In Vivo Accumulation of a Turnip Crinkle Virus Defective Interfering RNA Is Affected by Alterations in Size and Sequence

XIAO HUA LI† and ANNE E. SIMON‡

Department of Plant Pathology† and Department of Biochemistry and Program in Molecular and Cellular Biology‡, University of Massachusetts, Amherst, Massachusetts 01003

Received 1 April 1991/Accepted 23 May 1991

Turnip crinkle virus is one of several single-stranded RNA plant viruses associated with defective interfering (DI) RNAs. A complete cDNA copy of a 344-base DI RNA (DI RNA G) was cloned downstream from a T7 RNA polymerase promoter. Transcripts synthesized in vitro were infectious when inoculated with helper virus on turnip plants. Studies of the infectivity of DI transcripts containing deletions, insertions, and single-base changes suggest that (i) in general, only the 5’ two-thirds of the molecule can tolerate mutations; (ii) between 52 and 67 bases of terminal 5’ sequence are required for infectivity; (iii) nucleotides in positions 68 to 138 are not specifically involved in RNA infectivity; (iv) DI RNA G molecules smaller than 327 bases are not amplified efficiently in plants.

Defective interfering (DI) RNAs are curtailed versions of viral genomic RNAs that have lost essential coding sequences required for viral movement, replication, maturation, or packaging and are infectious only in the presence of a nondefective helper virus (16). The formation of detectable DI RNAs in a virus population involves two major events. The first is the generation of defective molecules from a viral genome by aberrant replicative events; the viral replicase terminates synthesis of the complementary strand prematurely and then reinitiates synthesis on a second template or different portion of the same template. The second event required in DI RNA formation is amplification of the aberrant molecule; this supposes that the molecule contains cis-acting sequences required for replicase recognition as well as signals necessary for interaction with packaging factors. Since amplification of DI RNAs can interfere with the amassing of the helper virus from which they were derived, analysis of sequences required for DI RNA amplification can provide insight into regions of the viral genome necessary for replication and packaging. For example, studies on Sindbis virus DI RNAs have suggested that only 162 nucleotides at the 5’ terminus and 19 nucleotides at the 3’ terminus are specifically required for replication and packaging of the DI molecule (10). The 19 nucleotides at the 3’ terminus are highly conserved among the genomic RNAs of other viruses in the same genus (15). Thus, these nucleotides are thought to be necessary and sufficient to define the recognition site for initiation of Sindbis virus genomic negative-strand synthesis.

Besides specific sequences required for replication and packaging of DI RNAs, the size of the DI molecule has been proposed to be important for efficient amplification (8). If synthesis of complementary strands proceeds at similar rates using large or small templates in a viral system, then smaller templates should accumulate more quickly than larger templates. Studies on Sendai virus DI RNAs indicate that infection with mixed populations of DI RNAs ranging from 1,600 to 3,500 bases resulted in the selective amplification of the smallest species (18). However, DI RNAs smaller than 1,600 bases were outcompeted in mixed infections with larger DI RNAs owing to decreased efficiency of virion assembly.

At present, the importance of cis-acting signals and the size of the molecule in DI RNA amplification has been exclusively derived from studies of animal viruses. Recently, several monopartite, positive-stranded plant RNA viruses have also been found associated with DI RNAs (1, 7, 9, 11, 20; for a review, see reference 19). One of these viruses, turnip crinkle virus (TCV), is a member of the carmovirus group and has a genomic RNA of 4,051 bases (3) which is encapsidated in an icosahedral particle. TCV is associated with several different classes of small subviral RNAs including DI RNAs, satellite (sat-) RNAs, and molecules with properties of both DI and sat-RNAs (11, 22, 26). DI RNAs associated with TCV have the intriguing property of intensifying the symptoms of TCV while interfering with replication of the viral genomic RNA (11). Transcripts synthesized in vitro from full-length cDNA clones of both the TCV-B and TCV-M isolates are infectious (2, 6), and DI RNAs originating de novo or associated naturally with the TCV-B isolate of the virus have been characterized (11). A copy choice model in which the replicase targets one of three dissimilar ~20-nucleotide motifs when reinitiating replication on internal viral genomic RNA sequence has been proposed for the generation of TCV DI RNAs based on an analysis of sequences present at the right side of all junctions between discontinuous RNAs including DI RNAs, chimeric DI/sat-RNAs, and recombinant sat-RNAs (4, 26).

In this study, full-length cDNA specifying the DI RNA naturally associated with TCV-B (DI RNA G) was cloned, and transcripts generated in vitro were infectious when coinoculated with helper virus on turnip plants. Mutations generated throughout the DI RNA sequence were analyzed for their effect on accumulation of the DI RNA in vivo. Results suggest that segments at both the 5’ and 3’ ends of the DI RNA as well as the size of the molecule are critical for efficient amplification in plants.

* Corresponding author.
† Present address: Department of Biology, Texas A&M University, College Station, TX 77843-3258.
MATERIALS AND METHODS

Virus isolates and plant inoculations. Plant growth conditions and TCV strains (TCV-M and TCV-B) have been described previously (11). The helper virus used for infectivity assays (TCV-m + D) was a total RNA preparation from turnip plants previously infected with TCV-M genomic RNA and sat-RNA D (23). sat-RNA D is included in the inoculum since it is difficult to maintain the virus free of this sat-RNA (12).

Cloning a full-length cDNA of DI RNA G. Nearly full length cDNA clones of DI RNA G have been described previously (11). To obtain a complete cDNA of DI RNA G, we cloned the junction region of DI RNA G dimers (Fig. 1). An oligonucleotide complementary to positions 90 to 100 of DI RNA G (3 μg) was hybridized to approximately 5 μg of gel-purified DI RNA G dimer in 20 μl of hybridization buffer (0.4 M NaCl, 10 mM PIPES [piperazine-N,N' bis(2-ethanesulfonic acid)] pH 6.8) at 68°C for 10 min and then slowly cooled to 25°C. First-and second-strand cDNAs were synthesized as previously described (22). After treatment with Escherichia coli DNA polymerase large fragment (Klenow; Bethesda Research Laboratories), the double-stranded cDNA was ligated into Smal-digested calf intestinal phosphatase (Boehringer Mannheim)-treated pUC18 and then transformed into a decm strain of E. coli. cDNA clones extending through the 3' and 5' junction between monomeric units of DI RNA G were selected by colony hybridization (5) using as a probe an oligonucleotide complementary to the 20 nucleotides at the 3' terminus of DI RNA G and labeled with polynucleotide kinase (New England Biolabs) and [γ-32P]ATP. cDNA clones containing the junction sequence between monomeric units were sequenced by dideoxynucleotide-chain termination (T7 DNA polymerase; Pharmacia) with standard sequencing primers (Synthetic Genetics).

Full-length DI RNA G was obtained by ligating the 3' region of DI RNA G supplied by the partial monomer cDNA clone pUCM4 (11) to the 5' region of DI RNA G supplied by a partial dimer cDNA clone. The 92-base Smal-EcoRI fragment corresponding to the 5' region of DI RNA G was excised from the dimer cDNA clone and ligated to the 252-base EcoRII-EcoRI fragment from pUCM4. The resulting 344-base Smal-EcoRI fragment was then ligated to pPM2 (21) after Smal-EcoRI digestion of the vector, generating pPM2G. This construct contains an E. coli RNA polymerase promoter directly upstream of the DI cDNA insert. To later take advantage of the higher transcription capability of T7 RNA polymerase promoters, the complete DI RNA G cDNA insert was removed from pPM2G by digestion with SmalI and inserted into the in vitro transcription vector pT7E19 (17), generating pT7G. pT7E19 was first digested with SstI and then treated with Klenow and calf intestinal phosphatase, which resulted in a single G residue immediately downstream of the promoter. At all stages of the cloning procedure, proper ligation junction sequences and fragment orientations were confirmed by dideoxynucleotide-chain termination sequencing.

Insertion of an Apal site into pPM2G. pPM2G was partially digested with RsaI, and Apal linkers (GGGGGCC; New England Biolabs) were ligated onto the linear DNA. The plasmid was next digested with an excess of Apal, purified by electroelution, and then self-ligated with T4 DNA ligase (Bethesda Research Laboratories) overnight at 25°C. Following transformation, a clone harboring an Apal site at position 99 in the cDNA insert was identified by restriction mapping after Apal and EcoRI digestion and named

5' __ 3' 5' __ 3' 3' 5' 5' 3' 3'
Reverse Transcriptase

5' __ 3' 5' __ 3' 3' 5' 5' 3' 3'
RNase H DNA Polymerase

SmaI

Klenow

T4 DNA Ligase

pUCG6

pUCGM4

SmaI EcoRII

EcoRI

SmaI EcoRI

EcoRI

SmaI EcoRI

EcoRI

T4 DNA Ligase

pPM2

pPM2G

pT7E19

pT7G

FIG. 1. Cloning a full-length cDNA copy of DI RNA G. See Materials and Methods for details. Broken line denotes gel-purified DI RNA G dimer molecules. Small filled boxes represent oligonucleotide primers for reverse transcription. Shaded circles indicate the junction between monomeric units. Juxtaposition of the 3' and 5' ends of DI RNA G cDNA create an Smal site which, when digested with the enzyme, produces exact 3' and 5' ends. pUCGM4 contains an incomplete cDNA copy of DI RNA G lacking the 5'-terminal 52 nucleotides (11). Solid and open arrowheads denote E. coli RNA polymerase or T7 RNA polymerase promoters, respectively.
pPM2GA. The complete cDNA insert from pPM2GA was later excised after digestion with SmaI and ligated into pT7E19, as described above, generating pT7GA. DI RNA derived from plants inoculated with transcripts from pT7GA is referred to as DI RNA GA.

**Gapped-duplex mutagenesis.** To create a unique NcoI site in DI RNA GA, we altered two G residues at positions 212 and 213 (numbered according to DI RNA GA sequence) to CA residues by oligonucleotide mutagenesis (14). pPM2 (vector) and pPM2GA (vector plus DI cDNA insert) were digested with EcoRI and ScaI, respectively, treated with calf intestinal phosphatase, and hybridized to an oligonucleotide with the alterations described above. The gaps were filled in with Klenow and T4 DNA ligase, and clones harboring an NcoI restriction site were identified by NcoI digestion of plasmid minipreparations and confirmed by DNA sequencing. The cDNA insert was excised by SmaI digestion and inserted into pT7E19 as described above.

**Generation of mutations in the central and 3′ regions of DI RNA G.** To make deletions in the 3′ one-third of DI RNA G, pT7G was digested with SpeI or BstEII and then treated with the slow-moving form of exonuclease Bal 31 (IBI) for various times ranging from 5 to 20 min. The ends of the DNA were ligated directly, creating deletions in both directions from the respective restriction sites. Four-nucleotide insertions at the SpeI and StyI sites were generated by digesting pT7G with the respective restriction enzyme, treating with Klenow to generate flush ends, and subsequently ligating. One mutant with a 1-base insertion at the SpeI site was unexpected, probably resulting from an aberrant Klenow reaction. Deletions and insertions were all confirmed by DNA sequencing.

**Deletion mutations in the 5′ region of DI RNA G.** Stepwise deletions from the Apal site toward the 5′ end of DI RNA G, the 3′ end of DI RNA G, or spanning the Apal site were generated in pT7GA. Apal-digested pT7GA was treated with the slow-moving form of Bal 31 for 10 to 60 min at 37°C. The size of the deletions was checked by polyacrylamide gel electrophoresis after digesting one-half of the DNA from each of the time points with BstEII, which cleaves 161 bases downstream of the Apal site. The ends of the remaining plasmid from each time point were made flush with Klenow, after which an Apal linker was reinterted. The plasmids were digested with KpnI, which cleaves the DNA at a single site downstream from the cDNA insert. To construct deletions upstream of the Apal site (toward the 5′ end of DI RNA G), the large KpnI-Apal fragment containing the deletions was gel purified and ligated to the small Apal-KpnI fragment from pT7GA. To construct deletions downstream of the Apal site toward the 3′ end of DI RNA G, the small Apal-KpnI fragment harboring the deletions was gel purified and ligated to the large KpnI-Apal fragment from pT7GA. Random deletions in both directions were generated by ligating the large and small deletion-containing fragments together.

**Insertion of foreign sequences into truncated DI RNAs.** DI RNA GA cDNA clones with deletions from the Apal site toward the 5′ end of pT7GAL-31, 31-base deletion; pT7GAL-47, 47-base deletion; pT7GAL-50, 50-base deletion; pT7GAL-53, 53-base deletion) or from the Apal site toward the 3′ end (pT7GAR-32, 32-base deletion) were digested with Apal, and the ends of the DNA were made flush and dephosphorylated by Klenow and calf intestinal phosphatase treatment. Foreign sequences excised from different positions in pUC19 or pUC8 by respective restriction digestions were gel purified, treated with Klenow, and ligated into each of the linearized plasmids described above. Foreign sequences were as follows. SH was the 31-base Smal-HindIII polylinker fragment of pUC8 into which four nucleotides had been inserted at the BamHI site. EH was the 51-base EcoRI-HindIII polylinker fragment of pUC19 in which the KpnI site had been destroyed by removing four nucleotides after Klenow treatment. KH was the 39-base KpnI-HindIII polylinker fragment of pUC19. HH65 was a 68-base Klenow-treated HindIII fragment (positions 642 to 707) of pUC19. HH75 was a 78-base Klenow-treated HindIII fragment (positions 707 to 781) of pUC19. Insertions and fragment orientations were confirmed by DNA sequencing.

To replace the sequence between the two StyI sites, pT7GA was digested with StyI, the ends of the fragments were made flush with Klenow, and the larger fragment containing the entire vector sequence and DI RNA GA cDNA positions 1 to 89 and 172 to 352 (numbering according to DI RNA GA sequence) at the two ends was purified and ligated to HH75. Following transformation, a clone harboring a 78-base foreign sequence that replaced the 82-base fragment (including 8 bases of Apal linker) between the two StyI sites in pT7GA was identified by DNA sequencing.

**Combining deletions upstream and downstream from the Apal site.** pT7GAL-22 (22-base deletion from the Apal site toward the 5′ end) and pT7GAR-18 (18-base deletion from the Apal site toward the 3′ end) were digested with KpnI and Apal. The large Apal-KpnI fragment of pT7GAL-22 and the small KpnI-Apal fragment of pT7GAR-18 were purified and ligated to form an intact plasmid. Following transformation, a clone harboring a 78-base foreign sequence that replaced the 82-base fragment (including 8 bases of Apal linker) between the two StyI sites in pT7GA was identified by DNA sequencing.

**In vitro RNA synthesis.** A 5-μg sample of pT7G, pT7GA, or plasmids derived therefrom was linearized by KpnI digestion and then extracted with phenol-chloroform and precipitated with ethanol. Linear plasmids were incubated in 46 μl of reaction solution containing 0.5 mM of the four ribonucleotides, 10 mM dithiothreitol, 40 units of RNAse (Promega), and buffer conditions suggested by the manufacturer for 5 min at 37°C, and this was followed by the addition of 100 U of T7 RNA polymerase (Bethesda Research Laboratories). The reaction mixture was incubated at 37°C for 45 min and then 5 μl of 0.5 M EDTA (pH 8.0) was added. This final mixture was combined with an equal volume of infection buffer (14) to which 15 μg of TCV-m D helper virus inoculum had been added for use in infecting six turnip plants.

**RESULTS**

In vitro transcripts of pT7G, a full-length cDNA clone of DI RNA G, are infectious on turnip plants. DI RNA G is a 342- to 346-base heterogeneous species originally isolated from TCV isolate TCV-B (11). The relationship between DI RNA G and the single TCV genomic RNA is shown in Fig. 2A. The DI RNA is composed of 10 nucleotides at the 5′ end unrelated to viral genomic RNA, but identical to the 5′ terminal 10 nucleotides of sat-RNA sequence, joined to 10 nucleotides of unknown origin. This sequence is followed by 99 nucleotides found near the 5′ end of TCV (beginning with the base at TCV position 43) joined to 225 nucleotides of TCV 3′-terminal sequence including an imperfect repeat of 36 bases (4, 11).

Full-length DI RNA G cDNA was constructed as described in Materials and Methods and as diagrammed in Fig. 1 by joining bases in positions 1 to 92 obtained from a partial DI RNA G dimer cDNA to bases in positions 93 to 344 of a partial DI RNA G monomer cDNA. The sequence of the
EFFECT OF SIZE ON ACCUMULATION OF TCV DI RNAs

VOL. 65, full-length insert

G RNA and DI RNA

plasmid residue polymerase with was inoculated onto both helper

for found

species RNA occurring present removed during were we

for accumulation and base

1

43 140/3863

TCV Genomic RNA (4051 b)

43

3863

4051

DI RNA G (344 b) 100 b

B

GGGATAAAAAGGAGCTTACCGACATTCTCTTATTGACAGTCCCTTC 50

AACGGTGGCAGCACTGTCTAGCTGCGGGCATTAGACTGGAAAACTAGTGC

GGGATMAAMAGGAGGCTTACCMACCTTCTCTATTCACGATGCCTCTTC 50

TACACACATCAAACAGGCGTTCAGTGGAACTCCGTTGGACAGGAC 100

3031

TACCCCGAAGGCTTCAAAATCGACCCCCATGTCGCTTTACTTTGAAATG 150

TGTTAGAAAGGCCCCAGGTCTATTTCTGGACCTGTTAGACGAAA 200

AACGGTGGCAGCCTGTCAGTGGCGGATTAGACTGCAAGATGCTCG 250

TCTCTGCTATACCACTAAAAACCCGAGAGGCGCCTGTCGGGACCCCTTC 300

GAACCTAAAAGGAGCCCTCCCTTCCTGCGGGGGGGGGGGGGGCTCC 344

FIG. 2. (A) Diagram of the relationship between the TCV genomic RNA and DI RNA G. Similar regions have similar shading. Arrows represent a 36-base imperfect repeat in DI RNA G. The solid block represents satellite-related sequence. The hatched box contains sequence of unknown origin. Numbers refer to residue positions in the TCV genomic RNA. (B) Sequence of full-length DI RNA G cDNA. Numbers above the bases refer to positions of TCV genomic RNA-similar sequence. Arrows denote the 36-base imperfect direct repeat. The position of the 8-nucleotide Apal linker inserted in pT7GA is indicated (see text).

full-length DI RNA G cDNA is presented in Fig. 2B. The DI cDNA was inserted into p17E19, a plasmid containing a T7 RNA polymerase promoter (17) generating pT7G. Runoff transcripts synthesized in vitro from constructs derived from this plasmid were used for all the experiments in this study. All transcripts differ from natural DI RNA G by an additional G residue at the 5' end and two extra bases at the 3' end. Transcripts synthesized in vitro from pT7G were combined with helper virus (derived from the TCV-M isolate) and inoculated onto six turnip seedlings. Total leaf RNA was extracted 2 weeks later and subjected to electrophoresis on denaturing 4% acrylamide gels. Only plants inoculated with both helper virus and DI RNA transcripts accumulated an RNA species which was clearly visible in ethidium bromide-stained gels and which comigrated with natural DI RNA G (data not shown). Plants accumulating the transcript-derived DI RNA also exhibited the severe symptoms associated with naturally occurring DI RNA G (data not shown). Although we did not investigate for this study whether the few additional nucleotides found at the ends of the transcripts were removed during replication in planta, we previously found that additional nucleotides at the 3' end of transcripts for another TCV subviral RNA (sat-RNA C) were not present in the RNA which accumulated in planta (23).

Effect of mutations in DI RNA G transcripts on infectivity in vivo. To analyze domains in DI RNA G which are important for accumulation in vivo, we constructed deletions, insertions, and base changes throughout the DI RNA G cDNA insert in pT7G (Fig. 3). DI RNA G with 4-base insertions in the two StyI sites were infectious, indicating that DI RNA G can tolerate insertions in the midregion of the molecule, as can the similar-sized subviral RNA sat-RNA C (21). However, with the exception of a 1-base insertion in the SpeI site, DI RNA G transcripts harboring all other deletions or insertions in either the SpeI or BstEII sites, ranging from 4-base insertions to 17-base deletions, were not infectious on turnip plants. These results were again similar to those obtained during analysis of sat-RNA C transcripts, which share a nearly identical 3' domain with DI RNA G (21); sat-RNA C transcripts with a deletion of 16 bases at the BstEII site or an insertion of 4 bases at the SpeI site were also not infectious. Curiously, 10 plants inoculated with transcripts from pT7G with a 1-base insertion at the SpeI site accumulated wild-type levels of the DI RNA, and in every case, infected plants showed slightly more intense symptoms than plants infected with wild-type transcripts (data not shown).

To construct DI RNA G cDNA with unique restriction sites in the 3' and 5' portions of the molecule for further mutagenesis studies, an NcoI site was generated at position 212 and an Apal site was introduced at position 99. To insert the Apal site at position 99, pT7G was partially digested with Rsal, which cleaves the plasmid at any one of three positions. An 8-bp Apal linker (GGGCCCCC) was inserted into the Rsal-cleaved sites, and a clone which harbored an Apal linker at position 99 of DI RNA G cDNA was identified and named pT7GA. Plasmids having a DI RNA synthesized in vitro from pT7GA (DI RNA GA), along with the helper inoculum, developed symptoms characteristic of DI RNA G (data not shown) and accumulated levels of RNA similar to those of wild-type DI RNA G (Fig. 4B, lanes 2 and 3). The migration of DI RNA GA in denaturing polyacrylamide gels was consistent with a size increase of 8 nucleotides. The NcoI site was constructed in pT7GA by altering two G residues at positions 212 and 213 (numbered according to DI RNA GA) to CA residues by oligonucleotide-directed mutagenesis. Transcripts containing the two nucleotide changes were synthesized in vitro and inoculated with helper virus onto six plants. RNA extracted from these plants 2 weeks postinoculation showed no evidence of DI RNA accumulation, again indicating the sensitivity of DI RNA G to mutations in the vicinity of the 3' one-third of the molecule.

Deletions from the Apal site toward the 5' and 3' ends of DI RNA GA. To identify sequences required for DI RNA GA infectivity, we generated a series of deletions within the molecule and monitored the infectivity of mutants on turnip plants. Stepwise deletions toward the 3' and 5' ends from the Apal site at position 99 were generated in pT7GA with the slow-acting form of Bal 31. There are two possible reasons...
why transcripts containing the deletions are not amplified in vivo: (i) the deletions might eliminate specific cis-acting sequences required for interaction with either the viral replicase or proteins involved in packaging, or (ii) the deletions might decrease the size of the molecule such that it is no longer stable in the plant cell or able to be packaged in the icosahedral capsid. To aid in distinguishing between possible effects resulting from deletion of essential cis-acting sequences or reduction in the size of the molecule, we reinserted an ApaI linker at the deletion point to facilitate the reinsertion of random sequences. Deletions constructed from the ApaI site of pT7GA toward the 5' end are described in Fig. 4A and Table 1. Transcripts containing each of these deletions were synthesized in vitro and inoculated with helper virus onto three to six turnip plants. Total RNA was extracted from uninoculated leaves of individual plants 3 weeks later and subjected to electrophoresis on denaturing 4% acrylamide gels (Fig. 4B). DI RNA GA transcripts harboring deletions of 25 bases (positions 74 to 98) or less were amplified in plants. However, deletion of a single additional base (positions 73 to 98) abolished infectivity on all six plants tested. The migration of the truncated DI RNA GA molecules was consistent with the number of bases deleted. Different levels of accumulation reflect normal variation between individual plants and are not due to inherent differences in the accumulation of the truncated transcripts. Although several DI RNAs appear as doublet species on the gel presented in Fig. 4B, migration patterns in other experiments suggest that a single species is accumulating in each plant.

Deletions were also constructed from the ApaI site toward the 3' end of the molecule (Table 1) as well as random sizes spanning the ApaI site. DI RNA GA with deletions of bases in positions 107 to 124 (18 bases downstream of the ApaI insertion, numbered according to DI RNA GA) or less were infectious. However, DI RNA GA harboring deletions of bases 107 to 138 (32 bases downstream of the ApaI insertion) or more were not infectious. DI RNA GA with deletions of up to 21 bases surrounding the ApaI site (positions 88 to 98 and positions 107 to 116, clone pT7GA2W-21) were infectious. DI RNA GA harboring deletions of 41 bases (positions 86 to 126, pT7GA2W-33) or more were not infectious. pT7GA2W-33 differed from other deletion mutants by lacking the ApaI linker at the deletion site. Therefore, the 41 deleted bases included 33 bases of DI RNA G sequence and 8 bases of ApaI linker.

**Larger deletions generated by combining smaller deletions of infectious RNAs abolish RNA infectivity.** From the results of the deletion analysis above, it was not possible to distinguish whether transcripts with deletions exceeding 25 bases were more infectious because of the elimination of cis-acting sequences required for DI RNA accumulation or whether large deletions abolished infectivity because of size constraints. To determine possible size effects on maintenance of DI RNA GA infectivity, plasmids harboring deletions extending toward either the 5' or the 3' terminus from the ApaI site which do not singly affect normal accumulation in vivo were combined to generate a plasmid which produced transcripts with one large deletion spanning the ApaI site. As described above, RNA transcripts harboring a 22-base deletion upstream of the ApaI site (pT7GA-L22) or an 18-base deletion downstream of the ApaI site (pT7GA-L18) were infectious (Fig. 5B, lanes 2 and 3). Transcripts from a construct which combined both deletions (pT7GA(L22)-(R18)) were used to infect turnip plants along with helper virus (Fig. 5B, lane 4). No accumulation of DI RNA with a 40-base deletion was detected in infected plants, suggesting that size, or distance between two points on the molecule, is important in RNA replication, stability, packaging, or movement through the plant.

**Insertion of foreign sequences into DI RNA GA deletion mutants.** The results above strongly suggest that size constraints and not deletion of essential cis-acting signals are responsible for the lack of infectivity of some of the DI RNA GA truncated transcripts described in Table 1. To investi-
EFFECT OF SIZE ON ACCUMULATION OF TCV DI RNAs

FIG. 5. Relationship between infectivity and size of the DI RNA GA transcript. (A) Diagram of the deletions in the DI RNA GA cDNA insert. For experimental details, see the legend to Fig. 4. (B) Total RNA was extracted from uninoculated leaves 3 weeks after inoculation with helper virus and transcripts from plasmids described in panel A. The migration position of DI RNA GA is indicated.

To further investigate the importance of size versus sequence on infectivity of DI RNA GA, we inserted random fragments into the Apal site at the deletion endpoint of several constructs to increase the size of the DI back to nearly normal (DI RNA G) length (352 bases). One or two copies of a 31-base sequence (SH) derived from a modified Smal-HindIII fragment from the pUC8 polylinker were inserted into the Klenow-treated Apal site of pT7GAL-31 (31-base deletion upstream of the Apal site) or pT7GAR-32 (32-base deletion downstream of the Apal site; Fig. 6A). Transcripts from the resulting plasmids were inoculated with helper virus onto turnip plants. Two weeks postinoculation, infected plants showed the same level of symptom severity (data not shown) and accumulated sufficient DI RNA to be clearly visible on ethidium bromide-stained gels (Fig. 6B). Transcripts containing two copies of the 31-base cassette were only infectious if the inserts were arranged head to tail;
FIG. 6. Replacing the sequence deleted in an uninfected DI RNA GA transcript restores infectivity. (A) Positions of bars and numbers above bars indicate the locations and sizes of the deletions in the cDNAs. Names of the plasmids, expected size of the DI RNAs accumulating in plants, and infectivity of the transcripts are shown. The inserted sequence SH, denoted by a box, is the Smal-HindIII polylinker fragment of pUC8. Relative orientation and size of the inserted sequence are indicated. Asterisk (*) indicates that the size of the DI RNA after insertion of plasmid-derived sequence is shortened by 4 bases as a result of Klenow treatment. (B) Total RNA was extracted from un inoculated leaves 3 weeks after inoculation with helper virus and transcripts from plasmids described in panel A. The migration position of DI RNA GA is indicated.

FIG. 7. Infectivity of DI RNA GA transcripts containing deletions of 47 bases or more is not restored by sequence replacements. Bars indicate the location and extent of the deletions. Boxes denote replacement sequences. Number of bases deleted and inserted is indicated. Replacement sequences used in each construct are identified in the names of the plasmids. Orientation of the replacement sequences is shown within the boxes. The inserted sequences contain the following segments of pUC19: KH and EH are the KnlI-HindIII and EcoRI-HindIII polylinker fragments, respectively; HH65 and HH75 are the Hind fragments in positions 643 to 707 and 707 to 781 of pUC19, respectively (25). Asterisk (*) indicates that the size of the DI RNA after insertion of plasmid-derived sequence is shortened by 4 bases as a result of Klenow treatment.

head-to-head arrangement of the two inserted fragments did not restore infectivity (data not shown). Foreign sequences were also inserted into larger deletion mutants of DI RNA GA (Fig. 7). RNA species harboring foreign sequences replacing deletions of 47 bases (positions 52 to 98), 50 bases (positions 49 to 98), or 53 bases (positions 46 to 98) in DI RNA GA were not infectious. Similarly, RNA transcripts harboring a 78-base bacterial sequence (HH75) replacing an 82-base deletion (positions 90 to 171) in DI RNA GA were not infectious. These results indicate that while nucleotides in positions 68 to 138 of DI RNA GA can be replaced by foreign sequences of similar or greater length, replacement of sequences located at positions 52 to 67 and 139 to 171 (base numbering according to DI RNA GA) did not restore infectivity. Bases in positions 68 to 138 of DI RNA GA are therefore likely to be involved in maintaining the RNA molecule at the necessary size rather than having a specific cis-acting function in DI amplification in vivo.
DISCUSSION

In this report, we describe the establishment of a system for analyzing TCV DI RNA domains involved in amplification processes in vivo. We constructed a full-length cDNA copy of TCV DI RNA G downstream from a T7 RNA polymerase promoter. Transcripts containing one additional 5' residue and two extra 3' residues were infectious when inoculated with helper virus on turnip plants. To determine sequence requirements at the 5' end of the molecule, we constructed deletions from an engineered *ApaI* site at position 98 in such a way as to allow for the reinsertion of foreign sequence into the deleted region. This strategy permitted us to determine whether loss of infectivity was due to deletion of specific *cis*-acting sequences or due to a decrease in the size of the molecule. Uninfected transcripts containing deletions of bases in positions 68 to 98 or 107 to 138 could be amplified in turnip plants if the deleted sequences were replaced with bacterial plasmid sequences of similar or greater size. Replacement of bases in positions 52 to 98 did not restore infectivity, indicating that between 52 and 67 bases at the 5' end of DI RNA G are required for accumulation in vivo.

With the exception of a 1-base insertion at position 244, all mutations constructed in the vicinity of the 3' one-third of the molecule resulted in uninfected transcripts, suggesting a specific involvement of this portion of the DI RNA in either replication or encapsidation. Nucleotides in positions 264 to 292 of DI RNA G were recently proposed to participate in encapsidation: this sequence, which is highly conserved among sat-RNA C, DI RNA G, and de novo-generated DI RNAs, was protected by coat protein subunits from RNase A degradation in vitro (24). Although this 29-base sequence was not altered in the present study, it is possible that changing surrounding sequences disrupts the secondary or tertiary structure of the region, thereby affecting packaging. Studies of animal virus DI RNAs have indicated that DI RNAs which have an advantage in packaging have advantages for survival, replication, and infection (13).

DI RNA GA with a deletion of 25 bases (positions 74 to 98) was amplified, while a 26-base deletion (positions 73 to 98) eliminated transcript infectivity. Replacement of these sequences with unrelated nucleotides (in a construct with a 31-base deletion) restored infectivity to the molecule. This suggests that the size of the RNA is important in making DI RNA GA an infectious unit. Based on the results presented in this report, the minimum length of an infectious TCV DI RNA is 327 bases, while molecules of at least 379 bases (transcripts derived from p7GAL-31SH2X; Fig. 6) can be amplified in vivo. Other TCV subviral RNAs, such as the composite sat-RNA/DI RNA molecules sat-RNA C (356 bases [23]) and sat-RNA CX (420 bases [5]) as well as DI RNAs which arise de novo (DI RNA, 383 bases; also see Fig. 4 in reference 11), also have similar lengths. Presumably, nonspecific nucleotides may function in DI infectivity by holding the molecule in such a way that recognition signals for replication and encapsidation are appropriately exposed. Alternatively, these nonspecific nucleotides may also function as filler to physically keep DI molecules at certain sizes to meet possible encapsidation constraints. The importance of size effects has also been demonstrated for other DI RNAs (10). While multiple deletion analysis showed that 90% of a Sindbis virus DI RNA played no role in replication or processing, long internal deletions resulted in amplification of heterogeneous DI RNAs larger than the deleted size. Furthermore, DI transcripts containing only the minimal sequences required for biological activity (19 nucleotides at the 3' end and 162 nucleotides at the 5' end) were not infectious.

Presently, it is not known what processes associated with amplification of inoculated DI RNA transcripts are affected by the mutations described in this study which eliminate infectivity. We have recently developed a turnip protoplast system which allows synchronous replication of viral genomic and DI transcripts (27). This system will allow us to determine whether size decreases and other mutations in DI RNA G affect replication or packaging.

ACKNOWLEDGMENTS

We thank Clifford Carpenter for critical reading of the manuscript. This work was supported by National Science Foundation grants DMB 8803533 and DMB 9004665 to A.E.S., the Massachusetts Agricultural Experiment Station, and a Sigma Xi award to X.H.L.

REFERENCES

Sendai virus DI RNA size on survival: inefficient envelopment of small nucleocapsids. Virology 165:331-337.


