JOURNAL OF VIROLOGY, Jan. 1996, p. 148–158 Vol. 70, No. 1
0022-538X/96/$04.00+0

A DNA Structure Is Required for Geminivirus Replication Origin Function

B. M. OROZCO* AND L. HANLEY-BOWDOIN

Department of Biochemistry, North Carolina State University, Raleigh, North Carolina 27695-7622

Received 5 June 1995/Accepted 4 October 1995

The genome of the geminivirus tomato golden mosaic virus (TGMV) consists of two single-stranded circular DNAs, A and B, that replicate through a rolling-circle mechanism in nuclei of infected plant cells. The TGMV origin of replication is located in a conserved 5′ intergenic region and includes at least two functional elements: the origin recognition site of the essential viral replication protein, AL1, and a sequence motif with the potential to form a hairpin or cruciform structure. To address the role of the hairpin motif during TGMV replication, we constructed a series of B-component mutants that resolved sequence changes from structural alterations of the motif. Only those mutant B DNAs that retained the capacity to form the hairpin structure replicated to wild-type levels in tobacco protoplasts when the viral replication proteins were provided in *trans* from a plant expression cassette. In contrast, the same B DNAs replicated to significantly lower levels in transient assays that included replicating, wild-type TGMV A DNA. These data established that the hairpin structure is essential for TGMV replication, whereas its sequence affects the efficiency of replication. We also showed that TGMV AL1 functions as a site-specific endonuclease in vitro and mapped the cleavage site to the loop of the hairpin. In vitro cleavage analysis of two TGMV B mutants with different replication phenotypes indicated that there is a correlation between the two assays for origin activity. These results suggest that the in vivo replication results may reflect structural and sequence requirements for DNA cleavage during initiation of rolling-circle replication.

Geminiviruses are a unique family of eukaryotic DNA viruses with single-stranded genomes that replicate through double-stranded DNA intermediates in the nuclei of infected plant cells (for reviews, see references 37 and 66). They represent one of only two families of plant viruses with true DNA replication cycles. Geminivirus genomes are small, replicate to high copy number, and contain a well-defined origin of replication. They encode only one protein essential for their replication and recruit the rest of the replication machinery from their hosts (10, 23, 43). Geminivirus replication proceeds through a rolling-circle mechanism with separate leading- and lagging-strand DNA synthesis steps (25, 57, 58). These properties suggest that geminiviruses are excellent models for studying DNA replication in plants and for resolving the various steps involved in leading- and lagging-strand DNA synthesis.

The *Geminiviridae* family consists of three subgroups that differ with respect to insect vector, host range, and genome structure (4, 45, 50). Subgroup I includes leafhopper-transmitted viruses that generally infect monocot plants and have monopartite genomes. Subgroup III includes whitefly-transmitted viruses that infect dicot plants and most commonly have bipartite genomes. Subgroup II viruses are transmitted by leafhoppers and have monopartite genomes like those of subgroup I, but they infect dicot plants like subgroup III viruses do. Tomato golden mosaic virus (TGMV), a subgroup III geminivirus, has a bipartite genome consisting of two 2.6-kb circular components designated A and B (21). The A and B genomic DNAs are arranged similarly with bidirectional transcription units separated by a 5′ intergenic region that is 95% conserved between the two components (21). The A component encodes all the viral proteins required for replication and transcription (9, 19, 49, 60, 63), including the essential viral replication protein AL1 (10, 23). The B DNA, which cannot replicate in the absence of the A component, contributes functions involved in virus movement and symptom development (5).

The TGMV origin of replication is located in the conserved 5′ intergenic region or common region of the A and B components. Lazarowitz et al. (38) used deletion analysis to map the TGMV replication origin to the common region. Transient replication studies with tobacco protoplasts showed that an 89-bp sequence from the TGMV common region is sufficient to support episomal replication when AL1 is supplied in *trans* (16). The 89-bp sequence includes a high-affinity binding site for AL1 (14) and an inverted-repeat motif with the potential to form a hairpin or cruciform structure (21). Mutation of the AL1 binding site or the hairpin motif interferes with TGMV replication (12, 13, 47), indicating that both elements are essential for origin activity. There is evidence that the replication origins of other geminiviruses are also located in their 5′ intergenic regions. The origin of replication for squash leaf curl virus (SqLCV) has been mapped to its 5′ intergenic region (38). Analyses of replication intermediates and genomic recombinants indicated that replication of African cassava mosaic virus (ACMV) DNA initiates in the 5′ intergenic region (53, 57). Comparison of the 5′ intergenic regions of subgroup II and III geminiviruses uncovered sequences that resemble the TGMV AL1 binding site (1, 12). The hairpin motif is found in the 5′ intergenic regions of all known geminivirus genomes and contains a sequence in the putative loop region, 5′ TAAT ATTAC, that is evolutionarily conserved (49). As with TGMV, mutations or deletions in the hairpin motifs of SqLCV and ACMV interfere with viral replication (38, 48).

Many DNA replication origins contain inverted-repeat sequences, some of which have the capacity to form hairpin

* Corresponding author. Mailing address: Department of Biochemistry, Box 7622, North Carolina State University, Raleigh, NC 26795-7622. Phone: (919) 515-5736. Fax: (919) 515-2047. Electronic mail address: linda@bchserver.bch.ncsu.edu.

structures. DNA secondary structure has been shown to be important for origin activity of some bacterial plasmids, bacteriophages, and eucaryotic viruses (27, 39, 42, 44, 70) and may be involved in initiation of DNA synthesis in mammalian cells (71). The initiation and termination sites for geminivirus positive-strand replication of beet curly-top virus (58), wheat dwarf virus (WDV) (25), and ACMV (57) have been mapped in vivo to the invariant loop sequence of the putative hairpin. The AL1 protein, or its C1 homolog in monopartite geminiviruses, mediates initiation of positive-strand DNA synthesis during geminivirus replication. Previous studies showed that the AL1 proteins of TGMV and bean golden mosaic virus specifically interact with their respective positive-strand origins and are responsible for origin recognition (13). Recent experiments showed that the C1 proteins of tomato yellow leaf curl virus (TYLCV) and WDV have sequence-specific cleavage activity in vitro (26, 36), a function necessary for initiation of positive-strand DNA synthesis. The in vitro cleavage sites have been mapped to the hairpin motifs of both viral genomes. In this study, we examined the structural and sequence requirements of the hairpin motif during TGMV replication in vivo and asked if TGMV AL1 functions as a site-specific endonuclease in vitro.

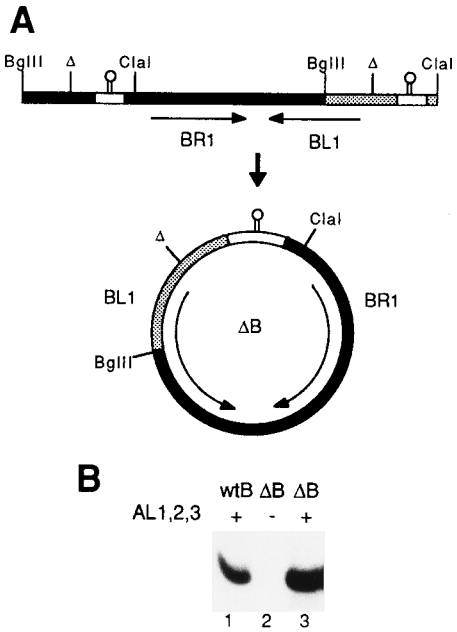

FIG. 1. Tandem partial repeats of the TGMV B genome and the monomeric component released during replication. (A) The TGMV B replicon contains 1.4 tandem repeats of the B component. A 0.7-kb duplication of the genome from the *Bgl*II site to the *Cla*I site includes the hairpin motif (♀) and a 20-bp deletion (△) that facilitated cloning of the hairpin mutants. Solid regions represent sequences from the full-length B DNA, stippled regions represent sequences from the 0.4-copy B component, and open regions mark the common region. Vector sequences are not shown. Arrows mark the positions of the open reading frames for BL1 and BR1. A circular, full-length copy of the B component is released by a nonrandom replicative mechanism. The replicated B component containing the 20-bp deletion is designated ΔB to distinguish it from the wild-type B DNA. (B) Replication levels of wild-type B and ΔB constructs were compared in transient assays with tobacco protoplasts. NT-1 cells (1.6×10^6) were electroporated with 30 μg of the B replicons and 15 μg of the AL1,2,3 expression cassette, pMON1548. Total DNA was isolated 2 days posttransfection, digested with *Bgl*II and *Dpn*I, resolved on 1% agarose gels, and analyzed by DNA gel blot hybridization with a ^{32}P-labeled TGMV B probe. Lane 1, wtB replication with pMON1548; lane 2, ΔB replication alone; lane 3, ΔB replication with pMON1548.

MATERIALS AND METHODS

The position numbers used to describe the clones listed below refer to the nucleotide coordinates of the TGMV sequence determined by Hamilton et al. (22), as corrected by MacDowell et al. (41), von Arnim and Stanley (67), and Schaffer (54).

Protoplast TGMV replicon plasmids. The pUC18-based vector pNSB109 (9) was modified by insertion of a *Cla*I linker into a *Hin*dIII site repaired by using *Escherichia coli* DNA polymerase I (Klenow fragment) to create pNSB226. pNSB226 was further modified by *Nde*I digestion, repair, and religation and by insertion of an *Ssp*I linker into a repaired *Xba*I site to give pNSB228. A *Bgl*II fragment containing a full-length copy of TGMV B from pMON393 (11) was cloned into pNSB109 to generate pNSB233. A *Bgl*II-*Cla*I fragment (TGMV B positions 1892 to 322), which contained a 0.4-copy TGMV B DNA including the common region, was cloned into *Bgl*II-*Cla*I-digested pNSB226 to give pNSB229. pNSB232 was created by deletion of 20 bp (TGMV B positions 2280 to 2299) between *Nae*I and *Ssp*I sites in the BL1 promoter of the 0.4-copy insert of pNSB229 (Fig. 1A). pNSB232 retained a single *Ssp*I site at TGMV B position 136 in the hairpin motif that was used in subsequent cloning and mutagenesis steps. The 186-bp *Ssp*I-*Cla*I (TGMV B positions 137 to 322) and 753-bp *Bgl*II-*Ssp*I (positions 1892 to 136) fragments from pNSB232 were cloned separately into pNSB228 at the same restriction sites to generate pNSB236 and pNSB237, respectively. These plasmids were used as templates for PCR-based mutagenesis.

Following mutagenesis, the *Ssp*I-*Cla*I-digested PCR products from pNSB236 or *Ssp*I-*Bgl*II-digested PCR products from pNSB237 were cloned into pNSB228 and verified by dideoxy DNA sequencing. The 0.4-copy TGMV B clones containing mutant hairpin motifs, pNSB232m, were created by combining mutant and wild-type fragments or two mutant fragments derived from pNSB236 and pNSB237, respectively. The clones are designated by the mutation number(s) in the hairpin motif. For example, pNSB232m1 contains mutation 1 shown in Fig. 2A. The 1.0-copy TGMV B mutants designated pNSB234m were generated by replacing the wild-type 0.4-copy fragment of pNSB233 with the corresponding mutant fragments from the pNSB232m plasmids. A *Sac*I site in the vector adjacent to the *Bgl*II site at TGMV B position 1892 and the *Cla*I site at position 322 were used for the replacement cloning. Replacement cloning was also used to construct pNSB234, a 1.0-copy plasmid with a wild-type hairpin sequence in the TGMV B background containing the 20-bp deletion in the BL1 promoter (Fig. 1A). The 1.4-copy pNSB235 and pNSB235m plasmids were then generated by insertion of the 1.0-copy *Bgl*II fragments from pNSB234 and pNSB234m into the corresponding 0.4-copy pNSB232 and pNSB232m plasmids previously digested with *Bgl*II. pMON1565 is a pUC18-based plasmid with a 1.5-copy wild-type TGMV A insert as described previously for pMON337 (11). pTG1.4B, which includes 1.4 copies of wild-type TGMV B, has been described elsewhere (13).

Agroinoculation plasmids. Agroinoculation vectors were constructed by digesting pNSB235, pNSB235m7+8, pNSB235m11+12, and pNSB235m13+14 with *Pvu*II and transferring the 1.4-copy TGMV B fragments into the *Agrobacterium* transformation vector pMON721 (35) previously digested with *Sma*I. The resulting clones, designated pNSB317, pNSB317m7+8, pNSB317m11+12, and pNSB317m13+14, were introduced by direct DNA transformation (30) into *Agrobacterium tumefaciens* ABI, which contains a chromosomal copy of pMP90RK (71). Plasmids pMON337 and pMON393 correspond to wild-type TGMV A and B agroinoculation vectors, respectively (11).

Plant expression cassette. pMON1548 (9) is a plant expression cassette with the TGMV AL1, AL2, and AL3 coding sequences under the control of the cauliflower mosaic virus 35S promoter with a duplicated enhancer region and the pea rbcS E9 3′ end (9).

Baculovirus transfer plasmids. A 0.7-kb *Sac*I-*Hin*dIII fragment containing the glutathione *S*-transferase (GST) coding sequence was isolated from pEG(KT) (56) and cloned into *Sac*I-*Hin*dIII-digested pUC119 to give pNSB310. A 1.2-kb fragment containing the AL1 coding sequence was isolated from pMON1548 (9) by digestion with *Nde*I and *Bam*HI, repaired with Klenow fragment, and inserted into the *Sma*I site of pNSB310. The resulting plasmid, pNSB312, contains an in-frame fusion of the GST and AL1 coding sequences. A 2.0-kb *Sac*I fragment including the GST-AL1 fusion was isolated from pNSB312, trimmed with T4 DNA polymerase, and inserted into a repaired *Hin*dIII site in the baculovirus transfer plasmid pMON27025 (40) to give pNSB314. The *Sac*I fragment from pNSB310, which contains the GST coding sequence, was also subcloned in pMON27025 to give pNSB313.

PCR mutagenesis. Plasmids pNSB236 and pNSB237 were used as templates for PCR-based mutagenesis of the left and right sides of the TGMV B hairpin, respectively. All mutations in the hairpin created G→T, C→A, T→G, and A→C changes, with the following exceptions: A^{125}/T^{126}→T/A (A at position 125 and T at position 126 were changed to T and A, respectively) (mutant 7), T^{145}/A^{146}→A/T (mutant 8), T^{130}/T^{131}→C/C (mutant 15), T^{130}/T^{132}→G/C (mutant 17), T^{130}/T^{132}→G/A (mutant 18), and A^{140}/C^{141}→G/G (mutant 19). The superscripts indicate TGMV B position numbers. The left side of the hairpin was modified by using the M13/pUC reverse sequencing primer −48 (New England Biolabs; Beverly, Mass.) and primers containing mutations in the wild-type sequence 5′-ctagaAATATTAAACGGATGGCCGCCACGTGC at TGMV B positions 121 and 122 (mutant 1), 122 and 123 (mutant 3), 123 and 124 (mutant 5),

125 and 126 (mutant 9), 127 and 128 (mutant 11), 130 (mutant 16), and 130 and 131 (mutant 15). Uppercase letters represent viral sequences, and lowercase letters represent vector sequences. The right side of the hairpin was modified by using an M13/pUC sequencing primer (−47; New England Biolabs) and primers containing mutations in the sequence 5′-ctagaAATATTACCGGATGGCCGCG CGAT at TGMV B positions 143 and 144 (mutant 12), 145 and 146 (mutant 10), 147 and 148 (mutant 6), 148 and 149 (mutant 4), and 149 and 150 (mutant 2). The PCR mixtures included 25 ng of plasmid DNA, 20 pmol of each primer, 0.25 mM deoxynucleoside triphosphate, 3 mM $MgSO_4$, 1× Vent polymerase buffer, and 0.5 μl of Vent *Taq* polymerase (New England Biolabs). The fragments were amplified for 30 cycles at 94°C for 1 min, 55°C for 30 s, and 72°C for 1 min. The PCR products were digested, cloned, and sequenced as described above.

Transient replication and whole-plant infectivity assays. Protoplasts were prepared from *Nicotiana tabacum* suspension cells (NT-1) and electroporated as described previously (12). Protoplasts (5×10^6) were transfected with 30 μg of sheared salmon sperm DNA, 30 μg of a TGMV B replicon, and either 20 μg of pMON1548 or 15 μg of pMON1565. Total DNA was isolated 72 h posttransfection (32), digested with *Bgl*II and *Dpn*I, and analyzed by DNA gel blot hybridization. Each replicon was tested in at least four independent experiments.

Nicotiana benthamiana plants were inoculated with mixed *A. tumefaciens* cultures carrying a 1.5-copy TGMV A vector (pMON337) and one of the 1.4-copy B vectors (pMON393, pNSB317, and pNSB317m13+14), as reported by Elmer et al. (11). Inoculated plants were observed for 3 weeks for viral symptoms. At 3 weeks postinoculation, each plant was analyzed for viral DNA accumulation in upper leaves by squash blot hybridization. Leaf discs were gently macerated on nylon membrane (Magnagraph; MSI, Westboro, Mass.) to release DNA, UV cross-linked, and analyzed by hybridization with a ^{32}P-labeled TGMV B probe. Total DNA was also isolated from upper leaves of representative plants inoculated with pMON337 and either pMON393, pNSB317, or pNSB317m13+14, and TGMV B DNA accumulations were compared by DNA gel blot hybridization.

N. benthamiana leaf discs were inoculated with mixed *A. tumefaciens* cultures carrying a 1.5-copy TGMV A vector (pMON337) and one of the 1.4-copy B vectors (pMON393, pNSB317, pNSB317m7+8, pNSB317m11+12, and pNSB317m13+14), according to the protocol of Elmer et al. (10). Total DNA was extracted from leaf discs 3 days postinoculation, and 5-μg aliquots were digested with *Bsa*WI and mung bean nuclease. The digested samples and 5 μg of undigested DNA were analyzed by DNA gel blot hybridization.

DNA gel blot hybridizations were performed as described by Thomashow et al. (64), except that 0.2% sodium PP_i was included in all solutions. DNA samples were resolved on 1% agarose gels, transferred to nylon membrane, and UV cross-linked. A 2.1-kb *Pst*I fragment of TGMV B isolated from pTG1.4B was randomly labeled with [α-^{32}P]dATP and the Klenow fragment (51) to generate a component-specific probe.

Expression and purification of GST-AL1. The GST and GST-AL1 expression cassettes from pNSB313 and pNSB314 were integrated into the parent bacmid vector, bMON14272, in *E. coli*, as described by Luckow et al. (40). Recombinant bacmid DNAs were purified from *E. coli* by using Qiagen DNA purification columns (Qiagen, Chatsworth, Calif.) and transfected into *Spodoptera frugiperda* Sf9 insect cells by calcium phosphate coprecipitation. Baculovirus-mediated expression of GST and GST-AL1 was confirmed by immunoblot analysis with polyclonal antibodies against GST (Upstate Biotechnology Inc., Lake Placid, N.Y.) and AL1 (23). High-titer lysates corresponding to both baculoviruses were prepared and used for subsequent infections.

Sf9 cells were infected with the recombinant baculoviruses at a multiplicity of infection of 1.0. Cells were harvested 48 h postinfection, washed twice with Grace's serum-free medium (Life Technologies, Gaithersburg, Md.), and lysed in extraction buffer (EB) containing 25 mM Tris-HCl, pH 7.5; 1 mM EDTA; 1 mM dithiothreitol; 1% Nonidet P-40; 10 μg of pepstatin per ml; 50 μg of leupeptin per ml; and 10 mM phenylmethylsulfonyl fluoride. After centrifugation at 5,000 × *g* for 10 min, GST or GST-AL1 protein was extracted from the supernatant by binding to reduced glutathione-Sepharose 4B (Pharmacia Biotech Inc., Piscataway, N.J.) (56), either by passage twice over a 1-ml column or by mixing for 1 h with the resin prior to pouring of a 1-ml column. The bound protein was washed with 10 column volumes of EB and eluted in 2 ml of 10 mM reduced glutathione–50 mM Tris-HCl, pH 7.5. The purified proteins were analyzed on sodium dodecyl sulfate (SDS)–12.5% polyacrylamide gels and then subjected to silver staining or immunoblotting with anti-AL1 and anti-GST antibodies.

In vitro cleavage assays. Purified GST (650 ng) and GST-AL1 (ca. 160 ng in 650 ng of total protein) were incubated with radiolabeled oligonucleotides (ca. 3 fmol; 10,000 cpm) in 7.5 μl of cleavage buffer (25 mM Tris-HCl, pH 7.5; 75 mM NaCl; 5 mM $MgCl_2$; 2.5 mM EDTA; and 2.5 mM dithiothreitol) at 37°C for 30 min. The reactions were terminated with 5 μl of gel loading buffer (95% formamide, 20 mM EDTA, 0.05% bromphenol blue), mixtures were heated to 90°C for 2 min, and the products were resolved on 15% polyacrylamide–8 M urea gels. Gels were dried onto Whatman 3MM paper and autoradiographed.

Four synthetic oligonucleotides corresponding to the loop and right side of the hairpin motif and an oligonucleotide corresponding to the entire positive-strand hairpin were used as substrates for cleavage (see Fig. 7). The sequences of the oligonucleotides were 5′-GTTTAATATTACCGGATGGCCGC (positive strand), 5′-GTTTAATATTggCGGATGGCCGC (positive strand, mutation 19), 5′-GTT TAATATTACCttcgttaaGC (positive strand, mutation 13), 5′-GCGGCCATCCG GTAATATTAAAC (negative strand), and 5′-GCGGCCTACCGTTTAATAT TACCGTAGGCCGC (hairpin). Uppercase letters are wild-type sequences, and lowercase letters represent mutations. The numbers correspond to mutation numbers used in site-directed mutagenesis of the hairpin. Oligonucleotides were radiolabeled with [γ-^{32}P]ATP and T4 polynucleotide kinase, purified by denaturing gel electrophoresis, and resuspended in water to 10,000 cpm/μl.

RESULTS

DNA secondary structure is required for TGMV replication. The 5′ intergenic regions of all known geminivirus genomes contain an inverted repeat with the potential to form a hairpin or cruciform structure (37). We constructed a series of TGMV B mutants that resolved sequence changes from structural alterations of the hairpin motif and asked what impact these mutations had on viral replication in vivo. For these experiments, we used constructs that contained 1.4 copies of the TGMV B genome with duplicated wild-type or mutant common regions. Previous studies have shown that unit-length, positive-strand geminivirus DNA between two common regions is released by rolling-circle replication and efficiently replicated (58).

To facilitate cloning of the mutants, we first constructed a 1.4-copy TGMV B plasmid (ΔB, pNSB235) with a 20-bp deletion between TGMV B positions 2280 and 2299 (Fig. 1A). The deletion destroyed an *Ssp*I site in the BL1 promoter such that ΔB contained a single *Ssp*I site at TGMV B position 136 in the common region. Levels of replication from ΔB and the equivalent wild-type TGMV B plasmid (pTG1.4B) were compared to determine if the deletion affected replication. Each of the B DNAs was coelectroporated into tobacco protoplasts with an AL1,2,3 plant expression cassette (pMON1548), which provided the viral replication proteins AL1 and AL3 in *trans* (9). The AL1,2,3 expression cassette was used because, unlike TGMV A, it does not contain a wild-type viral origin that might compete or recombine with mutant TGMV B origins. After a 72-h culture period, total DNA was isolated from the transfected cells, digested with *Bgl*II-*Dpn*I, and analyzed by DNA gel blot hybridization. *Bgl*II linearized TGMV B DNA, whereas *Dpn*I digested Dam-methylated, input DNA. A single 2.5-kb band corresponding to unit-length, double-stranded TGMV B diagnostic of viral replication was detected in cells transfected with the AL1,2,3 expression cassette and wild-type TGMV B (Fig. 1B, lane 1) or ΔB (lane 3). The relative intensities of the bands in lanes 1 and 3 were similar, indicating that the 20-bp deletion in ΔB did not impact TGMV B replication in transient assays. No replication of ΔB was detected in the absence of the AL1,2,3 expression cassette (Fig. 1B, lane 2), verifying that replication was dependent on expression of the AL1 and AL3 proteins.

Six 2-bp mutations were introduced at various positions on each side of the TGMV B stem sequence in the 1.4-copy ΔB background. To increase the probability that the mutations destabilized the putative stem structure, adjacent positions were modified and the modifications were designed to minimize the formation of stable non-Watson-Crick base pairs (29). Complementary mutations were also combined to give six mutants that contained compensatory changes on both sides of the putative stem which altered DNA sequence but maintained potential base pairing. Figure 2A shows the position of each mutation in the stem of the hairpin.

All of the stem mutants were analyzed for replication in tobacco protoplasts in the presence of the AL1,2,3 expression cassette (Fig. 2B). The left-side (no. 3, 5, 7, 9, and 11; Fig. 2B, lanes 4 to 8) and right-side (no. 2, 4, 6, 8, 10, and 12; lanes 3 to 8) mutants replicated to significantly lower levels than did ΔB (double, lane 2), with the exception of mutant 1 (left, lane 3). In contrast, the compensatory double mutants (mutants 1+2,

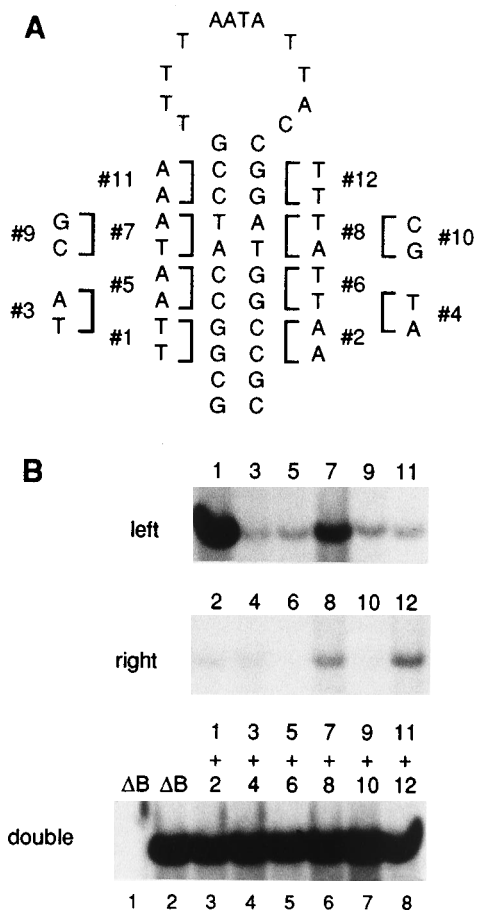

FIG. 2. Transient replication of TGMV B components with mutations in the stem of the hairpin. (A) The wild-type sequence of the TGMV B component from positions 119 to 152 is shown with its putative hairpin structure. Mutations generated in the stem sequence of the hairpin are shown in brackets adjacent to the wild-type sequence and designated by number. (B) *N. tabacum* protoplasts were cotransfected with the TGMV B replicons and pMON1548. Total DNA was isolated 2 days posttransfection, digested with *Bgl*II and *Dpn*I, and analyzed by DNA blot hybridization. Newly synthesized DNAs from the various replicons are indicated at the top of each gel by the mutation numbers shown in panel A. Lane numbers are given at the bottom. Lane 1, ΔB alone; lane 2, ΔB with pMON1548; lanes 3 to 8, mutant B components with pMON1548. (Top) DNA replication from constructs containing mutations in the left side of the hairpin (mutants 1, 3, 5, 7, 9, and 11); (center) replication from constructs containing mutations in the right side of the hairpin (mutants 2, 4, 6, 8, 10, and 12); (bottom) replication from constructs containing complementary mutations on both sides of the hairpin (mutants 1+2, 3+4, 5+6, 7+8, 9+10, and 11+12). Replication levels of the left-side, right-side, and double mutants were analyzed on the same autoradiogram.

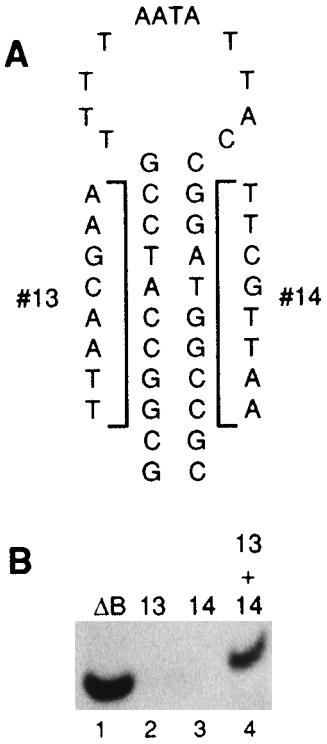

FIG. 3. Transient replication of TGMV B DNAs with 8-bp mutations in the stem of the hairpin. (A) The wild-type sequence of the TGMV B component from positions 120 to 152 is shown with its putative hairpin structure. The 8-bp mutations in the stem of the hairpin are shown in brackets and designated #13 (left-side mutations) and #14 (right-side mutations). (B) *N. tabacum* protoplasts were cotransfected with the TGMV B replicons and pMON1548. Total DNA was isolated 2 days posttransfection, digested with *Bgl*II and *Dpn*I, and analyzed by DNA blot hybridization with a ^{32}P-labeled TGMV B probe. Newly synthesized DNAs from the various replicons are indicated at the top by the mutation numbers shown in panel A. Lane numbers are given at the bottom. Lane 1, ΔB DNA; lanes 2 to 4, ΔB DNA containing mutations 13, 14, and both 13 and 14, respectively.

3+4, 5+6, 7+8, 9+10, and 11+12) replicated to the same level as ΔB did (Fig. 2B, double, lanes 2 to 8). The low levels of de novo synthesis seen with the left- and right-side mutants may reflect residual base pairing across the stem or the formation of stem structure with an internal bulge capable of supporting some replication. To address this possibility, mutants 13 and 14, which had alterations of 8 of the 11 nucleotides (nt) in the stem on the left and right sides, respectively, and the compensatory double mutant 13+14 (Fig. 3A) were made and analyzed for replication in tobacco protoplasts. No replication was detected for mutants 13 (Fig. 3B, lane 2) and 14 (lane 3), whereas mutant 13+14 (lane 4) replicated to near-wild-type levels relative to ΔB (lane 1). These results established the importance of the hairpin structure in TGMV replication and indicated that there is flexibility in the sequence of the stem and its AT-versus-GC content as long as base pairing is maintained.

It is not clear why mutant 1, which is modified at the base of the stem on the left side, replicated efficiently, unlike other left- and right-side mutants (Fig. 2B). One explanation is that the replication sequence requirement of the left side is less stringent than that of the right side. In general, right-side mutants (mutants 2, 4, 6, 8, and 10) replicated to lower levels than did left-side mutants (mutants 1, 3, 5, 7, and 9). The differences were slight but consistent over six to eight repetitions. A second possibility is that mutation 1 was less disruptive to the stem structure than was either mutation 2 or 3.

The loop structure and sequence display some flexibility in replication. Five mutations were introduced into the hairpin loop to examine the effect of structure and sequence of this region on TGMV replication. Mutations 15, 16, 17, and 18 were located in the variant loop sequence on the left side of the loop, whereas mutation 19 was in the invariant sequence on the right side (Fig. 4A). Mutations 16 and 17 are predicted to alter structure by reducing the loop size by 4 nt and increasing the stem length by 2 bp. Mutation 18 is predicted to reduce the loop size by 6 nt and increase the stem length by 3 bp. Mutations 15 and 19 should not alter structure. Mutants 15, 16, 17, and 18 replicated to the same level as did ΔB in the presence

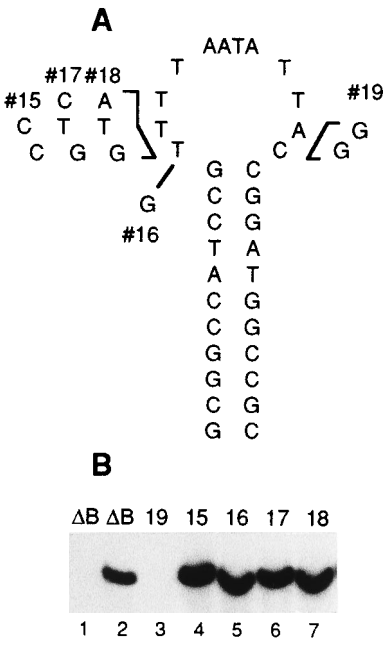

FIG. 4. Transient replication of TGMV B components with mutations in the loop of the hairpin. (A) The wild-type sequence of the TGMV B component from positions 119 to 152 is shown with its putative hairpin structure. Mutations generated in the loop sequence of the hairpin are shown in brackets and designated by number. (B) *N. tabacum* protoplasts were cotransfected with the TGMV B replicons and pMON1548. Total DNA was isolated 2 days posttransfection, digested with *Bgl*II and *Dpn*I, and analyzed by DNA blot hybridization with a ^{32}P-labeled TGMV B probe. Newly synthesized DNAs from the various replicons are indicated at the top by the mutation numbers shown in panel A. Lane numbers are given at the bottom. Lane 1, DNA B from ΔB alone; lanes 2 to 7, B components cotransfected with pMON1548: ΔB DNA (lane 2) and ΔB DNA containing hairpin mutations 19, 15, 16, 17, and 18 (lanes 3 to 7, respectively).

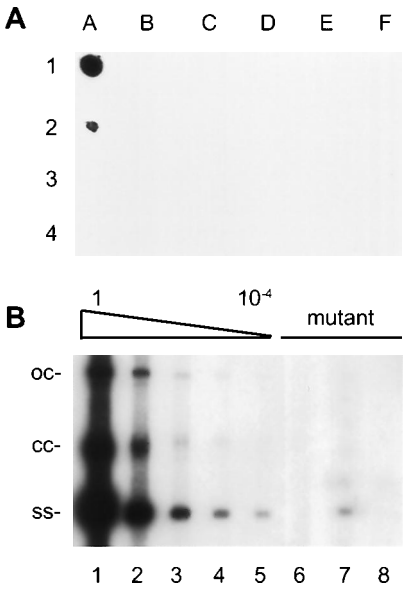

FIG. 5. Replication of TGMV B DNA in infected plants. (A) Squash blot analysis of plants infected with a wild-type A component and a wild-type or mutant B component. *N. benthamiana* plants were agroinoculated with the TGMV A component (pMON337) and a wild-type B component (pMON393) or ΔB component containing the hairpin mutation 13+14 (pNSB317m13+14). Leaf samples were macerated onto nylon membrane and analyzed by DNA blot hybridization with a ^{32}P-labeled TGMV B probe. Lane A, rows 1 and 2, wild-type B DNA accumulation; lane A, row 3, an uninfected plant; lane A, row 4, and lanes B to F, the 21 plants infected with the mutant B component. (B) A 10-fold dilution series of total DNA from a plant infected with the wild-type B component was compared with total DNA from three of the plants infected with the mutant B component by DNA blot hybridization with a ^{32}P-labeled TGMV B probe. Open-circle (oc), closed-circle (cc), and single-stranded (ss) forms of the DNA are indicated on the left. Lanes 1 to 5, total DNA (5, 0.5, 0.05, 0.005, and 0.0005 μg, respectively) from plants infected with the wild-type B component; lanes 6 to 8, total DNA (5 μg) from plants infected with the mutant B component.

of the AL1,2,3 plant expression cassette in tobacco protoplasts (Fig. 4B; cf. lane 2 and lanes 4 to 7). In contrast, no replication was detected for mutant 19 (Fig. 4B, lane 3). These results demonstrated that there are sequence requirements for TGMV replication in the invariant sequence of the loop. They also showed that a functional origin can accommodate sequence changes in the loop outside the conserved motif and structural alterations in the hairpin.

B components with mutant hairpin sequences are not infectious and do not replicate efficiently in leaf disc assays. Transient replication analysis of the TGMV B hairpin mutants in the presence of the AL1,2,3 expression cassette (Fig. 2 to 4) indicated that there is flexibility in sequence and structure as long as a stable hairpin can be formed and the invariant sequence is unchanged. This is surprising, because the sequence of the stem is highly conserved among subgroup II and III geminiviruses (48). A possible explanation for these observations is that the hairpin has a second role in the virus infection process. To address this hypothesis, *N. benthamiana* plants were agroinoculated with a wild-type A component (pMON337 [11]) and either wild-type TGMV B (pMON393 [11]), ΔB (pNSB317), or mutant 13+14 containing the compensatory 8-bp changes in the stem (pNSB317m13+14). Eighteen of the 21 plants inoculated with wild-type A and ΔB components developed symptoms that were indistinguishable from those of plants agroinoculated with wild-type TGMV A and B components. In contrast, none of the 21 plants inoculated with the wild-type A component and mutant 13+14 showed symptoms 28 days postinoculation. The assay was repeated twice more with similar results. TGMV B DNA could not be detected by squash blot analysis in upper leaves of any of the plants inoculated with mutant 13+14 but could be detected readily in equivalent, systemically infected leaves of plants inoculated with ΔB (Fig. 5A). Total DNA was also extracted from upper leaves of three plants inoculated with mutant 13+14 and one plant infected with ΔB. DNA gel blot hybridization showed at least a 1,000-fold reduction in TGMV B DNA accumulation in the plants inoculated with mutant 13+14 in comparison with the level in the plant inoculated with ΔB (Fig. 5B). These experiments showed that the sequence of the stem strongly impacts TGMV infectivity.

One explanation of the infectivity results is that the stem sequence impacts single-stranded DNA accumulation, which was not readily assayed in the protoplast replication system. This hypothesis was examined by using a leaf disc replication assay which efficiently detects both double- and single-stranded viral DNAs (10). *N. benthamiana* leaf discs were agroinoculated with wild-type TGMV A and either wild-type TGMV B (wtB), ΔB, mutant 7+8, mutant 11+12, or mutant 13+14. Total leaf disc DNA was isolated 2 days postinoculation and analyzed by DNA gel blot hybridization (Fig. 6A). Single-stranded DNA accumulations for mutants 7+8 (Fig. 6A, lane 3) and 11+12 (lane 4) were comparable to those observed for wtB (lane 1) and ΔB (lane 2). However, single-stranded DNA was significantly reduced with mutant 13+14 (lane 5). In con-

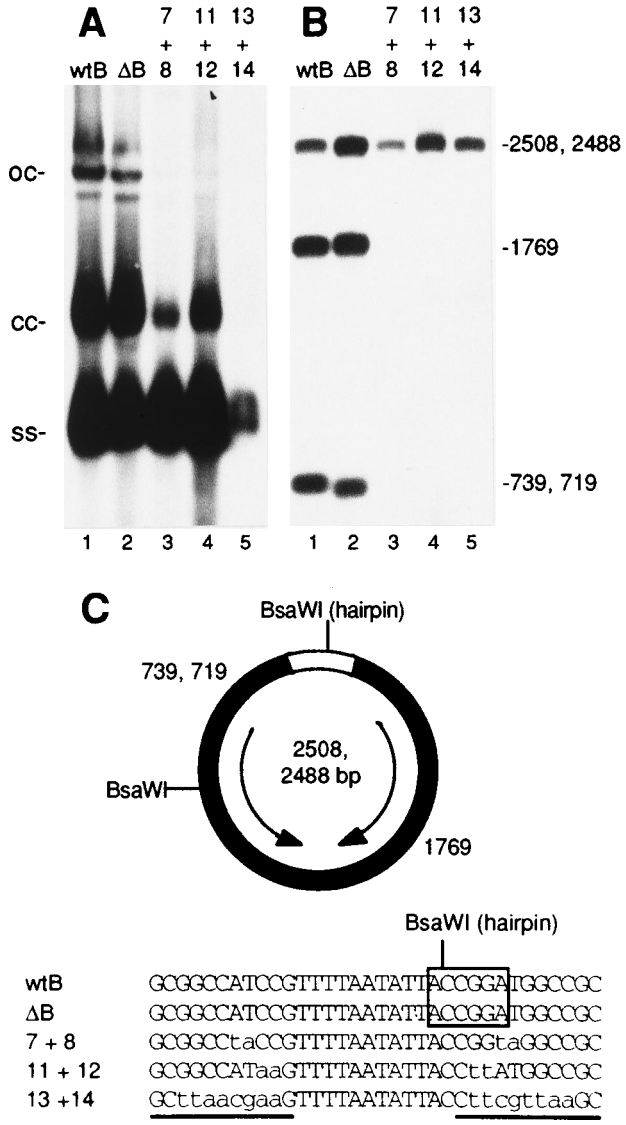

FIG. 6. DNA blot hybridization analysis of B-component accumulation in leaf discs. *N. benthamiana* leaf discs were agroinoculated with the TGMV A component (pMON337) and wild-type or mutant B components. Total DNA was isolated 2 days postinoculation and analyzed by DNA blot hybridization with a ^{32}P-labeled TGMV B probe. Newly synthesized DNAs from the various B replicons are indicated at the top of the gels by the mutation numbers shown in Fig. 2A or 3A. Lane numbers are given at the bottom. (A) Uncut total DNA. Positions of single-stranded (ss), closed-circular (cc), and open-circular (oc) forms of TGMV B DNA are indicated on the left. (B) Total DNA digested with mung bean nuclease and *Bsa*WI. The lengths of *Bsa*WI restriction fragments are shown on the right. A 20-bp deletion in ΔB compared with wtB is indicated by the second set of numbers. For both panels, lane 1, wild-type B DNA (wtB); lane 2, deletion derivative of wild-type B (ΔB); lane 3, ΔB DNA with mutations 7 and 8; lane 4, ΔB DNA with mutations 11 and 12; lane 5, ΔB DNA with mutations 13 and 14. (C) Diagram of TGMV B component showing positions of *Bsa*WI restriction sites and sizes of restriction fragments. The sequences of wild-type hairpins (B and ΔB) and mutant hairpins (7+8, 11+12, and 13+14) are shown. Mutations are indicated by lowercase letters. Horizontal lines mark the stem sequence. The boxed *Bsa*WI site in the wild-type sequence is destroyed in the three mutants.

trast to the replication levels observed to occur in protoplasts in the presence of the AL1,2,3 expression cassette, double-stranded DNA accumulation in leaf discs was severely impaired, with a reduction seen for all three hairpin mutants in comparison with levels for wtB and ΔB (Fig. 6A, cf. lanes 3 to 5 with lanes 1 and 2).

The reduced levels of viral DNA in leaf discs may reflect a delay in replication until the mutant B DNAs recombined with the coinoculated A component to restore the wild-type hairpin, as has been observed for ACMV (48). To determine whether recombination occurred between the mutant TGMV B components and the wild-type TGMV A component, DNA samples were digested with *Bsa*WI and mung bean nuclease. Mutations 7+8, 11+12, and 13+14 destroy a *Bsa*WI restriction site that can be used to distinguish between wild-type and mutant hairpin sequences (Fig. 6C). Mung bean nuclease digests single-stranded DNA which migrates to the same position as the *Bsa*WI restrictions fragments on agarose gels. Digestion of the wild-type hairpin sequence produced two *Bsa*WI fragments, 1,769 and 739 bp for wtB (Fig. 6B, lane 1) and 1,769 and 719 bp for ΔB (lane 2). The linear, full-length products (2,508 and 2,488 bp) in lanes 1 and 2 of the gel shown in Fig. 6B resulted from partial digestion of the samples. In contrast, *Bsa*WI digestion of mutants 7+8 (Fig. 6B, lane 3), 11+12 (lane 4), and 13+14 (lane 5) resulted in only the full-length, linear B DNA. The leaf disc assays indicated that none of the three hairpin mutants reverted to wild-type sequence through recombination with A-component DNA. They also showed that the hairpin mutants replicated with a lower degree of efficiency than did wtB and ΔB in leaf discs.

B components with mutant hairpin sequences do not replicate efficiently in the presence of TGMV A. The hairpin mutations had different effects on TGMV replication according to the assay system. Double-stranded DNA levels were reduced for mutants with compensatory changes in the hairpin motif in leaf disc and infectivity assays, whereas double-stranded DNA accumulation was not adversely affected by the same mutations in protoplast assays. A major difference between the assay systems is that a TGMV A component provided the viral replication proteins in leaf discs and infected plants, while an AL1,2,3 plant expression cassette produced the proteins in protoplasts. One possibility is that the wild-type origin on the TGMV A component competed with the TGMV B hairpin mutants and adversely affected their replication in leaf discs and infected plants.

This hypothesis was tested in protoplast replication assays by using a wild-type A component (pMON1565) in place of the AL1,2,3 expression cassette (Fig. 7). Mutations 5+6 (Fig. 7A, lane 4) and 11+12 (lane 7) and mutation 13+14 (lane 8) replicated to significantly lower levels than did ΔB (lane 1) in the presence of TGMV A. As a control, mutant and wild-type B DNA replication levels were compared in the presence of the AL1,2,3 expression cassette, the wild-type A component, or both. These results showed that mutant B DNA replication was reduced in comparison with replication of wild-type B DNA in the presence of the wild-type A component even when saturating levels of AL1 were provided from the expression cassette (data not shown). DNA replication of the A component was unaffected by mutations in the B component (data not shown). Two possible explanations for these results are that there are stem sequence constraints on replication and that the stem sequence may affect replication by influencing structural stability. In either case, the impact of the mutations was observed only in the presence of the wild-type A component.

Effects of mutations in the variant sequence of the loop were also tested in protoplast replication assays in the presence of the wild-type A component. Similar to the situation with the stem mutations, changes in the loop sequence had a detrimental effect on replication efficiency in the presence of the A component. Mutations 17 (Fig. 7B, lane 5) and 18 (lane 6),

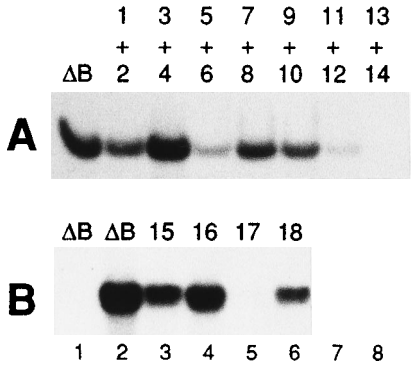

FIG. 7. Transient replication assays with the mutant TGMV B components and a wild-type TGMV A component. *N. tabacum* protoplasts were coelectroporated with 30 μg of the TGMV B replicons and 20 μg of the TGMV A replicon, pMON1565. Total DNA was isolated 2 days posttransfection, digested with *Bgl*II and *Dpn*I, and analyzed by DNA blot hybridization with a ^{32}P-labeled TGMV B probe. The newly synthesized B DNAs from the various replicons are indicated at the top by the mutation numbers shown in Fig. 2A, 3A, and 4A. Lane numbers are at the bottom. (A) DNA B accumulation in protoplasts from replicons containing compensatory mutations on both sides of the hairpin stem. Lane 1, ΔB DNA; lanes 2 to 8, ΔB with mutations 1+2, 3+4, 5+6, 7+8, 9+10, 11+12, and 13+14, respectively. (B) DNA B accumulation from replicons containing mutations in the hairpin loop. Lane 1, ΔB DNA alone; lanes 2 to 6, B components transfected with pMON1565: ΔB DNA (lane 2) and ΔB with mutations 15, 16, 17, and 18 (lanes 3 to 6, respectively).

which altered the loop sequence at both positions 130 and 132, reduced replication significantly in comparison with that of ΔB (lane 2). Mutations 15 (Fig. 7B, lane 3) and 16 (lane 4), which altered the sequence at the junction of the stem and loop at both positions 130 and 131, also impacted replication. Thus, both stem and loop sequences contribute to the activity of the hairpin motif during origin function when a wild-type competing replicon is present.

The hairpin structure and the stem sequence are not required for AL1-catalyzed DNA cleavage in vitro. The AL1 protein is likely to mediate initiation of positive-strand DNA synthesis by introducing a site-specific nick into the positive-strand origin of TGMV. On the basis of experiments with other geminiviruses (26, 36), the TGMV cleavage site is most likely located in the invariant sequence in the hairpin. We examined these possibilities by in vitro cleavage assays using TGMV AL1 purified as a GST fusion protein from insect cells and synthetic oligonucleotides corresponding to the TGMV hairpin sequence.

A GST-AL1 fusion protein was expressed in Sf9 insect cells by using a baculovirus expression system (40). GST-AL1 protein was purified from the cytoplasmic fraction by affinity chromatography with glutathione-Sepharose 4B. Four polypeptide bands were detected in the glutathione-eluted fraction resolved by SDS-polyacrylamide gel electrophoresis and visualized by silver staining (Fig. 8A, lane 3). Immunoblot analysis with an anti-AL1 antibody showed that the largest band, of ca. 65 kDa, corresponded to GST-AL1 (data not shown). GST polyclonal antibodies also cross-reacted with the GST-AL1 fusion protein and two of the low-molecular-weight, copurifying polypeptides (Fig. 8A, lane 4). Multiple low-molecular-weight bands were also detected in affinity-purified preparations of GST expressed in insect cells by silver staining (data not shown) and immunoblot analysis (Fig. 8A, lane 5). On the basis of silver staining, GST-AL1 accounted for 20 to 25% of the total protein eluted from the glutathione resin and endogenous GST proteins constituted most of the remaining 75 to 80%.

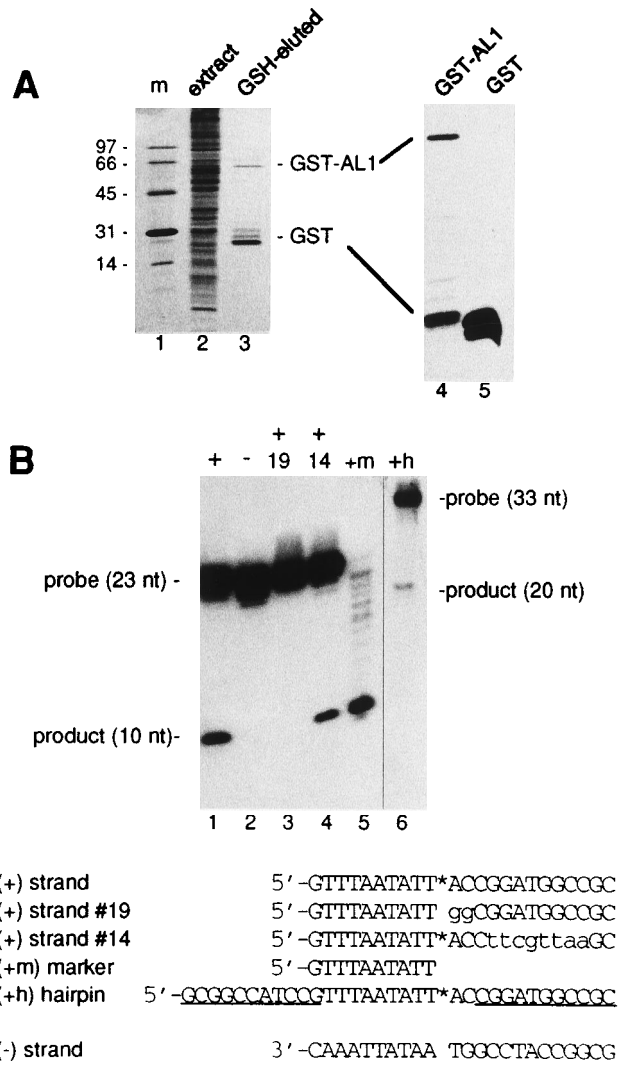

FIG. 8. In vitro cleavage assays of the hairpin sequence. (A) GST-AL1 was purified from Sf9 insect cells expressing the fusion protein from a recombinant baculovirus vector. Proteins were resolved on a 12.5% polyacrylamide–SDS gel and silver stained. Sizes of molecular mass markers in kilodaltons are shown on the left. Lanes 1 to 3, silver stain analysis; lanes 4 and 5, immunoblot analysis. Lane 1, molecular mass markers (m); lane 2, total cytoplasmic protein; lanes 3 and 4, glutathione-eluted proteins from an extract containing GST-AL1; lane 5, glutathione-eluted proteins from an extract containing GST. (B) Cleavage assays with purified GST-AL1. Oligonucleotides were end labeled with [γ-^{32}P]ATP and purified from 15% polyacrylamide gels. The radiolabeled substrates (3 fmol, ca. 10,000 cpm) were incubated with 650 ng of protein (ca. 160 ng of purified GST-AL1) for 30 min at 37°C. The cleavage products were resolved on a 15% polyacrylamide–8 M urea gel and autoradiographed. Lane 1, positive strand; lane 2, negative strand; lane 3, positive strand with mutation 19; lane 4, positive strand with mutation 14; lane 5, 10-nt marker corresponding to the 5′ end of the positive-strand oligonucleotide (+m); lane 6, positive-strand hairpin (+h). Sequences of the oligonucleotides, shown below, correspond to the positive strand and negative strand of the virus from positions 130 to 151. Oligonucleotides containing mutations in the positive-strand sequence are designated (+) strand #19 and (+) strand #14, with the numbers corresponding to the same mutations in the TGMV B replicons described in Fig. 3 and 4.

Four oligonucleotides corresponding to the loop and right side of the stem and an oligonucleotide containing the entire positive-strand hairpin (Fig. 8B) were incubated with GST-AL1 protein, and the products were resolved on 15% polyacrylamide–8 M urea gels. GST-AL1 cleaved two oligonucleotides with wild-type positive-strand sequences (Fig. 8B, lanes

1 and 6), although cleavage of the full hairpin sequence (lane 6) was reduced in comparison with cleavage of the partial hairpin sequence (lane 1). The position of the nick site was precisely mapped by comigration of the cleavage product with an oligonucleotide that corresponded to the first 10 nucleotides of the partial hairpin (Fig. 8B, cf. lanes 1 and 5). A preparation of GST-AL1 digested with thrombin to separate the GST and AL1 protein domains also specifically cleaved the positive-strand oligonucleotide (data not shown). In contrast, GST-AL1 did not cleave a negative-strand oligonucleotide (Fig. 8B, lane 2) and GST alone did not cleave the positive- or negative-strand oligonucleotides under similar conditions (data not shown). These results demonstrated that TGMV AL1 specifically cleaves the positive-strand hairpin sequence between TGMV B positions 139 and 140.

In vitro nicking assays also detected GST-AL1 cleavage of a positive-strand oligonucleotide with the 8-bp change in the stem sequence described for mutant 14 (Fig. 8B, lane 4). However, no cleavage by GST-AL1 of a positive-strand oligonucleotide containing mutation 19 (lane 3) was observed. Mutation 19 alters two nucleotides in the invariant sequence of the loop that also abolished TGMV B DNA replication in protoplasts (Fig. 4B, lane 3). These data established that AL1-mediated DNA cleavage in vitro does not require the hairpin structure or stem sequence, whereas sequences in the loop are essential for nicking.

DISCUSSION

Initiation of DNA synthesis is generally dependent on specific regions of DNA that act as origins of replication. The sequence and structural requirements necessary for origin function have been studied for a variety of plasmid, viral, and chromosomal DNAs in bacterial, fungal, and animal cells (24, 34, 59). In contrast, little is known about the *cis*-acting elements that are necessary for initiation of DNA synthesis in plant cells. This lack of information could be attributed, until recently, to the absence of an experimental system that was amenable to the study of plant DNA replication. However, recent studies of geminivirus replication have begun to provide insight into the character of DNA sequences that can function as replication origins in plant cells. In this paper, we demonstrated that DNA sequence and secondary structure contribute to geminivirus origin activity in vivo. In vitro nicking data suggested that the in vivo results reflect requirements for DNA cleavage during initiation of rolling-circle replication.

All known geminivirus genomes contain a 5′ intergenic region that includes an inverted-repeat sequence with the capacity to form a hairpin or cruciform structure (37). S1 nuclease digestion of TGMV DNA showed that the repeated sequence can form a hairpin structure in vitro (61). Other studies demonstrated that the hairpin motif is an essential element of geminivirus replication origins in vivo. Insertion of an 8-bp DNA linker into the invariant loop sequence of TGMV (47) or partial deletion of the stem sequence of SqLCV (38) abolished viral replication in leaf disc assays. In these studies, both the sequence and the structure of the hairpin motif were altered and thus their individual contributions to origin activity were not addressed. Recently, Roberts and Stanley (48) showed that ACMV B mutants with single-base substitutions in the stem sequence are infectious in whole-plant assays, whereas most mutants with changes in the loop sequence are only infectious after recombination that restores the wild-type sequence. These results confirmed the importance of the loop sequence in geminivirus replication but did not address the role of the hairpin structure, because single-base mismatches in the 11-bp stem are unlikely to significantly disrupt a hairpin structure.

In the study reported here, direct proof for the involvement of DNA structure in geminivirus replication was obtained by comparison of the effects of mutations that changed both sequence and structure with those of mutations that altered only the sequence of the hairpin motif. The replication of TGMV B mutants with 2-bp changes on the left or the right side of the stem that modified both sequence and structure was impaired, whereas compensatory double mutants that contained the same sequence changes but retained wild-type structure replicated efficiently when the viral replication proteins were provided by an AL1,2,3 plant expression cassette. These results showed that base pairing in the TGMV B hairpin is important for viral replication. However, it is probable that the 2-bp mutations only shortened or introduced bulges into the stem and did not fully disrupt the structure. Partial retention of the hairpin structure may explain the low levels of replication seen with the left- and right-side mutants. This hypothesis was confirmed by the absence of any detectable replication with two TGMV B mutants that do not have the potential to form a modified hairpin structure because of mutations at 8 of 11 positions on the left or the right side of the stem (mutants 13 and 14). In contrast, a TGMV B mutant with compensatory 8-bp mutations (mutant 13+14) in the stem replicated to near-wild-type levels in the presence of an AL1,2,3 expression cassette. These data unequivocally demonstrated that the hairpin structure is essential for geminivirus replication.

The TGMV B hairpin contains a 12-nt loop sequence that includes both variant and invariant sequences between TGMV A and B DNAs and among different geminiviruses. Mutation of two nucleotides in the invariant sequence abolished TGMV B replication (mutant 19), as was reported for ACMV (48). In contrast, mutation of the variant positions in TGMV B (mutants 14, 15, 16, and 17) demonstrated sequence and structural flexibility at these positions for replication with the AL1,2,3 expression cassette. These results were consistent with earlier experiments that showed that changing the equivalent sequences in BGMV B had no impact on replication (13). However, the mutations in the variant sequence of the loop decreased replication efficiency in the presence of the wild-type A component, suggesting that the sequence and/or structural requirements are more stringent in the presence of a competing replicon. This assay cannot distinguish between sequence and structural contributions of the loop in origin function since the structure of the loop region is unknown.

Similarly, many of the compensatory TGMV B stem mutants did not replicate to wild-type levels in the presence of a TGMV A component in whole-plant infectivity, leaf disc, and protoplast assays. This difference may have reflected lower levels of the AL1 protein as a result of autoregulation of the AL1 promoter in TGMV A (9, 62) versus constitutive expression from the cauliflower mosaic virus 35S promoter in the plant expression cassette. This hypothesis is not likely because of the strength of the AL1 promoter, the inefficiency of AL1-mediated repression in NT-1 suspension cells (9) used for the protoplast assays, and the amplification of AL1 gene copy number through TGMV A replication. In addition, TGMV replication levels in NT-1 protoplasts are comparable whether AL1 is provided from the A component or a 35S-AL1 expression cassette (62). Alternatively, there is precedence for mutant origins displaying impaired replication phenotypes in competition assays (31, 67). Further analysis showed that replication of the mutant B component was reduced in the presence of the A component alone or in the presence of the A component plus the AL1,2,3 expression cassette, demonstrating that the

mutant B origin does not replicate efficiently in the presence of a competing wild-type origin.

Independent of the reason for the effect, the presence of TGMV A revealed that the sequence of the stem, in addition to its structure, is important for efficient replication. One possibility is that the mutations reduced the stability of base pairing in the stem compared with the wild-type sequence. Generally, the stem sequences of geminivirus hairpin motifs are GC rich whereas the loop regions contain an invariant AT-rich motif. Most of the compensatory mutations involved GC-to-AT base pair changes predicted to reduce stability (29). Compensatory 2-bp mutations resulting in clustered AT base pairs in the middle of the stem were most detrimental to replication (mutants 5+6 and 11+12). The compensatory 8-bp mutations severely impacted TGMV B replication in the presence of TGMV A and reduced the predicted ΔG for the hairpin from a wild-type level of −19.7 to −13.2 kcal (−82.4 to −55.2 kJ)/mol (3). The calculations include ΔG values for RNA loops (15) and do not account for non-Watson-Crick base pairing in the loop, but these sequences should contribute equally to both structures. Alternatively, the stem may contain a recognition sequence involved in origin function. This is consistent with the observation that mutations on the right side of the hairpin were more detrimental than mutations on the left. Similar results were reported for the rolling-circle plasmid, pT181, whose hairpin structure included some nucleotides that were important for replication even when base pairing was maintained (42). Taken together, our results demonstrated that both sequence and structure of the hairpin contribute to TGMV origin activity in vivo.

The precise structural requirements for TGMV replication are not known, but the relative replication efficiencies of compensatory 2-bp mutants in the presence of TGMV A provided some insight into the constraints. Mutations that shortened the predicted stem and reduced replication (mutations 2, 3, 4, 11, and 12) suggested that a length of at least 9 to 10 bp is necessary for efficient replication. The precise limit for stem length could not be determined because mutations on opposite sides at the base of the stem did not impact replication equally. Mutations that extended the length of the stem and reduced the size of the loop (mutations 16, 17, and 18) also impacted replication. In this case, mutations may have altered the loop structure if non-Watson-Crick base pairing occurs between bases in the wild-type loop, as has been observed for RNA loop sequences (52).

The role of the hairpin structure in geminivirus replication is not known. A hairpin structure may be part of the negative-strand origin, as has been observed for a variety of bacterial rolling-circle replication systems (7, 18, 27, 55). Analysis of ACMV replication intermediates suggested that the negative-strand origin for subgroup III geminiviruses is located in the 5′ intergenic region, which includes the hairpin motif (53). However, it is unlikely that the hairpin structure is part of subgroup I negative-strand origins, which are thought to be located in the unique 3′ intergenic regions of these viruses (8). In contrast, initiation and termination sites for positive-strand DNA synthesis have been mapped to the hairpin region for members of all three geminivirus subgroups (25, 57, 58), thereby strongly suggesting that the hairpin structure is part of the positive-strand origin. There is precedence for the involvement of cruciform structures in initiation of positive-strand DNA synthesis of some bacterial plasmids (20, 69). DNA secondary structure has also been implicated in initiation of parvovirus replication, which proceeds through a related rolling-hairpin mechanism in eucaryotic cells (39).

The hairpin structure could serve related, but not identical, functions in positive-strand DNA replication in different geminivirus subgroups. High-molecular-weight genomic concatemers were observed for WDV, a subgroup I virus, when its hairpin was deleted (33). The initiation site for positive-strand synthesis in the WDV deletion mutant was mapped to a cryptic site that contained a partial copy of the hairpin motif unable to form a hairpin structure (25). Unlike the situation with WDV, we were unable to detect any high-molecular-weight intermediates for TGMV B mutants with changes on either the left or right side of the stem (data not shown). In TGMV replication, the hairpin structure appears to be required for both initiation and termination of positive-strand DNA synthesis, whereas in WDV replication, the structure may only be required for termination of positive-strand synthesis.

Initiation of positive-strand DNA synthesis requires cleavage at a specific site in the positive-strand origin. We showed that TGMV AL1 catalyzes site-specific cleavage of single-stranded oligonucleotides that corresponded to the positive-strand of the hairpin, analogous to the situation with TYLCV and WDV C1 proteins (26, 36). TGMV AL1 also cleaved a positive-strand oligonucleotide that contained an 8-bp mutation in the right stem sequence but failed to nick an oligonucleotide with a 2-bp change in the invariant loop sequence. These results showed that the sequence required for TGMV AL1 cleavage includes no more than 15 bp and established that at least one base in the loop sequence is specifically required for nicking. Similar sequence requirements have been described for DNA cleavage by the TYLCV C1 protein (36). Our in vitro cleavage results correlated well with in vivo replication results showing that the compensatory 8-bp TGMV B mutant (mutant 13+14) replicated in tobacco protoplasts whereas the 2-bp loop mutant (mutant 19) did not. However, no replication was detected for the TGMV B mutants with 8-bp changes on the left or the right sides of the stem even though both stem sequences were not required for nicking in vitro. This discrepancy may reflect a requirement for the hairpin or cruciform structure for cleavage in vivo. This hypothesis is consistent with the observation that TGMV AL1 preferentially binds single-stranded DNA from the TGMV common region (65) and does not cleave double-stranded DNA containing the hairpin motif in vitro (data not shown), similar to what is observed for TYLCV C1 (36). Together, these results strongly suggest that TGMV AL1 requires a single-stranded region for cleavage and that this region is generated in vivo by the formation of a cruciform structure in double-stranded viral DNA, the template for positive-strand DNA synthesis. There is precedence for the requirement of a single-stranded region for cleavage in rolling-circle systems (28).

We demonstrated previously that TGMV AL1 is a sequence-specific DNA-binding protein (14) that mediates positive-strand origin recognition by binding to a 13-bp repeated sequence in the 5′ end of the TGMV common region (12). In this paper, we showed that TGMV AL1 also cleaves in the hairpin to initiate positive-strand DNA synthesis. The relationship between these two events during positive-strand replication is not known. AL1-DNA binding may induce a conformational change in double-stranded viral DNA that results in extrusion and/or stabilization of the cruciform structure necessary for cleavage. Alternatively, DNA supercoiling (6, 17) or the binding of a host factor (2, 46) may facilitate the cruciform structure. These possibilities will be examined in future experiments using supercoiled substrates that include the full positive-strand origin and by the inclusion of host proteins in cleavage assays in vitro.

ACKNOWLEDGMENTS

We thank Verne Luckow (Monsanto Company) for invaluable help in setting up the baculovirus expression system. We are grateful to Jurgen Laufs and Bruno Gronenborn (CNRS, Gif sur Yvette, France) for making their TYLCV nicking results available to us prior to publication. We thank Sharon Settlage and Patricia Eagle for critical reading of the manuscript. We also thank Charles Hardin (North Carolina State University) for helpful discussions.

This research was supported by National Science Foundation grant DMB-9104150 (to L.H.-B.) and postdoctoral fellowship AI08775 (to B.M.O.) from the Public Health Service-National Institutes of Health.

REFERENCES

1. **Arguello-Astorga, G. R., R. G. Guevara-Gonzalez, L. R. Herrera-Estrella, and R. F. Rivera-Bustamante.** 1994. Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. Virology **203:**90–100.
2. **Bracale, M., M. G. Galli, C. Savini, and M. E. Bianchi.** 1994. Specific interaction of plant HMG-like proteins with cruciform DNA. J. Exp. Bot. **45:**1493–1496.
3. **Breslauer, K. J., R. Frank, H. Blocker, and L. A. Marky.** 1986. Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. USA **83:**3746–3750.
4. **Briddon, R. W., and P. G. Markham.** 1995. Geminiviridae, p. 158–165. *In* F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers (ed.), Virus taxonomy: sixth report of the International Committee on Taxonomy of Viruses. Springer-Verlag, Vienna.
5. **Brough, C. L., R. J. Hayes, A. J. Morgan, R. H. A. Coutts, and K. W. Buck.** 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.
6. **Dayn, A., S. Malkhosyan, and S. M. Mirkin.** 1992. Transcriptionally driven cruciform formation in vivo. Nucleic Acids Res. **20:**5991–5997.
7. **del Solar, G. H., A. Puyet, and M. Espinosa.** 1987. Initiation signals of the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Res. **15:**5561–5580.
8. **Donson, J., B. A. M. Morris-Krsinich, P. M. Mullineaux, M. I. Boulton, and J. W. Davies.** 1984. A putative primer for second-strand DNA synthesis of maize streak virus is virion-associated. EMBO J. **3:**3069–3073.
9. **Eagle, P. A., B. M. Orozco, and L. Hanley-Bowdoin.** 1994. A DNA sequence required for geminivirus replication also mediates transcriptional regulation. Plant Cell **6:**1157–1170.
10. **Elmer, J. S., L. Brand, G. Sunter, W. E. Gardiner, D. M. Bisaro, and S. G. Rogers.** 1988. Genetic analysis of tomato golden mosaic virus. II. Requirement for the product of the highly conserved AL1 coding sequence for replication. Nucleic Acids Res. **16:**7043–7060.
11. **Elmer, J. S., G. Sunter, W. E. Gardiner, L. Brand, C. K. Browning, D. M. Bisaro, and S. G. Rogers.** 1988. *Agrobacterium*-mediated inoculation of plants with tomato golden mosaic virus DNAs. Plant Mol. Biol. **10:**225–234.
12. **Fontes, E. P. B., P. A. Eagle, P. A. Sipe, V. A. Luckow, and L. Hanley-Bowdoin.** 1994. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. J. Biol. Chem. **269:**8459–8465.
13. **Fontes, E. P. B., H. J. Gladfelter, R. L. Schaffer, I. T. D. Petty, and L. Hanley-Bowdoin.** 1994. Geminivirus replication origins have a modular organization. Plant Cell **6:**405–416.
14. **Fontes, E. P. G., V. A. Luckow, and L. Hanley-Bowdoin.** 1992. A geminivirus replication protein is a sequence-specific DNA binding protein. Plant Cell **4:**597–608.
15. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83:**9373–9377.
16. **Gladfelter, H. J., and L. Hanley-Bowdoin.** Unpublished data.
17. **Glucksmann, M. A., P. Markiewicz, C. Malone, and L. B. Rothmandenes.** 1992. Specific sequences and a hairpin structure in the template strand are required for N4-virion RNA polymerase promoter recognition. Cell **70:**491–500.
18. **Greenbaum, J. H., and K. J. Marians.** 1985. Mutational analysis of primosome assembly sites: evidence for alternative DNA structures. J. Biol. Chem. **260:**12266–12272.
19. **Groning, B. R., R. J. Hayes, and K. W. Buck.** 1994. Simultaneous regulation of tomato golden mosaic virus coat protein and AL1 gene expression: expression of the AL4 gene may contribute to suppression of the AL1 gene. J. Gen. Virol. **75:**721–726.
20. **Gros, M. F., H. T. Riele, and S. D. Ehrlich.** 1987. Rolling circle replication of single-stranded DNA plasmid pC194. EMBO J. **6:**3863–3869.
21. **Hamilton, W. D. O., D. M. Bisaro, R. H. A. Coutts, and K. W. Buck.** 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.
22. **Hamilton, W. D. O., V. E. Stein, R. H. A. Coutts, and K. W. Buck.** 1984. Complete nucleotide sequence of the infectious cloned DNA components of tomato golden mosaic virus: potential coding regions and regulatory sequences. EMBO J. **3:**2197–2205.
23. **Hanley-Bowdoin, L., J. S. Elmer, and S. G. Rogers.** 1990. Expression of functional replication protein from tomato golden mosaic virus in transgenic tobacco plants. Proc. Natl. Acad. Sci. USA **87:**1446–1450.
24. **Held, P. G., and N. H. Heintz.** 1992. Eukaryotic replication origins. Biochim. Biophys. Acta **1130:**235–246.
25. **Heyraud, F., V. Matzeit, M. Kammann, S. Schaefer, J. Schell, and B. Gronenborn.** 1993. Identification of the initiation sequence for viral-strand DNA synthesis of wheat dwarf virus. EMBO J. **12:**4445–4452.
26. **Heyraudnitschke, F., S. Schumacher, J. Laufs, S. Schaefer, J. Schell, and B. Gronenborn.** 1995. Determination of the origin cleavage and joining domain of geminivirus rep proteins. Nucleic Acids Res. **23:**910–916.
27. **Hiasa, H., K. Tanaka, H. Sakai, K. Yoshida, Y. Honda, T. Komano, and G. N. Godson.** 1989. Distinct functional contributions of three potential secondary structures in the phage G4 origin of complementary DNA strand synthesis. Gene **84:**17–22.
28. **Higashitani, A., D. Greenstein, H. Hirokawa, S. Asano, and K. Horiuchi.** 1994. Multiple DNA conformational changes induced by an initiator protein precede the nicking reaction in a rolling circle replication origin. J. Mol. Biol. **237:**388–400.
29. **Hobza, P., and C. Sandorfy.** 1987. Nonempirical calculation on all the 29 possible DNA base pairs. Biochemistry **109:**1302–1307.
30. **Holsters, M., D. de Waele, A. Depicker, E. Messens, M. van Montagu, and J. Schell.** 1978. Transfection and transformation of Agrobacterium tumefaciens. Mol. Gen. Genet. **163:**181–187.
31. **Ishimi, Y., and K. Matsumoto.** 1993. Model system for DNA replication of a plasmid DNA containing the autonomously replicating sequence from *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **90:**5399–5403.
32. **Junghans, H., and M. Metzlaff.** 1990. A simple and rapid method for the preparation of total plant DNA. BioTechniques **3:**176.
33. **Kammann, M., H.-J. Schalk, V. Matzeit, S. Schaefer, J. Schell, and B. Gronenborn.** 1991. DNA replication of wheat dwarf virus, a geminivirus, requires two *cis*-acting signals. Virology **184:**786–790.
34. **Kornberg, A., and T. A. Baker.** 1992. DNA replication, 2nd ed. W. H. Freeman and Company, New York.
35. **Lanahan, M. B., H.-C. Yen, J. J. Giovannoni, and H. J. Klee.** 1994. The *never ripe* mutation blocks ethylene perception in tomato. Plant Cell **6:**521–530.
36. **Laufs, J., W. Traut, F. Heyraud, V. Matzeit, S. G. Rogers, J. Schell, and B. Gronenborn.** 1995. In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. Proc. Natl. Acad. Sci. USA **92:**3879–3883.
37. **Lazarowitz, S.** 1992. Geminiviruses: genome structure and gene function. Crit. Rev. Plant Sci. **11:**327–349.
38. **Lazarowitz, S. G., L. C. Wu, S. G. Rogers, and J. S. Elmer.** 1992. Sequence-specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. Plant Cell **4:**799–809.
39. **Lefebvre, R. B., S. Riva, and K. I. Berns.** 1984. Conformation takes precedence over sequence in adeno-associated virus DNA replication. Mol. Cell. Biol. **4:**1416–1419.
40. **Luckow, V. A., S. C. Lee, G. F. Barry, and P. O. Olins.** 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*. J. Virol. **67:**4566–4579.
41. **McDowell, S. W., R. H. A. Coutts, and K. W. Buck.** 1986. Molecular characterisation of subgenomic single-stranded and double-stranded DNA forms from plants infected with tomato golden mosaic virus. Nucleic Acids Res. **14:**7967–7984.
42. **Miao, D. M., Y. Honda, K. Tanaka, A. Higashi, T. Nakamura, Y. Taguchi, H. Sakai, T. Komano, and M. Bagdasarian.** 1993. A base-paired hairpin structure essential for the functional priming signal for DNA replication of the broad host range plasmid RSF1010. Nucleic Acids Res. **21:**4900–4903.
43. **Nagar, S., T. J. Pedersen, K. Carrick, L. Hanley-Bowdoin, and D. Robertson.** 1995. A geminivirus induces expression of a host DNA replication protein in terminally differentiated plant cells. Plant Cell **7:**705–719.
44. **Noirot, P., J. Batgonetti, and R. P. Novick.** 1990. Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. Proc. Natl. Acad. Sci. USA **87:**8560–8564.
45. **Padidam, M., R. N. Beachy, and C. M. Fauquet.** 1995. Classification and identification of geminiviruses using sequence comparisons. J. Gen. Virol. **76:**249–263.
46. **Pearson, C. E., M. T. Ruiz, G. B. Price, and M. Zannishadjopoulos.** 1994. Cruciform DNA binding protein in HeLa cell extraction. Biochemistry **33:**14185–14196.
47. **Revington, G. N., G. Sunter, and D. M. Bisaro.** 1989. DNA sequences essential for replication of the B genome component of tomato golden mosaic virus. Plant Cell **1:**985–992.
48. **Roberts, S., and J. Stanley.** 1994. Lethal mutations within the conserved stem-loop of African cassava mosaic virus DNA are rapidly corrected by

genomic recombination. J. Gen. Virol. **75:**3203–3209.
49. **Rogers, S. G., D. M. Bisaro, R. B. Horsch, R. T. Fraley, N. L. Hoffman, L. Brand, J. S. Elmer, and A. M. Lloyd.** 1986. Tomato golden mosaic virus A component DNA replicates autonomously in transgenic plants. Cell **45:**593–600.
50. **Rybicki, E. P.** 1994. A phylogenetic and evolutionary justification for three genera of Geminiviridae. Arch. Virol. **139:**49–77.
51. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
52. **SantaLucia, J., R. Kierzek, and D. H. Turner.** 1991. Stabilities of consecutive A · C, C · C, G · G, U · C, and U · U mismatches in RNA internal loops: evidence for stable hydrogen-bonded U · U and C · C+ pairs. Biochemistry **30:**8242–8251.
53. **Saunders, K., A. Lucy, and J. Stanley.** 1992. RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. Nucleic Acids Res. **20:**6311–6315.
54. **Schaffer, R. L.** 1992. Host range determination of geminiviruses. M.S. thesis. North Carolina State University, Raleigh.
55. **Seery, L., and K. M. Devine.** 1993. Analysis of features contributing to activity of the single-stranded origin of *Bacillus* plasmid pBAA1. J. Bacteriol. **175:**1988–1994.
56. **Smith, D. B., and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene **67:**31–40.
57. **Stanley, J.** 1995. Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. Virology **206:**707–712.
58. **Stenger, D. C., G. N. Revington, M. C. Stevenson, and D. M. Bisaro.** 1991. Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. Proc. Natl. Acad. Sci. USA **88:**8029–8033.
59. **Stillman, B.** 1994. Initiation of chromosomal DNA replication in eukaryotes—lessons from lambda. J. Biol. Chem. **269:**7047–7050.
60. **Sunter, G., and D. M. Bisaro.** 1992. Transactivation of geminivirus AR1 and geminivirus BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. Plant Cell **4:**1321–1331.
61. **Sunter, G., K. W. Buck, and R. H. A. Coutts.** 1985. S1-sensitive sites in the supercoiled double-stranded form of tomato golden mosaic virus DNA component B: identification of regions of potential secondary structure and regulatory function. Nucleic Acids Res. **13:**4645–4659.
62. **Sunter, G., M. D. Hartitz, and D. M. Bisaro.** 1993. Tomato golden mosaic virus leftward gene expression: autoregulation of geminivirus replication protein. Virology **195:**275–280.
63. **Sunter, G., M. D. Hartitz, S. G. Hormuzdi, C. L. Brough, and D. M. Bisaro.** 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.
64. **Thomashow, M. F., R. Nutter, A. L. Montoya, M. P. Gordon, and E. W. Nester.** 1980. Integration and organization of Ti plasmid sequences in crown gall tumors. Cell **19:**729–739.
65. **Timmermans, M. C. P., O. P. Das, and J. Messing.** 1994. Geminiviruses and their uses as extrachromosomal replicons. Annu. Rev. Plant Physiol. **45:**79–112.
66. **Ustav, E., M. Ustav, P. Szymanski, and A. Stenlund.** 1993. The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. Proc. Natl. Acad. Sci. USA **90:**898–902.
67. **von Arnim, A., and J. Stanley.** 1992. Determinants of tomato golden mosaic virus symptom development located on DNA B. Virology **186:**286–293.
68. **Wang, P. Z., S. J. Projan, V. Henriquez, and R. P. Novick.** 1993. Origin recognition specificity in pT181 plasmids is determined by a functionally asymmetric palindromic DNA element. EMBO J. **12:**45–52.
69. **Williams, D. L., and D. Kowalski.** 1993. Easily unwound DNA sequences and hairpin structures in the Epstein-Barr virus origin of plasmid replication. J. Virol. **67:**2707–2715.
70. **Wing, D., C. Koncz, and J. Schell.** 1989. Conserved function in *Nicotiana tabacum* of a single *Drosophila hsp70* promoter heat shock element when fused to a minimal T-DNA promoter. Mol. Gen. Genet. **219:**9–16.
71. **Zannis-Hadjopoulos, M., L. Frappier, M. Khoury, and G. B. Price.** 1988. Effect of anti-cruciform DNA monoclonal antibodies on DNA replication. EMBO J. **7:**1837–1844.