JOURNAL OF VIROLOGY, July 1996, p. 4370–4379 Vol. 70, No. 7
0022-538X/96/$04.00+0

Roles of the Sequence Encoding Tobacco Etch Virus Capsid Protein in Genome Amplification: Requirements for the Translation Process and a *cis*-Active Element

SUNITA MAHAJAN,[1] VALERIAN V. DOLJA,[2] AND JAMES C. CARRINGTON[1]*

Department of Biology, Texas A&M University, College Station, Texas 77843-3258,[1] and Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon 97331[2]

Received 4 December 1995/Accepted 2 April 1996

The roles of the capsid protein (CP) and the CP coding sequence of tobacco etch potyvirus (TEV) in genome amplification were analyzed. A series of frameshift-stop codon mutations that interrupted translation of the CP coding sequence at various positions were introduced into the TEV genome. A series of 3′ deletion mutants that lacked the CP coding sequence beyond each of the frameshift-stop codon mutations were also produced. In addition, a series of 5′ CP deletion mutants were generated. Amplification of genomes containing either frameshift-stop codon insertions after codons 1, 59, 103, and 138 or genomes containing the corresponding 3′ deletions of the CP coding sequence was reduced by 100- to 1,000-fold relative to that of the parental genome in inoculated protoplasts. In contrast, a mutant containing a frameshift-stop codon after CP position 189 was amplified to 27% of the level of the parental virus, but the corresponding 3′ deletion mutant lacking codons 190 to 261 was nonviable. Deletion mutants lacking CP codons 2 to 100, 2 to 150, 2 to 189, and 2 to 210 were amplified relatively efficiently in protoplasts, but a deletion mutant lacking codons 2 to 230 was nonviable. None of the amplification-defective frameshift-stop codon or deletion mutants was rescued in transgenic cells expressing TEV CP, although the transgenic CP was able to rescue intercellular movement defects of replication-competent CP mutants. Coupled with previous results, these data led to the conclusions that (i) TEV genome amplification requires translation to a position between CP codons 138 and 189 but does not require the CP product and (ii) the TEV CP coding sequence contains a *cis*-active RNA element between codons 211 and 246. The implications of these findings on mechanisms of RNA replication and genome evolution are discussed.

The potyviruses belong to the picornavirus-like supergroup of positive-stranded RNA viruses. The potyvirus genome is approximately 10 kb long, with a viral protein (VPg) attached covalently to the 5′ terminus (37). The viral genome contains a single open reading frame encoding a large polyprotein that is proteolytically processed by three virus-encoded proteinases (18). Mutational and biochemical analyses of several potyviruses, including tobacco etch virus (TEV) and tobacco vein mottling virus, revealed that most of the proteins function either directly or indirectly in genome replication (3, 24, 26, 28, 36, 44). Some proteins contribute essential enzymatic functions, whereas others appear to provide accessory, replication-enhancing functions. Only the potyvirus capsid protein (CP) has yet to be implicated in genome replication.

The TEV CP is 263 amino acid residues long and is encoded at the 3′ end of the polyprotein coding region between nucleotides 8518 and 9306 (1). Proteolytic cleavage of CP from NIb polymerase, which is positioned adjacent to the N terminus of CP within the polyprotein, is catalyzed by the NIa proteinase (11). The CP has three known functions: it encapsidates genomic RNA with a helical, flexuous, rod-shaped morphology (39); it is required for cell-to-cell and long-distance movement of the virus through plants (14, 15); and it is required for aphid transmissibility (4). The core region forms the structural domain involved in RNA encapsidation, while the N- and C-terminal domains are exposed on the virion surface (2, 14, 38). TEV mutants that contain substitutions affecting conserved residues within the core are assembly defective and unable to move between cells (14, 15). The cell-to-cell movement function of the core is a *trans*-active function, as the movement activity of core-defective mutants can be complemented by CP supplied in transgenic plants. The N- and C-terminal domains are not required for virion assembly or stability. However, mutants lacking either the N- or C-terminal domain are unable to move long distances (leaf to leaf) through plants (14, 15). Additionally, mutations affecting a conserved motif within the N-terminal domain possess defects in aphid transmissibility (4–6).

Although data on the functions of most potyvirus-encoded proteins have accumulated, little is known about the *cis*-acting RNA sequences required for genome replication, encapsidation, and intercellular movement. Presumably, sequences near the 5′ and 3′ termini of positive- and/or negative-sense RNAs provide essential recognition signals for initiation of synthesis of both strands, but these have yet to be characterized. During the course of a mutational analysis to dissect activities of the TEV CP, two distinct functions of the CP coding sequence in genome amplification were identified. Despite the location of the CP coding sequence at the 3′ end of the TEV open reading frame and the noninvolvement of CP in RNA replication, efficient genome amplification was found to require translation through at least one-half of the CP coding region. In addition, a *cis*-active RNA sequence essential for genome amplification was identified within a 105-base region near the 3′ end of the CP coding sequence.

MATERIALS AND METHODS

Mutagenesis of the TEV CP coding sequence. All mutations were introduced into the TEV CP coding sequence within plasmids based on pTL7SN-SP, which contains cDNA representing TEV nucleotides 7166 to 9495, including the 3′ poly(A) tail. This plasmid was constructed by transfer of the *Sal*I-*Pvu*II fragment from pTEV7DA (16) into vector pTL7SN (34). To generate pTL7SN-SP3′Stu, a

* Corresponding author. Fax: 409-845-2891. Electronic mail address: carrington@bio.tamu.edu.

*Stu*I restriction site (AGGCCT) was introduced by site-directed mutagenesis (27) into pTL7SN-SP. This was accomplished by changing nucleotides 9296 and 9297 within CP codons 260 and 261 from G residues to C residues, resulting in a *Stu*I site in the context TTA GGC CTC (the *Stu*I site is underlined; the mutagenized positions are in boldface type). These mutations resulted in a conservative amino acid change of Leu for Val at CP residue 261.

Five frameshift-stop codon insertion mutations were introduced into pTL7SN-SP3′Stu after CP codons 1, 59, 103, 138, and 189. Each mutation resulted in the deletion of one nucleotide, the insertion of a stop codon (TGA), and the insertion of a *Stu*I restriction site (AGGCCT). The resulting mutations were designated 1S, 59S, 103S, 138S, and 189S. Deletion mutations that eliminated the nontranslated CP sequences beyond the frameshifted sites were generated by removal of the *Stu*I-*Stu*I fragment in each of the frameshift-stop codon mutant plasmids. These mutations were designated 1D, 59D, 103D, 138D, and 189D.

A series of deletion mutations that removed various amounts of the 5′ coding sequence of CP in pTL7SN-SP were prepared by loop-out mutagenesis (27). The deletions began at CP codon 2 and extended to codons 100, 150, 189, 210, and 230. The resulting mutations were designated ΔN100, ΔN150, ΔN189, ΔN210, and ΔN230, respectively.

Each of the mutated CP sequences was inserted into pTEV7DAN-GUS (16) by replacement of the homologous *Sal*I-*Bgl*II restriction fragment consisting of the sequence between nucleotide 7166 and the 3′ end of the poly(A) tail. Plasmid pTEV7DAN-GUS contains the full-length TEV cDNA and the coding sequence for β-glucuronidase (GUS) fused in frame between the P1 and HC-Pro regions. Virus derived by inoculation of transcripts from pTEV7DAN-GUS was termed TEV-GUS. The presence of each mutation in the pTEV7DAN-GUS-derivatives was confirmed by nucleotide sequence analysis.

Several wild-type sequence reversion mutants were produced. In each case, the mutant CP sequence was subcloned from the full-length pTEV7DAN-GUS derivative into pTL7SN-SP3′Stu by using *Sal*I and *Bgl*II sites, and site-directed mutagenesis was conducted to restore the wild-type sequence. The resulting wild-type sequence was subcloned by inserting *Bgl*II and *Sal*I sites back into the pTEV7DAN-GUS vector from which the original mutant fragment was excised. Each reversion mutant was designated with the suffix R after the original mutant code. The 189AB revertant was generated by subcloning the *Stu*I-*Stu*I fragment from pTL7SN-SP189S into pTEV7DAN-GUS189D. In effect, this reverted the 189D mutant back to the 189S mutant.

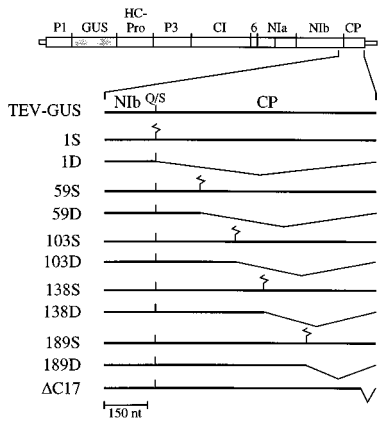
In vitro transcription and translation. Transcripts were produced from *Bgl*II-linearized plasmids by using SP6 RNA polymerase as previously described (16). Transcripts to be used for infectivity assays were capped with m7GpppG, whereas transcripts used in cell-free translation experiments were noncapped. In vitro translation was done with commercially prepared rabbit reticulocyte lysate or wheat germ extract (Promega) in the presence of [^{35}S]methionine (NEN-Du Pont). Some translation products were reacted with affinity-purified NIa proteinase as previously described (24, 35, 43). Radiolabeled reaction products were resolved by discontinuous gel electrophoresis and detected by autoradiography.

Inoculation of protoplasts and plants. Three types of *Nicotiana tabacum* were used in these experiments—nontransgenic Xanthi nc and Burley 49 cultivars and transgenic Burley 49 expressing CP. The transgenic line (FL3.3) was characterized previously (29). For inoculation of protoplasts from greenhouse-grown plants, synthetic transcripts from pTEV7DAN-GUS-derived plasmids were concentrated by precipitation with an equal volume of 4 M LiCl and resuspended in one-fifth of the original volume of deionized water. Approximately 10 μg of transcripts was used to inoculate 7.5×10^5 protoplasts by the polyethylene glycol-mediated procedure (32). Protoplasts were maintained at room temperature under continuous fluorescent light. For inoculation of plants, 10 μl of the original transcription mixture was applied manually to one or two leaves that had been dusted with carborundum. Inoculated plants were maintained on a 16-h light–8-h dark cycle at 21°C in a Conviron growth chamber.

Immunoblot analysis. Leaf tissue (25 mg) from plants inoculated with parental and ΔN150 mutant transcripts was ground in 8 volumes of GUS lysis buffer (40 mM sodium phosphate, 10 mM EDTA, 0.1% Triton X-100, 0.1% sodium lauryl sarcosine, 0.075% β-mercaptoethanol, pH 7.0) and clarified by centrifugation at $13,000 \times g$ for 10 min. An aliquot from this extract was used to measure protein concentration (10) and GUS activity (23). The remainder of the extract was diluted with an equal volume of protein dissociation buffer (0.625 M Tris-HCl, 2% sodium dodecyl sulfate [SDS], 10% 2-mercaptoethanol, 10% glycerol, pH 6.8) and subjected to SDS-polyacrylamide gel electrophoresis (PAGE), followed by immunoblot analysis with capsid and NIb monoclonal antibodies (19, 40).

GUS activity assays. Aliquots of inoculated protoplasts (2.5×10^5) were harvested at 24, 48, and 72 h postinfection (p.i.) by brief centrifugation, resuspended in 100 μl of GUS lysis buffer, and subjected to one freeze-thaw cycle. The lysate was subjected to centrifugation at $13,000 \times g$ for 10 min, and GUS activity in the supernatant was measured. Most leaf extracts were obtained by grinding tissue in 5 volumes of GUS lysis buffer and clarification by centrifugation. Fluorometric GUS assays using 4-methylumbelliferyl glucuronide as the substrate were conducted as previously described (23). To visualize GUS activity in infected leaves, the colorimetric substrate 5-bromo-4-chloro-3-indolyl-β-D-glucuronide was allowed to infiltrate the leaves as in previous studies (16).

RNA secondary-structure predictions. Secondary-structure predictions for the RNA sequence between TEV nucleotides 9145 and 9495 were made with the program mFOLD (22, 46).



Amplification in Protoplasts (% of TEV-GUS)	Xanthi nc	B49	FL3.3
TEV-GUS	100	100	100
1S	0.4 ± 0.1	0.5 ± 0.1	0.1 ± 0.1
1D	nd	0.1 ± 0.1	0.0 ± 0
59S	0.1 ± 0	0.4 ± 0.1	0.1 ± 0
59D	0.1 ± 0.1	0.0 ± 0	0.0 ± 0
103S	0.1 ± 0	0.2 ± 0.1	0.1 ± 0
103D	0.0 ± 0	0.0 ± 0	0.0 ± 0
138S	0.5 ± 0.2	0.7 ± 0.1	0.5 ± 0.2
138D	0.0 ± 0	0.1 ± 0	0.0 ± 0
189S	27.3 ± 6.4	26.1 ± 5.8	28.8 ± 1.4
189D	0.0 ± 0	0.2 ± 0.1	0.0 ± 0
ΔC17	83.2 ± 13	nd	nd

FIG. 1. Diagrammatic representation of TEV-GUS, frameshift-stop codon and 3′ deletion mutants, and relative genome amplification of mutants in protoplasts. The TEV-GUS genome, with protein coding sequences indicated above, is at the top left. The region corresponding to sequences encoding CP, part of NIb, and the proteolytic cleavage site (Q/S) is enlarged. Positions at which frameshift-stop codon mutations were introduced are indicated by zigzag vertical lines. Bent horizontal lines represent sequences that were removed in the 3′ deletion mutants. Mean relative genome amplification values (± the standard deviations) at 48 h p.i. from at least three contemporaneous inoculations of protoplasts are shown at the right for each mutant. Protoplasts were derived from the nontransgenic *N. tabacum* cv. Xanthi nc and Burley 49 (B49) and transgenic Burley 49 line FL3.3. Amplification values were scaled relative to that of TEV-GUS (100%). nt, nucleotides; nd, not done.

RESULTS

Frameshift-stop codon and 3′ deletion mutants. The TEV CP coding sequence is composed of 263 codons. Two series of mutations that each resulted in production of CPs with truncated C termini were generated. Frameshift-stop codon mutations (designated by the last codon translated and an S suffix) were introduced at five positions. Deletion mutations (designated by the last codon translated and a D suffix) that lacked the CP sequence after each of the corresponding frameshift mutations were also prepared (Fig. 1). Pairs of mutants containing the corresponding frameshift and deletion mutations encoded a similar truncated CPs, although the deletion mutant CPs each contained at their C termini residues 262 (Arg) and 263 (Gln), which were lacking in the frameshift mutant CPs.

The mutations were introduced first into an intermediate plasmid that encoded a polyprotein consisting of most of NIb and CP (Fig. 2A). To verify that the frameshift and deletion mutant pairs encoded similar-size polyproteins, transcripts produced from the intermediate plasmids were translated in vitro. Polyproteins encoded by the frameshift and deletion mutants were smaller than the polyprotein encoded by the parental plasmid (Fig. 2B). Each of the frameshift and deletion mutant pairs (e.g., 103S and 103D) encoded truncated polyproteins with the same relative electrophoretic mobility. To confirm that the NIb-CP cleavage site was active in each mutant polyprotein, the translation products were subjected to treatment with purified NIa proteinase. Processing was predicted to yield a partial NIb product of the same size for parental and mutant polyproteins but truncated CPs of various sizes for the different mutant pairs. All of the polyproteins were susceptible to NIa-mediated proteolysis, yielding the constant-size NIb product and the different-size CPs (Fig. 2C). The relative mobility of the truncated CP increased as the size of the deletion increased. The truncated CPs derived from the corresponding frameshift and deletion mutants were similar in electrophoretic mobility, although the 189S CP migrated slightly

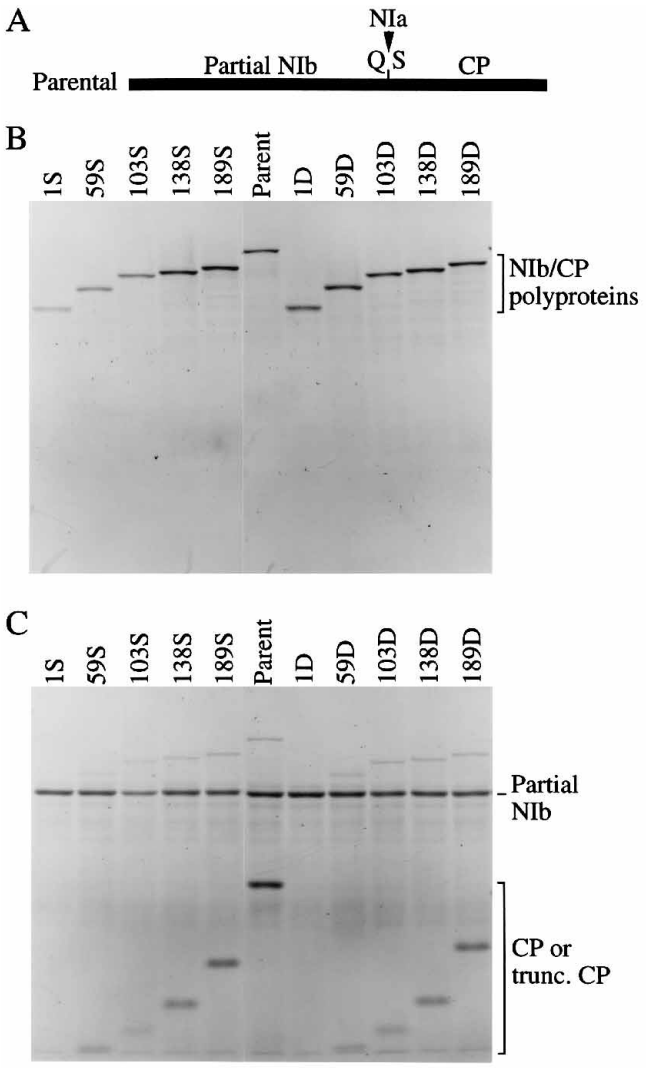


FIG. 2. In vitro synthesis and processing of polyproteins encoded by frameshift-stop codon and 3′ deletion mutants. (A) Diagrammatic representation of the partial NIb-CP polyprotein encoded by parental intermediate plasmid pTL7SN-SP3′Stu, including the position of the NIa-mediated cleavage site. (B) Transcripts encoding parental and mutant partial NIb-CP polyproteins, as indicated above each lane, were translated in a wheat germ system containing [^{35}S]methionine. The nonprocessed, radiolabeled polyproteins were detected by autoradiography after SDS-PAGE. (C) Radiolabeled, in vitro-synthesized polyproteins were treated with purified NIa proteinase and subjected to SDS-PAGE and autoradiography. trunc., truncated.

faster than the 189D CP. This difference might have been due to the two C-terminal residues in the 189D CP that were lacking in the 189S CP.

The CP mutations were transferred to a full-length TEV cDNA clone that contained the sequence for the reporter protein, GUS. Transcripts were prepared in vitro and used to inoculate *N. tabacum* cv. Xanthi nc and Burley 49 protoplasts. Genome amplification was assayed indirectly by measuring GUS activity at 24, 48, and 72 h p.i. Distinction between GUS activity due to the translation of input transcripts and activity due to amplified RNA was made possible by inoculating protoplasts in parallel with transcripts from replication-defective mutant TEV-GUS/VNN. This mutant contains substitutions resulting in a change of the conserved Gly-347–Asp-348–Asp-349 motif in NIb polymerase to Val-Asn-Asn (12, 28). For comparison, the 100% amplification level in each experiment was based on the mean GUS activity value from parental TEV-GUS-inoculated cells at 48 h p.i. Inoculation with each CP mutant was repeated at least four times in parallel with parental TEV-GUS and the VNN mutant.

The 1D, 59D, 103D, 138D, and 189D mutants were each nonviable (Fig. 1). However, a deletion mutant (ΔC17) lacking the sequence between codons 246 and 262 of CP was shown previously to amplify as efficiently as parental TEV-GUS (Fig. 1) (15). As the 189D deletion lacked the sequence between codons 190 and 261, it was concluded that at least part of the sequence between codons 190 and 245 is necessary for genome amplification. This could reflect either a requirement for the genomic RNA in this region, a requirement for functional CP, or the possibility that deletions spanning this region induce genome instability. To determine if the RNA sequence per se is required, the corresponding frameshift-stop codon mutant series was analyzed in Xanthi nc and Burley 49 protoplasts. Surprisingly, the 1S, 59S, 103S, and 138S mutants were amplified poorly, with GUS activity reaching only 0.1 to 0.5% of the level of activity in TEV-GUS-infected cells (Fig. 1 and 3A). The 189S frameshift mutant, on the other hand, was amplified to 27% of the level of the parental virus. These data suggest that relatively efficient genome amplification requires translation of the CP coding sequence to a position between codons 138 and 189 and that neither translation nor amino acid residues beyond position 189 are critical for amplification. They also suggest that the amplification defect of the 189D mutant was likely the result of disruption of an RNA sequence or structure beyond codon 189.

To be certain that the amplification-defective phenotypes summarized in Fig. 1 were due only to the specific CP mutations introduced, the wild-type sequence was restored by site-directed mutagenesis in several of the frameshift-stop codon mutants, including 59S, 103S, and 138S. The modifications yielded the 59R, 103R, and 138R restoration mutants. In each case, restoration of the wild-type sequence restored efficient genome amplification (Fig. 3A and data not shown). Additionally, the 189D deletion mutant was modified by reintroduction of the deleted sequence between codons 189 and 261, generating the 189AB mutant. In this case, the frameshifted sequence from the 189S mutant was transferred to the 189D plasmid. Viability was restored by addition of this sequence (Fig. 3B).

CP 5′ deletion mutants. To further investigate the possible requirements for CP, for translation per se to a position between codons 138 and 189, and for an amplification-enhancing region of RNA near the 3′ end of the coding sequence, a series of 5′ deletion mutants was constructed. Sequences between codons 2 and 100 (mutation termed ΔN100), 2 and 150 (ΔN150), 2 and 189 (ΔN189), 2 and 210 (ΔN210), and 2 and 230 (ΔN230) were deleted in the same intermediate plasmid that was used for the frameshift and 3′ deletion mutants (Fig. 4). Transcripts were prepared from each mutagenized plasmid and translated in vitro. As anticipated, the progressive deletions resulted in translation products with progressively faster electrophoretic mobility (Fig. 5A). The susceptibility of each NIb-CP polyprotein to cleavage by NIa proteinase was also tested. The parental polyprotein and the ΔN189, ΔN210, and ΔN230 mutant polyproteins were each processed nearly to completion, yielding the partial NIb and CP products (Fig. 5A). Because of their small sizes, the truncated CPs from these mutant polyproteins were not resolved on the gel. The ΔN100 mutant polyprotein was processed partially, while the ΔN150 polyprotein was resistant to proteolysis. The deletions in the ΔN100 and ΔN150 mutants resulted in the placement of Asn

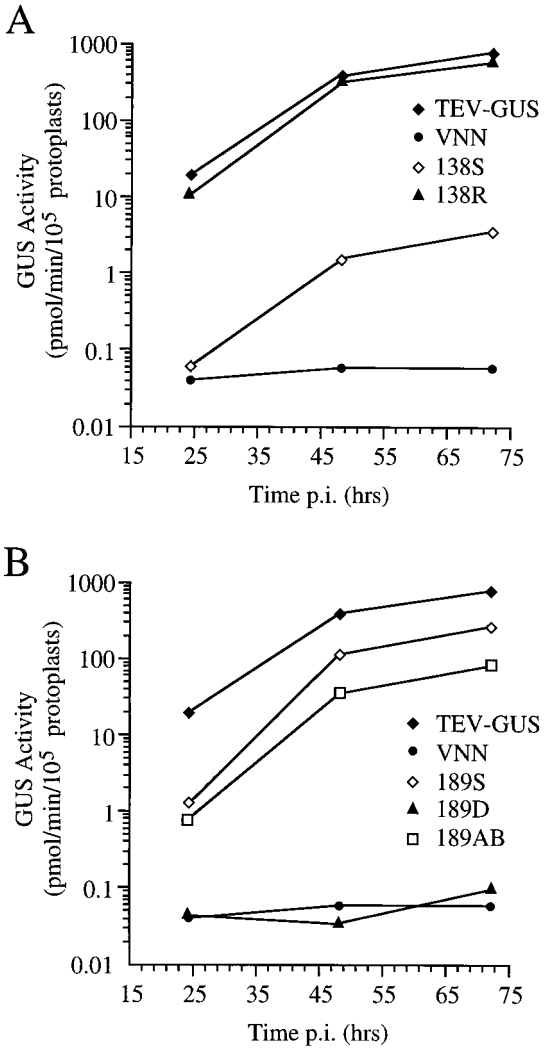


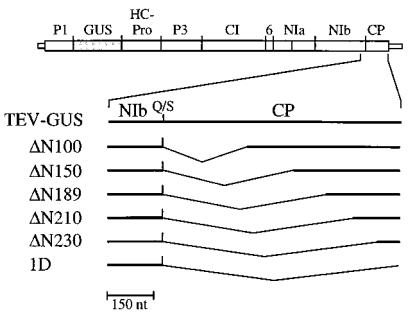
FIG. 3. Amplification of parental and selected mutant TEV-GUS genomes in protoplasts. (A and B) Plots of mean GUS activity values for two contemporaneous infections with the TEV-GUS derivatives indicated at 24, 48, and 72 h p.i. The VNN mutant contains a debilitating mutation in the NIb sequence and served as a replication-defective control. The 138R virus was derived by site-directed reversion of the 138S mutation back to the wild-type sequence.

and Pro residues at the P2′ position of the NIb-CP cleavage site, respectively. Reaction time course analysis by Dougherty et al. (17) revealed that Asn at this position has no effect on cleavage site utilization by NIa, although Pro at this position has a significant inhibitory effect. The inhibited cleavage of the ΔN100 and ΔN150 polyproteins in vitro may have been due to poor presentation of the cleavage site specifically because of the P2′ residue or an effect of the novel CP sequence adjacent to the cleavage site.

Amplification of each of the 5′ deletion mutants was tested in parallel with that of parental TEV-GUS, the VNN mutant, and the 1D deletion mutant in Burley 49 protoplasts. The ΔN100, ΔN150, ΔN189, and ΔN210 mutants were amplified to levels between 10 and 90% of the level of parental TEV-GUS (Fig. 4). Interestingly, there was no correlation between the efficiency of NIa-mediated processing in vitro and the efficiency of genome amplification. The ΔN150 mutant was amplified to levels similar to those of the ΔN189 and ΔN210 mutants and higher than that of the ΔN100 mutant, even though the ΔN150 NIb-CP cleavage site was processed least efficiently in vitro. (Processing of the ΔN150 NIb/CP cleavage site in vivo is addressed below.) Despite having had the shortest sequence deleted, the ΔN100 mutant was amplified to a lesser extent than were the ΔN150, ΔN189, and ΔN210 mutants. The ΔN230 and 1D mutants were amplified to 0.1% of the level of the parental virus and were indistinguishable from the VNN mutant control. These results strongly imply that neither the N-terminal 210 residues of CP nor the RNA sequence encoding the first 210 residues is necessary for relatively efficient genome amplification. They also suggest that genome amplification requires a sequence between CP codons 211 and 230.

Cell-to-cell movement defects, but not genome amplification defects, can be complemented by CP in *trans*. To determine whether the RNA amplification defects of the frameshift-stop codon and 3′ and 5′ deletion mutants could be complemented by CP in *trans*, protoplasts from nontransgenic Burley 49 and CP-expressing transgenic Burley 49 line FL3.3 were inoculated with transcripts from each of the mutants. The FL3.3 plants had been shown previously to express a functional CP that could complement cell-to-cell and long-distance movement defects of mutants with modifications in CP and that could incorporate into virions of some mutants (14, 15). The RNA amplification activities for all of the frameshift and deletion mutants were virtually identical in Burley 49 and FL3.3 protoplasts (Fig. 1 and 4). No significant stimulation or inhibition was detected in FL3.3 cells.

To determine if cell-to-cell movement activities of certain frameshift or deletion mutants could be rescued by the transgenic CP, Burley 49 and FL3.3 plants were inoculated with 189S and ΔN150 mutant transcripts, as well as with TEV-GUS transcripts as a control. Both the 189S and ΔN150 mutants encoded CPs that lacked part of the core domain, which was shown previously to be necessary for cell-to-cell movement (14). Inoculated plants were infiltrated by the GUS colorimetric substrate 5-bromo-4-chloro-3-indolyl-β-D-glucuronide at 3 days p.i., and the sizes of initial infection foci (expressed as the number of epidermal cells across the focus diameter) were measured microscopically. In Burley 49 plants, parental TEV-GUS formed infection foci with a mean diameter of 7.8 cells while the 189S and ΔN150 mutants were restricted to single cells (Table 1 and Fig. 6A, C, and E). In FL3.3 plants, the sizes of the parental TEV-GUS foci were similar to the sizes measured in Burley 49 (Table 1 and Fig. 6B). In contrast, the



	Amplification in Protoplasts (% of TEV-GUS)	
	B49	FL3.3
TEV-GUS	100	100
ΔN100	10.3 ± 4.8	10.9 ± 2.5
ΔN150	57.3 ± 25.7	25.0 ± 7.2
ΔN189	89.6 ± 44	103.8 ± 37.7
ΔN210	51.3 ± 11.5	44.1 ± 7.8
ΔN230	0.1 ± 0	0.0 ± 0
1D	0.1 ± 0	0.0 ± 0

FIG. 4. Diagrammatic representation of TEV-GUS and 5′ deletion mutants and relative genome amplification of mutants in protoplasts. The format for the diagrams at the left is identical to that of Fig. 1. Mean relative genome amplification values (± the standard deviations) at 48 h p.i. from at least three contemporaneous inoculations of protoplasts are shown at the right for each mutant. Protoplasts were derived from nontransgenic *N. tabacum* cv. Burley 49 (B49) and transgenic Burley 49 line FL3.3. Amplification values were scaled relative to that of TEV-GUS (100%). nt, nucleotides.

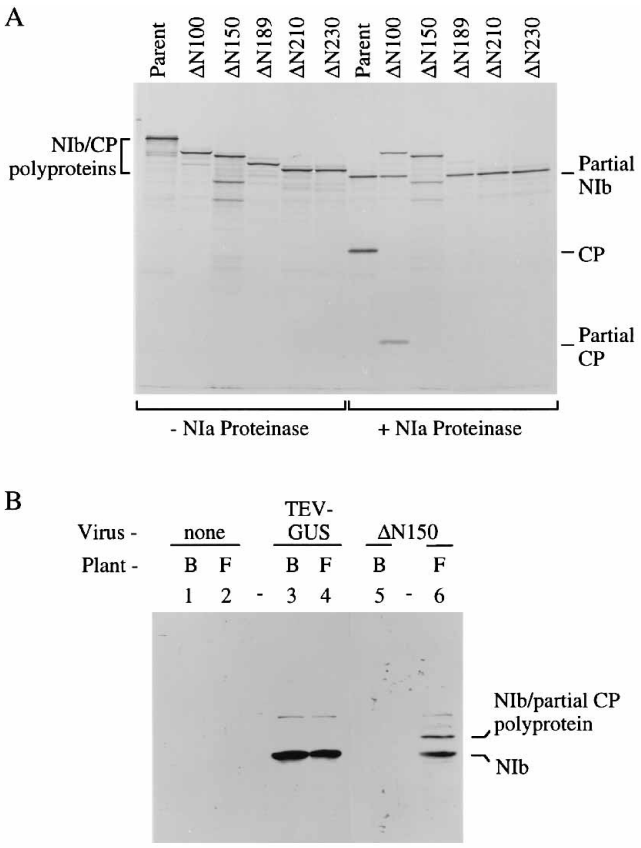


FIG. 5. In vitro synthesis and processing of polyproteins encoded by 5′ deletion mutants. (A) Transcripts encoding parental and mutant NIb-partial CP polyproteins, as indicated above the lanes, were translated in the rabbit reticulocyte lysate system containing [^{35}S]methionine. The nonprocessed (− NIa Proteinase) and processed (+ NIa Proteinase) proteins were detected by autoradiography after SDS-PAGE. (B) Immunoblot analysis of extracts from plants infected with parental TEV-GUS or the ΔN150 mutant. Total SDS-solubilized proteins from noninfected (lanes 1 and 2), parental TEV-GUS-infected (lanes 3 and 4), and mutant ΔN150-infected (lanes 5 and 6) plants were resolved by SDS-PAGE and subjected to immunoblot analysis with anti-NIb monoclonal antibodies. Samples were obtained from nontransgenic *N. tabacum* cv. Burley 49 (B lanes) or transgenic FL3.3 (F lanes) plants. The electrophoretic positions of the NIb-partial CP polyprotein and NIb are indicated at the right.

cell-to-cell movement activities of the mutants were restored partially for 189S and completely for ΔN150 (Table 1 and Fig. 6D and F). Long-distance movement of the mutants in Burley 49 and FL3.3 plants was also analyzed by measuring GUS activity in leaves two nodes (+2) above the inoculated leaves at 14 days p.i. In Burley 49 plants, activity was detected in +2 leaves infected by TEV-GUS only (Table 1). In FL3.3 plants, comparable levels of GUS activity were detected in +2 leaves infected by TEV-GUS and the ΔN150 mutant, but no activity was detected in plants inoculated with the 189S mutant. In situ analysis of FL3.3 leaves inoculated with the 189S mutant at 19 days p.i. revealed that unlike parental TEV-GUS, no secondary movement or vasculature-associated spread was evident (data not shown). These data indicate that the cell-to-cell and long-distance movement functions of the ΔN150 mutant were complemented efficiently by CP in *trans* but that the movement defects of the 189S mutant were complemented to only a limited extent.

Considering that the ΔN150 NIb-CP cleavage site was refractory to NIa-mediated proteolysis in vitro, the relative steady-state levels of NIb and the NIb-CP polyprotein in vivo were ascertained by immunoblot analysis of extracts from +2 leaves of FL3.3 plants infected by parental TEV-GUS and the ΔN150 mutant. Monoclonal antibodies specific for NIb failed to react with proteins from noninfected plants (Fig. 5B, lanes 1 and 2). Mature NIb was the predominant immunoreactive protein in TEV-GUS-infected Burley 49 and FL3.3 plants (lanes 3 and 4). No NIb was detected in +2 leaves of ΔN150 mutant-inoculated Burley 49 plants because of the cell-to-cell movement restriction. However, two predominant forms of NIb were detected in FL3.3 plants infected by the ΔN150 mutant. One migrated to a position identical to that of mature, TEV-GUS-derived NIb. The other migrated slightly more slowly and to a position consistent with a NIb-truncated CP polyprotein. In contrast to the in vitro processing experiments, we conclude that a limited amount of mature NIb polymerase was generated by inefficient NIa-mediated proteolysis in ΔN150 mutant-infected cells. The limited amount of NIb, however, was sufficient to support reasonably efficient genome amplification.

DISCUSSION

Roles for the TEV CP coding sequence in genome amplification were identified through analysis of the effects of deletion and frameshift-stop codon mutations. The CP coding sequence appears to stimulate genome amplification through two distinct mechanisms: the process of translation to a position between codons 138 and 189 is necessary, and the CP coding sequence contains one or more *cis*-active signals between codons 211 and 246 (Fig. 7).

Requirement for translation but not the translation product. The 1S, 59S, 103S, and 138S frameshift-stop codon mutants were each amplified to levels that were 100- to 1,000-fold lower than that of parental TEV-GUS. These mutants encoded truncated CPs but contained all of the CP coding sequence. We argue that the debilitating effects of premature translational termination are due to a requirement for the translation process to a point between codons 138 and 189 rather than to a requirement for a *trans*-active function of the CP product. First, the results of the 5′ and 3′ deletion series reveal that most of the CP is dispensable. Deletion of the CP RNA sequence between codons 2 and 210 (ΔN210 mutant) had little effect on genome amplification. This indicates not only that the CP encoded by this region is dispensable for amplification but that the RNA sequence between codons 2 and 210 is not necessary. This also indicates that the debilitating effects of the 1S, 59S, 103S, and 138S mutations were due to reasons other than disruption of essential *cis*-active control elements required for amplification. The 189S frameshift-stop codon mutant, which encoded a CP lacking residues 190 to

TABLE 1. Cell-to-cell and long-distance movement of TEV-GUS and CP mutants

Virus	Cell-to-cell movement (focus diam in no. of epidermal cells)[a]		Long-distance movement (GUS activity[b] in +2 systemic leaves)	
	B49	FL3.3	B49	FL3.3
TEV-GUS	7.8 ± 3.3	8.8 ± 4.1	11,625 ± 8,079	17,423 ± 11,947
189S	1.2 ± 0.4	3.6 ± 1.5	0.0 ± 0	0.0 ± 0
ΔN150	1.0 ± 0	9.3 ± 3.1	0.0 ± 0	21,146 ± 5504

[a] Infection focus diameter was determined by using inoculated leaves of nontransgenic (B49) or CP-expressing transgenic (FL3.3) plants.

[b] GUS activity (in picomoles per minute per microgram of protein) was measured by using extracts from +2 leaves at 14 days p.i.

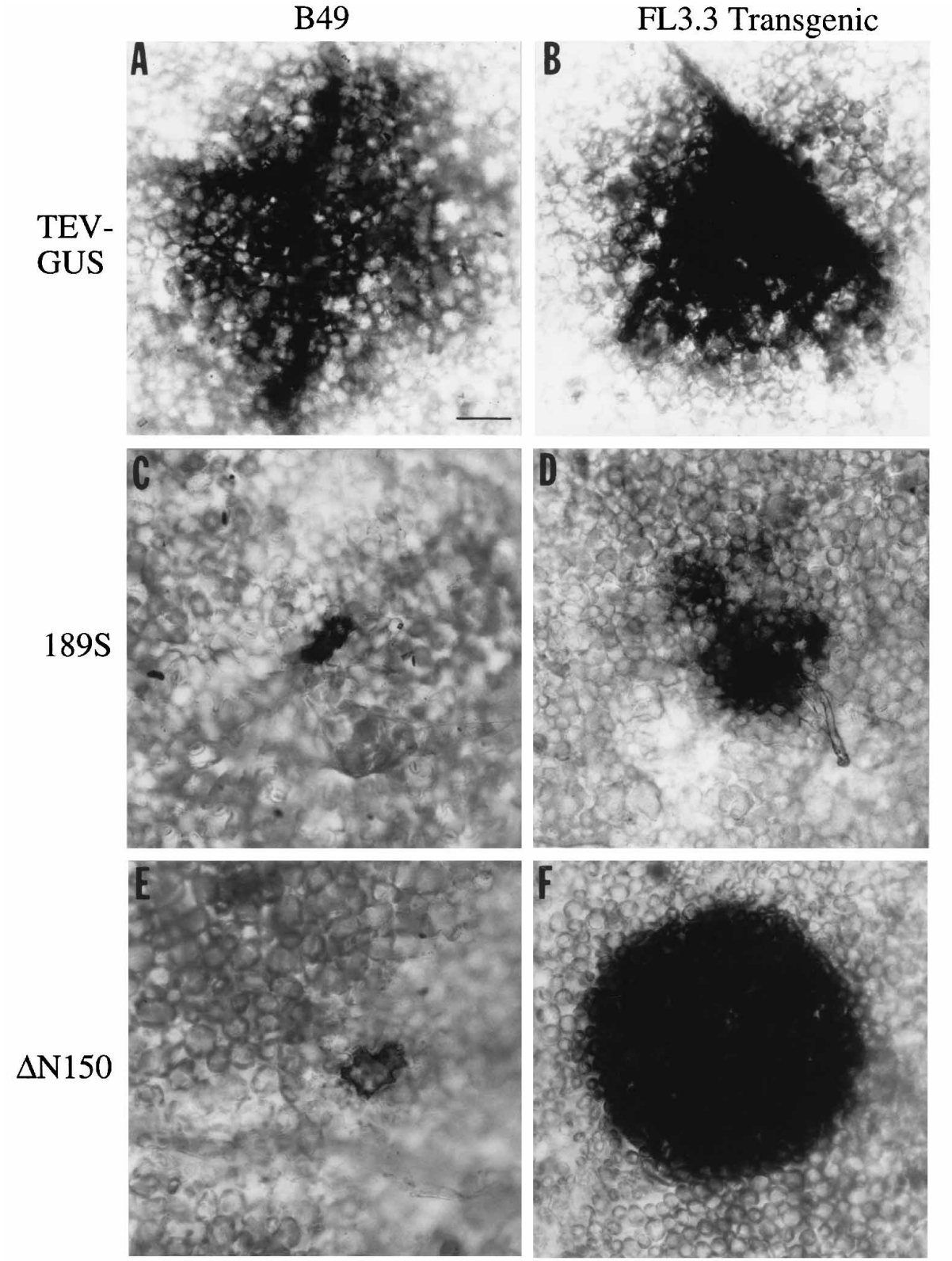


FIG. 6. In situ localization of GUS activity in leaves of plants inoculated with parental TEV-GUS (A and B), mutant 189S (C and D), or mutant ΔN150 (E and F) at 3 days p.i. Infections were done with two plant lines: nontransgenic *N. tabacum* cv. Burley 49 (B49) (A, C, and E) and transgenic FL3.3 (B, D, and F). GUS activity (dark areas) was visualized by infiltration of tissue by the colorimetric substrate 5-bromo-4-chloro-3-indolyl-β-D-glucuronide. Bar, 200 μm.

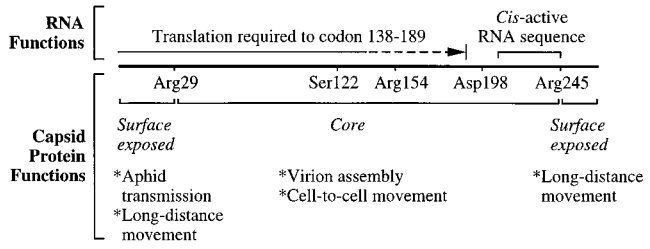


FIG. 7. Diagrammatic representation of CP RNA and protein functions. The thick line represents the CP coding sequence. *cis*-active RNA functions are indicated above the line. Known functions of the CP product are given below the line. The positions coding for Arg-29 and Arg-245, which define the boundaries of the surface-exposed N- and C-terminal domains, respectively, are shown. Sequences coding for residues essential for both encapsidation and cell-to-cell movement functions (Ser-122, Arg-154, and Asp-198) are also shown.

263, was amplified to approximately 27% of the level of parental TEV-GUS. Although moderately suppressed compared with the parental virus, the 189S mutant was amplified to a level approximately 100-fold greater than the 1S, 59S, 103S, and 138S mutants. Second, the amplification defects of the 1S, 59S, 103S, and 138S frameshift-stop codon mutants were not rescued by CP supplied in *trans* by transgenic cells, even though the transgenic CP was capable of assembly into virions (14) and able to rescue intercellular movement defects of certain mutants. Taken together, these data strongly imply that translation of at least the 5′ half of the CP coding sequence, but not the CP translation product, is a prerequisite for efficient genome amplification (Fig. 7). More precisely, we propose that ribosomes must simply be able to reach a critical point between codons 138 and 189 and that the CP coding sequence traversed on the way to this point is unnecessary.

How might the process of translation itself through a sequence encoding a nonessential protein be a stimulatory event? It is likely that secondary structure of RNA within the CP coding sequence is altered by passage of the translational apparatus. Likewise, secondary structure of any RNA segment that is base paired with sequences within the CP sequence would also be disrupted. If critical *cis*-active sequences within the TEV genome are present in "closed" secondary structures with the CP coding sequence upstream of codon 189, passage of ribosomes might activate these sequences. Activation of sequences through movement of ribosomes or activities of the translational apparatus has well-established precedence in positive-stranded RNA bacteriophage (8, 9). Alternatively, the process of translation through part of the CP coding sequence might be required for *cis*-preferential delivery of nascent replication proteins to a site near the genome 3′ end. Several plant- and animal-infecting positive-strand RNA viruses have been shown or postulated to couple genome translation and replication (33, 42, 45), although in each of these cases, the translation requirement was demonstrated through sequences coding for replication proteins.

Another hypothesis to explain the CP translational requirement is that active association of the translational apparatus with the 5′ half of the CP coding sequence may promote genome stability. Recent findings concerning the mechanisms involved in RNA-mediated (or homology-driven) resistance against TEV in transgenic plants support this idea. Transgenic plants expressing nontranslatable RNAs corresponding to the TEV CP sequence can show extreme TEV-specific resistance (30). This resistance is associated with accelerated degradation of the transgenic CP mRNA (41). Interestingly, the degradation pathway may involve preferential targeting of sequences between codons 138 and 189 (21). Transgenic plants expressing translatable TEV CP mRNA do not show extreme resistance, but they exhibit a slow-recovery phenotype characterized by gradual activation of the mRNA degradation process (31). Although the specific mechanisms of recognition and turnover of transgenic mRNAs and incoming homologous viral RNA in resistant plants has not been defined, it is likely that the nontranslatable TEV CP gene is a more potent activator of this cellular response. In view of these data, it seems plausible that the mutant TEV genomes containing frameshift-stop codons at or before codon 138 could rapidly activate a similar cellular response to specifically degrade the nontranslatable region and thus restrict virus amplification.

***cis*-active RNA sequence between CP codons 211 and 246.** The mutational analysis presented here revealed an essential function of the CP coding sequence near the 3′ end of the open reading frame (Fig. 7). The ΔN210 deletion mutant lacking CP codons 2 to 210 was amplified relatively efficiently, whereas the ΔN230 mutant lacking codons 2 to 230 was nonviable. Deletion of codons 246 to 262 (ΔC17 mutant) had no effect on genome amplification, while deletion of codons 190 to 261 (189D mutant) rendered the mutant amplification defective. The amplification-defective phenotype of the 189D mutant was clearly due to loss of the RNA sequence rather than to an effect on translation or the truncated CP product, because the 189S frameshift-stop codon mutant was viable. Infection of cells expressing the CP coding sequence was insufficient to rescue the amplification-defective genomes. These data are consistent with the presence of an essential *cis*-active RNA sequence that is no longer than 105 nucleotides, with a 5′ boundary downstream of codon 210 and a 3′ boundary upstream of codon 246.

This sequence may interact directly or indirectly with other *cis* or *trans* factors involved in genome amplification or may promote stability of the viral genome. *cis*-active sequences generally occur near the 5′ or 3′ ends of RNAs of positive-strand viruses. Internal *cis*-active sequences are not uncommon, however, as has been demonstrated for several viruses, including brome mosaic virus (20) and flock house virus (7), and in defective-interfering RNAs associated with some viruses (13, 25).

The *cis*-active function of the CP coding sequence may involve specific primary, secondary, or tertiary structures. In the sequence between nucleotides 9145 (CP codon 210) and 9495 [3′ terminus plus poly(A) tail], secondary-structure predictions based on free-energy minimization principles (22, 46) revealed a series of potential stem-loop structures. The two lowest-free-energy structures, fold A (−89.3 kcal [1 cal = 4.184 J]/mol) and fold B (−85.1 kcal/mol), are shown in Fig. 8. In each predicted structure, the 3′-terminal 150 nucleotides folded into two stem-loop structures, each containing several bulges and short loops. The *cis*-active sequence between codons 211 and 245 (Fig. 8, boldface type) was predicted to have an extended stem-loop structure containing a perfect 9-bp stem at the base. Sequences flanking this stem were predicted to base pair with extended sequences within the 3′ noncoding region, resulting in formation of a hammerhead-like structure. The 3′ noncoding sequences contributing to the base-paired region of the hammerhead differed in the two folds. The nonessential sequence between codons 246 and 262 (Fig. 8, lowercase type) was predicted to form two imperfect stem-loop structures. Deletion of these sequences, as in the viable ΔC17 mutant, was predicted not to disrupt the proposed hammerhead-like structure of either fold. Although direct biochemical data are not available to support or reject either structure, limited genetic data are consistent with some predictions of fold B. Insertion of either of two 12-nucleotide sequences immediately to the 5′ side of

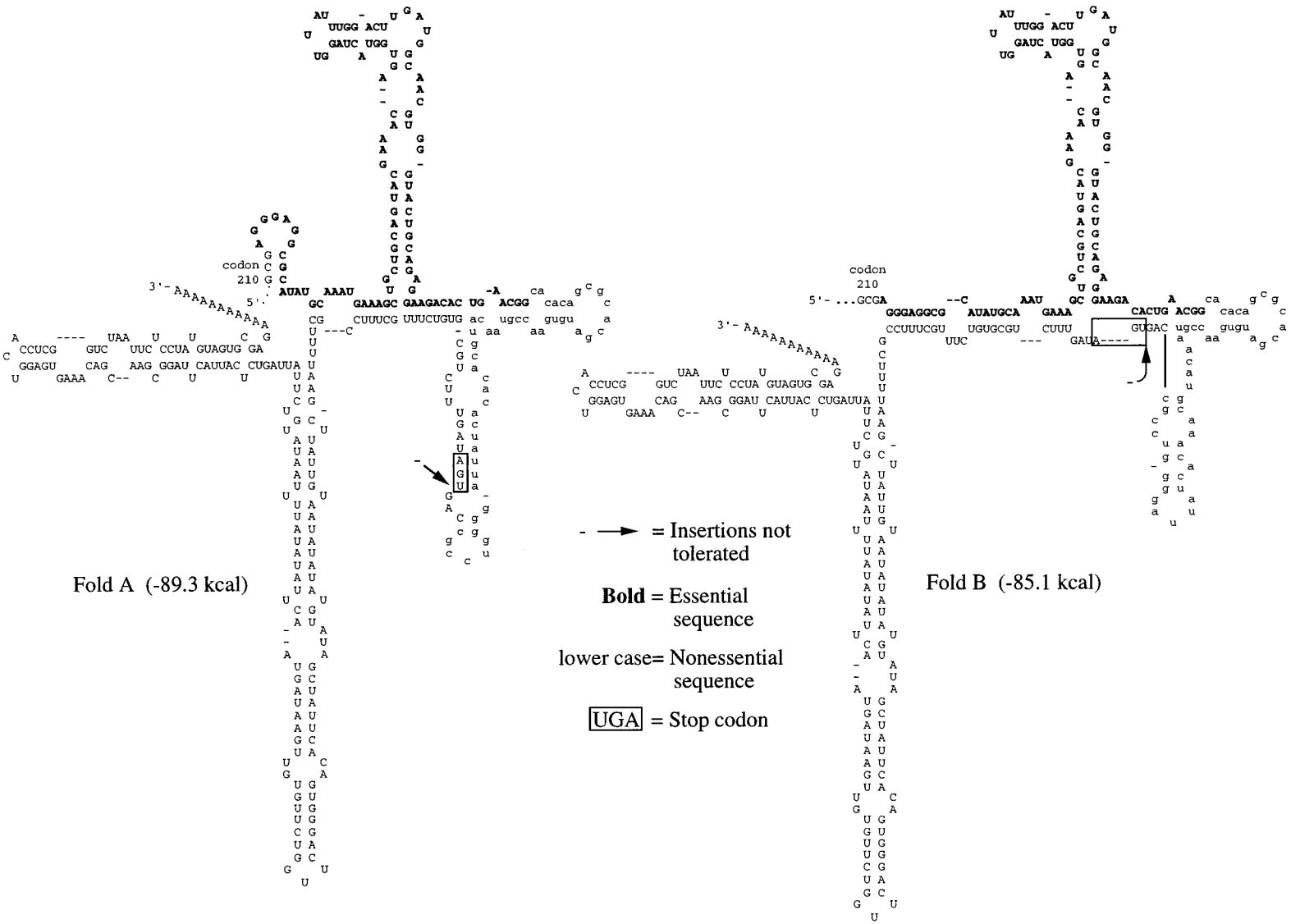


FIG. 8. Secondary-structure predictions of the 3′-terminal 351 nucleotides of the TEV genome. Fold A and fold B were the two lowest-free-energy structures calculated with the program mFOLD. The essential *cis*-active sequence identified between CP codons 211 and 245 is in boldface type. The nonessential sequence between CP codons 246 and 262 is in lowercase type. The position at which oligonucleotide insertions render the TEV genome nonviable is indicated by the arrows.

the translational stop codon (Fig. 8, arrow) rendered the genome nonviable (data not shown). In fold A, this position occurs in a stem-loop that involves a nonessential sequence, while in fold B, it occurs in a stem that involves the *cis*-active sequence. Insertions are predicted to destabilize this putative fold B stem. If this secondary structure is critical, it would explain the requirement for the *cis*-active CP coding sequence and the inactivating effects of the 12-nucleotide insertions.

Implications for genome stability and evolution. The requirements for translation through at least one-half of the CP coding sequence and for the 3′-proximal *cis*-active sequence have significant ramifications for the maintenance and evolution of the replicating virus population. Nearly all spontaneously arising mutant genomes that fail to maintain an intact open reading frame would be lost from the replicating pool. Likewise, any deletion mutant genomes lacking the 3′-proximal *cis*-active sequence within the CP coding region would be lost from the replicating population. These constraints would limit the types of genomes that could be sustained as defective-interfering RNAs. Unlike many other positive- and negative-strand RNA viruses, defective-interfering RNAs from potyvirus genomes have never been documented. Novak and Kirkegaard (33) postulated that defective poliovirus genomes carrying inactivating mutations in sequences coding for *cis*-preferential replication proteins would fail to be propagated in the presence of replication-competent helper viruses. Such mechanisms to ensure that RNA genomes maintain intact open reading frames and encode only functional replication proteins may provide a powerful stabilizing effect on a virus quasispecies population.

ACKNOWLEDGMENTS

We thank Ruth Haldeman-Cahill for excellent technical assistance and Patricia Valdez and Aaron Unterbrink for maintenance of plants.

This research was supported by grants from the U.S. Department of Agriculture (NRICG 95-37303-1867), the National Institutes of Health (AI27832), and the National Science Foundation (IBN9158559).

REFERENCES

1. **Allison, R., R. E. Johnston, and W. G. Dougherty.** 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology **154:**9–20.
2. **Allison, R. F., W. G. Dougherty, T. D. Parks, L. Willis, R. F. Johnston, M. Kelly, and F. B. Armstrong.** 1985. Biochemical analysis of the capsid protein gene and capsid protein of tobacco etch virus: N-terminal amino acids are located on the virion's surface. Virology **147:**309–316.
3. **Atreya, C. D., P. L. Atreya, D. W. Thornbury, and T. P. Pirone.** 1992. Site-directed mutations in the potyvirus HC-PRO gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. Virology **191:**106–111.
4. **Atreya, C. D., R. Raccah, and T. P. Pirone.** 1990. A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. Virology **178:** 161–165.
5. **Atreya, P. L., C. D. Atreya, and T. P. Pirone.** 1991. Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. Proc. Natl. Acad. Sci. USA **88:**7887–7891.
6. **Atreya, P. L., J. J. Lopez-Moya, M. Chu, C. D. Atreya, and T. P. Pirone.** 1995. Mutational analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids. J. Gen. Virol. **76:**265–270.
7. **Ball, L. A., and Y. Li.** 1993. *cis*-acting requirements for the replication of flock house virus RNA 2. J. Virol. **67:**3544–3551.
8. **Berkhout, B., and J. Van Duin.** 1985. Mechanism of translational coupling between coat protein and replicase genes of RNA bacteriophage MS2. Nucleic Acids Res. **13:**6955–6967.
9. **Berkhout, B., A. Van Strien, J. H. Van Boom, J. Van Westrenen, and J. Van Duin.** 1987. Lysis gene of bacteriophage MS2 is activated by translation termination at the overlapping coat gene. J. Mol. Biol. **195:**517–524.
10. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
11. **Carrington, J. C., and W. G. Dougherty.** 1987. Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. J. Virol. **61:**2540–2548.
12. **Carrington, J. C., R. Haldeman, V. V. Dolja, and M. A. Restrepo-Hartwig.** 1993. Internal cleavage and *trans*-proteolytic activities of the VPg-proteinase (NIa) of tobacco etch potyvirus in vivo. J. Virol. **67:**6995–7000.
13. **Chang, Y. C., M. Borja, H. B. Scholthof, A. O. Jackson, and T. J. Morris.** 1995. Host effects and sequences essential for accumulation of defective interfering RNAs of cucumber necrosis and tomato bushy stunt tombusviruses. Virology **210:**41–53.
14. **Dolja, V. V., R. Haldeman, N. L. Robertson, W. G. Dougherty, and J. C. Carrington.** 1994. Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. EMBO J. **13:**1482–1491.
15. **Dolja, V. V., R. Haldeman-Cahill, A. E. Montgomery, K. A. VandenBosch, and J. C. Carrington.** 1995. Capsid protein determinants involved in cell-to-cell and long distance movement of tobacco etch potyvirus. Virology **207:** 1007–1016.
16. **Dolja, V. V., H. J. McBride, and J. C. Carrington.** 1992. Tagging of plant potyvirus replication and movement by insertion of β-glucuronidase into the viral polyprotein. Proc. Natl. Acad. Sci. USA **89:**10208–10212.
17. **Dougherty, W. G., S. M. Cary, and T. D. Parks.** 1989. Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. Virology **171:** 356–364.
18. **Dougherty, W. G., and B. L. Semler.** 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol. Rev. **57:**781–822.
19. **Dougherty, W. G., L. Willis, and R. E. Johnston.** 1985. Topographical analysis of tobacco etch virus capsid protein epitopes. Virology **144:**66–72.
20. **French, R., and P. Ahlquist.** 1987. Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA3. J. Virol. **61:**1457–1465.
21. **Goodwin, J., K. Chapman, S. Swaney, T. D. Parks, E. Wernsman, and W. G. Dougherty.** 1996. Genetic and biochemical dissection of transgenic RNA-mediated virus resistance. Plant Cell **8:**95–105.
22. **Jaeger, J. A., D. H. Turner, and M. Zucker.** 1989. Improved predictions of secondary structure for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
23. **Jefferson, R. A.** 1987. Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol. Rep. **5:**387–405.
24. **Kasschau, K. D., and J. C. Carrington.** 1995. Requirement for HC-Pro processing during genome amplification of tobacco etch potyvirus. Virology **209:**268–273.
25. **Kim, Y. N., Y. S. Jeong, and S. Makino.** 1993. Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication. Virology **197:**53–63.
26. **Klein, P. G., R. R. Klein, E. Rodríguez-Cerezo, A. G. Hunt, and J. G. Shaw.** 1994. Mutational analysis of the tobacco vein mottling virus genome. Virology **204:**759–769.
27. **Kunkel, T. A., J. D. Roberts, and R. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
28. **Li, X. H., and J. C. Carrington.** 1995. Complementation of tobacco etch potyvirus mutants by active RNA polymerase expressed in transgenic cells. Proc. Natl. Acad. Sci. USA **92:**457–461.
29. **Lindbo, J. A., and W. G. Dougherty.** 1992. Pathogen-derived resistance to a potyvirus: immune and resistant phenotypes in transgenic tobacco expressing altered forms of a potyvirus coat protein nucleotide sequence. Mol. Plant-Microbe Interact. **5:**144–153.
30. **Lindbo, J. A., and W. G. Dougherty.** 1992. Untranslatable transcripts of the tobacco etch virus coat protein gene can interfere with tobacco etch virus replication in transgenic plants and protoplasts. Virology **189:**725–733.
31. **Lindbo, J. A., L. Silva-Rosales, W. M. Proebsting, and W. G. Dougherty.** 1993. Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. Plant Cell **5:**1749–1759.
32. **Negrutiu, I., R. Shillito, I. Potrykus, G. Biasini, and F. Sala.** 1987. Hybrid genes in the analysis of transformation conditions. 1. Setting up a simple method for direct gene transfer in plant protoplasts. Plant Mol. Biol. **8:**363–373.
33. **Novak, J. E., and K. Kirkegaard.** 1994. Coupling between genome translation and replication in an RNA virus. Genes Dev. **8:**1726–1737.
34. **Oh, C.-S., and J. C. Carrington.** 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. Virology **173:** 692–699.
35. **Parks, T. D., K. K. Leuther, E. D. Howard, S. A. Johnston, and W. G. Dougherty.** 1994. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. Anal. Biochem. **216:**413–417.
36. **Restrepo-Hartwig, M. A., and J. C. Carrington.** 1994. The tobacco etch potyvirus 6-kilodalton protein is membrane associated and involved in viral replication. J. Virol. **68:**2388–2397.
37. **Riechmann, J. L., S. Laín, and J. A. García.** 1992. Highlights and prospects of potyvirus molecular biology. J. Gen. Virol. **73:**1–16.
38. **Shukla, D. D., P. M. Strike, S. L. Tracy, K. H. Gough, and C. W. Ward.** 1988. The N and C termini of the coat proteins of potyviruses are surface-located and the N terminus contains the major virus-specific epitopes. J. Gen. Virol. **69:**1497–1508.

39. **Shukla, D. D., and C. W. Ward.** 1989. Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group. Adv. Virus Res. **36:**273–314.
40. **Slade, D. E., R. E. Johnston, and W. G. Dougherty.** 1989. Generation and characterization of monoclonal antibodies reactive with the 49-kDa proteinase of tobacco etch virus. Virology **173:**499–508.
41. **Smith, H. A., S. L. Swaney, T. D. Parks, E. A. Wernsman, and W. G. Dougherty.** 1994. Transgenic plant virus resistance mediated by untranslatable sense RNAs: expression, regulation, and fate of nonessential RNAs. Plant Cell **6:**1441–1453.
42. **van Bokhoven, H., O. Le Gall, D. Kasteel, J. Verver, J. Wellink, and A. van Kammen.** 1993. Cis- and trans-acting elements in cowpea mosaic virus RNA replication. Virology **195:**377–386.
43. **Verchot, J., and J. C. Carrington.** 1995. Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. J. Virol. **69:**1582–1590.
44. **Verchot, J., and J. C. Carrington.** 1995. Evidence that the potyvirus P1 protein functions as an accessory factor for genome amplification. J. Virol. **69:**3668–3674.
45. **Weiland, J. J., and T. W. Dreher.** 1993. Cis-preferential replication of the turnip yellow mosaic virus RNA genome. Proc. Natl. Acad. Sci. USA **90:** 6095–6099.
46. **Zucker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.