Distinct Viral Sequence Elements Are Necessary for Expression of *Tomato golden mosaic virus* Complementary Sense Transcripts That Direct AL2 and AL3 Gene Expression

Chia-Yi Shung, Janet Sunter, Shyam S. Sirasanagandla, and Garry Sunter

Department of Biology, The University of Texas at San Antonio, San Antonio, TX 78249, U.S.A.

Submitted 26 May 2006. Accepted 13 July 2006.

Transient expression studies using *Nicotiana benthamiana* protoplasts and plants have identified sequences important for transcription of complementary sense RNAs derived from *Tomato golden mosaic virus* (TGMV) DNA component A that direct expression of AL2 and AL3. Transcription of two complementary sense RNAs, initiating at nucleotides 1,935 (AL1935) and 1,629 (AL1629), is directed by unique sequences located upstream of each transcription initiation site. One element is located between 28 and 124 nucleotides (nt) upstream of the AL1935 transcription start site, which differs from a second element located 150 nt downstream, between 129 and 184 nt upstream of the AL1629 transcription start site. Transcription initiation at nucleotide 1,935 is lower than that at nucleotide 1,629 as determined by run-on transcription assays, and the resulting transcript is only capable of expressing AL3. The transcript initiating at nucleotide 1,629 is capable of directing expression of both AL2 and AL3, although expression of AL3 is up to fourfold greater than that for AL2. Nuclear factors purified from tobacco suspension cells bind to sequences upstream of both AL1935 and AL1629, correlating with the ability of these sequences to direct gene expression. Thus, in tobacco, regulatory sequences direct transcription of two unique TGMV messenger RNAs that differentially express AL2 and AL3.

Additional keywords: Geminivirus, promoter, regulation.

The *Geminiviridae* are unique eukaryotic plant viruses characterized by twin icosahedral particles encapsidating a circular, single-stranded (ss)DNA genome. Viral replication occurs in the nuclei of infected plant cells primarily by rolling circle replication (RCR) using host DNA polymerases, producing double-stranded (ds)DNA replicative form (RF) intermediates (Stenger et al. 1991). There also is evidence that replication can occur via a recombination-dependent mechanism, as demonstrated for *Abutilon mosaic virus* (AbMV), although it currently is not known whether this mechanism requires viral gene products (Preiss and Jeske 2003). Viral transcription occurs from RF DNA templates generated during RCR utilizing host

RNA polymerase II (Bisaro 1996; Hanley-Bowdoin et al. 1999). This reliance on cellular RNA polymerase makes geminiviruses a valuable model system for the study of transcription in plants.

Geminiviruses have small genomes, approximately 2.5 to 3.0 kb in size, and differ with respect to host range, insect vector, number and organization of genome components, and mechanisms of transcriptional regulation (Briddon and Markham 1995; Rybicki 1995). *Tomato golden mosaic virus* (TGMV) belongs to the genus *Begomovirus*, with two DNA components (A and B), both of which are required for infectivity (Hamilton et al. 1983). DNA A encodes functions required for replication (*AL1* and *AL3*) and encapsidation (*AL2* and *ARI*) of the virus (Rogers et al. 1986; Sunter et al. 1987). Genes on the B component (*BR1* and *BL1*) encode functions for viral movement (Brough et al. 1988; Jeffrey et al. 1996).

Geminiviruses conform to a general strategy of DNA virus transcription programs where early gene products activate expression of viral genes required later in the replication cycle. The AL2 gene product (also known as AC2) of TGMV is necessary for efficient transcription from the coat protein (*CP*) (*ARI*) and *BR1* nuclear shuttle protein gene promoters (Sunter and Bisaro 1991, 1992). Analysis of *CP* promoter constructs in transgenic plants show that AL2 directs *CP* expression in all plant cell types via activation in mesophyll cells and derepression in phloem cells (Sunter and Bisaro 1997). A 5' intergenic region of approximately 230 bp (conserved in both components of begomoviruses) separates divergent coding regions, and contains elements mediating bidirectional transcription (Fontes et al. 1994; Frischmuth et al. 1991; Hanley-Bowdoin et al. 1990; Lazarowitz 1992; Petty et al. 1988; Sunter and Bisaro 1989; Sunter et al. 1993; Townsend et al. 1985). In TGMV, virus-specific messenger (m)RNAs transcribed from both DNA components have been characterized, and are indicative of a bidirectional transcription strategy (Hanley-Bowdoin et al. 1989; Sunter and Bisaro 1989; Sunter et al. 1989). The viral sense transcription unit of TGMV DNA A comprises a single RNA species spanning the *CP* gene. In contrast, the complementary sense transcription unit of TGMV DNA A is more complex, consisting of multiple overlapping RNAs with different 5' ends, all of which are 3' co-terminal (Fig. 1A) (Hanley-Bowdoin et al. 1988; Sunter and Bisaro 1989). The only RNA capable of producing a functional AL1 protein initiates at nucleotide 62 (AL62), but this RNA also has the potential to code for the AL2 and AL3 proteins. Two smaller RNAs, initiating at nucleotides 1,935 and 1,629 (AL1935 and AL1629, respectively), specify both the AL2 and AL3 coding regions. Northern blot analysis and fine mapping of transcrip-

Corresponding author: G. Sunter; Telephone: (210) 458-5479; Fax: (210) 458-5561; garry.sunter@utsa.edu

Current address of S. S. Sirasanagandla: UT Southwestern Medical Center Department of Internal Medicine Dallas, TX 75235, U.S.A.

tion start sites suggest that AL1629 is the more abundant transcript in TGMV-infected plants (Hanley-Bowdoin et al. 1988; Sunter and Bisaro 1989). No RNA capable of coding for AL3 alone has been identified to date and no evidence of splicing has been found. All of the viral-specific complementary sense mRNAs identified to date in TGMV are potentially multicistronic.

A minimal sequence recently has been identified that is both necessary and sufficient for TGMV CP promoter activity (Sunter and Bisaro 2003). The sequence has a bipartite arrangement, with both elements being required for activation of the promoter by AL2. The TGMV AL1 protein negatively regulates expression from its own promoter by interaction with its cognate binding site that lies between the transcription start site for AL1 (AL62) and a consensus RNA polymerase II TATAA sequence (Eagle et al. 1994; Sunter et al. 1993). Repression of the AL1 promoter is not dependent on replication, because mutations in the AL1 protein that are deficient for DNA replication have no effect on repression (Eagle et al. 1994). Although the CP and AL1 promoters of TGMV have been studied in some depth, very little information is available on promoters that might regulate AL2 and AL3 expression. A previous study of *African Cassava mosaic virus* (ACMV) suggests that sequences 5' to the AC2 gene could direct expression of a reporter gene in protoplasts, although the level of activation observed

was approximately threefold lower than that of the AC1 promoter (Zhan et al. 1991). However, the data does not directly determine whether these sequences direct expression of AC2, AC3, or both, because the authors used a transcriptional fusion to the β -glucuronidase (*GUS*) reporter gene. Recently, a strong promoter was mapped within the AC1 coding region of *Mung bean yellow mosaic virus* (MYMV) that appears to direct expression of both AC2 and AC3 (Shivaprasad et al. 2005).

Geminivirus RF DNA is associated with host nucleosomes as a viral minichromosome that exists in at least two defined structures, with nucleosome-free gaps accessible for interactions with additional host factors (Pilartz and Jeske 1992, 2003). In AbMV DNA A, one gap co-localized with sequences corresponding to the AL1 promoter and the origin of replication. A second gap was located within 350 to 490 bp upstream of the translation initiation codon for AL2 and AL3, respectively, which could suggest the presence of promoter sequences in this region.

Considering the importance of the AL2 and AL3 genes in replication and transcriptional activation, it is important to understand how expression of these two genes is regulated in geminiviruses. Several questions relating to expression of the complementary sense genes in geminiviruses prompted this study. First, viral sequences required for expression of TGMV AL2 and AL3 have not been defined. Second, it is not known

Fig. 1. Deletion derivatives of the *Tomato golden mosaic virus* (TGMV) DNA A genome component. **A**, The linear map illustrates the wild-type TGMV DNA A genome component, indicating relevant restriction sites with nucleotide coordinates given in parentheses. Numbering is according to the TGMV sequence determined by Hamilton and associates (1984). Open boxes indicate coding regions and arrows designate transcripts and direction of transcription, with initiation site indicated. The hatched box represents the approximately 230-bp intergenic region, solid triangles represent polyadenylation signals, and open triangles indicate TATA boxes. **B**, The linear maps illustrate derivatives of TGMV DNA A, containing different 5' flanking regions of the TGMV AL2 and AL3 genes cloned as translational fusions with the β -glucuronidase (*GUS*) reporter gene (open bar) using the BamHI site. The nopaline synthase (*nos*) 3' polyadenylation signal is indicated. Deletion endpoints are indicated using relevant restriction sites. Clone designations are shown on the right-hand side of the diagram and indicate the amount of 5' flanking sequence relative to the translation initiation site for either AL2 or AL3. Arrows designate transcripts and direction of transcription for AL1935 and AL1629 only. Diagrams are not drawn to size.

whether the AL2 and AL3 proteins are expressed from separate RNAs that are differentially regulated by independent sequences, or are expressed from the potentially multicistronic RNAs. In this article, we present results that allow us to describe minimal viral sequences that are necessary for transcription of two complementary sense RNAs, AL1935 and AL1629, that differentially direct expression of TGMV AL2 and AL3.

RESULTS

Identification of viral sequences required for expression of TGMV AL2.

Both AL2 and AL3 potentially can be expressed from three TGMV complementary sense mRNAs that initiate at nucleotides 62, 1,935, and 1,629 (Sunter and Bisaro 1989). To define sequences necessary and sufficient for expression of TGMV AL2, a series of 5'-truncated promoters linked to the GUS reporter in a translational fusion were constructed (Fig. 1B). It was essential to use translational fusions to distinguish sequences required for AL2 and AL3 gene expression due to the fact that these two genes overlap in different open reading frames (ORFs) for approximately 60% of their length. Constructs were transfected into protoplasts prepared from a recently described *Nicotiana benthamiana* suspension culture cell line (Sunter and Bisaro 2003). A construct that generates replicating, native TGMV DNA A, lacking the GUS reporter gene (pTGA26), served as a background control. Extracts were prepared 3 days posttransfection and fluorometric GUS assays performed as described (Sunter and Bisaro 1991). Initial experiments determined that maximal expression was detected at 3 days posttransfection; therefore, expression from promoter-reporter constructs was analyzed at this timepoint, although similar results were obtained at other timepoints (data not shown).

As illustrated (Fig. 2), extracts from protoplasts transfected with AL2 promoter deletions exhibited varying GUS activity depending on the amount and context of remaining promoter sequence. Significantly greater activity than background (analysis of variance [ANOVA], $P < 0.05$) was detected in ex-

tracts from protoplasts transfected with constructs containing deletion end points at -1,778, -1,391, -654, or -462 (Fig. 2). In contrast, no significant increase over background (ANOVA, $P < 0.05$) was observed in extracts from protoplasts transfected with constructs containing 5' deletion end points at -215, or -63. Using ANOVA, activity directed by a construct with a 5' deletion end point at -364 did not group with the higher activity promoters (-1,778, -1,391, -654, or -462). However, the activity was significantly higher than that directed by constructs containing 5' deletion end points at -215 or -63, by Student's *t* test ($P < 0.05$). These results allow us to conclude that sequences important for AL2 expression lie within 462 bp of the translation start site for AL2. The data suggests that one or more sequence elements important for AL2 expression lie between -462 and -364 (*NcoI* to *EagI*) and, in addition, another element between -364 and -215 (*EagI* to *Clal*) may contribute to AL2 expression. Alternatively, there is a single element directing expression of AL2 that spans the *EagI* restriction site.

Identification of viral sequences required for expression of TGMV AL3.

To define sequences required for AL3 expression, a series of 5'-truncated promoters were constructed (Fig. 1B) and analyzed in protoplasts as described for AL2. Extracts from protoplasts transfected with AL3 promoter deletions exhibited varying GUS activity, depending on the amount and context of remaining 5' flanking sequence (Fig. 3). Significantly greater activity than background (ANOVA, $P < 0.05$) was detected in extracts from protoplasts transfected with constructs containing AL3 deletion end points at -1,914, -1,527, -790, -502, or -351 (Fig. 3). In contrast, no significant increase over background (ANOVA, $P < 0.05$) was observed in extracts from protoplasts transfected with constructs containing promoter 5' end points at -296 or -199. These results allow us to conclude that sequences required for AL3 expression lie within 351 bp of the translation start site for AL3. An element (or elements) important for AL3 expression lies between -351 and -296 (*Clal* to *Apal*) and differs from the sequence element defined as important for AL2 expression (Fig. 2). A second element between -502 and -351 (*EagI*

Fig. 2. Activity of AL2 promoter-reporter constructs. The ability of promoter-reporter constructs containing different AL2 5' flanking sequences to direct transcription was determined by comparing β -glucuronidase (GUS) activity to background (pTGA26). Protoplasts (5×10^5 cells) were transfected with 10 μ g of a promoter-reporter construct and GUS activity measured in extracts isolated 3 days post transfection. Columns represent mean GUS activity in picomoles (pm) of methylumbelliferone (MU) produced per minute per milligram (mg) of protein, and error bars represent the standard error of the mean.

Fig. 3. Activity of AL3 promoter-reporter constructs. The ability of promoter-reporter constructs containing different AL3 5' flanking sequences to direct transcription was determined by comparing β -glucuronidase (GUS) activity to background (pTGA26). Protoplasts (5×10^5 cells) were transfected with 10 μ g of a promoter-reporter construct and GUS activity measured in extracts isolated 3 days post transfection. Columns represent mean GUS activity in picomoles (pm) of methylumbelliferone (MU) produced per minute per milligram (mg) of protein, and error bars represent the standard error of the mean.

to *Clal*) also contributes to AL3 expression, and is located within the same region that also appears to contribute to AL2 expression (–462 and –364) (Fig. 2). A twofold increase in AL3 expression was observed when sequences were deleted between –1,914 and –1,527 (*ScaI* and *XhoI*), which was not detected for AL2 expression (–1,778 to –1,391), but the reason for this is not known.

TGMV DNA sequences important for AL2 and AL3 expression direct transcription of AL1935 and AL1629.

The previously described experiments define viral sequences that are necessary for AL2 and AL3 expression but do not address whether these sequences direct transcription of previously identified viral mRNAs. Promoter-reporter constructs with deletion end points at the *EcoRI*, *NcoI*, and *EagI* restriction sites contain the transcription initiation site for both AL1935 and AL1629, whereas a construct with a deletion end point at the *Clal* site contains the transcription initiation site for AL1629 only (Fig. 1). Therefore, we analyzed transcripts produced from selected constructs that directed expression of AL2 and AL3. Constructs with deletion end points at *NcoI*, *EagI*, and *Clal* were cloned into pMON521. Cloned DNAs directed expression of AL2 and AL3 in protoplasts, as determined by GUS assay, in a manner identical to that observed in the pUC-based vectors (data not shown). This demonstrates that regulatory sequences for AL2 and AL3 expression are active in this context and, because the two vectors (pUC and pMON521) contain different sequence backbones, this makes it unlikely that cryptic promoter elements in the vector contribute to the expression observed. *Agrobacterium* cultures containing each construct were used to infuse *N. benthamiana* leaves, and poly (A)+ RNA was isolated as described (Sunter and Bisaro 1989). Initial experiments again indicated that 3 days postinfusion was optimal for isolation of RNA (data not shown).

We initially performed reverse-transcription polymerase chain reaction (RT-PCR) using a primer pair (AL1629 + GUS) that would amplify a fragment from RNA transcribed from initiation at both AL1935 and AL1629, to confirm the presence of TGMV-specific RNA. As can be seen (Fig. 4A), a product of the predicted size (395 bp) was detected in RNA isolated from leaves containing constructs with deletion end points at *NcoI* (lane 1), *EagI* (lane 2), and *Clal* (lane 3), but not from a mock-infused control (lane 4). The DNA template containing a deletion end point at *Clal* lacks the transcription initiation site at nucleotide 1,935 (Fig. 1B); therefore, it is assumed that the PCR product identified in this sample is derived from TGMV AL1629 mRNA. Sequences within AL1629 also are contained within AL1935 (Fig. 1A); therefore, we tested whether AL1935 was produced from DNA templates containing deletion end points at *NcoI* or *EagI*. RT-PCR was performed using primers (AL1935 + GUS) that would anneal only to RNA derived from AL1935 (Fig. 4B). A product of the predicted size (703 bp) was detected only in RNA isolated from leaves containing constructs with deletion end points at *NcoI* (Fig. 4B, lane 1) and *EagI* (Fig. 4B, lane 2). This product was not detected in RNA isolated from leaves containing a construct with a deletion end point at *Clal* (Fig. 4B, lane 3) or in a mock-infused control (Fig. 4B, lane 4). As a control, RT-PCR was performed using primers designed to amplify a 150-bp fragment of EF1 α (Yang et al. 2004) and, in all cases, the predicted fragment was detected (Fig. 4C, lanes 1 to 4). No product was detected when the reverse transcriptase was omitted, indicating the absence of contaminating DNA template (Fig. 4D, lanes 1 to 4).

Although a product derived from AL1935 can be identified, deletion end points at *NcoI* and *EagI* contained transcription initiation sites for both AL1935 and AL1629; therefore, it was

not possible to determine whether transcripts initiating at AL1629 also were present using this approach. To test whether DNA containing a deletion end point at *NcoI* can produce RNA corresponding to AL1629, we used circularization RT-PCR (cRT-PCR) (Shivaprasad et al. 2005). Total RNA isolated from *N. benthamiana* leaves infused with a construct containing a deletion end point at *NcoI* was circularized and subjected to reverse transcription using random primers (Invitrogen, Carlsbad, CA, U.S.A.). An approximately 250-bp product was amplified by PCR using primers (1629cRT + GUScRT) that anneal to cDNA derived from transcripts initiating at both nucleotides 1,935 and 1,629 (data not shown). The cDNA product was isolated and cloned into pGL2-basic (Promega Corp., Madison, WI, U.S.A.) and the sequence determined. In all cases, sequencing confirmed that the 5' end of the cDNA clones mapped to nucleotide 1629 and was ligated to the 3' end that mapped to the nopaline synthase (*nos*) terminator sequence present in the cloned DNA. Together with RT-PCR, our data indicates that a DNA template containing a deletion end point at *NcoI* is capable of directing transcription of both AL1935 and AL1629. The results suggest that elements between *NcoI* and *EagI* could direct expression of AL1935, and elements between *EagI* and *Clal* could direct expression of AL1629. We could not detect a product representing the transcript initiating

Fig. 4. Expression of *Tomato golden mosaic virus* (TGMV) AL1935 and AL1629 RNA from promoter-reporter constructs. Each panel illustrates reverse-transcriptase polymerase chain reaction (RT-PCR) products detected in RNA isolated from *Nicotiana benthamiana* leaves infused with *Agrobacterium* cultures containing promoter-reporter constructs with deletion end-points at *NcoI* (lane 1), *EagI* (lane 2), *Clal* (lane 3), or a mock-infused control (lane 4). Each reaction contained the following primer pairs: **A**, 1629 + β -glucuronidase (GUS); **B**, 1935 + GUS; and **C**, EF1 α F and EF1 α R. **D**, PCR reactions in the absence of RT with primers 1629 + GUS.

at nucleotide 1,935 using cRT-PCR, even when a second round of PCR using nested primers was performed.

Nuclear run-on transcription assays.

The absence of cRT-PCR products corresponding to AL1935 could indicate low, steady-state RNA levels in our samples, especially because cRT-PCR involves many manipulations that could result in loss of sample compared with RT-PCR (Fig. 4). Therefore, to determine whether sequences that effect AL2 and AL3 expression influence transcription of AL1935 or AL1629, nuclear run-on transcription assays were performed. *N. benthamiana* leaves were infused with *Agrobacterium* cultures containing deletion end points at *NcoI*, *EagI*, and *ClaI* and nuclei isolated 3 days postinfection. Isolated nuclei were used in run-on transcription assays and ^{32}P -labeled transcripts used to probe nitrocellulose filters containing immobilized DNA fragments specific for AL1935. It is not possible to use a DNA fragment that would specifically detect AL1629 alone, because AL1629 sequences are contained entirely within AL1935. Therefore a DNA fragment that would detect ^{32}P -labeled transcripts from both AL1935 and AL1629 was included, along with a control DNA fragment spanning the AL1 and CP transcription initiation sites that are absent in the constructs tested. Signals were normalized using hybridization to a DNA fragment specific for 18S rRNA (Sunter et al. 2001) and EF1 α (Yang et al. 2004) by phosphorimager analysis. As can be seen (Fig. 5), ^{32}P -labeled transcripts isolated from leaves infused with DNA containing a deletion end point at *NcoI* hybridized to a DNA fragment specific for AL1935, although the signal was extremely low, comparable to the signal obtained for EF1 α . This signal was further reduced to background levels when sequences were deleted between *NcoI* and *EagI* (Fig. 5). In contrast, a strong signal was detected when ^{32}P -labeled transcripts isolated from leaves infused with DNA containing deletion end points at *NcoI*, *EagI*, and *ClaI* were hybridized to a DNA fragment that would detect both AL1935 and AL1629 (Fig. 5). The signal was comparable to (*ClaI*) and, in some cases (*NcoI* and *EagI*), exceeded the signal obtained for 18S rRNA. This suggests that transcription initiating at nucleotide 1,935 is extremely low in our constructs, and that AL1629 is the predominant RNA expressed in our promoter-reporter constructs. When sequences were deleted between *NcoI* and *EagI*, hybridization to the DNA fragment that would detect both AL1935 and AL1629 decreased approxi-

Fig. 5. Nuclear run-on transcription assays. Each panel represents phosphorimager analysis of DNA slot blots in which ^{32}P -labeled run-on transcripts were hybridized to 2 μg each of polymerase chain reaction (PCR) products specific for AL1935 (107 bp), AL1935/AL1629 (241 bp), AL1/CP (447 bp), EF1 α (150 bp), and 18S rRNA (315 bp). The ^{32}P -labeled transcripts were obtained from nuclear run-on transcription assays using nuclei isolated from *Nicotiana benthamiana* leaves infused with *Agrobacterium* cultures containing promoter-reporter constructs with deletion end-points at *NcoI*, *EagI*, and *ClaI*. Equivalent amounts of nuclei were used for all nuclear run-on transcription assays. Each slot blot was exposed to a phosphor screen for 5 days, except an 8-day exposure also is shown of the slot blot for ^{32}P -labeled run-on transcripts using nuclei containing the promoter-reporter construct with a deletion end-point at *NcoI*. The data are representative of two independent experiments.

mately 1.5-fold, and deletion of sequences between *EagI* and *ClaI* reduced transcription a further twofold, as determined by phosphorimager analysis. This is in agreement with our promoter-reporter gene assays where deletion of sequences between *NcoI* and *EagI* decreased AL2 expression fourfold, and deletion of sequences between *EagI* and *ClaI* resulted in loss of AL2 expression (Fig. 2) and a threefold reduction in AL3 expression (Fig. 3). Differences in the levels of reduction in expression observed for nuclear run-on transcription and promoter-reporter assays probably reflect differences in the measurement of transcription, or GUS enzyme activity that also could involve translation effects.

Identification of minimal sequences required for activation of a heterologous promoter core.

To further delineate elements required for expression of AL1935 and AL1629, putative promoter fragments were placed upstream of the heterologous *Cauliflower mosaic virus* (CaMV) 35S promoter core (−46 to +8) (Benfey and Chua 1990). This allowed us to test the ability of viral sequences to activate expression of a core promoter. Chimeric promoters (AL1935:35S core and AL1629:35S core) were linked as transcriptional fusions to the *GUS* reporter gene and used to transfect *N. benthamiana* protoplasts.

In contrast to data obtained from the deletion analysis, no significant GUS activity (Student's *t* test, $P < 0.05$) above basal levels from the 35S core promoter alone was detected in extracts from protoplasts transfected with an AL1935:35S chimeric promoter containing sequences from *NcoI* to *EagI*, shown to be important for AL2 expression (Fig. 6). These results suggest that additional sequences between *EagI* and *ClaI* are required for expression, based on data demonstrating that

Fig. 6. Activation of AL1935:35S and AL1629:35S core promoters. β -glucuronidase (GUS) activity from extracts isolated from protoplasts (5×10^5 cells) transfected with DNA (10 μg) containing *Tomato golden mosaic virus* (TGMV) sequences fused to the *Cauliflower mosaic virus* (CaMV) 35S promoter core is shown. The values given are represented as the ratio of activated activity relative to basal activity detected with the core promoter alone, which was arbitrarily assigned a value of 1. AL1935:35S core promoters tested included the *NcoI* to *EagI* and *NcoI* to *ClaI* fragments (three experiments). AL1629:35S core promoters tested included the *ClaI* to *ApaI*, *ClaI* to *NheI*, and *EagI* to *ApaI* fragments (three experiments). A construct containing sequences between *EagI* and *ClaI* common to both the AL1935 and AL1629 promoters also was tested (three experiments). Error bars represent the standard error of the mean.

deletion of these sequences results in loss of AL2 expression (Fig. 2). To test this possibility, protoplasts were transfected with an AL1935:35S chimeric promoter containing TGMV DNA sequences from *Nco*I to *Cla*I (Fig. 6). Significant GUS activity (Student's *t* test, $P < 0.05$), up to five- to sixfold greater than basal levels from the 35S core promoter, was detected (Fig. 6). The most direct interpretation of these data is that two elements are required for activation of the 35S core promoter, one between *Nco*I and *Eag*I and a second between *Eag*I and *Cla*I. Alternatively, sequences required for activation of the 35S core promoter could span the *Eag*I site. To determine whether the second sequence element was capable of independently activating expression of the 35S core promoter, protoplasts were transfected with a chimeric promoter containing sequences from *Eag*I to *Cla*I (Fig. 6). However, no significant GUS activity (Student's *t* test, $P < 0.05$) above basal levels was detected, indicating that this element is incapable of activating expression independently of sequences located between *Nco*I and *Eag*I. This is consistent with deletion data (Fig. 2) that demonstrates decreases in expression upon deletion of either of these sequences. The results identify a 247-bp fragment of viral DNA (*Nco*I to *Cla*I) capable of activating expression of the 35S core promoter in *N. benthamiana* protoplasts.

To further delineate the element or elements required for expression of AL1629, chimeric promoters (AL1629:35S core) were used to transfect *N. benthamiana* protoplasts. No significant GUS activity (Student's *t* test, $P < 0.05$) above basal levels was detected in extracts prepared from protoplasts transfected with an AL1629:35S chimeric promoter containing viral sequences from *Cla*I to *Apa*I (Fig. 6). These results suggest that additional sequences, either upstream or downstream, are required for AL3 expression. To test this possibility, protoplasts were transfected with additional chimeric AL1629:35S promoter constructs. No significant GUS activity (Student's *t* test, $P < 0.05$) was detected in extracts from protoplasts transfected with an AL1629:35S chimeric promoter containing sequences from *Cla*I to *Nhe*I (Fig. 6). However, a significant increase in GUS activity (Student's *t* test, $P < 0.05$), up to eight-

fold greater than basal levels, was detected in extracts from protoplasts transfected with the AL1629:35S chimeric promoter containing sequences from *Eag*I to *Apa*I (Fig. 6). Again, the most direct interpretation of these data is that two elements are required for activation of the 35S core promoter, one located between *Eag*I and *Cla*I and a second located between *Cla*I and *Apa*I, although a single sequence element could span the *Cla*I site. This is consistent with the deletion data that demonstrates that AL3 expression decreases upon deletion of either of these sequences. The results also identify a second, 206-bp fragment of viral DNA (*Eag*I to *Apa*I) capable of activating the 35S core promoter in *N. benthamiana* protoplasts.

Nuclear proteins bind to sequences that can confer expression from a heterologous promoter core.

To investigate whether TGMV DNA sequences shown to be required for activation of the 35S core promoter were capable of interacting with nuclear proteins, electrophoretic mobility shift analysis (EMSA) was performed, using the 98-bp *Nco*I to *Eag*I and 55-bp *Cla*I to *Apa*I DNA fragments (Fig. 1A). Experiments utilized a soluble protein fraction extracted from purified nuclei isolated from *N. benthamiana* suspension cells. As illustrated, nuclear proteins bound both the 98-bp *Nco*I to *Eag*I (Fig. 7A) and 55-bp *Cla*I to *Apa*I (Fig. 7B) DNA fragments, as evidenced by a shift in the labeled probe fragment. In the case of the 98-bp *Nco*I to *Eag*I DNA fragment, two products were observed which could reflect binding of different amounts of a single protein or binding of multiple proteins. Binding of nuclear proteins to both DNA fragments is dependent on protein concentration (data not shown). The specificity of binding was shown by competition using unlabeled competitor DNAs. The binding detected with the 98-bp *Nco*I to *Eag*I DNA fragment was efficiently competed with increasing amounts (50 to 200 molar excess) of cold *Nco*I to *Eag*I DNA (Fig. 7A, lanes 3 to 5), but was not competed using a nonspecific 78-bp DNA fragment from pUC even at a 200 molar excess (Fig. 7A, lanes 6 and 7). A similar observation was made for binding to the 55-bp *Cla*I to *Apa*I DNA fragment. Binding

Fig. 7. Electrophoretic mobility shift assays. DNA binding studies were carried out using double-stranded DNA fragments corresponding to the **A**, 98-bp *Nco*I-*Eag*I and **B**, 55-bp *Cla*I-*Apa*I sequences shown to be important for *Tomato golden mosaic virus* (TGMV) AL2 and AL3 gene expression, respectively. Soluble nuclear extracts isolated from tobacco suspension cells were incubated with labeled probe DNA in the absence or presence of cold competitor DNAs, and separated on 6% polyacrylamide retardation gels. The position of free, unbound probe (Free) and bound complexes (asterisks) are indicated. The competitor DNA and molar excess of each that was used is indicated below each panel.

was efficiently competed with increasing amounts (50 to 200 molar excess) of cold 55-bp *Cla*I to *Apa*I DNA fragment (Fig. 7B, lanes 3 to 5) but, again, was not competed using a nonspecific 78-bp DNA fragment from pUC even at a 200 molar excess (Fig. 7B, lanes 6 and 7). It was concluded from these experiments that DNA sequences shown to be necessary for activation of a 35S core promoter, and to be important for TGMV AL2 and AL3 gene expression, specifically bind nuclear proteins, which may reflect a host transcription factor. Although the larger fragments from *Nco*I to *Cla*I and *Eag*I to *Apa*I bind nuclear factors, the *Eag*I to *Cla*I fragment shown to contain sequences important for activation of the 35S core promoter (Fig. 6) did not demonstrate any ability to bind nuclear factors under these conditions (data not shown).

TGMV AL1629 mRNA expresses both AL2 and AL3, but AL1935 only expresses AL3.

Our analysis has identified viral sequences that contribute to expression of two previously identified viral transcripts (AL1935 and AL1629). However, a deletion end point at –215 (*Cla*I) can direct transcription of AL1629 but only appears to express AL3 (Figs. 2 and 3). To test whether AL1629 also can express AL2, we generated additional promoter-reporter constructs downstream of the transcription initiation site for AL1935 that only could direct transcription of AL1629. As seen (Fig. 8), both AL2 and AL3 expression was detected in extracts from protoplasts transfected with constructs containing a deletion end point at –380 or –244 relative to the AL3 and AL2 translation initiation sites, respectively. Therefore, our interpretation is that both AL2 and AL3 can be expressed from the AL1629 transcript.

Expression of AL3 directed by the deletion end point at –380 was not different from that at –502 (*Eag*I), but decreased approximately sevenfold upon deletion of additional sequences to the *Cla*I site (–351) (Fig. 8). A similar result was observed for AL2 expression, where a fivefold decrease in expression was detected when sequences between –244 and –215 (*Cla*I) were deleted (Fig. 8). Comparison of GUS activity directed by DNA templates with the same deletion end point indicates that

Fig. 8. Comparison of AL2 and AL3 expression from promoter-reporter constructs with comparable 5' flanking sequences. *Nicotiana benthamiana* protoplasts (5×10^5 cells) were transfected with 10 μ g of each construct and β -glucuronidase (GUS) activity measured in extracts isolated 3 days post transfection. Columns represent percent relative GUS activity compared with AL3[-502], which was arbitrarily assigned a value of 100%. Error bars represent the standard error of the mean.

AL3/GUS expression is approximately two- to fivefold higher than AL2/GUS (Fig. 8). This suggests differences in translation efficiencies for AL2 and AL3, although this was not directly measured in these experiments.

Cloned DNAs with deletion end-points at *Nco*I and *Eag*I contain transcripts initiating at nucleotides 1,935 and 1,629; therefore, it is not possible to determine whether AL1935 can express both AL2 and AL3 using promoter-reporter constructs. Therefore, we generated transgenic *N. benthamiana* plants capable of expressing the TGMV AL1935 transcript from the CaMV 35S promoter. The presence of the transgene was confirmed by PCR and expression of RNA corresponding to TGMV AL1935 was detected by Northern blot and RT-PCR (data not shown). To examine the ability of these transgenic plants to express AL2, AL3, or both, we used a complementation assay. Transgenic *N. benthamiana* plants were challenged by agroinoculation with wild-type TGMV or TGMV containing mutations in the AL2 or AL3 coding regions. Previous results have demonstrated that mutations in AL2 lead to loss of infectivity and mutations in AL3 result in a delay and attenuation in symptom formation (Elmer et al. 1988). When inoculated with wild-type virus, AL1935 transgenic *N. benthamiana* plants exhibited typical TGMV symptoms indistinguishable from those on nontransgenic *N. benthamiana* plants (Table 1), indicating that the transgene does not influence infectivity with respect to symptom severity or appearance of symptoms. Transgenic *N. benthamiana* plants expressing AL1935 inoculated with TGMV *al2* mutant virus were asymptomatic and no viral DNA was detected (data not shown), which was identical to the lack of symptoms observed for nontransgenic *N. benthamiana* plants (Table 1). This suggests that AL2 is not expressed from the AL1935 transgene, although it is possible, of course, that AL2 is expressed at levels that are not capable of complementing the mutation; however, we believe this to be unlikely. In contrast, the mutation in the AL3 coding region was efficiently complemented in AL1935 transgenic *N. benthamiana*. Nontransgenic plants inoculated with TGMV *al3* mutant virus exhibited extremely mild symptoms 11 days postinoculation (dpi) in 1 of 11 plants, as previously observed (Elmer et al. 1988). In contrast, TGMV symptoms were detected 6 dpi in 22 of 35 transgenic AL1935 *N. benthamiana* plants inoculated with TGMV *al3* mutant virus. Symptom severity and appearance of symptoms on transgenic AL1935 plants were comparable to that observed on nontransgenic plants inoculated with wild-type TGMV (Table 1). Complementation is not due to recombination of the virus with the transgene, or reversion, because viral DNA isolated from AL1935 transgenic plants retained an *Xba*I restriction site (data not shown) that is a result of the introduced mutation (Elmer et al. 1988). Complementation of a deficiency in the viral AL3 coding

Table 1. Complementation of Tomato golden mosaic virus (TGMV) *al2* and *al3* mutant viruses in transgenic *N. benthamiana* plants expressing TGMV AL1935 RNA

Inoculum ^a	<i>N. benthamiana</i> line ^b	
	Nontransgenic	AL1935 transgenic
Wild-type TGMV	9/12 (day 6)	7/13 (day 6)
<i>al2</i> mutant TGMV	0/12	0/36
<i>al3</i> mutant TGMV	1/12 (day 11)	24/35 (day 6)

^a Inoculum consisted of a mixture of *Agrobacterium tumefaciens* cells ($OD_{600} = 1.0$) containing binary Ti plasmids harboring 1.5 copies of wild-type or mutant TGMV A, and TGMV B.

^b Values represent the number of plants exhibiting symptoms typical of a TGMV infection compared with the total number of plants inoculated. The numbers in parentheses indicate the first day at which symptoms were observed.

region suggests that AL3 is expressed from an RNA corresponding to TGMV AL1935 in transgenic *N. benthamiana* plants. However, lack of complementation of a deficiency in the viral AL2 coding region indicates that AL2 is not expressed from AL1935.

DISCUSSION

Experiments described here using 5'-truncated promoters have revealed some notable features concerning TGMV AL2 and AL3 gene expression. Analysis of steady-state RNA levels and nuclear run-on transcription assays have determined that constructs with deletion end points at the *Nco*I and *Eag*I sites of TGMV DNA A can direct transcription of both AL1935 and AL1629. When sequences containing the transcription initiation site for AL1935 are deleted, constructs with end points at, or near, the *Cla*I restriction site direct transcription of AL1629. Sequences between *Nco*I and *Eag*I are 28 bp upstream of the transcription initiation site at nucleotide 1,935, but there is no consensus RNA polymerase II TATAA box present (Fig. 9). This could result in poor initiation of transcription, which may explain the extremely low levels of transcription detected (Fig. 5). There is a TATAA sequence 30 bp upstream of the initiation site at nucleotide 1,629 (Fig. 9) which could reflect the fact that transcription of AL1629 appears to be much higher than AL1935.

Sequences important for AL2 and AL3 expression lie between *Nco*I and *Apa*I, 160 to 462 bp upstream of the translation start site for AL2, or 296 to 502 bp upstream of the translation start site for AL3 (Fig. 9). One sequence element between *Nco*I and *Eag*I (A) appears to be important for AL1935 transcription, a second element between *Cla*I and *Apa*I (B) appears to be important for AL1629 transcription, and one element between *Eag*I and *Cla*I (A/B) appears to be contributing to the expression of both transcripts (Fig. 9). All sequence elements lie upstream of the transcription initiation site for AL1629, but only sequences between *Nco*I and *Eag*I are upstream of AL1935. However, sequences from *Nco*I to *Eag*I are insufficient to activate expression from a heterologous CaMV 35S promoter core and require sequences from *Eag*I to *Cla*I (Fig.

6). One explanation is that two elements contribute to activation, one between *Nco*I and *Eag*I and one between *Eag*I and *Cla*I. A second possibility, which we favor, is that a single sequence element that spans the *Eag*I restriction site is responsible for expression of AL1935. This conclusion is supported by data indicating that DNA sequences between *Nco*I and *Eag*I are capable of specifically binding plant nuclear proteins, whereas the DNA fragment between *Eag*I and *Cla*I does not appear to bind nuclear factors under the conditions used here. One reason for this could be that the majority of the binding sequence is contained within the *Nco*I to *Eag*I DNA fragment and extends into the *Eag*I to *Cla*I DNA fragment. However, the overlap is too small to allow stable binding of proteins to DNA sequences within the *Eag*I to *Cla*I DNA fragment independently. This might suggest that sequences between *Nco*I and *Eag*I represent the promoter for AL1935 transcription.

Sequences between *Eag*I and *Apa*I also can activate expression from a heterologous CaMV 35S promoter core. As discussed above, this sequence could comprise two elements, one between *Eag*I and *Cla*I and a second between *Cla*I and *Apa*I. However, because the sequence between *Cla*I and *Apa*I is capable of binding plant nuclear extracts in a specific manner, we favor a single sequence element important for expression of AL1629 that spans the *Cla*I restriction site, for the reasons outlined above. In this case, the data suggests that sequences surrounding the *Cla*I site represent the promoter for AL1629 transcription.

It has been shown that geminivirus DNA is associated with host nucleosomes as a viral minichromosome (Pilartz and Jeske 1992, 2003). The region of viral DNA shown here to be important for expression of the TGMV AL1935 and AL1629 viral RNAs is in close agreement to a nucleosome-free region of AbMV located upstream of the translation initiation codons for the AbMV AL2 and AL3 genes, respectively (Fig. 9). Therefore, it is likely that geminivirus minichromosomes have the potential to open a nucleosome-free gap upstream of the AL2 and AL3 genes, thereby becoming accessible for transcription factors. This region also is in a similar position relative to sequences upstream of the AC2 translational initiation site shown to be important for expression of AC2 and AC3 in

Fig. 9. Promoter structure for *Tomato golden mosaic virus* (TGMV) complementary sense transcription units. A linear map of the wild-type TGMV DNA A genome component is illustrated. Relevant restriction sites are indicated with nucleotide coordinates given in parentheses. Open boxes indicate coding regions and arrows designate transcripts and direction of transcription, with initiation site indicated. The hatched box represents the approximately 230-bp intergenic region, solid triangles represent polyadenylation signals, and open triangles indicate TATA boxes. Sequences capable of activating a 35S core promoter are indicated with boxes below the linear map. The hatched box (A) represents sequences upstream of the AL1935 transcription initiation site that are important for AL1935 expression. Sequences important for AL1629 expression are indicated with a stippled box (B), and sequences shared by both promoters (A/B) are shown with a crosshatched box. The approximate position of nucleosome-free regions detected in *Abutilon mosaic virus* (AbMV)-infected plants is indicated. Asterisks indicate the position of AUG codons, upstream of the AL2 coding region, that are contained within the AL1935 transcript. Diagrams are not drawn to scale.

MYMV (Shivaprasad et al. 2005), and to putative promoter sequences in ACMV (Zhan et al. 1991). However, there are some differences in complementary sense transcription among TGMV, MYMV, and ACMV. Two virus-specific mRNAs, AL1935 and AL1629, initiate within the AL1 coding region of TGMV (Sunter and Bisaro 1989). However, a single mRNA initiates within the AC1 coding region of ACMV (Townsend et al. 1985), TLCV (Mullineaux et al. 1993), and MYMV (Shivaprasad et al. 2005). This could represent differences in transcription strategies between New World (TGMV) and Old World (MYMV) geminiviruses, although the limited amount of information on transcription in geminiviruses as a whole makes this difficult to determine definitively.

Previous results have indicated that geminivirus proteins can be expressed from polycistronic viral mRNAs. Transgenic plants expressing an RNA corresponding to TGMV AL62 can complement viruses containing mutations in the AL1 and AL3 coding regions but not AL2 (Hanley-Bowdoin et al. 1989). These results indicate that a polycistronic mRNA corresponding to TGMV AL62 can express the AL1 and AL3 proteins but not the AL2 protein. In similar experiments, *N. benthamiana* plants expressing RNA equivalent to TGMV AL1935 are capable of complementing virus with a mutation in the AL3 coding region but not AL2. This is consistent with a conclusion that AL1935 expresses only AL3. An analysis of the arrangement of ORFs contained within AL1935 shows that there are three AUG codons upstream of the translation initiation site for AL2 (Fig. 9). In the scanning model for translation initiation, the ribosome enters at the 5' end of the mRNA and scans until it encounters the first AUG (Kozak 2005). Translation potentially could be initiated at one or more of the AUG codons within the leader sequence of AL1935 which could be used to deliberately inhibit translation of AL2, as has been observed for thrombopoietin (Kozak 2005). This raises the question of how AL3 is expressed from AL1935. Based on the arrangement of ORFs within AL1935 (Fig. 9), one possibility is reinitiation of translation, in which the ribosome remains bound to the mRNA after translation of the upstream ORF and resumes scanning (Kozak 2005) where it then would encounter the translation initiation codon for AL3. These aspects of AL2 and AL3 expression currently are under investigation.

In contrast to results for AL1935, promoter-reporter experiments suggest that both AL2 and AL3 proteins can be expressed from AL1629, although expression of AL2 in this context is two- to fivefold lower than that of AL3. This could be a consequence of different translation efficiencies for AL2 and AL3, which also is under investigation. For AL1629, the first AUG is the translation initiation codon for AL2; therefore, translation of AL3 always must initiate at an internal start codon. One possible mechanism to allow translation of one or both proteins from a single transcript could be reflected in differences in sequences surrounding the translation initiation site for the two genes, which exhibit different degrees of similarity to the consensus sequence for translation initiation (Joshi et al. 1997; Kozak 1984). The sequence surrounding the AL2 translation initiation site has a poor match to the consensus, whereas the sequence surrounding the AL3 translation initiation site shows strong homology to the consensus sequence for translation initiation. Therefore, it is possible that leaky ribosome scanning could result in expression of both AL2 and AL3 from bicistronic AL1629 mRNA, as has been suggested for a similar transcript identified in MYMV (Shivaprasad et al. 2005).

The use of different promoter sequences for AL1935 and AL1629 affords the virus a way to regulate expression of two important viral genes that influence the cellular environment. The AL3 gene product is important for replication of the virus and for conditioning the cell for replication through interaction

with cell cycle components (Settlage et al. 2001; Sunter et al. 1990). The AL2 gene product is involved in inactivating defense responses (Hao et al. 2003; Wang et al. 2003) and for activation of the coat protein and BR1 nuclear shuttle protein genes (Sunter and Bisaro 1991). Independent regulation of these two genes would appear to be crucial to ensure successful completion of the viral life cycle, but it currently is unclear why geminiviruses would express AL3 from multiple transcripts. One possibility is that large amounts of AL3 are required for its multiple functions, and it is more critical to regulate the level of AL2 protein. Alternatively, AL3 could be required at several stages of the viral life-cycle requiring expression from different RNAs transcribed at different times. Experiments are in progress to further define the promoter sequences for AL1935 and AL1629, to identify nuclear factors that bind these sequences, and to determine the role of sequences surrounding the translation initiation codon in production of these two important viral proteins.

MATERIALS AND METHODS

DNA techniques.

The map locations and restriction endonuclease sites cited here refer to the published DNA sequence of TGMV (Hamilton et al. 1984). All restriction endonucleases and DNA modifying enzymes were used as recommended by the manufacturers. DNA and RNA manipulations, PCR, and electrophoretic mobility shift analyses were performed essentially as described by Ausubel and associates (2001) unless otherwise stated.

Promoter-reporter constructs.

A construct capable of generating a replicating TGMV genome component (pTGA26) has been described previously (Sunter et al. 1990). Briefly, the replicating construct contains a partial tandem dimer of the viral DNA A genome component within a bacterial plasmid vector. In plant cells, a monomeric viral genome is released from the plasmid, primarily by a replicative mechanism (Stenger et al. 1991).

A series of AL2 promoter deletions, beginning from -1,778, was constructed. The 872-bp *HindIII*-*Bam*HI fragment of pBI221, which contains the CaMV 35S promoter, was replaced with the 2,031-bp *HindIII*-*Bam*HI fragment from pTGA26 to generate AL2[-1778]-GUS. This construct contains a translational fusion between the *GUS* reporter gene and the N-terminal 83 amino acids of the AL2 coding region, with the TGMV viral sequence from -1,778 to +1 (AL2 translation start site), and the *nos* 3' end. Deletions corresponding to -1,391 and -654 were generated by replacing the 2,031-bp *HindIII*-*Bam*HI fragment of AL2[-1778]-GUS with the 1,638-bp *XhoI*-*Bam*HI (AL2[-1391]-GUS) and 895-bp *EcoRI*-*Bam*HI (AL2[-654]-GUS) fragments of pTGA26 respectively. Deletion end points corresponding to -462, -364, -215, and -63 were generated by deleting the 1,329-bp *HindIII*-*NcoI* (AL2[-462]-GUS), 1,425-bp *HindIII*-*EagI* (AL2[-364]-GUS), 1,575-bp *HindIII*-*Clal* (AL2[-215]-GUS), and 1,728-bp *HindIII*-*NheI* (AL2[-63]-GUS) fragments of AL2[-1778]-GUS. The promoter-reporter construct AL2[-244]-GUS containing a deletion end point at -244 was generated by PCR. A primer (5'-ACTCAAACAAGG TTG-3') that anneals to nucleotides 1,830 to 1,844 of the TGMV genome was used with a primer that anneals to the *GUS* coding sequence (5'-CCCACCAACGCTGATCAATTCC AC-3') to amplify a 612-bp fragment using AL2[-1778]-GUS as template. After restriction with *HindIII*-*Bam*HI, the resulting 483-bp fragment was used to replace the 2,032-bp *HindIII*-*Bam*HI fragment of AL2[-1778]-GUS.

A series of AL3 promoter deletions, beginning from -1,914, was constructed. Cloned AL2[-1778]-GUS, AL2[-1391]-GUS,

AL2[-654]-GUS, and AL2[-244]-GUS DNAs were cleaved with *Bam*HI, end filled with Klenow, and religated, resulting in a +1 frame shift, to generate AL3[-1914]-GUS, AL3[-1527]-GUS, AL3[-790]-GUS, and AL3[-380]-GUS, respectively. These constructs contain a translational fusion between the *GUS* reporter gene and the N-terminal 36 amino acids of the AL3 coding region with the TGMV viral sequence from -1,914 to +1 (AL3 translation start site) and the *nos* 3' end. Deletions corresponding to -502, -351, -296, or -199 were generated by deleting the 1,425-bp *Hind*III-*Eag*I (AL3[-502]-GUS), 1,575-bp *Hind*III-*Cla*I (AL3[-351]-GUS), 1,629-bp *Hind*III-*Apa*I (AL3[-296]-GUS), and 1,728-bp *Hind*III-*Nhe*I (AL3[-199]-GUS) fragments of AL3[-1914]-GUS.

The CaMV 35S core promoter fragment from -46 to +8 (Benfey and Chua 1990) cloned upstream of the *GUS* reporter gene has been described previously (Sunter and Bisaro 2003). The following fragments of TGMV DNA A were fused upstream of the core promoter sequence at a *Pml*I site. AL1935:35S core promoters were generated using the 247-bp *Nco*I-*Cla*I and 98-bp *Nco*I-*Eag*I fragments of TGMV DNA A. AL1629:35S core promoters were generated using the 152-bp *Cla*I-*Nhe*I, 56-bp *Cla*I-*Apa*I, and 206-bp *Eag*I-*Apa*I fragments of TGMV DNA A. The 149-bp *Eag*I-*Cla*I DNA fragment also was cloned upstream of the CaMV 35S core promoter.

Protoplast transfection and analysis.

Protoplasts were isolated from a *N. benthamiana* suspension culture cell line and transfected with various DNAs as described (Sunter and Bisaro 2003). After incubation in the dark for 3 days, protoplasts were harvested and fluorometric GUS assays performed as described (Sunter and Bisaro 1991), using equivalent amounts of protein as determined by the method of Bradford (1976). GUS activities were compared using ANOVA followed by Student-Neuman-Keuls test, or by Student's *t* test.

Leaf infusion assays.

Cloned DNA containing the promoter-reporter constructs with deletion end points at *Nco*I (AL2[-462]-GUS), *Eag*I (AL2[-364]-GUS), and *Cla*I (AL2[-215]-GUS), were restricted with *Hind*III-*Eco*RI, and the resulting 2,846-, 2,750-, and 2,600-bp fragments cloned into the binary plasmid vector pMON521 (Rogers et al. 1987). Binary plasmid constructs then were mobilized into *Agrobacterium* strain GV3111SE, containing the disarmed Ti plasmid pTiB36SE, by triparental mating (Horsch and Klee 1986). *Agrobacterium* cultures containing promoter-reporter constructs were delivered to *N. benthamiana* leaves by leaf infusion as described (Johansen and Carrington 2001; Wang et al. 2005).

Construction of transgenic plants.

A binary plasmid construct containing DNA capable of constitutively expressing the TGMV AL1935 transcript from the CaMV 35S promoter was generated. An 884-bp PCR product was amplified with primers specific for the AL1935 transcription initiation site (5'-GCGCCATGGCGATAGTCGGACGGGAAAGAC-3') and the 3' end of the TGMV AL3 coding region (5'-GCGAGATCTATGATTCGATAAC-3') using pTGA26 as template. The amplified DNA product was restricted with *Nco*I and *Bgl*II (underlined in the primer sequences) inserted between the 35S promoter and *nos* 3' end of the binary plasmid vector pMON530 (Rogers et al. 1987). The presence of TGMV DNA sequences corresponding to AL1935 was confirmed by sequencing. Binary plasmid constructs were mobilized into *Agrobacterium* spp. as described above and cultures used to transform *N. benthamiana* leaf discs (Horsch and Klee 1986). Transformed plants were selected as described previously (Sunter et al. 2001).

RNA isolation and analysis.

Total RNA was isolated from *N. benthamiana* leaf tissue using Plant RNA Purification Reagent according to the manufacturer's instructions (Invitrogen) and poly (A)+ RNA purified as described previously (Sunter and Bisaro 1989). Up to 500 ng of poly (A)+ RNA was used for detection of TGMV viral RNAs in RT-PCR reactions. Primers used for amplification were AL1629 (5'-TCCACTAAAGAACTGGACTTTCCATAA TGCG-3'), GUS (5'-CCCACCAACGCTGATCAATTCCAC-3'), AL1935 (5'-GGCGATAGTCGGACGGGAAAGACTAT GTGGGC-3'), EF1 α F (5'-TGGTGTCTCAAGCCTGGTAT GGTTGT-3'), and EF1 α R (5'-ACGCTTGAGATCCTTAAC CGCAACATTCTT-3'). cRT-PCR was performed as described (Shivaprasad et al. 2005), using RNA isolated from *N. benthamiana* leaf tissue. Following circularization, reverse transcription was performed using random primers. Samples of cDNA were used for PCR amplification using primers 1629cRT (5'-CGTCGCCCTTCTAATAGCTC-3') and GUScRT (5'-CGCTG GACTGGCATGAACCTTCG-3'), which are specific for the AL1629/GUS and AL1935/GUS transcription units. PCR products were purified by gel extraction and introduced into a TA cloning vector (Promega Corp.). Sequencing of individual clones (Plant Microbe Genomics Facility, The Ohio State University, Columbus, OH, U.S.A.) was carried out using the pUC forward primer.

Nuclear run-on transcription assays.

Nuclei were prepared from *N. benthamiana* leaves or protoplasts essentially as described (Sunter and Bisaro 1992) 3 days postinfection or posttransfection. Isolated nuclei were resuspended in nuclear resuspension buffer and total nuclear DNA was isolated from a small aliquot (10 μ l) and quantified by spectrophotometer. The remaining nuclei were centrifuged and resuspended in transcription buffer (10 mM Tris-HCl, pH 8.5, 5 mM MgCl₂, 300 mM KCl, 1 mM UTP, 1 mM GTP, 1 mM ATP) and 100 μ Ci α -³²P-CTP (800 Ci/mmol). Following incubation at 30°C for 50 min, reactions were terminated and ³²P-labeled transcripts isolated using plant RNA purification reagent according to the manufacturer's instructions (Invitrogen). PCR was used, with TGMV DNA A (pTGA26) as template, to generate TGMV-specific DNA probe fragments. Primers 1629 (5'-ACTCCACTAAAGAACTGGAC-3') and TGBam (5'-GCTGATTATCTTGAAATAGAGG-3') were used to amplify a 241-bp fragment (AL1935/AL1629). Primers 1935 (5'-CGATAGTCG GACGGGAAAGAC-3') and AL3comp (5'-CATCATCGATG ACGTTATAC-3') were used to amplify a 107-bp fragment. Primers AL1pxt (5'-GGCATTTATTTGAAACCG-3') and CPpxt (5'-ATCACGCTTAGGCAAACCTTC-3') were used to amplify a 447-bp fragment. RT-PCR was performed on total RNA isolated from healthy *N. benthamiana* leaf tissue to generate a 150-bp DNA fragment specific for EF1 α (EF1 α F and EF1 α R primers), and a 315-bp DNA fragment specific for 18S rRNA using Universal 18S primers (QuantumRNA 18S Internal Standards; Ambion, Austin, TX, U.S.A.). PCR fragments were denatured and immobilized onto nitrocellulose using a slot blot apparatus (BioRad, Hercules, CA, U.S.A.), hybridized with ³²P-labeled transcripts, and membranes washed as described (Sunter and Bisaro 1992). Hybridization was assessed by phosphorimager analysis (BioRad).

Purification of soluble nuclear protein extracts.

Protoplasts were isolated from 250 ml of tobacco suspension cells as described (Sunter and Bisaro 2003). Nuclei were isolated using the Plant Nuclei Isolation/Extraction Kit (Sigma-Aldrich, St. Louis) essentially according to the protocol. All procedures were carried out at 4°C unless otherwise stated. Briefly, protoplasts (40-ml packed-cell volume) were resus-

pended in 20 ml of nuclei isolation buffer (NIB) containing protease inhibitor cocktail (Sigma-Aldrich). Nuclei were released from the cells using a Dounce homogenizer and integrity of intact nuclei confirmed by microscopy using methyl green stain. Nuclei were collected by centrifugation at $2,500 \times g$ for 15 min, washed with NIB containing protease inhibitor cocktail, and finally resuspended in 2 volumes of extraction buffer containing 5 mM dithiothreitol (DTT). Nuclei were mixed gently, ammonium acetate slowly added to a final concentration of 0.4 M, and extracted for 90 min with gentle stirring. Extracts were centrifuged at $100,000 \times g$ for 1 h, and the supernatant containing soluble nuclear proteins (S100) dialyzed against 2×500 ml of dialysis buffer (20 mM Hepes-KOH, pH 7.9, 100 mM KCl, 12.5 mM $MgCl_2$, 0.2 mM EDTA, 20% glycerol, 1 mM phenylmethylsulfonyl fluoride, and 2 mM DTT), for 3 h. Soluble nuclear extracts were concentrated using Amicon Ultra protein concentrators (Millipore, Bedford, MA, U.S.A.) as per the manufacturer's instructions, and protein concentration determined by Bradford Assay (Bradford 1976).

Electrophoretic mobility shift analysis.

Binding of plant nuclear proteins to promoter fragments was examined by EMSA. Various DNA fragments were end labeled with $\gamma^{32}P$ -ATP and T4 polynucleotide kinase as described (Hartitz et al. 1999). Radiolabeled probe DNA and unlabeled competitor DNA were incubated with 15 μg of soluble nuclear protein extracts. Each binding reaction contained 375 ng of poly (dI:dC), 50 fmol of probe, and 2.5 to 10 pmol of competitor DNA in 15 μl of binding buffer (50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 2.5 mM EDTA, 2.5 mM DTT, 250 mM NaCl, and 20% glycerol). Reactions were incubated at room temperature for 30 min and complexes resolved by electrophoresis in a 6% polyacrylamide DNA retardation gel (Invitrogen). The gel was dried by vacuum onto Whatman 3M paper prior to exposure using a phosphor storage screen and analyzed using a phosphorimager (BioRad).

ACKNOWLEDGMENTS

The research was supported in part by a National Institutes of Health MBRS/SCORE grant (GM-08194). We thank G. Lacatus for generating the *N. benthamiana* transgenic lines and S. Baliji for critical reading of the manuscript.

LITERATURE CITED

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Siedman, J. G., Smith, J. A., and Struhl, K. 2001. Current Protocols in Molecular Biology. John Wiley and Sons, New York.
- Benfey, P. N., and Chua, N.-H. 1990. The *Cauliflower mosaic virus* 35S promoter: Combinatorial regulation of transcription in plants. *Science* 250:959-966.
- Bisaro, D. M. 1996. Geminivirus replication. Pages 833-854 in: *DNA Replication in Eukaryotic Cells*. M. DePamphilis, ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* 72:248-254.
- Briddon, R. W., and Markham, P. G. 1995. Geminiviridae. Pages 158-165 in: *Sixth Report of the International Committee on Taxonomy of Viruses*. C. M. Fauquet, ed. Springer-Verlag, New York.
- Brough, C. L., Hayes, R. J., Morgan, A. J., Coutts, R. H. A., and Buck, K. W. 1988. Effects of mutagenesis *in vitro* on the ability of cloned *Tomato golden mosaic virus* DNA to infect *Nicotiana benthamiana* plants. *J. Gen. Virol.* 69:503-514.
- Eagle, P. A., Orozco, B. M., and Hanley-Bowdoin, L. 1994. A DNA sequence required for geminivirus replication also mediates transcriptional regulation. *Plant Cell* 6:1157-1170.
- Elmer, J. S., Brand, L., Sunter, G., Gardiner, W. E., Bisaro, D. M., and Rogers, S. G. 1988. Genetic analysis of *Tomato golden mosaic virus* II. The product of the AL1 coding sequence is required for replication. *Nucleic Acids Res.* 16:7043-7060.
- Fontes, E. P. B., Eagle, P. A., Sipe, P. S., Luckow, V. A., and Hanley-Bowdoin, L. 1994. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *J. Biol. Chem.* 269:8459-8465.
- Frischmuth, S., Frischmuth, T., and Jeske, H. 1991. Transcript mapping of *Abutilon mosaic virus*. *Virology* 185:596-604.
- Hamilton, W. D. O., Bisaro, D. M., Coutts, R. H. A., and Buck, K. W. 1983. Demonstration of the bipartite nature of the genome of a single-stranded DNA plant virus by infection with the cloned DNA components. *Nucleic Acids Res.* 11:7387-7396.
- Hamilton, W. D. O., Stein, V. E., Coutts, R. H. A., and Buck, K. W. 1984. Complete nucleotide sequence of the infectious cloned DNA components of *Tomato golden mosaic virus*: Potential coding regions and regulatory sequences. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:2197-2205.
- Hanley-Bowdoin, L., Elmer, J. S., and Rogers, S. G. 1988. Transient expression of heterologous RNAs using *Tomato golden mosaic virus*. *Nucleic Acids Res.* 16:10511-10528.
- Hanley-Bowdoin, L., Elmer, J. S., and Rogers, S. G. 1989. Functional expression of the leftward open reading frames of the A component of *Tomato golden mosaic virus* in transgenic plants. *Plant Cell* 1:1057-1067.
- Hanley-Bowdoin, L., Elmer, J. S., and Rogers, S. G. 1990. Expression of functional replication protein from *Tomato golden mosaic virus* in transgenic tobacco. *Proc. Natl. Acad. Sci. U.S.A.* 87:1446-1450.
- Hanley-Bowdoin, L., Settlege, S., Orozco, B. M., Nagar, S., and Robertson, D. 1999. Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. *Crit. Rev. Plant Sci.* 18:71-106.
- Hao, L., Wang, H., Sunter, G., and Bisaro, D. M. 2003. Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. *Plant Cell* 15:1034-1048.
- Hartitz, M. D., Sunter, G., and Bisaro, D. M. 1999. The geminivirus transactivator (TRAP) is a zinc-binding phosphoprotein with an acidic activation domain. *Virology* 263:1-14.
- Horsch, R. B., and Klee, H. J. 1986. Rapid assay of foreign gene expression in leaf discs transformed by *Agrobacterium tumefaciens*: Role of the T-DNA borders in the transfer process. *Proc. Natl. Acad. Sci. U.S.A.* 83:4428-4432.
- Jeffrey, J. L., Poama, W., and Petty, I. T. D. 1996. Genetic requirements for local and systemic movement of *Tomato golden mosaic virus* in infected plants. *Virology* 223:208-218.
- Johansen, L. K., and Carrington, J. C. 2001. Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiol.* 126:930-938.
- Joshi, C. P., Zhou, H., Huang, X., and Chiang, V. L. 1997. Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* 35:993-1001.
- Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12:857-872.
- Kozak, M. 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* 361:13-37.
- Lazarowitz, S. G. 1992. Geminiviruses: Genome structure and function. *Crit. Rev. Plant Sci.* 11:327-349.
- Mullineaux, P. M., Rigden, J. E., Dry, I. B., Krake, L. R., and Rezaian, M. A. 1993. Mapping of the polycistronic RNAs of *Tomato leaf curl virus*. *Virology* 193:414-423.
- Petty, I. T. D., Coutts, R. H. A., and Buck, K. W. 1988. Transcriptional mapping of the coat protein gene of *Tomato golden mosaic virus*. *J. Gen. Virol.* 69:1359-1365.
- Pilartz, M., and Jeske, H. 1992. *Abutilon mosaic geminivirus* double-stranded DNA is packed into minichromosomes. *Virology* 189:800-802.
- Pilartz, M., and Jeske, H. 2003. Mapping of *Abutilon mosaic geminivirus* minichromosomes. *J. Virol.* 77:10808-10818.
- Preiss, W., and Jeske, H. 2003. Multitasking in replication is common among geminiviruses. *J. Virol.* 77:2972-2980.
- Rogers, S. G., Bisaro, D. M., Horsch, R. B., Fraley, R. T., Hoffmann, N. L., Brand, L., Elmer, J. S., and Lloyd, A. M. 1986. *Tomato golden mosaic virus* A component DNA replicates autonomously in transgenic plants. *Cell* 45:593-600.
- Rogers, S. G., Klee, H. G., Horsch, R. B., and Fraley, R. T. 1987. Improved vectors for plant transformation: Expression cassette vectors and new selectable markers. *Methods Enzymol.* 153:253-277.
- Rybicki, E. P. 1995. A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Arch. Virol.* 139:49-77.
- Settlege, S. B., Miller, A. B., Gruissem, W., and Hanley-Bowdoin, L. 2001. Dual interaction of a geminivirus replication accessory protein and a plant cell cycle regulator. *Virology* 279:570-576.
- Shivaprasad, P. V., Akbergenov, R., Trinks, D., Rajeswaran, R., Veluthambi, K., Hohn, T., and Pooggin, M. M. 2005. Promoters, transcripts, and regulatory proteins of *Mungbean yellow mosaic geminivirus*. *J. Virol.* 79:8149-8163.

- Stenger, D. C., Revington, G. N., Stevenson, M. C., and Bisaro, D. M. 1991. Replicational release of geminivirus genomes from tandemly repeated copies: Evidence for rolling circle replication of a plant viral DNA. *Proc. Natl. Acad. Sci. U.S.A.* 88:8029-8033.
- Sunter, G., and Bisaro, D. M. 1989. Transcription map of the B genome component of *Tomato golden mosaic virus* and comparison with A component transcripts. *Virology* 173:647-655.
- Sunter, G., and Bisaro, D. M. 1991. Transactivation in a geminivirus: AL2 gene product is needed for coat protein expression. *Virology* 180:416-419.
- Sunter, G., and Bisaro, D. M. 1992. Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. *Plant Cell* 4:1321-1331.
- Sunter, G., and Bisaro, D. M. 1997. Regulation of a geminivirus coat protein promoter by AL2 protein (TrAP): Evidence for activation and repression mechanisms. *Virology* 232:269-280.
- Sunter, G., and Bisaro, D. M. 2003. Identification of a minimal sequence required for activation of the *Tomato golden mosaic virus* coat protein promoter in protoplasts. *Virology* 305:452-462.
- Sunter, G., Gardiner, W. E., and Bisaro, D. M. 1989. Identification of *Tomato golden mosaic virus*-specific RNAs in infected plants. *Virology* 170:243-250.
- Sunter, G., Gardiner, W. E., Rushing, A. E., Rogers, S. G., and Bisaro, D. M. 1987. Independent encapsidation of *Tomato golden mosaic virus* A component DNA in transgenic plants. *Plant Mol. Biol.* 8:477-484.
- Sunter, G., Hartitz, M. D., and Bisaro, D. M. 1993. *Tomato golden mosaic virus* leftward gene expression: Autoregulation of geminivirus replication protein. *Virology* 195:275-280.
- Sunter, G., Hartitz, M. D., Hormuzdi, S. G., Brough, C. L., and Bisaro, D. M. 1990. Genetic analysis of *Tomato golden mosaic virus*. ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication. *Virology* 179:69-77.
- Sunter, G., J. L. Sunter, J. L., and Bisaro, D. M. 2001. Plants expressing *Tomato golden mosaic virus* AL2 or *Beet curly top virus* L2 transgenes show enhanced susceptibility to infection by DNA and RNA viruses. *Virology* 285:59-70.
- Townsend, R., Stanley, J., Curson, S. J., and Short, M. N. 1985. Major polyadenylated transcripts of *Cassava latent virus* and location of the gene encoding coat protein. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:33-37.
- Wang, H., Hao, L., Shung, C-Y., Sunter, G., and Bisaro, D. M. 2003. Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. *Plant Cell* 15:3020-3032.
- Wang, H., Buckley, K. J., Yang, X., Buchmann, R. C., and Bisaro, D. M. 2005. Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. *J. Virol.* 79:7410-7418.
- Yang, S.-J., Carter, S. A., Cole, A. B., Cheng, N-H., and Nelson, R. S. 2004. A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by *Nicotiana benthamiana*. *Proc. Natl. Acad. Sci. U.S.A.* 101:6297-6302.
- Zhan, X. C., Haley, A., Richardson, K., and Morris, B. 1991. Analysis of the potential promoter sequences of *African cassava mosaic virus* by transient expression of the β -glucuronidase gene. *J. Gen. Virol.* 72:2849-2852.