JOURNAL OF VIROLOGY, May 1992, p. 3131–3139
0022-538X/92/053131-09$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 66, No. 5

The Full-Length Transcript of a Caulimovirus Is a Polycistronic mRNA Whose Genes Are *trans* Activated by the Product of Gene VI

HERMAN B. SCHOLTHOF,† SIDDARAME GOWDA, FANG C. WU, AND ROBERT J. SHEPHERD*
Department of Plant Pathology, University of Kentucky, Lexington, Kentucky 40546-0091

Received 20 December 1991/Accepted 6 February 1992

Gene expression of figwort mosaic virus (FMV), a caulimovirus, was investigated by electroporation of *Nicotiana edwardsonii* cell suspension protoplasts with cloned viral constructs in which a reporter gene was inserted at various positions on the genome. The results showed that the genome of FMV contains two promoters; one is used for the production of a full-length RNA and another initiates synthesis of a separate monocistronic RNA for gene VI. Evidence is provided that the full-length transcript, the probable template for reverse transcription, can serve as a polycistronic mRNA for translation of genes I through V and perhaps also gene VI. Expression of all the genes on the polycistronic mRNA is *trans* activated by the gene VI protein. Reporter gene expression appears most efficient when its start codon is in close proximity to the stop codon of the preceding gene, as for the native genes of caulimoviruses. We propose that the gene VI product enables expression of the polycistronic mRNA by promoting reinitiation of ribosomes to give translational coupling of individual genes.

Figwort mosaic virus (FMV) is a member of the caulimovirus group, which, together with badnaviruses (29) and animal hepadnaviruses, represents the pararetroviruses (47). FMV is a small isometric virus which infects plants belonging to the families Scrophulariaceae, Chenopodiaceae, and Solanaceae (45). The circular double-stranded DNA genome of FMV is 7,743 bp. Its genomic organization (34) (Fig. 1) strongly resembles that of cauliflower mosaic virus (CaMV) (12), carnation etched ring virus (22), and soybean chlorotic mottle virus (SoyCMV) (19).

Caulimoviruses contain a large intergenic region (L-IR) with a promoter for the synthesis of an RNA that, in CaMV, has been shown to span the entire genome (18). This transcript can serve as a template for reverse transcription, which is the process by which these viruses replicate, as reviewed by Mason et al. (28). A separate promoter in the small intergenic region (S-IR) of CaMV between genes V and VI directs transcription of a smaller RNA that spans gene VI (18).

The mechanism by which caulimoviruses express genes I through V has been a matter of speculation. These genes are usually spaced with only one or two nucleotides between the stop codon of one gene and the start codon of the next; occasionally they overlap by a few nucleotides. In this respect they resemble prokaryotic genes that are translated in a coupled manner, as discussed by Gronenborn (17). There is no evidence that the full-length positive-sense transcript is spliced in a regulatory manner (2, 10, 42). Neither have internal transcription units for genes I through V been identified, although a subgenomic transcript for gene V of CaMV has been postulated (21, 33) and a sequence inside gene III of SoyCMV can reportedly promote expression of a downstream gene in protoplasts (19). However, no subgenomic RNAs other than the gene VI transcript have been well documented, and consequently none of the major open reading frames (ORFs) appear in a 5′-proximal position on mRNAs. There is considerable evidence that the consensus rule of translation of eukaryotic monocistronic mRNAs (25) does not apply to genes I through V of caulimoviruses such as CaMV and FMV. Apparently, alternative strategies are used, allowing either internal initiation of ribosomes for translation or reinitiation by scanning ribosomes. A protein which might regulate the unusual translation mechanism is encoded by gene VI, which was shown to be involved in the posttranscriptional expression of reporter genes positioned downstream of the caulimovirus leader sequence (2, 15).

Previous investigations of gene expression with CaMV and FMV were usually performed with prematurely terminated transcripts (2, 11, 15), which could not conclusively demonstrate the polycistronic nature of the full-length caulimovirus transcript for translation of all its genes. In addition, the importance of gene VI in expression of genes downstream of ORF I has been documented only in a preliminary report with CaMV and again with prematurely terminated transcripts (21). Therefore we sought further evidence for a polycistronic messenger and the role of gene VI in expression of several genes by using FMV DNA constructs which would produce authentic viral transcripts during transient-expression studies in protoplasts. The results revealed that FMV DNA is transcribed as two transcripts, a gene VI RNA and a full-length RNA; each has a separate promoter. Several lines of evidence suggest that the full-length RNA serves as a polycistronic mRNA; the expression of all the genes on this transcript is efficiently *trans* activated by gene VI protein. We also compared the expression of a reporter gene in various positions on the genome and postulate that the translation of genes I through V is coupled.

MATERIALS AND METHODS

Standard protocols. Standard molecular biology techniques were used in this study; for protocols, see the manual of Maniatis et al. (27). Electroporation and chloramphenicol acetyltransferase (CAT) assay of protoplasts from *Nicotiana*

* Corresponding author.
† Present address: Department of Plant Pathology, University of California, Berkeley, CA 94720.

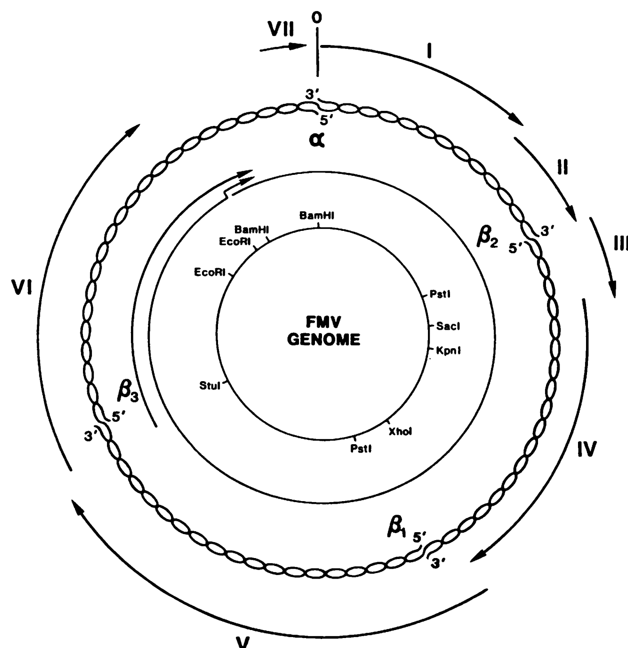

FIG. 1. Genetic organization of the FMV genome. The circular double-stranded DNA genome of FMV is 7,743 bp in length and is indicated by the interwoven lines. Single-stranded discontinuities (α and β_1 to β_3) are shown at specific points. The peripheral arrows indicate the positions of the major ORFs (Roman numerals). The two major RNAs which are transcribed from the genome are depicted by the thin innermost arrows.

edwardsonii cell suspensions were described by Gowda et al. (15).

Analyses of RNA. *Datura stramonium* seedlings at the 3- to 4-leaf stage were inoculated with an infectious partially redundant clone of FMV, pFiPR (42). At 3 weeks postinoculation, total RNA was isolated from 3 to 5 g of leaf material by the method of Verwoerd et al. (48). After phenol extraction of the aqueous phase, RNA was extracted with phenol-chloroform, precipitated with ethanol, resuspended in 1% Sarkosyl at 65°C, and again extracted and precipitated. After resuspension in sterile water, the RNA was treated with RNase-free DNase I (10 U/ml at 37°C for 30 min), extracted, and precipitated. Finally the RNA was resuspended in sterile water containing 1 U of RNasin per μl. RNA was subjected to electrophoresis through a denaturing agarose gel containing formaldehyde and transferred to a membrane as described by Maniatis et al. (27); it was then hybridized with nick-translated probes by the method of Jones et al. (24).

Recombinant plasmids. In the following section the nucleotide numbers refer to the position on the wild-type genome (34). Most of the plasmids are derived from a deletion mutant clone (pKF4) of FMV in which the major part of ORF IV and the 5′ end of ORF V are deleted but the remainder of the genome is identical to the wild type (42).

For construction of pH13, the *Eco*RI fragment of pSX103 (42), which encompasses most of the IV/V fusion gene as well as the S-IR and the 5′ end of the gene VI-CAT fusion gene, was inserted into the *Eco*RI window of pRCAT (15). As a result, the promoter of gene VI in pH13 drives CAT gene expression and the transcription is terminated by the polyadenylation signal of the gene (RT) for the small subunit of ribulose 1,5-diphosphate carboxylase (rubisco) immediately downstream of the CAT gene (see Fig. 3). In pH15 the *Nde*I fragment of pH13 from nucleotide (nt) 4690 to the site in pUC119 is deleted; consequently only a 680-bp promoter fragment remains upstream of the CAT gene (see Fig. 3). The *Sna*BI (nt 4795, at the 3′ end of the IV/V fusion)-*Stu*I (nt 5379, 5′ end of VI) fragment is rotated in pH16 with respect to the orientation in pH15 (see Fig. 3). As a result, the putative promoter fragment is in the opposite orientation but the ORF for CAT is intact. A 300-bp *Eco*RI-*Sac*I fragment from plasmid pBI131.1 (23), which contains the polyadenylation signal of the nopaline synthase gene (NT), was inserted in the compatible window of pUC120 to create p3′NT. Subsequently pH13 was digested with *Pst*I (nt 3632) and *Bam*HI (to cut just downstream of the ORF for CAT), and this fragment was inserted into p3′NT to create pH32 (see Fig. 3). In this plasmid the promoter of gene VI and downstream CAT gene are positioned under the control of NT.

The *Nsi*I (nt 6603)-*Sal*I fragment of pKF4 was inserted in the *Pst*I-*Sal*I window of pJAW130 (pUC119 without an *Eco*RI site) to give pHS4. The viral *Sal*I fragment of either pH22, pH23, pH14, pH24, or pH25 (41) was inserted at the *Sal*I site of pHS4 to give pH62, pH63a, pH63b, pH64, and pH65, respectively (see Fig. 4). For construction of pH61, the CAT gene of pCM-1 (Promega) was inserted at the *Sal*I site of pDFiPR, a partially redundant clone of the deletion mutant of FMV (42). For construction of pH66, an *Msc*I fragment, from position 1804 to the middle of the CAT gene in an equivalent of pH63a (39), was replaced with a *Sma*I linker (CCCCGGGG) and then the CAT gene from pSX103 (42) with flanking *Stu*I sites was inserted in the newly created *Sma*I site; as a result the CAT gene was positioned in frame with ORF III. Plasmid pH24 (42) was used for construction of pH44-3 and its derivative pH44ΔE as described previously (16). For construction of pH61ΔS through pH66ΔS, the *Sna*BI-*Stu*I fragment (nt 4795 and 5379, respectively) of pH60 through pH66 was removed, resulting in plasmids which lack the promoter and 5′ end of ORF VI (see Fig. 5).

RESULTS

Identification of two viral transcripts and two promoters. Northern (RNA) blot analyses of RNA isolated from FMV-infected leaves of *D. stramonium* showed that two major virus-specific transcripts are produced, as shown in Fig. 2. The larger RNA, of ca. 7,900 nt, spans the entire genome as indicated by its hybridization pattern to four specific probes that cover the major ORFs of FMV. The smaller transcript, estimated to be ca. 1,700 nt, spans only the gene VI region of the genome and did not hybridize to probes a, b, and c (Fig. 2), which differentiate other ORFs on the genome. The area on the genome covered by the two transcripts is schematically diagrammed as thin arrows in Fig. 1.

Previous studies demonstrated that the L-IR of FMV contained an active promoter (Fig. 3) comparable in strength to the 35S promoter of CaMV (15, 37, 49). It is likely that this promoter is responsible for the production of the ca. 7,900-nt transcript in Fig. 2. By analogy to CaMV, we postulated that FMV also contains a promoter within the S-IR for transcription of gene VI (34). The transient gene expression data obtained after electroporation of *N. edwardsonii* cell suspension protoplasts with plasmids pH13, pH15, and pH32 (illustrated in Fig. 3) confirm that the segment between nt 4690 and 5379 can stimulate expression of CAT. In pH16, when the *Sna*BI-*Stu*I fragment from nt 4795 to 5379 of pH15 was inserted in the reverse orientation, the expression of CAT

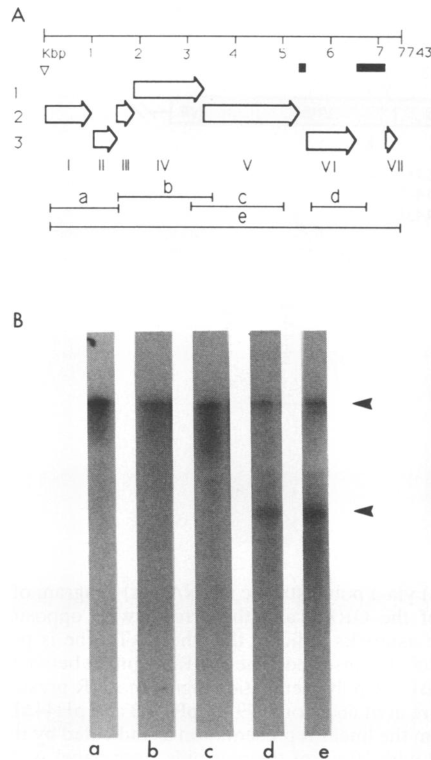

FIG. 2. Characterization of two virus-specific RNAs in plants. (A) Linear representation of the FMV genome. The genome is divided in segments of 1 kbp, as indicated by the single digit numbers, and the total size in base pairs is given at the right-hand side. The zero position is indicated (▽), and a *Sal*I site at this position is present in most of the plasmids described in Materials and Methods and connects the zero site on the circular genome to position 7743. The three reading frames are indicated by Arabic numerals, and the locations of the ORFs on the genome and their notation are represented by the boxed arrows and Roman numerals, respectively. Small and large intergenic regions are indicated by the respective black boxes. Letters a through e represent gene-specific probes in the Northern analysis in panel B. DNA segments: a, *Bgl*II-*Pst*I (nt 106 to 1564); b, *Pst*I-*Pst*I (nt 1564 to 3632); c, *Xho*I-*Sna*BI (nt 3166 to 4795); d, *Stu*I-*Pvu*II (nt 5379 to 6692); e, the full-length genome. (B) Northern analyses of RNA isolated from FMV-infected *D. stramonium* plants. Hybridization was performed with the probes in panel A as stated below each lane. The arrows indicate the positions of the two transcripts.

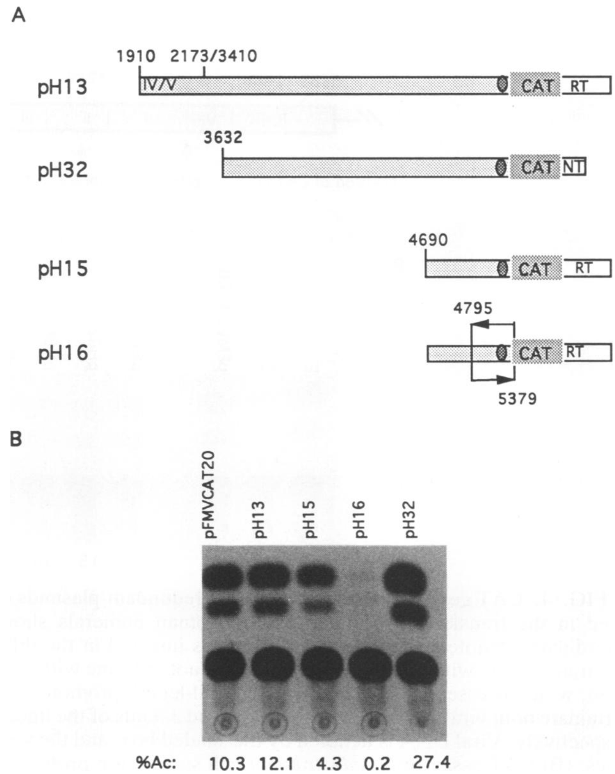

FIG. 3. Identification of a promoter upstream of gene VI. (A) Diagram of plasmids. Numbers indicate the positions on the (boxed) viral genome. The 2173/3410 position indicates the location at which the 1,237-bp deletion has occurred in the deletion mutant (42). The dark ovals denote the location of the promoter, and the position of CAT is indicated. The white boxes show whether the polyadenylation signal is derived from the gene for the small subunit of rubisco (RT) or the nopaline synthase gene (NT). The fragment from nt 4795 to 5379 in pH15 is rotated in pH16, as indicated by the arrows. (B) CAT assay of *N. edwardsonii* cell suspension protoplasts 24 h after electroporation with the plasmids denoted above each lane. To enable quantitative comparisons, equimolar amounts of DNA were used for electroporation in each case (i.e., 15 μg of pFMVCAT20, 5,217 bp; 20 μg of pH13, 7,129 bp; 15 μg of pH15 and pH16, 5,389 bp; 17 μg of pH32, 6,047 bp). The numbers below the lanes give the percentage of [^{14}C]chloramphenicol that is acetylated (42).

was abolished, although the coding region for CAT was intact. The CAT gene expression from pH13, pH15, and pH32 was at least as active as the CAT expression from the full-length promoter in pFMVCAT20. Plasmid pH32 repeatedly demonstrated the highest level of CAT expression (Fig. 3). The above results strongly suggest that the small transcript in Fig. 2 is the result of a separate promoter positioned upstream of gene VI.

CAT expression at various positions on the genome via a polycistronic mRNA. In previous studies of gene expression of FMV and CaMV, a reporter gene was positioned as the first or second major ORF with the remainder of the viral DNA deleted (2, 11, 15). To study gene expression from authentic full-length viral transcripts, we generated the DNA constructs shown in Fig. 4. These redundant-end FMV plasmids can produce both full-length and gene VI transcripts in vivo from their respective promoters. Various plasmids with the CAT gene in several downstream positions (Fig. 4) were used in transient-expression assays with protoplasts. CAT expression was obtained with every member of this series of plasmids, indicating that CAT is expressed well at a variety of downstream positions.

In pH61 the CAT gene is positioned as a separate cistron between ORF VII and ORF I. It is efficiently expressed at a consistently higher level than when the CAT gene is inserted into gene II as a separate cistron (i.e., pH62) (Fig. 4). Expression from pH62 was reduced two- to fourfold compared with expression from pH61, pH66, pH63b, and pH64. The expression of pH63a was lower than that of pH63b. Both plasmids have the CAT gene in frame with gene IV/V, but in pH63b the CAT gene was inserted further downstream (100 bp) than in pH63a. In pH66 the CAT gene was inserted at the same position as in pH63a, but in the former plasmid it was positioned in frame with ORF III and not in frame with the ORF IV/V fusion gene as in pH63a. The CAT expression of pH66 is ca. ninefold higher than that of pH63a. The CAT gene in pH64 is inserted as a separate cistron in the distal

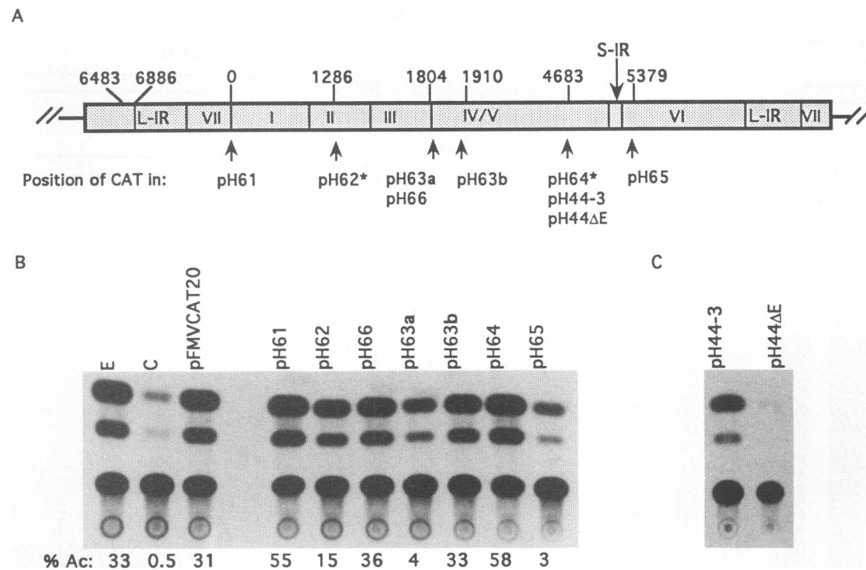

FIG. 4. CAT gene expression of partially redundant plasmids of FMV (1.2-mers) via a polycistronic mRNA. (A) Diagram of the 1.2-mers used in the transient-expression assays. Roman numerals show the positions of the ORFs, and the arrows with opposing nucleotide coordinates denote where the CAT gene was inserted in the different clones. The asterisks indicate that the CAT gene is positioned as a separate cistron with its own start codon and not in frame with the viral ORF in which it is inserted. The *Eco*RI segment between nt 6483 and 6886, which is essential for activity of the full-length promoter, is deleted in pH44ΔE. A polyadenylation signal in L-IR presumably acts to terminate both viral transcripts (34). The 5′ and 3′ ends of the linear representation are at nt 6603 (or 5379 for pH44-3 and pH44ΔE) and 0/7743, respectively. Viral DNA is denoted by the shaded box, and the vector is omitted from the linear representation as indicated by the interrupted lines. (B) CAT assay of *N. edwardsonii* cell suspension protoplasts electroporated with 50 μg of the plasmids from panel A. Lanes: E, 0.1 U of CAT enzyme; C, mock-electroporated protoplasts. pFMVCAT20 (20 μg) was used as a positive control. (C) CAT assay of *N. edwardsonii* cell suspension protoplasts electroporated with 50 μg of pH44-3 or pH44ΔE.

half of the IV/V ORF fusion, and CAT expression from this plasmid is comparable to that from pH61 (Fig. 4). CAT expression from pH65, with the CAT gene at the 5′ end of ORF VI and in frame with this gene, is lower than that of any other position.

In plasmid pH44ΔE (Fig. 4) a segment essential for the activity of the full-length promoter (15, 49) was deleted. This plasmid and its parent plasmid (pH44-3), with an intact promoter, both have the CAT gene as a separate cistron at the 3′ end of ORF IV/V. The deletion practically abolished the otherwise high level of CAT expression (Fig. 4C). Evidently, expression of ORF IV/V is dependent on the activity of the promoter for the full-length transcript. Northern analyses of RNA from plants (Fig. 2) and protoplasts (data not shown) showed no evidence of spliced transcripts. In a preliminary report (41) it was shown that full-length FMV transcripts also give rise to readily detectable levels of CAT expression in an in vitro translation system. This will be described in more detail in a separate communication.

Gene VI *trans* activates expression of CAT at every position on the genome. The participation of gene VI in expression of genes on the partially redundant plasmids was investigated through introduction of a deletion (ΔS) in plasmids of the pH60 series (Fig. 5). In Fig. 3 it was shown that such a deletion would inactivate the expression of gene VI. Electroporation of these plasmids into *N. edwardsonii* protoplasts resulted in a very poor expression when compared with the expression of the parental constructs that contained an intact gene VI (Fig. 5). This low level of expression could be dramatically increased when gene VI was provided in *trans* from pGS1RVI, a plasmid with gene VI between the 35S promoter of CaMV and the polyadenylation signal of the gene for the small subunit of rubisco (15). This coelectroporation with pGS1RVI increased the CAT expression 5- to 44-fold when compared with the data after coelectroporation with the control plasmid pGS1 (Fig. 5), which contains the promoter and terminator of the expression cassette but is devoid of gene VI.

DISCUSSION

Two major RNAs are transcribed from separate promoters on the FMV genome. Two major transcripts were detected in FMV-infected cells; the larger of these is about 7,900 bases and spans all of the major protein-encoding regions of the viral genome (Fig. 2). This RNA is very probably a redundant-end full-length copy of the genome which starts in the large intergenic region and terminates after a full round of transcription, in the same manner as the 35S transcript of CaMV (6, 8). As predicted from the sequence, the promoter for this transcript is located within the distal end of gene VI (34, 49). This promoter can be used to direct a high level of gene expression comparable to that directed by the 35S promoter of CaMV (15, 37, 49), as can also be seen in Fig. 3 and 4. Preliminary results of primer extension experiments locate the start site of transcription from the full-length promoter in the close vicinity of nt 6900 (unpublished observations).

The transient-expression data in Fig. 3 demonstrate the presence of a strong promoter upstream of gene VI which is very probably responsible for the smaller transcript of about 1,700 nt which spans coding region VI (Fig. 2). Hence, FMV is similar to CaMV, in which a 19S transcript (6, 8) is generated by a promoter positioned in the small intergenic region between genes V and VI. The FMV genome has a similar small intergenic region of 116 bases between these

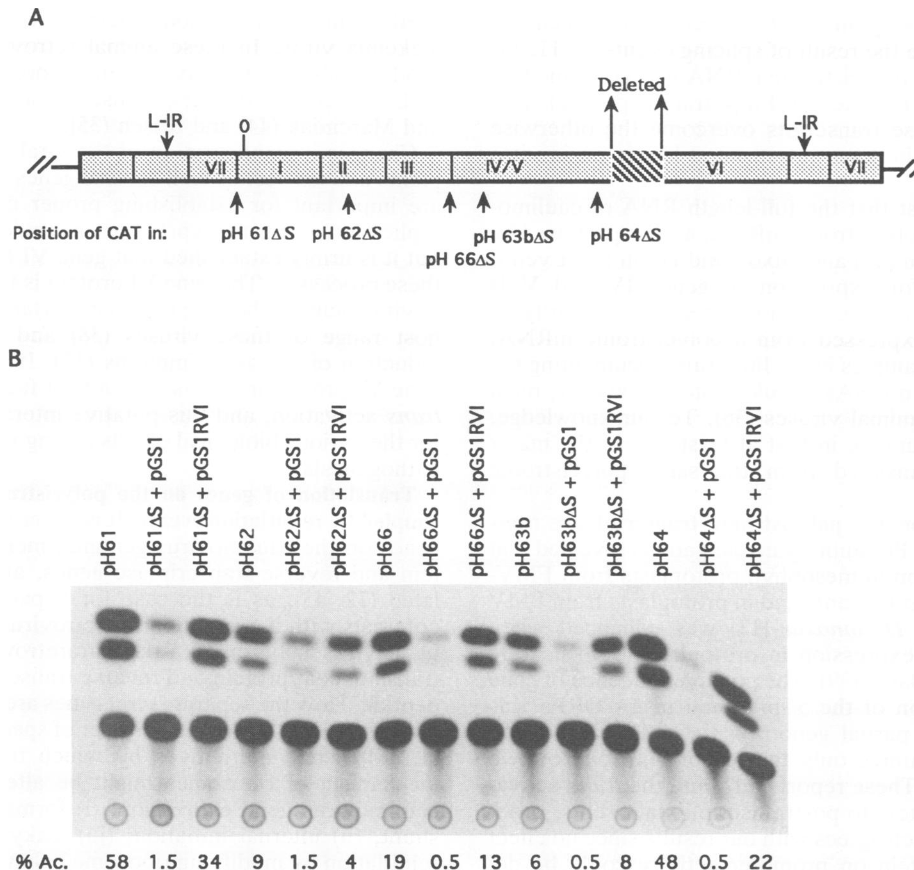

FIG. 5. Enhancement of CAT expression in several downstream positions on the FMV genome by gene VI. (A) Diagram of plasmids used for transient CAT expression assay in protoplasts. Symbols are as in Fig. 4. The *Sna*BI-*Stu*I fragment containing the promoter and 5′ end of gene VI from nt 4795 to 5379 (box with diagonal lines) of the parental clones pH61 through pH66 is deleted in the plasmid series designated pH61ΔS through pH66ΔS. (B) CAT assays of *N. edwardsonii* cell suspension protoplasts electroporated with 50 μg of each plasmid indicated above the lane. Numbers (% Ac.) give the percentage of [^{14}C]chloramphenicol acetylated by CAT.

two coding regions (34). It is generally believed that the strength of the 19S promoter of CaMV is inferior to that of the 35S or full-length promoter. The results of transient-expression assays with the FMV promoters (Fig. 3) do not show this distinction since expression of both promoters appears equally strong (compare pH13 and pFMVCAT20 in Fig. 3). However, the data do suggest that regions within ORF V enhance expression from the gene VI promoter since the CAT expression of pH13 and pH32 (which contain these sequences) is higher than that of pH15.

The expression of CAT from pH65 (Fig. 4) was consistently lower than the expression from pH13, pH15, and pH32 (Fig. 3), which have the CAT gene at the same position. One could attempt to ascribe this effect to differences in 3′ processing efficiencies, because in pH65 transcription is terminated by the FMV poly(A) signal whereas it is regulated by RT in pH13 and pH15 and by NT in pH32 (Fig. 3). However, a derivative of pH13 in which RT was replaced with the FMV polyadenylation signal (as for pH65) also expressed CAT at a higher level than pH65 did (data not shown). This suggests that the low expression of pH65 is not caused by the FMV poly(A) signal. The data are in agreement with preliminary results which indicate that the expression of the full-length RNA from pH65 interferes with the expression from the gene VI promoter, thereby hindering CAT expression (39). This interference does not occur in pH13, pH15, and pH32 since the promoter for the full-length RNA is not present on these plasmids.

At least five major genes are translated from the full-length transcript. Partially redundant clones of FMV expressed CAT in protoplasts when this reporter gene was inserted in a variety of positions (Fig. 4). This expression is very probably due to the action of the full-length FMV promoter upstream of the L-IR, as demonstrated for expression of CAT in ORF V (Fig. 4C). This provides the first evidence that internal transcription units are not required for gene expression of genes I to V of FMV, although such cryptic promoters might be active in SoyCMV (19) and perhaps CaMV (21). The Northern blot analyses of transcripts from infected plants (Fig. 2) and from protoplasts (data not shown) do not detect spliced RNA products. This is in agreement with results obtained with CaMV (2, 10), and it confirms the notion that although splicing might occur (20, 42), it is probably not part of the regulatory process for gene expression of caulimoviruses. Plasmids pH64 and pH44-3 (Fig. 4) and pH24 (42) all have the CAT gene as a separate cistron located about 1,430 bp within ORF V. The high level of CAT expression from this configuration also indicates that splicing is not involved, because a legitimate splicing signal in the middle of ORF V is very unlikely. Neither does CAT gene insertion result in illegitimate splicing for production of subgenomic mRNAs, because the ΔS constructs in Fig. 5

express CAT very poorly and the *trans* activation by gene VI is very unlikely to be the result of splicing events (2, 11, 14, 15, 40). The use of a polycistronic mRNA was confirmed by in vitro translation of full-length FMV transcripts (41). It is not known how these transcripts overcome the otherwise inhibitory effect of the long untranslated leader on in vitro translation (1, 9).

Our results suggest that the full-length RNA of caulimoviruses serves as a polycistronic mRNA, as was proposed by Sieg and Gronenborn (46) and Dixon and Hohn (7). Even if a spliced product for expression of genes IV and V is produced at levels that we did not detect, the majority of genes will still be expressed from a polycistronic mRNA. There are several examples in the literature documenting the use of polycistronic mRNAs in eukaryotes. However, most examples apply to animal viruses (36). To our knowledge, caulimoviruses are unique in that at least five of the major viral genes are translated from the same polycistronic mRNA.

All major genes on the polycistronic transcript are *trans* activated by gene VI. Preliminary investigations revealed that FMV gene expression in mesophyll protoplasts from FMV-infected *N. edwardsonii* plants and in protoplasts from FMV gene VI transgenic *D. innoxia* (13) was enhanced when compared with the expression in protoplasts from healthy and nontransgenic plants (39). The positive influence of gene VI on the expression of the 5′-proximal major ORFs was also observed with partial genomes of CaMV (2, 11) and FMV (15, 16) containing only the leader with one or two downstream genes. These reports attribute the *trans* activation by gene VI protein to posttranscriptional events, probably translation, which agrees with our results since no effect of the gene VI protein on promoter activity could be detected (data not shown). As suggested by recent preliminary reports (40, 41), the polycistronic FMV transcript responds better to *trans* activation when transcription is not prematurely terminated. For this reason and because it was not established whether gene VI regulated the expression of all genes on the full-length transcript, especially the 3′-proximal ORFs, we tested *trans* activation with plasmids that would produce transcripts closely resembling the native caulimovirus RNAs.

The results (Fig. 5) show that a deletion spanning the promoter and the 5′ end of gene VI, in the partially redundant clones, practically abolishes expression of the CAT gene upstream of gene I and within genes II, III, IV, and V. Expression of CAT is restored when these plasmids are coelectroporated with a plasmid that provides gene VI in *trans* (Fig. 5). These observations firmly establish the importance of gene VI in the expression of genes I through V on the full-length polycistronic RNA. The low CAT gene expression of pH65 (Fig. 4), a plasmid with a nonfunctional gene VI, can also be dramatically enhanced by providing a functional gene VI either on the same plasmid (42) or by coelectroporation with a gene VI-expressing plasmid (39, 40). Apparently expression of gene VI is *trans* activated by its own protein. It is possible that during *trans* activation gene VI is translated from the full-length RNA in vivo, as it is in vitro (39, 41).

Recent studies with FMV have revealed that *cis* elements within region VII (14) and at the 3′ end of the viral RNA (40) are required in addition to gene VI protein for efficient *trans* activation of the viral genes. Both these regions are present on the full-length transcripts used in our experiments. The *cis-trans* elements involved in gene expression in this plant pararetrovirus may be analogous to gene regulation systems used by human immunodeficiency virus and human T-cell leukemia virus. In these animal retroviruses certain gene products also *trans* activate gene expression by interaction with elements of the viral transcript, as reviewed by Sharp and Marciniak (44) and Rosen (35).

Changes in configuration of the viral genome in different plants and tissues, as well as viral genes other than gene VI, are important for establishing proper conditions for virus replication and gene expression, as reviewed by Covey (5), but it is firmly established that gene VI has a pivotal role in these processes. The gene VI protein is the main constituent of viral inclusion bodies; it is an important determinant of the host range of these viruses (38) and has a role in the induction of disease symptoms (13). It is possible that the gene VI protein interacts with a host factor(s) which allows *trans* activation, and this putative interaction may account for the various biological effects during virus replication and pathogenesis.

Translation of genes on the polycistronic mRNA may be coupled by reinitiation events. It is generally believed that the genes on the caulimovirus genome, including the coat protein and reverse transcriptase genes, are separately translated (32, 43), as is the case for hepadnaviruses (4). This contrasts with the situation for retroviruses and perhaps the badnavirus subgroup of the pararetroviruses (29), which translate coat protein and reverse transcriptase as one polypeptide. How the separate viral genes are translated from the polycistronic RNA is still a matter of speculation. Futterer et al. (10) listed alternatives by which the normal scanning mechanism of ribosomes might be altered to account for caulimovirus gene expression: (i) formation of a ribosome shunt, (ii) internal initiation, (iii) leaky scanning, and (iv) reinitiation by modifying ribosome behavior. On the basis of data obtained by dissection of certain regions on the leader of the full-length RNA of CaMV, a modified ribosome shunt model was favored to be the mechanism by which ribosomes bypass certain leader sequences that would otherwise inhibit translation (10). However, results describing the translational *trans* activation of the downstream reporter gene on a bicistronic transcript indicated that an alternative mechanism(s) was used for translational events downstream of the small ORFs in the CaMV leader (11). In the discussion below, our findings concerning the expression of the CAT gene in various positions on the FMV genome will be evaluated with respect to the proposed models. For this purpose, CAT expression data are summarized adjacent to diagrams of the relevant portions of the different plasmids in Fig. 6. The data represent reporter gene expression, and although confirmation from quantitative RNA analyses is needed, it seems reasonable to assume that the data reflect translation efficiencies since all the evidence suggests that the CAT gene is expressed from the full-length polycistronic mRNA in all instances.

In pH61 the stop codon for gene VII is preceded by the nearby downstream start codon of the CAT gene, which was inserted inside the small intergenic sequence between ORF VII and ORF I (Fig. 1). Apparently, this configuration, which resembles the close packing of FMV genes, is sufficient for an efficient level of CAT expression. In pH62, ORF II is not fused in frame with the CAT gene, but because of inherent characteristics of the CAT gene cassette, gene II was positioned in frame with a 170-bp alternative 5′-proximal reading frame (designated X in Fig. 6) that overlaps the 5′ end of the CAT coding sequences. Similarly, in pH63a the CAT gene is fused in frame with the first start codon for ORF IV at a site that overlaps with the 3′ end of ORF III (34).

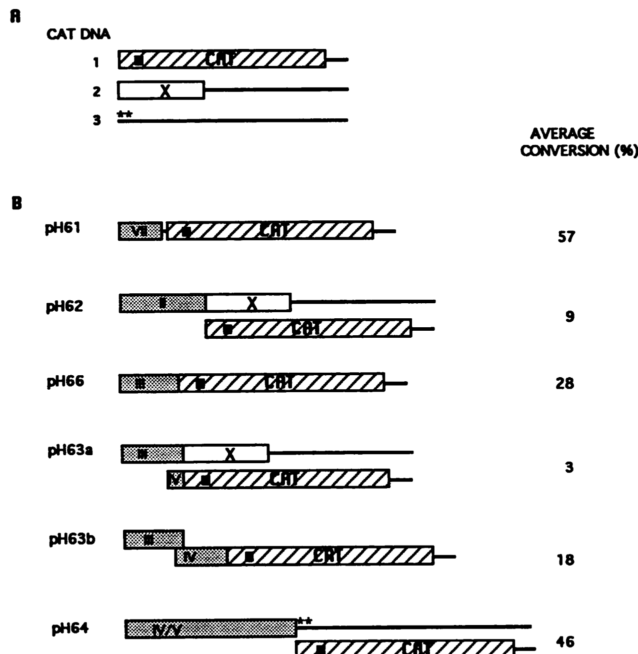

FIG. 6. Positional effects of the CAT gene, with respect to the viral ORFs, on gene expression. (A) Schematic representation of the three reading frames (Arabic numbers) that are present on the DNA segment containing the CAT gene. The box with diagonal lines represents the ORF for the CAT protein, and the small black box indicates the endogenous start codon. The open box denotes a small ORF (designated X) on the DNA, which overlaps the ORF for CAT for 170 bp. The third reading frame (no. 3) contains two stop codons (*) which are positioned immediately upstream of the start codon for CAT (black box). (B) Relative expression of CAT at the different positions on the polycistronic mRNA produced by plasmids pH61 through pH66; only the relevant portion of these plasmids is shown. The shaded boxes with the Roman numerals represent the viral ORFs. The reading frame of the CAT DNA that is positioned in frame with the viral ORF is indicated in each case. The column on the right contains the average acetylation ratios from three (pH62, pH63b, and pH64) or two (pH61, pH66, and pH63a) separate experiments.

Consequently ORF III is fused with ORF X (Fig. 6). The data in Fig. 6 show that pH62 and pH63a produce low levels of CAT expression. The low expression of pH63a is not due to an overall low expression of ORF IV, because when the CAT gene is inserted 100 bp farther downstream (pH63b), in frame with ORF IV, it is expressed efficiently. Neither is it likely that structural alterations, due to the CAT gene insertion, are inhibiting translation of the polycistronic mRNA transcribed from pH63a, because when the reporter gene is located at the identical site but in frame with ORF III in pH66, gene expression is not inhibited (Fig. 6). The relatively low CAT expression from pH62 and pH63a indicates that an upstream ORF (ORF X [Fig. 6]) overlapping CAT for a long distance has a strong inhibitory effect on CAT expression. Similar results were recently obtained with CaMV (11). Relatively high levels of expression are apparent for pH64 and its derivatives (Fig. 4 and 5) (42) in which the CAT gene is positioned as a separate cistron 1,430 bp into ORF IV/V (Fig. 6). In these constructs the start codon of the CAT gene is preceded by two stop codons located 4 and 13 bp upstream. The short spacer sequence between tandem genes resembles that of the native genes of FMV.

The results suggest that the downstream gene of two consecutive genes in FMV is most efficiently expressed when its start codon is close to the stop codon of the preceding ORF, resulting in a tight arrangement of genes, as is the case for pH61, pH63b, pH64, and pH66 but not for pH63a and pH62. These data are in agreement with the observation that a foreign gene in ORF II of CaMV was translated efficiently in vivo when the start codon was positioned a few nucleotides downstream from the termination codon of gene I (3).

A ribosome shunt as well as internal initiation would not per se require a tight arrangement of the major genes because overlapping genes or spacer regions could potentially be tolerated provided that internal signals are present for the entry of ribosomes. However, caulimovirus genomes do not contain extensive gene overlaps or spacers, and our transient-expression data indicate that a tight head-to-tail arrangement of closely packed, nonoverlapping genes is favored for high levels of FMV gene expression in vivo. In addition, it is highly unlikely that the proper sequences for a ribosomal jump or internal initiation are present ca. 1,430 bp into the coding region of gene V, as would be required for efficient expression of CAT of pH64 and its derivatives. The results also suggest that nonspecific or nondiscriminating internal initiation at randomly distributed AUGs is not likely, because the CAT expression from different plasmids would have been the same in that case.

It is very probable that in pH61, pH62, and pH64 the endogenous AUG of the CAT gene is used for initiation of translation, although this start codon is not located within the context of FMV initiation sites. It is therefore unlikely that particular characteristics of the initiation sites and surrounding sequences of the major FMV ORFs would be required to allow translation of the viral genes from the polycistronic mRNA. This finding that foreign contexts are efficiently recognized for translation of genes from a caulimovirus polycistronic mRNA is in agreement with results obtained for CaMV by Futterer and Hohn (11).

Leaky scanning of ribosomes seems to occur most frequently when upstream AUGs are positioned in a less favorable context (25). This mechanism is suggested to be used for translation of a tricistronic mRNA of a coronavirus which is characterized by the absence of internal AUGs on the first two ORFs (26). Start codons of most ORFs in caulimoviruses are in a favorable context (1, 22, 34), and AUGs are scattered throughout the cistrons. Therefore, if leaky scanning of ribosomes were to occur on the full-length transcript of these viruses, it would probably involve a polar effect of gene expression in which the most distal ORFs receive the fewest ribosomes, and consequently the expression of downstream genes should be lower than the expression of those further upstream. Such a trend is not apparent from our data with FMV, since the expression of CAT upstream of ORF I in pH61 is comparable to that of pH64, which has CAT within ORF IV/V (Fig. 6). However, in the absence of the *trans* activator (gene VI) (Fig. 5, lanes containing ΔS plasmids plus pGS1), the CAT gene is expressed less efficiently in the most 3′-proximal ORFs than when it is inserted in ORFs toward the 5′ end of the full-length transcript. This indicates that a polarity of translation occurs only when the *trans*-activating gene VI protein is absent. In that case the translation of downstream ORFs seems to depend solely on the efficiency of reinitiation by ribosomes. Reinitiation is known to become less efficient the farther downstream the gene is positioned on the mRNA (25). However, with the full-length transcript of FMV, it is

important to note that this putative polar effect is lost in the presence of gene VI (Fig. 4 and 5, lanes containing ΔS plasmids plus pGS1RVI).

Our results are in good agreement with the "relay-race" model for translation of the major genes on caulimovirus RNA, as was originally proposed by Sieg and Gronenborn (46), in which ribosomes do not dissociate after translation of one ORF but continue, or reach back over small distances, for reinitiation of translation of the subsequent ORF. In eukaryotic cells the following characteristics, as observed in the present investigation, support this model: (i) the presence of a polycistronic mRNA on which tightly packed but nonoverlapping genes are expressed most efficiently; closely packed genes are also a requirement for reinitiation of translation in mammalian cells (30, 31); (ii) the high expression of CAT as a separate but closely spaced cistron ca. 1,430 bp inside gene V; (iii) the negative effect of gene overlaps on expression of the downstream ORF in such an arrangement; (iv) the lack of specific viral sequence requirements for translational initiation of downstream cistrons; and (v) the absence of a polar effect of gene expression.

We speculate that at least one component of the *trans* activation by the gene VI product involves an interaction which allows a coupled translation of each of the genes without disengagement of ribosomes from the mRNA. It is not clear whether the gene VI product has a quantitative or temporal regulatory function for differential gene expression during the replication cycle. It is possible that the small overlaps which are present for a few caulimovirus genes have a regulatory role in controlling the level of gene expression.

ACKNOWLEDGMENTS

We thank Jennifer Kiernan for providing the cell suspension cultures, Joe Wolff for modified pUC119 plasmids, and Karen-Beth Goldberg Scholthof for providing the gene VI transgenic *D. innoxia* and for critical reading of the manuscript. We are grateful to the reviewers for their constructive comments.

This work was supported by U.S. Department of Agriculture Competitive Grants 88-37263-3847, 87-CRCR-1-2528, and 90-37267-5307. We are also indebted to Kentucky State THRI project 5-41086 for facilities and support.

REFERENCES

1. **Baughman, G., and S. H. Howell.** 1988. Cauliflower mosaic virus 35S RNA leader region inhibits translation of downstream genes. Virology **167:**125–135.
2. **Bonneville, J. M., H. Sanfacon, J. Futterer, and T. Hohn.** 1989. Posttranscriptional *trans*-activation in cauliflower mosaic virus. Cell **59:**1135–1143.
3. **Brisson, N., J. Paszkowski, J. R. Penswick, B. Gronenborn, I. Potrykus, and T. Hohn.** 1984. Expression of a bacterial gene in plants using a viral vector. Nature (London) **310:**511–514.
4. **Chang, L.-J., P. Pryciak, D. Ganem, and H. E. Varmus.** 1990. Biosynthesis of the reverse transcriptase of hepatitis B viruses involves *de novo* translational initiation not ribosomal frameshifting. Nature (London) **337:**364–367.
5. **Covey, S. N.** 1991. Pathogenesis of a plant pararetrovirus: CaMV. Semin. Virol. **2:**151–159.
6. **Covey, S. N., G. P. Lomonosoff, and R. Hull.** 1981. Characterization of cauliflower mosaic virus DNA sequences which encode major polyadenylated transcripts. Nucleic Acids Res. **9:**6735–6747.
7. **Dixon, L. K., and T. Hohn.** 1984. Initiation of translation of the cauliflower mosaic virus genome from a polycistronic mRNA: evidence from deletion mutagenesis. EMBO J. **3:**2731–2736.
8. **Dudley, R. K., J. T. Odell, and S. N. Howell.** 1982. Structure and 5′-termini of the large and 19S RNA transcripts encoded by the cauliflower mosaic virus genome. Virology **117:**19–28.
9. **Futterer, J., K. Gordon, P. Pfeiffer, H. Sanfacon, B. Pisan, J.-M. Bonneville, and T. Hohn.** 1989. Differential inhibition of downstream gene expression by the cauliflower mosaic virus 35S RNA leader. Virus Genes **3:**45–55.
10. **Futterer, J., K. Gordon, H. Sanfacon, J.-M. Bonneville, and T. Hohn.** 1990. Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. EMBO J. **9:**1697–1707.
11. **Futterer, J., and T. Hohn.** 1991. Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. EMBO J. **10:**3887–3896.
12. **Gardner, R. C., A. J. Howarth, P. Hahn, M. Brown-Luedi, R. J. Shepherd, and J. Messing.** 1981. The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. Nucleic Acids Res. **9:**2871–2888.
13. **Goldberg, K.-B., J. Kiernan, and R. J. Shepherd.** 1991. A disease syndrome associated with expression of gene VI of caulimoviruses may be a nonhost reaction. Mol. Plant Microbe Interact. **4:**182–189.
14. **Gowda, S., H. B. Scholthof, F. C. Wu, and R. J. Shepherd.** 1991. Requirement of gene VII in *cis* for the expression of downstream genes on the major transcript of figwort mosaic virus. Virology **185:**867–871.
15. **Gowda, S., F. C. Wu, H. B. Scholthof, and R. J. Shepherd.** 1989. Gene VI of figwort mosaic virus (caulimovirus group) functions in posttranscriptional expression of genes on the full-length RNA transcript. Proc. Natl. Acad. Sci. USA **86:**9203–9207.
16. **Gowda, S., F. C. Wu, H. B. Scholthof, and R. J. Shepherd.** 1990. Gene VI of figwort mosaic virus activates expression of internal cistrons of the full-length polycistronic RNA transcript, p. 79–88. *In* T. P. Pirone and J. G. Shaw (ed.), Viral genes and plant pathogenesis. Springer-Verlag, New York.
17. **Gronenborn, B.** 1987. The molecular biology of cauliflower mosaic virus and its application as plant gene vector, p. 1–29. *In* T. Hohn and J. Schell (ed.), Plant DNA infectious agents. Springer-Verlag, New York.
18. **Guilley, H., R. K. Dudley, G. Jonard, E. Balazs, and K. Richards.** 1982. Transcription of cauliflower mosaic virus DNA: detection of promoter sequences, and characterization of transcripts. Cell **30:**763–773.
19. **Hasegawa, A., J. Verver, A. Shimada, M. Saito, R. Goldbach, A. Van Kammen, K. Miki, M. Kameya-Iwaki, and T. Hibi.** 1989. The complete sequence of soybean chlorotic mottle virus DNA and the identification of a novel promoter. Nucleic Acids Res. **17:**9993–10013.
20. **Hiroshika, H., H. Takatsuji, A. Ubasawa, and J.-E. Ikeda.** 1985. Site-specific deletion in cauliflower mosaic virus DNA: possible involvement of RNA splicing and reverse transcription. EMBO J. **4:**1673–1680.
21. **Hohn, T., J.-M. Bonneville, J. Futterer, K. Gordon, J. Jiricny, S. Karlsson, H. Sanfacon, M. Schultze, and M. de Tapia.** 1990. The use of 35S RNA as either messenger or replicative intermediate might control the cauliflower mosaic virus replication cycle, p. 67–78. *In* T. P. Pirone and J. G. Shaw (ed.), Viral genes and plant pathogenesis. Springer-Verlag, New York.
22. **Hull, R., J. Sadler, and M. Longstaff.** 1986. The sequence of carnation etched ring virus DNA: comparison with cauliflower mosaic virus and retroviruses. EMBO J. **5:**3083–3090.
23. **Jefferson, R. A., T. A. Kavanagh, and M. W. Bevan.** 1987. GUS fusions: β-glucoronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:**3901–3907.
24. **Jones, R. W., A. O. Jackson, and T. J. Morris.** 1990. Defective-interfering RNAs and elevated temperature inhibit replication of tomato bushy stunt virus in inoculated protoplasts. Virology **176:**539–545.
25. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
26. **Liu, D. X., D. Cavanagh, P. Green, and S. C. Inglis.** 1991. A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. Virology **184:**531–544.
27. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

28. **Mason, W. S., J. M. Taylor, and R. Hull.** 1987. Retroid virus genome replication. Adv. Virus Res. **32:**35–96.
29. **Medberry, S. L., B. E. L. Lockhardt, and N. O. Olszewski.** 1990. Properties of commelina yellow virus's complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. Nucleic Acids Res. **18:**5505–5513.
30. **Peabody, D. S., and P. Berg.** 1986. Termination-reinitiation occurs in the translation of mammalian cell mRNAs. Mol. Cell. Biol. **6:**2695–2703.
31. **Peabody, D. S., S. Subramani, and P. Berg.** 1986. Effect of upstream reading frames on translation efficiency in simian virus 40 recombinants. Mol. Cell. Biol. **6:**2704–2711.
32. **Penswick, J., R. Hubler, and T. Hohn.** 1988. A viable mutation in cauliflower mosaic virus, a retroviruslike plant virus, separates its capsid protein and polymerase genes. J. Virol. **62:**1460–1463.
33. **Plant, A. L., S. Covey, and D. Grierson.** 1985. Detection of a subgenomic mRNA for gene V, the putative reverse transcriptase gene of cauliflower mosaic virus. Nucleic Acids Res. **13:**8305–8321.
34. **Richins, R. D., H. B. Scholthof, and R. J. Shepherd.** 1987. Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Res. **15:**8451–8466.
35. **Rosen, C. A.** 1991. Regulation of HIV gene expression by RNA-protein interactions. Trends Genet. **7:**7–14.
36. **Samuel, C. E.** 1989. Polycistronic animal virus mRNAs. Prog. Nucleic Acid Res. Mol. Biol. **37:**127–153.
37. **Sanger, M., S. Daubert, and R. M. Goodman.** 1990. Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous promoter from cauliflower mosaic virus and the regulated mannopine synthase promoter. Plant. Mol. Biol. **14:**433–443.
38. **Schoelz, J. E., and R. J. Shepherd.** 1988. Host range control of cauliflower mosaic virus. Virology **162:**30–37.
39. **Scholthof, H. B.** 1990. Regulatory mechanisms and elements involved in gene expression of figwort mosaic virus, a caulimovirus. Ph.D. dissertation. University of Kentucky, Lexington.
40. **Scholthof, H. B., F. C. Wu, S. Gowda, and R. J. Shepherd.** 1991. Figwort mosaic virus (caulimovirus) RNA contains *cis*-acting sequences within gene VI involved in transactivation of upstream genes. Phytopathology **81:**1185.
41. **Scholthof, H. B., F. C. Wu, R. S. Ranu, S. Gowda, J. M. Kiernan, and R. J. Shepherd.** 1990. Expression of a polycistronic transcript of figwort mosaic virus (caulimovirus) *in vivo* and *in vitro*, abstr. W82-8, p. 126. Abstr. VIIIth Int. Congr. Virol. 1990.
42. **Scholthof, H. B., F. C. Wu, R. D. Richins, and R. J. Shepherd.** 1991. A naturally occurring deletion mutant of figwort mosaic virus (caulimovirus) is generated by RNA splicing. Virology **184:**290–298.
43. **Schultze, M., T. Hohn, and J. Jiricny.** 1990. The reverse transcriptase gene of cauliflower mosaic virus is translated separately from the capsid gene. EMBO J. **9:**1177–1185.
44. **Sharp, P. A., and R. A. Marciniak.** 1989. HIV TAR: an RNA enhancer. Cell **59:**229–230.
45. **Shepherd, R. J., R. D. Richins, J. E. Duffus, and M. K. Handley.** 1987. Figwort mosaic virus: properties of the virus and its adaptation to a new host. Phytopathology **77:**1668–1673.
46. **Sieg, K., and B. Gronenborn.** 1982. Evidence for a polycistronic messenger RNA encoded by CaMV, p. 184. Abstr. NATO/FEBS Adv. Course: Struct. Funct. Plant Genomes 1982.
47. **Temin, H. M.** 1985. Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts. Mol. Biol. Evol. **2:**455–468.
48. **Verwoerd, T. C., B. M. M. Dekker, and A. Hoekema.** 1989. A small scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Res. **17:**2362.
49. **Wu, F. C., S. Gowda, H. B. Scholthof, and R. J. Shepherd.** 1988. Comparative analysis of caulimovirus promoters in protoplasts. Phytopathology **78:**1517.