The EMBO Journal vol.7 no.6 pp.1589 - 1596, 1988

The intergenic region of maize streak virus contains promoter elements involved in rightward transcription of the viral genome

Carmen Fenoll, Diane M.Black[1] and Stephen H.Howell

Biology Department C016, University of California San Diego, La Jolla, CA 92093, USA
[1]Present address: Hoefer Scientific Instruments, 654 Minnesota Street, San Francisco, CA 94107, USA

Communicated by F.Garcia-Olmedo

Maize streak virus (MSV), a geminivirus with a one-component genome, encodes a major coat protein RNA which accumulates in infected plants. Using a maize protoplast cell transient expression system, we have defined and studied the promoter which drives rightward transcription of the RNA encoding the coat protein. We have identified a 122 bp upstream segment that enhances promoter activity and functions as an upstream activating sequence (UAS). The UAS lies in the starting intergenic region of the viral genome and includes a region which is similar in all geminiviruses. The 122 bp UAS activates the MSV core promoter in an orientation, but not position, independent fashion. The MSV promoter UAS is interchangeable with a similar element in the cauliflower mosaic virus (CaMV) 35S RNA core promoter, that is the MSV UAS will activate the CaMV 35S core promoter and vice versa. However, the MSV promoter UAS specifically binds proteins in maize nuclear extracts which appear to differ from those bound by the functionally equivalent region of the CaMV 35S promoter.

Key words: plant virus/geminivirus/promoter upstream activating sequences/transient expression system/DNA-protein complexes

Introduction

Eucaryotic cellular and viral RNA polymerase II promoters are modular units composed of *cis*-acting elements which bind *trans*-acting factors that regulate and determine the precise location of initiation events of transcription (for reviews, see Picard, 1985; Dynan and Tjian, 1985; McKnight and Tjian, 1986; Kadonaga *et al.*, 1986; Struhl, 1987). Plant cell and virus promoters which have been examined in detail show a modular organization typical of other eucaryotic promoters (Odell *et al.*, 1985; Timko *et al.*, 1985; Morelli *et al.*, 1985; Fluhr *et al.*, 1986; Simpson *et al.*, 1986; Gurley *et al.*, 1986; Kaulen *et al.*, 1986; Chen *et al.*, 1986; Ow *et al.*, 1987; Nagy *et al.*, 1987; Kuhlemeier *et al.*, 1987a; Green *et al.*, 1987; Ebert *et al.*, 1987; Jofuku *et al.*, 1987; Ellis *et al.*, 1987a,b). Some of the more interesting plant promoters include those that are regulated in different tissues and respond to environmental cues and others, which include plant DNA virus promoters, that are strong, constitutive promoters (for review, see Kuhlemeier *et al.*, 1987b).

There are two general groups of plant DNA viruses in which promoters can be studied—the single-stranded geminiviruses and the double-stranded caulimoviruses. The structure of the 35S RNA promoter from the caulimovirus, cauliflower mosaic virus (CaMV), has been studied in detail (Odell *et al.*, 1985; Ow *et al.*, 1987). In this paper we examine the structure of a geminivirus promoter. Geminiviruses are small plant DNA viruses with an unusual twinned particle morphology and genomes composed of one or two circular ssDNAs of about 2.6 kb (Stanley and Davies, 1985; Lazarowitz, 1987). The leaf hopper-transmitted geminiviruses, such as maize streak virus (MSV; Mullineaux *et al.*, 1984; Howell, 1984, 1985); wheat dwarf virus (WDV; MacDowell *et al.*, 1985) and beet curly top (BCTV; Stanley *et al.*, 1986), have genomes composed of one component while the whitefly transmitted viruses, such as cassava latent virus (CLV; Stanley and Gay, 1983; Stanley, 1983), tomato golden mosaic virus (TGMV; Hamilton *et al.*, 1984) and bean golden mosaic virus (BGMV; Morinaga *et al.*, 1983; Howarth *et al.*, 1985), have genomes composed of two components.

The MSV genome is organized in a manner similar to other geminiviruses (Figure 1; Stanley and Davies, 1985; Lazarowitz, 1987). Major open reading frames (ORFs) diverge rightward and leftward from a central or starting intergenic region (SIR) and, with the exception of two ORFs (Mullineaux *et al.*, 1984), they converge on the other side of the genome at another small intergenic region called the terminal intergenic region (TIR; Howell, 1984, 1985). Likewise, transcription appears to be bidirectional (Morris-Krsinich *et al.*, 1985). The most abundant RNA in infected cells results from rightward transcription of the coat protein gene (Morris-Krsinich *et al.*, 1985). Within the intergenic region of the two-component geminivirus genomes is a region of homology between the two components called the 'common region'. The common region contains sequences capable of forming a hairpin loop and a 'conserved sequence' within the loop found in all geminiviruses (Lazarowitz, 1987). Because the region is highly conserved, it is thought to contain fundamental elements required for expression of the genome. The intergenic region or TIR on the opposite side of the MSV genome may be involved in non-virion strand DNA replication. A short DNA fragment, found in virions, complementary and presumably base-paired to a region of the TIR, may serve as a primer for non-virion strand DNA synthesis (Donson *et al.*, 1984; Howell, 1984, 1985).

In this paper, we have defined the structure of the MSV rightward promoter that drives the synthesis of the major RNA transcript encoding the coat protein gene. We have found that the MSV rightward promoter is constructed in a modular fashion and have demonstrated that an upstream activating sequence (UAS) is essential for rightward transcription in a maize transient expression system. The MSV rightward promoter UAS lies in the SIR and includes structures common to all geminiviruses.

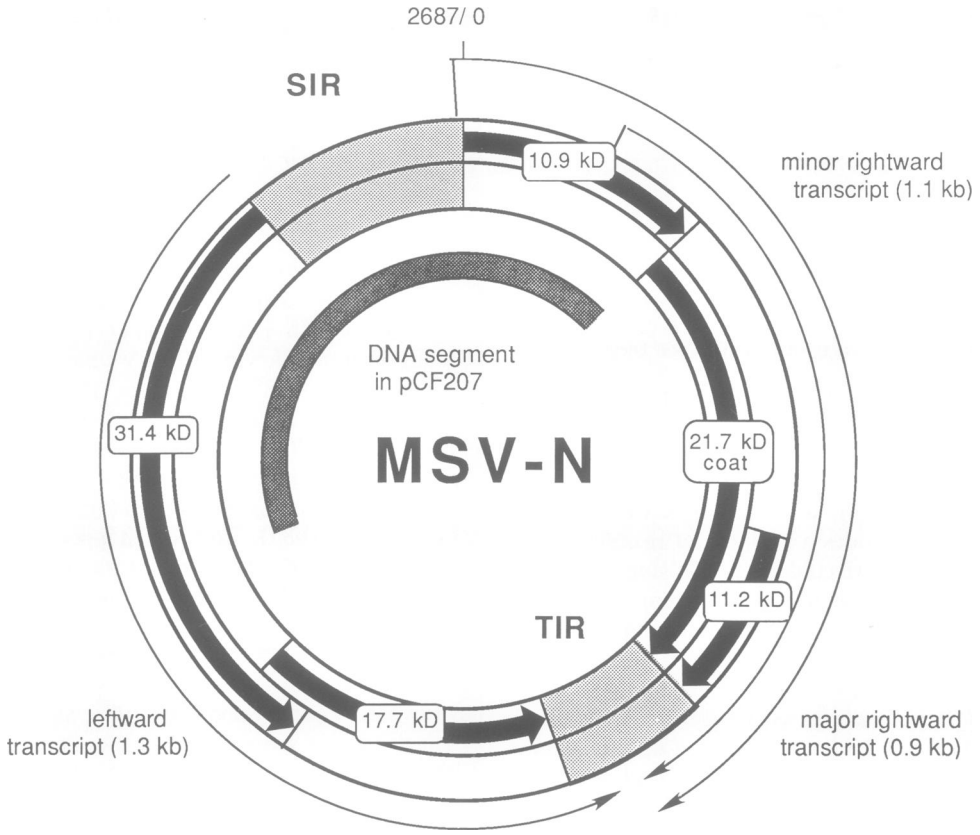

Fig. 1. Map of the MSV-N genome derived from sequence data of Mullineaux *et al.* (1984) and the transcriptional mapping data of Morris-Krsinich *et al.* (1985). The *Bam*HI site in the SIR is set at zero and genome is oriented with the virion strand in a clockwise direction (5′ → 3′). Position and direction of RNA transcripts are indicated by thin arrows; ORFs by thick arrows. Shaded segment indicates DNA fragment in pCF207 (see Figure 2).

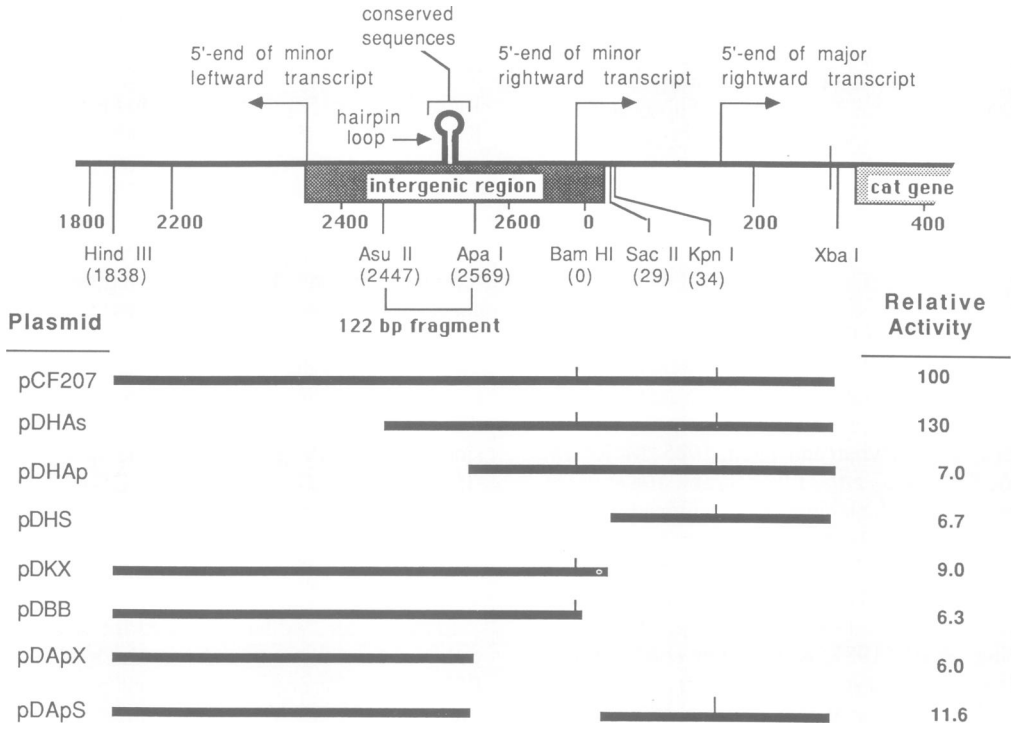

Fig. 2. Deletion map of the MSV DNA segment in pCF207. The CAT gene (with the *nos* −3′) was substituted for the MSV coat protein gene and relative CAT activity for the various constructs (normalized to the activity of pCF207) was determined in a maize protoplast transient expression system. Coordinates are expressed in terms of the genome map in Figure 1. The 5′ ends of the two rightward and one leftward transcripts which accumulate in MSV infected maize plants were determined by Morris-Krsinich *et al.* (1985). The 122 bp *Asu*II−*Apa*I fragment (UAS) required for full promoter activity is indicated. The broad black lines represent the DNA segments present in the indicated plasmids.

Results

To study rightward transcription of MSV genome and to identify the promoter responsible for expression of the coat protein gene, the 1138 base DNA fragment upstream from the coat protein gene (from position 1838 to 289, Figure 2) was linked to the chloramphenicol acetyltransferase gene (*cat* gene) in the plasmid pCF207. The MSV DNA fragment was derived from a complex region of the viral genome which included the intergenic region and sites corresponding to the 5′ ends for the major and minor rightward transcripts (Morris-Krsinich *et al.*, 1985). The plasmid pCF207 was introduced into maize protoplasts, using a polyethylene glycol transfection method, and CAT activity was determined in cell extracts 24 h later. The MSV DNA insert in pCF207 was an active promoter in the maize cell system, but only about 20% as active as the CaMV 35S RNA promoter in the same system (data not shown). The activity of the CaMV 35S RNA promoter in a maize cell transient expression system had been demonstrated previously by Fromm *et al.* (1985).

To determine what region(s) within the MSV DNA segment was responsible for promoter activity, a series of deletions was produced bounded by convenient restriction sites, and the resulting constructs were tested for CAT expression. Full activity of the MSV rightward promoter was retained in a 5′ deletion up to the *Asu*II site (as in pDHAs, Figure 2) 403 bases upstream from the start of the major rightward transcript. However, further 5′ deletions (from the *Hin*dIII site) to the *Apa*I site (in pDHAp) or to the *Sac*II site (in pDHS), 281 and 134 bases respectively, upstream from the start of the major transcript, reduced promoter activity to less than 10% of full expression. The 122 base DNA segment between the *Asu*II and *Apa*I sites was the furthest upstream fragment in this deletion series required for promoter function and hereafter we refer to that segment as the UAS. 3′ deletions, extending from the *Bst*NI site (at position 290) upstream to the *Kpn*I site (in pDKX), to the *Bam*HI site (in pDBB) or to the *Apa*I site (in pDApX), 129, 163 and 281 bases upstream from the start of the major transcript respectively, were all inactive. The deletions in pDKX and pDBB eliminated the start site of the major rightward transcript (and flanking sequences) but left intact the start of the minor rightward transcript and upstream sequences. Because the constructs were largely inactive, it is unlikely that rightward transcription in the transient expression system was initiated from the start site of the minor transcript which was left intact in these two plasmid constructs. Nonetheless, the internal deletion of the *Apa*I−*Sac*II segment (in pDApS), which removed the start site of the minor transcript and flanking sequences, was relatively inactive and, therefore, demonstrated indispensability of this region for full rightward promoter activity. However, it was not clear from this deletion whether it was the minor transcript start site *per se* that was required for CAT expression or whether the deleted region contained elements necessary for the function of a promoter which starts transcription elsewhere, such as the start site for the major transcript.

We explored in greater detail the functional properties of the 122 bp *Asu*II−*Apa*I UAS fragment, which is required for full activity of the MSV promoter. To study and manipulate this fragment, we inserted *Xho*I and *Sal*I linkers at the *Asu*II (position 2447) and *Apa*I sites (position 2569) respectively which flank the fragment (Figure 2). This permitted

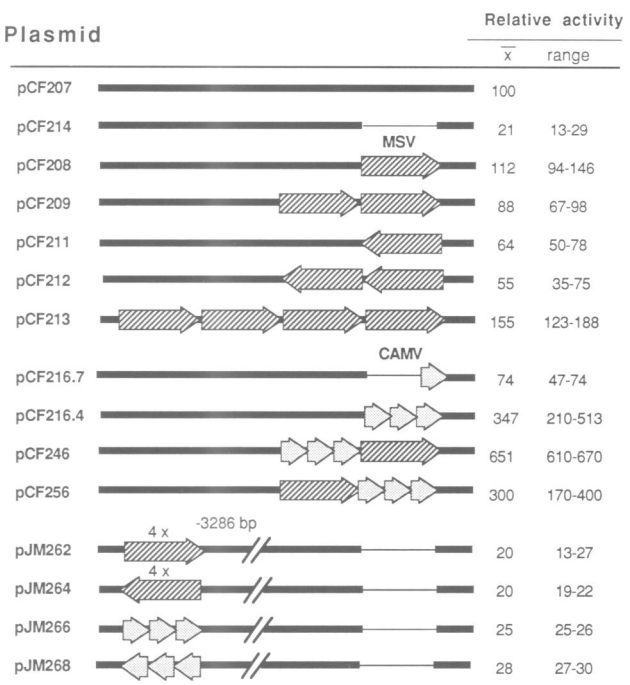

Fig. 3. Effects of inversion, multimerization and substitution of the MSV 122 bp (hatched arrow) on MSV core promoter activity. CAT activity from various constructs [the mean ($\bar{x}$) and range of at least three independent determinations and normalized to the standard pCF207] was determined in a transient expression system. pCF208, the starting plasmid for making these constructs, was produced from pCF207 by inserting *Xho*I and *Sal*I linkers at the *Asu*II site and *Apa*I site respectively. Thus, pCF208 differs from pCF207 by the presence of linkers flanking the UAS. The 60 bp promoter activating sequence (DR fragment) from the CaMV 35S RNA promoter (shaded arrow) was described by Ow *et al.* (1987). A thin line represents a region that is deleted.

the 122 bp UAS fragment to be inverted and reinserted as multimers into the rightward 'core' promoter. (The 'core' promoter was the term given to the promoter construct with a 5′ deletion to the *Apa*I site, 281 bp upstream from the start site of the major transcript.) Insertion of linkers on either side of the *Asu*II and *Apa*I sites (in pCF208, Figure 3) had no demonstrable effect on promoter activity when compared with the standard pCF207. As expected from the deletion analysis described above, when the 122 bp fragment was deleted and the flanking *Xho*I to *Sal*I sites rejoined (as in pCF214) promoter activity dropped, in this case to 21% the level of the standard pCF207 activity.

Insertion of multiple copies of the 122 bp fragment into the 'core' promoter had limited stimulatory effects. The promoter with two copies of the UAS (pPC209, Figure 3) actually had less activity than the promoter with one copy. However, insertion of four copies (pCF213) stimulated expression to 155% the level of the standard pCF207 activity. The direction of the 122 bp element was reversed (in pCF211) to determine whether the element could function in an orientation-independent fashion. We found that it could, although the construct with one reversed copy had only 64% the activity of the standard construct. Insertion of two copies of the 122 b fragment in the reverse orientation (pCF212) did not stimulate expression further. Thus, the 122 base fragment is limitedly orientation-independent in its ability to stimulate the MSV core promoter.

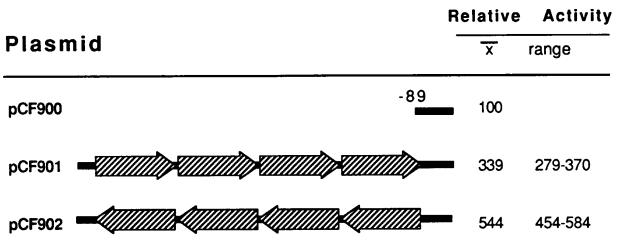

Fig. 4. Effects of the MSV 122 bp UAS (hatched arrows) on a heterologous core promoter, a −89 truncated CaMV 35S promoter (Ow *et al.*, 1987). The mean CAT activity ($\bar{x}$) is expressed relative to pCF900 containing the truncated CaMV promoter.

The CaMV 35S RNA promoter also has an upstream segment called the distal region or DR segment which is required for full activity of the 35S RNA promoter (Ow *et al.*, 1987); therefore, we attempted to activate the MSV 'core' promoter with the 60 bp DR segment from CaMV 35S RNA promoter. Insertion of one CaMV DR segment (pCF216.7, Figure 3) stimulated transcription from the MSV promoter but not to the level of the pCF207 standard. Insertion of a triplet of CaMV DR segments (pCF216.4) further enhanced transcription to levels 3.5 times that of the standard MSV promoter. Hence, the CaMV DR segment appears to substitute for the function of the 122 bp MSV UAS and in multiple copies can stimulate expression of the MSV core promoter at even higher levels than multiple copies of the MSV UAS. We also asked whether the effect of the MSV and CaMV activating segments was additive. When the triplet of CaMV DR segments was inserted in the MSV promoter in addition to and immediately upstream from the MSV 122 bp UAS (pCF236), we observed the highest level of MSV promoter activity, 6.5 times that of the standard pCF207 promoter. However, when the order of the segments was reversed such that the 122 bp MSV segment was upstream from the triplet of CaMV segments, the promoter activity was no greater than with the triplet of CaMV segments alone. Thus, the effect of the combination of MSV and CaMV segments is more than additive when the MSV segment is proximal to the start of transcription.

To test whether these segments could act from a distant site, we relocated the set of multiple copies of both the MSV and CaMV segments to a site 3286 bp upstream from the normal insertion site (Figure 3). (Because the plasmid is circular, the site can also be considered to be 2093 bases downstream from the normal insertion site.) Neither four copies of the MSV segment (pJM262 or pJM264) nor three copies of the CaMV segment (pJM266 or pJM268) in either orientation had any effect on the activity of the MSV 'core' promoter. Hence, neither the MSV nor CaMV promoter activating segment appears to act from a distance or in a position-independent fashion.

We sought to determine whether the MSV UAS could activate a heterologous promoter, in particular whether it was functionally equivalent to the DR segment of the CaMV 35S RNA promoter. To test this, we appended MSV UASs onto a CaMV 35S RNA core promoter fragment. The CaMV core promoter fragment is a −89 base fragment containing a TATA and CAAT box which can be activated by the CaMV DR segment (Ow *et al.*, 1987). We found that multiple MSV UASs could activate the CaMV core promoter such as in pCF901 (Figure 4), which contains four copies of the

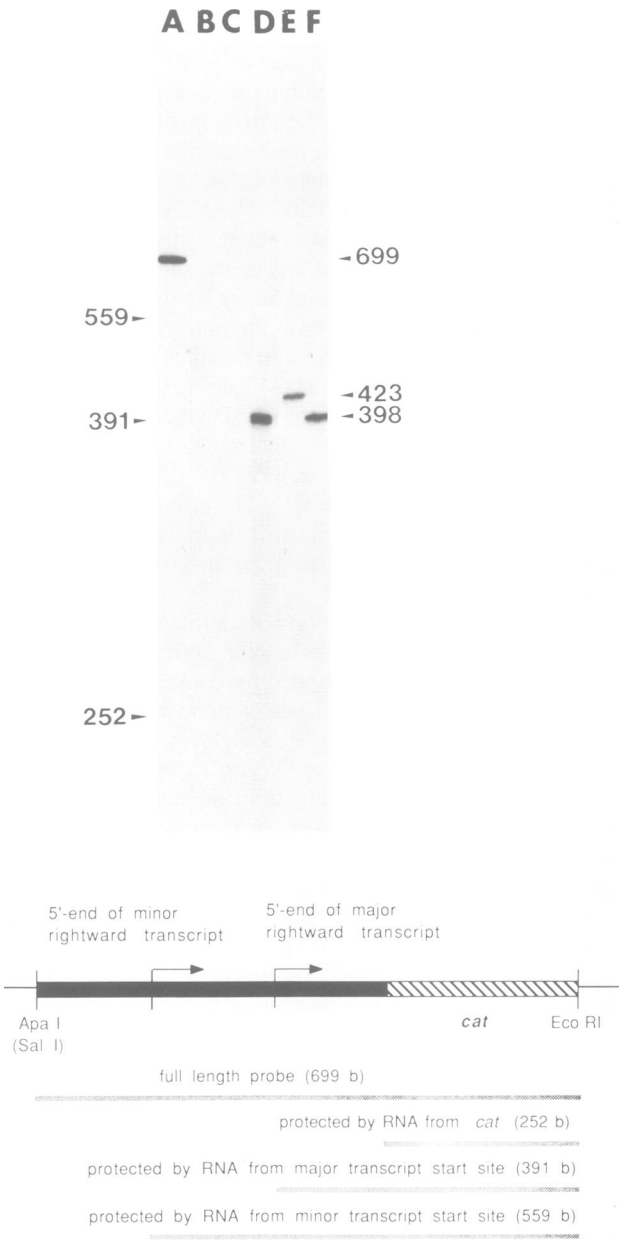

Fig. 5. Mapping the 5′ end of the RNA transcripts synthesized in the maize protoplast transient expression system. Protoplasts were transfected with pCF216.4 (Figure 3) in which multimers of the CaMV 35S promoter activating segments were substituted into the MSV promoter. The 5′ end was determined by RNase I protection of a labeled 699 base RNA probe synthesized *in vitro*. Probe was not digested with S1 nuclease (**lane A**), or digested with S1 nuclease following hybridization to RNA from mock-transfected cells (**lane B**), from cells transfected with an inactive promoter deletion construct, pDApX (**lane C**) or from cells transfected with pCF216.4 (**lane D**). **Lanes E** and **F** are RNAs protected by *in vitro* synthesized, complementary transcripts of a size to give RNA marker fragments of 423 and 398 bases in length respectively. The gel migration positions of the predicted RNA fragments shown in the diagram at the bottom are indicated on the left side of the gel. The migration positions were determined from the sequence of a known DNA fragment loaded on the same gel.

MSV UAS oriented in the forward direction. UAS copies oriented in the opposite direction (pCF902) also activated the CaMV core promoter and at this copy number level were more effective in the reverse orientation than in the forward direction.

Determining the start site for transcription in the transient expression system

Morris-Krsinich *et al.* (1985) mapped the sites corresponding to the 5′ ends of the polyadenylated viral RNAs from infected maize plants on the MSV genome. The most abundant RNA was a 0.9 kb transcript starting at nucleotide position 163 and extending rightward through the coat protein gene (Figure 2). A minor rightward transcript was also found starting 168 base upstream at position 2682. In the deletion analysis described above, we observed that the regions surrounding both start sites were indispensable for rightward transcription. Hence, it was unclear which site (if either) was the start of transcription in the transient expression system.

To determine the start of transcription in the transient expression system, we took advantage of the enhanced expression of the MSV promoter with multiple CaMV promoter activating elements. RNA extracted from cells transfected with pCF216.4 (Figure 3) was analyzed using S1 nuclease protection procedures. Extracted RNA was hybridized to a 699 base cRNA probe (Figure 5, bottom and lane A) synthesized *in vitro* from a DNA fragment extending from an *Eco*RI site in the *cat* gene upstream to the *Apa*I site in the MSV DNA fragment (position 2569 in the MSV genome, converted to a *Sal*I site by insertion of a *Sal*I linker). RNA extracted from cells that were mock-transfected (lane B) or transfected with the inactive pDApX promoter deletion construct which still contains the *cat* gene (lane C) did not protect the probe to any appreciable extent. (If RNA was synthesized in cells transfected with pDApX, the RNA would protect a fragment of 252 bases.) RNA from cells transfected with pCF216.4 protected a 390–395 base fragment (lane D). This corresponded to the expected size of a fragment protected by RNA originating at the major transcript start site (Figure 5, bottom). There was no evidence of a larger fragment (expected size of 559 bases) protected by RNA starting at the minor transcript start site. Thus, barring any undetected processing events, it appears that in the transient expression system transcription from the modified MSV promoter in pCF216.4 starts at the same site as the major transcript starts *in vivo*.

The MSV UAS binds proteins from maize cell nuclear extracts

We sought to determine whether proteins in a nuclear extract from maize cells used in the transient expression system recognized and bound specifically to the MSV 122 bp UAS. The labeled MSV DNA fragment was incubated with the nuclear extract and subjected to electrophoresis on polyacrylamide gels. Incubation with nuclear extracts led to the formation of DNA–protein complexes which migrated as several retarded bands on the gel (Figure 6). The DNA–protein complexes resolved into five major retarded bands—a group of three closely spaced, slowly migrating bands (slow bands) and another group of two more rapidly migrating bands (fast bands). The complexes appeared to be specific for the 122 bp DNA fragment because they were formed in the presence of 50 ng/μl dI/dC (Figure 6, lane 2), but could be competed away by the addition of excess unlabeled 122 bp MSV UAS DNA (lanes 3–6). The three slow bands were competed away by a lower concentration of the unlabeled fragment than was required to compete away the two fast bands. About 50% of each of the slow bands

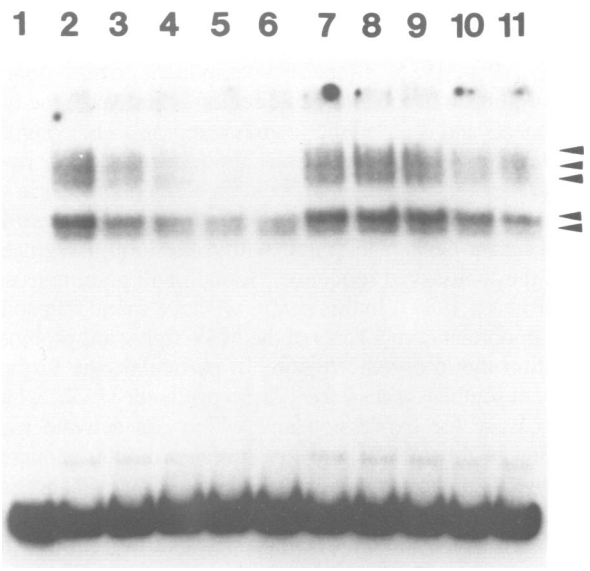

Fig. 6. Gel retardation assay for protein–DNA complexes. Maize crude nuclear extract was incubated with 0.1 ng of end-labeled MSV UAS (*Xho*I–*Sal*I fragment from pCF208) in the presence of 1 μg poly (dI–dC). **Lane 1**, no nuclear extract; **lanes 2–11**, incubated with nuclear extract. **Lanes 3–6** have 0.1, 1, 5 and 10 ng respectively of the unlabeled UAS as competitor. **Lanes 8–11** have 0.1, 1, 5 and 10 ng respectively of the CaMV DR fragment (*Xho*I–*Sal*I fragment from pCF216.4). Lanes 2 and 7 have no unlabeled competitor DNA. Arrows indicate position of five major retarded bands.

were competed away with an equivalent amount of unlabeled MSV DNA fragment (compare lanes 2 and 3, Figure 6), while ~50% of the fast band material was competed away with a 10-fold excess of the unlabeled fragment (compare lanes 2 and 4, Figure 6). We do not know whether the different complexes were formed by different proteins associating with the MSV UAS or by the same proteins associating with the MSV UAS in a different way. None the less, the factors which contributed to the formation of the slow complexes were required in higher concentration in the incubation reaction than those forming fast complexes.

Because the DR promoter activating segment from the CaMV genome was able to replace the function of the 122 bp fragment in the MSV promoter, we asked whether the CaMV DR segment would compete with the 122 bp MSV fragment in the formation of DNA–protein complexes. The CaMV DR segment competed more weakly (and, perhaps, nonspecifically) than the MSV UAS competed with itself. The CaMV DR segment competed poorly for the formation of the fast bands even at a 50-fold excess of CaMV DNA (Figure 6, lanes 8–11). One might expect that if the CaMV and MSV promoter fragments bound the same factors that the CaMV fragment might have greater binding factor affinities, given its greater promoter activation abilities. However, it is probably more reasonable to conclude that the CaMV DR segment binds different proteins than the MSV fragment, particularly in the formation of complexes constituting the MSV fast bands.

Discussion

The genomes of geminiviruses, whether they are composed of one or two components, are organized in a similar manner (Stanley and Davies, 1985; Lazarowitz, 1987). MSV, a single component geminivirus, has a central or starting inter-

genic region (SIR) with ORFs diverging bidirectionally around the genome (Figure 1; Mullineaux *et al.*, 1984; Howell, 1984, 1985). In the two component geminiviruses, a major portion of the intergenic region is similar in the two components and is called the 'common region'. The common (or intergenic) regions of various geminiviruses differ from one another but, nonetheless, contain elements found in all other geminiviruses. Most notable is a region that can be folded into a hairpin loop and within the loop a segment, called the 'conserved sequence', found in all geminiviruses (Lazarowitz, 1987). In this study, we have found that some of the important components of the MSV rightward promoter lie within the intergenic region. In particular, the furthest upstream segment (called the 122 bp promoter UAS), which is necessary for expression and which can activate transcription of the MSV rightward 'core' promoter, contains the conserved elements described above. However, preliminary experiments in our laboratory have demonstrated that, within the UAS, the hairpin loop may not be required for full expression, but that GC-rich segments immediately upstream from the hairpin loop are essential. It is interesting that in another geminivirus, squash leaf curl virus (SqLCV), variations in this region are associated with differences in host range (Lazarowitz, 1987).

Like the MSV promoter, the CaMV 35S promoter also contains an upstream segment called the distal region (DR), necessary for promoter expression. The CaMV 35S promoter DR can substitute for the MSV UAS and drive expression of an MSV 'core' promoter (281 bp MSV DNA fragment extending upstream from the start of transcription). Likewise, the MSV UAS can activate a CaMV core promoter. Because maize is not a host for CaMV, it was unexpected when Fromm *et al.* (1985) reported that the CaMV 35S RNA promoter was active in a maize transient expression system and when we found that the *apparent* strength of the CaMV promoter was even greater than that of the MSV rightward promoter in the same system. (This conclusion pertains to the level of CAT expression obtained from the particular constructs we have used in our experiments. However, it should be pointed out that not only do the promoters differ in these constructs, but also the untranslated leader sequences of the RNA transcripts encoding CAT. Therefore, we cannot discount the fact that the apparent promoter strength may also reflect differential stability or translation of the RNAs in the transient expression system.) Furthermore, the CaMV DR segment (in multiple copies) is more effective than the MSV UAS in empowering the MSV core promoter. It is possible that the difference is due to a greater tissue-specific limitation in expression of the MSV promoter. MSV is normally a phloem-limited virus and the expression of the rightward promoter may be greater in phloem cells than in suspension culture cells used in the transient expression assay. Another possibility is that MSV encodes an early transactivating protein that stimulates rightward transcription in virus-infected cells.

Even though the CaMV and MSV promoter activating segments are functionally equivalent, they are not similar in sequence. Furthermore, the segments apparently bind different nuclear factors since the CaMV DR fragment does not compete as effectively for the factors that bind to the MSV UAS. Alternatively, the segments may bind similar factors but with different affinities. The MSV UAS forms several different complexes which are resolvable as at least five major bands (and several minor bands) on retardation gels. Preliminary experiments demonstrate that the MSV UAS does not have as many factor binding sites as it has resolvable bands and that several of these bands may represent different forms of the same basic complex. The MSV rightward promoter appears to be less compact and more complex than the CaMV 35S RNA promoter. The critical elements of the CaMV 35S RNA promoter lie within 134 bases of the start of transcription (Ow *et al.*, 1987), while those of the MSV promoter extend about 400 bases upstream. Furthermore, the MSV promoter appears to have multiple interactive elements, because the UAS is only capable of activating a relatively large MSV 'core' promoter (281 base DNA fragment upstream from the start of transcription). The UAS does not activate a smaller MSV core promoter as demonstrated by the inactivity of the pDApS construct (Figure 2) which has an intact UAS but is missing the *ApaI*−*Sac*II fragment. Thus, it seems likely that the *ApaI*−*Sac*II fragment contains an upstream element or elements (between 134−281 bp upstream) required in concert with the 122 base UAS for function of the MSV rightward promoter.

Three RNA transcripts which accumulate in MSV infected plants originate within or near to the intergenic region. We have studied the synthesis of only one of these transcripts—the major rightward RNA, which presumably serves as a message for synthesis of the coat protein (21.7 kd protein). The 5′ end of the major transcript from infected plants lies outside the intergenic region within the first rightward ORF (10.9 kd protein, Figure 1). We have shown that the start site of the MSV RNA synthesized in the transient expression system corresponds to the start site of the major rightward transcript in infected plants. Two other transcripts start within the intergenic region—a minor rightward transcript and a minor leftward transcript. The minor rightward transcript starts just upstream from the first rightward ORF (10.9 kd protein) and we should have been able to detect this RNA if it was synthesized in the transient expression system; however, we did not. A possibility that we cannot discount is that minor rightward transcript is a precursor of the major transcript. If that is so, then it is surprising that we find no evidence of the unprocessed form in the transient expression system. Internal deletions at the start site of the minor rightward transcript (and flanking sequences) reduce rightward expression to low levels. However, the reduction in rightward expression may not be due to the loss of the start site of transcription but may result from the deletion of critical elements in the major rightward promoter which lie at or near the minor transcript start site, as discussed above. In any case, the intergenic region may have multiple promoter functions, certainly one of which is acting as a bidirectional promoter.

Until recently, it has not been possible to mechanically inoculate plants or infect cells in culture with MSV or its DNA in order to study the replication cycle of the virus. As a result, it has not been possible to determine whether different parts of the MSV genome are expressed in a temporal pattern. However, Grimsley *et al.* (1986) have reported on a technique for the inoculation of maize plants with MSV DNA introduced via the *Agrobacterium tumefaciens* Ti-plasmid. Thus, it may be possible to study the mode of operation of this complex promoter region in MSV infected plant cells in the near future.

Materials and methods

Construction of plasmids

The basic plasmid used in this study from which most other plasmids were derived is pCF207, an MSV promoter-*cat*-nopaline synthetase (*nos*) 3′ fusion in pUC19. Plasmid pCF207 contains a 1138 bp *Hin*dIII−*Eco*RII fragment from an MSV-N genome as shown in Figures 1 and 2. The MSV-N (Nigerian isolate) genome was cloned from virion DNA independently of Mullineaux *et al.* (1984), and the fragment used in pCF207 was derived from a tandem dimer of the genome in pMSV-7N-4D (unpublished). The map numbering system of Mullineaux *et al.* (1984) was used and was identical, except for 1 base change, for the region under study of this MSV genome (M.Schneider, unpublished observations). The MSV DNA fragment in pCF207 (Figures 1 and 2) contains the region immediately upstream from the coat protein gene, the two putative start sites for the rightward major and minor RNA transcripts (Morris-Krsinich *et al.*, 1985), the entire starting intergenic region and 537 bp from the first leftward ORF (Mullineaux *et al.*, 1984). The *cat*−*nos* 3′ component of the plasmid was a *Bam*HI−*Stu*I fragment obtained from pNCAT4 (kindly provided by J.Schell) and subcloned into pDO372 (D.W.Ow, unpublished).

Plasmid pCF208 is a derivative of pCF207 in which linkers were inserted at sites flanking the 122 bp MSV rightward promoter UAS—an *Xho*I linker was inserted at the *Asu*II site (position 2448) and a *Sal*I linker at an *Apa*I site (position 2573). Linkers were inserted so that the site occupied by the UAS could be easily accessed and so that the UAS could be reinserted, inverted or multimerized.

Plasmids containing the 60 bp activating segment (DR segment) from the CaMV 35S RNA promoter (Ow *et al.*, 1987) were constructed by inserting 1 or 3 CaMV DR segments at the *Xho*I linker site (pCF236, pCF256), *Sal*I linker site (pCF226, pCF246) or between the *Xho*I and *Sal*I linker sites (pCF216.7, pCF216.4) in pCF208. The CaMV 60 bp promoter activating segment was obtained from a derivative of pJO62 (Ow *et al.*, 1987) containing an *Xho*I site at −148 and a *Sal*I site at −89 with respect to the start of the 35S RNA transcription.

Plasmids pJM262−pJM268 with the UAS relocated a distance upstream (or downstream) from the promoter site were constructed by James Masuoka. These plasmids were constructed by inserting the *Xho*I−*Sal*I fragment from pCF213, containing four copies of the 122 bp MSV promoter activating segment, or a similar fragment from pCF216.4, containing three copies of the 60 bp CaMV DR segment, into the *Xho*I site of a derivative of pCF214 (Figure 3) which has an *Xho*I linker at the *Kpn*I site at the end of the *nos* 3′ segment.

Plasmid pCF900, a *cat*−*nos* 3′ gene construct which bears a truncated or 'core' CaMV 35S RNA promoter, was produced by replacing the luciferase gene in pJD4x (Ow *et al.*, 1987) with *cat*. Derivatives pCF901 and pCF902 were obtained by inserting the 122 bp MSV UAS from pCF213 at the *Xho*I site flanking the upstream side of the truncated CaMV 35S promoter in pCF900.

Transient expression assay

Maize protoplasts were obtained from a *Zea mays* var. Black Mexican Sweet cell suspension (P.Chourey, University of Florida) maintained continuously in liquid MSZM medium (Chourey and Zurawski, 1981). Cells growing at 26°C were transferred to new medium at regular 5 day intervals and harvested when the cell density reached 10−15 ml of packed cells per 200 ml of culture. Protoplasts were produced by digesting harvested cells for 4 h at 26°C in PWS (0.2 M mannitol, 80 mM $CaCl_2$, 26 mM MES, pH 5.6) containing 1% (w/v) cellulase (Calbiochem) and 0.5% (w/v) hemicellulase (Sigma Chemicals) and pectinase (Calbiochem). Digested cells were filtered successively through 125 and 62 μm polypropylene mesh filters and the resulting protoplasts were sedimented, washed and resuspended in PWS at a final density of 5×10^6 protoplasts/ml.

Cells were transfected using a polyethylene glycol method modified from Krens *et al.* (1982). Protoplasts (2×10^6) were mixed with 100 μg of CsCl gradient-purified plasmid DNA and 25 μg of sonicated carrier calf thymus DNA. Polyethylene glycol 6000 (Matheson Scientific) dissolved in F medium (Krens *et al.*, 1982) was added to a final concentration of 25%, and the mixture (1.5 ml total vol.) was incubated at room temperature for 30 min. After stepwise dilution with 8.5 ml of F medium, the transfected protoplasts were sedimented, resuspended in CMM [0.2 M mannitol, 1.4 mM glucose, 5 mM $CaCl_2$, 2% coconut water (Gibco) in MSZM medium] at 2×10^5 protoplasts/ml and incubated at 26°C for 20−24 h. (Incubation time course experiments showed that expression reached a maximum by 20−24 h.) Incubated cells were harvested by centrifugation, suspended in 250 mM Tris−HCl (pH 7.8) to a final volume of 500 μl and lysed by three cycles of freeze−thawing and 10 s sonication with a Branson sonifer equipped with a microtip. The extract was clarified by centrifugation for 5 min in a microcentrifuge at 4°C. Aliquots (150−180 μl) of the clarified extract were heated to 63°C for 10 min, and the extracts were tested for CAT activity essentially as described by Gorman *et al.* (1982) except that reactions were incubated for 4 h at 37°C. The amount of ^{14}C-labeled acetylated products produced in each reaction and separated by TLC was measured by scintillation counting.

Preparation of maize nuclear extracts

Nuclei were isolated from maize suspension culture cells used in the transient expression assays by a procedure modified from Paul *et al.* (1987). Cells from 200 ml of culture in log growth were collected by centrifugation and resuspended in 100 ml of nuclear isolation buffer (NIB) containing 50 mM Tris−HCl (pH 7.5), 0.3 M sucrose, 5 mM $MgCl_2$, 15 mM KCl, 1 mM dithiothreitol (DTT), 0.1 mM EDTA and 0.2 mM phenylmethylsulfonyl fluoride (PMSF). Resuspended cells were collected on a 125 μm nylon mesh filter and washed with an additional 100 ml of NIB to remove debris. Washed cells (10 g) were resuspended in about 5 ml of NIB and ground for 2 min on ice in a motor-driven homogenizer with a Teflon pestle. The homogenate was diluted to about 50 ml with NIB and filtered through a 125 μm mesh filter. A crude nuclear pellet was obtained by centrifuging the filtrate at 4300 *g* for 10 min in a Sorvall SS-34 rotor, removing the supernatant layer, centrifuging the pellet again for 1 min at the same speed and carefully removing all the supernatant fluid.

Proteins were extracted on ice from the crude nuclear pellet in 1 ml of protein extraction buffer (PEB) containing 10 mM Tris−HCl (pH 7.5), 1 mM DTT, 0.4 M NaCl and 0.2 mM PMSF. To do so, the pellet was resuspended using a glass rod and the mixture was mixed on a vortex mixer 2−3 times over a 30 min interval. The extract was clarified by centrifugation at 4°C in a microcentrifuge for 15 min and frozen at −80°C after adding glycerol to 16%. The usual protein concentration obtained in the extracts was 1−2 μg/μl.

Gel retardation assay

Gel retardation assays for the detection of DNA−protein complexes formed between maize nuclear proteins and the 122 bp MSV UAS were performed as modified from Fried and Crothers (1981). The binding reactions (20 μl) which were incubated for 15 min at room temperature contained 15 μl with 0.5−1 ng of ^{32}P-end labeled DNA fragment, 1 μg poly(dI−dC), 13 mM Tris− HCl (pH 7.5), 1.3 mM DTT. A 2-fold dilution of the protein extract in 5 μl of PEB with 16% glycerol was added as indicated without further mixing to avoid shearing formed complexes. Following the reaction, the sample was loaded on a 1.5 mm thick standard 4% polyacrylamide gel (30:1::polyacrylamide:bisacrylamide) in 0.25 × TEB buffer (Maniatis *et al.*, 1982) and run at 10 V/cm.

RNA extraction and S1 nuclease protection

RNA was extracted from cells 20 h after transfection by a modification of the method of Chirwin *et al.* (1979). Cells (~10^7) were ground under liquid N_2 with a mortar and pestle and the frozen powder was added to extraction buffer containing 5 M guanidinium thiosulfate, 5% β-mercaptoethanol (v/v) and 50 mM Hepes buffer (pH 7.4). The RNA-containing mixture was clarified by centrifugation at 12 000 *g* for 20 min and to the supernatant fluid was added the detergent Sarkosyl to 2% and CsCl to 0.57 M final concentration. The mixture was heated to 65°C for 10 min and the RNA sedimented through a cushion containing 5.7 M $CsCl_2$ for 6 h at 106 000 *g*. The pelleted RNA was resuspended in water, phenol extracted and ethanol precipitated.

Nuclease S1 protection procedures were carried out using cRNA or riboprobes as described by Quarless and Heinrich (1986). The cRNA probes were prepared and ^{32}P-labeled by *in vitro* transcription of the 690 bp *Sal*I−*Eco*RI fragment of pCF208 cloned into pGEM-1 (Promega). The DNA template was digested with RNase-free DNase and the probe was phenol extracted and ethanol precipitated. Labeled probe (2×10^6 c.p.m.) was added to 20 μl of hybridization buffer (80% formamide, 0.4 M NaCl, 1 mM EDTA and 40 mM MOPS, pH 7) containing 40 μg of total cellular RNA plus 40 μg of sheared calf thymus DNA, and the mixture was denatured by heating at 85°C for 15 min and then incubated under mineral oil at 63°C for 3 h. Following hybridization, the samples were diluted 10-fold with ice-cold S1 nuclease digestion buffer [100 mM NaCl, 2 mM $ZnSO_4$, 5% glycerol (v/v) and 30 mM Na acetate, pH 4.6] and digested with 2000 U of S1 nuclease for 1 h at 37°C. The samples were phenol extracted and the S1 resistant hybrids were ethanol precipitated in the presence of 50 μg carrier tRNA and dissolved in 80% formamide. The samples were heated to 85°C for 15 min and loaded onto a 6% polyacrylamide sequencing gel (Maniatis *et al.*, 1982). The protected fragments of the radioactive probe were detected by autoradiography of the dried gel.

Acknowledgements

We would like to thank Michel Schneider, David W.Ow and James Masuoka for their contribution to this project. This work was supported by the Competitive Research Grants Program of the US Department of Agriculture, the US National Science Foundation and a Fulbright Postdoctoral Fellowship to C.F.

References

Chen,Z.-L., Schuler,M.A. and Beachy,R.N. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8560–8564.

Chirwin,J.M., Przybyla,A.E., MacDonald,R. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.

Chourey,P.S. and Zurawski,D.B. (1981) *Theor. Appl. Genet.*, **59**, 341–344.

Donson,J., Morrish-Krsinich,B.A.M., Mullineaux,P.M., Boulton,M.I. and Davies,J.W. (1984) *EMBO J.*, **3**, 3069–3073.

Dynan,W.S. and Tjian,R. (1985) *Nature*, **316**, 774–778.

Ebert,P.R., Ha,S.B. and An,G. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5745–5749.

Ellis,J.G., Llewellyn,D.J., Dennis,E.S. and Peacock,W.J. (1987a) *EMBO J.*, **6**, 11–16.

Ellis,J.G., Llewellyn,D.J., Dennis,E.S. and Peacock,W.J. (1987b) *EMBO J.*, **6**, 3203–3208.

Fluhr,R., Kuhlemeier,F., Nagy,F. and Chua,N.-H. (1986) *Science*, **232**, 1106–1112.

Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.

Fromm,M., Taylor,L.P. and Walbot,V. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5824–5828.

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.

Green,P.J., Kay,S.A. and Chua,N.-H. (1987) *EMBO J.*, **6**, 2543–2549.

Grimsley,N., Hohn,T., Davies,J.W. and Hohn,B. (1987) *Nature*, **325**, 177–179.

Gurley,W.B., Czarnecka,E., Nagao,R.T. and Key,J.L. (1986) *Mol. Cell. Biol.*, **6**, 559–565.

Hamilton,W.D.O., Stein,V.E., Coutts,R.H.A. and Buck,K.W. (1984) *EMBO J.*, **3**, 2197–2205.

Howarth,A.J., Caton,J., Bossert,M. and Goodman,R.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3572–3576.

Howell,S.H. (1984) *Nucleic Acids Res.*, **12**, 7359–7375.

Howell,S.H. (1985) *Nucleic Acids Res.*, **13**, 3018–3019.

Jofuku,K.D., Okamuro,J.K. and Goldberg,R.B. (1987) *Nature*, **328**, 734–737.

Kadonga,J.T., Jones,K.A. and Tjian,R. (1986) *Trends Biol. Sci.*, **11**, 20–23.

Kaulen,H., Schell,J. and Kreuzaler,F. (1986) *EMBO J.*, **5**, 1–8.

Krens,F.A., Molenduk,L., Wullems,G.J. and Schildperoot,R.A. (1982) *Nature*, **296**, 72–74.

Kuhlemeier,C., Fluhr,R., Green,P.J. and Chua,N.-H. (1987a) *Genes Devl.*, **1**, 247–255.

Kuhlemeier,C., Green,P.J. and Chua,N.-H. (1987b) *Annu. Rev. Plant Physiol.*, **38**, 221–257.

Lazarowitz,S.G. (1987) *Plant Mol. Biol. Reporter*, **4**, 177–192.

MacDowell,S.W., Macdonald,H., Hamilton,W.D.O., Coutts,R.H.A. and Buck,K.W. (1985) *EMBO J.*, **4**, 2173–2180.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McKnight,S. and Tjian,R. (1986) *Cell*, **46**, 795–805.

Morelli,G., Nagy,F., Fraley,R.T., Rogers,S.G. and Chua,N.-H. (1985) *Nature*, **315**, 200–204.

Morinaga,T., Ikegami,M. and Miura,K. (1983) *Proc. Jap. Acad.*, **59**, 363–366.

Morris-Krsinich,B.A.M., Mullineaux,P.M., Donson,J., Boulton,M.I., Markham,P.G., Short,M.N. and Davies,J.W. (1985) *Nucleic Acids Res.*, **13**, 7237–7256.

Mullineaux,P.M., Donson,J., Morris-Krsinich,B.A.M., Boulton,M.I. and Davies,J.W. (1984) *EMBO J.*, **3**, 3063–3068.

Nagy,F., Boutry,M., Hsu,M.-Y., Wong,M. and Chua,N.-H. (1987) *EMBO J.*, **6**, 2537–2542.

Odell,J.T., Nagy,F. and Chua,N.-H. (1985) *Nature*, **313**, 810–812.

Ow,D.W., Jacobs,J.D. and Howell,S.H. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4870–4874.

Paul,A.-L., Vasil,V., Vasil,I.K. and Ferl,R.J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 799–803.

Picard,D. (1985) *Oxford Surveys on Eucaryotic Genes*. Oxford University Press, Oxford, Vol. 2, pp. 24–48.

Quarless,S.A. and Heinrich,G. (1986) *Biotechniques*, **4**, 434–438.

Simpson,J., Schell,J., Van Montagu,M. and Herrera-Estrella,L. (1986) *Nature*, **323**, 551–554.

Stanley,J. (1983) *Nature*, **305**, 643–645.

Stanley,J. and Gay,M.R. (1983) *Nature*, **301**, 260–262.

Stanley,J. and Davies,J.W. (1985) In Davies,J.W. (ed.), *Molecular Plant Virology*. CRC Press, Boca Raton, Florida, Vol. II, pp. 161–218.

Stanley,J., Markham,P.G., Callis,R.J. and Pinner,M.S. (1986) *EMBO J.*, **5**, 1761–1767.

Struhl,K. (1987) *Cell*, **49**, 295–297.

Timko,M.P., Kausch,A.P., Castrensana,C., Fassler,J., Herrera-Estrella,L., Van den Broeck,G., Van Montagu,M., Schell,J. and Cashmore,A.R. (1985) *Nature*, **318**, 579–582.

Received on December 16, 1987; revised on February 22, 1988