RNA (1999), *5*:550–561. Cambridge University Press. Printed in the USA.

Subgenomic mRNA regulation by a distal RNA element in a (+)-strand RNA virus

GUICHANG ZHANG, VIOLETTA SLOWINSKI and K. ANDREW WHITE
Department of Biology, York University, Toronto, Ontario M3J 1P3, Canada

ABSTRACT

Subgenomic (sg) mRNAs are synthesized by (+)-strand RNA viruses to allow for efficient translation of products encoded 3′ in their genomes. This strategy also provides a means for regulating the expression of such products via modulation of sg mRNA accumulation. We have studied the mechanism by which sg mRNAs levels are controlled in tomato bushy stunt virus, a small (+)-strand RNA virus which synthesizes two sg mRNAs during infections. Neither the viral capsid nor movement proteins were found to play any significant role in modulating the accumulation levels of either sg mRNA. Deletion analysis did, however, identify a 12-nt-long RNA sequence located approximately 1,000 nt upstream from the site of initiation of sg mRNA2 synthesis that was required specifically for accumulation of sg mRNA2. Further analysis revealed a potential base-pairing interaction between this sequence and a sequence located just 5′ to the site of initiation for sg mRNA2 synthesis. Mutant genomes in which this interaction was either disrupted or maintained were analyzed and the results indicated a positive correlation between the predicted stability of the base-pairing interaction and the efficiency of sg mRNA2 accumulation. The functional significance of the long-distance interaction was further supported by phylogenetic sequence analysis which revealed conservation of base-pairing interactions of similar stability and relative position in the genomes of different tombusviruses. It is proposed that the upstream sequence represents a *cis*-acting RNA element which facilitates sg mRNA accumulation by promoting efficient synthesis of sg mRNA2 via a long-distance RNA–RNA interaction.

Keywords: *cis*-element; gene expression; RNA–RNA interaction; RNA polymerase; RNA secondary structure; RNA synthesis; TBSV; tombusvirus; transcription

INTRODUCTION

Eukaryotic (+)-strand RNA viruses employ a wide variety of strategies for expression and regulation of their encoded genes (Maia et al., 1996). These include genome segmentation, polyprotein processing, translational frameshifting, termination codon suppression, and subgenomic (sg) mRNAs synthesis. The latter strategy represents an important mechanism used commonly for the expression of proteins encoded 3′ in polycistronic genomes. Translational events involving such genomes generally conform to the scanning model for translational initiation (Kozak, 1978); thus 5′-most open reading frames (ORFs) are translated efficiently whereas those downstream are not. To circumvent this problem, sg mRNAs, which represent 5′-truncated 3′-coterminal versions of the genome, are synthesized during infections. In these messages, ORFs originally located downstream in the genome are positioned 5′-proximally, thereby allowing for their efficient translation. Consequently, the timing of production and level of accumulation of these proteins can be regulated via modulating sg mRNA synthesis and/or stability.

A large number of diverse (+)-strand RNA viruses, including those of animals (e.g., alphaviruses, rubiviruses, coronaviruses, and nodaviruses) and plants (e.g., bromoviruses, potexviruses, closteroviruses, and carmoviruses) synthesize sg mRNAs (reviewed in Koonin, 1991a). Several models involving viral RNA-dependent RNA polymerases (RdRp) and *cis*-elements have been proposed to describe how sg mRNAs are generated. These include (1) “internal initiation” of sg mRNA synthesis on a full-length minus strand of the genome (Miller et al., 1985); (2) “discontinuous transcription,” in which a short leader sequence corresponding to the 5′ terminus of the genome is fused via RdRp-mediated recombination to the body of a sg mRNA (Spaan et al., 1983; Makino et al., 1986; Sawicki & Sawicki, 1998); and (3) “premature termination” during (−)-strand synthesis of the genome, followed by use of the truncated nascent RNA as a template for sg mRNA synthesis (Miller et al., 1997; Sit et al., 1998).

Reprint requests to: K. Andrew White, Department of Biology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada; e-mail: kawhite@yorku.ca.

Early evidence for the internal initiation model came from studies on brome mosaic virus (BMV) that demonstrated sg mRNA promoter (Psg) activity in vitro from a (−)-strand template (Miller et al., 1985). Studies carried out on alfalfa mosaic virus (van der Kuyl et al., 1990) and turnip crinkle virus (Wang & Simon, 1997) have also provided important experimental support for this model in other viral systems. Recently, the BMV core Psg was shown to be recognized principally on the basis of primary structure, a feature that appears to extend to animal-infecting members of the alphavirus-like superfamily (Siegel et al., 1997). This sequence-dependent specificity, in addition to other features, has prompted mechanistic analogies between viral RdRps and DNA-dependent RNA polymerases (Adkins et al., 1998).

For the discontinuous transcription model, the fusion of 5′ genomic (or leader) sequences with the body of sg mRNAs is well established (Lai et al., 1994), and *cis*-acting sequences and cellular proteins important for this activity have been identified (Joo & Makino, 1992; Jeong et al., 1996; Lin et al., 1996; Li et al., 1997). Although significant progress has been made toward understanding the mechanism of discontinuous transcription, general aspects of this process still remain unresolved. For instance, it has not yet been established whether the fusion event occurs during plus-strand or minus-strand synthesis (Sawicki & Sawicki, 1998). In addition, a debate continues as to whether the sg mRNAs synthesized can be amplified further via replication (Sethna & Brian, 1997; An & Makino, 1998).

At present, only weak experimental evidence exists for the premature-termination model. The most encouraging data come from recent studies on red clover necrotic mosaic *Dianthovirus* (RCNMV) (Sit et al., 1998). This virus possesses a bipartite genome consisting of RNA1 and RNA2, and an sg mRNA is synthesized from the former (Xiong & Lommel, 1991). Interestingly, it was found that synthesis of the sg mRNA required *trans*-activation by an RNA hairpin structure located in RNA2 (Sit et al., 1998). A critical factor in the *trans*-activation event was base pairing of 8 nt in the loop of the hairpin with a complementary sequence in RNA1 located just upstream of the initiation site for sg mRNA synthesis. It was proposed that this interaction physically blocks negative strand synthesis on the RNA 1 template, thereby promoting premature termination (Sit et al., 1998); the involvement of a protein factor(s) in this process, however, was not excluded. The truncated nascent (−)-strand RNA generated would then serve as a template for the synthesis of sg mRNA. Additional support for the importance of such RNA–RNA interactions comes from studies on the *Potexvirus* potato virus X (PVX) where the potential for base-pairing interactions between the 5′ end of the genome and sequences upstream of the start site of the two sg mRNAs were identified (Kim & Hemenway, 1997). It is likely that the proposed interactions are functionally relevant because disruptive modifications introduced into either the 5′-terminal or internal segments reduced the level of sg mRNA accumulation (Kim & Hemenway, 1996, 1997). A recent preliminary report further supports this notion by providing additional experimental evidence for the interaction (Kim & Hemenway, 1998).

Tomato bushy stunt virus (TBSV), the prototype member of the genus *Tombusvirus*, has a 4.8-kb (+)-strand RNA genome encoding five functional ORFs (Hearne et al., 1990; Fig. 1). The 5′ proximally encoded viral replication components, p33 and p92 (Oster et al., 1998), are translated from the genome, whereas products encoded more 3′ in the genome are translated from two sg mRNAs (Hearne et al., 1990; Fig. 1). The p41 product is the capsid protein (CP) and it is translated from the larger 2.1-kb sg mRNA1. The overlapping products p22 and p19 are movement proteins that facilitate local and systemic spread of the infection in plants (Scholthof et al., 1995), respectively, and are both translated from the 0.9-kb sg mRNA2. Only limited information exists regarding the sequences involved in synthesis of sg mRNAs for members of this genus. Studies on cucumber necrosis *Tombusvirus* (CNV) have defined sequences necessary for synthesis of sg mRNA2 to a region approximately 20 nt upstream and 6 nt downstream (−20/+6) from the site of initiation (Johnston & Rochon, 1995). The requirement for additional viral sequences and the mechanism by which tombusvirus Psgs function remain unknown. Here we report the identification and analysis of a novel distal RNA element in the TBSV genome that facilitates efficient accumulation of sg mRNA2.

RESULTS

Viral capsid and movement proteins do not influence sg mRNA accumulation

Figure 2 shows a typical accumulation profile of viral RNAs during an infection of protoplasts with T100 tran-

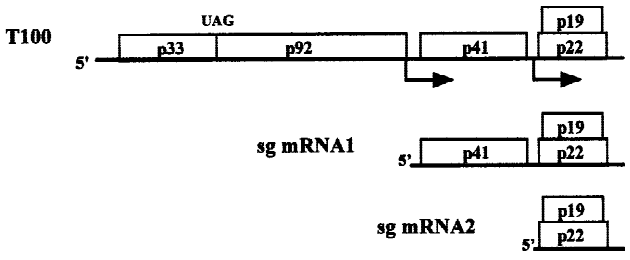

FIGURE 1. Schematic representation of the TBSV genome. The genome (T100) is shown at the top as a horizontal line, with coding regions depicted as boxes that include the approximate molecular mass values (in thousands) of the encoded proteins (Hearne et al., 1990). Arrows indicate the approximate positions of the initiation sites for sg mRNA synthesis within the genome. The two sg mRNAs synthesized during infections are depicted below.

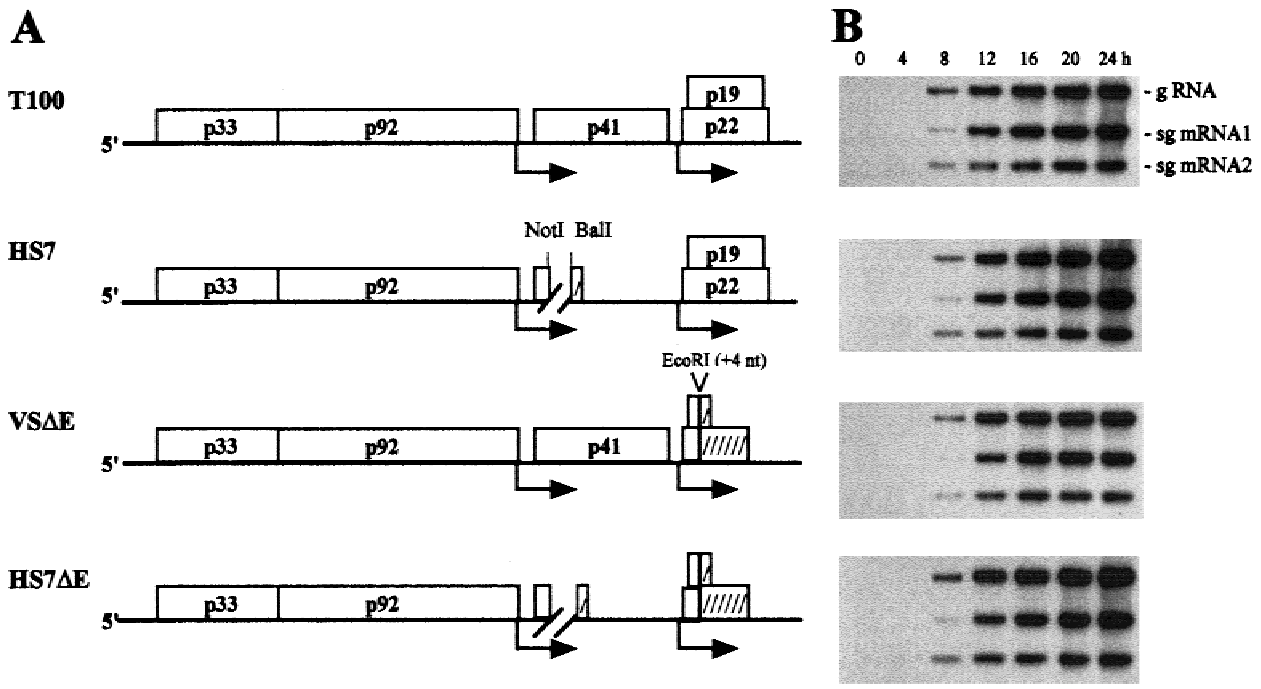

FIGURE 2. **A**: Schematic representation of the mutant genomes in which the capsid protein and/or movement proteins have been inactivated. For comparison the wild-type genome (T100) is shown at the top. The capsid protein mutant, HS7, contains a deletion (shown as a gap) inactivating the capsid protein ORF via a frameshift (hatched box). The movement protein mutant, VSΔE, contains a 4-nt insertion (indicated by a V) that inactivates both movement protein ORFs via frameshifts. The double mutant, HS7ΔE, contains the modifications in HS7 and VSΔE. **B**: Northern blot analysis showing progeny viral RNA accumulation kinetics for capsid protein and movement protein mutant genomes. The identity of the transcript used in the infection is shown to the left in **A** while the positions of various viral RNAs are indicated on the right. The times of isolation of the nucleic acids postinoculation are indicated above the lanes in hours. Total nucleic acids were isolated from infected protoplasts, separated in a 1.4% agarose gel, transferred to nylon membrane, and hybridized with a [^{32}P]-end-labeled oligonucleotide probe (P9) complementary to the 3′-terminal 23 nt of the genome.

scripts corresponding to the wild-type (wt) TBSV genome. Both sg mRNAs were detectable as early as 8 h postinoculation and their initial appearance coincided with that of the genome. At the 8-h time point, the level of sg mRNA2 tended to be greater than that of sg mRNA1; however this situation was reversed at later times. This accumulation profile is in agreement with those reported for other tombusviruses (Tavazza et al., 1994; Johnston & Rochon, 1995) and is consistent with the concept that the movement proteins (p19 & p22) are required early in the infection to promote invasion of adjacent cells whereas capsid protein (p41) is necessary at a later stage for encapsidation of progeny genomes (Johnston & Rochon, 1995).

To determine if viral proteins other than those involved in RNA replication play any role in regulating the synthesis of sg mRNAs, mutant viral genomes were constructed in which capsid protein and/or the movement proteins were rendered nonfunctional by a series of frameshifts and/or deletions in their corresponding ORFs. The inactivation of the encoded products was confirmed in whole-plant assays in which genomes with mutated capsid protein or movement protein ORFs were unable to mediate assembly (i.e., no virus was detected) or invade adjacent cells (i.e., no lesions were observed), respectively (data not shown). Inactivation of these proteins either independently in HS7 and VSΔE or concurrently in HS7ΔE did not result in any consistent alteration in the accumulation profiles of the sg mRNAs (Fig. 2), suggesting no significant role for these products in the regulation of sg mRNA accumulation.

An upstream sequence influences sg mRNA2 accumulation

To test whether the production of the two sg mRNAs was coupled, the synthesis of each was inactivated independently. This was accomplished by mutating the respective promoter regions using the defined −20/+6 core promoter of CNV as a guide. Alterations were introduced into T100 that either removed or modified the sequences surrounding the initiation sites for the two sg mRNAs (Fig. 3A). These modifications also disrupted the capsid protein and movement protein ORFs; however we have already shown that the encoded proteins do not influence sg mRNA accumulation (Fig. 2). Because the predicted position for Psg1 overlaps the ORF for p92, a combination of five nucleotide substitutions within the coding region (designed to not alter coding identity) and a deletion of a downstream seg-

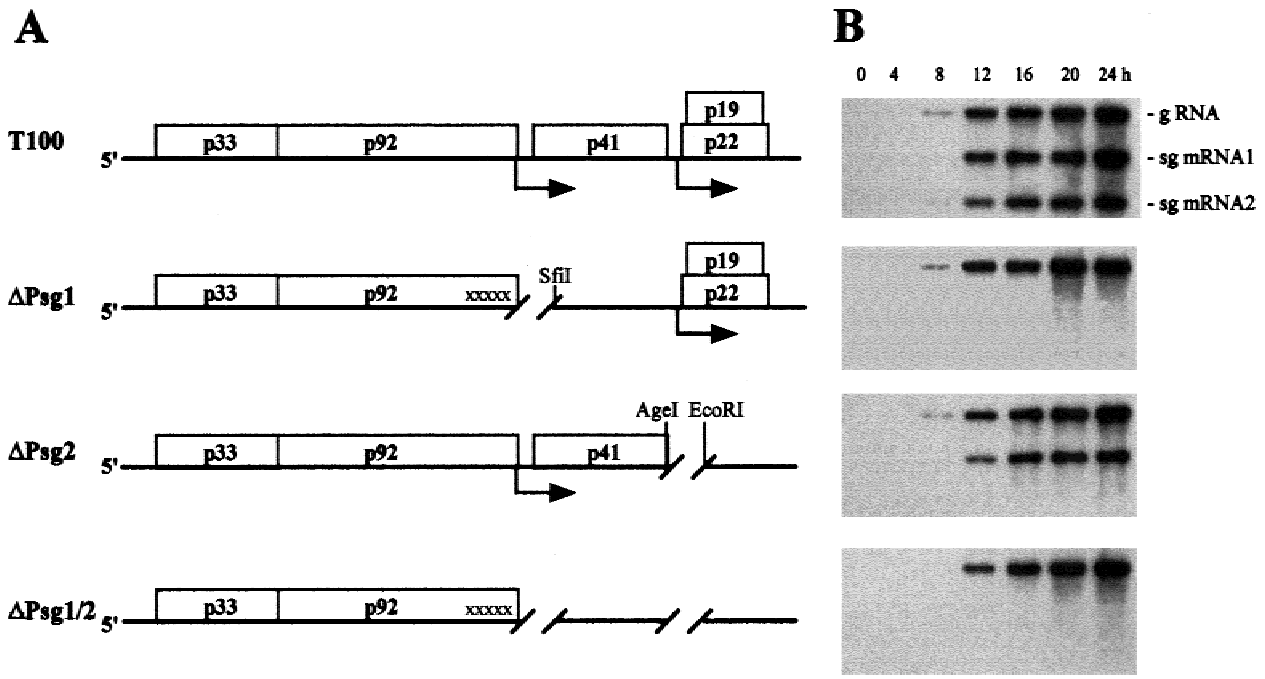

FIGURE 3. **A**: Schematic representation of the mutant genomes in which sg mRNA synthesis has been inactivated. For comparison the wild-type genome (T100) is shown at the top. The Psg1 mutant, ΔPsg1, contains five nucleotide substitutions within the p92 coding region (each represented as an x) and a deletion of a downstream segment (shown as a gap). The Psg2 mutant, ΔPsg2, contains a deletion spanning the initiation site of sg mRNA2. The double mutant ΔPsg1/2 contains the modifications in ΔPsg1 and ΔPsg2. **B**: Northern blot analysis showing progeny viral RNA accumulation kinetics for Psg1 and Psg2 mutant genomes. The identity of the transcript used in the infection is shown to the left in **A** and the positions of various viral RNAs are indicated on the right. The times of isolation of the nucleic acids postinoculation are indicated above the lanes in hours. Total nucleic acids were isolated and analyzed as described in the legend to Figure 2.

ment were used to inactivate it. Interestingly, in ΔPsg1 (Fig. 3A), the inactivation of sg mRNA1 synthesis resulted in a corresponding dramatic decrease in accumulation of sg mRNA2, suggesting possible coupling of these processes (Fig. 3B). Alternatively, the inactivation of sg mRNA2 synthesis in ΔPsg2 (Fig. 3A), via deletion of a segment spanning its site of initiation, did not notably influence the accumulation levels of sg mRNA1 (Fig. 3B). Neither of the sg mRNAs were detectable when both promoters were inactivated simultaneously in ΔPsg1/2 (Fig. 3A,B).

To determine if the mutation in ΔPsg1 was functioning in a dominant negative manner, a 161-nt segment containing a functional Psg1 (+100/-61) was introduced into ΔPsg1 at an *Msc*I site located downstream of the existing modifications, thus generating ΔPsg1+1 (Fig. 4A). The inserted promoter mediated the accumulation of a predicted sg mRNA1-like molecule (sg mRNA1*) and restored the accumulation of sg mRNA2 substantially (Fig. 4). Insertion of the same Psg1 promoter segment at a more 3′ position (i.e., the *Age*I site; Fig. 4A) also led to accumulation of a predicted sg mRNA1-like molecule and restoration of sg mRNA2 accumulation (data not shown). These results indicate that the defect is not dominant and suggest that the absence of sg mRNA1 and/or the modifications in ΔPsg1 are inhibitory to sg mRNA2 accumulation. To address these possibilities, a mutant genome containing only the five nucleotide substitutions was constructed (Psg1S5; Fig. 4A). When Psg1S5 was tested, efficient sg mRNA2 accumulation was observed, but no sg mRNA1 accumulation was detected at any of the time points analyzed (Fig. 4B). This result demonstrates that efficient accumulation of sg mRNA2 does not require the accumulation of sg mRNA1 and implicates the segment deleted in ΔPsg1 as an important element for sg mRNA2 accumulation.

Identification of RNA sequences important for sg mRNA2 accumulation

Deletion mutants were constructed to further delineate the region responsible for the increased accumulation levels of sg mRNA2 (Fig. 5A). Analysis of the mutant genomes Psg1-19, Psg1-20, and Psg1-21 indicated that a 14-nt sequence from the intergenic region between the p92 and capsid protein ORFs was essential for this activity (Fig. 5A,B, cf. Psg1-19 with Psg1-20). One possible way in which this sequence could influence sg mRNA2 accumulation is by modulating its synthesis. If this were the case, one would predict some form of communication between the regulatory element and Psg2. Interestingly, analysis of the region just 5′ to the initiation site for sg mRNA2 synthesis revealed a 12-nt

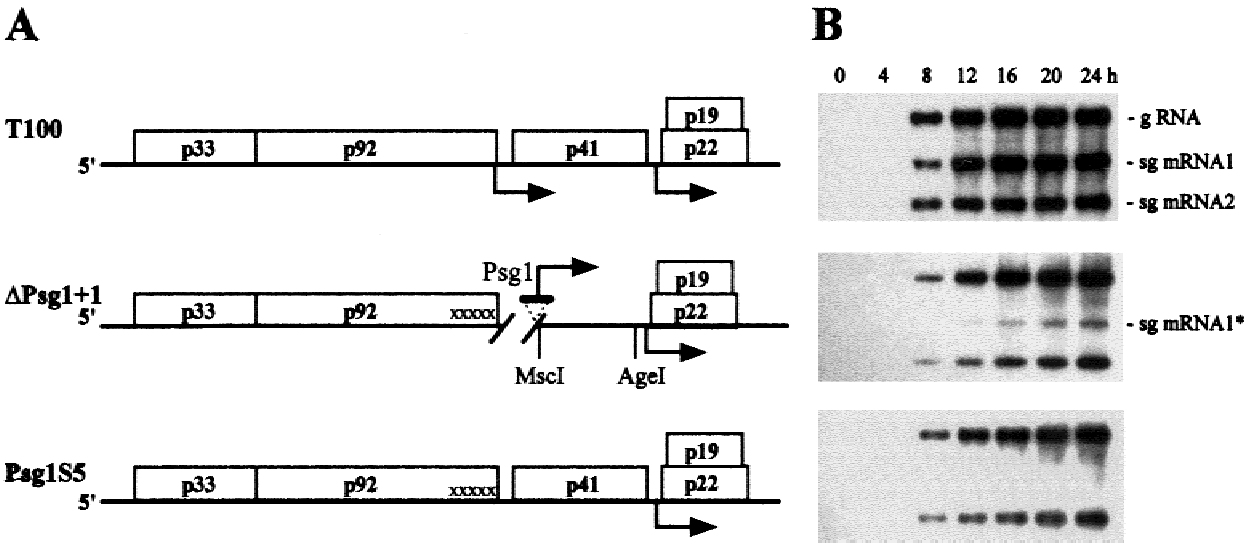

FIGURE 4. A: Schematic representation of the mutant genomes ΔPsg1+1 and Psg1S5. For comparison the wild-type genome (T100) is shown at the top. Mutant ΔPsg1+1 is a derivative of ΔPsg1 that contains a sequence insertion that harbors a functional Psg1. Mutant Psg1S5 contains 5-nt substitutions within the p92 coding region. **B**: Northern blot analysis showing progeny viral RNA accumulation kinetics for ΔPsg1+1 and Psg1S5 mutant genomes. The identity of the transcript used in the infection is shown to the left and the positions of various viral RNAs are indicated on the right (the sg mRNA1-like molecule generated by ΔPsg1+1 is indicated by an asterisk). The times of isolation of the nucleic acids postinoculation are indicated above the lanes in hours. Total nucleic acids were isolated and analyzed as described in the legend to Figure 2.

sequence (in bold; see Psg2-26 in Fig. 5A) in which 11 of the 12 nt were complementary to residues in the delineated upstream sequence. When this complementary downstream sequence was deleted independently in Psg2-27, no sg mRNA2 accumulation was detected (Fig. 5B), indicating that sg mRNA2 accumulation requires the presence of both complementary segments. The relative accumulation levels of sg mRNA2 were quantified and are presented in Figure 5C. The notably lower levels of sg mRNA2 observed for Psg1-20, Psg1-21 and Psg2-26, as compared with wt T100, are addressed in the last section of Results.

Maintenance of base-pairing potential is required for efficient sg mRNA2 accumulation

To test whether the putative base-pairing interaction between the upstream and downstream 12-nt segments was functionally relevant, mutants were constructed in which the predicted interaction was disrupted (Fig. 6A). To facilitate the introduction of such modifications, the construct Psg20/26, containing wild-type upstream and downstream elements along with the deletions found in Psg1-20 and Psg2-26, was generated (Fig. 6A). When tested, the levels of sg mRNA1 and sg mRNA2 generated from Psg20/26 were similar to those observed for T100 (Fig. 6B). In Psg28/26, modifications were introduced into the upstream segment that were predicted to disrupt the base-pairing interaction (Fig. 6A). The effect of this modification was to significantly reduce sg mRNA2 accumulation (Fig. 6B). A similar result was observed for Psg20/29 (Fig. 6B), in which base pair-disrupting substitutions were introduced into the downstream segment (Fig. 6A). However, when the upstream and downstream substitution mutations were combined in Psg28/29 (Fig. 6A), so as to restore base-pairing potential, sg mRNA2 accumulation was restored to levels similar to that observed for Psg20/26 (Fig. 6B). These results were further confirmed by quantifying the relative accumulation levels of sg mRNA2 (Fig. 6C). The notable correlation between regained activity and regenerated base-pairing potential supports the functional importance of the proposed long-distance interaction.

Conservation of a long-distance RNA–RNA interaction in tombusviruses

Alignment of the sequences spanning the initiation sites for sg mRNA2 synthesis from different tombusviruses revealed a high degree of identity near the initiation site (Fig. 7A). However, regions located more 5′, which include the TBSV downstream segment implicated in the base-pairing interaction (Fig. 7A, underlined), displayed significantly less identity. Further analysis of the TBSV sequences near the upstream and downstream elements (Fig. 7B) revealed the possibility of an additional base-pairing interaction (DE-B/CE-B; Fig. 7C). The entire upstream segment has been defined as the distal element (DE) and the entire downstream complementary segment as the core element (CE; Fig. 7B,C). Accordingly, the segments within these regions that

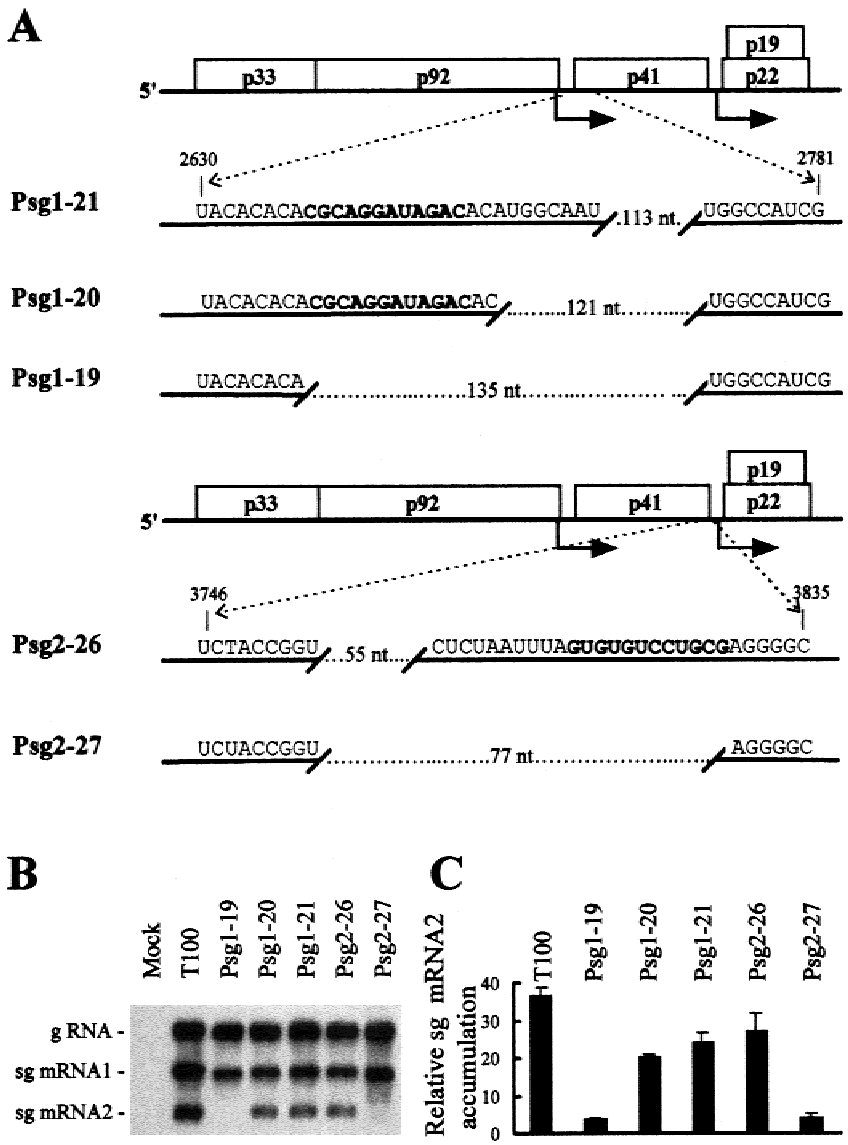

FIGURE 5. **A**: Schematic representation of mutant genomes containing deletions that define sequences important for efficient sg mRNA2 accumulation. The region deleted in each mutant genome is shown in relationship to the wild-type genome and is represented by a gap with the number of residues deleted. The upstream (Psg1-21 and Psg1-20) and downstream (Psg2-26) sequences that are complementary to each other are shown in bold. **B**: Northern blot analysis showing progeny viral RNA accumulation for various deletion mutants. The identity of the transcript used in the infection is shown above each lane and the positions of various viral RNAs are indicated on the left. Total nucleic acids were isolated and analyzed after a 24-h incubation as described in the legend to Figure 2. **C**: Relative accumulation levels of sg mRNA2. The relative accumulation levels, which represent sg mRNA2 levels standardized to the levels of their corresponding genomes, were quantified by radioanalytical scanning of Northern blots. The values shown represent means (with standard errors) from two separate protoplast infections.

correspond to the two distinct sets of base-pairing interactions are designated as A and B (Fig. 7C). The B regions may play an auxiliary role in sg mRNA2 accumulation since mutants that maintain both DE-A and CE-A but lack DE-B (Psg1-20, Psg1-21; Fig. 5C) or CE-B (Psg2-26; Fig. 5C) or both DE-B and CE-B (Psg20/26; Fig. 6C) showed reduced levels of sg mRNA2, as compared to T100 (containing all elements). The DE-B/CE-B interaction could facilitate more efficient sg mRNA2 accumulation by stabilizing the essential DE-A/CE-A interaction. Conversely, the sequence just 5′ to DE-A shows limited complementarity with the sequence just 3′ to CE-A (Fig. 7C). Additional studies are underway to address the roles of these flanking sequences.

A search for corresponding distal element and core element segments in the genomes of other sequenced tombusviruses revealed maintenance of the base-pairing potential and of its position relative to the site of initiation (CyRSV being the apparent exception to the latter; Fig. 7D). This conservation supports further the proposed functional interaction between the core element and the distal element. In addition, the maintenance of a significant portion of the relevant interaction in the core promoter defined for CNV (Johnston & Rochon, 1995; Fig. 7D, underlined) is consistent with the proposed function.

DISCUSSION

Our results support a functionally relevant base-pairing interaction between DE-A and CE-A . The data generated from sequence comparisons of different tombusvirus genomes is particularly compelling in that the interaction is preserved despite significant variation in primary sequence (Fig. 7C,D). The experimentally dem-

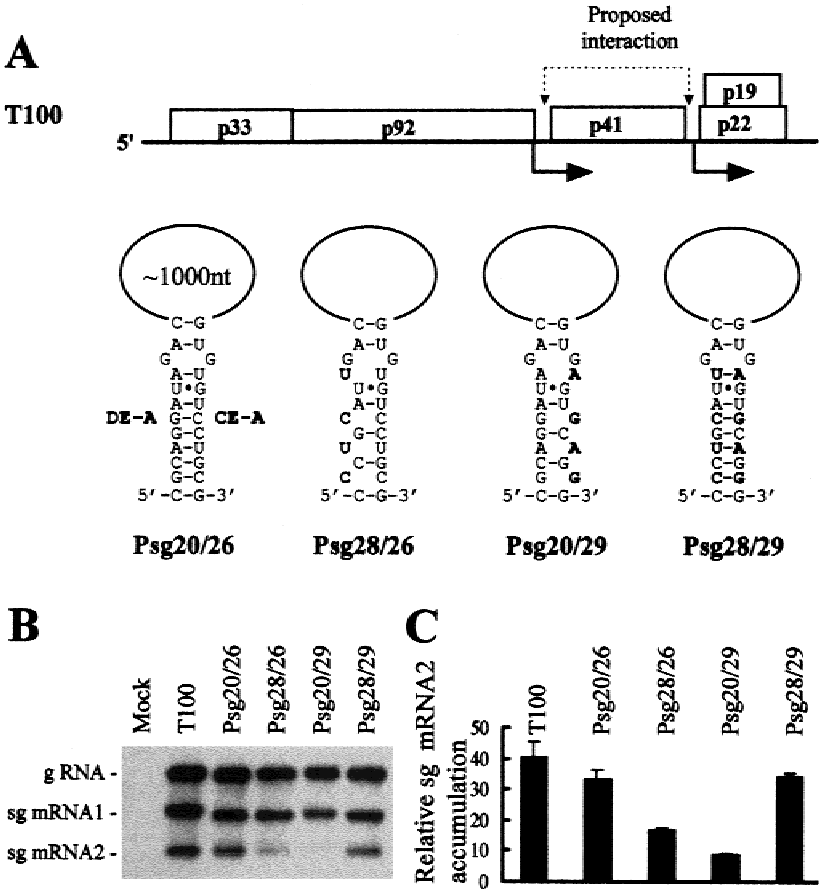

FIGURE 6. **A**: Schematic representation of the wild-type genome showing the relative positions of the sequences proposed to interact via base pairing. Below, the predicted base-pairing interaction between the upstream and downstream sequence (DE-A and CE-A, respectively; see Fig. 7 for details) for each mutant is shown. **B**: Northern blot analysis showing progeny viral RNA accumulation for mutant genomes with modifications in the predicted base-pairing interaction. The identity of the transcript used in the infection is shown above each lane and the positions of various viral RNAs are indicated on the left. Total nucleic acids were isolated and analyzed after a 24-h incubation as described in the legend to Figure 2. **C**: Relative accumulation levels of sg mRNA2. The relative accumulation levels, which represent sg mRNA2 levels standardized to the levels of their corresponding genome, were quantified by radioanalytical scanning of Northern blots. The values shown represent means (with standard errors) from three separate protoplast infections.

onstrated correlation between the predicted stability of the base-pairing interaction and the degree of activity is also supportive of a functional interaction (Fig. 6). For TBSV, the maintenance of this interaction conferred significantly higher levels of sg mRNA2 accumulation, a consequence of either enhanced synthesis or increased stabilization of the message. The latter mechanism seems less likely for four reasons. (1) The destabilization of sg mRNA2 via perturbation of its sequence can be excluded because none of the mutant genomes analyzed were modified 3′ to its initiation site. (2) The functionally relevant interaction occurs in a region outside that of sg mRNA2; therefore stabilization via RNA–RNA interactions between DE-A and sg mRNA2 is unlikely. Also consistent with this concept is the finding that maintenance of DE-A (in the absence of CE-A, i.e., Psg2-27) did not allow for efficient sg mRNA2 accumulation. (3) sg mRNA1, which contains sg mRNA2 within it, was not destabilized in mutant genomes lacking DE-A (i.e., Psg1-19). (4) Although DE-A mutations also perturbed the expression of capsid protein (which could potentially stabilize sg mRNAs), the capsid protein was shown to play no significant role in regulating sg mRNA accumulation. The features of the interaction are instead more consistent with a role in synthesis, and several properties of DE-A are similar to those of DNA enhancers: (1) it acts to stimulate the accumulation of an mRNA; (2) it interacts with elements near the site of initiation of mRNA synthesis; (3) it functions from a distal location; and (4) its activity is maintained when it is repositioned in the genome. Taken together, our data more strongly support DE-A as a *cis*-acting enhancer-like element of sg mRNA2 synthesis whose activity is mediated, at least in part, by RNA–RNA interactions.

How could an RNA–RNA interaction mediate enhanced sg mRNA synthesis? For RCNMV, it has been proposed that a *trans*-interaction between two viral plus-strand RNAs forms a structure that is able to stall the RdRp during minus strand synthesis (Sit et al., 1998). Subsequent premature termination then generates a truncated negative strand that serves as a template for sg mRNA synthesis. A *cis*-acting version of this premature-termination model could also explain the DE-A/CE-A mediated increase in sg mRNA levels. If this were the case, the relevant interaction would occur in the plus strand, as depicted in Figures 7B and 7C. Interestingly, studies on tombusvirus RNA replication and recombination have provided evidence supporting secondary structure-mediated stalling and/or dissociation of the actively copying viral RdRp during minus strand synthesis (White & Morris, 1995; Havelda et al.,

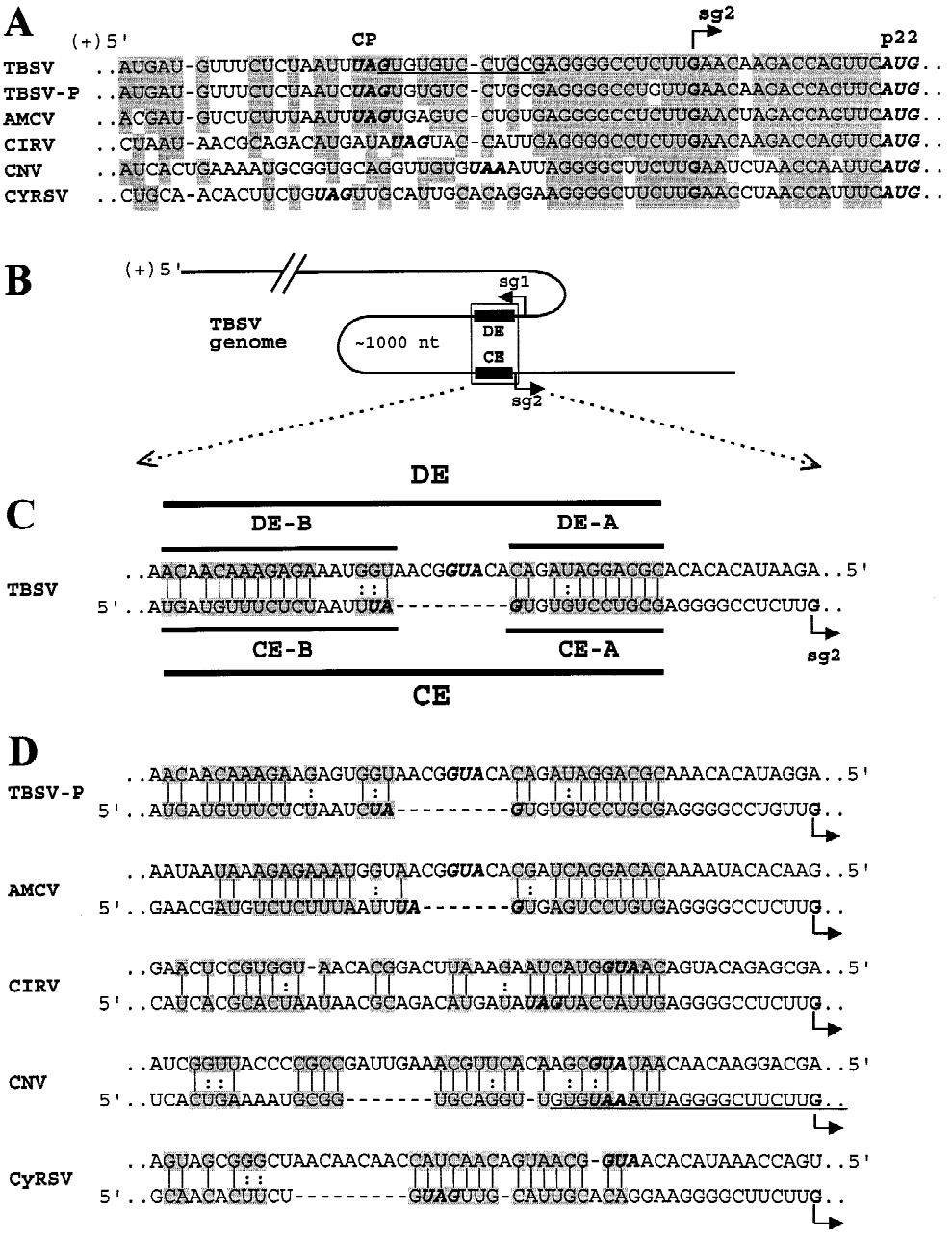

FIGURE 7. **A**: Alignment of RNA sequences corresponding to the sg mRNA2 promoters from different tombusviruses. Residues are shaded at positions where a minimum of 4 of the 6 nt are identical. Single nucleotide gaps were introduced in order to maximize alignment of identical residues. The site of initiation of synthesis of sg mRNA2 is indicated by an arrow labeled sg2, and the termination and initiation codons corresponding to capsid protein and p22, respectively, are in bold italics. For TBSV, the 12-nt sequence predicted to interact with the upstream element is underlined. The sequences presented are those of TBSV (Hearne et al., 1990), TBSV-pepper isolate (TBSV-P; Havelda et al., 1998), artichoke mottle crinkle virus (AMCV; Tavazza et al., 1994), carnation Italian ringspot virus (CIRV; Rubino et al., 1995), CNV (Rochon & Tremaine, 1989), and cymbidium ringspot virus (CyRSV; Grieco et al., 1989). **B**: Schematic representation of the TBSV genome showing the relative positions of the distal element and core element (not to scale). **C**: The distal element and core element in **B** (boxed) have been enlarged to include their corresponding sequences and proposed base-pairing interactions (shaded residues). The distal element and core element are delineated by thick bars while their subregions A and B are defined by the thinner bars. The initiation and termination codons for capsid protein are indicated in bold italics in the upper and lower strands, respectively. **D**: Base-pairing interactions similar to that shown in **C** are presented for various tombusviruses. The sequence underlined corresponds to the minimal promoter defined for CNV (Johnston & Rochon, 1995).

1997: Wu & White, 1998). In addition, negative strands of tombusvirus sg mRNAs have been detected in infections (Tavazza et al., 1994; Ray & White, 1999). Furthermore, examination of the DE-A/CE-A interactions in the various tombusviruses (Fig. 7C,D) reveals the potential for several noncanonical base pairs (e.g., G-U, G-A) that could potentially stabilize or not significantly destabilize the plus strand interaction (Limmer, 1997).

However, in the minus strand, the corresponding pairs could be more disruptive. Collectively, these observations are consistent with a premature-termination-type model; however, more definitive studies addressing this possibility are clearly warranted. It is worth noting that if this type of premature-termination model does apply, the possibility would also exist for a similar DE-A/CE-A interaction within sg mRNA1. However, the relatively efficient production of sg mRNA2 in the absence of sg mRNA1 accumulation (i.e., Psg1S5, Fig. 3) suggests that such sg mRNA1-mediated production would be rather limited.

Another option is that the DE-A/CE-A interaction could itself represent part of a recognition motif for a replication factor. For TCV, a localized secondary structure in the minus strand is believed to be important for sg mRNA promoter activity (Wang & Simon, 1997; Wang et al., 1999). For TBSV, a DE-A/CE-A interaction in the minus strand could fulfill a similar role. Alternatively, the interaction could recruit sequences associated with DE-A to the core promoter. Such sequences could (1) interact with proteins already bound to the core promoter, (2) independently bind and recruit proteins to the core promoter, or (3) interact with other RNA elements important for sg mRNA synthesis. For instance, the DE-A/CE-A interaction would bring Psg1 sequences close to the initiation site for sg mRNA2, a role more consistent with the positioning of the element. However, the production of sg mRNA2 from a genome with an inactive Psg1 (i.e., Psg1S5; Fig. 4) suggests that such an event would not require Psg1 to be competent for initiation (however, other properties such as polymerase binding may still be functional under such conditions). A strategy akin to that just described is utilized by bacteriophage Qβ (Klovins et al., 1998). In the (+)-strand RNA genome of this phage the replicase binding site is positioned approximately 1,000 nt from the initiation site for minus strand synthesis. The enigmatic placement of these elements is dealt with by long-distance base-pairing interactions which colocalize the two elements in a higher order RNA structure (Klovins et al., 1998).

Long-distance interactions between termini of viral genomes has been suggested as a mechanism providing a safety check for complete genomes prior to initiation of RNA synthesis (Lai, 1998). The internal location of the TBSV elements in the genome precludes this role for the distal element/core element interaction, however, a DE-A-like element can be identified near the 5′ terminus of the TBSV genome (5′-UCCAGGAUUUCU, coordinates 8 to 19), and in other tombusvirus genomes (data not shown), which can potentially base pair (underlined residues) with CE-A. This terminal location in the genome is similar to that of the element described for PVX (Kim & Hemenway, 1996, 1997). Interestingly, this sequence is also complementary to a segment positioned just 5′ to the initiation site for sg mRNA1 (5′-CAGCAUCCUCGA, coordinates 2595–2602; underlined residues are complementary to the DE-A-like sequence). Since DE-A does not appear to influence sg mRNA1 accumulation levels, the 5′-proximal DE-A-like sequence may represent an enhancer-like element for Psg1.

The growing list of different functional long-distance interactions in a variety of (+)-strand RNA viruses suggests that this mechanism represents a general strategy used for regulating RNA synthesis and/or gene expression. A comparison of the features of characterized distal sg mRNA enhancer elements of TBSV, RCNMV, and PVX reveals some similarities, but also several differences (Table 1). For instance, the TBSV and PVX interactions are predicted to occur *in cis*, whereas that of RCNMV occurs *in trans*. The spacing from the interaction to the site of initiation of sg mRNA synthesis in both TBSV and PVX is 11 nt; however, this intervening sequence is highly conserved in tombusviruses but is divergent in potexviruses. Unlike DE-A, which appears to influence only sg mRNA2 accumulation, the PVX element appears to regulate two sg mRNAs. The differences in these systems likely reflect mechanistic variations; thus, further studies on these regulatory RNA elements will undoubtedly uncover new properties that will aid in understanding their modes of action.

MATERIALS AND METHODS

Oligonucleotides

The following oligonucleotides were used in this study (underlined residues correspond to nonviral sequence, whereas those not underlined correspond to viral sequence). P9: 5′-GGCGGCCCGCATGCCCGGGCTGCATTTCTGCAATGTTCC (minus sense, 4754–4776); P56: 5′-TACACACGAGCCGTGGAGAGTCTGC (plus sense, 1441–1465); PB30: 5′-CTTGATTCGAATTCGTCTCATC (minus sense, 4026–4047); PF4: 5′-GCGCGTCTAGAGAAACGGGAAGCTCGCTCG (plus sense, 1285–1304); PG1: 5′-CCTCCTCTTCTCCGCACTGCTTTGTACGC (minus sense, 2811–2839); PG10: 5′-GCCCCCACCGGTCTTGTTCCCTAGCGTCGTTTC (minus sense, 3907–3927); PG11: 5′-GGCCCCACCGGTGAGCTCCTGGCTCTGGAGGACCGTCTG (plus sense, 2521–2541); PG12: 5′-GGCCCCACCGGTCGTGTTGTTGTTTCTCTTTAC (minus sense, 2661–2681); PG17: 5′-CTTGGTTACGCGACTGCCGAGTCGAGGATG (minus sense, 2598–2627); PG18: 5′-AATACACACACGCAGGATAGACAC (plus sense, 2628–2651); PG19: 5′-CCAACCGGTTCTAGATGTGTGTATTCTTGGTCAAGCTAC (minus sense, 2614–2637); PG20: 5′-CCAACCGGTTCTAGAGTGTCTATCCTGCGTGTGTGTATTC (minus sense, 2627–2650); PG21: 5′-CCAACCGGTTCTAGAATTGCCATGTGTCTATCCTGCGTG (minus sense, 2636–2659); PG26: 5′-GCCCCCACCGGTCTCCAGCTCTAATTTAGTGTGTCCTGC (plus sense, 3808–3828); PG27: 5′-GCCCCCACCGGTCTGCAGAGGGGCCTCTTGAACAAGACCAG (plus sense, 3830–3852); PG28: 5′-CCAACCGGTTCTAGAGTGTCAATGCAGGGTGTGTGTATTCTTG

TABLE 1. Virus properties related to sg mRNA synthesis.

Property	TBSV	RCNMV	PVX
Classification			
Genus	Tombusvirus	Dianthovirus	Potexvirus
Polymerase supergroup[a]	II	II	III
Features of sg mRNAs			
Number synthesized	2[b]	1[c]	3[d]
Coupled synthesis/regulation	possible partial[e]	N.A.[f]	yes[g]
Accumulation of (−)strand sg mRNAs	yes[h]	yes[c]	yes[i]
Base-pairing interaction			
In cis or *in trans*	*cis*[j]	*trans*[k]	*cis*[g]
Relative position (structure) of element	internal (unknown)[j]	internal (hairpin)[k]	terminal (unknown)[g]
Length of base pairing sequence (nt)	12[j]	8[k]	11[g]
Length of intervening sequence (nt)	~1,000 (sg mRNA 2)[j]	N.A.[k]	~5600 (CP sg mRNA)[g] ~4300 (TGB sg mRNA)[g]
Distance to initiation site (nt)	11[j]	2[k]	11[g]

[a]Designation according to Koonin, 1991b.
[b]Hearne et al., 1990.
[c]Zavriev et al., 1996.
[d]Verchot et al., 1998.
[e]Production of sg mRNA2 from sg mRNA1 remains a possibility.
[f]Not applicable.
[g]Kim & Hemenway, 1996, 1997, 1998.
[h]D. Ray & K.A. White (1999).
[i]Chapman et al., 1992.
[j]This study.
[k]Sit et al., 1998.

GTC (minus sense, 2621–2652); PG29: 5′-GCCCCCACCGGTCTGCAGCTCTAATTTAGTGAGTGCAGGGAGGGGCCTCTTGAACAAG (plus sense, 3808–3847); PV2: 5′-GTAGTCTACTGGCCCGCGCGGCCTTGGTTACGCGACTGCCGAGTCGAGGATGCTGGG (minus sense, 2593–2627); PV4: 5′-CAGTAGACTAGGCCGCGCGGGCCCTGGCTCTGGAGGACCGTCTGG (plus sense, 2521–2542); PV5: 5′-CCACGTGTTGTTGTTTCTCTTTACCATT (minus sense, 2656–2681).

Plasmid construction

All constructs used in this study were derivatives of clone T100 that contains a full-length cDNA copy of the wild-type TBSV genome (Hearne et al., 1990). TBSV genomes with inactivated capsid protein or movement proteins were acquired or generated as follows. pHS7, containing a deletion in ORF 3 (capsid protein) resulting in a frameshift and premature termination, has been described previously (Scholthof et al., 1993; Fig. 2A). VSΔE, in which frameshifts were introduced into both ORFs 4 and 5 (movement proteins), was made by digesting T100 with *Eco*RI, filling in the 5′ overhangs, and religating the product (Fig. 2A). A similar approach was used to generate the capsid protein/movement proteins double mutant pHS7ΔE (Fig. 2A), except that pHS7, rather than T100, was used.

To inactive Psg1, 5-nt substitutions [C → G (2610), C → A (2613), A → C (2616), T → G (2619), G → A (2621)] and a 104-nt segment deletion (2627–2730) were introduced into T100 using PCR. The amplification was carried out using oligonucleotide pair PF4/PV2 and T100 as template. The PCR product generated was then digested with *Bam*HI and *Sfi*I and ligated with *Bam*HI/*Sfi*I-digested T100, creating ΔPsg1. ΔPsg2 (Fig. 3A) was constructed by digesting T100 with *Age*I and *Eco*RI, isolating the larger fragment, filling in the termini, and self-ligating the product. ΔPsg1/2 (Fig. 3A), which combines the mutations of ΔPsg1 and ΔPsg2, was made by replacing a *Msc*I/*Sph*I fragment of ΔPsg1 with that from ΔPsg2.

The construct ΔPsg1+1 (Fig. 4A), containing a reintroduced Psg1 fragment, was generated by using a *Sfi*I/*Msc*I-digested PCR fragment, generated using oligonucleotide pair PV4/PV5 and T100 as template, to replace a *Sfi*I/*Msc*I fragment in ΔPsg1. The construct ΔPsg1S5 (Fig. 4A) containing only the 5-nt substitutions was constructed using two PCR-generated fragments. The first PCR product, PCR17, was generated with oligonucleotide pair P56/PG17 using T100 as template. The second PCR product, PCR18, was generated with oligonucleotide pair PG18/PG1 using T100 as template. In the PCR reactions described, oligonucleotides PG17 and PG18 were phosphorylated. PCR17 was digested with *Bam*HI and PCR18 with *Sfi*I and were subsequently used in a three-part ligation with *Bam*HI/*Sfi*I-digested T100.

To map the distal element, a series of deletion constructs was generated using PCR. The product PCR19 was generated with oligonucleotide pair P56/PG19, PCR20 with P56/PG20 and PCR21 with P56/PG21. In all cases the template T100 was used. The products were subsequently digested with *Bam*HI and ligated individually into *Bam*HI/*Msc*I-digested T100, generating Psg1-19, Psg1-20, and Psg1-21 (Fig. 5A), respectively. To map the core element, Psg2-26 and Psg2-27 (Fig. 5A) were constructed using products PCR26 and PCR27, respectively, generated using primer pairs PG26/PG10 and PG27/PG10. The PCR products were then digested with *Age*I and *Nco*I and individually ligated into *Age*I/*Nco*I-digested T100.

Four constructs were made to test the putative base-pairing interaction between the distal element and the core element. Psg20/26 (Fig. 6A) was created by digesting both Psg1-20 and Psg2-26 with *Msc*I and *Nco*I and using the 1,060-nt fragment released from Psg2-26 to replace the corresponding fragment in Psg1-20. Psg28/26 (Fig. 6A) was generated using the product PCR28, generated with oligonucleotide pair P56/PG28 and T100 as template. The PCR product was subsequently digested with *Bam*HI and *Xba*I and used to replace a *Bam*HI/*Xba*I fragment in Psg20/26. Psg20/29 (Fig. 6A) was generated using the product PCR29, generated with oligonucleotide pair PG29/PB30 and T100 as template. The PCR product was subsequently digested with *Pst*I and *Eco*RI and used to replace a *Pst*I/*Eco*RI fragment in Psg20/26. Psg28/29 (Fig. 6A) was constructed by digesting Psg20/29 with *Pst*I and *Sph*I and using the excised fragment to replace the corresponding *Pst*I/*Sph*I fragment in Psg20/29.

In vitro transcription

Viral transcripts were generated in vitro via transcription of *Sma*I-linearized template DNAs using the Ampliscribe T7 RNA polymerase transcription kit (Epicentre Technologies) as described previously (White & Morris, 1994). Following the transcription reaction, DNA templates were removed by treatment with DNase I (Epicentre Technologies) and unincorporated nucleotides were removed via column chromatography using a Sephadex G-25 spin column (Pharmacia). Ammonium acetate was added to the flowthrough to a final concentration of 2 M, the transcripts were extracted twice with equal volumes of phenol-chloroform-isoamyl alcohol, and then precipitated with ethanol. Subsequently, the transcripts were quantified spectrophotometrically and an aliquot was analyzed by agarose gel electrophoresis to verify integrity.

Isolation and inoculation of protoplasts

Protoplasts were prepared from 6–8-day-old cucumber cotyledons (var. Straight 8) as described previously (White & Morris, 1994). Quantification was carried out by bright field microscopy using a hemacytometer. Purified protoplasts (approximately 4×10^5) were inoculated with 5 μg of each viral RNA transcript and were incubated in a growth chamber under fluorescent lighting at 22 °C for 24 h or as otherwise specified.

Analysis of viral RNAs

Total nucleic acid was harvested from protoplasts at various times postinoculation as described previously (White & Morris, 1994). Aliquots of the total nucleic acid preparation (a sixth) were separated in 1.4% agarose gels and viral RNAs were detected by Northern blot analysis using [^{32}P]-end-labeled oligonucleotide P9 complementary to the 3′-terminal 23 nt of the TBSV genome. Radioisotopic quantification was performed using an InstantImager (Packard Instrument Company).

ACKNOWLEDGMENTS

We thank members of our laboratory for reviewing the manuscript. This work was supported by grants to KAW from the Natural Sciences and Engineering Research Council of Canada.

Received November 23, 1998; returned for revision January 6, 1999; revised manuscript received January 19, 1999

REFERENCES

Adkins S, Stawicki SS, Faurote G, Siegel RW, Kao CC. 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.

An S, Makino S. 1998. Characterizations of coronavirus *cis*-acting RNA elements and the transcription step affecting its transcription efficiency. *Virology 243*:198–207.

Chapman S, Hills G, Watts J, Baulcombe D. 1992. Mutational analysis of the coat protein gene of potato virus X: Effects on virion morphology and viral pathogenicity. *Virology 191*:223–230.

Grieco F, Burgyan J, Russo M. 1989. The nucleotide sequence of cymbidium ringspot virus RNA. *Nucleic Acids Res 17*:6383.

Havelda Z, Dalmay T, Burgyan J. 1997. Secondary structure-dependent evolution of cymbidium ringspot virus defective interfering RNA. *J Gen Virol 78*:1227–1234.

Havelda Z, Szittya G, Burgyan J. 1998. Characterization of the molecular mechanism of defective interfering RNA-mediated symptom attenuation in tombusvirus-infected plants. *J Virol 72*:6251–6256.

Hearne PQ, Knorr DA, Hillman BI, Morris TJ. 1990. The complete genome structure and synthesis of infectious RNA from clones of tomato bushy stunt virus. *Virology 177*:141–151.

Jeong YS, Repass JF, Kim YN, Hwang SM, Makino S. 1996. Coronavirus transcription mediated by sequences flanking the transcription consensus sequence. *Virology 217*:311–322.

Johnston JC, Rochon DM. 1995. Deletion analysis of the promoter for the cucumber necrosis virus 0.9-kb subgenomic RNA. *Virology 214*:100–109.

Joo M, Makino S. 1992. Mutagenic analysis of the coronavirus intergenic consensus sequence. *J Virol 66*:6330–6337.

Kim KH, Hemenway C. 1996. The 5′ nontranslated region of potato virus X RNA affects both genomic and subgenomic RNA synthesis. *J Virol 70*:5533–5540.

Kim KH, Hemenway C. 1997. Mutations that alter a conserved element upstream of the potato virus X triple block and coat protein genes affect subgenomic RNA accumulation. *Virology 232*:187–197.

Kim KH, Hemenway C. 1998. Complementarity of sequence elements at the 5′ end of the potato virus X genome and upstream of subgenomic RNA is important for RNA accumulation. 17th Annual Meeting of the American Society for Virology. Vancouver, British Columbia, Canada, 1998. [Abstract]. pp 132.

Klovins J, Berzins V, van Duin J. 1998. A long-range interaction in Q_β RNA that bridges the thousand nucleotides between the M-site and the 3′ end is required for replication. *RNA 4*:948–957.

Koonin EV. 1991a. Genome replication/expression strategies of positive-strand RNA viruses: A simple version of a combinatorial classification and prediction of new strategies. *Virus Genes 5*:273–282.

Koonin EV. 1991b. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol 72*:2197–2206.

Kozak M. 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell 15*:1109–1123.

Lai MMC. 1998. Cellular factors in the transcription and replication of viral RNA genomes: A parallel to DNA-dependent RNA transcription. *Virology 244*:1–12.

Lai MMC, Liao CL, Lin YJ, Zhang X. 1994. Coronavirus: How a large RNA viral genome is replicated and transcribed. *Infect Agents Dis 3*:98–105.

Li HP, Zhang X, Duncan R, Comai L, Lai MM. 1997. Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. *Proc Natl Acad Sci USA 94*:9544–9549.

Limmer S. 1997. Mismatch base pairs in RNA. *Prog Nucleic Acid Res Mol Biol 57*:1–39.

Lin YJ, Zhang X, Wu RC, Lai MM. 1996. The 3′ untranslated region of coronavirus RNA is required for subgenomic mRNA transcription from a defective interfering RNA. *J Virol 70*:7236–7240.

Maia IG, Seron K, Haenni A, Bernardi F. 1996. Gene expression from viral RNA genomes. *Plant Mol Biol 32*:367–391.

Makino S, Stohlman SA, Lai MM. 1986. Leader sequences of murine coronavirus mRNAs can be freely reassorted: Evidence for the role of free leader RNA in transcription. *Proc Natl Acad Sci USA 83*:4204–4208.

Miller WA, Brown CM, Wang S. 1997. New punctuation for the genetic code: Luteovirus gene expression. *Semin Virol 8*:3–13.

Miller WA, Dreher TW, Hall TC. 1985. Synthesis of brome mosaic virus subgenomic RNA in vitro by internal initiation on (−) sense genomic RNA. *Nature 313*:68–70.

Oster SK, Wu B, White KA. 1998. Uncoupled expression of p33 and p92 permit amplification of tomato bushy stunt virus RNAs. *J Virol 72*:5845–5851.

Ray D, White KA. 1999. Enhancer-like properties of an RNA element that modulates tombusvirus RNA accumulation. *Virology*. In press.

Rochon DM, Tremaine JH. 1989. Complete nucleotide sequence of the cucumber necrosis virus genome. *Virology 169*:251–259.

Rubino L, Burgyan J, Russo M. 1995. Molecular cloning and complete nucleotide sequence of carnation Italian ringspot tombusvirus genomic and defective interfering RNAs. *Arch Virol 140*: 2027–2039.

Sawicki SG, Sawicki DL. 1998. A new model for coronavirus transcription. *Adv Exp Med Biol 440*:215–219.

Scholthof HB, Morris TJ, Jackson AO. 1993. The capsid protein gene of tomato bushy stunt virus is dispensible for systemic movement and can be replaced for localized expression of foreign genes. *Mol Plant-Microbe Interact 6*:309–322.

Scholthof HB, Scholthof KB, Kikkert M, Jackson AO. 1995. Tomato bushy stunt virus spread is regulated by two nested genes that function in cell-to-cell movement and host-dependent systemic invasion. *Virology 213*:425–438.

Sethna PB, Brian DA. 1997. Coronavirus genomic and subgenomic minus-strand RNAs copartition in membrane-protected replication complexes. *J Virol 71*:7744–7749.

Siegel RW, Adkins S, Kao CC. 1997. Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proc Natl Acad Sci USA 94*:11238–11243.

Sit TL, Vaewhongs AA, Lommel SA. 1998. RNA-mediated *trans*-activation of transcription from a viral RNA. *Science 281*:829–832.

Spaan W, Delius H, Skinner M, Armstrong J, Rottier P, Smeekens S, van der Zeijst BA, Siddell SG. 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J 2*:1839–1844.

Tavazza M, Lucioli A, Calogero A, Pay A, Tavazza R. 1994. Nucleotide sequence, genomic organization, and synthesis of infectious transcripts from a full-length clone of artichoke mottle crinkle virus. *J Gen Virol 75*:1515–1524.

van der Kuyl AC, Langereis K, Houwing CJ, Jaspars EM, Bol JF. 1990. *cis*-acting elements involved in replication of alfalfa mosaic virus RNAs in vitro. *Virology 6*:346–354.

Verchot J, Angell SM, Baulcombe DC. 1998. In vivo translation of the triple gene block of potato virus X requires two subgenomic mRNAs. *J Virol 72*:8316–8320.

Wang J, Carpenter CD, Simon AE. 1999. Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. *Virology 253*:327–336.

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

White KA, Morris TJ. 1994. Nonhomologous RNA recombination in tombusviruses: Generation and evolution of defective interfering RNAs by stepwise deletions. *J Virol 68*:14–24.

White KA, Morris TJ. 1995. RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. *RNA 1*:1029–1040.

Wu B, White KA. 1998. Formation and amplification of a novel tombusvirus defective RNA which lacks the 5′ nontranslated region of the viral genome. *J Virol 72*:9897–9905.

Xiong ZG, Lommel SA. 1991. Red clover necrotic mosaic virus infectious transcripts synthesized in vitro. *Virology 182*:388–392.

Zavriev SK, Hickey CM, Lommel SA. 1996. Mapping of the red clover necrotic mosaic virus subgenomic RNA. *Virology 216*:407–410.