

# Encapsidation of Turnip Crinkle Virus Is Defined by a Specific Packaging Signal and RNA Size

FENG QU AND T. JACK MORRIS\*

*School of Biological Sciences, University of Nebraska—Lincoln, Lincoln, Nebraska 68588-0118*

Received 20 September 1996/Accepted 4 November 1996

**A protoplast infection assay has been used to reliably examine the viral RNA encapsidation of turnip crinkle virus (TCV). Analysis of the encapsidation of various mutant viral RNAs revealed that a 186-nucleotide (nt) region at the 3' end of the coat protein (CP) gene, with a bulged hairpin loop of 28 nt as its most essential element, was indispensable for TCV RNA encapsidation. When RNA fragments containing the 186-nt region were used to replace the CP gene of a different virus, tomato bushy stunt virus, the resulting chimeric viral RNAs were encapsidated into TCV virions. Furthermore, analysis of the encapsidated chimeric RNA species established that the RNA size was an important determinant of the TCV assembly process.**

The specific recognition of viral genomic RNA by capsid protein (CP) is a critical event in the infection cycle of all single-stranded RNA viruses. This recognition event plays a crucial role in specific encapsidation of the viral genome as well as other steps in the virus life cycle. In the RNA phage, for example, the binding of CP to specific RNA elements regulates translation of the replicase as well as assembly initiation (43). In hepatitis B pararetrovirus, the packaging of the RNA pregenome is a prerequisite for the reverse transcription of the pregenome into genomic DNA (2, 25). In most plant viruses, RNA packaging has been shown to be essential for the systemic spread (6, 33) and, in some cases, essential for cell-to-cell movement (7, 34).

The high specificity of viral RNA encapsidation is reflected in the observation that cellular RNAs are only rarely packaged into virions (27). In addition, viral RNA is not usually packaged by the CP of an unrelated virus upon coinfection of the same cells (references 5, 21, and 22 and our unpublished observation). Despite the importance of this RNA-protein recognition event, the specific packaging signal or origin of assembly (OAS) has not been well characterized for very many viruses. The most thoroughly characterized OAS is that of tobacco mosaic virus (46, 47). Sequence elements involved in specific packaging of the viral RNAs of hepatitis B virus (19), Sindbis virus (38, 39), and the yeast double-stranded RNA virus L-A have also been identified (10). The identification of a specific packaging signal for an unenveloped, icosahedral plant virus has not yet been accomplished despite intensive investigations spanning more than 30 years (see reference 9 for a review). We report here the identification and functional analysis of the first OAS for an icosahedral plant virus.

Turnip crinkle virus (TCV) is a 30-nm-diameter isometric plant virus which consists of a single copy of a 4,051-nucleotide (nt) single-stranded RNA genome and 180 copies of the 38-kDa CP arranged in T=3 icosahedral symmetry (4, 17). TCV is a member of the *Carmovirus* genus in the *Tombusviridae*. It is physically similar but genetically distinct from the family's type member, tomato bushy stunt virus (TBSV). The structures of both TCV and TBSV have been elucidated by X-ray crystallography and shown to be quite similar (13, 17). TCV became

the model system of choice for studying assembly because of the detailed in vitro dissociation and reassembly studies of Sorger et al. (32) and because of the availability of a genomic cDNA clone from which infectious viral RNA could be transcribed and on which in vitro mutagenesis could be performed (16). Sorger et al. (32) showed that at elevated pH, TCV would dissociate into a ribonucleoprotein (rp) complex that consisted of the viral RNA tightly attached to six CP subunits. They further demonstrated that this rp complex could serve as a nucleation complex for reassembly in vitro. We subsequently purified this rp complex, treated it with RNase A and RNase T<sub>1</sub>, and determined the sequences of a set of undigested RNA fragments presumably protected in the complex by the remaining CP subunits (35). Several of the RNA fragments were found to be able to fold together into ordered structures in two general locations in the genome. One cluster (Ff/Fa) was located in the polymerase gene, and the second set (Fd and Fe) was located in the CP gene. The Ff/Fa structure was initially chosen for detailed investigation because it formed a long hairpin of 15 bp that surrounded the first stop codon within the RNA polymerase gene. We suspected that it might play a role in the regulation of the readthrough of the polymerase gene. Although preliminary in vitro binding and gel shift assays suggested that the Ff/Fa element might bind TCV CP specifically (36), we were unable to confirm a higher  $K_d$  of binding for the viral RNA regions by using quantitative gel shift and filter binding assays (31). Furthermore, we could not directly confirm a role for the Ff/Fa element in the assembly in vivo because mutants constructed with a disrupted hairpin structure also failed to replicate (37).

Consequently, we have not been able to definitively identify a specific packaging signal in TCV RNA by using the in vitro approaches that proved successful for TMV and small RNA phages. These in vitro approaches have also had limited success for cowpea chlorotic mottle virus, another spherical plant RNA virus for which intensive in vitro assembly studies have been undertaken (see reference 9 for a review). In this report, we describe an alternative approach that involves the analysis of the assembly efficiency of complementing mutants coinoculated into protoplasts. We have used this in vivo analysis to identify a 186-nt fragment at the 3' end of the CP-coding region responsible for specific packaging of viral RNA. We also provide evidence to show that the size of the viral RNA being packaged is a critical factor for stable assembly of virions.

\* Corresponding author. Mailing address: School of Biological Sciences, 348 Manter Hall, University of Nebraska-Lincoln, Lincoln, NE 68588-0118. Phone: (402) 472-6676. Fax: (402) 472-2083. E-mail: jmorris@unlinfo.unl.edu.

## MATERIALS AND METHODS

**Construction of plasmids.** The restriction enzymes were from New England Biolabs, and all DNA manipulations were done as described by Sambrook et al. (29). The oligodeoxynucleotides (oligos) were from the DNA Synthesis Lab at the University of Nebraska—Lincoln. Plasmids T1d1, APA, and RT were described previously (11, 16). dA was constructed by using the site-directed mutagenesis procedure described by Hearne et al. (14) and oligo 5'-TCTTTTGTCCGCTAG\_ACTCAgATCTCCAGGGGTGAA-3'; the dash identifies the location of the deletion, and the lowercase nucleotide was changed to create a *Bgl*II site to facilitate screening. RT-dNar and dNar were obtained by deleting the *Nar*I fragment from the RT and T1d1 clones, respectively. NarR was created by religating the *Nar*I fragment in the reverse direction.

Plasmid TCR was constructed by ligating a part of the rice U3 snRNA gene (nt 1 to 125 of its coding region [26]) into *Nhe*I-digested and blunt-ended plasmid T1d1. The deletion mutants identified in Fig. 3A were constructed by cutting TCR with the indicated restriction enzymes followed by Klenow treatment and religation. For some of these mutants, additional manipulations were required to avoid complications associated with multiple digestion sites in the clone.

Plasmids TCR-ST, -dI, -dII, -dIII, -dIV, -dV, -dIVa, -dIVb, -dIVc, and -dIVd were constructed using the Transformer site-directed mutagenesis kit (Clontech, Palo Alto, Calif.) and oligos ST (+*Nsi*I) (5'-CAACACTGGAAATG\_cATGGC GCCAATG-3'), dI (+*Avr*II) (5'-CCGGAACTTCTC\_ctAGGCGCCGGTC-3'), dII (-*Nar*I) (5'-CATTTGTTCTACGAGA\_CCGCAGCAGGTA-3'), dIII (+*Hind*III) (5'-CCTGGGAGAAAGC\_tGGGCAGGAGTG-3'), dIV (+*Hind*III) (5'-CAGGTAGTGTCCAA\_gcTTAGTACGGTAATAGT-3'), dV (+*Ngo*MI) (5'-TGGGAGAAGCCG\_GCCAAAGGGTAAAT-3'), dIVa (+*Hind*III) (5'-CAACTGAGGAGCA\_AgcTTAGTACGGTAATAGT-3'), dIVb (+*Mun*I) (5'-GAGGAGCAGCCAA\_TTGCAAGCACTAG-3'), dIVc (+*Pst*I) (5'-AATGG TCACAACCTG\_cAGGGTAAATTGCAAG-3'), and dIVd (+*Ssp*I) (5'-GCCAA AGGGTAAAT\_ ATTTAGTACGGTAATA-3'). A dashed line denotes the sequence deleted; the restriction enzymes in parentheses refer to the corresponding restriction sites created (marked as +) or eliminated (marked as -) to facilitate screening; nucleotides mutated or added are indicated by lowercase letters.

Plasmid TBTCep was constructed from the TBSV infectious clone T100 (14). T100 was modified by using the Transformer site-directed mutagenesis kit and two oligos, TBS-2650C (5'-ACGCGAGATAGACcCATGGCAATGGTA-3') and TBS-3826CT (5'-TTTAGTGTGTCCcCGAGGGGCCTCT-3'), to produce TB-M3, which had a new *Nco*I site around the CP start codon and an *Xho*I site 8 nt downstream of the CP stop codon. The original *Nco*I site in T100 was eliminated simultaneously by oligo TBS-3888G (5'-CAAGTCAATAAACCGTG GAACGAGCTA-3'). The TCV CP gene was amplified from T1d1 by using oligos TCV5'2743 (5'-CGGCGCCCATGGAAATGATCCTAGAGTCCGGGA-3' and TCV3'3805 (5'-CCCGTGTCTCGAGTACCGTACTAAATCTGAGTG CTTG-3'), in which the new *Nco*I and *Xho*I sites, respectively, are underlined. The amplified TCV CP gene was ligated directly into TB-M3 digested with *Nco*I and *Xho*I. TBTCep-PB was created by deleting the portion between *Pf*MI (nt 3856 of TBSV) and *B*stBI (nt 4388 of TBSV) of plasmid TBTCep (shown in Fig. 6A). TBTCep-NM and TBTCep-NH were created by deleting the portions of TBTCep between *Nco*I (start codon of TCV CP) and *M*scI (nt 3386 of TCV) and between *Nco*I and *H*inII (at nt 3606 of TCV), respectively. TBTCep-PB-HA was created by deleting the HA fragment (see Fig. 3A and 7A) from TBTCep-PB. An additional plasmid, TBTCep-NH/RU, was created by inserting a 152-nt fragment of a rice U3 snRNA gene (from nt -27 to +125 of the gene [26]) into the *Acc*65I site (at nt 3961 of TBSV) of TBTCep-NH.

**Encapsidation assay.** In vitro transcripts were made using an AmpliScribe T7 transcription kit (Epicentre Technologies), and 6 to 7  $\mu$ g of each transcript was inoculated into cucumber protoplasts as described previously (18, 40). Total nucleic acids were extracted from protoplasts and analyzed on a 1.2% agarose-0.5 $\times$  Tris-borate-EDTA gel at 45 V for 2.5 h prior to staining with ethidium bromide and electrotransfer.

Virus particles were isolated from the protoplast infections by resuspension of pelleted protoplasts in 200  $\mu$ l of 0.2 M sodium acetate (NaOAc; pH 5.2). Sterile glass beads (0.1-mm diameter) were then added, and the mixture was shaken in a Mini-Beadbeater (Biospec Products) for two 15-s intervals. The aqueous phase was recovered by centrifugation, and the solid phase was reextracted with an additional 200  $\mu$ l of 0.2 M NaOAc (pH 5.2). The aqueous phases were then combined and kept on ice for 1 h prior to centrifugation for 5 min at 11,000 rpm. The supernatant was then mixed with 1/4 volume of 40% polyethylene glycol-1 M NaCl and kept on ice overnight. The virus precipitate was collected by centrifugation for 30 min at 15,000 rpm and dissolved in 20  $\mu$ l of 10 mM NaOAc (pH 5.5). Virus content was analyzed by electrophoresis of 5  $\mu$ l of the virus solution on a 1% agarose gel prepared as described by Heaton (15).

Agarose gels of both the nucleic acids and viruses were stained with ethidium bromide prior to electrotransfer to nylon membranes, using a Bio-Rad apparatus as specified by the manufacturer. Two kinds of nylon membranes were used: Hybond-N (Amersham) and Zeta-Probe GT (Bio-Rad). The hybridization procedures recommended by the manufacturers were followed.

Northern blot analysis was performed with the <sup>32</sup>P-labelled TCV-specific probe, made by using an Oligolabelling kit (Pharmacia Biotech), and pT1d1 as the template. A NarR-specific probe labelled by using a Genius 4 kit (Boehringer

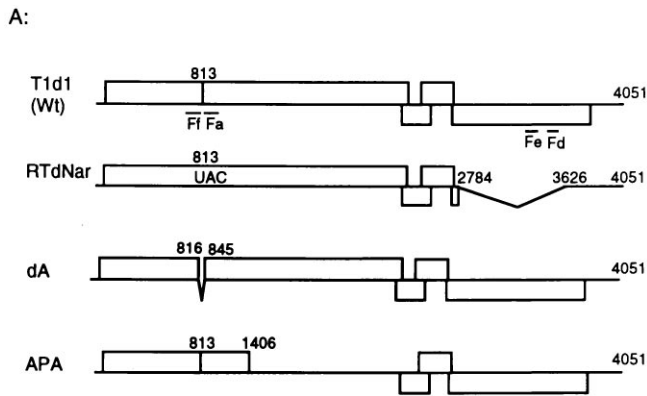
Mannheim) was made by in vitro transcription of a PCR fragment amplified by using oligos TCV5'2865 (5'-GCGCTAATACGACTCACTATAGGGATCAAG CTCTCTCTGT-3' (the underlined region represents the T7 promoter) and TCV3'3293 (5'-AAACCTGTCCAAGGCACGCTAGA-3'). A <sup>32</sup>P-labelled TBSV-specific probe was made by oligolabelling (Pharmacia Biotech) a 400-bp 3'-end fragment of TBSV cDNA (PCR amplified by using oligos 9 and 42 in reference 41). A <sup>32</sup>P-labelled TCR-specific probe was made by oligolabelling the PCR fragment generated with oligos U3P and U3N (26) and pRU3-5.

## RESULTS

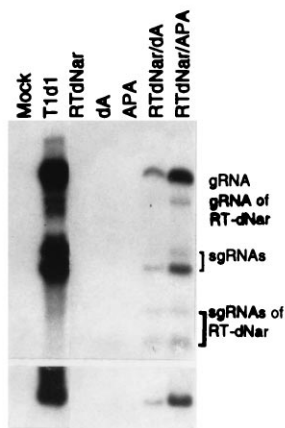
**The Fa element is not essential for packaging.** Our inability to conclusively implicate the RNA regions identified in the rp complex in assembly by using the in vitro procedures prompted us to develop the in vivo approach detailed in Materials and Methods. Briefly, infections of wild-type and mutant viruses were initiated by inoculating cucumber protoplasts with RNA transcripts made from TCV infectious clones. After 24 h of incubation, the inoculated protoplast set was divided into two aliquots: one-fourth of the protoplasts were extracted for total RNA, and the remainder were processed for virion purification. The RNA sample was analyzed on a 1.2% agarose gel in 0.5 $\times$  Tris-borate-EDTA, and the intact virus preparation was electrophoresed in a 1% agarose gel in 5 mM Tris-38 mM glycine (pH 8.3) as described by Heaton (15). Both gels were then stained with ethidium bromide prior to blotting and hybridization with appropriate probes.

We initially investigated the putative role of the Ff/Fa structure in virus assembly by deleting the Fa fragment from the genome. Previously, White et al. (42) demonstrated that both products of the TCV polymerase gene, P28 and its readthrough product P88, were essential for viral RNA replication. The mutant RT was produced with the readthrough stop codon (UAG) for P28 changed to UAC so that P88 would be the only product of the polymerase gene. Although this mutant was replication defective, it did complement other P88-defective mutants that were able to produce a functional P28 product (e.g., the APA mutant in Fig. 1). To test for the role of Fa in assembly, we produced a complementable mutant (dA) with a functional P28 open reading frame (ORF) in which the Fa fragment (nt 816 to 845) was completely deleted. A second, complementable P88 mutant was constructed by deleting the *Nar*I fragment (nt 2784 to 3626) from the CP gene of mutant RT (RT-dNar). RT-dNar was then coinoculated into protoplasts with mutants dA and APA. As shown in Fig. 1B, all three mutants were replication defective unless coinoculated with the complementing mutant. Viral RNAs accumulated in both coinoculations, and as expected (42), the RT-dNar mutant replicated less efficiently. Although the dA mutant replicated to a lesser extent than the APA mutant in the coinoculations, a proportional level of virus particle accumulation was detected in both sets of infections, as determined by comparing the relative levels of total viral RNA and virus (Fig. 1B). This result suggested that there was no significant difference in packaging efficiency between dA and APA RNAs. Furthermore, in an experiment described below, we also found that the mutants of RT-dNar size were not packaged. We thus conclude that the fragment Fa is not essential for TCV viral RNA packaging and therefore cannot be an essential component of the TCV OAS in vivo.

**The CP gene contains the OAS of TCV.** We next examined the role of the CP coding region in the RNA packaging process by testing the following replication-competent mutants: (i) dNar, a mutant in which the *Nar*I fragment (nt 2784 to 3626) within the CP gene was deleted; and (ii) NarR, a mutant in which the *Nar*I fragment was reversed (Fig. 2A and reference 11). Both of these mutants replicated to wild-type levels when

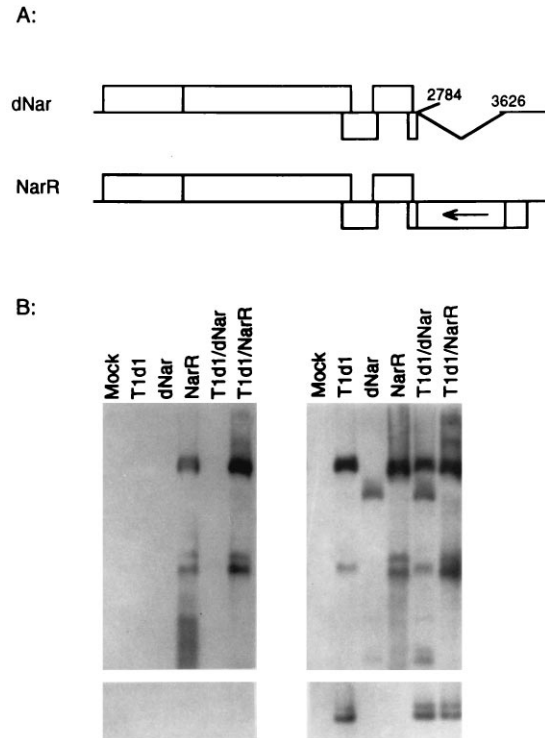


**B:**



**FIG. 1.** Packaging of TCV RNAs modified in the polymerase gene. (A) Schematic representation of the wild-type (Wt; T1d1) TCV genome and derived mutants. The ORFs are depicted by the boxes, and the encoded ORFs are identified as specified in reference 11. The most 5'-proximal ORF is the polymerase gene which includes the P28 and P88 ORFs punctuated by an amber terminator at nt 813. The two small ORFs encode the movement proteins, P8 and P9. The CP gene is the most 3'-proximal ORF. The small bars below the diagram identify the locations of regions of RNA protected from nuclease digestion in an rp complex analyzed by Wei et al. (35). (B) Northern hybridization analysis of a total RNA fraction (top panel) and virus particles (bottom panel) separately isolated from aliquots of the same protoplast sample 24 h after inoculation. The protoplasts were inoculated with transcripts of the T1d1 and mutant clones alone and in various combinations. Both blots were hybridized with the same  $^{32}$ P-labelled TCV-specific probe. The relative accumulation of the virus-specific RNAs (top panel) compared to the amount of viral RNA in the virus particles (bottom panel) reflects proportional recovery of total viral RNA and virus from the same infection. The encapsidation efficiency for each infection can therefore be directly interpreted by comparing the relative intensities of the hybridization signals in both panels.

inoculated alone, and as expected, no virus particles were detected in the single infections (Fig. 2B). Moreover, both mutants accumulated to the same level as the wild type when coinoculated with TCV helper virus. However, when the blots were reprobed with a NarR-specific probe to differentially detect the NarR mutant from the wild-type helper in the coinoculations, it was evident that NarR RNA was not packaged into virions in the doubly infected cells (Fig. 2B, bottom left panel). Similarly, the smaller dNar genomic RNA was not detected in virions when RNAs extracted from viruses purified from the coinoculations were analyzed (data not shown). The fact that the dNar and NarR mutants were unable to assemble into virions in the presence of sufficient CP provided in *trans* by



**FIG. 2.** Packaging of TCV RNAs mutated in the CP gene. (A) Schematic representation of two mutants modified in the CP gene. Mutant dNar has a majority of the CP gene deleted between the two *NarI* sites at nt 2784 and 3626. Mutant NarR has the sequence between the two *NarI* sites inverted. (B) Northern hybridization analysis of total RNA fractions (top panels) and virus particles (bottom panels) of protoplast infections processed essentially as described for Fig. 1. The lanes on the right were hybridized with a TCV-specific probe; the lanes on the left show the same membranes hybridized with the NarR-specific minus-sense probe.

a helper virus suggested that the RNA packaging signal was most likely within the CP gene itself.

**Delineation of the packaging signal in the CP gene.** We next made a series of mutants with small deletions that spanned the CP-coding region to further characterize the size and location of the putative packaging signal (Fig. 3A). Because the assembly competency of the deletion mutants had to be assayed in mixed infections with wild-type virus, we created a distinctive TCV mutant with a marker sequence (TCR) and then constructed the deletion mutants in the TCR background. The TCR mutant was obtained by replacing a portion of P8 movement gene (a 12-nt *NheI* fragment from TCV nt 2457 to 2469 [19a]) with a 125-nt foreign sequence derived from a rice U3 snRNA gene (26). Transcripts made from TCR were able to replicate and package like wild-type virus in protoplast infections, and the mutant RNA genome could be specifically detected with a rice U3-specific probe (Fig. 3). All of the CP deletion mutants were shown to replicate normally in protoplasts alone, and as expected, none could produce detectable virus (data not shown). In coinoculations with TCV helper, all but one of these deletion mutants were able to assemble into virions (Fig. 3B). The mutant that failed to package (TCR-HA) contained a 186-nt deletion that extended from nt 3606 to 3792. This HA fragment was at the 3' end of the CP gene and included a portion of the Fd fragment identified in the rp complex in previous studies. Additional smaller deletions were made in this region to further delineate the packaging signal within the 186-nt HA fragment (Fig. 4). As seen in Fig. 4B, all

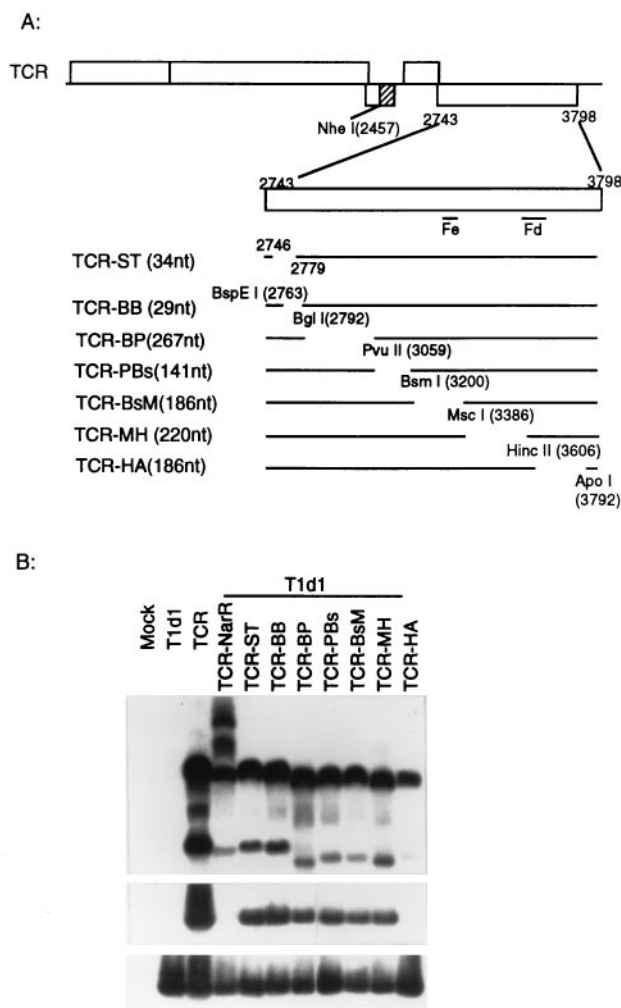


FIG. 3. Packaging of mutant RNAs with deletions in the TCV CP gene. (A) Schematic representation of TCR clone and the CP gene deletion mutants derived from TCR. The box below the genome diagram shows the enlarged CP gene; the lines symbolize the different deletion mutants, with the locations of the deleted portion and the restriction sites used to construct the mutants shown. TCR-ST was obtained by *in vitro* mutagenesis. The numbers in parentheses next to the mutant names show the sizes (in nucleotides) of the deletions. (B) Northern hybridization analysis of total RNA fractions (top panel) and virus particles (bottom two panels) of protoplast infections processed essentially as described for Fig. 1. The top two panels show analyses of the viral RNAs and virions of the TCR-based deletion mutants, using a TCR-specific probe; the bottom panel shows the accumulation of both mutant and wild-type helper virions by reprobing the same virus gel blot with a TCV-specific probe.

of the TCR-dI to -dV mutants replicated to the same extent as TCR, but all appeared to be packaged less efficiently than mutant TCR-MH, which contained a 220-nt deletion just upstream of the HA fragment. Among these mutants, TCR-dIV appeared to be the only one that was completely deficient in assembly. By comparing the deleted portions of TCR-dIV and -dV, we narrowed the essential packaging element to a 28-nt region of TCV RNA spanning nt 3765 to 3792. It is important to note that since mutants TCR-dI, -dII, -dIII, and -dV all were packaged less efficiently than TCR-MH, the sequences surrounding the 28-nt region probably play a significant role in the overall efficiency of assembly initiation.

The RNAFOLD program (Genetics Computer Group, Madison, Wis.) was used to predict the RNA secondary structure for the 186-nt HA fragment (Fig. 5A). The 28-nt packag-

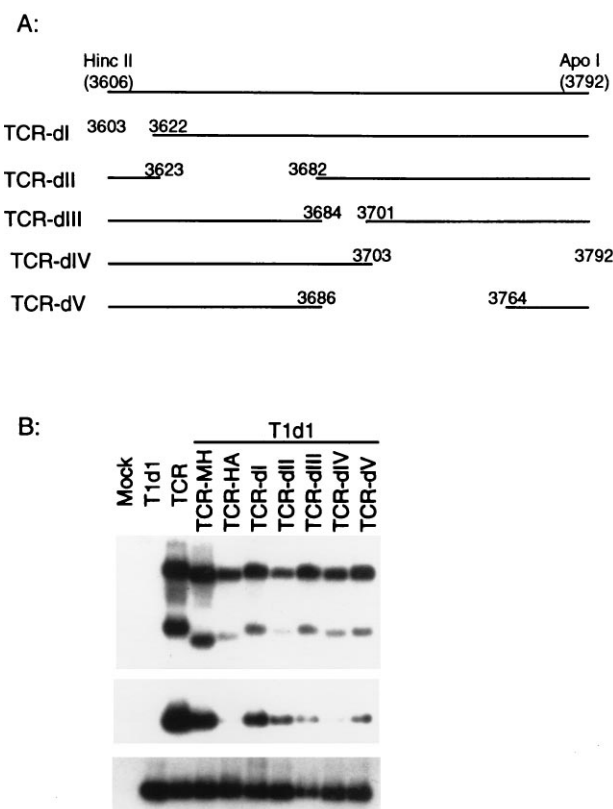


FIG. 4. Packaging of mutant RNAs with deletions within HA fragment. (A) Schematic representation of the additional deletion mutants constructed within the HA fragment in the CP gene. (B) Northern hybridization analysis of total RNA fractions (top panel) and virus particles (bottom two panels) of protoplast infections probed as described for Fig. 3B.

ing element could be readily identified in a bulged hairpin loop structure that extended from nt 3755 to 3791 (Fig. 5B). This same bulged hairpin loop appeared in all of the structures generated by analyzing additional RNA fragments with different lengths of flanking sequence. Furthermore, we evaluated the minimal size of the element needed for assembly by creating four additional deletion mutants: TCR-dIVa, with nt 3765 to 3792 deleted; TCR-dIVb, with the loop from nt 3770 to 3777 deleted; and TCR-dIVc and -dIVd, with the upper and lower arms of the stem deleted (Fig. 5B). None of these deletion mutants were assembled into virions (Fig. 5C), and we therefore could not define an essential sequence element smaller than 28 nt as the minimal RNA element necessary for assembly of RNA into virions.

**The TCV CP gene promotes assembly of an unrelated viral genome.** Confirmatory evidence for the location of the TCV OAS in the CP gene came with the demonstration that TCV CP could assemble around a chimeric virus genome consisting of the TCV CP-coding region and the genetically unrelated viral RNA genome of TBSV. First, we constructed a TBSV-TCV chimeric virus in which the TBSV CP gene was precisely replaced with the TCV CP gene (TBTCcp in Fig. 6A). We also produced a smaller chimeric viral genome, closer in size to that of the TCV genome (TBTCcp-PB, 4,133 nt), by deleting a portion of TBSV movement genes (nt 3856 to 4388 of TBSV [14]) shown previously to be dispensable for replication in protoplasts (30). Both chimeric mutants replicated to about the same level as wild-type TBSV in protoplasts, and only the

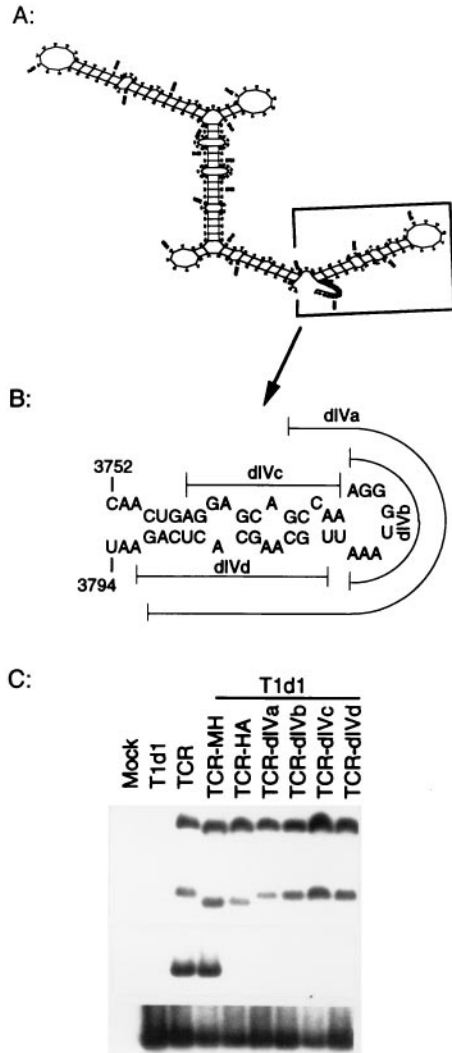


FIG. 5. Secondary structure of the HA fragment and further delineation of the packaging signal in the HA fragment. (A) Predicted secondary structure of the region of the TCV sequence encompassing the HA fragment. The essential element enlarged in panel B is identified as the boxed portion. (B) Predicted secondary structure of the stem-loop element extending from nt 3752 to 3794 in the HA fragment. The lines identify the four smaller regions deleted in an effort to further delineate the packaging signal (dIVa, nt 3765 to 3791; dIVb, nt 3770 to 3777; dIVc, nt 3758 to 3769; and dIVd, nt 3779 to 3792). (C) Northern hybridization analysis of total RNA fractions (top panel) and virus particles (bottom two panels) isolated from the protoplast infections probed as described for Fig. 3B.

shorter chimera (TBTCcp-PB) appeared to be capable of limited assembly into TCV virions (Fig. 6B, lane 5). In contrast, both chimeric mutants were packaged to detectable levels into TCV virions when they were coinoculated with wild-type TCV helper, which was presumably needed to supply a sufficient amount of the TCV CP *in trans* (Fig. 6B, left, lanes 6 and 7). (It is worth noting here that TBSV and TCV virions are distinguished by the different mobilities evident in the ethidium bromide-stained gel presented in the bottom left panel of Fig. 6B. Note as well that the TBSV- and TCV-specific probes directly detect which genomic RNA is packaged into virions of appropriate mobility.) In control experiments, we found no evidence for cross-packaging in coinoculations of TCV and TBSV or of TCV and a TCV-size deletion mutant of the TBSV

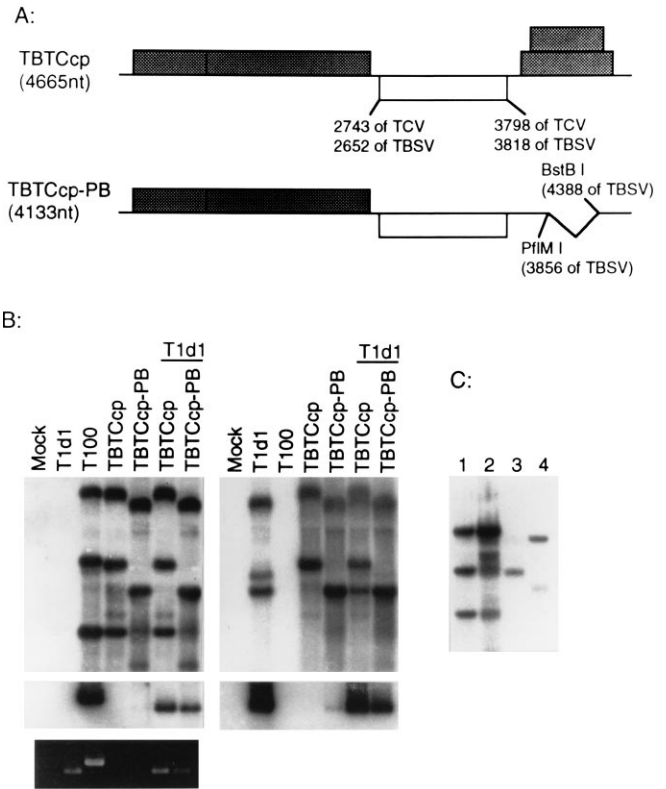


FIG. 6. Packaging of the TBSV-TCV chimeric viral RNAs. (A) Schematic representation of the TBTCcp and TBTCcp-PB chimeric clones. The shadowed boxes represent the ORFs identified in the genomic clone of TBSV (T-100 in reference 14). The open box identifies the location of the TCV CP gene that was used to replace the TBSV CP gene. The numbers below the boxes identify the locations of the start and stop codons in the genomes. The bent line in TBTCcp-PB identifies the location of the deletion. The two 5'-proximal ORFs of TBSV encode the polymerase genes needed for replication, and the two 3'-proximal ORFs, deleted in TBTCcp-PB, have been shown not to be necessary for replication. (B) Northern hybridization analysis of total RNA fractions (top panels) and virus particles (bottom panels) of protoplast infections processed essentially as described for Fig. 1. The left and right panels show the same set of membranes hybridized with the TBSV-specific and TCV-specific probes, respectively. The bottom panel on the left is the virus gel stained with ethidium bromide to illustrate the distinctive mobilities of the intact TCV and TBSV virions. (C) Northern blot analysis to show the TBSV-specific viral RNAs isolated from the virion preparations shown in panel B. Lane 1, total RNA isolated from protoplasts infected with the T100 transcript; lane 2, viral RNAs from the TBSV virions purified from the T100-infected protoplasts; lane 3, viral RNAs from virions isolated from protoplasts inoculated with helper TCV (T1d1) and chimeric TBTCcp; lane 4, viral RNAs from virions isolated from protoplasts inoculated with T1d1 and chimeric TBTCcp-PB.

genome (data not shown here). These results confirm that the TCV CP-coding region contains the element necessary for specific packing of a foreign viral RNA into TCV virions.

**RNA size as a constraint in virus assembly.** Although the size limit of RNA that can be assembled into single-stranded RNA viruses with icosahedral symmetry has not been investigated directly, we had some preliminary indication that there would be an RNA size limit for the TCV virion (31). It was unexpected, therefore, that TBTCcp would assemble because it was more than 600 nt larger than the TCV genome. When we examined the viral RNA species in virions purified from the protoplast infections (Fig. 6C), we discovered that only the 2,044-nt subgenomic RNA (sgRNA) and not the larger 4,665-nt genomic RNA of TBTCcp was packaged in the TBTCcp/T1d1 coinfections (lane 3). In contrast, the smaller

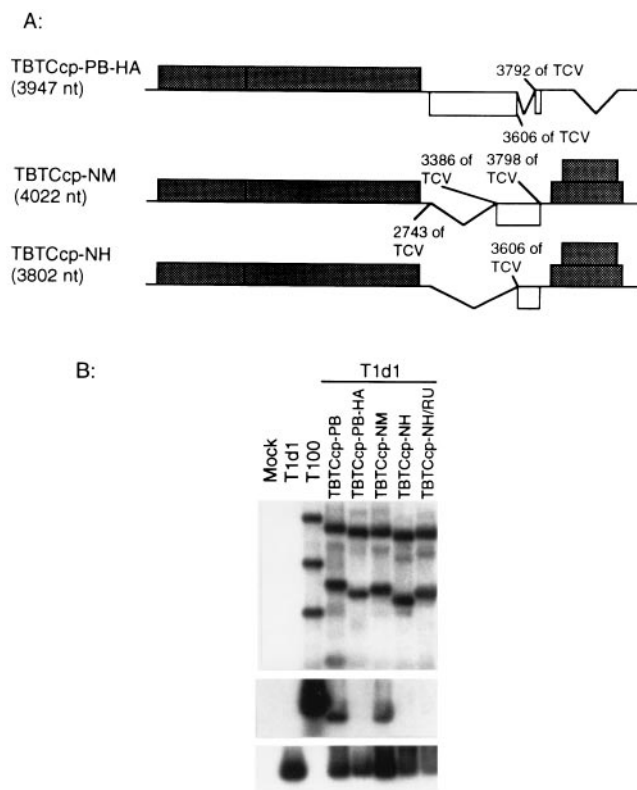


FIG. 7. Effect of the HA fragment on the packaging of the chimeric virus mutants. (A) Schematic representation of deletion mutants of TCVP gene in the chimeric genome background. (B) Northern hybridization analysis of total RNA fractions (top panel) and virus particles (bottom two panels) of protoplast infections processed essentially as described for Fig. 1. The top two panels were analyzed for the accumulation of viral RNAs and viruses by using a TBSV-specific probe. The bottom panel shows the relative accumulation of the helper TCV in the same infections by using a TCV-specific probe.

genomic RNA of TBTCcp-PB (4,133 nt) was the primary species packaged into virus in the TBTCcp-PB/T1d1 coinoculation (lane 4). We conclude from these results that RNA larger than 4,600 nt cannot be assembled into intact TCV virions and that the viruses assembled in the TBTCcp/T1d1 infections, aside from the wild-type TCV, likely consisted of virions that contained two copies of the 2,044-nt sgRNA in order to produce a full particle with an RNA content close to that of wild-type virus (4,088 nt, versus 4,051 for TCV). These results, coupled with the observation that wild-type TCV sgRNAs and sgRNA of TBTCcp-PB (1,512 nt) were not packaged efficiently, suggest that there is a strict RNA size preference for assembly of complete TCV virions.

**The TCV OAS mediates assembly of the chimeric viral genome.** The demonstration that the TCV CP would mediate assembly of the chimeric TBSV RNA provided the opportunity to test the limits of the packaging signal identified in TCV in an unrelated viral RNA genome. To do so, we produced a deletion mutant of the chimeric virus (TBTCcp-PB) in which the HA fragment was deleted (TBTCcp-PB-HA [Fig. 7A]). Protoplast infections of this mutant with helper TCV demonstrated that removal of the HA fragment in the chimeric genome was sufficient to prevent the packaging of the chimeric RNA (Fig. 7B). This experiment showed unambiguously that the same element necessary for specific assembly of native TCV RNA was needed for assembly of the chimeric TBSV genome. Although the same element could be shown to be necessary for

assembly, it was not clear how much of the TCV CP gene sequence would be required to mediate assembly of the foreign viral RNA. To test this, we constructed two chimeric mutants with large deletions in the TCV CP gene: (i) TBTCcp-NM, in which only the 3'-proximal 412 nt of the CP gene remained; and (ii) mutant TBTCcp-NH, in which only the last 192 nt remained (Fig. 7A). Coinoculation of these mutants with TCV demonstrated that TBTCcp-NM was able to assemble whereas TBTCcp-NH was not (Fig. 7B). We ruled out the possibility that the genome of the TBTCcp-NH mutant (3,802 nt) was too small for efficient assembly by inserting an additional 152-nt nonviral fragment (TBTCcp-NH/RU) to increase the RNA size to 3,954 nt. This larger mutant also failed to assemble in the coinoculations. We conclude from these results that the TCV assembly origin is functional in the foreign environment of the TBSV genome but that there might be some secondary folding constraints that necessitate the presence of a larger region of TCV RNA around the OAS for it to function.

## DISCUSSION

We report here the first successful identification of an RNA element essential for the process of assembly initiation for a small icosahedral plant RNA virus of T=3 symmetry. Localization of the essential 28-nt element within a larger 186-nt fragment at the 3' end of the CP gene was assisted by previous *in vitro* studies designed to identify high-affinity CP binding regions on the viral genome (31, 35, 36). The definitive identification of this OAS was not forthcoming, however, until we applied the current assay system, which permitted studying the process of assembly *in vivo*. This approach allowed us to identify the packaging signal and also to demonstrate that the size of the RNA species being packaged is a critical factor in the assembly of intact virus particles.

Structural as well as biochemical analyses have revealed the highly cooperative nature of the RNA-CP interactions for many small spherical plant RNA viruses (8, 12, 31). *In vitro* assembly of virions by using independently generated RNA and CP has been also accomplished in some cases (44). However, a specific RNA packaging signal has not been identified for any of these T=3 icosahedral plant viruses, probably because the *in vitro* assembly systems used so far do not reflect the high level of specificity encountered *in vivo*. Although the *in vitro* approaches have facilitated the identification of RNA packaging signals for TMV and some double-stranded RNA viruses of bacteria and yeast (10; see reference 1 for a review), they have not proved as reliable for similar studies with other icosahedral viruses. For example, *in vitro* studies with bacteriophage R17 identified a specific RNA fragment of 20 nt that was both a packaging signal and a translational operator (3, 28). However, the relative importance of this signal in the encapsidation of RNA *in vitro* was shown to decrease as the length of nonspecific RNA attached to it increased. It was also shown that R17 CP was no more effective at packaging its own RNA *in vitro* than packaging the related Q $\beta$  RNA which lacked the R17 signal (43). Furthermore, it is not clear if the R17 signal would be sufficient to facilitate the packaging of heterologous RNA *in vivo*. These results support our claim that any definitive analysis of assembly specificity must include supportive *in vivo* studies such as the one we described here.

We have now shown that deletion of as little as a 28-nt fragment (mutant TCR-dIVa) from the CP gene is sufficient to abolish TCV RNA packaging. We also demonstrated that the deletion of other similar-size fragments in the vicinity of the 28-nt fragment have quantitative effects on packaging efficiency but do not completely prevent assembly. We conclude

from this demonstration that this 28-nt region must contain the essential TCV packaging signal or OAS. This conclusion is further supported by the observation that some of the defective interfering (DI) RNAs (DI 1) and satellite RNAs (satellite RNA C) associated with TCV infections also contain the OAS region (20). Although we have yet to evaluate the possible role of the OAS sequences in these molecules, it is worth noting that the specific region of the OAS present in satellite RNA C extends from TCV nt 3764 to 3779 and includes only 16 nt that form the top of the structure proposed in Fig. 5B. Although the smaller size of the OAS region in satellite RNA C suggests that the size of the essential sequence could be delineated further, the observation that the 186-nt HA fragment was not sufficient to promote assembly of the foreign TBSV RNA *in trans* leads to the hypothesis that the RNA sequence around the OAS must also be important in stabilizing the functional structure. Further support for this hypothesis came from the fact that NarR, although possessing the OAS, was not packaged. It is also noteworthy that the HA fragment in which the OAS resides overlaps one of the RNase T<sub>1</sub>-protected fragments originally postulated to be involved in assembly initiation (Fd, nt 3586 to 3637 [35]). It seems reasonable that the Fd fragment may well be included in the stabilizing element of the OAS. Unfortunately, when we tested the Fd region for CP binding affinity *in vitro*, we did not include any sequence 3' to Fd and hence did not include the 28-nt OAS identified here (31). Therefore, we did not directly evaluate the 28-nt OAS region in our previous *in vitro* studies.

The results presented here also raise questions about the effect of RNA size on the TCV assembly process. The model of Sorger et al. (32) proposes that assembly would initiate at one site (the OAS) and proceed until the entire genome was packaged. This model predicts that RNAs smaller or larger than the 4,051-nt genome would make incomplete virions or oversized and incompletely closed particles. The presence of sgRNA in virions suggests, however, that assembly of smaller than genome-size RNAs does occur. The location of the OAS suggests that both the 1.7- and 1.45-kb sgRNAs possess a packaging signal. Our unpublished results showed that the 1.7-kb sgRNA is not packaged whereas the 1.45-kb sgRNA is packaged, although less efficiently than genomic RNA (data not shown). This finding suggests that the size of the sgRNA plays an important role: three of the small sgRNA molecules would be close to the packaging capacity of the TCV virion (4.35 kb), while the large sgRNA would be under capacity at two molecules per shell (3.4 kb) and over capacity at three molecules per shell (5.1 kb). Our most definitive evidence for the importance of the RNA size in stable virus assembly comes from the results of the assembly of the two chimeric viruses shown in Fig. 6. The inability of the TBTCcp genomic RNA (4,665 nt) to be packaged coupled with the exclusive assembly of the sgRNA of 2,044 nt in this infection strongly argues that the genomic RNA exceeded the capacity whereas the two sgRNAs (2,044 nt  $\times$  2 = 4,088 nt) were ideally sized to be packaged into a single shell. This conclusion was reinforced with the assembly of the smaller chimeric genome (TBTCcp-PB, 4,133 nt) which was reduced to a size closer to that of native TCV RNA (4,051 nt). These results strongly suggest that the size of the RNA to be packaged is a critical factor in TCV assembly into stable virus. We suggest as well that this observation is probably true for many other similarly structured icosahedral plant viruses that require RNA for shell assembly. Genome size has been shown to be a limiting factor for packaging of other viruses, including bacteriophage  $\phi$ 6, a double-stranded RNA virus (23, 24). It has also been shown that flock house virus, a bipartite insect RNA virus, packages only the DI RNA of a definite size

(45), but the authors did not resolve if the size limit was a function of the DI RNA or the packaging capacity of the specific virions.

Our observations provide the first real evidence that the size of the RNA is as critical as the packaging signal for assembly into stable virus. The suggestion that two half-size RNA molecules, each with a packaging signal, can be assembled into a single stable virus particle implies that the protein-protein interaction between two TCV half shells is sufficient to form a stable virus shell. This possibility expands on the previous model of Sorger et al. (32) in which only a single RNA with one packaging signal would be included in each particle. Although it is possible that two half-size RNAs could interact with each other to form an RNA complex with a combined OAS, this seems less likely. It will be interesting to resolve this question for the structurally similar bipartite genome member of the *Tombusviridae*, red clover necrotic mosaic virus. Vaewhongs and Lommel (33) have presented evidence that both RNAs are needed for assembly of stable virus particles, although they have provided no direct evidence for how this might be achieved. Understanding the mechanism of assembly and the size limits of the RNA that can be assembled into virus will be important in understanding the assembly process and the effect of RNA size on selection and evolution of the many satellite and DI RNAs that have been discovered in these viruses.

#### ACKNOWLEDGMENTS

We thank Roy French, Mark Young, and Steve Lommel for critically reading the manuscript.

We also thank S. C. Harrison for the reminder that the OAS sequence described here was tentatively identified among the RNA fragment sequences provided to us during the preliminary characterization of the TCV rp-complex described by Sorger et al. (32).

This research was supported in part by USDA CRGO grant 93-37303-9033.

#### REFERENCES

- Bamford, D. H., and R. B. Wickner. 1994. Assembly of double-stranded RNA viruses: bacteriophage  $\phi$ 6 and yeast virus L-A. *Semin. Virol.* 5:61-69.
- Bartenschlager, R., and H. Schaller. 1992. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. *EMBO J.* 11:3413-3420.
- Beckett, D., H.-N. Wu, and O. C. Uhlenbeck. 1988. Roles of operator and non-operator RNA sequences in bacteriophage R17 capsid assembly. *J. Mol. Biol.* 204:939-947.
- Carrington, J. C., L. A. Heaton, D. Zuidema, B. I. Hillman, and T. J. Morris. 1989. The genome structure of turnip crinkle virus. *Virology* 170:219-226.
- Cuilliel, M., M. Herzog, and L. Hirth. 1979. Specificity of *in vitro* reconstitution of bromegrass mosaic virus. *Virology* 95:146-153.
- 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.
- Dolja, V. V., R. Haldeman-Cahill, A. E. Montgomery, K. A. Vandebosch, and J. C. Carrington. 1995. Capsid protein determinants involved in cell-to-cell and long distance movement of tobacco etch potyvirus. *Virology* 206:1007-1016.
- Duggal, R., and T. C. Hall. 1993. Identification of domains in brome mosaic virus RNA-1 and coat protein necessary for specific interaction and encapsidation. *J. Virol.* 67:6406-6412.
- Fox, J. M., J. E. Johnson, and M. J. Young. 1994. RNA/protein interactions in icosahedral virus assembly. *Semin. Virol.* 5:51-60.
- Fujimura, T., R. Esteban, L. M. Esteban, and R. B. Wickner. 1990. Portable encapsidation signal of the L-A double-stranded RNA virus of *S. cerevisiae*. *Cell* 62:819-828.
- Hacker, D. L., I. T. D. Petty, N. Wei, and T. J. Morris. 1992. Turnip crinkle virus genes required for RNA replication and virus movement. *Virology* 186:1-8.
- Harrison, S. C. 1990. Common features in the structures of some icosahedral virus: a partly historical overview. *Semin. Virol.* 1:387-403.
- Harrison, S. C., A. J. Olson, C. E. Schutt, F. K. Winkler, and C. Bricogne. 1978. Tomato bushy stunt virus at 2.9 Å resolution. *Nature* 276:368-373.
- Hearne, P. Q., D. A. Knorr, B. I. Hillman, and T. J. Morris. 1990. The complete genome structure and synthesis of infectious RNA from clones of

- tomato bushy stunt virus. *Virology* **177**:141–151.
15. **Heaton, L. A.** 1992. Use of agarose gel electrophoresis to monitor conformational changes of some small, spherical plant virus. *Phytopathology* **82**:803–807.
  16. **Heaton, L. A., J. C. Carrington, and T. J. Morris.** 1989. Turnip crinkle virus infection from RNA synthesized in vitro. *Virology* **170**:214–218.
  17. **Hogle, J. M., A. Maeda, and S. C. Harrison.** 1986. Structure and assembly of turnip crinkle virus. I. X-ray crystallographic structure analysis at 3.2 Å resolution. *J. Mol. Biol.* **191**:625–638.
  18. **Jones, R. W., A. O. Jackson, and T. J. Morris.** 1990. Defective-interfering RNAs and elevated temperatures inhibit replication of tomato bushy stunt virus in inoculated protoplasts. *Virology* **176**:539–545.
  19. **Junker-Niepmann, M., R. Bartenschlager, and H. Schaller.** 1990. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. *EMBO J.* **9**:3389–3396.
  - 19a. **Li, W. Z.** 1996. Gene expression of turnip crinkle virus and characterization of the two movement genes. Ph.D. thesis. University of Nebraska–Lincoln, Lincoln, Neb.
  20. **Li, X. H., L. A. Heaton, T. J. Morris, and A. E. Simon.** 1989. Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated de novo. *Proc. Natl. Acad. Sci. USA* **86**:9173–9177.
  21. **Ling, C. M., P. P. Hung, and L. R. Overby.** 1970. Interdependent assembly of  $\phi$ 8 and MS2 phages in doubly infected *Escherichia coli*. *Virology* **40**:920–929.
  22. **Linial, M. L., and A. D. Miller.** 1990. Retroviral RNA packaging: sequence requirements and implication. *Curr. Top. Microbiol. Immunol.* **157**:125–152.
  23. **Mindich, L., X. Qiao, and J. Qiao.** 1995. Packaging of multiple copies of reduced-size genomic segment by bacteriophage  $\phi$ 6. *Virology* **212**:213–217.
  24. **Onodera, S., X. Qiao, J. Qiao, and L. Mindich.** 1995. Acquisition of a fourth genomic segment in bacteriophage  $\phi$ 6, a bacteriophage with a genome of three segments of dsRNA. *Virology* **212**:204–212.
  25. **Pollack, J. R., and D. Ganem.** 1994. Site-specific RNA binding by a hepatitis B virus reverse transcriptase initiates two distinct reactions: RNA packaging and DNA synthesis. *J. Virol.* **68**:5579–5587.
  26. **Qu, F., W. Zhai, H. Chen, L.-H. Zhu, and T. J. Morris.** 1996. Cloning, characterization and transient expression of the gene encoding a rice U3 small nuclear RNA. *Gene* **172**:217–220.
  27. **Rochon, D'A., and A. Siegel.** 1984. Chloroplast DNA transcripts are encapsidated by tobacco mosaic virus coat protein. *Proc. Natl. Acad. Sci. USA* **81**:1719–1723.
  28. **Romaniuk, P. J., P. Lowary, H.-N. Wu, G. Stormo, and O. C. Uhlenbeck.** 1987. RNA binding site of R17 coat protein. *Biochemistry* **26**:1563–1568.
  29. **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.
  30. **Scholthof, H. B., T. J. Morris, and A. O. Jackson.** 1993. The capsid protein gene of tomato bushy stunt virus is dispensable for systemic movement and can be replaced for localized expression of foreign genes. *Mol. Plant-Microbe Interact.* **6**:309–322.
  31. **Skuzeski, J. M., and T. J. Morris.** 1995. Quantitative analysis of the binding of turnip crinkle virus coat protein to RNA fails to demonstrate binding specificity but reveals a highly cooperative assembly interaction. *Virology* **210**:82–90.
  32. **Sorger, P. K., P. G. Stockley, and S. C. Harrison.** 1986. Structure and assembly of turnip crinkle virus. II. Mechanism of in vitro assembly. *J. Mol. Biol.* **191**:375–383.
  33. **Vaewhongs, A. A., and S. A. Lommel.** 1995. Virion formation is required for the long-distance movement of red clover necrotic mosaic virus in movement protein transgenic plants. *Virology* **212**:607–613.
  34. **Van der Vossen, E. A. G., L. Neeleman, and J. F. Bol.** 1994. Early and late functions of alfalfa mosaic virus coat protein can be mutated separately. *Virology* **202**:891–903.
  35. **Wei, N., L. A. Heaton, T. J. Morris, and S. C. Harrison.** 1990. Structure and assembly of turnip crinkle virus. VI. Identification of coat protein binding sites on the RNA. *J. Mol. Biol.* **214**:85–95.
  36. **Wei, N., and T. J. Morris.** 1991. Interactions between viral coat protein and a specific binding region on turnip crinkle virus. *J. Mol. Biol.* **222**:437–443.
  37. **Wei, N., D. L. Hacker, and T. J. Morris.** 1992. Characterization of an internal element in TCV RNA involved in both coat protein binding and replication. *Virology* **190**:346–355.
  38. **Weiss, B., H. Nitschko, I. Ghattas, R. Wright, and S. Schlesinger.** 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. *J. Virol.* **63**:5310–5318.
  39. **Weiss, B., U. Geigenmueller-Gnirke, and S. Schlesinger.** 1994. Interactions between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site. *Nucleic Acids Res.* **22**:780–786.
  40. **White, K. A., and T. J. Morris.** 1994. Nonhomologous RNA recombination in tombusviruses: generation and evolution of defective interfering RNAs by stepwise deletions. *J. Virol.* **68**:14–24.
  41. **White, K. A., and T. J. Morris.** 1995. RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. *RNA* **1**:1029–1040.
  42. **White, K. A., J. M. Skuzeski, W. Li, N. Wei, and T. J. Morris.** 1995. Immunodetection, expression strategy and complementation of turnip crinkle virus P28 and P88 replication components. *Virology* **211**:525–534.
  43. **Witherell, G. W., J. M. Gott, and O. C. Uhlenbeck.** 1991. Specific interaction between RNA phage coat protein and RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **40**:185–220.
  44. **Zhao, X., J. M. Fox, N. H. Olson, T. S. Baker, and M. J. Young.** 1995. In vitro assembly of cowpea chlorotic mottle virus form coat protein expressed in *Escherichia coli* and in vitro transcribed viral cDNA. *Virology* **207**:486–494.
  45. **Zhong, W., R. Dasgupta, and R. Rueckert.** 1992. Evidence that the packaging signal for nodaviral RNA2 is a bulged stem-loop. *Proc. Natl. Acad. Sci. USA* **89**:11146–11150.
  46. **Zimmern, D., and P. J. G. Butler.** 1977. The isolation of tobacco mosaic virus RNA fragments containing the origin for viral assembly. *Cell* **11**:455–462.
  47. **Zimmern, D.** 1977. The nucleotide sequence at the origin for assembly on tobacco mosaic virus RNA. *Cell* **11**:463–482.