JOURNAL OF VIROLOGY, May 1992, p. 3069–3076
0022-538X/92/053069-08$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 66, No. 5

# Coding Capacity Determines In Vivo Accumulation of a Defective RNA of Clover Yellow Mosaic Virus

K. ANDREW WHITE,[1] J. B. BANCROFT,[2] AND GEORGE A. MACKIE[1]*

*Departments of Biochemistry[1] and Plant Sciences,[2] University of Western Ontario, London, Ontario, Canada N6A 5C1*

Received 5 December 1991/Accepted 6 February 1992

**Naturally occurring defective RNAs (D RNAs) derived from the potexvirus clover yellow mosaic virus (CYMV) contain large internal deletions yet maintain a single open reading frame (ORF) representing the in-frame fusion of 5′ and 3′ terminal ORFs. Capped transcripts of the prototype 1.2-kb D RNA of CYMV were synthesized in vitro and used to inoculate broad bean plants. Progeny D RNA accumulated only if synthetic D RNA transcripts were coinoculated with CYMV RNA. Several experiments showed that helper-dependent accumulation of the D RNA in vivo depended on the maintenance of its encoded fusion ORF. (i) D RNAs with six-residue deletions introduced early in the fusion ORF accumulated, whereas those with four-residue out-of-frame deletions at the same sites were nonviable. (ii) Analysis of D RNAs containing termination codons at different locations showed that only the most 3′ stop codon (maintaining over 93% of the fusion ORF) was permissive for D RNA accumulation. (iii) D RNAs with small in-frame deletions and insertions in their 3′ coding regions were viable. (iv) Nonviable D RNAs containing disrupted fusion ORFs could not be complemented by the presence in the infection of a D RNA encoding a complete fusion ORF. Taken together, the results indicate that the process of translation, rather than the encoded product, modulates an event(s) which influences the propagation and/or accumulation of this RNA in vivo. This represents a unique requirement among plant virus D RNAs.**

Defective RNAs (D RNAs) contain portions of the parental virus genome but are incapable of self-replication. These RNAs do, however, maintain *cis*-acting elements which contain viral replication and encapsidation signals. As a result, D RNAs accumulate only in mixed infections with nondefective helper virus when essential components are supplied in *trans* (see reference 21 for a review). Clover yellow mosaic virus (CYMV), which contains a single positive-sense genomic RNA (gRNA) is the only potexvirus (potato virus X group virus) with which D RNAs are associated (25). Four different D RNAs containing only the 5′- and 3′-terminal regions of the CYMV 7.0-kb gRNA have been characterized (25). The prototype D RNA is 1,172 nucleotides long [excluding a poly(A) tail] and is composed of 757 nucleotides of the 5′ terminus and 415 nucleotides of the 3′ terminus of CYMV gRNA. Like CYMV gRNA and subgenomic RNAs (22, 26), the D RNAs contain a 5′ cap structure and a 3′ poly(A) tail and are encapsidated by CYMV coat protein (25). Interestingly, symptom development in CYMV-infected broad bean plants is not altered by the presence of these RNAs. This feature distinguishes CYMV D RNAs from plant virus defective interfering RNAs (DI RNAs), which can have a similar primary structure but are able to increase (12) or decrease (5, 10) the severity of symptoms.

In contrast to the DI RNAs described for other plant viruses, the CYMV D RNAs possess considerable coding capacity and are found in polyribosomes (25). Polypeptides encoded by CYMV D RNAs contain the N terminus of the 191-kDa nonstructural protein (encoded by open reading frame 1 [ORF1]; the putative replicase) fused in frame to the C terminus of the 23-kDa coat protein (encoded by ORF5), as shown for the prototype D RNA in Fig. 1. The coding strategy of the CYMV D RNAs closely resembles that of the DI RNAs of mouse hepatitis coronavirus (15, 16, 23) and poliovirus (9), in which junctions in the DI RNA are fused in frame to maintain a single ORF. All four species of CYMV D RNA examined encode an ORF1-ORF5 fusion product (25). We have proposed that conservation of this fusion ORF may be important for survival of this RNA within the plant (25).

We have sought to define the structural features of a CYMV D RNA which are necessary for its accumulation in vivo. A full-length cDNA clone of the prototype D RNA from which biologically active RNA could be transcribed in vitro was constructed (25). Modified synthetic transcripts were used to determine the effects of 5′ capping, internal deletions, and reduced coding capacity on the viability of this RNA in vivo.

## MATERIALS AND METHODS

**Preparation of virus and RNA.** Virus and RNA were extracted from the upper (uninoculated) leaves of broad bean (*Vicia faba*) plants 2 weeks postinoculation. Virus, virion RNA and polyribosomal RNA were purified as previously described (25). Total RNA was prepared by grinding 2 g of tissue in liquid nitrogen with a mortar and pestle. The powdered tissue was stirred thoroughly with 9 ml of a buffer containing 50 mM Tris-Cl (pH 8.8), 1% sodium dodecyl sulfate, and 200 mM EDTA and extracted twice with 9 ml of $H_2O$-saturated phenol. Following a further extraction with phenol-chloroform-isoamyl alcohol, the nucleic acids were twice precipitated with ethanol. RNAs were separated electrophoretically in agarose gels (17) following treatment with 10 U of RNase-free DNase I (Pharmacia) for 20 min at 37°C. RNAs were transferred to nylon and hybridized with 5′ $^{32}P$-labeled oligonucleotides as previously described (14), except for exclusion of formamide from the annealing buffer. Hybridization and washing of the blots were performed at ambient temperature.

**Plasmid construction and cDNA synthesis.** The strategy

* Corresponding author.

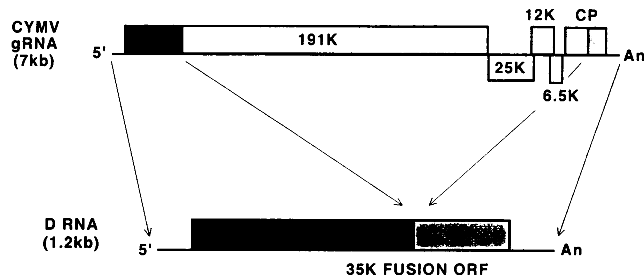


FIG. 1. Diagram of the general structure of the CYMV prototype 1.2-kb defective RNA (25). The organization of coding regions on the 7-kb gRNA of CYMV is shown with the approximate $M_r$ values of the encoded proteins (22). Coding regions from which the 1.2-kb D RNA (shown below) is derived are shaded. ORF1 encodes the putative replicase, and CP denotes the 23-kDa coat protein. An denotes the poly(A) tract. Reprinted from *Virology* (25) with permission of Academic Press, Inc., Orlando, Fla.

used to clone a biologically active cDNA (pAW1) of the prototype CYMV D RNA by using reverse transcription (RT)-polymerase chain reaction (PCR) has been described previously (25). Basically, oligonucleotides containing sequences identical and complementary, respectively, to the 5′ and 3′ termini of CYMV gRNA were designed to include a T7 RNA polymerase promoter and a poly(dT) tract, respectively. These oligonucleotides were used for RT-PCR of RNA preparations containing D RNA, the product of which was cloned (pAW1). By using this construct, synthetic D RNAs which contained authentic capped 5′ and polyadenylated 3′ termini could be produced in vitro (Fig. 2A).

The synthetic oligonucleotides used in this study are listed in Table 1. All coordinates given correspond to the sequence of the CYMV prototype D RNA (25). pAW1-M was produced by linearization of pAW1 at a unique *Ava*I site (position 933), followed by insertion of a short double-stranded DNA composed of annealed CY-16 and CY-17 (Fig. 2B).

Constructs pAW1-MΔ1 and pAW1-MΔ2, containing four-residue deletions (residues 280 to 283 and 463 to 466, respectively), were produced by linearization of pAW1-M at a unique *Pst*I (position 279) or *Bst*XI (position 459) site, followed by blunt ending with T4 DNA polymerase and religation. Six-residue deletions in pAW1-MΔ3 (residues 278 to 283) and pAW1-MΔ4 (residues 461 to 466) were introduced by oligonucleotide-directed mutagenesis with oligonucleotides CY-20 and CY-21, respectively, in accordance with the protocol of the Bio-Rad Mutagene Kit (11, 18).

Large internal deletions in pAW1-M (pAW1-MΔ5 to pAW1-MΔ10) were created by digestion of pAW1-M with restriction enzymes (*Pst*I, position 279; *Bst*XI, position 459; *Stu*I, position 853; or *Ava*I, position 933) in different combinations of two, followed by blunt ending with T4 DNA polymerase and religation (see also Fig. 6A). Internal deletions produced by using restriction endonuclease *Ava*I (pAW1-MΔ8 to pAW1-MΔ10) also removed a portion of the inserted oligonucleotide marker because of the presence of an *Ava*I site in it (underlined in CY-16 and CY-17; Table 1).

Additional internal stop mutations were introduced into pAW1-M by using oligonucleotide-directed mutagenesis. For all constructs (pAW1-MS1 to pAW1-MS6), in-frame ochre stop codons were created through substitution of thymine for the base 5′ to two adenosines (pAW1-MS1, position 317 [G→T]; pAW1-MS2, position 521 [A→T];

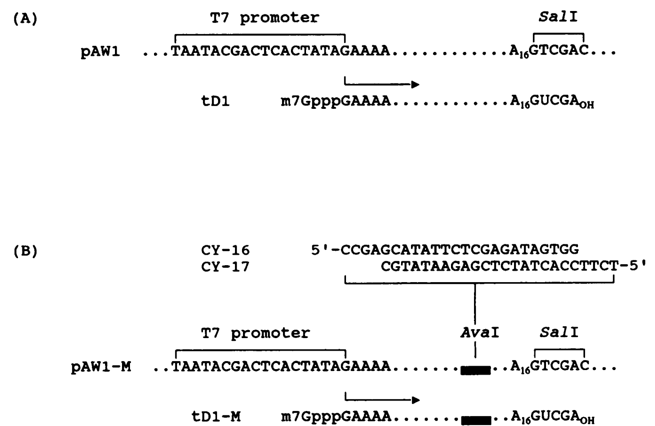


FIG. 2. Partial nucleotide sequences of plasmids containing a cloned cDNA of the prototype 1.2-kb D RNA and partial sequences of corresponding synthetic RNA transcripts. The cDNA of the prototype 1.2-kb D RNA was amplified by PCR by using primers designed to introduce an upstream T7 RNA polymerase promoter and maintain a 3′ poly(A) tract (see Materials and Methods). The position of the *Sal*I restriction site used to linearize the constructs before in vitro transcription is shown. Transcription of these constructs with T7 RNA polymerase initiates at the authentic 5′ nucleotide of native 1.2-kb D RNA, and partial 5′ and 3′ sequences of capped transcripts derived from the two constructs are shown beneath. (A) Construct pAW1, showing 5′ and 3′ regions of the cDNA of the prototype 1.2-kb D RNA. The synthetic RNA transcript produced from pAW1, tD1, is shown below. (B) Construct pAW1-M, containing cloned cDNA of the prototype 1.2-kb D RNA with an in-frame double-stranded oligonucleotide insert (composed of annealed CY-16 and CY-17) at the *Ava*I site (position 933). The synthetic RNA transcript derived from pAW1-M, tD1-M, is shown below.

pAW1-MS3, position 635 [G→T]; pAW1-MS4, position 755 [G→T]; pAW1-MS5, position 866 [G→T]; pAW1-MS6, position 974 [A→T]). The respective oligonucleotides used to construct pAW1-MS1 through pAW1-MS6 were CY-22 through CY-27. All mutations were confirmed by dideoxynucleotide chain termination sequencing of the modified region with T7 DNA polymerase (Pharmacia).

The 3′ region of progeny 1.2-kb D RNAs was cloned by using two oligonucleotides, CY-17 and CY-18, for its RT-PCR amplification. RT-PCR was performed under conditions previously described (25), except for the annealing temperature, which was increased from 50 to 65°C.

**In vitro transcription.** In vitro transcription was carried out on *Sal*I-digested pAW1 or its derivatives by using T7 RNA polymerase (Pharmacia). A standard reaction contained 10 μg of the linearized template; 40 mM Tris-HCl (pH 7.5); 10 mM NaCl; 6 mM $MgCl_2$; 5 mM dithiothreitol; 2 mM spermidine; 0.8 U of RNAguard (Pharmacia) per μl; ATP, CTP, and UTP, each at 0.5 mM; 0.02 mM GTP; 0.5 mM $m^7$GpppG (New England BioLabs, Inc.); and 0.7 U of T7 RNA polymerase per μl in a final volume of 200 μl. Following incubation at 37°C for 15 min, the GTP concentration was raised to 0.27 mM and incubation at 37°C was resumed for 2 h. Additional T7 RNA polymerase (70 U) was added midway through this incubation. The template was removed by digestion with 0.1 U of RNase free-DNase I (Pharmacia) per μl at 37°C for 20 min. Transcripts lacking a cap structure were produced as described above, except that $m^7$GpppG was replaced by 0.5 mM GTP in the initial reaction and no further GTP was added. Reactions were terminated by

TABLE 1. Synthetic oligonucleotides used in this study

| Oligonucleotide | Sequence | Binding site[a] |
|---|---|---|
| CY-6 | 5'-GAAAACAAAACGAAAACAA | 1–19 (+) |
| CY-16 | 5'-CCGAGCATATTCTCGAGATAGTGG[b] | NA |
| CY-17 | 5'-TCGGCCACTATCTCGAGAATATGC[b] | NA |
| CY-18 | 5'-GCCTGCGCGAATTCTTTTTTTTTTTTTT | Poly(A) tail (−) |
| CY-20 | 5'-GCAACTAACCCATTCGAACCACACACACAT | 263–298 (+) |
| CY-21 | 5'-CAAATTGTGGAGCCCGTGGCGAGATACGAC | 446–481 (+) |
| CY-22 | 5'-AAAGCTATATAAAACGAC | 308–325 (+) |
| CY-23 | 5'-ATAGAGACTTAAACAGTC | 512–529 (+) |
| CY-24 | 5'-TACCCACGTAAGCACTAT | 627–644 (+) |
| CY-25 | 5'-AAGGAGCTTAAAACTGGG | 747–764 (+) |
| CY-26 | 5'-CTAATCCGATAACCAAGC | 857–874 (+) |
| CY-27 | 5'-TAGCCACCTAAGGAGCTT | 966–983 (+) |

[a] Numbers correspond to coordinates of the CYMV prototype D RNA (25). The oligonucleotides are of either the same (+) or opposite (−) polarity with respect to the D RNA. NA, not applicable.
[b] The underlined sequence denotes an *Ava*I site.

extraction with phenol-chloroform-isoamyl alcohol, and the products were precipitated with ethanol. The quantity and quality of synthetic 1.2-kb D RNA produced were determined relative to standards on polyacrylamide gels stained with ethidium bromide.

**Inoculation of plants.** Each transcript was inoculated by using a glass spatula on 2-week-old broad bean plants dusted with aluminum oxide (600 grit). A typical inoculum contained 0.05 μg of CYMV RNA (free of native 1.2-kb D RNAs) per μl; 0.05 or 0.15 μg of a synthetic 1.2-kb transcript per μl, and 100 mM Na phosphate ($Na_2HPO_4$-$NaH_2PO_4$, pH 7.0) in a volume of 30 μl.

## RESULTS

**Biological activity of synthetic D RNAs.** We previously constructed a plasmid (pAW1) containing a cloned cDNA of the CYMV prototype 1.2-kb D RNA downstream of a T7 RNA polymerase promoter to allow for in vitro production of synthetic 1.2-kb D RNA (25). The transcript produced, termed tD1 (Fig. 2A), contains the authentic capped 5' terminus of the prototype 1.2-kb D RNA because of the design of the 5' primer used for PCR amplification of the cDNA and the inclusion of $m^7GpppG$ in the in vitro transcription reaction. The 3' terminus of the transcript contains a poly(A) tract of 16 residues followed by an additional five ribonucleotides contributed by the *Sal*I restriction site used to linearize the DNA template (Fig. 2A). By using T7 RNA polymerase, approximately 1.0 to 1.5 μg of a full-length 1.2-kb transcript was synthesized in vitro per μg of the linear DNA template.

Capped tD1 was inoculated singly or with CYMV RNA on broad bean leaves. The total CYMV virion RNA used in the coinoculations as a helper was determined to be free of native D RNA by both Northern (RNA) blotting and RT-PCR (data not shown). Polyribosomal RNA was purified from upper leaves harvested from control or infected plants and analyzed by Northern blotting (Fig. 3). No CYMV-specific RNAs were detectable in uninfected plants (Fig. 3, lane 1), whereas CYMV RNA inoculated alone resulted in production of gRNA [approximately 7.2 kb, including the poly(A) tail] and the 1.0-kb coat protein message, as expected (Fig. 3, lane 2). When tD1 was coinoculated with CYMV RNA, we detected an additional abundant product with a relative mobility of 1.2 kb which was not present when CYMV RNA was inoculated alone (Fig. 3, compare lane 3 with lane 2) or when tD1 was inoculated without the helper (Fig. 3, lane 4). The presence of tD1 progeny in RNA extracted from polyribosomes indicates that they are translated in vivo. For comparison, an aliquot of tD1 used as the inoculum in this experiment was separated in the same gel (Fig. 3, lane 5). The difference in mobility between the progeny D RNA and tD1 suggests that the progeny are modified in vivo in some manner which increases their relative molecular mass (see below). These results indicate

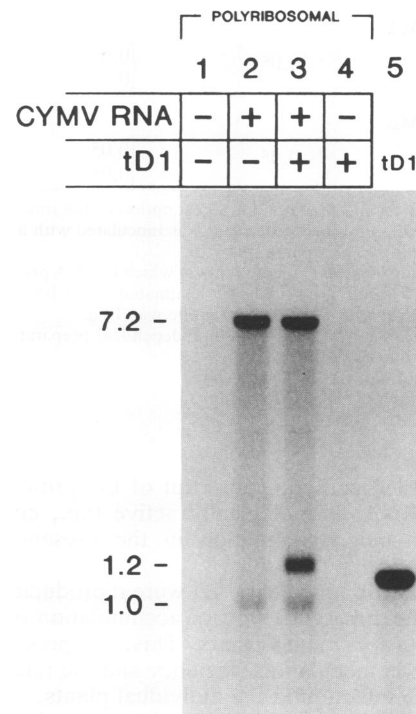


FIG. 3. Accumulation of progeny CYMV D RNAs in broad bean plants from inoculations containing the tD1 synthetic transcript. Total polyribosomal RNAs (5 μg) isolated from infected or uninfected plants were denatured, resolved by electrophoresis, and blotted onto nylon. The blot was probed with an oligonucleotide (CY-1) complementary to the 23 residues immediately 5' to the poly(A) tail of CYMV gRNA. RNAs which were present in (+) or absent from (−) the inoculum are shown above lanes 1 to 4. Lane 5 contained 0.1 μg of tD1.

TABLE 2. Biological activity of CYMV D RNAs

| Transcript[a] | Inoculum[b] | | In vivo accumulation of D RNA[c] |
|---|---|---|---|
| | CYMV RNA (μg/μl) | Synthetic D RNA (μg/μl) | |
| tD1 | | | |
| CAP | 0.01 | 0.1 | 2/2 |
| CAP | 0.01 | 0.1[d] | 1/1 |
| tD1-M | | | |
| CAP | 0.05 | 0.05 | 3/3 |
| CAP | 0.05 | 0.15 | 2/2 |
| CAP | 0.05 | 0.05[d] | 4/4 |
| No CAP | 0.05 | 0.05[d] | 0/4 |
| tD1-MΔ1 | | | |
| CAP | 0.05 | 0.05 | 0/4 |
| CAP | 0.05 | 0.15 | 0/2 |
| tD1-MΔ2 | | | |
| CAP | 0.05 | 0.05 | 0/4 |
| CAP | 0.05 | 0.15 | 0/2 |
| tD1-MΔ3 | | | |
| CAP | 0.05 | 0.05 | 3/4 |
| CAP | 0.05 | 0.15 | 2/2 |
| tD1-MΔ4 | | | |
| CAP | 0.05 | 0.05 | 2/4 |
| CAP | 0.05 | 0.15 | 1/2 |
| tD1 + tD1-MΔ1 | | | |
| CAP | 0.05 | {0.05<br>0.05 | {4/4[e]<br>0/4[f] |
| tD1 + tD1-MΔ2 | | | |
| CAP | 0.05 | {0.05<br>0.05 | {4/4[e]<br>0/4[g] |

[a] See Materials and Methods for a description of the transcripts.
[b] Each 2-week-old broad bean plant was inoculated with a total volume of 30 μl in 100 mM Na phosphate (pH 7.0).
[c] Values represent numbers of plants in which D RNA progeny accumulated/numbers of plants inoculated. The estimated lower limit of detection of progeny D RNAs with CY-17 as the probe was 1 ng.
[d] Synthetic D RNA derived from an independent preparation.
[e] Accumulation of tD1 progeny.
[f] Accumulation of tD1-MΔ1 progeny.
[g] Accumulation of tD1-MΔ2 progeny.

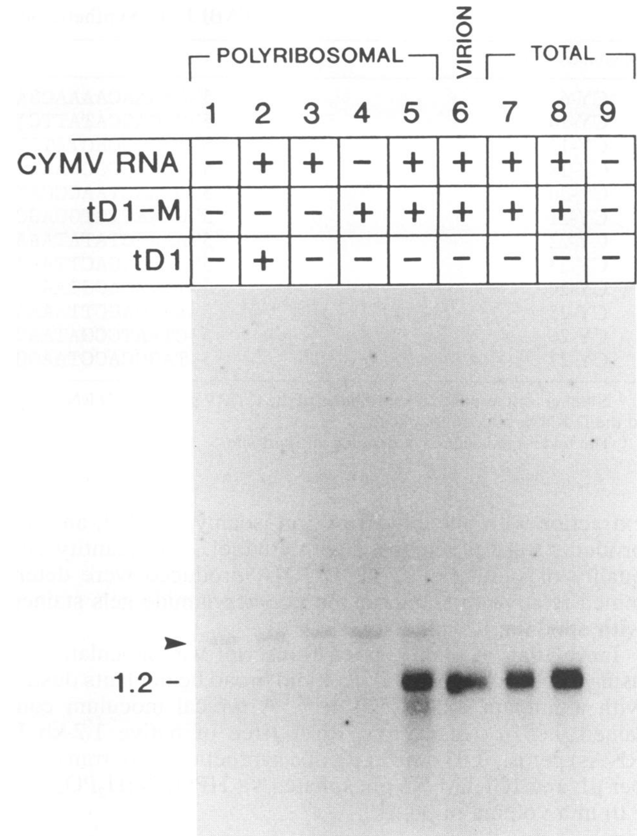


FIG. 4. Accumulation of CYMV progeny D RNAs in broad bean plants from inoculations containing the tD1-M synthetic transcript. Polyribosomal (10 μg; lanes 1 to 5), virion (0.5 μg; lane 6), or total (20 μg; lanes 7 to 9) RNAs isolated from infected or uninfected plants were denatured, resolved by electrophoresis, and blotted onto nylon. The blot was probed with an oligonucleotide (CY-17) complementary to an inserted marker sequence present in tD1-M (Fig. 2B). The class of RNA separated in each lane is indicated at the top, and RNAs which were present in (+) or absent from (−) the inoculum are shown for lanes 1 to 9. Lane 7 contained total RNA from a primary infection, and lane 8 contained total RNA from the first passage using sap from the primary infection. The arrowhead indicates 18S rRNA.

that tD1, the synthetic transcript of the prototype CYMV 1.2-kb D RNA, is biologically active (i.e., competent for replication and accumulation in the presence of helper functions).

The results obtained with tD1 were reproducible (Table 2), although the apparent levels of accumulation of progeny D RNAs did vary among plants. This, we presume, reflects differences in inoculation efficiency and the rate of development of the infection in the individual plants.

To preclude the possibility that the accumulation of 1.2-kb D RNA in vivo resulted from contamination of the helper RNA or through de novo production, a second transcript was designed. It included an insertion of a synthetic 24-residue double-stranded oligonucleotide into the *Ava*I site (position 933) present in the 3′ coding region of the prototype D RNA in pAW1. The insertion did not disrupt the reading frame of the fusion ORF, and we named the resulting plasmid pAW1-M and its transcript tD1-M (Fig. 2B). Progeny from tD1-M could now be unambiguously identified in RNA samples by probing with a labeled oligonucleotide (CY-17; Fig. 2B) complementary to the inserted region. Inoculations were carried out by using tD1-M, and various RNA preparations from infected plant leaves were examined by Northern blotting by using labeled CY-17 as the probe (Fig. 4). tD1-M was not detected in polyribosomal RNA fractions from uninoculated plants, from plants inoculated with CYMV RNA and tD1, or from single inoculations with CYMV RNA or tD1-M (Fig. 4, lanes 1 to 4, respectively). In contrast, an RNA of approximately 1.2 kb was identified in the polyribosomal, virion, and total RNA fractions with the tD1-M-specific probe when tD1-M was coinoculated with CYMV RNA (Fig. 4, lanes 5 to 7, respectively). The presence of tD1-M progeny in polyribosomes indicates that they are translated in vivo, as are naturally occurring CYMV D RNAs (25) and tD1 progeny (Fig. 3, lane 3). tD1-M was also stable during serial passages using sap, and a preparation of total RNA from the first such passage is shown in lane 8 of Fig. 4. These results prove that the synthetic transcript

is the progenitor of the accumulated 1.2-kb product and confirm its biological activity. It is also apparent that an in-frame 24-residue insertion at position 933 of the D RNA does not impair its ability to be encapsidated or accumulate.

Capped tD1-M consistently accumulated when inoculated at concentrations as low as 0.05 μg/μl in the presence of 0.05 μg of CYMV RNA per μl (approximately a 6-to-1 molar ratio of D RNA to gRNA; Table 2). tD1-M, however, did not accumulate when the transcript was not capped (Table 2).

**Sequence of the 3′ region of progeny D RNA.** The size of the 1.2-kb D RNAs produced in vivo was found to be approximately 60 to 80 residues longer than tD1 in the inoculum (Fig. 3, compare lane 3 with lane 5). Since it has been demonstrated that CYMV gRNA contains a poly(A) tail of 75 to 100 residues (1), we hypothesized that the increased size of the in vivo product was due to an increase in poly(A) tail length. To test this idea, the 3′ region of progeny 1.2-kb D RNAs was amplified by RT-PCR. Total RNA extracted from plants coinoculated with CYMV RNA and tD1-M served as the template for cDNA synthesis primed with CY-18 (Table 1). The poly(dT) tract in this primer allowed it to anneal to various regions of the poly(A) tail of tD1-M-derived D RNAs (as well as to other polyadenylated messages). Subsequent PCR amplification was carried out with an additional primer (CY-17) specific for the inserted region of tD1-M, which allowed for precise targeting of only tD1-M-derived D RNAs. The annealing temperature during the PCR reaction was 65°C to confer maximum specificity to the non-poly(dT) region of CY-18. The length and composition of the poly(A) tail was determined in six cloned cDNAs derived from PCR amplification. The sequence revealed poly(A) tracts of 14, 39, 40, 42, 50, and 67 residues (data not shown), suggesting relatively random hybridization of CY-18 to the poly(A) tail during cDNA synthesis. In general, the average length of these tails is consistent with the size difference between the synthetic 1.2-kb transcript and its in vivo-produced counterpart. Interestingly, the additional 5 residues present 3′ to the 16 poly(A) residues in tD1-M (Fig. 2B) were not maintained in any of the progeny analyzed.

As a further test of the fidelity with which tD1-M was replicated, the region 5′ to, and including, the oligonucleotide marker in the progeny of tD1-M was also amplified by RT-PCR (by using oligonucleotides CY-6 and CY-16 [Table 1]) and the product was compared with the similar region amplified from tD1-M produced in vitro. The two products exhibited equivalent electrophoretic mobilities when separated in an agarose gel (data not shown), indicating that there was no significant modification of tD1-M progeny 5′ to the insertion.

**D RNAs containing disrupted fusion ORFs.** We have previously reported the conservation of a single fusion ORF (ORF1-ORF5) in all native CYMV 1.2-kb D RNAs examined and have proposed that this coding organization may play an important role in the survival of this species of RNA within plants (25). We tested this concept by constructing tD1-M transcripts containing four-residue deletions in the 5′ region of the fusion ORF which cause a shift in the reading frame and premature termination at the first stop codon in the new reading frame. tD1-MΔ1 and tD1-MΔ2 contain four-residue deletions (coordinates 280 to 283 and 463 to 466, respectively) at *Pst*I (position 279) and *Bst*XI (position 459) restriction sites, and their encoded ORFs are predicted to terminate at stop codons in the new reading frame 32 and 47 residues, respectively, downstream from the deletions. The biological activities of tD1-MΔ1 and tD1-MΔ2 were tested under conditions which consistently gave positive results for tD1-M (Table 2). Progeny RNAs from either tD1-MΔ1 or tD1-MΔ2 were undetectable in total RNAs extracted from plants coinfected with CYMV RNA (Table 2), indicating that these RNAs, which contain disrupted fusion ORFs, did not accumulate under these conditions. Mutant transcripts tD1-MΔ3 and tD1-MΔ4 contain six-residue deletions (coordinates 278 to 283 and 461 to 466, respectively) which maintain a fused ORF and lack virtually the same region as in tD1-MΔ1 and tD1-MΔ2, respectively. When coinoculated with CYMV RNA, tD1-MΔ3 and tD1-MΔ4 were amplified in several of the plants tested (Table 2), suggesting that deletion of similar regions in tD1-MΔ1 and tD1-MΔ2 had the primary effect of reducing coding capacity rather than altering a critical *cis*-acting element. RT-PCR and sequencing of several progeny of tD1-MΔ3 and tD1-MΔ4 confirmed that the descendants did maintain their respective six-residue deletions (data not shown). Premature termination of the fusion ORF, therefore, appears to be detrimental to accumulation of the D RNA.

The ORF1-ORF5 fusion product encoded by the 1.2-kb D RNA may play a direct role in the accumulation of its message. Accordingly, we coinoculated plants with either tD1-MΔ1 or tD1-MΔ2 along with CYMV RNA and tD1. If only the full-length fusion product were capable of acting in *trans* to allow the D RNA to accumulate, then tD1-MΔ1 and tD1-MΔ2 (containing disrupted fusion ORFs) should accumulate in the presence of tD1 (containing a complete fusion ORF). Although tD1 did accumulate to detectable levels in all of the plants coinoculated (Table 2), its presence could not rescue either tD1-MΔ1 or tD1-MΔ2 (Table 2), suggesting that the fusion product does not act in *trans* to allow D RNA accumulation. The accumulation of tD1 in these experiments also serves as an internal control for inoculation efficiency and further verifies the inviability of tD1-MΔ1 and tD1-MΔ2.

To determine the extent of the fusion ORF which must be maintained in the D RNA for it to accumulate in vivo, we synthesized a series of modified transcripts (tD1-MS1 through tD1-MS6) containing termination codons introduced at various positions along the fusion ORF (Fig. 5A). The termination codons were introduced into pAW1-M (tD1-M) by single-base substitutions to minimize the chances of interference with *cis*-acting elements. The mutant transcripts, therefore, represent full-length D RNAs differing only in the 3′ extent of the encoded ORF (Fig. 5A). Synthetic transcripts were coinoculated with CYMV RNA, and total RNA from the plants was analyzed by Northern blotting for the accumulation of progeny D RNA. Each mutant transcript (tD1-MS1 through tD1-MS6) was initially tested on five plants. No RNAs of the predicted 1.2-kb length were detected with the CY-17 oligonucleotide probe in total RNA from uninfected plants, although there was slight cross-hybridization with 18S rRNA (Fig. 5B, lane a). tD1-M, containing the full-length ORF, accumulated in the presence of a helper (Fig. 5B, lane b), whereas mutant transcripts (tD1-MS1 through tD1-MS5) containing ORFs truncated at the 3′ end by 55 or more codons were unable to accumulate in any of the plants inoculated with them (Fig. 5B, lanes 1 to 5). The termination codons present in tD1-MS1 and tD1-MS2 are both located within 8 nucleotides of the predicted termination codons in tD1-MΔ1 and tD1-MΔ2, respectively. Lack of biological activity of tD1-MS1 and tD1-MS2 is, therefore, consistent with earlier results obtained with tD1-MΔ1 and tD1-MΔ2 (Table 2). Only tD1-MS6, containing an ORF reduced by only 19 codons at its 3′ end, was viable in three of five plants coinoculated with it (Fig. 5B, lane 6). The progeny of tD1-MS6 were amplified by RT-PCR, cloned,

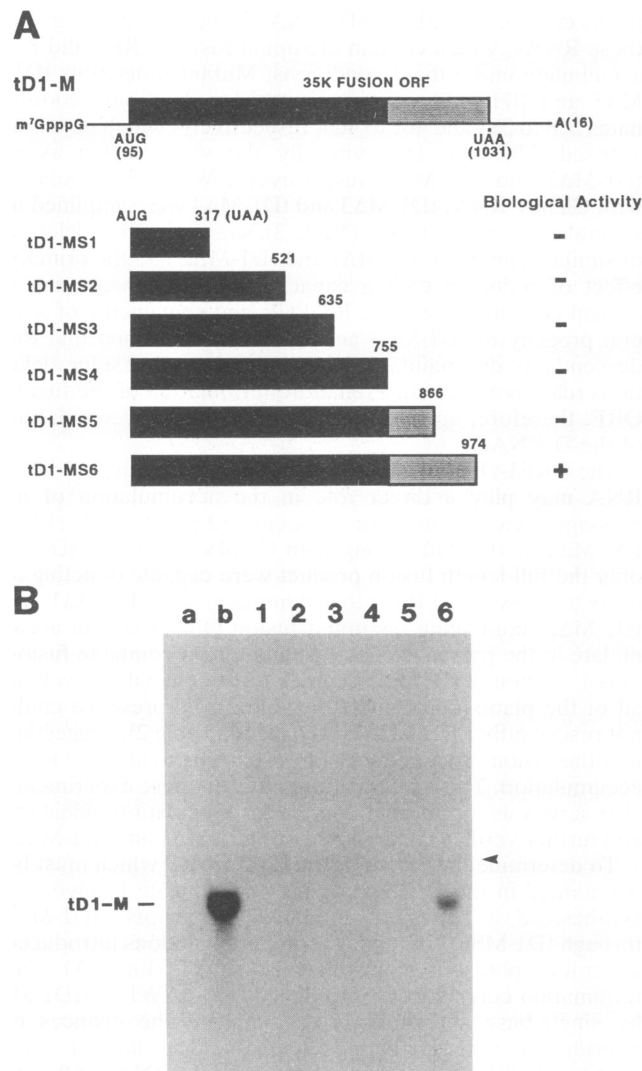


FIG. 5. Analysis of full-length D RNAs containing prematurely terminated fusion ORFs. (A) tD1-M and the extent of its coding region are shown at the top. Mutant transcripts (tD1-MS1 through tD1-MS6) containing ORFs reduced at their 3′ ends are shown below. The extents of the ORFs present in the mutant transcripts relative to that in tD1-M are indicated schematically, and the numerical value at the end of each partial ORF represents the position of the first residue of the introduced termination codon. (B) Total RNA from broad bean plants coinoculated with CYMV RNA and mutant transcripts containing reduced ORFs. Lanes: a, total RNA from uninfected plants; b, total RNA from inoculations with CYMV RNA and tD1-M; 1 through 6, total RNAs from plants coinoculated with CYMV RNA and transcripts tD1-MS1 through tD1-MS6, respectively. For each mutant transcript (tD1-MS1 through tD1-MS6), five plants were inoculated. An additional four plants (per mutant) were tested with independently prepared transcripts of tD1-MS3 through tD1-MS5 (see the text). Inoculations were carried out as described in Materials and Methods. Representative samples for tD1-MS1 through tD1-MS5 (for which no progeny could be detected in any of the plants analyzed) and for tD1-MS6 (for which progeny accumulated to detectable levels in three of the five plants tested) are shown. The arrowhead indicates 18S rRNA.

sequenced, and found to contain the stop codon introduced into the progenitor, tD1-MS6 (data not shown). Mutants tD1-MS3, tD1-MS4, and tD1-MS5, which showed no biological activity, were further tested with independently prepared inocula in four additional plants each. None of these mutant D RNAs accumulated to detectable levels in any of the plants tested. These results demonstrate that the potential for translation of all but the extreme 3′ portion of the fusion ORF must be maintained in a viable D RNA.

**Deletion analysis of D RNAs.** We have introduced several large deletions into the tD1-M transcript to identify internal regions of the D RNA which are important for its accumulation in vivo. Four restriction sites were used in the construction of the deletion mutants (Fig. 6A). Internal fragments were removed from pAW1-M (tD1-M) by digestion with two restriction enzymes, followed by blunt ending with T4 DNA polymerase and religation (see Materials and Methods). Deletions involving the *Ava*I site (tD1-MΔ8 to tD1-MΔ10) also removed 12 nucleotides of the inserted oligonucleotide marker (see Materials and Methods), leaving 16 residues complementary to the CY-17 oligonucleotide probe. Mutant transcripts containing such deletions were, however, clearly detectable by using the CY-17 oligonucleotide probe (see below). The deletion in tD1-MΔ5 did not maintain the reading frame of the fusion ORF, whereas the other five deletions (tD1-MΔ6 to tD1-MΔ10) did (Fig. 6A). Each mutant transcript (tD1-MΔ5 through tD1-MΔ10) was initially tested for biological activity by coinoculation with CYMV RNA on three plants, followed by Northern blot analysis of extracted total RNA by using labeled oligonucleotide CY-17 as the probe. Uninoculated plants contained no specific D RNAs, whereas plants coinoculated with tD1-M and CYMV RNA resulted in accumulation of an RNA of approximately 1.2 kb, as expected (Fig. 6B, compare lanes a and b). Of the six deletion mutants tested, only tD1-MΔ9, containing the smallest deletion, could be identified in total RNA extracted from coinfected plants (Fig. 6B, lane 9; tD1-MΔ9 accumulated in all three of the plants tested). The increase in mobility of the tD1-MΔ9 progeny D RNA (Fig. 6B, lane 9) compared with wild-type tD1-M progeny D RNA (Fig. 6B, lane b) is consistent with a deletion of 90 nucleotides (78 nucleotides of viral origin and 12 nucleotides of the oligonucleotide insertion). The progeny of tD1-MΔ9 were also found to be encapsidated (data not shown). tD1-M is, therefore, able to tolerate an in-frame deletion of 78 viral nucleotides (between coordinates 856 and 933) in the 3′ region of its fusion ORF. Independently prepared transcripts of mutants tD1-MΔ5 through tD1-MΔ8 and tD1-MΔ10 were further tested for biological activity, each in four additional plants. No progeny from these mutant D RNAs was detected in any of the plants tested.

## DISCUSSION

We have used a biologically active transcript of a defective 1.2-kb RNA of CYMV to (i) investigate properties of this RNA which allow it to accumulate in vivo and (ii) gain insight into the replicative strategy of CYMV. Synthetic 1.2-kb D RNA accumulates in vivo only when coinoculated with CYMV RNA, confirming that this RNA is in fact defective and requires elements provided in *trans* by the virus. The replication and encapsidation of the 1,172-nucleotide-long [excluding the poly(A) tail] synthetic prototype D RNA also indicates that the *cis*-acting elements necessary for these processes are present in the corresponding 5′ 757 and 3′ 415 residues of CYMV gRNA. Accumulation of

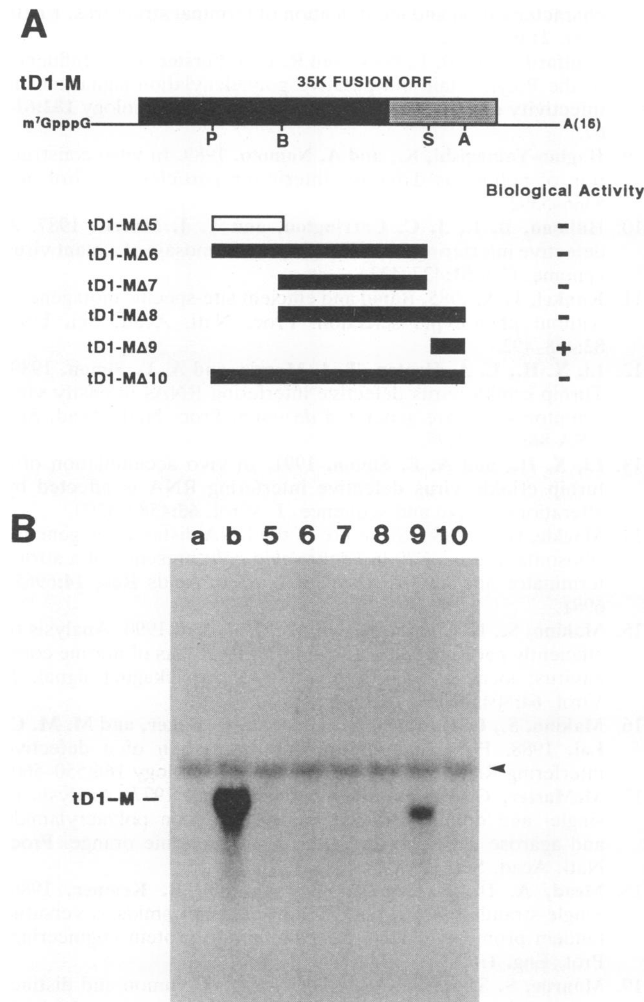


FIG. 6. Deletion analysis of CYMV D RNAs. (A) Various deletions were introduced into tD1-M (shown at the top) by removal of internal regions in its cloned DNA by digestion with two restriction enzymes (P, *Pst*I; B, *Bst*XI; S, *Stu*I; or A, *Ava*I), followed by religation. Regions removed in the various mutant transcripts (tD1-MΔ5 through tD1-MΔ10) are shown as solid (maintaining the fused ORF) or open (disrupting the fused ORF) bars. (B) Total RNAs from broad bean plants coinoculated with CYMV RNA and mutant transcripts containing large internal deletions. Lanes: a, total RNAs from uninfected plants; b, total RNAs from plants coinoculated with CYMV RNA and tD1-M; 5 through 10, total RNAs from coinoculations of CYMV RNA with transcripts tD1-MΔ5 through tD1-MΔ10, respectively. Initially, three plants were inoculated with each transcript as described in Materials and Methods. Representative samples for tD1-MΔ5 through tD1-MΔ8, tD1-MΔ10 (for which no progeny was detected in the plants analyzed), and tD1-MΔ9 (for which progeny accumulated to detectable levels in all three of the plants tested) are shown. This assay was repeated with independently prepared transcripts of tD1-MΔ5 through tD1-MΔ8 and tD1-MΔ10, each in four plants (see the text). The arrowhead indicates 18S rRNA.

progeny D RNA in vivo is dependent on $m^7$Gppp capping of the synthetic transcript used to inoculate the plants. Infectivity of full-length synthetic transcripts from several other plant viruses also absolutely requires a cap structure (e.g., tobacco mosaic virus [6], barley stripe mosaic virus [20], and turnip yellow mosaic virus [24]). Uncapped transcripts of the gRNA of the potexvirus white clover mosaic virus were found to have about 4% of the infectiousness of the capped version (3). In the case of the CYMV D RNA, the need for a cap structure is probably linked to its strict requirement for translation.

The synthetic D RNAs inoculated on plants contained poly(A) tails of 16 residues with 5 additional 3′-terminal residues contributed by the *Sal*I site used to linearize the DNA templates. The poly(A) tail in potexvirus RNAs may play a role not only in RNA stability and translational efficiency but also in the replication of viral RNA (7, 8). Synthetic transcripts of white clover mosaic virus with short poly(A) tails are less infectious than those with long tails (8). The poly(A) tail of 16 residues in several synthetic CYMV D RNAs was sufficient to allow their initial replication. Poly(A) tail length, however, subsequently increased in progeny produced in vivo. The putative polyadenylation signal identified in the 3′ noncoding region of CYMV gRNA (2, 8), which is also present in the D RNA, may account for the in vivo polyadenylation of the progeny. The five non-poly(A) residues at the end of the tD1-M transcript, however, were not maintained in progeny, suggesting either that they are not replicated in vivo or that they are removed prior to replication.

The remarkable feature of the D RNAs of CYMV is that maintenance of an encoded fusion ORF is essential for their accumulation in vivo. Several lines of evidence support this notion. (i) All of the natural D RNAs examined contain in-frame junctions fusing ORF1 with ORF5 (25). (ii) Small deletions early in the reading frame which cause premature termination of the ORF inactivate the D RNA, whereas similar deletions in the same region which preserve the reading frame allow the D RNA to accumulate. (iii) Maintenance of all but the extreme 3′ end of the fusion ORF is essential for D RNA viability. (iv) Small in-frame deletions and insertions in the 3′ end of the ORF are tolerated.

The requirement of a fusion ORF indicates that either the process of translation or the encoded product is essential to D RNA accumulation. Our results indicate that the fusion product does not act in *trans* to allow D RNA accumulation, but we cannot formally discount the possibility that it acts in *cis*. Certain poliovirus DI RNAs also require in-frame deletions for viability (9). It was suggested that the in-frame prerequisite was imposed by the *cis*-acting nature of at least one of the poliovirus nonstructural proteins encoded 3′ on the viral gRNA (and encoded 3′ to the junction in the poliovirus DI RNAs). The situation for CYMV D RNA, however, differs in that no complete authentic viral protein is produced from it. The encoded fusion product could conceivably retain a domain(s) required in *cis*, but its highly truncated and hybrid character argues against this. Additionally, the viability of D RNAs containing a variety of in-frame deletions and insertions in both 5′ and 3′ regions of the fusion ORF (which modify the fusion product) also suggests that the product is nonfunctional.

It is more likely that the process of translation is the essential factor which promotes the viability of the CYMV D RNA. Translation could possibly be coupled with replication or encapsidation if the factors participating in these viral processes were ribosome associated or required ribosome-associated viral RNA as a substrate. Alternatively, translation of the fused ORF may confer stability to the message, for example, by protecting a nuclease-sensitive region of the D RNA. It is, however, unclear whether translation of a specific viral sequence is necessary or translation in general is sufficient.

The requirement for translation in the D RNA also raises the question of whether translation of the corresponding regions in the CYMV gRNA is needed in normal infections. Analysis of products synthesized in vitro from CYMV gRNA suggested that the 191-kDa ORF-1 product is much more efficiently expressed than the coat protein (4). This result does not, however, preclude the possibility that low levels of coat protein translation would be sufficient to meet such a requirement.

Viable D RNAs are capable of tolerating a short in-frame insertion (tD1-M) or both a short in-frame insertion and a small deletion (tD1-MΔ9) in the 3′ coding region. This indicates that essential *cis* elements are not inactivated by these modifications. Larger in-frame deletions render the D RNA inactive, suggesting that the regions removed contain elements important for replication and/or encapsidation of D RNAs. Alternatively, deletion of a region may cause suboptimal spacing between elements which work cooperatively or destabilize the RNA. Replacement of deleted regions in nonviable DI RNAs of turnip crinkle carmovirus with nonviral sequences similar in length resulted in restored viability (13); we have not attempted similar experiments.

It has been postulated that DI RNAs which lack messenger activity may possess an evolutionary advantage by retaining only essential sequences required for replication and encapsidation (19). In CYMV D RNAs, poliovirus DI RNAs (9), and possibly mouse hepatitis coronavirus DI RNAs (15, 16, 23), it is, however, critical that certain coding regions be maintained. The requirement for messenger activity in some subgenomic replicons of both plant and animal viruses demonstrates that, in some cases, translatability can increase competitiveness.

## ACKNOWLEDGMENTS

This work was supported by a grant to G.A.M. from the Medical Research Council (MRC) and a Natural Sciences and Engineering Research Council of Canada grant to J.B.B. and G.A.M. K.A.W. is the recipient of an MRC studentship.

## REFERENCES

1. **AbouHaidar, M. G.** 1983. The structure of the 5′ and 3′ ends of clover yellow mosaic virus RNA. Can. J. Microbiol. **29:**151–156.
2. **Bancroft, J. B., M. Rouleau, R. Johnston, L. Prins, and G. A. Mackie.** 1991. The entire sequence of foxtail mosaic virus RNA. J. Gen. Virol. **72:**2173–2181.
3. **Beck, D. L., R. L. S. Forster, M. W. Bevan, K. A. Boxen, and S. C. Lowe.** 1990. Infectious transcripts and nucleotide sequence of cloned cDNA of the potexvirus white clover mosaic virus. Virology **177:**152–158.
4. **Bendena, W. G., J. B. Bancroft, and G. A. Mackie.** 1987. Molecular cloning of clover yellow mosaic virus RNA: identification of coat protein coding sequences *in vivo* and *in vitro*. Virology **157:**276–284.
5. **Burgyan, J., F. Grieco, and M. Russo.** 1989. A defective interfering RNA molecule in cymbidium ringspot virus infections. J. Gen. Virol. **70:**235–239.
6. **Dawson, W. O., D. L. Beck, D. A. Knorr, and G. L. Grantham.** 1986. cDNA cloning of the complete genome of tobacco mosaic virus and production of infectious transcripts. Proc. Natl. Acad. Sci. USA **83:**1832–1836.
7. **Dolja, V. V., D. P. Grama, S. Y. Morozov, and J. G. Atabekov.** 1987. Potato virus X-related single- and double-stranded RNAs: characterization and identification of terminal structures. FEBS Lett. **214:**308–312.
8. **Guilford, P. J., D. L. Beck, and R. L. S. Forster.** 1991. Influence of the Poly(A) tail and putative polyadenylation signal on the infectivity of white clover mosaic potexvirus. Virology **182:**61–67.
9. **Hagino-Yamagishi, K., and A. Nomoto.** 1989. In vitro construction of poliovirus defective interfering particles. J. Virol. **63:**5386–5392.
10. **Hillman, B. I., J. C. Carrington, and T. J. Morris.** 1987. A defective interfering RNA that contains a mosaic of a plant virus genome. Cell **51:**427–433.
11. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488–492.
12. **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.
13. **Li, X. H., and A. E. Simon.** 1991. In vivo accumulation of a turnip crinkle virus defective interfering RNA is affected by alterations in size and sequence. J. Virol. **65:**4582–4590.
14. **Mackie, G. A.** 1986. Structure of the DNA distal to the gene for ribosomal protein S20 in *Escherichia coli*: presence of a strong terminator and an IS1 element. Nucleic Acids Res. **14:**6965–6981.
15. **Makino, S., K. Yokomori, and M. M. C. Lai.** 1990. Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. J. Virol. **64:**6045–6053.
16. **Makino, S., C. K. Shieh, L. H. Soe, S. C. Baker, and M. M. C. Lai.** 1988. Primary structure and translation of a defective interfering RNA of murine coronavirus. Virology **166:**550–560.
17. **McMaster, G. K., and G. G. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. USA **74:**4835–4838.
18. **Mead, A. D., E. Szczesna-Skorupa, and B. Kemper.** 1986. Single-stranded DNA 'blue' T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. Prot. Eng. **1:**67–74.
19. **Monroe, S. S., and S. Schlesinger.** 1984. Common and distinct regions of defective-interfering RNAs of Sindbis virus. J. Virol. **49:**865–872.
20. **Petty, I. T. D., B. G. Hunter, N. Wei, and A. O. Jackson.** 1989. Infectious barley stripe mosaic virus RNA transcribed *in vitro* from full-length genomic cDNA clones. Virology **171:**342–349.
21. **Roux, L., A. E. Simon, and J. J. Holland.** 1991. Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. Adv. Virus Res. **40:**181–211.
22. **Sit, T. L., K. A. White, S. Holy, U. Padmanabhan, M. Eweida, M. Hiebert, G. A. Mackie, and M. G. AbouHaidar.** 1990. Complete nucleotide sequence of clover yellow mosaic virus RNA. J. Gen. Virol. **71:**1913–1920.
23. **van der Most, R. G., P. J. Bredenbeek, and W. J. M. Spaan.** 1991. A domain at the 3′ end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. J. Virol. **65:**3219–3226.
24. **Weiland, J. J., and T. W. Dreyer.** 1989. Infectious TYMV RNA from cloned cDNA: effects *in vitro* and *in vivo* of point substitutions in the initiation codons of two extensively overlapping ORFs. Nucleic Acids Res. **17:**4675–4687.
25. **White, K. A., J. B. Bancroft, and G. A. Mackie.** 1991. Defective RNAs of clover yellow mosaic virus encode nonstructural/coat protein fusion products. Virology **183:**479–486.
26. **White, K. A., and G. A. Mackie.** 1990. Control and expression of 3′ open reading frames in clover yellow mosaic virus. Virology **179:**576–584.