

Gene VI of figwort mosaic virus (caulimovirus group) functions in posttranscriptional expression of genes on the full-length RNA transcript

(RNA 5' leader/caulimovirus gene VII/translational transactivation)

SIDDARAME GOWDA, FANG C. WU, HERMAN B. SCHOLTHOF, AND ROBERT J. SHEPHERD

Department of Plant Pathology, University of Kentucky, Lexington, KY 40546

Contributed by Robert J. Shepherd, September 1, 1989

ABSTRACT Experimental evidence for a molecular function for gene VI of the caulimoviruses is presented. Based on experiments with the figwort mosaic virus (FMV), it appears that gene VI has a role in the posttranscriptional expression of the closely packed genes (VII and I–V), which appear on the larger, full-length RNA transcript of this virus. Gene VI with its flanking 5'/3' expression signals included as a separate plasmid during electroporation of DNA into protoplasts of *Nicotiana edwardsonii* shows an unusual type of transactivation of a chloramphenicol acetyltransferase (CAT) gene fused at its 5' end to a small open reading frame (gene VII) of the long 5' leader of the full-length RNA transcript of the FMV genome. The level of activity of the CAT gene is increased up to 20-fold over the activity of control plasmids when gene VI is included in the electroporation mixture. Mutagenesis of the coding portions of gene VI of pGS1 RVI, a transactivating plasmid used in the electroporation experiments, demonstrated that it was probably the polypeptide product of gene VI that was responsible for the transactivating effect. Experiments with various portions of the 5' leader of the large, full-length RNA of FMV showed that the coding region of gene VII is necessary for the transactivation event. Clones of cauliflower mosaic virus (CaMV) or FMV with intact gene VI were found to reciprocally transactivate gene VII–CAT fusions (FMV) or gene I–CAT fusions (CaMV) located downstream of the 5' leader sequences of either viral genome.

The mechanism by which the caulimoviruses express the five closely packed genes that appear on the large, full-length viral RNA transcript has been the subject of much speculation. These genes are spaced with only one or two nucleotides between the stop codon of one region and the start codon of the next, or they have short overlaps of a few shared nucleotides between successive genes (1–4). Moreover, the sequence ATGA, in which the stop codon of the upstream region overlaps by one nucleotide the start codon of the neighboring downstream region, is common at the junction of successive genes of the full-length RNA transcript (3). These features of the genomes of caulimoviruses plus the influence of nonsense mutations on the stability of mutant genomes have suggested to some investigators that ribosomes may not dissociate from the messenger RNA at the stop codon of an open reading frame but instead scan for short distances for a start codon on which translation can be reinitiated (5, 6). Observations bearing on translation of the full-length transcript using a newly described member of the caulimovirus group [figwort mosaic virus (FMV)] (7) are described here.

The FMV is a double-stranded DNA virus with a genome of ≈ 8 kilobase pairs (kbp) (4, 7). The genetic organization of this virus is similar to that of cauliflower mosaic virus

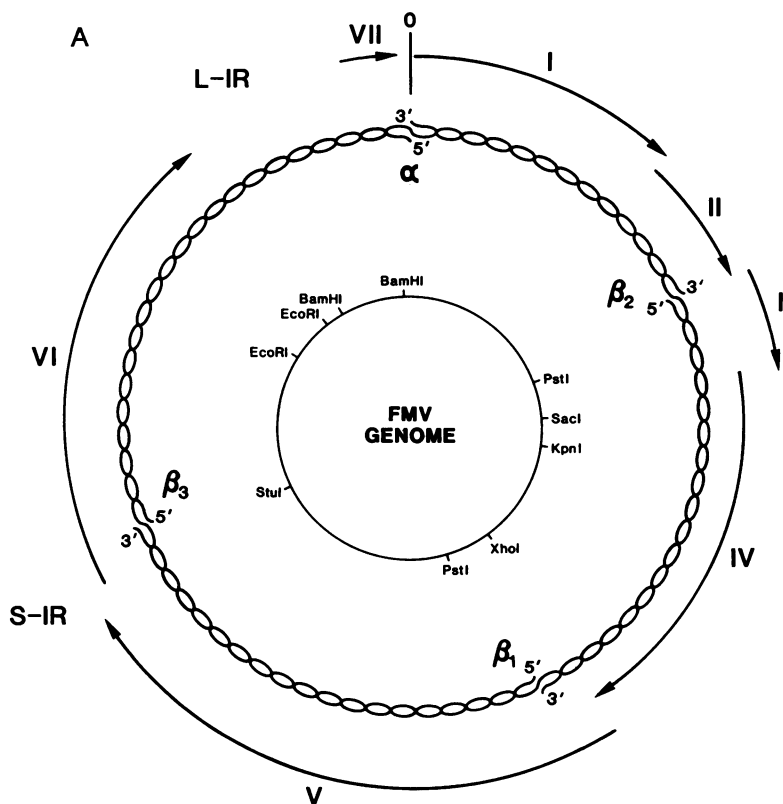
(CaMV) (1, 2), the type member of the caulimovirus group. The genome consists of six major genes (numbered I–VI), which are conserved to a considerable degree in other viruses of the group (Fig. 1A), plus a smaller open reading frame (gene VII), which is not conserved. The latter gene occurs downstream of a large intergenic region of ≈ 600 bp (Fig. 1A). The large intergenic region, like that of CaMV, contains a promoter that leads to the production of a full-length RNA transcript that spans the entire virus genome (8). A smaller RNA transcript arises from a promoter in the smaller intergenic region (117 bp) between genes V and VI (Fig. 1A). The latter RNA spans only the gene VI region of the genome.

In spite of the well-documented biological roles of gene VI of the caulimoviruses as a determinant of disease and host range (7, 9–13), its molecular function is obscure. However, it is the only viral gene transcribed as a separate transcript from its own promoter (14, 15). This suggests that gene VI may have some crucial early role in the infection process, perhaps as a forerunner of the expression of other viral genes. In this report, we present evidence that gene VI has such a role in the posttranscriptional expression of the closely spaced genes of the full-length RNA transcript of FMV.

MATERIALS AND METHODS

Construction of Recombinant Clones. Plasmids pFMV Sc3 and pCaMV10 are infectious full-length clones of FMV and CaMV in pUC7 and pBR322, respectively (2, 4). pFMV RVI contains region VI of FMV with its homologous promoter and termination sequences. This plasmid was constructed by cloning the *EcoRV* fragment [positions 4436–7314 of the FMV genomic map (4)] into the *Sma* I site of pUC119 (16). pGS1 RVI contains the coding region of FMV gene VI between the CaMV 35S promoter and a ribulose-bisphosphate carboxylase gene 3' terminator sequence. The latter was constructed as follows: A segment of DNA containing the CaMV 35S promoter and ribulose-bisphosphate carboxylase 3' termination sequence from the transformation vector pKYLX 7 (17) was cut with *EcoRI* and *Cla* I and cloned into pJAW60 (a derivative of pUC119 with *Pst* I/*Hind*III restriction sites deleted in the polylinker region) at the *EcoRI*/*Acc* I window. The resulting plasmid pGS1 was digested with *Hind*III and *Bam*HI and ligated with the *Hind*III and *Bam*HI fragment of pKB29-6 μ . The latter was generated by creating a *Hind*III restriction site at position 5310 (of the FMV genomic map) by oligonucleotide mutagenesis (18) of pKB29. pKB29 was constructed by cloning a *Hind*III fragment of the FMV genome (positions 4960–7142) into pUC119.

pFMV20 CAT contains the chloramphenicol acetyltransferase (CAT) gene attached to the promoter of the full-length RNA transcript of FMV. This plasmid expresses CAT ac-



B

```

CAGCTGGCTGTGGGGACCAGACAAAAAAGGAATGCTGCAGAAATGTTAG 6739
GCCACCTACCAAAAGCATCTTTGCCTTTATTTGCAAGATAAAGCAGATT 6789
CCTCTAGTACAAGTGGGGAACAAAATAACCTGGAAGAGCTGTCTCTGAC 6839
AGCCCACTCACTAATGCGTATGACGAACGCGATGACGACCACAAAAGAAT 6889
TCCTCTATATAAGGAGCCATTCCCATTTGAAAGGATCATCAGATAC 6939
TGAAACAATATTTCTCACTAAGAAATTAAGAGCTTTGGATTCTTCAAT 6989
GAGAGGCTAAGACCCATAAGAGTTTCGAAAGAGAAATGTAGTATAGTAAG 7039
AGTCTCCAGTCCGGGAGATTGTAATAAGAGATCTTGTAAATGGATCCA 7089
AGTGTCTGTAATTTTGGAAAAATGATCTATAAAATATTCAATCTTCT 7139
TTAAGCTTATCAAGAACAACATACTATCTATCATCCAAATCCACAGA 7189
GTGACAGAGAGAAAATGGTCTGTGTTGTGGATCTGAAGTACCCCGAG 7239
GCAGGAGCCGTTAGGGAAAAGGACTGTTTGGACCTCAAAGTATCAG 7289
GCTGGCTTAGGAAGGAAGATGAAGATATCAGGATTTGGTTTATGTTCTA 7339
AAAAATAAGTAATAAAGAAAAGTTTATTAAGAAAGAAAATTTTATCAAG 7389
AGCAAAATFACATGTCTAGAGGATACCTAGATCTATATTACAATACTTA 7439
CTTACATGTTTTATTCGTGACTCAAATTAAGAAAATTTGTTAATGTTT 7489
ATTCAAACATGTCAGGACTAACCCAGCAAGAGTATATACTCTTAG 7539
CACACCTTATTTTCAGGTAAGTCAAGCAAGCAGGTACAACCTCAT 7589
TCAGGAGACTTCCAGTTTCTCAGAACTCTATATGCTAGGCTTAAACGGCT 7639
TCGGTCACACCAAGCTCATCTCCAAGCAGAAATTCAGCTGTTTCTCAAC 7689
ACGGCAATTCAGGCTCATGAACCTCAAGAAAATCTCGGATCC 7732

```

FIG. 1. (A) Physical map of the FMV genome (4). The circular double-stranded DNA of 8 kbp is indicated by the interwoven lines. Virion DNA has four interruptions, one in the minus strand (designated α) and three in the complementary strand (designated β_1 , β_2 , and β_3). The peripheral arrows indicate the location of the open reading frames and positions of the small (S-IR) and the large (L-IR) intergenic regions. (B) Sequence of the large intergenic region of FMV, including the 3' end of gene VI and the whole of gene VII. "TATA" box, poly(A) signal, and the conserved 35-bp sequence (4) are underlined. The boxed codons represent the translational start and stop sites for gene VII. The numbers on the right correspond to the nucleotide sequence numbers of the FMV genomic map (4).

tivity very well in protoplasts (see Fig. 2). Consequently, it was used as a positive CAT expression plasmid throughout this study. pFMV20 CAT was constructed by cloning a FMV DNA fragment from positions 6690–7003 into pRCAT, which contained the CAT coding region and the ribulose-bisphosphate carboxylase gene 3' termination sequence cloned into pUC119. The other CAT plasmids (see Fig. 3) contain FMV sequences (indicated in parentheses) cloned into pRCAT—i.e., pFMV15 CAT (positions 6690–7105), pFMV19 CAT (positions 6690–7223), pFMV1 CAT (positions 6690–7316), pFMV10 CAT (positions 6690–7504), and pFMV32 CAT (positions 6690–7667).

pGS1 RVI ΔBgl II and pGS1 RVI ΔNsi I (see Fig. 2B) represent truncated versions of gene VI and were constructed by taking advantage of the presence of *Bgl* II and *Nsi* I restriction sites in FMV gene VI at positions 6119 and 6599, respectively. The plasmid pKB29-6 μ was cut with *Hind*III/*Bgl* II (positions 5310–6119) and *Hind*III/*Nsi* I (positions 5310–6599) and the gel-isolated fragments were cloned into pGS1 at *Hind*III/*Bam*HI and *Hind*III/*Pst* I restriction sites to generate pGS1 RVI ΔBgl II and pGS1 RVI ΔNsi I, respectively. By changing the in-frame ATG codons in gene VI by oligonucleotide mutagenesis (18) at position 5163 (to a CGG codon), at positions 5163 and 5430 (to CGG and AGA codons), and at positions 5163, 5430, and 5748 (to CGG, AGA, and GTC codons), respectively, the pGS1 RVI plasmid mutants designated M1, M2, and M3 were constructed (see Fig. 2B).

pS10 CAT 4 was constructed by cloning the *Stu* I/*Ssp* I fragment [positions 6654–43 on the CaMV genomic map (2)] of plasmid pCaMV10 into the *Sma* I site of pRCAT. In this case, the CAT coding region is fused to the first 67 nucleotides of gene I of the CaMV genome.

Electroporation and Assay of CAT Activity. Protoplasts were isolated from cell suspensions of *Nicotiana edwardsonii* cultured in DM-1 medium (19) by described procedures (20). Normally, 20 μ g of supercoiled CAT plasmid DNA was mixed with 2×10^6 protoplasts for electroporation. In coelectroporation experiments, 20 μ g of CAT expression con-

trol plasmid and 50 μ g of transactivating plasmid DNA were used. After 24 hr, 2×10^5 protoplasts were harvested and CAT assays were carried out as described (21). Relative activity of the CAT enzyme was expressed as the ratio of acetylated form to the total acetylated plus unacetylated chloramphenicol. The level of enhancement of CAT activity, referred to as the "activation level" in the figures, refers to the enzyme activity in the protoplasts coelectroporated with the CAT plasmid plus the transactivating plasmid minus the CAT activity of the CAT plasmid alone. This value was then divided by that of the CAT activity obtained with the CAT plasmid alone. The values given are those for 2×10^5 surviving cells.

RESULTS

Transactivation by Gene VI. Preliminary electroporation experiments comparing the expression levels of CAT placed immediately downstream of the FMV promoter for the full-length RNA transcript (pFMV20 CAT) versus that obtained when the reporter gene was placed downstream of the entire large intergenic region of FMV (pFMV32 CAT) show that the 5' leader largely abolished the high levels of CAT expression obtained with the promoter alone (Fig. 2A, lanes 3 and 6). A subsequent experiment with the CAT gene placed downstream at various positions in a series of lengthening FMV intergenic region constructs demonstrated this effect more convincingly (Fig. 3). Inclusion of even short segments of the FMV intergenic region downstream of the polyadenylation signal showed inhibition of CAT activity (Fig. 3). However, in preliminary experiments in which pFMV32 CAT was coelectroporated with pFMV Sc3, a full-length uncut clone of FMC cloned at the *Sac* I site in gene IV (coat protein gene), considerably higher levels of CAT activity were expressed. Also, when pFMV32 CAT was coelectroporated with uncut pFMV M3, cloned at the *Sal* I site in gene VI of FMV (7) (the cloning vector interrupted gene VI), no enhancement of CAT expression was observed. This indicated that an intact gene VI was required for enhanced expression of the CAT fusion

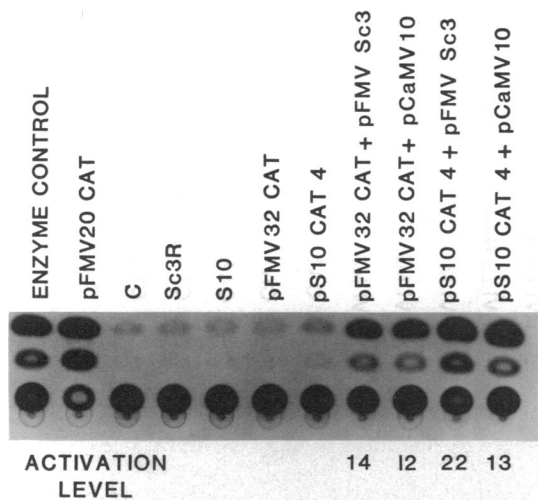


FIG. 4. Reciprocal transactivation between FMV and CaMV. The compositions of pFMV20 CAT and pFMV32 CAT are shown in Fig. 3. The gene VI-containing plasmid of FMV is pFMV Sc3, a full-length genomic clone of FMV in pUC13 (7). The gene VI-containing plasmid of CaMV is pCaMV10, a full-length genomic clone of CaMV in pBR322 (2).

Presumably, the ribosomes initiate well at the next start codon (at position 5430). When ATG codons at positions 5163 and 5430 were changed to CCG and AGA codons, as in plasmid pGS1 RVI M2, the capacity to transactivate was markedly reduced (Fig. 2B, lane 15). Furthermore, when the third in-frame ATG as well, at position 5748, was changed to a GTC codon (along with ATGs at positions 5163 and 5430), the resulting plasmid pGS1 RVI M3 did not induce any detectable transactivation of pFMV32 CAT (Fig. 2B, lane 16).

Effect of 5' Leader and Gene VII Sequences on Transactivation. FMV contains a long 5' leader of ≈ 600 nucleotides similar to that of CaMV (2, 4). To understand whether or not these sequences are required for transactivation by gene VI, a series of recombinant DNAs were constructed containing various lengths of the FMV large intergenic region attached to the CAT gene cartridge (Fig. 3). These were compared in gene VI transactivation experiments with pFMV32 CAT, which contained the entire FMV large intergenic region and the 5' end of gene VII fused to the CAT cartridge at nucleotide 7677 of the FMV genome (Figs. 1 and 3). The 5' end of all these constructs starts at nucleotide 6690 of the FMV genome, which is in the 3' end of gene VI. This region has been observed to contain the enhancer elements of the promoter for the major RNA transcript (unpublished observations). pFMV20 CAT exhibited very high CAT activity, and when coelectroporated with pGS1 RVI, gave no further enhancement of the high level of CAT expression. Coelectroporation of the gene VI plasmid (pGS1 RVI) failed to transactivate any of the other intergenic region constructs (Fig. 3).

pFMV10 CAT, which contains CAT fused to the first five nucleotides of gene VII, exhibited a higher basal level of CAT activity compared to the other plasmids (Fig. 3). When pFMV10 CAT was coelectroporated with pGS1 RVI, it did not show any increase in CAT activity. In contrast, pFMV32 CAT, which has the entire FMV intergenic region and the first 178 nucleotides of gene VII, gave a 10-fold enhancement of CAT activity when coelectroporated with the pGS1 RVI plasmid (Fig. 3). Hence, it is not the 5' leader or gene VII start site but the coding region of gene VII that accounts for the enhancement of gene expression (Fig. 3).

Reciprocal Transactivation Between CaMV and FMV. The coding regions of gene VI of FMV and CaMV show 26%

direct homology in the deduced amino acid sequences. In addition, they share a highly conserved 47-amino acid element (47% identical sequence) near the center of each gene product (4). These similarities suggested the proteins might mutually activate the expression of genes on the full-length RNA transcript. When this was tested by including the plasmid pFMV Sc3, a complete FMV genomic clone with an intact gene VI (7) in the electroporation mixture with pS10 CAT 4, a 22-fold higher level of CAT gene expression in protoplasts was obtained (Fig. 4, lane 10). This same FMV gene VI plasmid when mixed with pFMV32 CAT gave a 14-fold enhancement of CAT expression (lane 8). In the reciprocal case in which a complete genomic clone of CaMV with an intact gene VI (i.e., pCaMV10) was added to the electroporation mixture with pFMV32 CAT, a 13-fold increase in CAT expression was obtained (lane 9). The same level of increase in CAT expression was observed when the pCaMV10 plasmid was added to its homologous pS10 CAT 4 construct and electroporated into cells of *N. edwardsonii* (Fig. 4).

DISCUSSION

Our results document a role for gene VI of the caulimoviruses as an activator for the posttranscriptional expression of the major conserved genes of the polycistronic full-length viral transcript. Expression of a reporter gene fused with open reading frames downstream from the long 5' leader of the full-length transcript is greatly enhanced when plasmids containing gene VI are coelectroporated simultaneously into plant protoplasts. Elevated levels of expression of the reporter gene up to 20-fold above controls without gene VI have been obtained in many of our experiments.

Enhanced expression with CAT as gene fusions at two major sites downstream from the long 5' leader of either FMV (gene VII) or CaMV (gene I), have been obtained. In additional work in this laboratory with the CAT gene in positions further downstream, in fact as far removed as gene V, efficient expression is observed only when gene VI is included during the electroporation experiment. Consequently, it appears that the activation of response can occur with genes in virtually any downstream position from the 5' leader.

Hohn *et al.* (24) have also obtained an enhanced expression of CAT as either a gene VII or gene I fusion placed downstream of the intergenic region of CaMV when gene VI constructs were simultaneously electroporated into plant protoplasts. In their tests, even higher levels of expression were obtained with gene VI under the control of the CaMV 35S promoter compared to its native 19S promoter. In our experiments with FMV, the homologous promoter of gene VI gave higher levels of transactivation than the 35S promoter of CaMV (Fig. 2A). In other trials with the gene VI promoter of FMV fused to the CAT gene, higher levels of expression were obtained than with CAT fused to the promoter of the full-length RNA of FMV.

Mutagenesis of gene VI on plasmids prior to use in coelectroporation experiments suggests that it is the polypeptide product that is active in transactivation rather than the RNA transcript of gene VI. Mutations that remove the carboxyl terminus of the protein, or, alternatively, cause truncation beginning at the amino terminus of the polypeptide, lead to loss of transactivating potential of the gene. It seems rather doubtful that the relatively minor nucleotide changes made in removal of start codons toward the 5' terminus of the gene would destroy the transactivation response if the RNA transcript of gene VI was the active molecule.

Our experiments suggest that gene VII of FMV is the relevant portion of the 5' leader that responds to gene VI protein during transactivation. A DNA construct with CAT

fused to the start codon of gene VII (pFMV10 CAT) shows no enhancement of gene expression when gene VI plasmids are included during electroporation. However, the basal level of expression with this plasmid is consistently higher when electroporated alone into cells (Fig. 3). In contrast, CAT fused to the distal 3' end of gene VII (pFMV32 CAT) shows high levels of transactivation by gene VI (Fig. 3), suggesting that gene VII participates in the transactivation event. The observation (Fig. 4) that gene VI of CaMV and FMV can participate in elevated expression of genes on the full-length transcript of either virus argues that the mechanism of enhancement is the same in both cases. Consequently, it will be interesting to determine which sequences are responsible for transactivation in the 5' leader of CaMV.

It is difficult to reconcile the requirement for the coding region of gene VII for transactivation of FMV and various observations on the dispensability of gene VII for infectivity of CaMV. Gene VII of CaMV is reportedly not required for virus infectivity (25). However, such mutants of CaMV show long delays in symptom development following inoculation to plants (30–60 days postinoculation). Moreover, mutants of CaMV in which the start codon of gene VII is mutated (ATG → ACG) are infectious but revert at high frequency to the wild-type viral sequence. This suggests that gene VII may have some regulatory role in the replication cycle (26).

The FMV genome, like that of CaMV, has numerous small open reading frames in the 5' leader of the full-length RNA transcript. Five of these, varying from 3 to 23 amino acids, occur between the polyadenylation signal and the start codon for gene VII (Fig. 1B). These may be largely responsible for the depressing effect of the leader on gene expression similar to that observed with CaMV (24, 27). These reading frames may interrupt the scanning process of 40S ribosome subunits, which must move across the leader before initiating translation on downstream genes. However, the increased expression with pFMV10 CAT, with a slightly longer portion of the 5' leader, compared to the expression of pFMV1 CAT (Fig. 3), suggests that the products of some of these other small open reading frames may influence the translation of downstream genes.

The authors are indebted to Jennifer Kiernan for assistance in the suspension cell cultures and to Karen-Beth Goldberg, Richard Richins, and Dr. Joseph Wolff for some of the virus plasmids used in the investigation. This research was supported by U.S. Department of Agriculture CRGO Grants 87-CRCR-1-2528 and 88-37263-3847.

1. Franck, A., Guilley, H., Jonard, G., Richards, K. & Hirth, L. (1980) *Cell* **21**, 285–294.
2. Gardner, R. C., Howarth, A. J., Hahn, P., Brown-Leudi, M., Shepherd, R. J. & Messing, J. (1981) *Nucleic Acids Res.* **9**, 2871–2888.
3. Hull, R., Sadler, J. & Longstaff, M. (1986) *EMBO J.* **5**, 3083–3090.
4. Richins, R. D., Scholthof, H. B. & Shepherd, R. J. (1987) *Nucleic Acids Res.* **15**, 8451–8466.
5. Gronenborn, B. (1987) in *Plant DNA Infectious Agents*, eds. Hohn, T. & Schell, J. (Springer, New York), pp. 1–29.
6. Gronenborn, B. (1987) in *Plant DNA Infectious Agents*, eds. Hohn, T. & Schell, J. (Springer, New York), pp. 1–29.
7. Shepherd, R. J., Richins, R. D., Duffus, J. E. & Handley, M. K. (1987) *Phytopathology* **77**, 1668–1673.
8. Gowda, S., Wu, F. C. & Shepherd, R. J. (1989) *J. Cell. Biochem. Suppl.* **13-D**, 301 (abstr.).
9. Daubert, S., Shepherd, R. J. & Gardner, R. C. (1983) *Gene* **25**, 201–208.
10. Daubert, S. D., Schoelz, J., Li, D. & Shepherd, R. J. (1984) *J. Mol. Appl. Genet.* **2**, 537–547.
11. Schoelz, J. E., Shepherd, R. J. & Daubert, S. (1986) *Mol. Cell. Biol.* **6**, 2632–2637.
12. Schoelz, J. E. & Shepherd, R. J. (1988) *Virology* **162**, 33–42.
13. Baughman, G. A., Jacobs, J. D. & Howell, S. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 733–737.
14. Covey, S. & Hull, R. (1981) *Virology* **111**, 463–474.
15. Odell, J. & Howell, S. H. (1980) *Virology* **102**, 349–359.
16. Vieira, J. & Messing, J. (1987) *Methods Enzymol.* **153**, 3–11.
17. Schardl, C. L., Byrd, A. D., Benzo, G., Althschuler, M. A., Hildebrand, D. F. & Hunt, A. G. (1987) *Gene* **61**, 1–11.
18. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382.
19. Ranch, J. P. & Giles, K. L. (1980) *Ann. Bot.* **46**, 667–683.
20. Wilson, H. M., Styer, D. J., Conrad, P. L., Durbin, R. D. & Helgeson, J. P. (1980) *Plant Sci. Lett.* **18**, 151–154.
21. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1051.
22. Kozak, M. (1986) *Cell* **44**, 282–292.
23. Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. & Scheele, G. A. (1987) *EMBO J.* **6**, 43–48.
24. Hohn, T., Bonneville, J. M., Fütterer, J., Gordon, K., Pisan, B., Sanfacon, H., Schultze, M. & Jiricny, J. (1989) in *Molecular Biology of Plant-Pathogen Interactions*, UCLA Symposium on Molecular and Cellular Biology, eds. Staskawicz, B., Ahlquist, P. & Yoder, O. (Liss, New York), Vol. 101, pp. 153–165.
25. Dixon, L. K. & Hohn, T. (1984) *EMBO J.* **3**, 2731–2736.
26. Dixon, L., Jiricny, J. & Hohn, T. (1986) *Gene* **41**, 225–231.
27. Baughman, G. & Howell, S. A. (1988) *Virology* **167**, 125–135.