JOURNAL OF VIROLOGY, Dec. 1996, p. 8411–8421 Vol. 70, No. 12
0022-538X/96/$04.00+0

Efficient Transcription from the Rice Tungro Bacilliform Virus Promoter Requires Elements Downstream of the Transcription Start Site

GANG CHEN,[1]† HELEN M. ROTHNIE,[1] XIAOYUAN HE,[1] THOMAS HOHN,[1]*
AND JOHANNES FÜTTERER[2]

Friedrich Miescher-Institut, CH-4002 Basel,[1] and Institute for Plant Science, ETH Zentrum, CH-8092 Zürich,[2] Switzerland

Received 28 March 1996/Accepted 6 September 1996

Elements downstream of the transcription start site enhance the activity of the rice tungro bacilliform virus (RTBV) promoter in protoplasts derived from cultured rice cells. This enhancer region was located to the first 90 nucleotides of the RTBV leader sequence. Within this region, at least two components which act together to enhance expression from the RTBV promoter could be identified. One is a position- and orientation-independent DNA element within a CT-rich region, and the other is a position-dependent element. Either element was found to be capable of acting independently on a heterologous promoter. The enhancer activity of the DNA element correlates with specific binding of nuclear proteins. Nuclear proteins also recognize an RNA transcript covering the first 90 nucleotides of the RTBV leader.

Rice tungro bacilliform virus (RTBV) belongs to the badnaviruses (type member *Commelina* yellow mottle virus), a group of nonenveloped bacilliform viruses with a circular double-stranded DNA genome infecting both mono- and dicotyledonous plants (44). On the basis of sequence homology and similarity in genome organization with caulimoviruses (type member cauliflower mosaic virus [CaMV]), badnaviruses have been suggested to represent a second group of plant pararetroviruses, i.e., nonintegrating DNA viruses replicating via reverse transcription (30, 50, 56). This classification is further supported by recent studies detecting reverse transcription intermediates in infected rice plants (2) and reverse transcriptase activity associated with expressed RTBV proteins (42).

Pararetroviruses, like true retroviruses, produce a terminally redundant transcript which serves as the template for reverse transcription during replication of the viral genome (reviewed in reference 33). This pregenomic RNA also serves as mRNA to produce primary viral products via a variety of different translation mechanisms and can be spliced to produce additional mRNA(s). Thus, production of this transcript plays a central role in the virus infection cycle. Each of these viruses has a "genomic" promoter directing transcription of this RNA (reviewed in reference 57).

The best-studied plant pararetroviral promoter is the CaMV 35S promoter, which is by far the most widely used promoter in the field of plant genetic engineering. This promoter directs the production of a viral transcript of greater than genome length, the 35S RNA. It is highly active in the majority of plant tissues in both dicots and monocots (17, 29, 53). Its strong constitutive expression is based on the additive and synergistic interaction of a set of discrete *cis* elements located upstream of the transcription initiation site (reviewed in references 3 and 32). Nuclear factors binding to various regulatory elements of this promoter have now been identified that are involved in the regulation of expression in specific tissues (reviewed in reference 43).

Similarities in the genomic organization of RTBV and CaMV (30, 56) allowed the probable location of the RTBV promoter to be deduced, and two independent analyses of RTBV transcripts in infected plants mapped the transcription start site to genomic position 7405 (1, 66). In the context of reporter gene fusions, sequences upstream of genomic position 7390 or 7449 were reported to direct phloem-specific expression in young transgenic rice plants and in bombarded tissue (4, 66), consistent with the observation that phloem is the site of virus accumulation in infected plants (4).

In our initial study of the RTBV promoter (11), we defined the active TATA box and demonstrated the absence of strong enhancer elements upstream of the transcription initiation site, at least in the protoplast systems tested. A region of only 100 bp immediately upstream of the transcription start site retained almost full activity, and considerable residual activity is still observed even with only 50 bp of upstream sequence (11). In contrast, a region located at the 5′ end of the RTBV leader sequences, i.e., downstream of the transcription initiation site, was shown to have a dramatic effect on reporter gene expression in protoplasts (11). In this work, we have further characterized this region in our transient-expression system.

Analysis of sequences downstream of the transcription initiation site that control expression is complicated by the multitude of levels at which they might act. The situation is particularly complex in the case of RTBV and other plant pararetroviruses because of the multiple roles of the long leader sequence of the pregenomic virus RNA in translational regulation (20, 30), splicing (21, 41), RNA 3′-end formation (30, 60, 61), replication (33), and possibly packaging (57). The leader sequences in these viruses have evolved extensive and complex secondary structures which may be required to mediate one or more of these processes (31). In RTBV, the positive influence of the 5′ end of the leader varied depending on the upstream promoter sequences used, suggesting that a transcriptional effect rather than a posttranscriptional process was responsible (11). Here we report that this region is indeed involved in RNA accumulation and present evidence that at

* Corresponding author. Mailing address: Friedrich Miescher Institute, P.O. Box 2543, CH-4002 Basel, Switzerland. Phone: (61) 697 72 66. Fax: (61) 697 39 76. Electronic mail address: hohn@fmi.ch.

† Present address: Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, PA 19111.

least two overlapping elements act synergistically to greatly increase the efficiency of transcription from the RTBV promoter in rice protoplasts.

MATERIALS AND METHODS

Construction of plasmids. Plasmids R · I-CAT, C · I-CAT, Cc · I-CAT, RΔCl83 · I-CAT, CΔCl83 · I-CAT, and CcΔCl83 · I-CAT have been described previously (11). R · 8-CAT and R · 25-CAT were derived from R · I-CAT by digestion with *Nco*I (at the 5′ end of the chloramphenicol acetyltransferase [CAT] gene) and either *Cla*I or *Sac*I (within the RTBV leader), respectively, followed by treatment with Klenow or T4 DNA polymerase and religation of the large fragment.

For the construction of R · 50-CAT, a *Cla*I-*Hin*dIII fragment (filled in with Klenow polymerase) corresponding to the first 48 bp of the 5′ end of the RTBV leader plus a *Hin*dIII tail was excised from a cloned PCR amplification product and ligated to the large fragment of plasmid R · I-CAT (*Cla*I-*Nco*I, filled in).

To obtain R · 90-CAT, the context of the ATG codon of short open reading frame (sORF) 1 in the RTBV leader was changed to an *Nco*I site by PCR-mediated point mutagenesis and the mutated fragment was inserted as a *Cla*I-*Nco*I fragment between the *Cla*I and *Nco*I sites of R · I-CAT.

R · Ω-CAT was obtained by PCR amplifying the tobacco mosaic virus Ω leader, introducing *Cla*I and *Nco*I sites at the 5′ and 3′ ends, respectively, and inserting the *Cla*I-*Nco*I fragment from the cloned PCR product between the *Cla*I and *Nco*I sites of R · 90-CAT.

C · 8-CAT was created by generating a PCR product using a 5′ primer corresponding to CaMV 35S promoter sequences upstream of the *Eco*RV site and a 3′ primer immediately upstream of the transcription start site, with the 3′ primer including 8 nucleotides (nt) of the 5′ end of the RTBV leader and an *Nco*I site after the CaMV transcription start site. The PCR product was inserted between the *Eco*RV and *Nco*I sites of C · I-CAT.

C · 50-CAT was produced by cloning the *Cla*I-*Xho*I fragment from R · 50-CAT into the corresponding sites of C · I-CAT.

RΔBaBa · I-CAT (here called R-218 · I-CAT), a promoter deletion mutant of R · I-CAT, has been described previously (11). Cm · I-CAT was generated by inserting a PCR fragment obtained by using C · I-CAT as the template, flanked by *Bam*HI and *Xho*I sites, and with a 5′ end corresponding to the −46 position of the CaMV 35S promoter between the corresponding sites of R · I-CAT. CR · I-CAT was obtained by the ligation of the small *Xho*I–filled-in *Bam*HI fragment of R-218 · I-CAT with the large *Eco*RV-*Xho*I fragment of C · I-CAT.

R-218ΔCl83 · I-CAT, Cm · IΔCl83-CAT, and CR · IΔCl83-CAT were constructed by replacing the original *Cla*I-*Bgl*II fragments of R-218 · I-CAT, Cm · I-CAT, and CR · I-CAT, respectively, with the corresponding fragment from R · I-CATΔCl83.

C · IΔCl50-CAT, R-218ΔCl50 · I-CAT, CcΔCl50 · I-CAT, and CRΔCl50 · I-CAT were produced in the same way as leader deletion mutants ΔCl83 (11): R · I-CAT was linearized at the *Cla*I site, digested with *Bal* 31 exonuclease, and religated in the presence of a *Cla*I linker (CTATCGATAG). From the resulting mixture, a *Cla*I-*Bgl*II fragment of suitable size was selected and introduced into the corresponding sites in plasmids C · I-CAT, R-218 · I-CAT, Cc · I-CAT, Cm · I-CAT, and CR · I-CAT, respectively.

To generate CΔ50-83 · I-CAT, R-218Δ50-83 · I-CAT, CcΔ50-83 · I-CAT, CmΔ50-83 · I-CAT, and CRΔ50-83 · I-CAT, double-stranded oligonucleotides covering the 5′ end of the RTBV leader from positions +20 to +50 (relative to the transcription start site) flanked with *Sac*I and *Cla*I sites, were inserted between the *Sac*I and *Cla*I sites of plasmids CΔCl83 · I-CAT, etc.

The plasmids with the *dps* fragment located upstream of the promoter were created in the following way: *dps* DNA fragments were produced by PCR amplification of R · I-CAT sequences between positions upstream of the *Cla*I site and position +90 or +191, respectively, with a 3′ end primer containing a *Bgl*II site. These fragments were cut with *Cla*I, treated with Klenow polymerase, and then cut with *Bgl*II and inserted between the *Bgl*II and *Xba*I (filled-in) sites of R · I-CAT to create precursor plasmids. To produce the *dps* sense orientation constructs, the precursor plasmids were opened with *Bgl*II and *Xho*I, and a *Bam*HI-*Xho*I fragment from the leader deletion series of R-218 · I-CAT was introduced. To produce *dps* reverse-orientation constructs, *Hin*dIII-*Hin*cII fragments from the precursor plasmids were inserted into *Hin*dIII-*Bam*HI (filled-in) sites of the corresponding leader deletion mutants of R-218 · I-CAT.

For the plasmids with the *dps* subregion (positions +50 to +89) upstream of the R-218 promoter, synthesized oligonucleotides corresponding to original or mutated sequences were inserted between the *Hin*dIII and *Bam*HI sites of the corresponding leader deletions of R-218 · I-CAT.

The substrate for preparation of the antisense RNA probe, RTPA-L, was derived from the expression plasmid R-CAT* (58). A *Sph*I-*Hin*dIII fragment of R-CAT* covering the CaMV-R sequence was replaced with a *Sph*I-*Hin*dIII fragment covering RTBV promoter and leader sequences (positions 7121 to 7859).

A *Hin*dIII-*Bgl*II fragment cloned from the described precursor plasmid with 3′ RTBV fragment ending at position +90 was inserted into the polylinker of plasmid vector pGEM 2 to generate the substrate, P-*dps*, for preparation of both DNA and sense RNA *dps* probes.

All steps in the cloning procedures were performed according to conditions recommended by the suppliers of the respective enzymes and by standard protocols (59). Plasmids were characterized by restriction analysis and/or sequencing of double-stranded DNA with a Sequenase sequencing kit (U.S. Biochemicals). Plasmids were propagated in *Escherichia coli* NM522 or JM109 and purified with a plasmid kit from QIAGEN.

Suspension cultures, protoplast preparation, and transfection. Conditions for growth of suspension cultures of the *Oryza sativa* line Oc and *Orychophragmus violaceus* and preparation of protoplasts have been described previously (11, 18). Protoplasts were transfected by electroporation (*O. violaceus*) or polyethylene glycol-mediated transfection (*O. sativa*) as described previously (17–19, 64). Routinely, 5 to 8 μg of test plasmid DNA was used per transfection, together with 3 μg of a plasmid expressing β-glucuronidase (GUS) under the control of the CaMV 35S promoter to serve as an internal standard.

Transient-expression assays. After incubation overnight (around 20 h post-transfection), CAT activities in protoplast protein extracts were determined and normalized to the GUS activity of the same extract (6, 26, 34). For quantification of CAT activity, assay conditions were chosen such that a maximum of 30% of the radioactive chloramphenicol substrate was acetylated. Relative expression levels did not vary more than ±20% with the given constructs in this study, although the absolute values of the activities measured in protoplasts varied with the season for unknown reasons. All constructs were tested in at least three independent experiments from at least two different isolates of each clone.

RNase A + T_1 mapping. *O. sativa* protoplasts were transfected with C/R · I-CAT or C/RΔCl83 · I-CAT and an internal control plasmid (pDES7) (24). Total RNA was isolated from the protoplasts 18 h after transfection and subjected to RNase A + T_1 protection analysis according to published protocols (25). Radioactively labelled antisense RNA probes were synthesized by in vitro transcription (in the presence of [α-^{32}P]UTP) with T7 or SP6 RNA polymerase using *Pst*I-linearized RTPA-L or *Eco*RI-linearized pGS7 (24) as templates. Protected fragments were resolved on a 6% polyacrylamide denaturing gel and visualized by autoradiography. Radioactivity in the protected fragments was quantified by using a PhosphorImager (Molecular Dynamics), and the relative abundance of transcripts produced from each of the four constructs was calculated relative to the internal control (taking into account the differences in specific activities due to base composition of the protected fragments).

Primer extension. Total RNA was isolated from *O. sativa* protoplasts transfected with C/Cc · I-CAT or C/CcΔCl83 · I-CAT 18 h after transfection (25) and subjected to primer extension analysis as follows. A DNA oligonucleotide (5′-CATATTACACAGAGGTTCC-3′) complementary to positions 7536 to 7554 of the RTBV genomic sequence was end labelled with T4 polynucleotide kinase (Boehringer, Mannheim, Germany) and [γ-^{32}P]ATP. The RNA pellet was resuspended in 10 μl of annealing buffer (20 mM Tris-Cl (pH 8.0), 0.2 mM EDTA, 0.25 M KCl) and 1 μl of labelled oligonucleotide (50,000 cpm) was added. The primer was annealed to the RNA by incubation for 5 min at 65°C and for 30 min at 55°C and then 5 min at room temperature. Reverse transcriptase buffer (20 μl) (20 mM Tris-Cl [pH 7.0], 10 mM $MgCl_2$, 100 μg actinomycin D per ml, 0.5 mM [each] deoxynucleoside triphosphate, and 12.5 mM dithiothreitol [DTT]), 0.5 μl of RNase inhibitor (40 U/μl; Boehringer), and 0.5 μl of avian myeloblastosis virus reverse transcriptase (25 U/μl; Boehringer) were added. Primer extension was carried out at 42°C for 1 h before stopping the reaction with the addition of 150 μl of ethanol. After precipitation, the pellet was resuspended in 5 μl of a 2:1 mixture of formamide dye mixture: 0.1 M NaOH and heated for 3 to 5 min at 95 to 100°C before loading on a 6% sequencing gel.

Preparation of nuclear extracts. Crude nuclear extract was prepared from cell suspension cultures of *O. sativa* line Oc as follows: around 4×10^8 suspension cells or protoplasts were disrupted in 30 ml of homogenization buffer (20 mM morpholineethanesulfonic acid [MES] [pH 6.0], 5 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 10 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride [PMSF], 1 μg of leupeptine per ml, 1 μg of antipain per ml) with around 10 strokes in a Dounce homogenizer. The slurry was filtered sequentially through 60-, 40-, and 10-μm-pore-size nylon mesh filters. The nuclei were pelleted by centrifugation at 1,200 × *g* for 5 min and resuspended in 50 ml of homogenization buffer. The solution was filtered again through 10-μm-pore-size nylon mesh, and the nuclei were collected by centrifugation at 1,500 × *g* for 5 min and taken up in 20 ml of lysis buffer (20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES] [pH 7.5], 50 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 10% [vol/vol] glycerol, 2 mM DTT, 1 mM PMSF, 1 μg of leupeptine per ml, 1 μg of antipain per ml). The nuclei were broken by a few strokes in a Dounce homogenizer and diluted to an optical density at 260 nm of 10. Saturated ammonium sulfate (1/9 of the volume) was added and shaken to mix, and the solution was left to stand for 30 min. After centrifugation at 100,000 × *g* for 90 min, ammonium sulfate (0.33 g/ml) was added to the supernatant. After the solution was gently shaken for 15 min to dissolve the ammonium sulfate completely, the solution was left to stand for 30 min and then centrifuged for 30 min at 20,000 × *g*. The pellet was resuspended in 1 ml of dialysis buffer (same as lysis buffer but without leupeptine and antipain) and dialyzed against 200 ml of dialysis buffer 4 times for 1 h. The dialyzed solution was spun for 10 min at 10,000 × *g*, and the supernatant was collected and stored in aliquots at −80°C. All of the above steps were performed at 4°C.

Crude nuclear extracts from *O. sativa* plant seedlings were prepared by published protocols (48).

Gel retardation assays. To prepare the DNA probe covering the *dps* region, plasmid P-*dps* was linearized with *Bgl*II, labelled with ^{32}P by using the Klenow fragment of DNA polymerase as previously described (51), and then cut by *Hin*dIII and purified by excision from a 5% polyacrylamide gel. Competitor DNA fragments were prepared either by purifying restriction fragments from 1% agarose gels with a gel extraction kit (QIAGEN) or by directly annealing two synthetic complementary oligonucleotides.

Radioactively labelled RNA probes covering the *dps* region were synthesized by in vitro transcription with SP6 RNA polymerase in the presence of [α-^{32}P]UTP with *Bgl*II-linearized P-*dps* as a template. Unlabelled probes and RNAs used for competition assays were produced as described above except that only a trace amount of [α-^{32}P]UTP was used in the reaction mixture, allowing quantification.

Binding reaction mixtures contained crude nuclear extract (3 μl) and 5,000 cpm of labelled DNA or RNA probes (DNA probes are around 0.05 to 0.1 pM and RNA probes are around 5 to 10 fmol) in a final volume of 15 μl of buffer containing 10 mM HEPES (pH 7.6), 8 mM $MgCl_2$, 1 mM DTT, 4 mM spermidine, and 5% (vol/vol) glycerol (62a). Ten micrograms of poly(dI-dC) · poly(dI-dC) (Pharmacia) for DNA gel shift or tRNA (Boehringer) for RNA gel shift were included as unspecific competitors. RNA gel shift assay mixtures also contained 30 U of RNase inhibitor (Boehringer). For the competition assays, variable amounts of competitor as indicated in the figure legends were included. Binding reactions were begun by addition of nuclear extract to the buffer. This reaction mixture was preincubated for 10 min at room temperature and incubated for a further 20 min at room temperature after addition of the probe and the competitors. In RNA gel shift assays, heparin was added to a final concentration of 5 mg/ml after incubation, and the reaction mixture was placed at 4°C for 5 min. Samples were then loaded on a 5 or 7% native polyacrylamide gel (40:1 polyacrylamide-bisacrylamide) in 1× TBE (Tris-borate-EDTA) buffer for DNA or RNA gel shift, respectively. Following electrophoresis at 30 mA for 2 to 3 h at 4°C, gels were dried and autoradiographed.

RESULTS

The first 90 bp of the transcribed region of the RTBV genome affect transcription from the RTBV promoter. In our previous experiments (11), an effect of the first 90 nt of the RTBV leader on transcription was inferred from measuring CAT expression arising from constructs with different upstream promoter sequences fused to the RTBV leader and the CAT open reading frame in transiently transfected protoplasts. To more directly demonstrate the influence of this region on RNA production, transcript levels were quantified by RNase A + T_1 mapping. In the context of RTBV upstream promoter sequences, deletion of positions +8 to +83 from the RTBV leader greatly reduced the steady-state RNA level (Fig. 1). In contrast, if CaMV 35S upstream promoter sequences were used, the same deletion had little effect on the amount of RNA (note that the CaMV 35S promoter is very much stronger than the RTBV promoter in rice protoplasts). In all cases, RNA levels correlated well with the CAT expression data, supporting the conclusion that transcription from the RTBV promoter was specifically affected by the deletion. Although the extreme 5′ ends of the RNA molecules produced from the RTBV and CaMV 35S promoter constructs differ by a few nucleotides, an effect on RNA stability and/or differential translatability is considered unlikely since the region from positions +8 to +83 has a differential effect on promoters directing identical RNA transcripts (11; also see below). Taken together, these results strongly suggest that elements enhancing the activity of the RTBV promoter are present within the first 90 nt of the transcribed region. From now on, we will refer to this region as *dps* (*dps* for downstream promoter sequence).

It could be argued that an effect of *dps* in the context of the CaMV 35S promoter might be masked as a result of saturation of the transient-expression system with this strongly expressing construct. However, the amount of input DNA used in all our experiments was within the range of linear relationship between input DNA and CAT activity, i.e., nonsaturating conditions (data not shown).

The 90-bp *dps* fragment is sufficient to restore RTBV promoter activity. To determine if the *dps* fragment defined above

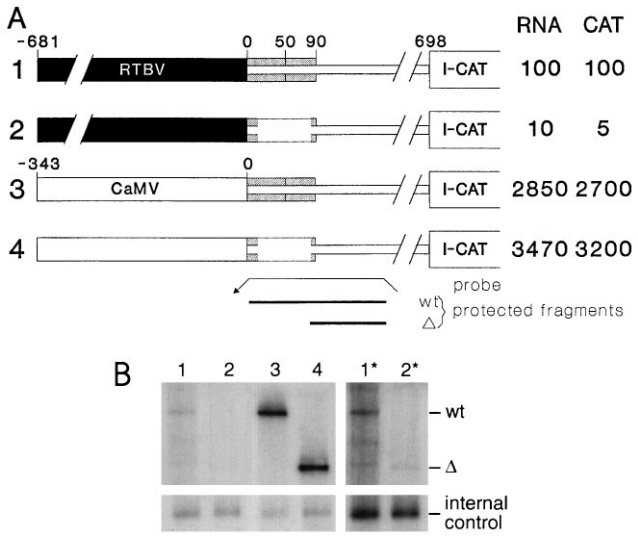


FIG. 1. Downstream sequences are required for RTBV promoter activity. (A) Constructs consisting of the RTBV (to position −681) or CaMV 35S (to position −343) upstream promoter sequences, the RTBV leader sequence (either complete or with a deletion between +8 and +83), and the CAT reporter gene fused to RTBV open reading frame I are shown schematically. All constructs are terminated by the CaMV polyadenylation signal. The range defined as downstream promoter sequence (positions 0 to +90; *dps*) is indicated by a shaded box. Relative RNA values from RNase protection experiments and CAT activity values obtained from transfected *O. sativa* protoplasts are given. The locations of the antisense probe used for RNase A + T_1 mapping and the protected fragments (wild type [wt] and with the deletion [Δ]) are indicated. The plasmids used were R · I-CAT (row 1), RΔCl83 · I-CAT (row 2), C · I-CAT (row 3), and CΔCl83 · I-CAT (row 4). (B) RNase protection analysis. Total RNA isolated from transfected *O. sativa* protoplasts was subjected to RNase A + T_1 protection analysis with the probe shown in A. Lanes 1 to 4, RNA from protoplasts transfected with the internal control plus the four constructs shown above (in order), Lanes 1* and 2* are longer exposures of lanes 1 and 2. Protected fragments corresponding to RTBV and internal control transcripts are indicated.

is sufficient to account for the observed effect on expression or whether other leader sequences are also involved, the entire leader sequence was deleted from our expression construct and 5′-end RTBV leader sequences of different lengths were reintroduced (Fig. 2). These constructs also allowed assessment of the effect of the *dps* fragment independently of the possible posttranscriptional effects of the entire leader.

Reintroducing the first 8, 25, or 50 bp of the RTBV leader between the RTBV upstream promoter sequences and the CAT gene did not restore CAT activity (Fig. 2, rows 2 to 4). However, when the 90-bp *dps* fragment was reintroduced, CAT activities returned to wild-type levels (Fig. 2, row 5). A detrimental effect of the length or structure of the shorter RNA leaders on CAT expression can be excluded, since high levels of CAT activity were observed with equivalent CaMV 35S promoter constructs (Fig. 2, rows 7 to 9), demonstrating that even a leader as short as 8 nt is sufficient to allow efficient translation. Also, replacement of RTBV sequences by the tobacco mosaic virus Ω leader sequences, which are unstructured and act as a translational enhancer (reference 22 and references therein), did not result in detectable CAT activity (Fig. 2, row 6). The results presented in Fig. 2 reveal that the 90-bp *dps* fragment has a dramatic effect on CAT expression in the context of the RTBV promoter, while the effect on the CaMV 35S promoter is minimal, thus providing further evidence that the RTBV leader contains sequences that specifically enhance RTBV promoter activity and that these sequences are contained within the 90-bp *dps* fragment.

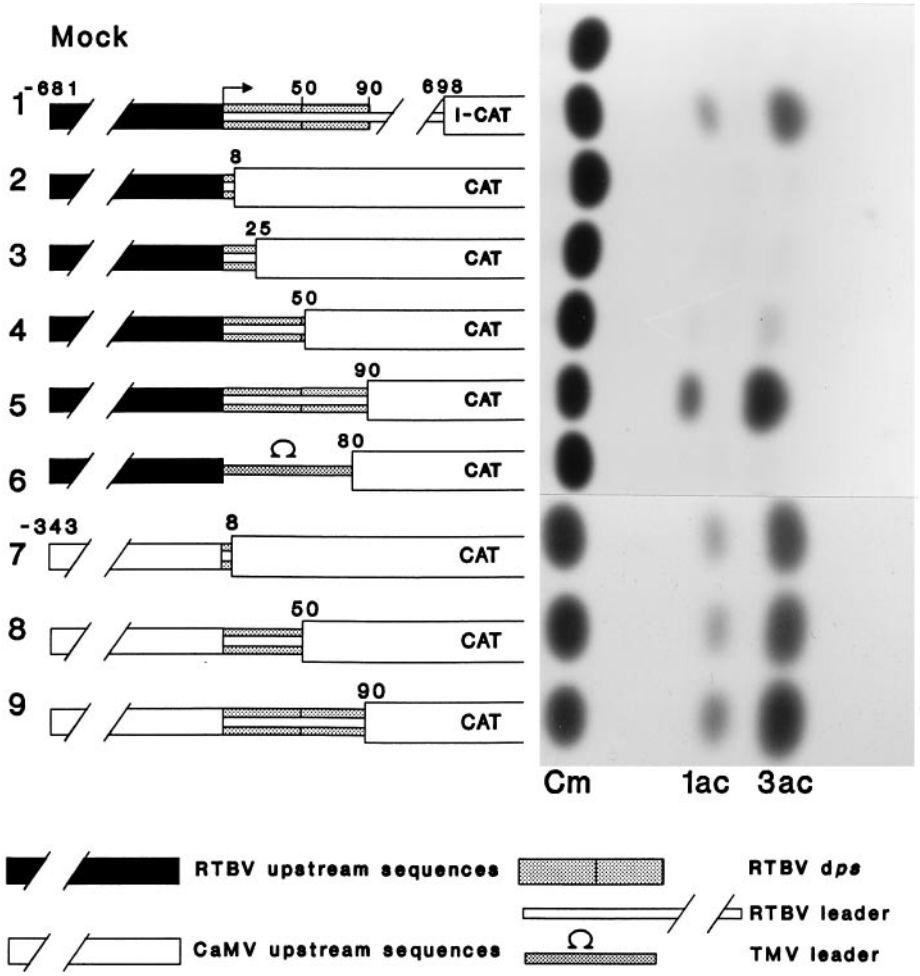
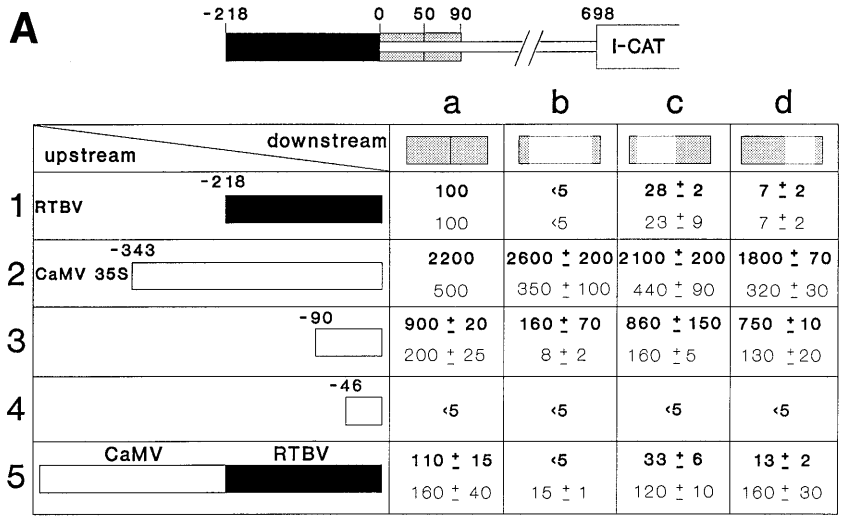


FIG. 2. A 90-nt *dps* fragment restores promoter activity. *O. sativa* protoplasts were transfected with the plasmids shown (R · I-CAT [row 1], R · 8-CAT [row 2], R · 25-CAT [row 3], R · 50-CAT [row 4], R · 90-CAT [row 5], R · Ω-CAT [row 6], C · 8-CAT [row 7], C · 50-CAT [row 8], and C · 90-CAT [row 9]). The number of nucleotides from the extreme 5′ end of the RTBV leader sequences which have been reintroduced in each construct is indicated. Ω represents the tobacco mosaic virus (TMV) leader. CAT assays performed with protoplast extracts are shown on the right, with a mock-transfected extract at the top. The positions of the substrate ([^{14}C]chloramphenicol [Cm]) and the acetylated products (1ac and 3ac) are indicated. The amount of extract used for the CaMV 35S promoter series is 30 times less than that for the RTBV series.

The *dps* region may consist of more than one element. To locate the active element(s) within the *dps* region more precisely, the effect of various *dps* deletions was tested in the context of the whole leader and either RTBV or CaMV 35S upstream promoter sequences. RTBV promoter sequences to position −218 were used in these constructs. We have shown previously that the region between positions −218 and −681 does not influence expression from the RTBV promoter in either *O. sativa* or *O. violaceus* protoplasts (11). In CaMV 35S promoter constructs (Fig. 3A, row 2), none of the leader sequence deletions strongly affected expression, again showing that *dps* has little effect on expression from this promoter at either the transcriptional or posttranscriptional level. In contrast, in the context of the RTBV promoter (Fig. 3A, row 1), deletion of the region from positions +50 to +83 almost abolished CAT activity, suggesting that the major determinants of *dps* function are contained within this region. Deletion of nt +8 to +49 retained around 25% of the original activity, indicating that the region upstream of nt +50 still contains sequence elements required for complete *dps* function.

Although the full-length CaMV 35S upstream promoter sequence is not enhanced by the RTBV *dps* region (Fig. 3A, boxes 2a and 2b), it remained possible that the strong enhancer elements of the 35S promoter might mask an effect of *dps* on the basic or core elements of the 35S promoter. Therefore, the effects of the full and partial *dps* deletions were also tested in the context of short 35S promoter derivatives lacking upstream enhancers. A promoter derivative truncated to position −90 retains around 40% of the activity of the full 35S promoter (Fig. 3A, cf. boxes 2a and 3a). This relatively high residual activity greatly depends on the presence of the entire *dps* sequence, its removal leading to a sixfold reduction in *O. sativa* protoplasts and to an even greater reduction in *O. violaceus* (Fig. 3A, boxes 3a and b; bold and normal typeface, respectively). However, in both cell types, either of the partial *dps* deletions was almost as effective as the whole *dps* sequence in enhancing the −90 CaMV promoter (Fig. 3A, boxes 3c and 3d). This indicated that elements within *dps* can act together with elements in the truncated 35S promoter. A CaMV promoter truncated to position −46 (see reference 53) was inactive and could not be rescued by *dps* (Fig. 3A, row 4), showing that while *dps* elements can augment or reinforce enhancer elements in the −90 promoter, they cannot functionally re-



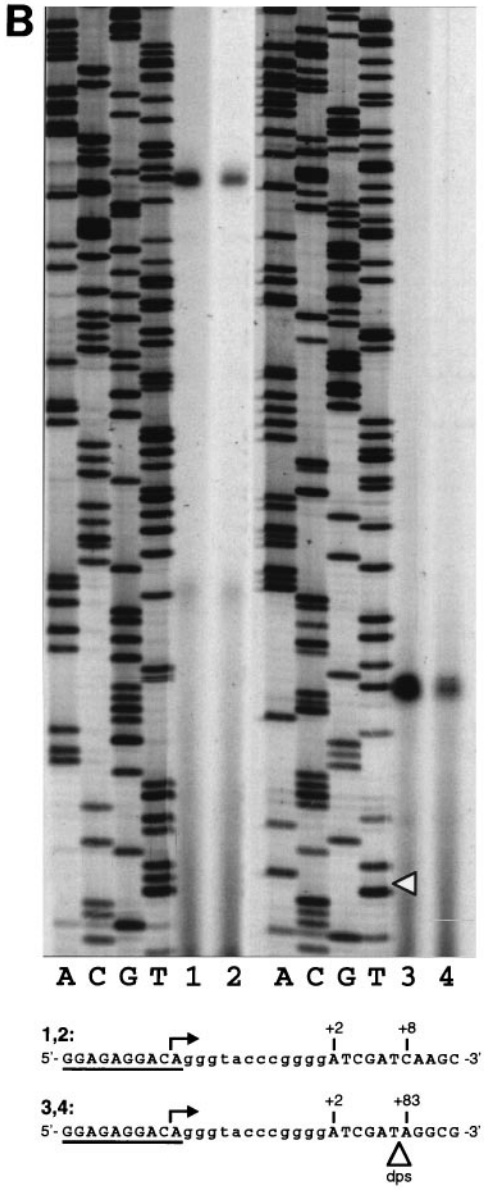


FIG. 3. Effect of RTBV *dps* in the context of different upstream promoter sequences. (A) The basic construct design is shown at the top. Combinations of different upstream promoter sequences (rows) and *dps* fragments (columns) are shown. Constructs are symbolized as in Fig. 2. Results from *O. sativa* and *O. violaceus* protoplasts are shown in bold and normal typeface, respectively. CAT activities (the averages from at least three experiments ± standard deviation) are expressed relative to that of the standard RTBV promoter construct (set to 100 for each protoplast type). The plasmids used (with row and column designations in parentheses) were R · I-CAT (1a), RΔCl83 · I-CAT (1b), RΔCl50 · I-CAT (1c), RΔ50-83 · I-CAT (1d), C · I-CAT (2a), CΔCl83 · I-CAT (2b), CΔCl50 · I-CAT (2c), CΔ50-83 · I-CAT (2d), Cc · I-CAT (3a), CcΔCl83 · I-CAT (3b), CcΔCl50 · I-CAT (3c), CcΔ50-83 · I-CAT (3d), Cm · I-CAT (4a), CmΔCl83 · I-CAT (4b), CmΔCl50 · I-CAT (4c), CmΔ50-83 · I-CAT (4d), CR · I-CAT (5a), CRΔCl83 · I-CAT (5b), CRΔCl50 · I-CAT (5c), and CRΔ50-83 · I-CAT (5d). (B) Primer extension analysis of total RNA isolated from *O. sativa* protoplasts transfected with C · I-CAT, Cc · I-CAT, CΔCl83 · I-CAT, or CcΔCl83 · I-CAT (lanes 1 to 4, respectively). Primer extension products are shown next to sequencing ladders produced by using the same primer (see Materials and Methods) with plasmids C · I-CAT and CΔCl83 · I-CAT (left and right sequencing ladders, respectively). The sequence around the transcription start site (bent arrow) is indicated below, showing the relevant parts of the CaMV 35S promoter sequence (underlined), polylinker (lower case), and RTBV leader sequence. Triangles indicate the position of the *dps* deletion.

place them, at least not from a position downstream of the transcription start site.

The fact that *dps* has a differential effect on the full-length 35S promoter and the −90 derivative is consistent with our hypothesis that the effect is transcriptional, since these two promoters should direct the production of identical RNA transcripts, thus ruling out possible translational and/or RNA stability effects. To verify that neither the 35S promoter truncation nor the *dps* deletion altered the transcription start site, primer extension analysis was performed on RNA isolated from rice protoplasts transfected with 35S or −90 promoter-driven constructs, with or without the *dps* deletion in the RTBV leader. The results confirmed that transcription starts at the same position in all four constructs (Fig. 3B). Primer extension analysis of RNA from RTBV promoter-driven constructs confirmed the published transcription start site (genomic position 7405) and verified that deletion of *dps* sequences does not alter this start site (data not shown).

To test whether the *dps* region can be functionally replaced by a classical DNA-based enhancer element, the CaMV 35S enhancer region (positions −343 to −90) (16, 54) was combined with the RTBV promoter (positions −218 to +8; Fig. 3A, row 5). In the absence of *dps*, no activity was observed; thus, the CaMV 35S enhancer alone is unable to activate

RTBV upstream promoter sequences. However, in *O. violaceus* protoplasts, the CaMV enhancer could functionally replace either of the *dps* subregions, as long as the other was still present. In *O. sativa* protoplasts, there was only a very slight effect of the CaMV enhancer on the partial *dps* deletion constructs.

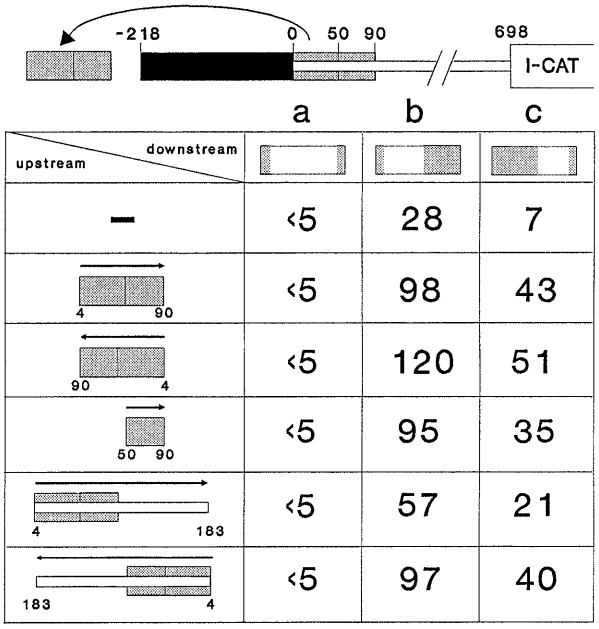
These results suggest that the *dps* region may be composed of at least two (possibly overlapping) elements which act synergistically to confer full enhancement of RTBV promoter activity but which can each interact independently with elements present in a heterologous promoter, even if the latter are in an upstream position.

The *dps* region contains both position-dependent and -independent elements. Because of its location within the transcribed region, the *dps* region could contain either DNA- or RNA-based elements or both. DNA-based enhancer elements are often position and orientation independent. To determine if this holds for the RTBV *dps* region, sequences from positions +8 to +83 were deleted from our standard construct (see top of Fig. 4) and a slightly extended version of this region (positions +4 to +90) was inserted in either forward or reverse orientation at the −218 position (Fig. 4, boxes 2a and 3a). Regardless of the orientation, the translocated *dps* was unable to enhance RTBV promoter activity in *O. sativa* protoplasts (Fig. 4, column a). To be sure that we had not inadvertently interrupted a functional element, we also tested a longer fragment (positions +4 to +183) inserted at this position. Again these constructs were inactive (Fig. 4, boxes 5a and 6a).

Although these results initially suggested that *dps* is position dependent, by moving the entire *dps* sequence to an upstream position, we may have inadvertently destroyed a position-dependent component which would thus mask the effect of moving a position-independent element. Therefore, we tested a series of constructs in which *dps* sequences (positions +4 to +90) were placed at position −218 upstream of the RTBV promoter, while leaving partial *dps* sequences in their original downstream position. All of these plasmids resulted in an increase of CAT activity compared with the corresponding constructs without the upstream insertion (Fig. 4, columns b and c). CAT activity was restored completely if the region from positions +50 to +90 was still present (Fig. 4, column b) and partially in the presence of nt +8 to +49 (Fig. 4, column c). Essentially the same results were obtained when a longer fragment (positions +4 to +183) was inserted upstream (Fig. 4, boxes 5b, 5c, 6b, and 6c). The orientation of the upstream fragment caused only slight differences in expression levels. These results show that, provided partial *dps* sequences remain in a downstream position, the *dps* sequence contains at least one component which is active in a position- and orientation-independent manner. Since inserting *dps* sequences from positions +50 to +90 in an upstream position (Fig. 4, row 4) has essentially the same effect as that of the full-length *dps* fragment, it seems likely that the position-independent element is contained within this region (see below).

The result is not specific for *O. sativa*, since similar results were also observed in *O. violaceus* protoplasts (data not shown). Additional evidence supporting the presence of a position-independent enhancer element within *dps* came from combining *dps* with the −90 version of the CaMV 35S promoter. In this promoter context, deletion of *dps* from its downstream position results in a dramatic drop in activity (Fig. 3A, boxes 3a and 3b), but replacing *dps* sequences upstream of the −90 35S promoter restores around two-thirds of the original activity (not shown).

Nuclear protein binds to *dps* DNA. To search for *dps*-binding proteins, gel shift experiments were performed. Formation of a DNA-protein complex was observed when a labelled *dps* fragment from positions +4 to +90 was incubated with nuclear protein extracts prepared from protoplasts obtained from *O. sativa* cell suspension cultures. The protein-DNA interaction was sequence specific since it could be competed only with unlabelled *dps* but not with other DNA fragments of similar length (Fig. 5A and B). We tentatively refer to the binding protein(s) as *dps* DNA-binding factor (DDBF). The same binding activities were seen in nuclear extracts prepared from *O. sativa* seedlings (not shown).



		a	b	c
1	—	<5	28	7
2	4 90	<5	98	43
3	90 4	<5	120	51
4	50 90	<5	95	35
5	4 183	<5	57	21
6	183 4	<5	97	40

FIG. 4. Translocation of RTBV *dps* to an upstream position identifies a position- and orientation-independent component within *dps*. Rows indicate RTBV *dps* fragments inserted at position −218 of the RTBV promoter, and columns indicate the *dps* fragments remaining in the leader. CAT activities relative to that of the complete standard construct (=100; see Fig. 3A, box 1a) in transfected *O. sativa* protoplasts are shown.

To define more precisely the sequence involved in interaction with DDBF, *dps* was dissected into three parts (I, II, and III; Fig. 5C). Only fragment III, which corresponds to positions +50 to +90, was able to compete with *dps* binding (Fig. 5D), suggesting that DDBF specifically recognizes sequences within this region.

Identification of a sequence motif involved in enhancer function. Deletion of fragment III (positions +50 to +90) from its original position resulted in a dramatic drop in CAT activity (7% of wild-type activity; Fig. 3A, row 1d). To test enhancer function of fragment III, it was inserted in the sense orientation at position −218 of the RTBV promoter. This insertion resulted in almost the same enhancer effect as that of the complete *dps* sequence (Fig. 4, cf. boxes 2c and 4c), strongly suggesting that the DNA-based enhancer element is indeed contained within the fragment which interacts with DDBF.

Fragment III contains a CT-rich region with two YCCTCTY repeats. Similar motifs are found in the promoter of the maize phosphoenolpyruvate carboxylase gene, where they have been shown to have an enhancer function and to bind nuclear proteins (38, 49, 62, 65). To determine whether this motif and/or other regions within *dps* are important for enhancer activity in RTBV, we constructed a series of mutations (Fig. 6). In m1, m2, and m3, some of the pyrimidines of the CT-rich motif were

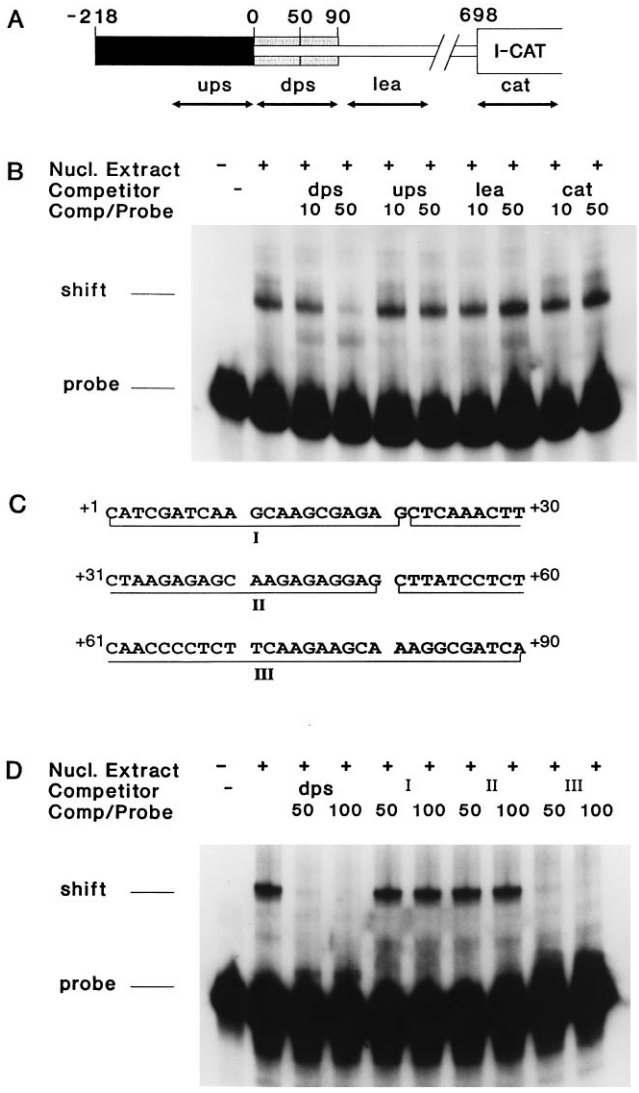


FIG. 5. Nuclear proteins specifically bind *dps* DNA sequences. (A) Map of our standard plasmid showing the positions of the probe (dps) and competitor DNA fragments. (B) A labelled *dps* DNA fragment was preincubated in the presence of carrier DNA either alone (−) or with a nuclear extract (Nucl. Extract) prepared from *O. sativa* cell suspensions (+). Competition was with 10 and 50 molar excess (Comp/Probe) of either unlabelled *dps* or fragments of comparable size from the upstream promoter sequences (ups), the leader (lea), or the CAT open reading frame (cat). The positions of free probe and the shifted complex in the gel are indicated. (C) Nucleotide sequence of the *dps* region from positions +1 to +90, indicating subfragments I, II, and III. (D) Competition experiments using *dps* as the labelled probe and subfragments I, II, and III as competitors. The molar excess of competitor is indicated (Comp/Probe).

mutated to A residues, and in m4, three A residues further downstream were each mutated to T. The abilities of these mutant fragments to compete with *dps* for DDBF binding in gel shift experiments and their effects on expression when inserted at position −218 of the RTBV promoter were assayed. The m4 mutation had no effect on the enhancer activity (Fig. 6D), and the mutant fragment could compete as well as the wild-type sequence in gel shifts (not shown). Mutations m1, m2, and m3 all had reduced enhancer and competition activities, with m2 showing the most dramatic effect (Fig. 6C and D).

In summary, results obtained with enhancer activity in transient-expression experiments and with DDBF binding in vitro are consistent with the interpretation that the CT-rich sequences within *dps* are part of an enhancer and are involved in DDBF binding. This in turn suggests that DDBF might be a transcription factor positively regulating transcription from the RTBV promoter.

Nuclear protein binds the *dps* transcript. The protein-binding, position-independent, DNA element within *dps* is not sufficient to enhance RTBV promoter activity fully (Fig. 3A and 4), indicating that an additional element is included within *dps*, which acts only in the downstream position. This might consist of another DNA-based element with a structural function or an RNA-based element. To search for potential RNA-protein interactions within *dps*, RNA gel shift experiments were performed. A ^{32}P-labelled transcript from the *dps* region was incubated with rice nuclear extracts. Following electrophoresis in native polyacrylamide gels, a smeared shift was observed, within which two weak bands (USB and LSB) shifted to a position above the free probe could be discerned (Fig. 7). No band shift was observed if the nuclear extract had been treated with proteinase K or replaced by a similar amount of bovine serum albumin, and the shift was clearly reduced if the nuclear extract was heated for 2 min to 85°C.

The *dps* transcript exists in two conformations of different mobility (PLB and PUB) in native gels (Fig. 7). After denaturation by heating to 85°C for 5 min before loading on the gel, only the lower band was seen and this did not interact with the nuclear extract (Fig. 7A). Consistent with this observation, only the upper band was capable of forming the complex after the two conformations were separated prior to the gel shift experiment (data not shown). This suggests that the structure of the *dps* transcript is important for protein binding.

Competition assays were performed to test the specificity of the RNA-protein interaction. Complex formation was reduced when an unlabelled *dps* transcript was used as competitor, whereas no effect was observed with RNA competitors derived from other regions of the construct (Fig. 7B), indicating a specific interaction between the *dps* transcript and protein(s) present in the nuclear extract.

DISCUSSION

A region downstream of the transcription start site of the RTBV pregenomic RNA was previously identified as an important determinant affecting efficiency of expression from the RTBV promoter (11). In this report, we show that sequence made up of the first 90 nt of the transcribed region (downstream promoter sequence [*dps*]) is sufficient to enhance RTBV upstream promoter activity. Three lines of evidence support the conclusion that this effect is transcriptional and that *dps* contains promoter elements. First, deletion of *dps* dramatically lowers steady-state RNA levels in the context of RTBV upstream promoter sequences, but not if the full-length CaMV 35S promoter is used. Second, *dps* was also found to enhance the truncated −90 version of the CaMV 35S promoter, so an effect on translation or RNA stability can be ruled out because production of an identical RNA molecule under the control of the full-length CaMV 35S promoter was not affected. Third, a DNA element within *dps* functions as an enhancer if moved to a position upstream of the transcription start site of either the RTBV promoter or the CaMV −90 promoter.

Our results indicate that *dps* comprises at least two components, one of which is a position- and orientation-independent DNA element. The second component is position dependent, and there are preliminary indications that it might be RNA based.

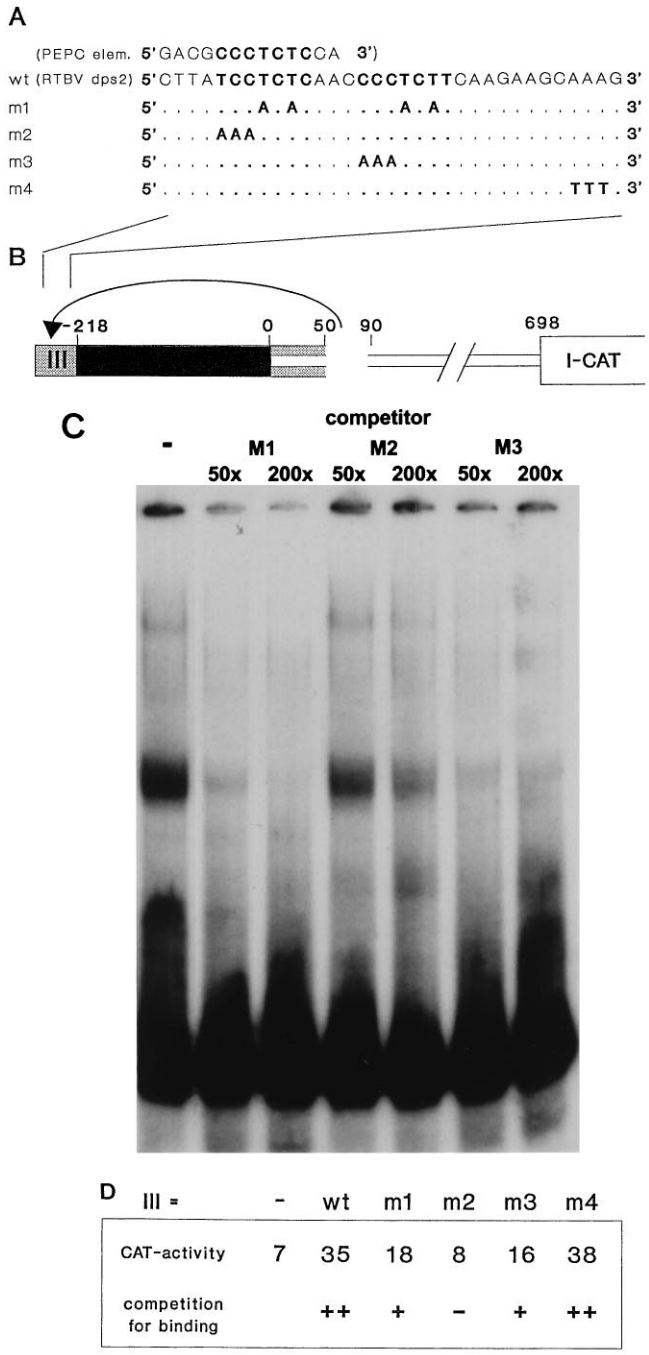


FIG. 6. Mutations in *dps* fragment III reduce both enhancer activity and DDBF binding. (A) The CT-rich motif in the phosphoenolpyruvate carboxylase gene promoter (PEPC elem.) and in RTBV *dps* fragment III (dps2) are compared in the first two rows. Mutant sequences are shown underneath (dots represent wild-type sequence). (B) In the construct used to test the mutations, fragment III was deleted from *dps* in its downstream position and this sequence or mutant derivatives were reinserted at position −218 of the RTBV promoter. (C) DNA gel shift using a labelled *dps* DNA probe, and unlabelled m1, m2, and m3 fragments, at the molar excess indicated, as competitors. −, no competitor (control). (D) Mutants m1 to m4 were tested in the upstream position for enhancer function and for their ability to compete with the complete *dps* region for DDBF binding in gel shift assays. CAT activities relative to that of the complete standard construct (=100, see Fig. 3A, box 1a) and relative strength of competition for binding compared with the wild-type (wt) sequence are shown. −, no competitor (control). Symbols for competition for binding: ++, strong competition at 50× molar excess; +, weak competition at 50× molar excess and strong competition at 200×; −, no or very weak competition at 50× molar excess and weak competition at 200×).

Binding of a DDBF to the DNA element depends on the presence of a region (positions +50 to +70) containing a repeated YCCTCTY motif. Mutations in this sequence reduce both the enhancer activity of *dps* and binding of DDBF, which is consistent with the hypothesis that DDBF represents a transcription factor. The same motif in the promoter of the maize phosphoenolpyruvate carboxylase gene has been shown to act as an upstream enhancer element (49, 62) and to bind the nuclear protein PEP-1 from maize (38, 65) and rice (48), but it remains to be seen if there is any relationship between DDBF and PEP-1.

Downstream regions believed to be important for efficient expression have been found in a number of plant promoters (7, 9, 10, 12, 13, 45), but in most cases, the level at which expression is affected has not been determined. Only in the case of the *Arabidopsis thaliana* ferredoxin gene has the effect been localized to a specific sequence region (in this case a CA-rich motif). In a number of light-regulated genes, quantitative determinants of transcription are contained within a $(CT)_5$ element downstream of the transcription start site (5). A CT-rich region at the beginning of the CaMV 35S transcribed region has a positive effect on expression (19). Part of this enhance-

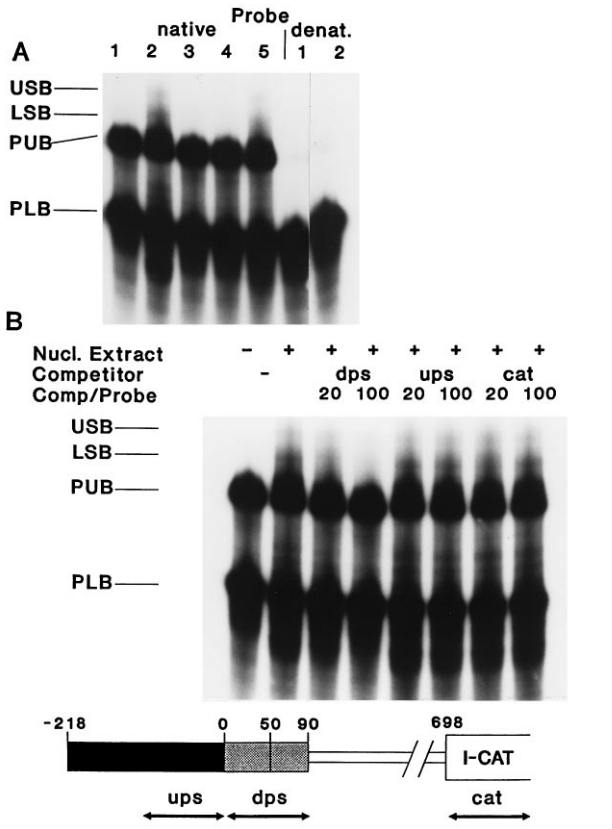


FIG. 7. Nuclear proteins bind to the *dps* transcript. (A) RNA gel shift experiment. PLB and PUB represent the faster- and slower-migrating forms of the probe, LSB and USB, the lower and upper shifted bands. Native or heat-denatured (denat.) *dps* RNA probes were used. Lanes: 1, RNA probe alone; 2 to 5, RNA probe plus nuclear extract (lanes 2), bovine serum albumin (lane 3), proteinase K-treated nuclear extract (lane 4), or heat-treated nuclear extract (lane 5). (B) Competition experiment. Native *dps* RNA probes were incubated with (+) or without (−) nuclear extract (Nucl. Extract) and with competitor transcripts as indicated over the gel. RNA fragments used as competitors were comparable in size to the *dps* transcript; the positions on the construct from which they were derived are shown on the map. Competition was with 20 and 100 molar excess of the indicated fragment as shown over the gel.

ment is probably due to translational effects (14), but our ongoing work has revealed that this region also affects transcription (unpublished data). CT- or GA-rich regions can be found in the vicinity of the transcription start sites of many plant genes, particularly in monocots, and in some cases can be correlated with increased gene expression. Enhancement of expression of the rice actin 1-, the maize ubiquitin 1-, and the maize shrunken 1-promoters (8, 46, 47) by such downstream elements was interpreted as the effect of introns contained within these regions because no enhancement was seen if these sequences were moved to outside the transcribed region. However, the RTBV promoter also becomes inactivated if the entire *dps* region is moved to an upstream position. Thus, functional dissection of the downstream region of the aforementioned promoters might reveal that they too contain transcriptional control elements downstream of the start site.

The CT-rich DNA enhancer element does not account for the full effect of the RTBV *dps* region. When placed upstream of the promoter at position −218, this DNA element was able to exert its effect only if at least part of the *dps* sequence was retained downstream of the transcription start site. While this might be explained by duplication of the DNA enhancer element in some constructs, with a copy both upstream and downstream of the transcription start site (Fig. 4, column b), in other constructs the additional enhancing effect was mediated by sequences from positions +1 to +50, which have no effect in an upstream position (Fig. 4, cf. rows 2 and 4). These results indicate that *dps* also contains a position-dependent component that can act synergistically with the position-independent component. The position-dependent element could be a DNA element functioning either as an architectural component or contributing to basal promoter activity, as in the case of a number of different downstream elements in animal promoters (23, 27, 28, 40, 52, 55). Although we could not detect any protein binding to the *dps* region besides the one around positions +50 to +70, we cannot exclude the possibility that an additional binding protein requires further upstream sequences. A protein recently reported to bind around positions −3 to +8 of the RTBV promoter (66), which was not detected in our gel shift experiments, might require additional upstream sequences (e.g., the TATA box) for binding. It is also possible that other proteins interacting with the DNA of this region are not detected by the gel shift assay for technical reasons. In fact, DDBF was found to be unstable in our extracts, even at −80°C, thus hampering further characterization, e.g., by DNA footprinting analysis.

Alternatively, or in addition to DNA sequences, RNA-based elements may influence expression efficiency. The *dps* region of RTBV RNA can be folded into a stable stem-loop structure (Fig. 8), and proteins in the nuclear extract interact with *dps* RNA in an apparently structure-dependent manner (Fig. 7). The results in Fig. 4 show that the first half of *dps* can act synergistically with the DNA enhancer, which would seem to speak against an involvement of the whole *dps* structure. However, as shown in Fig. 8, the sequence made up of the first 50 nt of *dps* has a predicted secondary structure that maintains several structural features of the entire *dps* region. These observations raise the possibility that RNA-binding proteins have some function in regulating RTBV promoter activity, and this is currently under further investigation.

The complexity of sequences that regulate expression and are located downstream of the transcription start site is exemplified by the human immunodeficiency virus type 1 (HIV-1). The leader sequences of HIV-1 contain overlapping elements (either DNA or RNA) influencing initiation of transcription, RNA elongation, and RNA utilization (35, 36, 39, 63). The

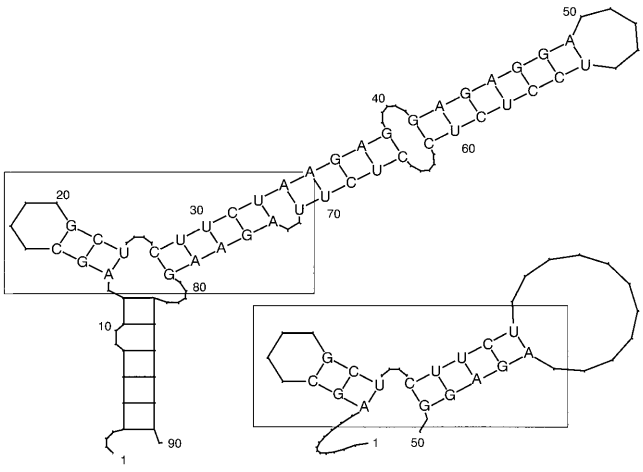


FIG. 8. Predicted secondary structure of the *dps* transcript. Potential secondary structure of RTBV *dps* sequences were calculated by using the MFOLD program with default parameters (Genetics Computer Group package, version 8) and visualized as SQUIGGLES output ($\Delta G = -21.2$ kcal/mol). The sequence is shown for only part of the structure; for the full sequence of *dps*, see Fig. 5C. The boxed area highlights the structure which is present in the predicted folding of both *dps* regions (positions 1 to 90 and 1 to 50).

best known of these is the well-studied RNA-based TAR element (reviewed in reference 37). It has proved difficult to dissect the individual effects of these elements. In the case of the RTBV promoter, the low level of transcription directed by even the complete RTBV promoter will make direct analysis of dynamic events during RTBV RNA synthesis difficult. However, constructs similar to those shown in Fig. 6, in which the DNA-based enhancer is placed upstream of the transcribed regions, will be useful in dissecting the position-dependent and -independent elements of the RTBV *dps* region.

The complexity of interactions within a promoter is illustrated by the results obtained with combinations of *dps* and different versions of the CaMV 35S promoter. The *dps* region was able to function with the CaMV 35S promoter truncated to −90, but not when the promoter was further shortened to position −46 (Fig. 3A). This indicates that *dps* requires other promoter sequences to exert its effect. The region between positions −90 and −46 in the CaMV 35S promoter contains the *as*-1 element, and this latter element may mediate the effect of *dps* as it mediates the effect of CaMV upstream enhancer sequences (3). It appears likely that the RTBV promoter also contains an upstream element that acts in concert with *dps*. This element must be located close to the transcription start site (up to position −55), since this short region supported substantial promoter activity together with *dps* (11). The differential effects of the *dps* deletions with the RTBV and the truncated CaMV promoters suggest that the interacting elements and their associated factors are promoter specific. Cell-specific differences were also observed. While in *O. violaceus* protoplasts, the CaMV enhancer region (positions −343 to −90) can replace either of the two *dps* subregions (Fig. 3A), no such effect was seen in *O. sativa* protoplasts, although the CaMV enhancer is active in these cells when combined with the CaMV 35S −90 promoter. Apparently, the final enhancing effect in all these combinations is accomplished by interactions of different *cis*- or *trans*-acting factors in the different promoter constructs and in the different cell types. However, the total overall effect of *dps* is similar in *O. sativa* and *O. violaceus* protoplasts.

The characterization of *dps* presented here was carried out in protoplast transient-expression systems. Studies on the RTBV promoter in transgenic rice plants has shown that *dps* sequences are apparently nonessential in that context; promoter fragments lacking most of the *dps* region can direct phloem-specific activity (4, 66). While these findings appear contradictory, they may in fact reflect intrinsic differences between stable- and transient-expression systems. In the case of pararetroviruses such as RTBV, neither system properly reflects the viral context. Protoplasts derived from cell cultures may not provide a full complement of the factors that restrict promoter activity to vascular tissue in planta, while in stable transformants, integration into the host chromosome may alter the activity and accessibility of promoter components by involving them in chromatin structures (see, for example, reference 15) that would not be present in the viral context; pararetroviral transcription occurs from nonintegrated copies of the viral DNA genome. Nevertheless, both approaches provide valuable insights into the workings of this complex viral promoter, furthering our understanding of viral regulation per se, and allowing the rational design of expression vectors exploiting viral components.

As mentioned above, we have evidence that elements downstream of the transcription start site are also involved in transcriptional regulation in CaMV. It will be interesting to determine if such elements are a general feature of the plant pararetroviruses.

ACKNOWLEDGMENTS

We are grateful for the gifts of an RTBV clone from R. Hull and his group and the rice Oc suspension culture from K. Shimamoto. We thank U. Klahre and W. Schmidt-Puchta for providing unpublished protocols concerning nuclear extracts and protein-nucleic acid binding conditions and R. Beachy for providing a manuscript prior to publication. We gratefully acknowledge the expert technical assistance of Matthias Müller and Sandra Corsten and the advice of all of our colleagues, particularly Diana Dominguez for help in setting up the gel shift experiment. The critical reading of the manuscript by Yvan Chapdelaine and Patrick Matthias is very much appreciated.

REFERENCES

1. **Bao, Y., and R. Hull.** 1993. Mapping the 5′-terminus of rice tungro bacilliform viral genomic RNA. Virology **197:**445–448.
2. **Bao, Y., and R. Hull.** 1994. Replication intermediates of rice tungro bacilliform virus DNA support a replication mechanism involving reverse transcription. Virology **204:**626–633.
3. **Benfey, P. N., and N. Chua.** 1990. The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. Science **250:**959–966.
4. **Bhattacharyya-Pakrasi, M., J. Peng, J. S. Elmer, G. Laco, P. Shen, M. B. Kaniewska, H. Kononowicz, F. Wen, T. K. Hodges, and R. N. Beachy.** 1993. Specificity of a promoter from the rice tungro bacilliform virus for expression in phloem tissues. Plant J. **4:**71–79.
5. **Bolle, C., S. Sopory, T. Lübberstedt, R. G. Herrmann, and R. Oelmüller.** 1994. Segments encoding 5′-untranslated leaders of genes for thylakoid proteins contain cis-elements essential for transcription. Plant J. **6:**513–523.
6. **Bonneville, J. M., H. Sanfaçon, J. Fütterer, and T. Hohn.** 1989. Posttranscriptional *trans*-activation in cauliflower mosaic virus. Cell **59:**1135–1143.
7. **Bovy, A., C. Van Den Berg, G. De Vrieze, W. F. Thompson, P. Weisbeek, and S. Smeekens.** 1995. Light-regulated expression of the *Arabidopsis thaliana* ferredoxin gene requires sequences upstream and downstream of the transcription initiation site. Plant Mol. Biol. **27:**27–39.
8. **Callis, J., M. Fromm, and V. Walbot.** 1987. Introns enhance gene expression in cultured maize cells. Genes Dev. **1:**1183–1200.
9. **Caspar, T., and P. H. Quail.** 1993. Promoter and leader regions involved in the expression of the *Arabidopsis* ferredoxin A gene. Plant J. **3:**161–174.
10. **Chappell, J., and K. Hahlbrock.** 1984. Transcription of plant defence genes in response to UV light of fungal elicitor. Nature (London) **311:**76–78.
11. **Chen, G., M. Müller, I. Potrykus, T. Hohn, and J. Fütterer.** 1994. Rice tungro bacilliform virus: transcription and translation in protoplasts. Virology **204:**91–100.
12. **Curie, C., M. Axelos, C. Bardet, R. Atanassova, N. Chaubet, and B. Lescure.** 1993. Modular organization and developmental activity of an *Arabidopsis thaliana* EF-1α gene promoter. Mol. Gen. Genet. **238:**428–436.
13. **Douglas, C. J., K. D. Hauffe, M.-E. Ites-Morales, M. Ellard, U. Paszkowski, K. Hahlbrock, and J. L. Dangl.** 1991. Exonic sequences are required for elicitor and light activation of a plant defense gene, but promoter sequences are sufficient for tissue specific expression. EMBO J. **10:**1767–1775.
14. **Dowson Day, M. J., J. L. Ashurst, S. F. Mathias, and J. W. Watts.** 1993. Plant viral leaders influence expression of a reporter gene in tobacco. Plant Mol. Biol. **23:**97–109.
15. **El Kharroubi, A., and M. A. Martin.** 1996. *cis*-Acting sequences located downstream of the human immunodeficiency virus type 1 promoter affect its chromatin structure and transcriptional activity. Mol. Cell. Biol. **16:**2958–2966.
16. **Fang, R., F. Nagy, S. Sivasubramaniam, and N. Chua.** 1989. Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell **1:**141–150.
17. **Fromm, M., L. P. Taylor, and V. Walbot.** 1985. Expression of genes transferred into monocot and dicot plant cells by electroporation. Proc. Natl. Acad. Sci. USA **82:**5824–5828.
18. **Fütterer, J., K. Gordon, P. Pfeiffer, H. Sanfaçon, B. Pisan, J. Bonneville, and T. Hohn.** 1989. Differential inhibition of downstream gene expression by the cauliflower mosaic virus 35S RNA leader. Virus Genes **3:**45–55.
19. **Fütterer, J., K. Gordon, H. Sanfaçon, J. 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.
20. **Fütterer, J., and T. Hohn.** 1992. Role of an upstream open reading frame in the translation of polycistronic mRNAs in plant cells. Nucleic Acids Res. **20:**3851–3857.
21. **Fütterer, J., I. Potrykus, M. P. Valles-Brau, I. Dasgupta, R. Hull, and T. Hohn.** 1994. Splicing in a plant pararetrovirus. Virology **198:**663–670.
22. **Gallie, D. R., and V. Walbot.** 1992. Identification of the motifs within the tobacco mosaic virus 5′-leader responsible for enhancing translation. Nucleic Acids Res. **20:**4631–4638.
23. **Ghazal, P., and J. A. Nelson.** 1991. Enhancement of RNA polymerase II initiation complexes by a novel DNA control domain downstream from the cap site of the cytomegalovirus major immediate-early promoter. J. Virol. **65:**2299–2307.
24. **Goodall, G. J., and W. Filipowicz.** 1989. The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58:**473–483.
25. **Goodall, G. J., K. Wiebauer, and W. Filipowicz.** 1990. Analysis of pre-mRNA processing in transfected plant protoplasts. Methods Enzymol. **181:**148–161.
26. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
27. **Guzowski, J. F., J. Singh, and E. K. Wagner.** 1994. Transcriptional activation of the herpes simplex virus type 1 U_L 38 promoter conferred by the *cis*-acting downstream activation sequence is mediated by a cellular transcription factor. J. Virol. **68:**7774–7789.
28. **Guzowski, J. F., and E. K. Wagner.** 1993. Mutational analysis of the herpes simplex virus type 1 strict late U_L 38 promoter/leader reveals two regions critical in transcriptional regulation. J. Virol. **67:**5098–5108.
29. **Harpster, M. H., J. A. Townsend, J. D. G. Jones, J. Bedbrook, and P. Dunsmuir.** 1988. Relative strength of the 35S CaMV, 1′2′, and nopaline synthase promoters in transformed tobacco, sugarbeet and oilseed rape callus tissue. Mol. Gen. Genet. **212:**182–190.
30. **Hay, J. M., M. C. Jones, M. L. Blakebrough, I. Dasgupta, J. W. Davies, and R. Hull.** 1991. An analysis of the sequence of an infectious clone of rice tungro bacilliform virus, a plant pararetrovirus. Nucleic Acids Res. **19:**2615–2621.
31. **Hemmings-Mieszczak, M.** Personal communication.
32. **Hohn, T., and J. Fütterer.** 1992. Transcriptional and translational control of gene expression in cauliflower mosaic virus. Curr. Opin. Genet. Dev. **2:**90–96.
33. **Hull, R.** 1992. Genome organization of retroviruses and retroelements: evolutionary considerations and implications. Sem. Virol. **3:**373–382.
34. **Jefferson, R. A., T. A. Kavanagh, and M. W. Bevan.** 1987. GUS-fusion:β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:**3901–3907.
35. **Jones, K. A.** 1993. Tat and the HIV-1 promoter. Curr. Opin. Cell Biol. **5:**461–468.
36. **Jones, K. A., P. A. Luciw, and N. Duchange.** 1988. Structural arrangements of transcription control domains within the 5′-untranslated leader regions of the HIV-1 and HIV-2 promoters. Genes Dev. **2:**1101–1114.
37. **Jones, K. A., and M. B. Peterlin.** 1994. Control of RNA initiation and elongation at the HIV-1 promoter. Annu. Rev. Biochem. **63:**717–743.
38. **Kano-Murakami, Y., I. Suzuki, T. Sugiyama, and M. Motsuoka.** 1991. Sequence-specific interaction of a maize factor with a GC rich repeat in the phosphoenolpyruvate carboxylase gene. Mol. Gen. Genet. **225:**203–208.
39. **Karn, J., and M. A. Graebel.** 1992. New insights into the mechanism of HIV-1 trans-activation. Trends Genet. **8:**365–368.
40. **Kashanchi, F., J. F. Duvall, P. F. Lindholm, and M. F. Radonovich.** 1993.

Sequences downstream of the RNA initiation site regulate human T-cell lymphotropic virus I basal gene expression. J. Virol. **67:**2894–2902.
41. **Kiss-László, Z., S. Blanc, and T. Hohn.** 1995. Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. EMBO J. **14:**3552–3562.
42. **Laco, G. S., and R. N. Beachy.** 1994. Rice tungro bacilliform virus encodes reverse transcriptase, DNA polymerase, and ribonuclease H activities. Proc. Natl. Acad. Sci. USA **91:**2654–2658.
43. **Lam, E.** 1994. Analysis of tissue-specific elements in the cauliflower mosaic virus 35S promoter, p. 181–196. *In* L. Nover (ed.), Results and problems in cell differentiation: plant promoters and transcription factors. Springer-Verlag, Berlin.
44. **Lockhart, B. E. L.** 1990. Evidence for a double-stranded circular DNA genome in a second group of plant viruses. Phytopathology **80:**127–131.
45. **Lois, R., A. Dietrich, and K. Hahlbrock.** 1989. A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive *cis*-acting elements. EMBO J. **8:**1641–1648.
46. **Luehrsen, K. R., and V. Walbot.** 1991. Intron enhancement of gene expression and the splicing efficiency of introns in maize cells. Mol. Gen. Genet. **225:**81–93.
47. **Maas, C., J. Laufs, S. Grant, C. Korfhage, and W. Werr.** 1991. The combination of a novel stimulatory element in the first exon of the maize shrunken-1 gene with the following intron 1 enhances reporter gene expression up to 1000-fold. Plant Mol. Biol. **16:**199–207.
48. **Martino-Catt, S. J., and S. A. Key.** 1994. Optimization of DNase 1 footprinting experiments, p. I2S. *In* B. Gelvin and R. A. Schilperoort (ed.), Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht, The Netherlands.
49. **Matsuoka, M., J. Kyozuka, K. Shimamoto, and Y. Kano-Murakami.** 1994. The promoter of two carboxylases in a C_4 plant (maize) direct cell-specific, light-regulated expression in a C_3 plant (rice). Plant J. **6:**311–319.
50. **Medberry, S. L., B. E. L. Lockhart, and N. E. Olszewski.** 1990. Properties of Commelina yellow mottle virus's complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. Nucleic Acids Res. **18:**5505–5513.
51. **Mikami, K., H. Takase, and M. Iwabuchi.** 1994. Gel mobility shift assay, p. I1S. *In* B. Gelvin and R. A. Schilperoort (ed.), Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht, The Netherlands.
52. **Nakatani, Y., M. Horikoshi, M. Brenner, T. Yamamoto, F. Besnard, R. G. Roeder, and E. Freese.** 1990. A downstream initiation element required for efficient TATA box binding and *in vitro* function of TFIID. Nature (London) **348:**86–88.
53. **Odell, J. T., F. Nagy, and N. Chua.** 1985. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature (London) **313:**810–812.
54. **Ow, D. W., J. D. Jacobs, and S. H. Howell.** 1987. Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by use of the firefly luciferase gene as reporter of promoter activity. Proc. Natl. Acad. Sci. USA **84:**4870–4874.
55. **Purnell, A. B., and D. D. Gilmour.** 1993. Contribution of sequences downstream of the TATA element to the protein-DNA complex containing the TATA-binding protein. Mol. Cell. Biol. **13:**2593–2603.
56. **Qu, R., M. Bhattacharyya, G. S. Laco, A. de Kochko, B. L. S. Rao, M. B. Kaniewska, J. S. Elmer, D. E. Rochester, C. E. Smith, and R. N. Beachy.** 1991. Characterization of the genome of rice tungro bacilliform virus: comparison with Commelina yellow mottle virus and caulimoviruses. Virology **185:**354–364.
57. **Rothnie, H. M., Y. Chapdelaine, and T. Hohn.** 1994. Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv. Virus Res. **44:**1–67.
58. **Rothnie, H. M., J. Reid, and T. Hohn.** 1994. The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3′-end formation in plants. EMBO J. **13:**2200–2210.
59. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
60. **Sanfaçon, H.** 1994. Analysis of figwort mosaic virus (plant pararetrovirus) polyadenylation signal. Virology **198:**39–49.
61. **Sanfaçon, H., and T. Hohn.** 1990. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. Nature (London) **346:**81–84.
62. **Schäffner, A. R., and J. Sheen.** 1992. Maize C_4 photosynthesis involves differential regulation of phosphoenolpyruvate carboxylase genes. Plant J. **2:**221–231.

62a. **Schmidt-Puchta, W.** Personal communication.

63. **Sheldon, M., R. Ratnasabapathy, and N. Hernandez.** 1993. Characterization of the inducer of short transcripts, a human immunodeficiency virus type 1 transcriptional element that activates the synthesis of short RNAs. Mol. Cell. Biol. **13:**1251–1263.
64. **Shillito, R. D., M. W. Saul, J. Paszkowski, M. Müller, and I. Potrykus.** 1985. High efficiency direct gene transfer to plants. Biotechnology **3:**1099–1103.
65. **Yanagisawa, S., and K. Lzui.** 1990. Multiple interactions between tissue-specific nuclear proteins and the promoter of the phosphoenolpyruvate carboxylase gene for C_4 photosynthesis in *Zea mays*. Mol. Gen. Genet. **224:** 325–332.
66. **Yin, Y., and R. N. Beachy.** 1995. The regulatory regions of the rice tungro bacilliform virus promoter and interacting nuclear factors in rice (*Oryza sativa* L.). Plant J. **7:**969–980.