Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Article No. jmbi.1999.3297 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 294, 667–682

JMB

RNA Sequence and Secondary Structural Determinants in a Minimal Viral Promoter that Directs Replicase Recognition and Initiation of Genomic Plus-strand RNA Synthesis

K. Sivakumaran[1], Chul-Hyun Kim[2], Robert Tayon, Jr[1] and C. Cheng Kao[1]*

[1]*Department of Biology Indiana University Bloomington, IN, 47405, USA*

[2]*Department of Chemistry University of California Berkeley, and Physical Bioscience Division, Lawrence Berkeley National Laboratory Berkeley, CA, 94720, USA*

**Corresponding author*

Viral RNA replication provides a useful system to study the structure and function of RNAs and the mechanism of RNA synthesis from RNA templates. Previously we demonstrated that a 27 nt RNA from brome mosaic virus (BMV) can direct correct initiation of genomic plus-strand RNA synthesis by the BMV replicase. In this study, using biochemical, nuclear magnetic resonance, and thermodynamic analyses, we determined that the secondary structure of this 27 nt RNA can be significantly altered and retain the ability to direct RNA synthesis. In contrast, we find that position-specific changes in the RNA sequence will affect replicase recognition, modulate the polymerization process, and contribute to the differential accumulation of viral RNAs. These functional results are in agreement with the phylogenetic analysis of BMV and related viral sequences and suggest that a similar mechanism of RNA synthesis takes place for members of the alphavirus superfamily.

Keywords: RNA replication; viral replicase; RNA promoter; RdRp; brome mosaic virus

Introduction

RNA-dependent RNA polymerases (RdRps) are an interesting class of enzymes for several reasons: (1) RdRps are key enzymes in the replication of viral pathogens, including the causative agents of measles, the common cold, and viral hepatitis (Van Dyke & Flanegan, 1980; Neulfeld *et al*., 1991; Poch *et al*., 1990; Behrens *et al*., 1996; Yuan *et al*., 1997); (2) cellular RdRps affect gene expression in enucleated cells and in gene silencing (Volloch *et al*., 1996; Bingham, 1997; Schiebel *et al*., 1998; Cogoni & Macino, 1999); and (3) RdRps are the only class of template-dependent polymerases that can initiate synthesis *de novo* from the end of the template (for reviews, see Buck, 1996; O'Reilly & Kao, 1998). We study the mechanism of RNA synthesis and promoter recognition using virally encoded replicases.

The virally encoded RdRp, along with viral and cellular proteins make up the replicase of positive-sense RNA viruses (Hardy *et al*., 1979; Quadt & Jaspars, 1990; Buck, 1996). In one cycle of RNA replication, the replicase uses the genomic plus-strand RNA as the template to generate the complementary minus-strand RNA, which in turn serves as the template for generating multiple copies of genomic, and in some viral species, subgenomic plus-strand RNAs. Viral RNA replication requires the specific recognition of RNA features by the replicase, a process that is only beginning to be elucidated.

We study RNA synthesis using the model system brome mosaic virus (BMV), a plant-infecting member of the alphavirus-like superfamily of positive-strand RNA viruses (Goldbach *et al*., 1991). BMV has three genomic RNAs, designated as RNA1 (3.2 kb), RNA2 (2.8 kb), and RNA3 (2.1 kb). RNA1 and RNA2 are monocistronic and encode the 1a and 2a proteins, respectively, which are essential for the replication of the BMV genome (Ahlquist, 1992). RNA3 is dicistronic and encodes the 3a movement protein and the coat protein that is expressed *via* a subgenomic RNA. An *in vitro*

Abbreviations used: RdRp, RNA-dependent RNA polymerase; BMV, brome mosaic virus; NOESY, nuclear Overhauser effect spectroscopy.

E-mail address of the corresponding author: ckao@bio.indiana.edu

0022-2836/99/480667–16 $30.00/0

system using highly enriched BMV replicase preparations from infected barley has been developed to study initiation of RNA synthesis from minimal templates. Using this system, accurate initiation of minus-strand, genomic plus-strand and subgenomic RNAs *in vitro* has been demonstrated (Dreher & Hall, 1988; Miller *et al*., 1985; Sun *et al*., 1996; Adkins *et al*., 1997; Siegel *et al*., 1997, 1998; Chapman *et al*., 1998; Sivakumaran & Kao, 1999).

The initiation of genomic plus-strand RNA synthesis from the 3′ end of the minus-strand template requires one nucleotide 3′ of the initiation nucleotide (Sivakumaran & Kao, 1999). This minus-strand template is recognized by the BMV replicase in a species-specific manner; the BMV replicase does not efficiently recognize the (−)-strand RNA of the closely related cucumber mosaic virus (CMV) or tobacco mosaic virus (Sivakumaran & Kao, 1999; Chapman & Kao, 1999). In this paper, the secondary structure and sequence determinants governing the specific recognition of the RNA directing genomic plus-strand RNA synthesis was investigated. Using biochemical analysis and nuclear magnetic resonance spectroscopy (NMR), we found that efficient RNA-directed genomic plus-strand RNA synthesis can take place with more than one RNA secondary structure. Furthermore, the template sequence appears to determine the efficiency of RNA synthesis.

Results

Proposed secondary structure required to direct genomic plus-strand synthesis

We have demonstrated that accurate initiation of genomic plus-strand synthesis from input minus-strand templates of 27 nt was comparable to those from longer templates (Sivakumaran & Kao, 1999). The removal of two additional 5′ template nucleotides from the 27 nt RNA reduced synthesis to 30 %. Therefore, the 27 nt RNA B2(−)26G represents a minimal functional promoter and template (to be named an endscript since the promoter is also the template). We seek to use endscript B2(−)26G to determine the features that enable efficient plus-strand RNA synthesis.

Multidimensional NMR is a powerful method for detailed structural analysis of molecules in solution. By virtue of its small size (27 nt), B2(−)26G is amenable to studies with proton NMR. NMR analysis requires the RNA to be in one predominant conformation. Since the regions forming intramolecular base-pairing can have high potential to form a dimeric conformation through intermolecular base-pairing, it is important to ensure that the RNA forms a stable monomeric conformation. To examine this, UV melting experiments were done with B2(−)26G over a 30-fold concentration range. Melting temperature (t_m) for the major transition was estimated at 55.6(±0.5) °C, and did not change appreciably over a concentration ranging from 2 to 60 μM, indicating that B2(−)26G has a monomeric conformation. In addition, native gel analysis was performed to determine the number of electrophoretically distinct conformations of B2(−)26G that may exist. As a control, a 17 nt RNA that exists both as a monomer and a dimer was used (Figure 1(a), lane 1). B2(−)26G was found to migrate predominantly as a single band (Figure 1(a), lane 2), indicating that it does not form a dimer. Also, when tested in salt concentrations ranging from 50 to 200 mM, the mobility did not change, indicating that its conformation is stable in the range of salt conditions used for NMR studies.

The most stable secondary structure predicted for B2(−)26G by the Zuker mfold program (Jaeger *et al*., 1989) is shown in Figure 1(b). We sought to confirm or modify this predicted structure using the results obtained from NMR analysis. Although solving a complete tertiary structure at atomic resolution using NMR is a time-consuming process, one-dimensional (1D) imino NMR and two dimensional NMR spectra of an RNA molecule can be used to determine RNA secondary structure (Varani & Tinoco, 1991). When the imino protons are involved in base-pairing or forming an unusual conformation, they are usually protected from fast proton exchange with the water solvent, resulting in sharp imino proton peaks. If those protected imino protons are located within a 5 Å space, the crosspeaks between imino peaks can be observed by two-dimensional nuclear Overhauser effect spectroscopy (2D NOESY), where protected imino protons in guanine and uracil bases will result in a NOE in a NOESY spectrum (Variani & Tinoco, 1991) and allow the assignment of nucleotides in base-pairs. As seen in Figure 1(b), the proposed secondary structure contains eight base-pairs with a U·U mismatch and a CUU tri-loop. If this structure is the major conformation in solution, we should be able to see major imino peaks from the two A·U, four G·C, and one G·U base-pair. The terminal A·U base-pair might not be stable due to increased flexibility caused by the adjacent 5′ and 3′ overhangs. More than eight imino peaks are observed in the 1D imino spectrum (Figure 1(c)), which indicates that the 27 nt RNA may have unusually protected imino peaks or some minor conformation that can be clarified by analyzing 2D water NOESY crosspeak pattern.

The crosspeak pattern shown in the 2D H_2O NOESY (Figure 1(d)) would allow us to discern between G·C and A·U base-pairs using the characteristic crosspeaks between the uracil imino and adenine H2 in the A·U base-pair, and the crosspeaks between guanine imino and cytosine amino protons in the G·C base-pair, respectively. The expected G·U wobble base-pair was discerned by its characteristic strong crosspeaks between G and U imino protons. The connectivity between weaker A·U base-pairs was also observed (gray circles with dotted lines in Figure 1(d)), which might be either the minor conformation of A·U base-pairs or unstable A·U base-pairs from the terminal stem or

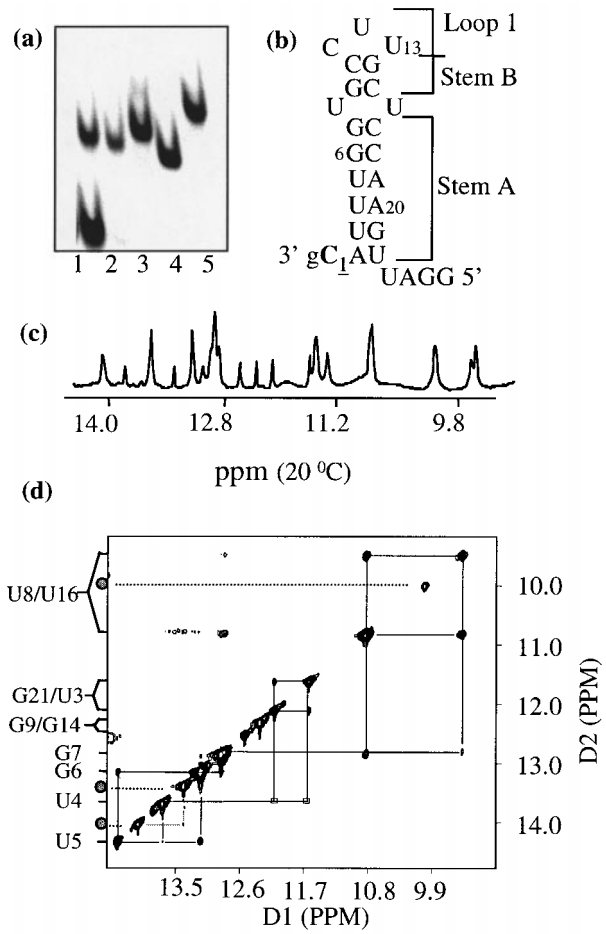


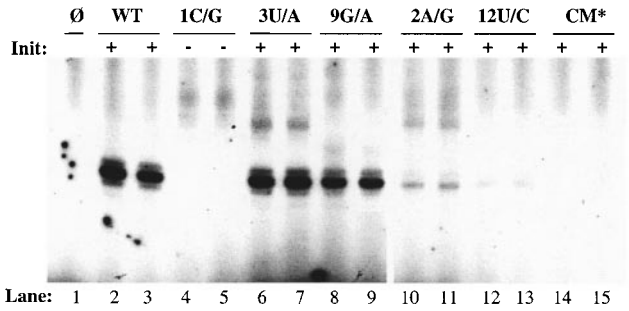
Figure 1. Analysis of the secondary structure of endscript B2(−)26G. (a) Visualization of RNA conformations by electrophoresis through a non-denaturing 12% polyacrylamide gel. Between 30 and 60 μM of RNA purified by denaturing gel electrophoresis were heated to 90 °C for two minutes, then allowed to cool over a 30 minute period to room temperature. The samples were then electrophoresed on a 17 cm × 15 cm × 0.8 cm gel at 200 V until the bromophenol blue migrated to the middle of the gel. The RNAs are SL13, an RNA that exists in equilibrium between an intramolecularly folded hairpin and an intermolecularly folded dimer (lane 1), B2(−)26G (lane 2), +3 U/A (lane 3), +6,7 G/C, +9G/A. (b) The secondary structure of B2(−)26G as predicted by the mfold program. (c) One-dimensional NMR spectrum of the imino protons from B2(−)26G taken at 20 °C by a Bruker 600 mHz spectrometer. B2(−)26G, at 0.7 mM, was purified and prepared in 10% deuterium/90% water as described in Experimental Procedures. (d) The two-dimensional NOESY spectrum of a 2.0 mM B2(−)26G sample taken at 20 °C taken from a Bruker 500 mHz spectrometer. The identity of the nucleotide responsible for the signals in the diagonal plane is indicated on the left-hand side of the Figure. Where noted with a gray spot, the origin of the signal was not determined with certainty and could have arisen from nucleotides in the loop of unbase-paired regions. The vertical and horizontal lines were added to the output containing the NMR signals to identify the connectivity of two nucleotides. The two crosspeaks associated with U4 and indicated by squares were observed in spectra with different mixing time than the one shown.

5′ overhang. Interestingly, the U·U mismatched base-pairs were observed and confirmed by its characteristic strong crosspeaks between the N3 proton of base-paired uracil bases and its connectivity with guanine 7 imino proton. Overall, with the possible presence of its local minor conformation, the assignment of imino peaks based on 1D and 2D NMR is well matched with the predicted secondary structure of B2(−)26G (Figure 1(b)).

Effect of mutations on synthesis from B2(−)26G endscript

Mutational analysis of B2(−)26G was performed to determine the features required for recognition by the BMV replicase. Mutant endscripts were all quantified after transcription and purification as described by Adkins *et al*. (1998). In essence, one picomole of each endscript was subjected to RNA synthesis assay and the amount of product synthesized was normalized to that produced by wild-type B2(−)26G. B2(−)26G yielded a predominant product of 26 nt, and some minor amounts of RNA one nucleotide longer and shorter due to non-templated nucleotide addition and premature termination, respectively (Figure 2(a), lanes 2-3). All three products initiated from the +1 cytidylate, since its change to a guanylate abolished all synthesis (Figure 2(a), lanes 4-5). Products directed by several of the mutant endscripts are shown in Figure 2(a) (lanes 6-15). Approximately 50 mutant endscripts were characterized according to the following scheme. To get an impression of the sequences important for RNA synthesis, a series of three nucleotide transitions starting at the 3′ end up to position +21 were carried out. In general, mutations within the first seven positions relative to the 3′ end of the template have a more severe effect on RNA synthesis than mutations from position +8 to +15 (Figure 2(b)). Some mutations near the 5′ terminus also decreased RNA synthesis. We have previously analyzed the effect of changes in positions between +17 and +24 and found that sequence within this region is required for efficient RNA synthesis in a yet undetermined mechanism (Sivakumaran & Kao, 1999). We decided to focus on position +2 to +8 because this sequence is likely to be directly involved in the initiation process (Sun *et al*., 1996; Sun & Kao, 1997a,b). We consequently tested every possible single nucleotide substitution from position +1 to +9, along with at least one change at position +10 to +21. In contrast to the BMV subgenomic promoter, where changes at key positions drastically reduced RNA synthesis (Siegel *et al*., 1997), we found that none of the substitutions prevented RNA synthesis except for a C/G change at the initiation site. However, position +4 to +7 were comparatively sensitive to substitutions. Interestingly, while single nucleotide substitutions of either the +6 or +7 guanylates with cytidylates resulted in low levels of synthesis, replacement of both guanylates with cytidylates in

(a)



(b)

	1		3		5		7		9		11		13		15		17		19		21
3′	G	C	A	U	U	U	G	G	U	G	C	C	U	U	G	C	U	C	C	A	A	G...5′
3-nt TS:				4			12			16			7			49			35			
Change to: A	40		100	80	12	21	16	13	110	73			24		115					100	100	
C	85	100	56	18	42	30	5	9	56	99	100	100	7	106		100	35	100	100			
G	100	0	11	18	25	44	75		191	100	157	86	110	88	100	25	44	5	12	67	87	100
U	45	7	77	100	100	100	7	12	100	88	180	116	100	100		64	100			13	20	102

Figure 2. Effect of nucleotide changes on plus-strand RNA synthesis. (a) An autoradiogram of RNAs produced by select endscripts. The reaction products were separated by denaturing 12 % PAGE and visualized by autoradiography. Ø represents the products of a control reaction with no added template. WT denotes the wild-type sequence of B2(−)26G. Changes to the B2(−)26G template are indicated above the autoradiogram of the RdRp products. CM* denotes an endscript with a +3-5 substitutions (UUU changed to CCC) plus a substitution of position +19-21 (AAG changed to GGG). Endscripts that have an initiation-competent cytidylate at the +1 position are indicated by a +, while endscripts with an initiation incompetent +1 guanylate are indicated with a −. The predominant product is 26 nt; lesser amounts of a 27 nt product and a 25 nt product are also observed. (b) Summary of the mutational analysis. The wild-type sequence is written near the top of the Figure. The changes at each of the positions are indicated on the left of the Figure. The effect of the nucleotide substitutions on RNA synthesis is denoted as a percentage relative to the amount of synthesis directed by the B2(−)26G wild-type. All results presented are from at least three independent trials with a standard deviation ⩽11 %. Boxes left blank indicate that the effect of that nucleotide substitution was not tested.

endscript +6,7 G/C restored synthesis to 75 % of that of B2(−)26G (Figure 2(b)).

The mutations in B2(−)26G include some that are expected to affect the two stems (A and B), the 3 nt loop, and the mismatched region within the predominant structure of B2(−)26G (Figure 1(b)). Mutations that should affect at least a portion of the B2(−)26G structure while retaining the ability to direct synthesis are summarized in Figure 3(a). A change to an adenylate (+3U/A) resulted in RNA synthesis of 80 % despite a predicted destabilizing effect on stem A (Figure 2(a), lanes 6-7, 2(b), and 3(a)). A change of the +9 guanylate to an adenylate (+9 G/A), expected to weaken stem B, resulted in 73 % synthesis while changes at positions +10 and +14 expected to weaken stem B resulted in increased synthesis (Figure 3(a)). The results indicate that the RNA secondary structure of B2(−)26G can be changed without debilitating plus-strand RNA synthesis by the BMV replicase.

Several mutations that are expected to preserve the secondary structure in B2(−)26G were found to have a detrimental effect on RNA synthesis. A change of +2 adenylate to a guanylate, which should retain a wobbled G·U base-pair instead of a Watson-Crick A·U base-pair, resulted in 8 % synthesis (Figures 2(b) and 3(b)). A transition of the +3 uridylate to a cytidylate resulted in 18 % synthesis, while a replacement of three base-pairs in stem A with three G·C base-pairs abolished synthesis. Changes in loop 1 also greatly reduced RNA synthesis (Figure 3(b)). Lastly, since nucleotide substitutions that disrupted stem A had detrimental effects on RNA synthesis, we made a compensatory change to the +6 G/C endscript by replacing the +18 cytidylate with a guanylate. This double mutant, SS1, should maintain a predicted secondary structure similar to B2(−)26G. However, RNA synthesis from SS1 was at 21 %. A similar change of +7 G/C along with the complementary change at +17 C/G in endscript SS2 resulted in only 26 % of the activity of B2(−)26 (Figure 4(a)). Consistent with the findings that mutations may alter B2(−)26G structure without significant loss of activity (Figure 3(a)), these results indicate that the preformed, predominant secondary structure of B2(−)26G is not essential for directing plus-strand RNA synthesis.

The structures predicted by the mfold program for select mutant endscripts were examined to better understand the RNA determinants in B2(−)26G required for initiating RNA synthesis. The predicted secondary structures of the endscripts directing the initiation of BMV RNA1 and RNA3

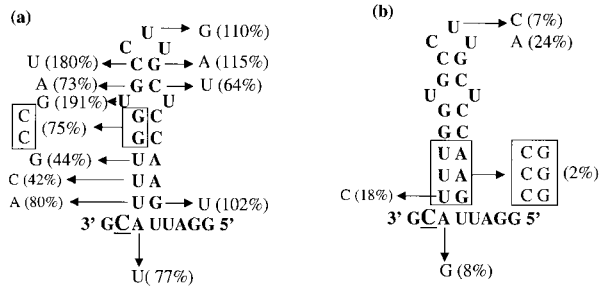


Figure 3. Correlation between the secondary structure and the ability of mutant endscripts to direct RNA synthesis. The amount of RNA synthesis directed by select mutant endscripts are shown as a percentage relative to that of B2(−)26G. (a) Schematic highlighting the nucleotide changes expected to disrupt the RNA secondary structure and their effect on RNA synthesis. (b) Schematic highlighting the nucleotide changes expected to maintain the RNA secondary structure and their effect on RNA synthesis.

are shown along with several derivatives of B2(−)26G (Figure 4(a)). All endscripts capable of directing RNA synthesis tend to have either an exposed 3′ initiation nucleotide, or have it present in a relatively weak portion of a stem that is rich in A·U and G·U base-pairs. Also, all except B3(−)26G possess a loop region rich in pyrimidines. If such a loop motif is necessary for RNA synthesis, it is not sufficient since SS1, SS2, and several other mutant endscripts that retain such a motif are debilitated for RNA synthesis. Finally, we note that while the RNAs are predicted to vary in the lengths of the stems from six (assuming a uridylate-uridylate base-pair in several of the endscripts) to nine base-pairs, they are all capable of directing RNA synthesis (Figure 4(a)).

To examine whether the predicted stability of the secondary structure is correlated to the ability to direct RNA synthesis, a plot of the calculated ΔG from the mfold program for all mutants endscripts were plotted against their ability to direct RNA synthesis (Figure 4(b)). Although the calculated ΔG based on the mfold program is only partially accurate due to the lack of thermodynamic information of the loop region, we should be able to determine whether there is an overall correlation between stability of endscript secondary structures and the ability to direct RNA synthesis. Two extremely unstable endscripts did result in inefficient RNA synthesis (Figure 4(b)). Nonetheless, no consistent correlation between RNA stability and the ability to direct RNA synthesis was observed. For example, endscript +10 C/U ($\Delta G = 3.4$ kcal/mol) is predicted to have lower stability than B2(−)26G ($\Delta G = 3.9$ kcal/mol), but was nonetheless more efficient in directing RNA synthesis. In addition, +8 U/G ($\Delta G = 8.8$ kcal/mol), was observed to direct RNA synthesis nearly twice as efficiently as B2(−)26G. Hence the stability of the RNA secondary structure was not a consistent indicator of the ability to direct synthesis.

Biochemical and biophysical studies on the conformations of endscript RNAs

Our mutational results are inconsistent with the possibility that the interaction between the BMV replicase and B2(−)26G is primarily determined by the predominant solution structure of the RNA. Therefore, we wanted to perform in-depth analyses of select endscripts using NMR and RNA melting experiments. Three mutant endscripts were selected from those described in Figure 2(b): (1) +3U/A contains a change of the +3 uridylate to adenylate; (2) +6,7 G/C has both the +6 and +7 guanylates changed to cytidylates; and (3) RNA +9 G/A has the +9 guanylate changed to an adenylate. Although each mutation should affect the stability of the different regions within B2(−)26G, the ability to direct RNA synthesis was at least 70 % of that of B2(−)26G (Figure 3(a)).

Native gel electrophoresis was performed to examine the structures of the three mutant endscripts. In a 12 % native gel, each of the three mutant endscripts exists predominantly in a single electrophoretic conformation (Figure 1(a)). However, the mobility of the three mutant RNAs differ from each other and from that of B2(−)26G. The mobility of +3 U/A, +9 G/A, and +6,7 G/C in comparison to B2(−)26G were, respectively: slightly reduced, drastically reduced, and increased (Figure 1(a), lanes 2-5). In a denaturing gel, all the three RNAs co-migrated with B2(−)26G. These results suggest that the differences in mobility of the RNAs are likely due to changes in the RNA conformation.

Next, we examined the melting profiles of the RNAs. The shape of the melting curve for B2(−)26G, +3 U/A and +9 G/A suggests that they melted with multiple step transition, while +6,7 G/C undergoes a single step transition between two states (Figure 5(a)). With a crude estimate of the relative increase in absorbance and the magnitude of the slope for each transition, it seems that B2(−)26G, +3U/A and +9G/A underwent a relatively sharp major transition and then a broad minor transition at higher temperatures (Figure 5(a)). The multiple transitions might be due to differential melting of secondary structure elements or the unstacking of the bases in the single-stranded portions of the RNAs. The absence of multiple transitions in +6,7G/C clearly shows that it has distinct thermodynamic properties from the other endscripts, implying a different conformation. In addition, even among B2(−)26G, +3 U/A and +9 G/A, the overall shapes of the melting curves look different, which indicates that these RNAs undergo different unfolding processes. These results are consistent with the native gel analysis and suggest that the three mutant endscripts are altered in their secondary structures in comparison to B2(−)26G and to each other.

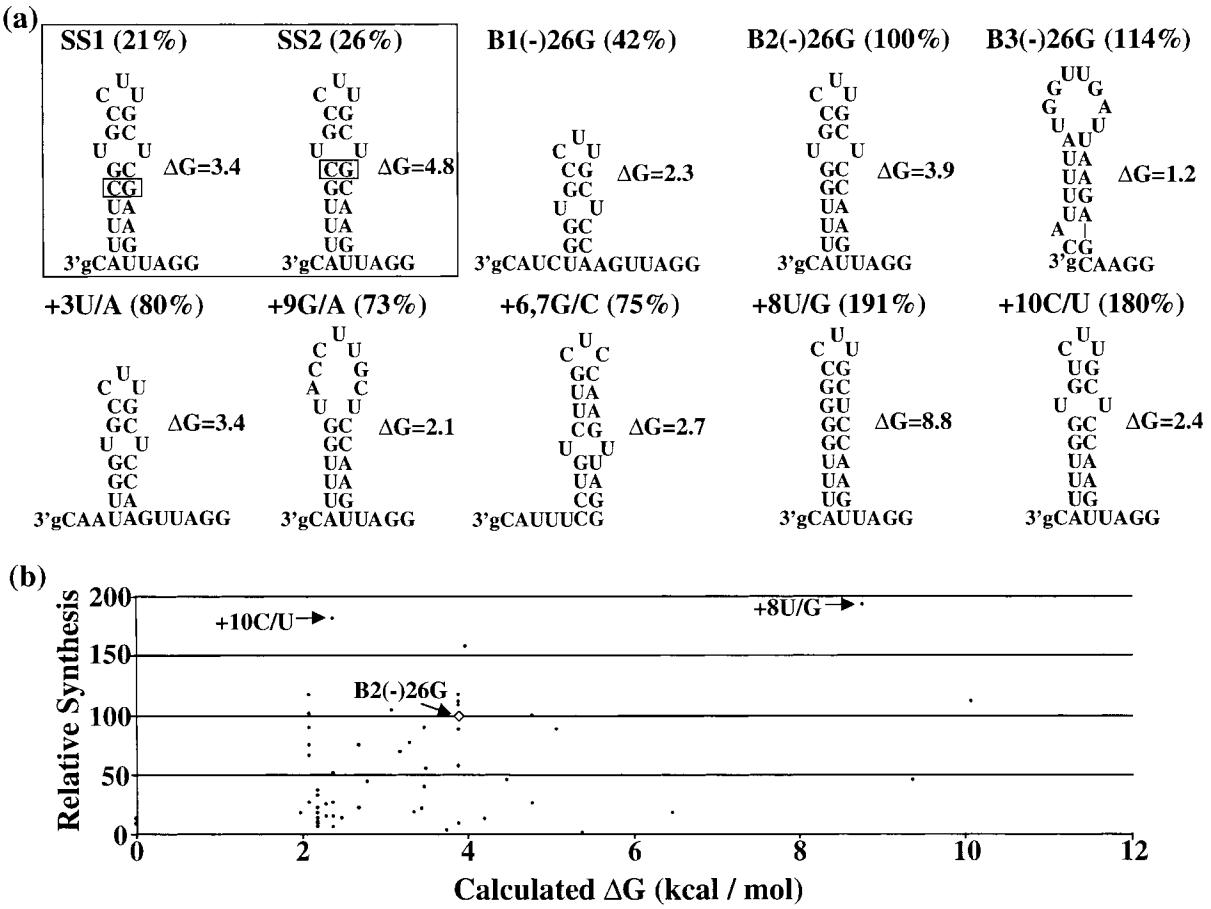


Figure 4. Correlation between the predicted stability of RNA secondary structures and their ability to direct RNA synthesis. (a) Schematic of the secondary structures of several endscripts predicted by the mfold program (Jaeger *et al.*, 1989). These structures were selected because they contain stems and loops that are useful to compare with those from B2(−)26G. The amount of synthesis directed by each endscript is shown as a percentage relative to that of B2(−)26G. The Δ*G* in kcal/mol of the predicted structure was calculated by the mfold program. Endscripts SS1 and SS2, (within the box), denote B2(−)26G RNA derivatives with compensatory nucleotide changes on both sides of the main stem. These were the only two mutant derivatives of B2(−)26G shown here that were significantly debilitated for RNA synthesis. The Other RNA derivatives are described in the legend to Figure 2(b) and in the text. (b) A plot of relative RNA synthesis *versus* the calculated Δ*G* value of the predicted secondary structures for all of mutant endscripts tested in this study. The value for B2(−)26G is denoted by a white diamond and the two endscripts with the least and most stable predicted secondary structures able to direct efficient RNA synthesis are indicated by arrows.

To characterize the changes in the structures of the mutant endscripts more precisely, 1D imino spectra were taken at different temperatures (Figure 5(b)). The quality of the ca 12 peaks characteristic of B2(−)26G were significantly affected by temperature. The peaks were relatively sharp at 20 °C (data not shown), but became increasingly broad at 30 °C and 40 °C, indicating that exchange with the environment of some of the imino protons increased at higher temperatures (Figure 5(b)). For ease of description, several of the imino peaks have been named in the spectra of B2(−)26G (Figure 5(b)). Of interest is the U8-U16 interaction at approximately 11 and 9.7 ppm. This interaction appeared to be more sensitive to temperature changes compared to interactions associated with G·C base-pairs at approximately 13-13.4 ppm. Endscript +3U/A has only five of the peaks associated with B2(−)26G (Figure 5(b)). At 20 °C, the U8-U16 interaction was retained in +3 U/A, as were the peaks associated with G·C base-pairs adjacent to the U8-U16 interaction. However, missing were the base-pairs at the bottom of stem A, as exemplified by the absence of the U5 peak at 14.4 ppm. At 40 °C, all the peaks were significantly broadened in comparison to those from B2(−)26G, indicating that the mutation in +3U/A RNA destabilizes the structure of the RNA. Endscript +6,7 G/C gave a number of sharp peaks that do not correspond to the ones observed in B2(−)26G (Figure 5(b)), indicating that +6,7 G/C exists in a different structural conformation than B2(−)26G. The structure predicted for +6,7 G/C by the mfold program was different from that for B2(−)26G (Figure 4(b)). In addition, since the peaks from +6,7 G/C remain well defined at 40 °C, its conformation is more

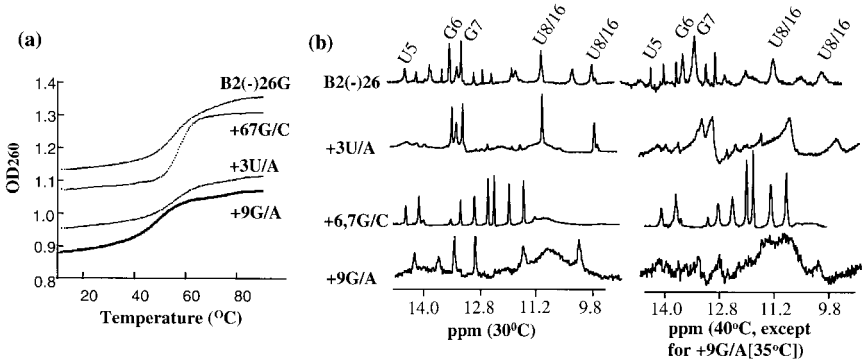


Figure 5. Characterization of the structures of endscript B2(−)26G and three mutant derivatives. (a) Change in absorbencies (at A_{260}) as a function of temperature. The absorbencies were recorded as described in Experimental Procedures. The identities of the RNAs are labeled next to the plot for each of the RNAs. (b) One dimensional NMR spectra of B2(−)26G and three mutant derivatives. The RNAs used ranged from 0.2 mM to 0.7 mM and the spectra were taken with a Bruker 600 mHz spectrometer. Similar regions of the spectrum from the samples are shown. The identities of several of the peaks were determined using 2D-NOESY data shown in Figure 1(d) and indicated above the peaks. All of the spectra in the left panel were taken at 30 °C. The spectra in the right-hand panel were taken at 40 °C for B2(−)26G, +3U/A, and +6,7G/C, and at 35 °C for +9G/A.

stable than that of B2(−)26G. This observation is consistent with its UV melt properties (Figure 5(a)). Endscript +9 G/A RNA, like +3 U/A RNA, retained a subset of the imino peaks observed in B2(−)26G, mostly corresponding to those in stem A (Figures 1(d) and 5(b)). However, these peaks are different from those observed in +3 U/A RNA. Most notably, the U8-U16 signal was drastically broadened, indicating the absence of the U8-U16 interaction adjacent to the mutation at position +9. Compared with the spectrum taken at 20 °C (data not shown), most peaks were shifted and broadened significantly at 30 °C (Figure 5(b)). At 35 °C, all the peaks were severely broadened, indicating a rapid exchange of the imino protons with the environment due to destabilization of its conformation. Taken together, the biophysical characterization strongly suggest that the predominant structure in B2(−)26G RNA was significantly altered by the introduced mutations. Therefore, the ability of endscripts to direct RNA synthesis by the BMV replicase under conditions that affect RNA conformation is of interest.

We examined the effect of temperature on RNA synthesis using B2(−)26G and the three mutant endscripts. Since temperature should affect both the RNA conformation and the activity of the BMV replicase, all syntheses are normalized to that of B2(−)26G at each assay temperature. While increasing the temperature from 30 °C to 43 °C did result in reduced synthesis (Figure 6), none of the mutant endscripts were differentially affected in comparison to B2(−)26G. Altogether, these biochemical and biophysical results indicate that multiple RNA secondary structures are acceptable for efficiency of RNA synthesis by the BMV replicase. These results also suggest that the template sequence may contribute more significantly to the level of RNA synthesis than its specific RNA structure.

The effect of nucleotide changes on template-RdRp interaction

Mutations in B2(−)26G could affect RNA synthesis either through modulation of the polymerization process, or by preventing template-replicase interaction. To examine the latter possibility, competition assays were performed with several endscripts. This assay determines the ability of competitor RNAs to reduce synthesis directed by an efficiently recognized subgenomic RNA, −20/15 (Siegel *et al.*, 1998; Sivakumaran & Kao, 1999). Competitor RNAs were added at 10 and 20 molar excess of −20/15 and the amount of synthesis directed by −20/15 was quantified. As a

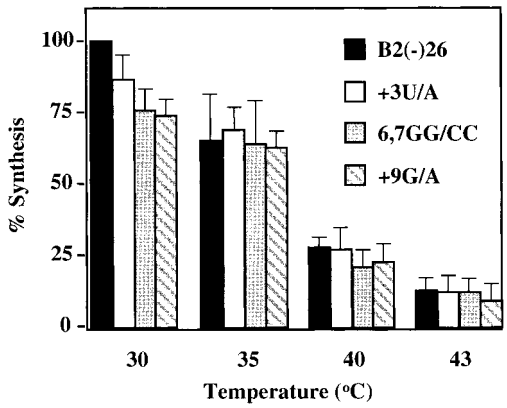


Figure 6. Effect of temperature on the ability of B2(−)26G and mutant endscripts to direct RNA synthesis. RdRp assays were carried out at 30 °C, 35 °C, 40 °C and 43 °C using B2(−)26G, +3 U/A, +6,7 G/C, and +9 G/A as templates. The amount of RNA synthesis directed by the various endscripts are presented as a percentage compared to B2(−)26G. All results presented are from at least three independent trials and the range for one standard deviation shown by the vertical lines.

negative control, we tested RNA B2(−)Δ3′, which has the three 3′ nucleotides deleted. In comparison to a reaction lacking a competitor, B2(−)Δ3′ reduced synthesis from −20/15 to 86% at ten molar excess and to 76% at 20 molar excess (Table 1). This mild reduction in −20/15 synthesis is consistent with our previous observations that the initiation site is required for recognition by the BMV replicase (Sivakumaran & Kao, 1999). In contrast, three functional RNAs, B1(−)26G, B2(−)26G, and B3(−)26G encoding the 3′ ends of BMV RNA1, RNA2, and RNA3, reduced synthesis from −20/15 more severely, to 52, 63, and 63% at ten molar excess, respectively (Table 1). CMV2(−)26G, the equivalent 3′ end of minus-strand RNA2 of CMV, reduced synthesis to 49%, a more effective level of inhibition than seen with BMV RNAs (Table 1). While CMV2(−)26G has the same 3′ three nucleotides as BMV RNAs and can interact with the BMV replicase, it was unable to direct efficient synthesis by the heterologous BMV replicase (Sivakumaran & Kao, 1999). Endscripts +3 U/A, +9G/A, +6,7 G/C, and B3(−) + 4U/C were at least as capable competitors as B2(−)26G (Table 1). These results suggest that the different levels of plus-strand RNA synthesis observed from the different endscripts is not due to a lack of recognition by the BMV replicase. Although it is possible that some RNA synthesis may be negatively affected by a replicase-endscript interaction that is too tight. Nonetheless, these competition results suggest that polymerization of nucleotides following the initial binding of the replicase to the initiation sequence contributes to the efficiency of RNA synthesis.

Template guanylates can reduce plus-strand RNA synthesis

We had previously reported that the BMV replicase could utilize DNA templates (Sivakumaran & Kao, 1999; Siegel *et al*., 1999). Furthermore, guanylates present within the first four positions of the initiation cytidylate were found to be detrimental to RNA synthesis from DNA endscripts (Sivakumaran & Kao, 1999). In our mutational analysis we observed that guanylates introduced near the initiation cytidylate of RNA B2(−)26G had a similar inhibitory effect (Figure 2(b)). Replacement of the native nucleotide with a guanylate from the +2 to +5 positions allowed synthesis between 8 to 44% of B2(−)26G. The detrimental effect of guanylates decreased as distance from the initiation cytidylate increased (Figure 7). From position +6 to +14, guanylates did not significantly affect RNA synthesis and sometime increased synthesis, as evidenced by guanylates at position +8 and +10 (Figure 7). However, guanylates at position +15 to +18 reduced RNA synthesis (Figure 7). The changes within position +15 to +18 likely affected a sequence previously observed to be important for RNA synthesis (Sivakumaran & Kao, 1999). Altogether, these results suggest that the identity of nucleotides at specific positions may affect the efficiency of RNA synthesis by the BMV replicase *in vitro*.

Table 1. Effect of several mutations on interaction with RdRp

RNA competition	1:10	1:20
None (100%)		
B2(−) Δ −1,1,2	86 ± 11	76 ± 7.7
B2(−)26G WT	63 ± 10	57 ± 7.5
B2(−)4U/C	52 ± 8	47 ± 4
B2(−)3U/A	62 ± 4.1	56 ± 2.6
B2(−)9G/A	39 ± 2.1	30 ± 3.5
B2(−)6,7GG/CC	60 ± 3.5	58 ± 9.8
B3(−)26G WT	63 ± 2.3	49 ± 8
B3(−)4U/C	76 ± 11	46 ± 11
CMV2(−)26G WT	49 ± 1.2	37 ± 5

Phylogenetic analysis of viral template sequence

To determine whether guanylates near the initiation site have a possible effect *in vivo*, we examined the distribution of guanylates in the sequences of BMV and other representatives of the alphavirus family. Sequences immediately following the initiation of genomic and subgenomic plus-strand RNAs generally lacked a guanylate in the first four positions (Table 2). At the +5 position, some animal-infecting, but no plant-infecting RNA viruses, had a guanylate in their RNAs. Guanylates were then found in increasing frequency starting at the +6 position.

The absence of template guanylates near the initiation nucleotide may be due to a requirement for replicase-promoter recognition, or because the

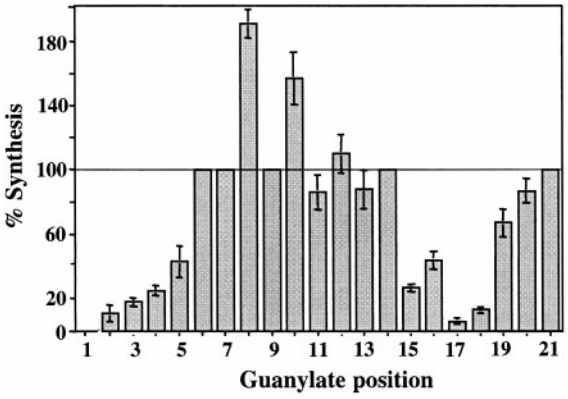


Figure 7. Correlation between guanylate substitutions and their effect on RNA synthesis. Guanylates were introduced at the different positions of the template as described in Experimental Procedures. All results are normalized to the amount of RNA synthesis directed by B2(−)26G (100%). The effect of three independently tested guanylate substitutions on viral RNA synthesis is shown with error bars denoting one standard deviation of the mean. Guanylates present in the wild-type sequence are shown without error bars.

Table 2. Presence of template guanylates in the initiation of genomic and subgenomic (+)-strand RNA synthesis in members of alphavirus-like superfamily

Virus	Nucleotide position +1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A. *Plant Infecting*															
BMV[a]	-	-	-	-	-	1,2	1,2	-	1,2,3	3	-	-	3	1,2,4	-
CCMV[a]	-	-	-	-	-	1,2,3	1,2	-	1,2	4	3	3	-	-	1,2,4
CMV[a]	-	-	-	-	-	3	-	-	4	1,2,3	3	4	3	-	-
AMV[a]	-	-	-	-	-	-	-	-	1,2	3	3	-	1	2	1
BBMV[b]	-	-	-	-	-	1	4	-	-	-	4	3	-	1,2	1,4
TMV[c]	-	-	-	-	-	-	-	-	-	1	1,5	-	4	5	
B. *Animal Infecting*															
Sindbis[d]	-	-	-	-	-	1,4	-	-	1,4	-	-	-	-	-	-
Semliki forest[d]	-	-	-	-	1	-	-	4	-	-	-	-	-	-	-
Middleburg[d]	-	-	-	-	-	-	-	4	-	-	-	-	1	-	-
EEE[d]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VEE[d]	-	-	-	-	-	1	-	-	1	-	-	-	-	-	-
Aura[e]	-	-	-	-	1	4	-	-	1	-	-	-	1	4	-
RRV[f]	-	-	-	-	-	1	-	4	-	1	-	-	-	4	-
Ockelbo[g]	-	-	-	-	-	1,4	-	1,4	-	-	-	-	-	-	-
O'Nyong-nyong[h]	-	-	-	-	1	-	-	4	-	1,4	-	4	-	-	4

BMV, brome mosic virus; CCMV, cowpea chlorotic mottle virus; CMV, cucumber mosaic virus (fny); AMV, alfalfa mosaic virus; BBMV, broad bean mottle virus; TMV, tobacco mosaic virus; EEE, eastern equine encephalitis virus; VEE, Venezuelan equine encephalitis virus; RRV, Ross River virus. The numbers 1, 2, 3 represent genomic RNA1, 2, 3 respectively; 4 represents the subgenomic RNA encoding the capsid protein, and 5 represents the 30 K subgenomic RNA encoding the movement protein of TMV.

[a] Ahlquist *et al.* (1981).
[b] Gunn & Symons (1980).
[c] Goelet *et al.* (1982).
[d] Ou *et al.* (1982a,b, 1983).
[e] Rumenapf *et al.* (1995).
[f] Faragher & Dalgarno (1986). Faragher *et al.* (1988).
[g] Shirako *et al.* (1991).
[h] Levinson *et al.* (1990).

interaction of the nascent RNA with the template is affected by tighter base-pairing. In the latter case cytidylates near the initiation site should also be detrimental to synthesis. While we did observe that substitutions with cytidylates near the initiation cytidylate did have detrimental effects, the effects were not to the same degree as the guanylate substitutions. However, in examining the 3′ ends of genomic plus-strand templates of approximately 60 viral templates, we observed a similar striking bias against guanylates and cytidylates from position +2 to +5, followed by a more normal distribution of these two nucleotides (Figure 8). This trend appears to be unique for templates directing the synthesis of plus-strand RNAs, since minus-strand synthesis does not follow the same trend (Figure 8). Together, the phylogenetic and mutational results are consistent with the testable hypothesis that the strength of the interaction between the template and the nascent RNA near the initiation site contributes to the efficiency of plus-strand RNA synthesis.

Sequence affects RNA synthesis

The observation suggesting that the primary sequence of BMV RNAs can affect RNA synthesis prompted us to examine the BMV RNA sequences more closely. The templates for synthesis of BMV RNA2 and RNA3 are identical in the first five positions and RNA1 and RNA2 are identical at the first 26 positions except at +4, where RNA1 has a cytidylate and RNA2 a uridylate (Figure 9(a)). *In vitro*, B1(−)26G directs only 42 % the amount of product as B2(−)26G (Figure 9(b), lanes 1-2 and 3-4). *In vivo*, RNA1 is known to accumulate at 68 % of RNA2 (Marsh *et al.*, 1991). To determine whether the +4 position is important in regulating the efficiency of RNA synthesis, we made B3(−)26G with a +4U/C change and quantified the amount of synthesis relative to B2(−)26G. Synthesis directed by B3(−)26G was 114 % of that from B2(−)26G (Figure 9(b), lanes 1-2 and 5-6). B3(−)26G with a +4 U/C change decreased synthesis to 52 % of that from B3(−)26G (Figure 9(b), lanes 5-6 and 7-8). These results provide another example that the identity of a template nucleotide near the initiation site can affect the efficiency of RNA synthesis.

Discussion

An understanding of the rules guiding RNA synthesis is lacking in all RNA viruses. Determining these rules will provide powerful guides for developing strategies to inhibit viral infections and to compare the mechanism of RNA synthesis by other polymerases. We have previously demonstrated that initiation of BMV genomic plus-strand synthesis can be accurately directed by minimal

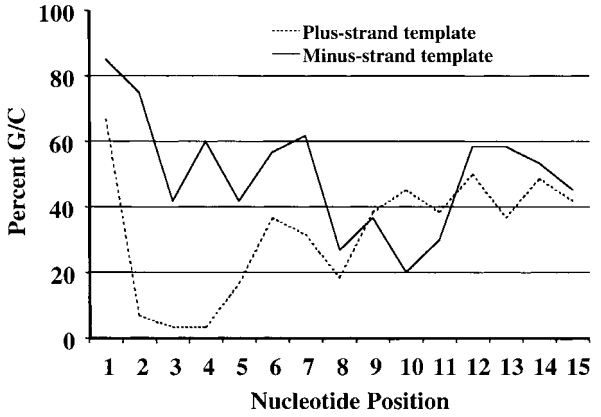


Figure 8. A selection of 60 plus-strand viral RNA templates reveals a propensity against guanylate and cytidylate proximal to the genomic plus-strand initiation site. Genomic minus-strand initiation sequences of the same templates do not exhibit such a preference. Sequence data were obtained from NCBI Entrez. Only plus-sense RNA viruses with complete genomic sequences and with plant tropism were included. For database accession numbers, see Experimental Procedures.

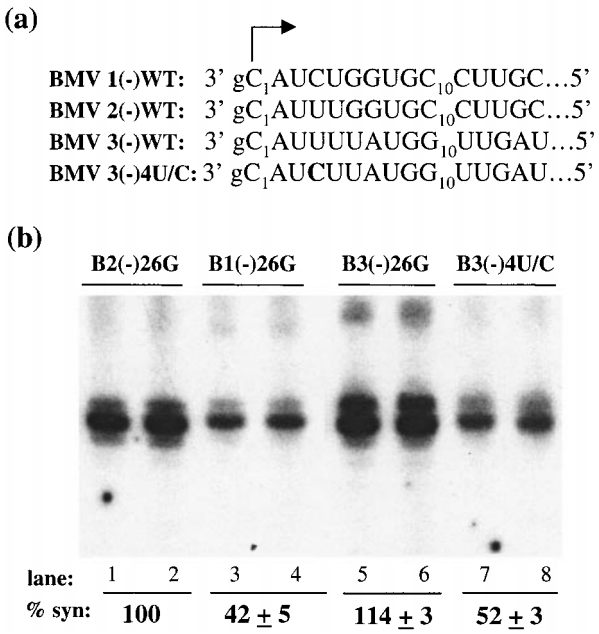


Figure 9. Effect of nucleotide changes on the differential accumulation of BMV RNAs. (a) Comparison of the 3′ minus-strand RNA sequences of BMV. The non-templated guanylate at the 3′ end is shown in lower case g. The arrow denotes the initiation cytidylate. The +4 U/C change in B3(−)26G endscript is shown in bold. (b) RNA synthesis directed by B1(−)26G, B2(−)26G, B3(−)26G and B3(−) + 4 U/C. RdRp reaction products were separated by 12% denaturing PAGE and visualized by autoradiography. The amount of RNA synthesis from various templates relative to B2(−)26G are summarized at the bottom of the autoradiogram. Each number presented is an average from three independent trials.

minus-strand templates *in vitro* (Sivakumaran & Kao, 1999). A more detailed characterization of a functional endscript of 27 nt, B2(−)26G, was undertaken in the present study. In this work we demonstrated that: (1) the preformed RNA secondary structure of B2(−)26G is not specifically required for template recognition by the BMV replicase; (2) the RNA promoter for RNA synthesis requires nucleotides −1, +1 and +2 as a determinant for replicase binding; (3) template sequence beyond the +3 position likely affects RNA synthesis by modulating the polymerization process; and (4) phylogenetic analysis of the viral sequences indicate that the *in vitro* requirements described likely reflect *in vivo* requirements. The results are placed below in context of the mechanism of viral RNA replication.

Genomic and subgenomic plus-strand syntheses

The BMV subgenomic promoter is recognized by the BMV replicase in a sequence-specific manner (Siegel *et al.*, 1997). Results herein demonstrate that the endscript directing genomic plus-strand synthesis also interacts with the viral replicase *via* a mechanism that does not depend on a specific preformed RNA structure. For the subgenomic promoter, the recognition nucleotides include the initiation cytidylate (+1) and at least four positions 3′ of the initiation site (positions −17, −14, −13, and −11) (Siegel *et al.*, 1998) in a region determined to be important for BMV subgenomic RNA synthesis in infected cells (French & Ahlquist, 1988). For the genomic endscript, interaction with the replicase requires the initiation cytidylate and the neighboring −1 and +2 nucleotides. The sequence 3′ to the subgenomic core promoter has been demonstrated to modulate the level of RNA synthesis (Marsh *et al.*, 1988; Adkins *et al.*, 1997). In the genomic endscript, 5′ sequence from +17 to +24 affect the level of RNA synthesis. Finally, both templates are rich in adenylates and uridylates. With the genomic endscript the A + U-rich nature of the template could affect synthesis, but not binding to the replicase (Figure 2(b) and Table 1).

The sequences directing plus-strand syntheses are part of the intermediate synthesized during viral replication. Given sufficient time, these intermediates will either associate intermolecularly with the complementary template to form stable double-stranded RNAs, or fold intramolecularly into a highly structured state (unless there are proteins that affect RNA structure). Since there are no sequences in common for genomic plus and subgenomic RNA synthesis except for the initiation cytidylate, it is likely that both promoters are recognized before stable RNA structures can be formed in the nascent RNA. Based on these and other observations (Bujarski *et al.*, 1994; Adkins

et al., 1997; Siegel *et al*., 1997, 1998; Sivakumaran & Kao, 1999), we propose that the initiation of BMV genomic and subgenomic plus-strand RNA synthesis takes place while the replicase is still associated with the nascent RNA (Figure 10). In this model, when minus-strand RNA synthesis reaches the BMV RNA3 intercistronic sequence encoding the subgenomic promoter, the replicase is faced with a roadblock that signals the replicase to either continue copying the template RNA or reverse course and *de novo* initiate subgenomic RNA synthesis from the nascent RNA. Overcoming this roadblock may depend on the ability of the replicase to successfully pass the intercistronic sequence, which in several bromoviruses contain an unusual homopolymeric sequence (Marsh *et al*., 1988; Allison *et al*., 1988, 1989; Romero *et al*., 1992) or contain stable stem-loop structures in other viruses (Zavriev *et al*., 1996; Koev *et al*., 1999), or stable RNA-RNA interactions (Kim & Hemenway, 1999; Sit *et al*., 1998). The replicases that cannot pass the intercistronic sequence would release the template RNA and initiate subgenomic RNA synthesis from the recently synthesized nascent RNA. The precise initiation position would be determined by the specific key nucleotides present in the subgenomic core promoter (Siegel *et al*., 1997). Bujarski *et al*. (1994) have noted that the BMV intercistronic sequence is a region for frequent template switching, and Nagy *et al*. (1999) have demonstrated that the AU content of a template may promote recombination, a process that depends on template switching. Further support for this hypothesis include the observations by French & Ahlquist (1988), who by placing a series of BMV subgenomic promoters at different positions found that promoters closer to the start site of minus-strand RNA synthesis directed more subgenomic RNA products.

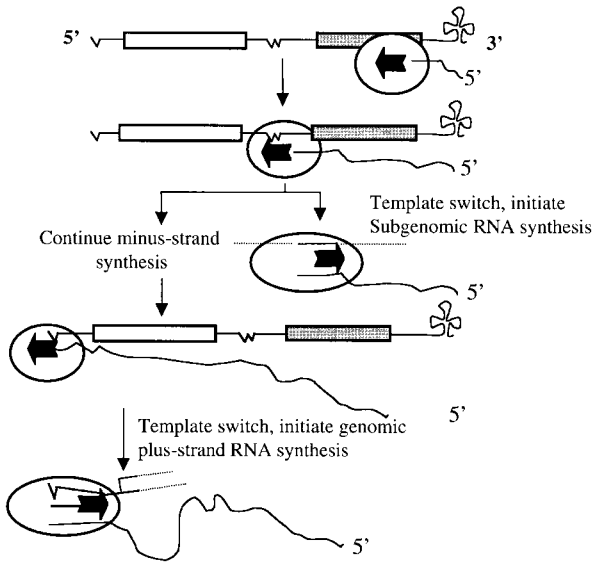


Figure 10. A model for the synthesis of genomic and subgenomic plus-strand RNAs by the BMV replicase. Schematic representation of BMV RNA3 (not intended to be at correct proportions) with the coding sequences denoted by long rectangles. The untranslated 5′, 3′ and the intercistronic region are shown as lines. The replicase is shown as an oval with the direction of synthesis indicated by an arrow. The replicase binds to the 3′ tRNA-like end of genomic RNA and initiates synthesis of the minus-strand RNA. The nascent strand is shown attached to the replicase. Upon reaching the intercistronic region the replicase could either continue minus-strand synthesis or reverse course to initiate subgenomic plus-strand RNA synthesis in a process that require replicase to use nascent minus-strand as template. The replicase could also pass through the homopolymeric region and complete minus-strand synthesis. Upon reaching the end of the template, the replicase adds a non-templated nucleotide by terminal transferase activity. The 3′ non-templated nucleotide, along with the initiation nucleotide signals the replicase to switch to the nascent-strand to initiate genomic plus-strand synthesis.

The replicase molecules that could successfully pass through the intercistronic sequence would proceed forward generating full-length minus-strand RNAs. Upon reaching the end of the viral template, the replicase adds a non-templated nucleotide to the nascent minus-strand RNA by a terminal transferase activity (Siegel *et al*., 1997). The recognition of the initiation site (nucleotides −1, +1 and +2) signals the replicase to initiate genomic plus-strand synthesis from the nascent minus-strand RNA template. The differential accumulation of subgenomic and genomic RNAs *in vivo* (Marsh *et al*., 1991) could be due to a larger number of replicase molecules stalling at the intercistronic promoter region and reversing course to initiate subgenomic RNA synthesis.

The hypothesis that plus-strand RNAs are synthesized from the nascent minus-strand RNA without releasing the polymerase has been proposed for several viral systems. In Coronaviruses, minus-strand RNAs approximately the length of subgenomic RNAs were produced during infection (Sethna *et al*., 1989) and these replication intermediates were found to be transcriptionally active (Stawicki & Stawicki, 1990), suggesting that incompletely synthesized minus-strand RNA may direct plus-strand RNA synthesis. In the polycistronic red clover necrotic virus, the interaction *in trans* between two RNA sequences may prevent the progression of the viral replicase and induce subgenomic RNA synthesis (Sit *et al*., 1998). For the polycistronic potato virus X RNA, a series of repeated sequences that resemble transcription pause/termination sites of DNA templates may act to slow the progression of the replicase during minus-strand synthesis and thereby induce the initiation of plus-strand RNA synthesis (Kim & Hemenway, 1999).

Minus *versus* plus-strand RNA synthesis

The sequence-specific requirements for plus-strand genomic and subgenomic syntheses are

apparently in contrast to the requirements for genomic minus-strand synthesis. Structures at the 3′ end of the viral genomic RNA that are required to direct genomic minus-strand synthesis have been demonstrated in several systems, including BMV (Dreher & Hall, 1988; Pogue & Hall, 1992; Philipenko *et al*., 1996; Song & Simon, 1995; Draper, 1995; Hong *et al*., 1998; Chapman & Kao, 1999). There may be major differences in the mode of recognition of promoters for minus *versus* plus-strand RNA synthesis. For plus-strand RNA viruses the genomic RNA is present in the infected cell prior to its translation and assembly of the replicase. Since all molecules prefer to exist in a lower energy state, the viral genomic RNA would possess highly ordered secondary and tertiary structures, which may be formed with the assistance of sequence-non-specific RNA-folding chaperons or sequence-specific RNA-binding cofactors (Weeks, 1997), such as the BMV 1a protein (Sullivan & Ahlquist, 1999). Hence, the replicase needs to recognize the highly structured genomic RNA to initiate minus-strand synthesis. The differences in the incidence of guanylates and cytidylates (Figure 8) provide another indication of the differences between minus and plus-strand RNA synthesis. The initiation of minus-strand RNA synthesis may also be affected by upstream sequence (Carpenter *et al*., 1995; Sullivan & Ahlquist, 1999; Chapman *et al*., 1998).

Template sequence for initiation and elongation

An interesting conclusion from our analysis of the BMV genomic minus-strand RNA endscript is that template sequence may regulate the amount of RNA synthesis in a manner independent of replicase-promoter interaction. This is different from the general notion of promoter-strength, which implies a modulation of the frequency of initiation events, and necessitates an emphasis on the frequency of successful transitions into processive elongation. Transcription studies using DNA templates indicate that the template sequence could affect the proportion of RNAs that are successfully transcribed to completion (Keene & Luse, 1999; Nudler *et al*., 1997; Von Hippel & Yager, 1991). A similar strategy may regulate RNA synthesis from RNA templates. An assumption of this hypothesis is that the viral sequence contains nucleotides that are selected at specified positions to enable proper transition from initiation to elongation phase of RNA synthesis and the optimal level of RNA accumulation. The difference in BMV RNA1 and RNA2 at the +4 position provides an example where a single-nucleotide change in the viral RNA can account for a difference in RNA accumulation (Figure 9).

The BMV replicase is known to increase its affinity for interaction with the template RNA as synthesis progresses (Sun & Kao, 1997a,b). The increased affinity may affect the stability of the template-nascent RNA duplex interaction, i.e. too stable an interaction during initiation may prevent the replicase from making the transition from the initiation site, resulting in decreased RNA synthesis. Consistent with this testable hypothesis, we observed that guanylates and cytidylates negatively affected RNA synthesis when they were within four positions of the initiation nucleotide. Inconsistent with this hypothesis, however, is the fact that substitutions with guanylate and cytidylates do not have identical effects (Figure 2(b)). Stability of the template-nascent strand duplex is also affected by the contributing effects of the neighboring nucleotides (Turner *et al*., 1988), and a cytidylate substitution at a given position may not have the same thermodynamic effect as a guanylate. Alternatively, thermodynamic parameters may not be the sole factor affecting viral RNA synthesis. Another possible effect could be that the template sequence might also have evolved to influence translation.

The transition from initiation to elongation phase during BMV RNA synthesis takes place at positions +8 to +10 (Sun & Kao, 1997a,b; Adkins *et al*., 1998). The observation that guanylates and cytidylates are naturally enriched between position +6 to +10 suggests that these positions are acting in a manner different than the positions required for initiation. Perhaps a tighter template-nascent RNA interaction is required for efficient transition from initiation to elongation (Figure 2(b)). Consistent with this, the introduction of a guanylate at position +8 or +10 in B2(−)26G increased RNA synthesis above wild-type levels (Figure 2(b)).

Concluding comments

Our results indicating that viral replicase-promoter interaction being independent of the initial RNA secondary structure may be interpretable by two general models. First, the recognition of the promoter may be directed by the sequence of the RNA. Sequence-specific recognition has been previously demonstrated for the bacteriophage T4 translational repressor RegA protein (Webster & Spicer, 1990; Brown *et al*., 1997). Second, it is possible that the recognition of the genomic minus-strand RNA by the viral replicase could take by inducing tertiary interaction between protein and RNA which is tolerant of some differences in the RNA secondary structure. An induced-fit mechanism has been suggested for a number of RNA-binding proteins, including the splicing factor, U1A, the ribosomal S15 protein, and the binding of the hepatitis B encapsidation signal by the P protein (Gubser & Varani, 1996; Batey & Williamson, 1996; Beck & Nassal, 1997). There is also evidence for an induced fit mechanism for the recognition of the BMV subgenomic promoter by the BMV replicase. However, in our studies, the contacts in the RNA will induce a conformational change in the replicase, resulting in different nucleotide moieties being recognized (Adkins & Kao, 1998; Stawicki & Kao, 1999). A subsequent change in the RNA sec-

ondary or tertiary structure after replicase binding to the RNA has not been ruled out. This induced fit mechanism is compatible with the ability of the replicase to recognize the subgenomic and genomic promoters without the release of the replicase during minus-strand RNA synthesis.

Experimental Procedures

NMR spectroscopy

Wild-type and mutant endscripts were prepared for NMR analysis by *in vitro* transcription using phage T7 RNA polymerase and synthetic DNA templates (Milligan *et al*., 1987; Wyatt *et al*., 1991). The transcribed RNAs were purified by denaturing polyacrylamide gel electrophoresis and eluted from the gel slices using a electroelution trap (Scheicher & Schuell). The eluted RNA was concentrated by ethanol precipitation and dialyzed against the NMR buffer (10 mM sodium phosphate, 0.1 mM EDTA, 100 mM NaCl, pH 6.5) for 48 hours. Finally, the samples were dried down and dissolved in the same volume of 90 % H_2O/10 % ^2H_2O. The RNA samples are at 0.4 to 1 mM for 1D spectra and 2.0 mM for 2D H_2O NOESY experiment.

Exchangeable proton 1D spectra were recorded on a Bruker AMX-600 spectrometer equipped with a Bruker triple resonance triple-axis gradient probe. Solvent suppression was achieved using the jump-return method with the excitation maximum set around the middle of the usual chemical shift range of the imino protons (12.5 ppm) (Plateau & Gueron, 1982). A spectral width of 12,000 Hz was used and 4096 complex data points were collected with a relaxation delay of two seconds. Depending on the sample concentration, 32 to 1024 scans were taken. In analysis of the mutant endscripts, all of the gains were set at the same level. The 2D H_2O NOESY spectra were collected independently using a Bruker AMX-600 spectrometer equipped with a Bruker triple resonance triple-axis gradient probe and a Bruker DRX-500 spectrometer equipped with a Nalorac triple resonance *z*-axis gradient probe. The results from the two spectrometers were consistent. H_2O NOESY spectra were collected at mixing times 200 ms and 300 ms at 20 °C. A total of 400 to 440 t_1 blocks were taken, for each block, 32 scans were taken with a relaxation delay of two seconds. Solvent suppression was accomplished by replacing the last pulse in NOESY pulse sequence with a jump-return pulse sequence and applying *z*-gradient pulse during the mixing time. Proton chemical shifts were indirectly referenced to DSS using the water resonance. All NMR spectra were processed using FELIX 95.0 (Biosym Technologies, Inc.).

Native gel analysis

RNA conformation native gel analysis was performed in 12 % polyacrylamide gels (40:2 acrylamide to bisacrylamide) containing 0.5× TBE (Maniatis *et al*., 1982). After electrophoresis for 20 minutes, 6 µl of a 6 µM sample was loaded in a dye containing 0.5× TBE, 5 % (v/v) glycerol and bromophenol blue. Electrophoresis was performed at a constant 200 volts until bromophenol blue migrated for 6-8 cm. The gel was then stained with Toluidine blue to visualize the RNAs.

RNA denaturation analysis

The protocol used was as described by Puglisi & Tinoco (1989) using a Gilford spectrophotometer model 250 and a model 2527 thermoprogrammer. Briefly, the sample concentration was determined by calculations using Beer-Lambert Law with a calculated extinction coefficient for B2(−)26G of 286. After adjusting the absorbance of the samples to be between 0.8 and 1.2 at 20 °C, data points were taken with a change in temperature of 0.5 deg. C/minute. The results were analyzed and plotted using a Kaleidagraph software package. t_m for each was calculated using a derivative of the raw data.

Synthesis of RNA templates for RdRp assay

cDNA copies of (−)-strand 3′ ends encompassing the first 26 nt of BMV RNA 1, 2 and 3 were generated by PCR amplification. Pairs of oligonucleotide primers were used, one of which contained a T7 promoter to enable transcription by T7 RNA polymerase. Following transcription, the RNAs were electrophorsed in a denaturing PAGE, and the band of correct molecular mass identified by staining with ethidium bromide and excised with a razor. The gel was then crushed to elute the RNA in a solution of 0.3 M sodium acetate. Concentrations of RNA were determined by spectrometry and staining with Toluidine blue following denaturing PAGE.

RNA synthesis assays

BMV replicase was prepared from infected barley as described by Sun *et al*. (1996). Replicase activity assays were carried out as described by Adkins *et al*. (1998). Briefly, each assay consisted of a 40 µl reaction mixture containing 20 mM sodium glutamate (pH 8.2), 12 mM dithiothreitol, 4 mM $MgCl_2$, 0.5 % (v/v) Triton X-100, 2 mM $MnCl_2$, 200 µM ATP, 500 µM GTP, 200 µM UTP, 242 nM [α-^{32}P]CTP (400 Ci/mmol, 10 mCi/ml, Amersham), the desired amount of template and 7 µl of RdRp. Following incubation for 90 minutes at 30 °C, the reaction products were extracted with phenol/chloroform (1:1, v/v) and precipitated with six volumes of ethanol, 10 µg of glycogen, and 0.4 M final concentration of ammonium acetate. Loading buffer (45 % (v/v) deionized formamide, 1.5 % (v/v) glycerol, 0.04 % (w/v) bromophenol blue, and 0.04 % (w/v) xylene cyanol) was added to the products and denatured by heating at 90 °C for three minutes prior to electrophoresis on 12 % acrylamide-7 M urea denaturing gels. All gels were dried and exposed to film at −80 °C and the amount of label incorporated into newly synthesized RNAs was determined with a Phosphorimager (Molecular Dynamics).

Phylogenetic analysis

The template sequences proximal to the plus and minus-strand initiation sites were found at the NCBI website (http://www.ncbi.nlm.nih.gov). Viral sequences used were those indicated as complete by the Entrez database. The species and accession numbers are: alfalfa mosaic virus RNA1 (L00163 J02000), alfalfa mosaic virus RNA2 (K02702 J02002), alfalfa mosaic virus RNA3 (K03542), apple chlorotic leaf spot virus (D14996), apple stem grooving virus (D14995 S47260), barley mild mosaic virus RNA1 (AJ242725), barley mild mosaic virus RNA2 (X75933), beet necrotic yellow vein mosaic

virus RNA1 (D00115), beet necrotic yellow vein mosaic virus RNA2 (X04197), beet necrotic yellow vein mosaic virus RNA3 (M36894), beet necrotic yellow vein mosaic virus RNA4 (M36896), barley stripe mosaic virus RNA alpha (J04342), barley stripe mosaic virus RNA beta (X03854), barley stripe mosaic virus RNA gamma (U13916), beet yellows virus (X73476), broad bean mottle virus RNA1 (M65138), broad bean mottle virus RNA2 (M64713), broad bean mottle virus RNA3 (M60291), brome mosaic virus RNA1 (X02380 K02706), brome mosaic virus RNA2 (X01678 K02707), brome mosaic virus RNA3 (J02042 J02043), carnation mottle virus (X02986), carrot mottle mimic umbravirus RNA1 (U57305), cowpea chlorotic mottle virus RNA1 (M65139), cowpea chlorotic mottle virus RNA2 (M28817), cowpea chlorotic mottle virus RNA3 (M28818), cowpea mosaic virus RNA B (X00206), cowpea mosaic virus RNA M (X00729), cucumber mosaic virus RNA1 (D12537 D01199), cucumber mosaic virus RNA2 (D12538 D01200), cucumber mosaic virus RNA3 (D10539), cymbidium mosaic virus (U62963), grapevine virus A (X75433), olive latent virus 1 (X85989), olive latent virus 2 RNA1 (X94346), olive latent virus 2 RNA2 (X94347), olive latent virus 2 RNA3 (X76993), olive latent virus 2 RNA4 (X77115), papaya mosaic virus (D13957 D00580), peanut clump virus RNA1 (X78602), peanut clump virus RNA2 (L07269), peanut mottle virus (AF023848), pepper mottle virus (M96425), potato mop-top virus RNA3 (D16193), potato virus X (D00344), potato virus Y (A08776), raspberry bushy dwarf virus RNA1 (S51557), ryegrass mosaic virus (AF035818), soil-borne wheat mosaic virus RNA1 (L07937), soil-borne wheat mosaic virus RNA2 (L07938), soybean mosaic virus (S42280), sweet clover necrotic mosaic virus (S46028), tobacco mosaic virus (D13438), tobacco necrosis virus (D00942), tobacco rattle virus RNA1 (D00155), tobacco rattle virus RNA2 (X03686), tobacco streak virus RNA3 (X00435), tomato bushy stunt virus (M21958 M31019), turnip yellow mosaic virus (J04373) and yam mosaic virus (U42596).

Acknowledgments

This work would not have been possible without the generous support of Professor Ignacio Tinoco Jr, in whose laboratory the NMR and RNA melting experiments were performed. We thank M. Schmitz, R. Gonzalus, for helpful discussions. Funding was provided in part by a National Health Institute grant GM10840 and Department of Energy grant DE-FG03-86ER60406 to I.T. and by the National Science Foundation grant MCB9507344 to C.K.

References

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

Adkins, S., Siegel, R., Sun, J. H. & Kao, C. C. (1997). Minimal templates directing accurate initiation of subgenomic RNA synthesis in vitro by the brome mosaic virus RNA-dependent RNA polymerase. *RNA,* **3**, 634-647.

Adkins, S., Stawicki, S. S., Faurote, G., Siegel, R. & 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.

Ahlquist, P. (1992). Bromovirus RNA replication and transcription. *Curr. Opin. Genet. Dev.* **2**, 71-76.

Ahlquist, P., Dasgupta, R. & Kaesberg, P. (1981). Near identity of 3′ RNA secondary structure in bromoviruses and cucumber mosaic virus. *Cell,* **23**, 183-189.

Allison, R. F., Janda, M. & Ahlquist, P. (1988). Infectious *in vitro* transcripts from cowpea chlorotic mottle virus cDNA clones and exchange of individual components with brome mosaic virus. *J. Virol.* **62**, 3581-3588.

Allison, R. F., Janda, M. & Ahlquist, P. (1989). Sequence of cowpea chlorotic mottle virus RNAs 2 and 3 and evidence of a recombination event during bromovirus evolution. *Virology,* **172**, 321-330.

Batey, R. T. & Williamson, J. R. (1996). Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16S RNA. II. Specificity determinants of protein-RNA recognition. *J. Mol. Biol.* **261**, 550-567.

Beck, J. & Nassal, M. (1997). Sequence and structure-specific determinations in the interaction between RNA encapsidation signals and reverse transcriptase of avian hepatitis B viruses. *J. Virol.* **71**, 4971-4980.

Behrens, S. E., Tomei, L. & De Franscesco, R. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **15**, 12-22.

Binghum, P. M. (1997). Cosuppression comes to the animals. *Cell,* **90**, 385-387.

Brown, D., Brown, J., Kang, C.-H., Gold, L. & Allen, P. (1997). Single-stranded RNA recognition by the bacteriophage T4 translational repressor, RegA. *J. Biol Chem.* **272**, 14969-14974.

Buck, K. W. (1996). Comparison of the replication of positive-strand RNA viruses of plants and animals. *Advan. Virus Res.* **47**, 159-251.

Bujarski, J. J., Nagy, P. D. & Flasinski, S. (1994). Molecular studies of genetic RNA-RNA recombination in brome mosaic virus. *Advan. Virus Res.* **43**, 275-302.

Carpenter, C. D., Oh, J.-W., Zhang, C. & Simon, A. E. (1995). Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. *J. Mol. Biol.* **245**, 608-622.

Chapman, M. R. & Kao, C. C. (1999). A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. *J. Mol. Biol.* **286**, 709-720.

Chapman, M., Rao, A. L. N. & Kao, C. C. (1998). Sequences 5′ of the conserved tRNA-like promoter modulate initiation of minus-strand synthesis by the brome mosaic virus RNA-dependent RNA polymerase. *Virology,* **252**, 458-467.

Cogoni, C. & Macino, G. (1999). Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature,* **399**, 166-169.

Draper, D. E. (1995). Protein-RNA recognition. *Annu. Rev. Biochem.* **64**, 593-620.

Dreher, T. W. & Hall, T. C. (1988). Mutational analysis of the sequence and structural requirements in brome mosaic virus for minus-strand promoter activity. *J. Mol. Biol.* **201**, 31-40.

Faragher, S. G. & Degarno, L. (1986). Regions of conservation and divergence in the 3′ untranslated

sequences of genomic RNA from Ross River virus isolates. *J. Mol. Biol.* **190**, 141-148.
Faragher, S. G., Meek, A. D. J., Rice, C. M. & Delgarno, L. (1988). Genome sequences of a mouse-avirulent and mouse-virulent strain of Ross river virus. *Virology,* **163**, 509-526.
French, R. & Ahlquist, P. (1988). Characterization and engineering of subgenomic RNA. *J. Virol.* **62**, 2411-2420.
Goelet, P., Lomonosoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J. & Karn, J. (1982). Nucleotide sequence of tobacco mosaic virus. *Proc. Natl Acad. Sci. USA,* **79**, 5818-5822.
Goldbach, R., LeGall, O. & Wellink, J. (1991). Alpha-like viruses in plants. *Semin. Virol.* **2**, 19-25.
Gubser, C. C. & Varani, G. (1996). Structure of the polyadenylation regulatory element of the human U1A pre-mRNA 3′-untranslated region and interaction with the U1A protein. *Biochemistry,* **35**, 2253-2267.
Gunn, M. R. & Symons, R. H. (1980). Sequence homology at the 3′-termini of the four RNAs of the alfalfa mosaic virus. *FEBS Letters,* **109**, 145-150.
Hardy, S. F., German, T. L., Loesch-Fries, L. S. & Hall, T. C. (1979). Highly active template-specific RNA-dependent RNA polymerase from barley leaves infected with brome mosaic virus. *Proc. Natl Acad. Sci. USA,* **76**, 4956-4960.
Hong, Y., Cole, T. E., Brasier, C. M. & Buck, K. W. (1998). Novel structures of two virus-like RNA elements from a diseased isolate of the Dutch elm disease fungus *Ophiostoma novo-ulmi*. *Virology,* **242**, 80-89.
Jaeger, J. A., Turner, D. H. & Zuker, M. (1989). Improved predictions of secondary structures for RNA. *Proc. Natl Acad. Sci. USA,* **86**, 7706-7710.
Keene, R. & Luse, D. (1999). Initially transcribed sequences strongly affect the extent of abortive initiation by RNA polymerase II. *J. Biol. Chem.* **274**, 11526-11534.
Kim, K.-H. & Hemenway, C. L. (1999). Long-distance RNA-RNA interactions and conserved sequence elements affect potato virus X RNA accumulation. *RNA,* **5**, 636-645.
Koev, G., Mohan, B. R. & Miller, W. A. (1999). Primary and secondary structural elements required for synthesis of barley yellow dwarf virus subgeomic RNA1. *J. Virol.* **73**, 2876-2885.
Levinson, R., Strauss, J. H. & Strauss, E. G. (1990). Determination of the complete nucleotide sequence of the genomic RNA of O'Nyong-nyong virus and its use in the construction of phylogenetic trees. *Virology,* **175**, 110-123.
Maniatis, S. T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Marsh, L. E., Dreher, T. W. & Hall, T. C. (1988). Mutational analysis of the core and modulator sequences of the BMV RNA3 subgenomic promoter. *Nucl. Acids Res.* **16**, 981-995.
Marsh, L. E., Huntley, C. C., Pogue, G. P., Connell, J. P. & Hall, T. C. (1991). Regulation of (+):(−)-strand asymmetry in replication of brome mosaic virus RNA. *Virology,* **182**, 76-83.
Miller, W. A., Dreher, T. W. & Hall, T. C. (1985). Synthesis of brome mosaic virus subgenomic RNA in vitro by internal initiation on (−)-sense genomic RNA. *Nature,* **313**, 68-70.
Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **15**, 8783-8798.
Nagy, P. D., Ogiela, C. & Bujarski, J. J. (1999). Mapping sequences active in homologous RNA recombination in brome mosaic virus: prediction of recombination hotspots. *Virology,* **254**, 92-104.
Neufeld, K. L., Richards, O. C. & Ehrenfeld, E. (1991). Purification, characterization, and comparison of poliovirus RNA polymerase from native and recombinant sources. *J. Biol. Chem.* **266**, 24212-24219.
Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell,* **89**, 33-41.
O'Reilly, E. K. & Kao, C. C. (1998). Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology,* **252**, 287-303.
Ou, J. H., Rice, C. M., Dalgarno, L., Strauss, E. G. & Strauss, J. H. (1982a). Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. *Proc. Natl Acad. Sci. USA,* **79**, 5235-5239.
Ou, J. H., Trent, D. H. & Strauss, J. H. (1982b). The 3′-noncoding regions of alphavirus RNAs contain repeating sequences. *J. Mol. Biol.* **156**, 719-730.
Ou, J. H., Strauss, E. G. & Strauss, J. H. (1983). The 5′-terminal sequences of the genomic RNAs of several alphaviruses. *J. Mol. Biol.* **168**, 1-15.
Philipenko, E. V., Poperechny, K. V., Maslova, S. V., Melchers, W. J., Slot, H. J. & Agol, V. I. (1996). *Cis*-elements, oriR, involved in the initiation of (−)-strand poliovirus RNA: a quasi-globular multidomain RNA structure maintained by tertiary (''kissing'') interactions. *EMBO J.* **15**, 5428-5436.
Plateau, J. R. & Gueon, M. (1982). Exchangeable proton NMR without base-line distortion, using strong-pulse sequences. *J. Am. Chem. Soc.* **104**, 7310-7311.
Poch, O., Blumberg, B. M., Bougueleret, L. & Tordo, N. (1990). Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J. Gen. Virol.* **71**, 1153-1162.
Pogue, G. P. & Hall, T. C. (1992). The requirement for a stem-loop structure in brome mosaic virus replication supports a new model for viral positive-strand RNA initiation. *J. Virol.* **66**, 674-684.
Puglisi, J. D. & Tinoco, I. (1989). Absorbence melting curves of RNA. *Methods Enzymol.* **180**, 304-325.
Quadt, R. & Jaspars, E. M. J. (1990). Purification and characterization of brome mosaic virus RNA-dependent RNA polymerase. *Virology,* **178**, 189-194.
Romero, J., Dzianott, A. M. & Bujarski, J. J. (1992). The nucleotide sequence and genome organization of the RNA2 and RNA3 segments in broad bean mottle virus. *Virology,* **187**, 671-681.
Rumenapf, T., Strauss, E. G. & Strauss, J. H. (1995). Aura virus is a new world representative of Sindbis-like viruses. *Virology,* **208**, 621-633.
Schiebel, W., Pélissier, T., Riedel, L., Thelmeir, S., Schiebel, R., Lottspeich, F., Sänger, H. L. & Wassenegger, M. (1998). Isolation of an RNA-directed RNA polymerase-specific cDNA clone from tomato. *Plant Cell,* **10**, 2087-2101.
Sethna, P. B., Hung, S. L. & Brian, D. A. (1989). Coronavirus subgenomic minus-strand RNAs and potential for mRNA replicons. *Proc. Natl Acad. Sci. USA,* **86**, 5626-5630.

Shirako, Y., Niklasson, B., Dalrymple, J. M., Strauss, E. G. & Strauss, J. H. (1991). Structure of the Ockelbo virus genome and its relationship to other Sindbis viruses. *Virology,* **182**, 753-764.

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

Siegel, R. W., Bellon, L., Beigelman, L. & 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.

Siegel, R. W., Bellon, L., Beigelman, L. & Kao, C. C. (1999). Use of DNA, RNA and chimeric templates by a viral RNA-dependent RNA polymerase. *J. Virol.* **73**, 6424-6429.

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

Sivakumaran, K. & Kao, C. C. (1999). Initiation of genomic positive strand synthesis from DNA and RNA templates by a viral RNA-dependent RNA polymerase. *J. Virol.* **73**, 6415-6423.

Song, C. & Simon, A. E. (1995). Requirement of a 3′-terminal stem-loop in *in vivo* transcription by an RNA dependent RNA polymerase. *J. Mol. Biol.* **254**, 6-14.

Stawicki, S. G. & Stawicki, D. L. (1990). Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.* **64**, 1050-1056.

Stawicki, S. S. & Kao, C. C. (1999). Spatial requirements for promoter recognition by a viral RNA-dependent RNA polymerase. *J. Virol.* **73**, 198-204.

Sullivan, M. L. & Ahlquist, P. (1999). A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. *J. Virol.* **73**, 2622-2632.

Sun, J. H. & Kao, C. C. (1997a). RNA synthesis by the brome mosaic virus RNA-dependent RNA polymerase: transition from initiation to elongation. *Virology,* **233**, 63-73.

Sun, J. H. & Kao, C. C. (1997b). Characterization of RNA products associated with or aborted by a viral RNA-dependent RNA polymerase. *Virology,* **236**, 348-353.

Sun, J. H., Adkins, S., Faurote, G. & Kao, C. C. (1996). Initiation of (−)-strand RNA synthesis catalyzed by the BMV RNA-dependent RNA polymerase: synthesis of oligonucleotides. *Virology,* **226**, 1-12.

Turner, D. H., Sugimoto, N. & Freier, S. M. (1988). RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 167-192.

Van Dyke, T. A. & Flanegan, J. B. (1980). Identification of poliovirus polypeptide p63 as a soluble RNA-dependent RNA polymerase. *J. Virol.* **35**, 732-740.

Varani, G. & Tinoco, I. (1991). RNA structure and NMR spectroscopy. *Quart. Rev. Biophys.* **24**, 479-532.

Volloch, V., Schweitzer, B. & Rits, S. (1996). Antisense globin RNA in mouse erythroid tissues: structure, origin, and possible function. *Proc. Natl Acad. Sci. USA,* **93**, 2476-2481.

Von Hippel, P. H. & Yager, T. D. (1991). Transcription elongation and termination are competitive kinetic processes. *Proc. Natl Acad. Sci. USA,* **88**, 2307-2311.

Webster, K. R. & Spicer, E. K. (1990). Characterization of bacteriophage T4 RegA protein-nucleic acid interaction. *J. Biol. Chem.* **265**, 19007-19014.

Weeks, K. M. (1997). Protein-facilitated RNA folding. *Curr. Opin. Struct. Biol.* **7**, 336-342.

Wyatt, J. R., Chastin, M. & Puglisi, J. D. (1991). Synthesis and purification of large amounts of RNA oligonucleotides. *BioTechniques,* **11**, 764-769.

Yuan, Z. H., Kumar, U., Thomas, H. C., Wen, Y. M. & Monjardino, J. (1997). Expression, purification and partial characterization of HCV RNA polymerase. *Biochem. Biophys. Res. Commun.* **232**, 231-235.

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

Edited by D. E. Draper

(Received 29 July 1999; received in revised form 11 October 1999; accepted 14 October 1999)