysis of the MP sgRNA promoter. (A) Prediction of putative stem–loop structures (SL1 and SL2) upstream of the MP sgRNA transcription start site (negative strand) by MFOLD (Zuker, 1989). The 3′-proximal wild-type sequence of −101/−84 is aligned with sequences for MRpA and the junctions' between the *Sal*I site and −95 (gS−95/+65) and −90 (gS−90/+65). Free energies were calculated at 25°C. Boxes contain nts substituted into the duplicated MP sgRNA promoter of TMVgS−285/+65. (B) Northern and Western blots were analyzed as described in Figs. 1 and 3, respectively. Positions of GFP sgRNA and GFP are indicated.

deletions (cpΔ−238/−27 and cpΔ−238/+1) did not produce detectable levels of CP sgRNA (Fig. 7A). Thus, the smallest portion of sequence upstream of the CP sgRNA transcription start site that directed transcription of detectable levels of CP sgRNA was 69 nts. Loss of CP sgRNA transcriptional activity in mutants cpΔ−238/−27 and cpΔ−238/+1 resulted in a dramatic increase in levels of MP sgRNA, suggesting that there may be competition between MP and CP sgRNA promoters (Fig. 7A). To determine whether sequences upstream of −238 affected transcriptional activity, mutants cpΔ−518/−245 and cpΔ−680/−245 were constructed. As shown in Fig. 7A, these deletions had little effect on CP sgRNA accumulation. Thus, the 3′ boundary of the full CP sgRNA promoter was mapped to −157.

To map the 5′ boundary of the CP sgRNA promoter, we created a series of deletions from +351 toward the transcription start site (Fig. 7B). To minimize the potential effect of RNA instability in the absence of CP, deletions were made in and compared to the free-RNA TMV mutant [−CP] (Culver *et al.,* 1993), which has the initiation codon of the CP ORF replaced with AGA. This enabled conclusions to be based only upon relative levels of CP sgRNA, while excluding the effect of encapsidation of genomic RNA on stability of the virus (Wang and Simon, 1997). Unlike the 3′ deletions (in the negative strand) that

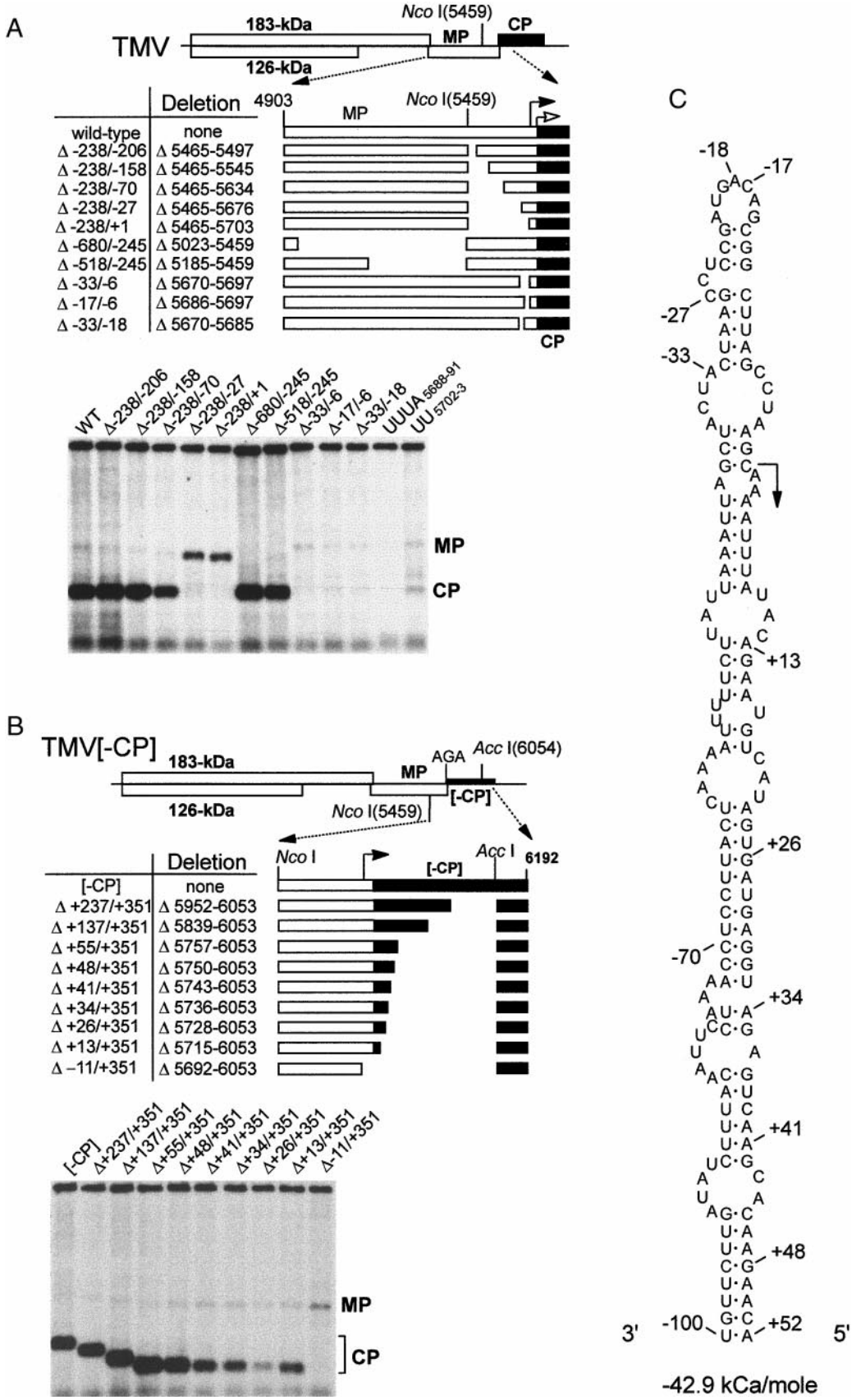

FIG. 7. Delineation of boundaries of the TMV CP sgRNA promoter and secondary structure prediction. Schematic diagram of strategies to map the 3′ border and putative structural elements (A) and 5′ border (B) of the CP sgRNA promoter and representative Northern blot. White boxes, MP ORF; black boxes, CP sequences; gaps, deletions; black arrowheads, transcription start site; empty arrowheads, translation start site. Northern blots were analyzed as described in Fig. 1. Positions of MP and CP sgRNAs are indicated. (C) Secondary structure prediction of the CP sgRNA promoter (negative strand). Stem–loop structure for the region −100 to +52 relative to the transcription start site was predicted by MFOLD (Zuker, 1989) at 25°C and was formatted using RnaViz (De Rijk and De Wachter, 1997).

showed a gradual decline in transcriptional activity, the smallest 5′ deletions had increased transcriptional activity. cpΔ+137/+351, which lacked nts 5839–6053, accumulated twofold more sgRNA than [−CP], and cpΔ+55/+351 resulted in the highest level of CP sgRNA (~250% of the [−CP] levels, Fig. 7B). Larger 5′ deletions led to decrease in CP sgRNA levels relative to [−CP] (Fig. 7B). Our data indicated that the core CP sgRNA promoter extends to +12, and 5′ boundary of the full CP sgRNA promoter extends into the CP ORF to +54. To determine whether nts at positions −1 and +1 are essential for the CP sgRNA promoter activity, we created a double mutant containing a −1/+1 GC to AA substitution in the complement of nts 5702–5703. These changes ($UU_{5702–3}$) nearly abolished transcription (Fig. 7A). Thus, we delineated the boundaries of the fully active CP sgRNA promoter between −157 and +54, and the core promoter was localized between −69 and +12, relative to the transcription start site.

Putative secondary structure of the CP sgRNA promoter

To investigate the possible role of RNA secondary structure in CP sgRNA synthesis, we analyzed the CP sgRNA promoter sequence (in the negative strand) using MFOLD (Zuker, 1989). A large stem–loop structure was predicted between −100 and +52 (complement of TMV nts 5603–5754) (Fig. 7C). Deletion of −33 to −6, sequences that were predicted to form the top portion of the stem plus the loop, abolished transcription (Fig. 7A). The computer-predicted folding of the cpΔ−33/−6 sgRNA promoter retained most of the original base pairing predicted for the wild-type sgRNA promoter (Fig. 7C), except that the upper stem–loop was deleted and a new loop was predicted. Deletion of −33 to −18 or −17 to −6 was predicted to retain base pairing in the bottom of the stem–loop (below the deletion). These deletions resulted in nearly complete (cpΔ−17/−6) or total (cpΔ−33/−18) loss of transcriptional activity (Fig. 7A).

In an attempt to more subtly disrupt the base pairing in the top of the putative stem–loop structure, the complement of nts 5688–5691 (−12 to −15) was changed from 3′GCGG to 3′AAAU. This mutation ($UUUA_{5688–91}$), which was predicted to create a larger loop, abolished transcription (Fig. 7A). There was a correlation between disruption of the computer-predicted structure and loss of transcriptional activity. The larger 3′ deletions and all 5′ deletions that resulted in reduced accumulation of CP sgRNA were predicted to shorten the overall length of the base-paired region, yet not disrupt the uppermost part of the stem–loop. The presence of the adjacent sequence that resulted from the deletions did not affect the base pairing within the stem–loop structure above the junction point. cpΔ−238/−27, cpΔ−238/+1, and cpΔ−11/+35 were predicted to disrupt the entire stem–loop structure (data not shown), which is consistent with the *in vivo* loss of transcriptional activity (Fig. 7). Thus, our data suggest that the upper part of the predicted stem–loop structure located upstream of the CP sgRNA transcription start site is required for CP sgRNA promoter activity.

DISCUSSION

Because of the discrepancy in the position of the MP sgRNA transcription start site (Watanabe *et al.,* 1984; Lehto *et al.,* 1990), we reexamined the 5′-end of the MP sgRNA. Two different approaches, point mutagenesis of published transcription start positions and primer extension analysis, were applied. Our results were consistent with the result of Watanabe *et al.* (1984) that position 4838 is the TMV MP sgRNA transcription start site. The misidentification of the transcription start site at position 4828 by Lehto *et al.* (1990) may have resulted from the use of a primer that bound further from the putative transcription start site than the primers used in this study or the use of MMLV reverse transcriptase at 37°C. Unlike the earlier studies, we utilized mutants with elevated levels of MP sgRNA and also used TMV derivatives with a heterologous ORF. Initiation at 4838 supports the "pyrimidine-adenylate rule" (Adkins *et al.,* 1998) for nts at the +1/+2 positions in sgRNA promoters (in the negative strands) of members of the alphavirus-like supergroup, whereas previously published data (Lehto *et al.,* 1990) indicated a purine G at +1.

Unexpectedly, we observed a doublet of bands of equal intensity corresponding to positions 4838 and 4837 produced from primer extension reactions on RNAs transcribed from the MP sgRNA promoter. Doublets were observed for TMV genomic RNA and CP sgRNA, which are known to be capped, but not for uncapped TMV *in vitro* transcripts (Lewandowski and Dawson, 2000). Previously, indirect evidence suggested that the MP sgRNA was not capped (Hunter *et al.,* 1983; Joshi *et al.,* 1983). A recent study of Skulachev *et al.* (1999) suggested that the MP might be translated from its sgRNA via an internal ribosome initiation pathway, which might be expected if the MP sgRNA was uncapped. However, Watanabe *et al.* (1984) suggested that the retarded migration of a trimmed product of S1 nuclease mapping of the MP sgRNA might indicate the existence of a cap. Although we have not biochemically confirmed the existence of a cap structure at the 5′-end of the MP sgRNA, our primer extension data suggest that MP sgRNA is capped.

We determined that the core and the full MP sgRNA promoters were between −35 and +10 and between −95 and +40, respectively, relative to the transcription start site. The 3′ border of the core CP sgRNA promoter was mapped between −26 and −69 and the 5′ border was +12, whereas the fully active CP sgRNA promoter was mapped between −157 and +54. Although the sizes

of core and fully active TMV MP and CP sgRNA promoters were similar, we observed that larger 3′ or 5′ deletions in the CP sgRNA promoter caused gradual declines in the sgRNA levels, whereas addition of only five 3′ or ten 5′ nts to the minimally active MP sgRNA promoter caused a sharp increase in transcription. However, additional 3′ or 5′ sequences did not result in further increases in MP sgRNA levels. We cannot rule out the possibility that some of the 5′ MP or CP sgRNA promoter deletions may have affected sgRNA stability, which could have affected sgRNA accumulation.

Alignment of the TMV MP sgRNA promoter with the sequences of the putative MP sgRNA promoters of seven tobamoviruses revealed 55% sequence homology with the highest identity in the region −5 to +5. Point mutations within this region demonstrated the importance of certain nts for MP sgRNA promoter activity *in vivo.* The only substitution that completely abolished transcription was a C to G transversion (in the negative strand) at +1, consistent with results obtained for *Alfalfa mosaic virus* (van der Vossen *et al.,* 1995) and BMV (Siegel *et al.,* 1997; Stawicki and Kao, 1999). The importance of individual nts around transcription start sites has been investigated in detail by Kao and co-workers using BMV proscripts *in vitro* (Siegel *et al.,* 1997; Stawicki and Kao, 1999; Adkins *et al.,* 1998). However, we found that different nts surrounding the transcription start site were important for the TMV MP sgRNA promoter activity. Unlike BMV, TMV nts at positions −5, −4, and −1 were important, but a substitution at +2, which is important for BMV sgRNA promoter activity, had little effect on the TMV MP sgRNA promoter. These differences are consistent with the lack of significant homology between TMV sgRNA promoters and consensus sequences from sgRNA promoters of many members of the alphavirus-like supergroup (Marsh *et al.,* 1988; French and Ahlquist, 1988; Grakoui *et al.,* 1989). Surprisingly, alterations at positions −3, −2, +3, and +4 caused twofold higher MP sgRNA levels relative to the wild-type promoter sequence. This increase suggests that there are restrictions on sgRNA promoter sequences used for other functions (183-kDa coding sequence) besides transcription, because three of these substitutions led to alteration of the 183-kDa amino acid sequence. Alternatively, the MP sgRNA promoter might be negatively regulated, which is consistent with extremely low amounts of MP sgRNA relative to genomic RNA and CP sgRNA (Ooshika *et al.,* 1984; Moser *et al.,* 1988).

The importance of secondary structure for sgRNA promoter activity has been demonstrated for the *Red clover necrotic mosaic virus* (Zavriev *et al.,* 1996), TCV (Wang *et al.,* 1999) and BYDV (Koev *et al.,* 1999; Koev and Miller, 2000) sgRNA promoters. Although BMV proscripts directed *in vitro* transcription in a manner independent of secondary structure (Siegel *et al.,* 1997), a recent study by Haasnoot *et al.* (2000) demonstrated that secondary structure is required for BMV full sgRNA promoter activity *in vitro.* Two stem–loop structures (SL1 and SL2) were predicted upstream of the TMV MP sgRNA transcription start site. The sequence of SL1, its putative secondary structure, and location relative to the transcription start site were found to be highly conserved among eight tobamoviruses. Our data suggest that the secondary structure of SL1, but not its nt sequence is required for MP sgRNA promoter activity.

Surprisingly, deletion of the 39 nts of the SL2 region caused a fourfold higher accumulation of MP sgRNA, which is consistent with a hypothesis of negative regulation of MP sgRNA synthesis. Although deletion of SL2 indicated that this region was dispensable for MP sgRNA promoter activity, attempts to disrupt potential base pairing in the bottom of the putative SL2 structure greatly reduced sgRNA accumulation, suggesting a possible role for SL2 in bringing discontinuous elements closer together. The 3′ 12-nt region (−95 to −84) of the MP sgRNA promoter predicted by MFOLD as a single-stranded region might be one of these elements. Mutational analysis indicated the importance of the primary sequence of this 12-nt region for MP sgRNA promoter activity.

Whereas several separate structural elements were predicted within the full MP sgRNA promoter, nearly the entire CP sgRNA promoter sequence was folded into one long stem–loop structure. Although there is low homology (18%) between the CP sgRNA promoters of different tobamoviruses, most of them could be folded into similar long stem–loop structures. Progressively larger 3′ or 5′ deletions in the CP sgRNA promoter caused progressive unfolding of the base of this putative structure, whose length of base-paired sequence was proportional to CP sgRNA promoter activity. The core CP sgRNA promoter retained only the upper part of this stem–loop, whereas mutations designed to disrupt base pairing in the upper portion of the stem–loop abolished or nearly abolished transcription. Together these data suggest that the upper part of the putative structure is essential for CP sgRNA promoter activity, and the base of the stem has an enhancing role.

Independent regulation of multiple genes from sgRNA promoters suggests differences between these elements. Similar to TMV, other viruses that produce multiple sgRNAs such as *Cucumber necrosis virus* (Johnston and Rochon, 1995), BYDV (Koev *et al.,* 1999; Koev and Miller, 2000), and TCV (Wang and Simon, 1997) have low sequence similarity between their sgRNA promoters. However, the ability for the viral replicase to recognize more than one sgRNA promoter presumes that there must be common elements within the promoters. Although we did not find significant sequence homology (40%) between the TMV MP and CP sgRNA promoters, we suggest that the stem–loop structure, located upstream of the transcription start site in both promoters,

and similar sequence motifs surrounding the sgRNA transcription start sites could represent general requirements for TMV sgRNA promoter recognition. Koev *et al.* (1999) suggested that a structural element located upstream of the transcription start site might act as a replicase recognition site, separated from the transcription start site. However, it remains possible that distinct elements within MP and CP sgRNA promoters could be recognized by different *trans*-acting factors within the replicase complex, as for Qβ replicase holoenzyme, in which separate protein factors are responsible for positive- and negative-stranded RNA recognition by the replicase complex (Brown and Gold, 1996).

MATERIALS AND METHODS

Plasmid constructions

All TMV clones used in this study are the derivatives of an infectious wild-type TMV cDNA clone (Dawson *et al.,* 1986) and were constructed with standard recombinant DNA techniques (Sambrook *et al.,* 1989). pTMV004 is T7 wild-type TMV clone containing a *Kpn*I site at the 3′ terminus for linearization (Lewandowski and Dawson, 1998). pTMVΔMPΔCP was constructed by amplifying the 3′ third of the 183-kDa ORF of pTMV004 with PCR using primer D369, corresponding to TMV nts 3324–3345, and a primer containing a *Xho*I site followed by the complement of TMV nts 4902–4919. The resulting PCR product was digested by *Bam*HI and *Xho*I and ligated into *Bam*HI/*Xho*I-digested pT7S3-28 (Lewandowski and Dawson, 1998). pTMVΔ5713-5891 (Fig. 1A) was constructed by PCR amplification of the 3′ terminal 0.5-kb region of pTMV004 with a primer containing a *Xho*I site followed by TMV nts 5892–5909 and M28, a primer containing a *Kpn*I site and the complement of TMV nts 6381–6395. The resulting PCR product was digested with *Xho*I and *Kpn*I and ligated into *Xho*I/*Kpn*I-digested pTMVΔMPΔCP. pTTT-GFP/301 was constructed by ligating the *Xho*I/*Kpn*I fragment from pTMVΔ5713–5891 into *Xho*I/*Kpn*I-digested pTTT-GFP (Shivprasad *et al.,* 1999).

To construct pTMVgfp1, two DNA fragments were amplified by PCR. pTMV004 was amplified with D369 and a primer containing an *Nco*I site followed by the complement of TMV nts 4883–4903, and the resulting product was digested with *Bam*HI and *Nco*I. pTTT-GFP/301 was amplified with M28 and a primer containing a *Bsp*HI site followed by the first 20 nts of the GFP ORF including the start codon, and the resulting product was digested with *Bsp*HI and *Kpn*I. The *Bam*HI/*Nco*I- and *Bsp*HI/*Kpn*I-digested PCR products were ligated into *Bam*HI/*Kpn*I-digested pTMV004. To construct pTMV-S5017, a *Sal*I site was introduced into pTMV004 at position 5017 by overlap extension PCR (Higuchi *et al.,* 1988). To construct pTMVgfp2, pTTT-GFP/301 was amplified by PCR with M28 and a primer that introduced a *Sal*I site and an additional A residue 5′ of the GFP ORF, digested with *Sal*I and *Kpn*I, and ligated into *Xho*I/*Kpn*I-digested pTMVΔ5713–5891. To construct pTMVgfp3, the 3′-proximal PCR amplification product from pTTT-GFP/301 that was used to construct pTMVgfp1 was digested with *Kpn*I and ligated into *Eco*RV/*Kpn*I-digested pT7TE1, a T7 derivative of pTMVTE1 (Gera *et al.,* 1995) that contains an *Eco*RV site at position 4929. pTMV413 contains the TMV 126/183-kDa ORF followed by TMV nts 5460–5756, which contains the CP sgRNA promoter, the GFP ORF, and TMV nts 5892–6395.

To construct pTMVg(N)S−335/+65, which contains a *Sal*I site at the position 5468, pTMVgfp2 was amplified by PCR with M28 and a primer containing a *Sal*I site followed by TMV nts 4503–4523, digested with *Kpn*I, and ligated into the pTMV004, which was first digested with *Nco*I, end-filled with T4 DNA polymerase, and then digested with *Kpn*I. To introduce the second copy of the putative MP sgRNA promoter at position 5023, pTMVg(N)S−335/+65 was digested with *Sal*I and *Kpn*I, and the *Sal*I-*Kpn*I fragment was ligated into the *Sal*I/*Kpn*I-digested pTMV-S5017, creating pTMVgS−335/+65.

Smaller fragments of the putative MP sgRNA promoter were amplified from pTMVgfp2 with PCR using M28 and an upstream primer containing a *Sal*I site followed by sequences corresponding to different positions upstream of the MP sgRNA transcription start site. Resulting PCR products were digested with *Sal*I and *Kpn*I and substituted into *Sal*I/*Kpn*I-digested pTMVg(N)S−335/+65 or pTMVgS−335/+65.

Constructs containing deletions within the MP sgRNA leader or mutations near the transcription start site or within computer predicted secondary structures were created from pTMVgS−285/+65 by overlap extension PCR using pairs of appropriate primers. Sequence differences relative to pTMVgS−285/+65 are shown in Figs. 3, 5, and 6.

To construct a series of clones with deletions upstream of the CP sgRNA transcription start site, pTMV004 was amplified with M28 and an upstream primer containing an *Nco*I site followed by the sequence corresponding to different regions upstream of the CP sgRNA transcription start site. PCR products were digested with *Nco*I and *Kpn*I and ligated into *Nco*I/*Kpn*I-digested pTMV004. Resulting clones are named according to which nts were deleted relative to the transcription start site. To delete sequences upstream of the *Nco*I site (position 5459), a *Sal*I site was introduced at 5179 by overlap extension PCR to create pTMV-S5179. To create pTMVcpΔ−518/−245 and pTMVcpΔ−680/−245, pTMV-S5179 and pTMV-S5017, respectively, were digested with *Sal*I and *Nco*I, end-filled with T4 DNA polymerase, and religated. Clones with deletions in the putative terminal stem–loop structure of the CP sgRNA promoter (pTMVcpΔ−33/−6, pTMVcpΔ−33/−18, and pTMVcpΔ−17/−6), substitutions at 5688–5691 CGCC → UUUA (pTMV-$UUUA_{5688-91}$) and CG → UU at 5702–5703 (pTMV-UU_{5702-3})

were constructed by overlap extension PCR. To obtain constructs with deletions downstream of the CP sgRNA transcription start site, pT7TMV[−CP], a T7 derivative of pTMV[−CP] (Culver and Dawson, 1993) containing a translationally silent CP ORF was amplified with a common upstream primer corresponding to TMV nts 5154–5175 and a 3′ primer that introduced an *Acc*I site, followed by the complement of nts downstream of the CP sgRNA transcription start site. Resulting PCR products were digested with *Nco*I and *Acc*I and ligated into *Nco*I/*Acc*I-digested pT7TMV[−CP].

Protoplast preparation and inoculation

In vitro RNA transcripts were synthesized from ~1.6 μg of *Kpn*I-linearized plasmid DNA as previously described (Lewandowski and Dawson, 1998) and used to inoculate ~2×10^6 protoplasts prepared from a *Nicotiana tabacum* cv. Xanthi suspension cell line (Lewandowski and Dawson, 2000).

Analysis of RNA and GFP

Total RNA was extracted at 20–22 h postinoculation and analyzed by Northern blot hybridization with a TMV 3′-untranslated region-specific probe as described (Lewandowski and Dawson, 1998). Each lane on Northern blot contains RNA extracted from ~5×10^3 tobacco suspension cell protoplasts. Total RNA was subjected to primer extension analysis with AMV reverse transcriptase (US Biochemicals) at 42°C as previously described (Lewandowski and Dawson, 2000) except that 5–10 × 10^5 dpm of MP sgRNA leader- or GFP-specific end-labeled primer was used per reaction.

Protoplast pellets containing ~5×10^5 cells were resuspended in 50 μl of 2× loading buffer (Laemmli, 1970) and boiled for 3 min, and 12 μl was resolved by 12% SDS–PAGE and analyzed by Western blotting with GFP-specific antiserum (Clontech, Inc.) as previously described (Shivprasad *et al.,* 1999). Relative levels of RNA were quantified using scanning and densitometry as previously described (Lewandowski and Dawson, 1998) by directly comparing levels of sgRNA. Each construct was analyzed in a minimum of three independent experiments.

ACKNOWLEDGMENTS

The authors thank Cherie Sine and Cecile Robertson for excellent technical assistance. We thank Scott Adkins for helpful discussions and María A. Ayllón and Elizabeth Knapp for critical reading of the manuscript. This research was supported in part by USDA Grant 98-35505-6752 and an endowment in honor of J. R. and Addie S. Graves. This work is part of the Florida Agricultural Experiment Station Journal Series R-07575.

REFERENCES

Adkins, S., and Kao, C. C. (1998). Subgenomic RNA promoters dictate the mode of recognition by bromoviral RNA-dependent RNA polymerases. *Virology* **252,** 1–8.

Adkins, S., Stawicki, S. S., Faurote, G., Siegel, R. W., and Kao, C. C. (1998). Mechanistic analysis of RNA synthesis by RNA-dependent RNA polymerase from two promoters reveals similarities to DNA-dependent RNA polymerase. *RNA* **4,** 455–470.

Aguilar, I., Sanchez, F., Martin, A., Martinez-Herrera, and Ponz, F. D. (1996). Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic virus), a crucifer tobamovirus infectious on *Arabidopsis thaliana*. *Plant Mol. Biol.* **30,** 191–197.

Ahlquist, P., and Janda, M. (1984). cDNA cloning and in vitro transcription of the complete brome mosaic virus genome. *Mol. Cell. Biol.* **4,** 2876–2882.

Alonso, E., Garcia-Luque, I., de la Cruz, A., Wicke, B., Avila-Roncon, M. J., Serra, M. T., Castresana, C., and Diaz-Ruiz, J. R. (1991). Nucleotide sequence of the genomic RNA of pepper mild mottle virus, a resistance-breaking tobamovirus in pepper. *J. Gen. Virol.* **72,** 2875–2884.

Balmori, E., Gilmer, D., Richards, K., Guilley, H., and Jonard, G. (1993). Mapping the promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3. *Biochimie* **75,** 517–521.

Beachy, R. N., and Zaitlin, M. (1977). Characterization and in vitro translation of the RNAs from less-than-full-length, virus-related, nucleoprotein rods present in tobacco mosaic virus preparations. *Virology* **81,** 160–169.

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

Brown, D., and Gold, L. (1996). RNA replication by Qβ replicase: A working model. *Proc. Natl. Acad. Sci. USA* **93,** 11558–11562.

Culver, J. N., Lehto, K., Close, S. M., Hilf, M. E., and Dawson, W. O. (1993). Genomic position affects the expression of tobacco mosaic virus movement and coat protein genes. *Proc. Natl. Acad. Sci. USA* **90,** 2055–2059.

Dawson, W. O., Beck, D. L., Knorr, D. A., and Grantham, G. L. (1986). cDNA cloning of the complete genome of tobacco mosaic virus and production of infectious transcripts. *Proc. Natl. Acad. Sci. USA* **83,** 5043–5047.

Dawson, W. O., Lewandowski, D. J., Hilf, M. E., Bubrick, P., Raffo, A. J., Shaw, J. J., Grantham, G. L., and Desjardins, P. R. (1989). A tobacco mosaic virus-hybrid expresses and loses an added gene. *Virology* **172,** 285–292.

De Rijk, P., and De Wachter, R. (1997). RnaViz, a program for the visualization of RNA secondary structure. *Nucleic Acids Res.* **25,** 4679–4684.

Deom, C. M., He, X. Z., Beachy, R. N., and Weissinger, A. K. (1994). Influence of heterologous tobamovirus movement protein and chimeric-movement protein genes on cell-to-cell and long-distance movement. *Virology* **205,** 198–209.

Donson, J., Kearney, C. M., Hilf, M. E., and Dawson, W. O. (1991). Systemic expression of a bacterial gene by a tobacco mosaic virus-based vector. *Proc. Natl. Acad. Sci. USA* **88,** 7204–7208.

French, R., and Ahlquist, P. (1988). Characterization and engineering of sequences controlling *in vivo* synthesis of brome mosaic virus subgenomic RNA. *J. Virol.* **62,** 2411–2420.

Gargouri, R., Joshi, R. L., Bol, J. F., Astier-Manifacier, S., and Haenni, A. L. (1989). Mechanism of synthesis of turnip yellow mosaic virus coat protein subgenomic RNA in vivo. *Virology* **171,** 386–393.

Gera, A., Deom, C. M., Donson, J., Shaw, J. J., Lewandowski, D. J., and Dawson, W. O. (1995). Tobacco mosaic tobamovirus does not require concomitant synthesis of movement protein during vascular transport. *Mol. Plant Microbe Interact.* **8,** 784–787.

Giesman-Cookmeyer, D., Silver, S., Vaewhongs, A. A., Lommel, S. A., and Deom, C. M. (1995). Tobamovirus and dianthovirus movement proteins are functionally homologous. *Virology* **213,** 38–45.

Goelet, P., Lomonosoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J., and Karn, J. (1982). Nucleotide sequence of tobacco mosaic virus RNA. *Proc. Natl. Acad. Sci. USA* **79,** 5818–5822.

Grakoui, A., Levis, R., Raju, R., Huang, H. V., and Rice, C. M. (1989) A

cis-acting mutation in the Sindbis virus junction region which affects subgenomic RNA synthesis. *J. Virol.* **63,** 5216–5227.

Guilley, H., Jonard, G., Kukla, B., and Richards, K. E. (1979). Sequence of 1000 nucleotides at the 3′ end of tobacco mosaic virus RNA. *Nucleic Acids Res.* **6,** 1287–1308.

Haasnoot, P. C., Brederode, F. T., Olsthoorn, R. C., and Bol, J. F. (2000). A conserved hairpin structure in Alfamovirus and Bromovirus subgenomic promoters is required for efficient RNA synthesis in vitro. *RNA* **6,** 708–716.

Higuchi, R., Krummel, B., and Saiki, R. K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: Study of protein and DNA interactions. *Nucleic Acids Res.* **16,** 7351–7367.

Hunter, T., Jackson, R., and Zimmern, D. (1983). Multiple proteins and subgenomic mRNAs may be derived from a single open reading frame on tobacco mosaic virus RNA. *Nucleic Acids Res.* **11,** 801–821.

Ikeda, R., Watanabe, E., Watanabe, Y., and Okada, Y. (1993). Nucleotide sequence of tobamovirus Ob which can spread systemically in N gene tobacco. *J. Gen. Virol.* **74,** 1939–1944.

Ishikawa, M., Meshi, T., Watanabe, Y., and Okada, Y. (1988). Replication of chimeric tobacco mosaic viruses which carry heterologous combinations of replicase genes and 3′ noncoding regions. *Virology* **164,** 290–293.

Jackson, A. O., Zaitlin, M., Siegel, A., and Francki, R. I. (1972). Replication of tobacco mosaic virus. 3. Viral RNA metabolism in separated leaf cells. *Virology* **48,** 655–665.

Johnston, J. C., and Rochon, D. M. (1995). Deletion analysis of the promoter for the cucumber necrosis virus 0.9-kb subgenomic RNA. *Virology* **214,** 100–109.

Joshi, S., Pleij, C. W. A., Haenni, A. L., Chapeville, F., and Bosch, L. (1983). Properties of the tobacco mosaic virus intermediate RNA-2 and its translation. *Virology* **127,** 100–111.

Koev, G., Mohan, B. R., and Miller, W. A. (1999). Primary and secondary structural elements required for synthesis of barley yellow dwarf virus subgenomic RNA1. *J. Virol.* **73,** 2876–2885.

Koev, G., and Miller, W. A. (2000). A positive-strand RNA virus with very different subgenomic RNA promoters. *J. Virol.* **74,** 5988–5996.

Koonin, E. V., and Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28,** 375–430.

Kumagai, M., Turpen, T., Weinzettl, N., dellaCioppa, G., Turpen, A., Donson, J., Hilf, M., Grantham, G., Dawson, W., Chow, T., Paitak, M., Jr., and Grill, L. (1993). Rapid, high-level expression of biologically active alpha-trichosanthin in transfected plants by an RNA viral vector. *Proc. Natl. Acad. Sci. USA* **90,** 427–430.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.

Lartey, R. T., Voss, T. C., and Melcher, U. (1995). Completion of a cDNA sequence from a tobamovirus pathogenic to crucifers. *Gene* **166,** 331–332.

Lehto, K., Grantham, G. L., and Dawson, W. O. (1990). Insertion of sequences containing the coat protein subgenomic RNA promoter and leader in front of the tobacco mosaic virus 30K ORF delays its expression and causes defective cell-to-cell movement. *Virology* **174,** 145–157.

Lewandowski, D. J., and Dawson, W. O. (1998). Deletion of internal sequences results in tobacco mosaic virus defective RNAs that accumulate to high levels without interfering with replication of the helper virus. *Virology* **251,** 427–437.

Lewandowski, D. J., and Dawson, W. O. (2000). Functions of the 126- and 183-kDa proteins of tobacco mosaic virus. *Virology* **271,** 90–98.

Marsh, L. E., Dreher, T. W., and Hall, T. C. (1988). Mutational analysis of the core and modulator sequences of the BMV RNA3 subgenomic promoter. *Nucleic Acids Res.* **16,** 981–995.

Miller, W. A., Dreher, T. W., and Hall, T. C. (1985). Synthesis of brome mosaic virus subgenomic RNA in vitro by internal initiation on (−)-sense genomic RNA. *Nature* **313,** 68–70.

Moser, O., Gagey, M. J., Godefroy-Colburn, T., Ellwart-Tschurtz, M., Nitchko, H., and Mundry, K. W. (1988). The fate of the transport protein of tobacco mosaic virus in systemic and hypersensitive tobacco hosts. *J. Gen. Virol.* **69,** 1367–1373.

Ogawa, M., and Sakai, F. (1984). A messenger RNA for tobacco mosaic virus coat protein in infected tobacco mesophyll protoplasts. II. Time course of its synthesis. *Phytopathol. Z.* **109,** 193–203.

Ohno, T., Aoyagi, M., Yamanashi, Y., Saito, H., Ikawa, S., Meshi, T., and Okada, Y. (1984). Nucleotide sequence of the tobacco mosaic virus (tomato strain) genome and comparison with the common strain genome. *J. Biochem.* **96,** 1915–1923.

Ooshika, I., Watanabe, Y., Meshi, T., Okada, Y., Igano, K., Inouye, K., and Yoshida, N. (1984). Identification of the 30K protein of TMV by immunoprecipitation with antibodies directed against a synthetic peptide. *Virology* **132,** 71–78.

Ryu, K. H., and Park, W. M. (1995). The complete nucleotide sequence and genome organization of Odontoglossum ringspot tobamovirus RNA. *Arch. Virol.* **140,** 1577–1587.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning," 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sawicki, S. G., and Sawicki, D. L. (1998). A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440,** 215–219.

Shivprasad, S., Pogue, G. P., Lewandowski, D. J., Hidalgo, J., Donson, J., Grill, L. K., and Dawson, W. O. (1999). Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. *Virology* **255,** 312–323.

Siegel, A., Hari, V., Montgomery, I., and Kolacz, K. (1976). A messenger RNA for capsid protein isolated from tobacco mosaic virus-infected tissue. *Virology* **73,** 363–371.

Siegel, A., Hari, V., and Kolacz, K. (1978). The effect of tobacco mosaic virus infection on host and virus-specific protein synthesis in protoplasts. *Virology* **85,** 494–503.

Siegel, R. W., Adkins, S., and Kao, C. C. (1997). Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proc. Natl. Acad. Sci. USA* **94,** 11238–11243.

Siegel, R. W., Bellon, L., Beigelman, L., and Kao, C. C. (1998). Moieties in an RNA promoter specifically recognized by a viral RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. USA* **95,** 11613–11618.

Sit, T. L., Vaewhongs, A. A., and Lommel, S. A. (1998). RNA-mediated trans-activation of transcription from a viral RNA. *Science* **281,** 829–832.

Skulachev, M. V., Ivanov, P. A., Karpova, O. V., Korpela, T., Rodionova, N. P., Dorokhov, Y. L., and Atabekov, J. G. (1999). Internal initiation of translation directed by the 5′-untranslated region of the tobamovirus subgenomic RNA I_2. *Virology* **263,** 139–154.

Smirnyagina, E., Hsu, Y. H., Chua, N., and Ahlquist, P. (1994). Second-site mutations in the brome mosaic virus RNA3 intercistronic region partially suppress a defect in coat protein mRNA transcription. *Virology* **198,** 427–436.

Solis, I., and Garcia-Arenal, F. (1990). The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus. *Virology* **177,** 553–558.

Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B. A., and Siddell, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* **2,** 1839–1844.

Stawicki, S. S., and Kao, C. C. (1999). Spatial perturbations within an RNA promoter specifically recognized by a viral RNA-dependent RNA polymerase (RdRp) reveal that RdRp can adjust its promoter binding sites. *J. Virol.* **73,** 198–204.

Sulzinski, M. A., Gabard, K. A., Palukatis, P., and Zaitlin, M. (1985). Replication of tobacco mosaic virus. VIII. Characterization of a third subgenomic TMV RNA. *Virology* **145,** 132–140.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.

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

van der Kuyl, A. C., Neeleman, L., and Bol, J. F. (1991). Deletion analysis of cis- and trans-acting elements involved in replication of alfalfa mosaic virus RNA 3 in vivo. *Virology* **183,** 687–694.

van der Vossen, E. A., Notenboom, T., and Bol, J. F. (1995). Characterization of sequences controlling the synthesis of alfalfa mosaic virus subgenomic RNA in vivo. *Virology* **212,** 663–672.

Wang, J., and Simon, A. E. (1997). Analysis of the two subgenomic RNA promoters for turnip crinkle virus in vivo and in vitro. *Virology* **232,** 174–186.

Wang, J., Carpenter, C. D., and Simon, A. E. (1999). Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. *Virology* **253,** 327–336.

Watanabe, Y., Meshi, T., and Okada, Y. (1984). The initiation site for transcription of the TMV 30-kDa protein messenger RNA. *FEBS Lett.* **173,** 247–250.

Zavriev, S. K., Hickey, C. M., and Lommel, S. A. (1996). Mapping of the red clover necrotic mosaic virus subgenomic RNA. *Virology* **216,** 407–410.

Zimmern, D. (1975). The 5′ end group of tobacco mosaic virus RNA is m7G5′ppp5′Gp. *Nucleic Acids Res.* **2,** 1189–1201.

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science* **244,** 48–52.