The EMBO Journal vol.10 no.12 pp.3887-3896, 1991

Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein

Johannes Fütterer and Thomas Hohn

Friedrich Miescher-Institut, Basel, Switzerland

Communicated by T.Hohn

Polycistronic mRNAs containing an upstream β-glucuronidase (GUS) and a downstream chloramphenicol acetyltransferase (CAT) reporter open reading frame (ORF) were expressed in transfected plant protoplasts. CAT expression could be strongly induced by coexpression of the cauliflower mosaic virus encoded translation transactivator. Transactivation was abolished when an upstream ORF overlapped the CAT ORF for a long distance. No specific sequence elements were required for transactivation but the presence of a short ORF upstream of the GUS ORF strongly enhanced the process. The inhibitory effect of additional presumed stem structures inserted into various regions of the reporter mRNAs indicates that both ORFs are translated by ribosomes that associate with the RNA at the 5′ end and reach the ORFs by a linear migration mechanism.
Key words: caulimovirus/pararetrovirus/ribosome scanning/transient expression/translational control

Introduction

On most eukaryotic mRNAs, binding of 43S ribosome initiation complexes occurs at the capped 5′ end (for review, see Pain, 1986; Rhoads, 1988) and is followed by scanning of the ribosome along the RNA until a suitable translation initiation site is encountered (Kozak, 1989a). At this site, usually an AUG codon, the 80S ribosome is assembled and peptide synthesis begins. Upon translation initiation at least some of the originally bound initiation factors are released. In some cases mRNAs may contain several translation initiation sites. Initiation at a downstream site can only be explained with modifications of the simple scanning model or consideration of entirely different mechanisms for start site selection:

(i) Initiation competent ribosomes can bypass an AUG codon located close to the cap (Sedman *et al.*, 1989a) or in unfavourable context (Kozak, 1986a) and reach another one further downstream (leaky scanning); (ii) after translation of a reading frame a ribosome might continue scanning, regain a new set of initiation factors and reinitiate at a downstream ORF (Kozak, 1987; Abastado *et al.*, 1991); (iii) in certain cases ribosomes might enter at an internal ribosome entry site (IRES; Jang *et al.*, 1988; Pelletier and Sonenberg, 1988); and (iv) for translation in avian sarcoma virus (Darlix *et al.*, 1982), Sendai virus (Curran and Kolakofsky, 1988) and CaMV (Hull, 1984; Fütterer *et al.*, 1990a) a mechanism has been proposed that combines features of both the scanning model and the internal entry model. According to this 'shunt' model, ribosomes bind to the RNA 5′ end, start scanning normally but are then transferred from a 'take-off point' directly to more internal regions of the mRNA.

Internal ribosome entry (for review, see Jackson *et al.*, 1990; Jang *et al.*, 1991) and ribosome shunt (Fütterer *et al.*, 1990a) clearly are exceptions and both require RNA *cis*-acting primary and/or secondary structure elements and factors. Downstream ORFs on polycistronic RNAs lacking such specific sequence features are thought to be translated by either leaky scanning or reinitiation. Leaky scanning seems to be applicable only when the downstream cistron is preceded by just one, or at most very few, initiation sites (for review, see Kozak, 1989a; Schwarz *et al.*, 1990), whereas reinitiation would be the mechanism of choice for RNAs with two or more separate, long, AUG codon-containing ORFs. Such RNAs are rare in eukaryotes and so far have been described only for viruses (for review, see Kozak, 1986b; Wang *et al.*, 1987; Horvath *et al.*, 1990). Long ORFs on artificial dicistronic mRNAs in most cases inhibit downstream translation drastically (Kaufman *et al.*, 1987; Angenon *et al.*, 1989). However, exceptional reinitiation efficiencies of 20–100% have been described (Peabody and Berg, 1986; Peabody *et al.*, 1986; Sedman and Mertz, 1988; Sedman *et al.*, 1989b; Horvath *et al.*, 1990). Factors influencing the reinitiation frequency are largely unknown although the organization of the ORFs seems to be important (Dixon and Hohn, 1984; Liu *et al.*, 1984; Kozak, 1987; Thomas and Capecchi, 1986; Peabody and Berg, 1986; Peabody *et al.*, 1986). The reinitiation frequency increased with increasing length of the intercistronic region on a synthetic mRNA (Kozak, 1987). This result was explained by requirement for a certain timespan for recruitment of necessary initiation factors. It has been suggested that regulation of this required timespan controls GCN4 translation (Abastado *et al.*, 1991). In other systems, however, no equivalent distance requirement was observed (e.g. Peabody and Berg, 1986).

A particularly complex mRNA is the 35S RNA of cauliflower mosaic virus (CaMV). This RNA consists of a 600 nucleotide long leader sequence, containing several AUG codons, followed by seven to eight tightly arranged long ORFs encoding all of the viral proteins (Figure 1A; Mason *et al.*, 1987; Bonneville *et al.*, 1988). CaMV ORFs following a long upstream ORF on a polycistronic RNA are not translated efficiently either *in vitro* (Gordon *et al.*, 1988) or in plant protoplasts (Bonneville *et al.*, 1989; Gowda *et al.*, 1989). In protoplasts, expression of the downstream ORF could be activated by coexpression of the CaMV transactivator (TAV) protein encoded by the ORF VI (Bonneville *et al.*, 1989; Gowda *et al.*, 1989). Transactivation occurred post-transcriptionally and was specific for the downstream ORF (Bonneville *et al.*, 1989) but the mechanism by which ribosomes reach this ORF and initiate translation is still unknown.

In order to analyse potential *cis* requirements for

polycistronic translation and to determine the route by which ribosomes reach downstream ORFs on such mRNAs, we constructed a variety of polycistronic expression plasmids that contained a β-glucuronidase (GUS) ORF as an upstream long ORF and a chloramphenicol acetyltransferase (CAT) ORF as a downstream one. We found two sets of dicistronic plasmids and corresponding conditions that allowed translation of both of the cistrons. In the one case, to be described elsewhere, *cis*-acting sequences derived from the CaMV leader and host factor(s) are required. In the other case, described here, translation of the second ORF does not require specific *cis*-acting sequences but depends entirely on the CaMV-encoded transactivator.

Results

Dicistronic expression requires transactivation

To analyse the requirements for translation of a downstream ORF in plant protoplasts we constructed polycistronic expression units containing two reporter ORFs which are transcribed under the control of CaMV transcription signals. An upstream GUS ORF hinders normal downstream translation and also acts as an endogenous control for production of the polycistronic mRNA and the translation capacity of the protoplasts. Downstream translation is monitored by measuring expression of a CAT ORF. Most of our constructs contain only one potential CAT initiation codon; however, the ORFI–CAT fusions (pGC1 and its derivatives; Figures 1, 2 and 4), like ORFI itself, have two (Bonneville *et al.*, 1989); truncated CAT proteins originating from initiation at internal AUG codons are not active (Gordon *et al.*, 1991).

In most constructs, the region between the GUS and CAT reporter ORFs contains a third ORF that is opened by an AUG codon that overlaps the GUS UGA stop codon in an AUGA quadruplet. In the context of the bacterial chromosome the respective reading frame is supposed to encode a glucuronide permease (GP; Jefferson *et al.*, 1986). We took advantage of this GP ORF to construct CAT expression units in which the ORF arrangement between the GP′ ORF and the CAT ORF mimics original CaMV situations. This requires manipulation of the 3′ end of the GP′ ORF preceding the CAT ORF which—if performed directly with the GUS ORF—might cause alterations in GUS activity and abolish the convenient function of GUS expression as endogenous transfection control.

For plasmid pGC1 the resulting mRNA (Figure 1A) contains a leader consisting of the first 300 nucleotides of CaMV RNA including its first four small open reading frames (sORFs A, B, C and D′). The GUS ORF is opened by the start codon of sORF D. The two reporter ORFs are separated by 100 nucleotides derived from original bacterial sequences downstream of GUS and from a piece of CaMV RNA including the end of ORF VII (fused in frame to the GP ORF) and the start codon of ORF I (VII′, Figure 1A).

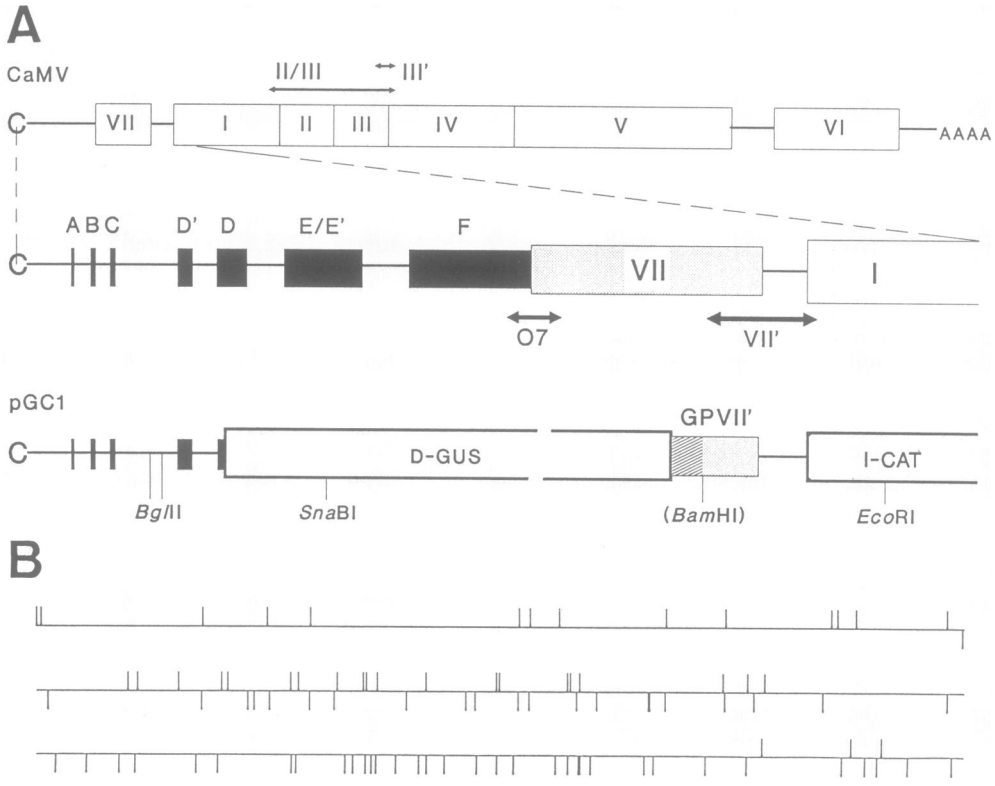

Fig. 1. (**A**) Regions of the CaMV 35S RNA used for construction of polycistronic reporter mRNAs. The top line shows a schematic map of the CaMV 35S RNA with the seven long open reading frames. The region containing the first 1000 nt starting from the cap site (C) is enlarged in the second line. Black boxes A–F indicate sORFs within the leader of the 35S RNA. The region from the cap to sORF D was used in the construction of expression plasmids as well as the regions denoted II/III, III′, O7 and VII′. The relevant region of the basal construct pGC1 is shown in the third line. The GUS reporter ORF is fused in frame to the start codon of sORF D (D-GUS) and the CAT ORF is fused to the start codon of ORF I (I-CAT). The ORF between GUS and CAT is derived partially from the coding sequence of the bacterial GP ORF and the CaMV ORF VII as indicated by differential shading. The location of restriction sites relevant for plasmid construction is indicated. The *Bam*HI site is shown in brackets because it is destroyed in pGC1. (**B**) Start and stop codon analysis of the GUS ORF. Graphical presentation is according to the 'Frames' programme of the 'GCG Sequence Analysis Software Package'. The lines represent the three reading phases, ticks above the lines the start codons and ticks below the lines the stop codons. The first line shows the phase of the GUS ORF (1809 nt).

The ORF I start codon becomes now the start codon of CAT. The ORF organization between the GPVII′ ORF and the I-CAT ORF is therefore identical to the organization of CaMV ORFs VII and I.

After transfection of pGC1 alone into protoplasts of *Orychophragmus violaceus*, a CaMV host plant, only GUS activity was detectable. However, upon co-transfection with pHELP7, a plasmid expressing the CaMV transactivator protein TAV (Bonneville *et al.*, 1989) CAT activity could also be observed. GUS activity was not significantly altered by pHELP7 (Figure 2). The quantity and quality of the pGC1 RNA was not changed by cotransfection of pHELP7 (not shown). This is in accordance with earlier results (Bonneville *et al.*, 1989) and demonstrates that pHELP7 acts on translation of the polycistronic mRNA. The level of GUS expression in the absence of the transactivator plasmid and the level of CAT expression in its presence were used as reference values (100%) in further experiments. A comparison with expression data from a similar plasmid with the reporter ORFs in the inverse order indicates that the downstream ORF is expressed upon transactivation at ~40% efficiency of the upstream ORF (not shown).

No special requirements for sequences between the reporter ORF

A series of pGC1 derivatives was made to determine if translation of the second ORF depends on specific sequences in the intercistronic region (Figure 2): plasmid pGC2 contains an intercistronic region derived from CaMV (III′, Figure 1A) resulting in a fusion of the GP′ ORF to the end of ORF III and CAT to the sixth codon of ORF IV. The 17 nt overlap of GPIII′ with IV−CAT corresponds to the natural overlap of CaMV ORFs III and IV. In pGC3 no CaMV derived sequences are present in the region between GUS and CAT ORF; the GP′ ORF in this construct terminates 130 nt downstream of the start codon of the CAT ORF. Frameshift mutations were introduced into the GP′ ORF of pGC3 either by filling-in the internal *Bam*HI site (pGC3.sh) or by inserting an oligonucleotide into it (pGC4). In the new phases, the GP′ ORF stops 40 and 10 nucleotides upstream of the CAT AUG start codon, respectively. In pGC4.sh the phase of GP′ was shifted back to restore the original overlap. In plasmid pGC5 the CAT ORF was directly fused to the GP′ start codon connecting GUS and CAT by the AUGA quadruplet, a configuration also found for all ORFs of the carnation etched ring virus (CERV), a close relative of CaMV (Hull *et al.*, 1986). In pGC6 the GUS ORF is truncated to eliminate the 3′ end of the GUS ORF including the AUGA quadruplet and the GP′ ORF (see also below).

Results reveal that for translation of the downstream CAT ORF neither an 'intercistronic' ORF (pGC5 and 6) nor CaMV specific sequences (pGC3.sh, pGC4, pGC5 and pGC6) are required, but also that none of the constructs allowed CAT expression in the absence of transactivator (Figure 2). The intercistronic ORF starting with the GP′ AUG, if present, is translated (pGC5) upon transactivation but this translation does not interfere with further downstream expression provided it either terminates upstream of the CAT ORF or overlaps it by a short distance, e.g. 17 nt (pGC2). However, a long overlap, i.e. of 130 nucleotides (pGC3 and pGC4.sh), did not allow CAT expression (Figure 2). This suggests that translation initiation at the CAT ORF is directly linked to active translation of the upstream ORF.

The GP start codon is probably part of a functional downstream translation initiation site in *Escherichia coli* (Jefferson *et al.*, 1986). This site may therefore contain

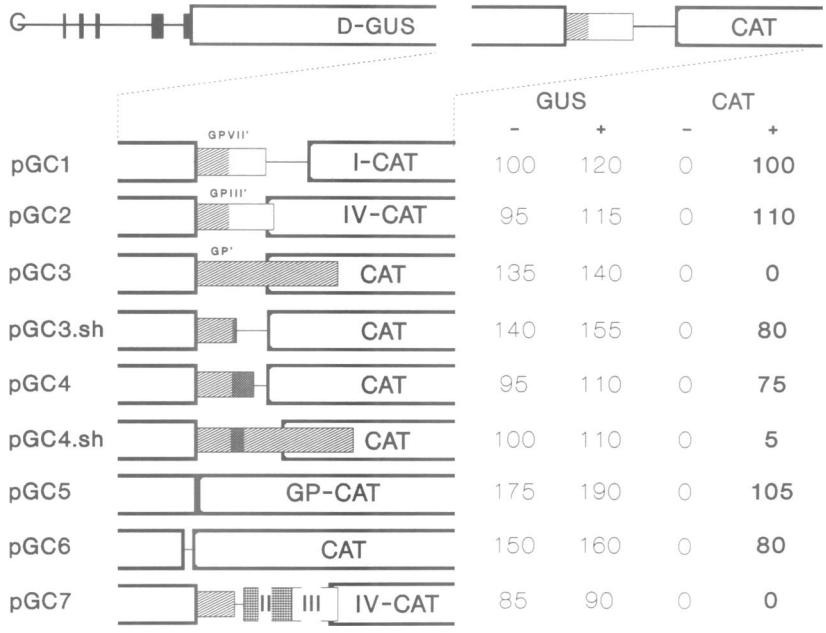

Fig. 2. Transient expression analysis of polycistronic mRNAs. The basal construct pGC1 is shown in the first line. Variants of the intercistronic region as described in the text and in Table IA are shown schematically. Plasmids were transfected into *O. violaceus* protoplasts and GUS and CAT activities were determined in the absence (−) or presence (+) of plasmid pHELP7 which expresses the CaMV ORF VI protein under control of the CaMV 35S promoter (Bonneville *et al.*, 1989). Values for reporter gene activity represent rounded averages of at least four independent transfection experiments. Values for non-transactivated expression varied by ±20% and values for transactivated expression by ±50% between different experiments. Values for GUS expression are presented relative to the non-transactivated expression of plasmid pGC1, values for CAT expression relative to transactivated expression from pGC1. Values below 2% were regarded as background activity and are shown as 0.

elements functioning in a special initiation mechanism. For instance, upstream of the GP start codon two motifs (CCGCAG and GCAGG) occur which could allow interaction of the mRNA with the conserved 3′ terminal stem–loop structure of prokaryotic 16S RNA and also eukaryotic 18S RNA. A similar motif has been implicated in 80S ribosome formation on some eukaryotic mRNAs (Azad and Deacon, 1980) and is found on GCN4 (Miller and Hinnebusch, 1989) and CaMV RNAs (Fütterer *et al.*, 1990b) in regions of potential translation reinitiation. It might be a good candidate for a reinitiation enhancing element. Removal of this element (pGC6, Figure 2), however, did not significantly reduce expression of the CAT ORF in the presence of transactivator, showing that the putative ribosome interactive sequence is not required.

Intercistronic regions containing more than one (s)ORF interfered with transactivated downstream translation. One example is construct pGC7 (Figure 2) in which CaMV ORFs II and III (region II/III, Figure 1A) are present in the intercistronic region in front of an ORF IV–CAT fusion.

Efficiency of mammalian translation reinitiation on a mRNA that contains an upstream minicistron was found to increase with the length of the intercistronic region. A distance of 79 nucleotides between translation stop and reinitiation site was sufficient for nearly 100% reinitiation (Kozak, 1987). By varying the intercistronic region of our pGC constructs we also observed an increase of non-transactivated translation of the downstream ORF; however, much larger distances were required (plasmid series pGC8, Figure 3). An intercistronic distance of 250 nt did not allow any significant CAT translation in the absence of transactivator, while distances of 400 and 690 nt allowed translation levels of 10 and 20%, respectively. In contrast, the levels of transactivated translation of the downstream ORF are much higher (i.e. 100%) and are not significantly affected by increasing the distance between the GUS and CAT ORFs (Figure 3).

Requirements for the leader

Omission of the leader sequence of pGC1 had little effect on GUS expression but reduced CAT expression drastically (pGC1.d9, Figure 4), showing that this part of the 35S RNA leader sequence contains an element that specifically enhances transactivated CAT expression. Deletion of the first 60 nucleotides of this region (defined as S1 by Fütterer *et al.*, 1990a) led to a decrease of ~2-fold in both reporter activities confirming earlier results (pGC1.δS1, Figure 3; Fütterer *et al.*, 1990a). Sequential deletion of the four sORFs (A, B, C and D′) in pGC1 yielded plasmids pGC1.d1, .d2, .d5 and .d9 (Figure 4). Translation of the second reporter ORF (CAT) was strongly reduced only when all four sORFs were deleted. In contrast, translation of the first reporter ORF (GUS) increased 2- to 3-fold following deletion of sORFs A, B, C and D′ (Figure 4), revealing the small inhibitory effect of these sORFs on direct downstream translation (Fütterer *et al.*, 1990a). The importance of an upstream sORF rather than another specific leader sequence element for downstream gene expression was also indicated by the effect of other sequences used as leaders. A sequence containing a sORF derived from the 5′ end of CaMV ORF VII (region O7, Figure 1A) stimulated transactivated CAT expression to a much higher level than a slightly shorter segment lacking the start codon (pGC1.O7 versus pGC1.N7, Figure 4). The reading frame opened by the ORF VII start

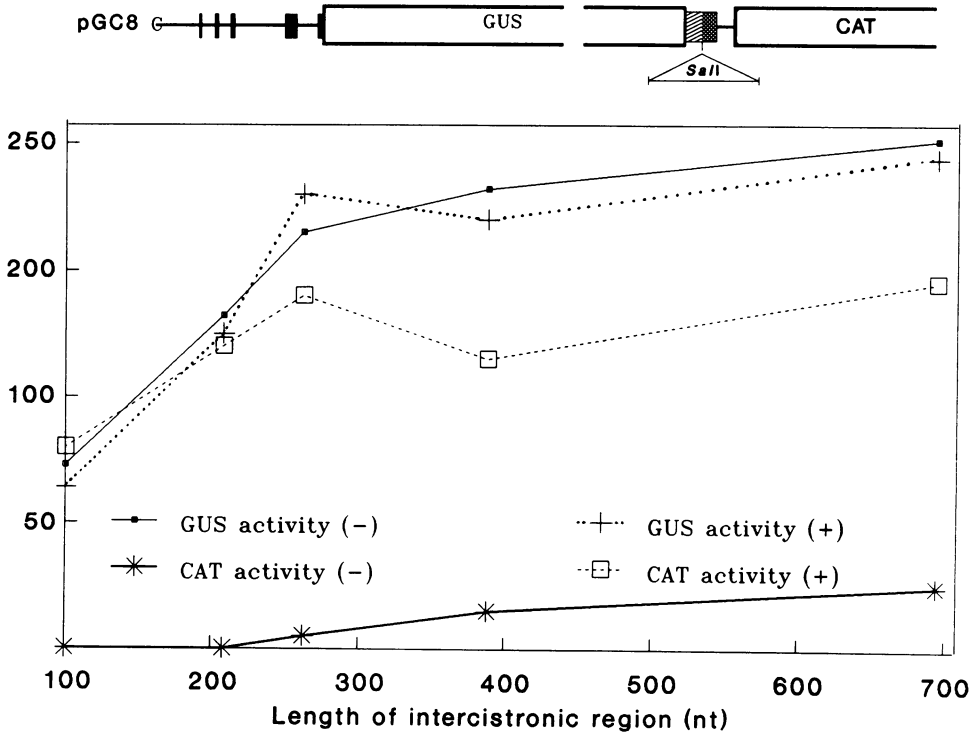

Fig. 3. Effect of the length of the intercistronic region on downstream translation. The intercistronic region of pGC8 was elongated by insertion of multiple copies of an 18 bp oligonucleotide into the *SalI* site (indicated in the schematic presentation of pGC8 in the top line). GUS and CAT activities in the presence (+) or absence (−) of transactivator are plotted against the length of the intercistronic region. Values are mean values taken from three series of experiments, in which also transactivator-independent CAT activity was discernible. They are presented relative to expression from plasmid pGC1 in the same series.

codon in pGC1.O7 overlaps the GUS ORF by 12 codons and drastically inhibits GUS expression. (In this case, low but reproducible transactivatability of GUS expression was observed.) A 278 bp piece of calf thymus DNA containing a 49 codon ORF (pGC1.OC) and a synthetic sequence containing a 4 codon ORF (pGC4.OS versus pGC4.NS with a point mutation destroying the sORF's ATG codon) also allowed transactivated CAT expression (Figure 4).

We conclude that an upstream sORF is indeed important for transactivated CAT expression in the dicistronic GUS−CAT constructs presented. This sORF can either be separated from, or overlap with the following reporter ORF. The variety of apparently functional upstream sORFs suggests that their translation *per se* is more important than the translation product; however, the different expression levels that result from the different sORFs suggest that the composition of codons and surrounding sequences also have an influence.

Ribosomes reach the downstream ORF by migration through the upstream sequences

To investigate the route by which ribosomes might reach the downstream ORF we inserted high energy stem−loop structures into pGC plasmids. Depending on their location, such stable stem−loop structures are thought to inhibit either binding at the capped 5′ end of mRNA or scanning of 40S ribosomes (Pelletier and Sonenberg, 1985; Chevrier *et al.*, 1988; Baim and Sherman, 1988; Kozak, 1989b).

When a stem structure ($\Delta G_{37°}$ = −29.7 kcal/mol calculated according to Freier *et al.*, 1986) was introduced directly at the cap site (pGC1.δS1.st, Figure 5) or when a more stable stem ($\Delta G_{37°}$ = −43.9 kcal/mol), previously shown to strongly inhibit translation of a downstream ORF (Kozak, 1989b), was introduced further downstream (pGC1.st), expression of both reporter ORFs was strongly reduced (Figure 5). In a similar construct (pLGC, described as pLC20-DG by Fütterer *et al.*, 1990a) where the downstream CAT ORF is translated in a scanning independent mechanism (J.Fütterer and T.Hohn, manuscript in preparation), the same stem structure acted selectively on the upstream GUS ORF (pLGC.st1, Figure 5), indicating that inhibition is a translational effect and not due to transcription or RNA stability. Insertion of the −43.9 kcal/mol stem downstream of the GUS ORF (pGC4.st) had no influence on GUS activity but reduced CAT activity 6-fold (Figure 5). These results are consistent with the scanning model, indicating that ribosomes translating the CAT ORF reached the initiation codon by migration from the RNA 5′ end.

Discussion

A variety of unusual mechanisms are involved in translation of the polycistronic CaMV 35S RNA: (i) The translational block of expression of an ORF downstream of the leader sequence (Baughman and Howell, 1988; Fütterer *et al.*,

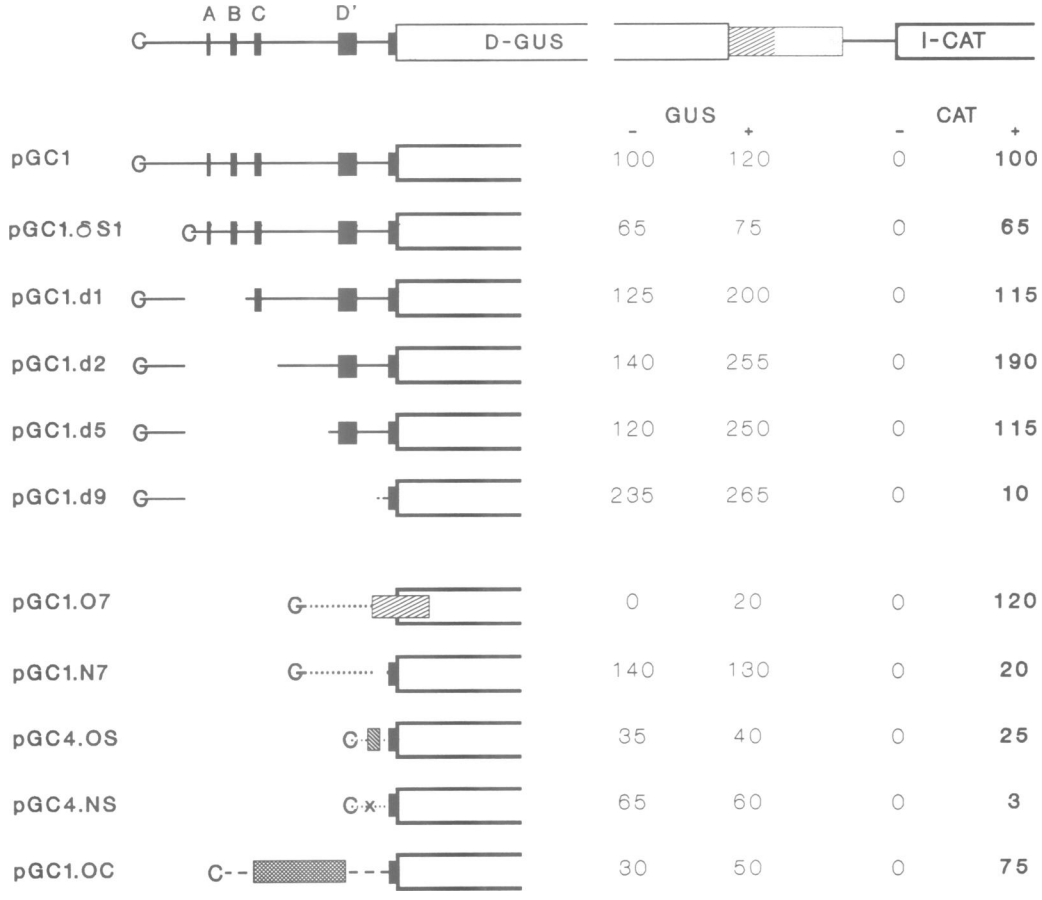

Fig. 4. Effect of the leader. The leader region of plasmid pGC1 (top line) or pGC4 (Figure 1A) was analysed by deletion mutagenesis or by replacement with other sequences. Deletions are indicated by gaps, replacement by differentially dotted or stippled lines. ORFs derived from the CaMV leader are shown as black boxes, ORFs from other sequences by different shading. × denotes a point mutation in the start codon of the leader sORF as the sole difference between plasmids pGC4.OS and pGC4.NS. Constructs are described in the text and in Table I. Values for GUS and CAT expression are depicted as in Figure 2.

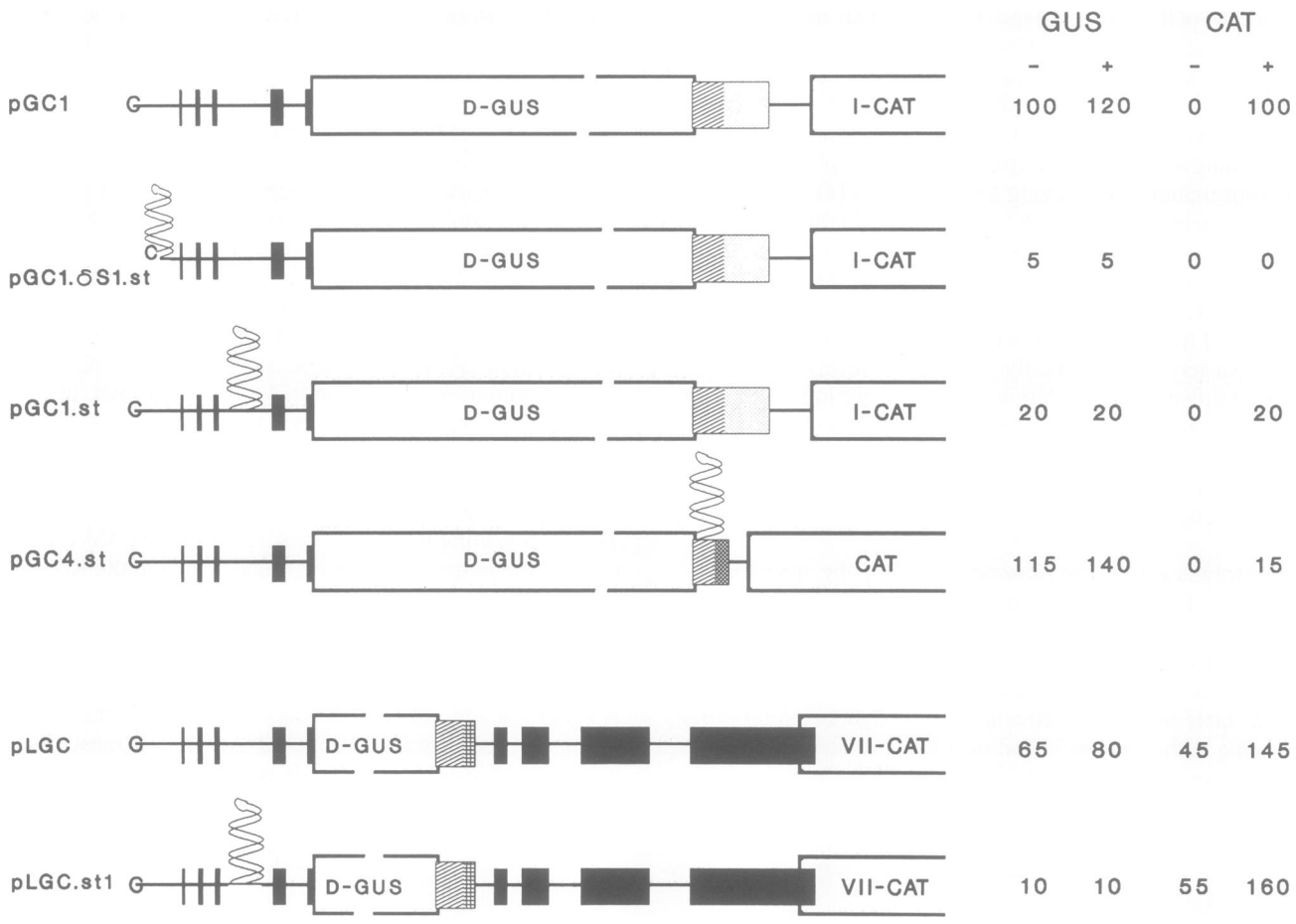

Fig. 5. Effect of additional secondary structure. Sequences supposed to form stable stem structures were inserted at the positions indicated by a double helix. Sequences are shown in Table I. GUS and CAT expression were monitored and are represented as in Figure 1A.

1989) is partially alleviated in host-plant protoplasts by a mechanism different from normal scanning (Fütterer *et al.*, 1990a; J.Fütterer and T.Hohn, manuscript in preparation); this mechanism depends on the presence of sequences from the CaMV 35S RNA leader and can be used to express simultaneously two reporter ORFs from one mRNA (Fütterer *et al.*, 1990a; see also pLGC, Figure 5). (ii) Translation of further downstream ORFs on the 35S RNA requires transactivation by the CaMV encoded TAV protein (Bonneville *et al.*, 1989; Gowda *et al.*, 1989).

To investigate the latter mechanisms in detail, we have constructed a series of polycistronic expression units and analysed the expression pattern of two consecutive reporter ORFs. The upstream GUS reporter ORF present either as the first ORF or preceded only by sORFs was expressed from all constructs constitutively, although the expression level varied slightly for different leader sequences and also for different downstream sequences, probably as a consequence of different inhibitory effects of leader sequences or long-range secondary structures. In contrast, with the exception of constructs containing extensively elongated intercistronic regions (Figure 3) the downstream ORF was not expressed unless TAV was present (Figure 2). As reported previously (Bonneville *et al.*, 1989) transactivation was specific for the downstream ORF. In an attempt to identify potential *cis*-acting signals that could be indicative of unusual translation mechanisms (see Introduction), we modified and finally replaced the CaMV derived sequences present in the original construct pGC1. It turned out that specific *cis* signals were not required; the salient characteristics of sequences allowing transactivation of downstream translation were a sORF in the leader and an ORF arrangement avoiding long overlaps. One additional ORF in the intercistronic region was permitted, provided that it terminated upstream or not very far downstream (e.g. 17 nt in pGC2, Figure 2) of the initiation codon of the following reporter ORF. It is noteworthy that an ORF arrangement in which the two ORFs are linked by an ATGA stop/start quadruplet (pGC5, Figure 2) also requires transactivation for downstream ORF translation. Such an arrangement is common for the ORFs of the caulimovirus CERV (Hull *et al.*, 1986) and might suggest a special translation mechanism. A tight arrangement of stop and start sites is observed in many translationally coupled prokaryotic operons where it allows reinitiation even in the absence of a functional Shine–Dalgarno sequence (Cone and Steege, 1985; Spanjaard and van Duin, 1989; Steege and Ivey-Hoyle, 1990). For prokaryotes no particular *trans*-acting factors involved in reinitiation have been identified so far but such factors may nevertheless exist and may also be present in chloroplasts where polycistronic translation occurs (Barkan, 1988). Redirection of such factors from action in the chloroplast to action in the cytoplasm could be a possible function of the CaMV TAV protein.

The positive influence of leader sORFs on translation of a far downstream reporter ORF in our constructs is

surprising and distinguishes our system from the described prokaryotic cases. The variety of functional sORFs, derived either from the leader of the 35S RNA or from unrelated sequences, suggests that the process of their translation is more important than the sORF's primary sequence. We note that the requirement for a conventional sORF is not absolute. Transactivation to a low level of CAT expression can also be observed in the absence of sORFs (plasmids pGC1.d9 and pGC1.N7, Figure 4).

The apparent lack of sequence specificity, the negative effects on downstream translation of a long overlap between upstream and downstream ORFs and the positive effect of a far upstream sequence all suggest that the downstream ORF is not translated by ribosomes that bind to the RNA internally. Rather it seems that the ribosomes which finally translate the downstream reporter ORF have scanned from the cap site and may also have translated upstream ORFs. Analysis of mutants with insertions forming stable stem structures at various positions of the mRNA confirm these conclusions (Figure 5): a strong inhibitory effect of a stem structure that includes the first nucleotides of the mRNA suggests ribosome entry at the cap site whereas reduction of all downstream translation by RNA internal stem structures is indicative of ribosome scanning. The effects of RNA stem−loop structures on ribosome scanning in plant protoplasts are quantatively comparable to those in mammalian cells (Kozak, 1989b). The observed inhibition acts only in *cis* (not shown) and is therefore probably not due to induction of a double-stranded RNA-induced kinase inactivating the function of eIF2. Such a mechanism was suggested for the action of the HIV-1 TAR region on internal initiation of translation in constructs combining TAR with the poliovirus internal ribosome entry site (Edery *et al.*, 1989). The observation that a long overlap between the intercistronic sORF and the downstream reporter ORF inhibits translation of the latter demonstrates a link between translation of the upstream ORF and initiation at the downstream one. This can for instance be explained by the assumption that the very same ribosome (or at least its 40S subunit) that translated the upstream ORF also translates the downstream one, complying with the modified scanning hypothesis (Kozak, 1989a) and that this ribosome cannot scan backwards for long distances. However, alternative explanations are conceivable:

(i) Ribosomes might reach the downstream initiation site without translating the upstream ORF. The inhibition of CAT expression in pGC3 and pGC4.sh (Figure 2) could then be explained by interference with 80S ribosome formation at the downstream ORF by the 80S ribosome translating the overlapping part of the upstream ORF. Such an inhibition should, however, be independent of the extent of the ORF overlap and therefore this model is not consistent with the toleration of short overlaps (plasmids pGC2 and pGC5). Interference between 80S ribosomes translating the two phases of the overlapping ORF region at different rates could also reduce expression of either one of the two ORFs as was suggested for translation of a bifunctional mRNA of a mammalian reovirus (Fajardo and Shatkin, 1990). However, a complete inhibition by such a mechanism appears to be unlikely.

(ii) 43S ribosomal subunits could be transported through the upstream ORF by physical coupling to a translating 80S ribosome. Coupling would prevent initiation until the 80S ribosome dissociates upon termination. This 'train' model would explain our data equally as well as the reinitiation model but has the advantage that it obviates the mysterious step of regaining initiation factors: the apparently reinitiating ribosome would still contain the original set of initiation factors when reaching the downstream initiation site. The model would be particularly attractive for cases of short ORF overlaps. In the reinitiation model, back scanning of ribosomal subunits from the termination site of the first ORF to the initiation site of the second must be postulated (Thomas and Capecchi, 1986; Peabody and Berg, 1986); in a 'train' model the 43S subunits would still be located upstream of the termination site when the leading 80S ribosome dissociates and could reach the start codon of the second ORF by normal forward scanning.

The reinitiation model and the train model are difficult to distinguish experimentally but the models offer different possibilities for the action of the CaMV translation transactivator. Our results indicate that TAV induces a change in the translation machinery during passage of ribosomes through the region upstream of the first reporter ORF, especially when this region contains a sORF. These observations resemble those obtained with the yeast GCN4 mRNA. Also in this case, translation of an upstream sORF is required to reinitiate translation downstream of an additional, otherwise inhibiting ORF; the upstream sORF can be derived from heterologous sequences (Williams *et al.*, 1988) and GCN4 translation is controlled by a variety of *trans*-acting factors (for review, see Hinnebusch, 1988). In the case of GCN4 the inhibiting ORF is a sORF in the 3′ portion of the leader sequence (uORF 3 and/or 4); in our case it is the first reporter ORF. It has been suggested that GCN4 translation is determined by precise regulation of the timespan required for ribosomes to regain initiation competence after translation of the first upstream sORF; depending on this timespan, ribosomes would reinitiate either at the AUG of the uORF 4 or of the GCN4 ORF (Abastado *et al.*, 1991). In our case this mechanism is unlikely, since the very long region between the termination codon of the sORF and the initiation codon of the CAT ORF contains a total number of 36 internal AUG codons either in the GUS reading phase or opening sORFs in the other two phases (Figure 1B). Furthermore, in contrast to GCN4 (Abastado *et al.*, 1991), downstream translation appears to be linked to upstream translation. It would nevertheless be interesting to compare the function of the CaMV transactivator (TAV) with that of the yeast GCN and GCD proteins required for regulation of translation of the GCN4 mRNA.

Some possible functions of TAV could be to enhance the formation or stability of a ribosome train or to generate reinitiation competent ribosomes either by functionally replacing or modulating initiation factors or by inhibiting their dissociation after the first initiation event. The latter mechanism might explain the different effects of an upstream sORF and long ORF. It has been speculated that loss of the initiation factors following initiation may be a slow process and may not be complete after translation of a sORF (Kozak, 1987). It therefore appears possible that the transactivator could fix such factors with higher efficiency to ribosomes that have translated a short ORF than to those that have translated a longer one.

Since TAV activates translation of a downstream ORF unspecifically, it may also act on some cellular mRNAs, e.g.

those that contain upstream sORFs. It remains to be investigated whether such an activity is responsible for the phenotypic effects of TAV expression in transgenic plants (Baughman *et al.*, 1988; Takahashi *et al.*, 1989; Goldberg *et al.*, 1991). The translation of the CaMV 35S RNA requires as a first step that ribosomes overcome the potential hindrance by the leader sequence; although this process is also affected by TAV, albeit slightly (Fütterer *et al.*, 1990b), it differs from the scanning related mechanism described here (Fütterer *et al.*, 1990a). It appears possible that for translation of other ORFs on the CaMV 35S RNA, specific combinations of these different mechanisms or even other mechanisms are active. This is suggested by the lack of CAT expression from pGC7 (Figure 2). Since the corresponding construct lacking the GUS ORF allowed TAV-induced CAT expression (Fütterer *et al.*, 1990b), either the GUS ORF might have a negative polar effect, or the CaMV sequences replaced by the GUS ORF might have a positive effect, on far downstream translation. Neither of these alternatives would have been predicted by our analysis of *cis*-requirements for transactivation. However, both would be in line with the observed polar effects of payload genes cloned into the CaMV genome and with the apparent requirement for a tight genome arrangement between the CaMV ORFs I–IV for CaMV genome stability (reviewed in Fütterer *et al.*, 1990c).

Materials and methods

Construction of plasmids: modification of the intercistronic region

Plasmid pGC was constructed by insertion of a GUS reading frame as an *Nco*I (filled in)–*Bam*HI fragment derived from plasmid RAJ275 (Jefferson, 1987) into plasmid pDW2 (Pietrzak *et al.*, 1986) opened by *Sma*I–*Bam*HI digestion. The ORF I upstream region (pGC1) was introduced into pGC by exchanging the small *Bam*HI–*Eco*RI fragment for the small *Bgl*II–*Eco*RI fragment from plasmid pCAT2 (Bonneville *et al.*, 1989). The ORF IV upstream region present in pGC2 was introduced similarly by inserting a *Bam*HI–*Eco*RI fragment from a plasmid (pCaMVIV-CAT, unpublished) that contains a CAT ORF fused to the sixth codon of the CaMV ORF IV by oligonucleotide directed mutagenesis (Zoller and Smith, 1983). In pGC7 an analogous *Bgl*II–*Eco*RI fragment derived from pCaMVIV-CAT was introduced.

Table IA. Sequences of intercistronic regions

GC (=GC3)
AAACCGCAGCAGGAGGCAAACAATG.AAT.CAA.CAA.CTC.TCC.TGG.CGC.ACC.ATC.GTC.GGC.TAC.AGC.CTC.GGG.AAT.
GUS ⟶ | GP′
*Bam*HI *Xba*I
TGG.GGA.TCC.TCT.AGA.GAG.CTT.CGA.CGA.GA.GAT.TTT.CAG.GAG.CTA.AGG.AAG.CTA.AAA.TGG.AG...
⟶ CAT

GC1
....GGA.TCT.AAA.$(NNN^a)_{24}$.TAA.ATCTCTCTGAGAATAGTACTCTAACGAGTATCCACAGGAAAAATAATCTTCTGTGTTGAG
GP′–VII ⟶|
ATGGGAT...
⟶ I-CAT

GC2
....GGA.TCC.TCT.GGA.$(NNN^b)_{12}$.GCA.GGA.TGG.CGG.AAT.CAA.TTT.TAG.AG...
⟶ IV–CAT GP′ ⟶|

*Bam*HI *Xba*I
GC4
....GGA.TCC.ACA.GGA.AAA.ATA.ACC.TTC.TGT.CTA.GAG.AGC.$(NNN^c)_6$.AGC.TAA.GGAAGCTAAAATGGAG...
GP′ ⟶| ⟶ CAT

*Xho*I
CG5
AAACCGCCCGGATCCCCGCAGCAGGAGGCAAACAATGAATCAACTCGAG...
GUS′ ⟶ | GP′–CAT

GC6
AAACCGCCCGGATCTTAGATCTACCATGGAACTCGAG...
GUS′ ⟶| ⟶CAT

GC7
...GGA.TCT.TTT.$(NNN^d)_{11}$.AAA.TGA.GC$(NNN^d)_{156}$GGCTGAAATG$(NNN^d)_{122}$GGATGGCC....(pGC2)
GP′–I ⟶ | II II ⟶ | ⟶ III ⟶IV–CAT

*Bam*HI *Sal*I *Xba*I
GC8
...GGA.TCC.TTA.GTC.(GAG.AAA.ACC.AAG.GAA.GTC.)$_n$GAC.TTC.CTC.TAG.A...(pGC)
[n = 0, 6, 9, 16, 32] GP′ ⟶|

GC3.sh
...GGA.TCG.ATC.CTC.TAG....(pGC)
GP′ ⟶|

GC4.sh
...GGA.TCG.ATC.CAC.AGG.AAA.AAT.AAC.CTT.CTG.TCT.AGA....(pGC)

GC: The intercistronic sequence is shown, starting close to the end of the GUS ORF, covering the GP′ ORF whose reading frame is indicated, and extending beyond the start codon of the CAT ORF. A sequence motif close to the end of the GUS ORF which could interact with the 18S rRNA is marked with a wavy line. Translation start and stop sites are indicated by arrows. The position of relevant restriction sites is indicated. In the following lines, derivatives of the GC intercistronic region are shown similarly, starting either at the same position (GC5 and GC6) or at the *Bam*HI site.

Sequences denoted as N are derived from [a]region VII′, [b]region III′, [c]region GC and [d] region II/III (Figure 1A).

Table IB. Sequences of leader regions.

GC1 ACACGCTGAAATCA N^{e}_{326}AAGATGCTCGACATGG...
⟶D−GUS

*Nde*I
GC1.δS1 ACACGCTGAAATCA−−TATGTGTGAG...
⟶sORF A

GC1.d1 to GC1.d5 see Fütterer *et al.* (1990a)

GC1.d9 ACACGC N^{f}_{36}CTCGATCCGGGGGAAAAGATGCTC...
⟶GUS

*Bam*HI
GC1.O7 ACAGGGTACCCCGAT N^{g}_{50}CATG. AAT.CGG.TTT.AAA.AAC.CAA.ACC.GGA.TCC...GGG.GGA.AAA.GAT.G CT.C...
⟶ VII′ ⟶GUS

GC1.N7 ACAGGGTACCCCGAT N^{h}_{39}GGATCCGGGGGAAAAGATGCTC...
⟶GUS

GC4.OS ACAGGGTACCCGGGCTCGAGAAAACCATG.GAA.GTC.GAC.TAA.GGATCCGGGGGAAAAGATGCTC...
⟶ sORF ⟶| ⟶GUS

GC4.NS ....(pGC4.OS).....AAAACCAAGGAA...(pGC4.OS)...

GC1.OC ACAGGGTACCCGGGGATC N^{i}_{89}ATG$(NNN^{i})_{48}$TGA N^{i}_{37}GGATCCGGGGGAAAAGATGCTC...
⟶ORF ⟶| ⟶GUS

The leader regions are shown starting at the transcription start site and extending to the start codon of the GUS ORF. Sequences denoted as N are derived either from the orginal leader sequence of the CaMV 35S RNA ([e]nucleotides +15 to +341; [f]nucleotides +15 to +51) or from region O7 (Figure 1A) [g]nucleotides +561 to +610 or [h]nucleotides +561 to +600 or [i]from a *Sau*3A fragment of calf thymus DNA. Dashes in GC·δS1 mark the position of the deletion in the CaMV 35S RNA leader sequence. The point mutation that distinguishes GC.OS and GC.NS is underlined.

Table IC. Sequences of stem insertions

*Nde*I
GC1.δS1.st ACAGGGTACCCGGGGATCCTT AAGGATCCTCGGGTACTCTGTAGATCTCATATG...
⟶⟵ ⟶sORF A

GC1.4.st ....GATCGGGCGCGTGGTGGCGGCTG CAGCCGCCACCACGCGCCCGATC...
⟶⟵

The region supposed to form stable stem structures is shown. Base pairing segments are underlined by arrows.

The intercistronic region was also varied by insertion of oligonucleotides between the *Bam*HI and *Xba*I sites (pGC4 and pGC8[n=0]), filling in the *Bam*HI site of plasmid pGC3 (resulting in pGC3.sh) and pGC4 (resulting in pGC4.sh) and by Bal31 deletion starting at the *Bam*HI site with subsequent insertion of a *Bam*HI linker (CCGGATCCGG). In one of the deletion mutants, pGCBal1, the last 20 nt of the GUS ORF (including the stop codon) were removed. The large pGCBal1 *Bam*HI−*Eco*RI fragment was recombined with the small *Xho*I−*Eco*RI fragment of plasmid pLC15 (Fütterer *et al.*, 1988) and oligonucleotides compatible with *Bam*HI and *Xho*I restriction sites (yielding plasmids pGC5 and pGC6). To create plasmid series pGC8, pGC8[n=0] was opened at its *Sal*I site and *Xho*I−*Sal*I fragments of different lengths created by reiteration of the respective fragment in the leader of pGC5.NS were inserted.

Sequences of the different intercistronic regions are shown in Table IA.

Construction of plasmids: modification of the leader region

Combination of various dicistronic expression units containing differing intercistronic regions and differing leader sequences was achieved by exchanging the small *Sna*BI−*Eco*RI fragments. The CaMV leader region and the sORF D−GUS fusion originate from plasmid pLC20-DG (Fütterer *et al.*, 1990a). The small *Bgl*II−*Eco*RI fragment of pGC1 was inserted between the corresponding sites of the leader deletion mutants pLC20-d1, -d2 and -d5 (Fütterer *et al.*, 1990a) yielding plasmids pGC1.d1, d2 and d5. The same fragment was also inserted into the large *Bgl*II−*Eco*RI fragment of pCAT2 (Bonneville *et al.*, 1989) and the resulting plasmid was opened at the *Bgl*II site, digested with Bal31 and religated in the presence of a *Bam*HI linker resulting finally in plasmids pGC1.O7 and N7. The small *Bam*HI−*Eco*RI fragment of pGC1.O7 was cloned between the *Bgl*II and *Eco*RI sites of pGC1-d5 (yielding pGC1.d9) or between the *Bam*HI and *Eco*RI sites of plasmid R2-CAT (Sanfaçon and Hohn, 1990) (yielding pGC1.OC). Plasmids pGC4.OS and pGC4.NS were obtained by inserting oligonucleotides between the *Kpn*I and *Bam*HI sites of pGC1.O7 and exchange of the small *Sna*BI−*Eco*RI fragment between the resulting plasmid and pGC4.

Plasmid pGC1.δS1 was produced by oligonucleotide directed mutagenesis of pGC1. Sequences of the different leader regions are shown in Table IB.

Construction of plasmids: insertion of stem structures

For the introduction of a 5′ terminal stem structure, the large *Xma*I−*Eco*RI fragment of pGC2.O7 was ligated simultaneously to an oligonucleotide with *Xma*I and *Nde*I compatible ends and to the small *Nde*I−*Eco*RI fragment of pGC1.δS1 (yielding plasmid pGC2.δS1.st).

The sequences forming RNA internal stems were introduced by inserting an oligonucleotide into the *Bgl*II site of pGC1 (yielding pGC1.st) or pLGC (described as pLC20-DGus in Fütterer *et al.*, 1990a; yielding pLGC.st1) or into the *Bam*HI site of pGC4 (yielding pGC4.st).

Sequences of relevant regions of plasmid constructs are shown in Table IC.

All steps in the cloning procedures were performed as recommended by the suppliers of the respective enzymes or as described in Maniatis *et al.* (1982). Important regions of plasmids were analysed by sequencing of

double-stranded plasmid DNA using a Sequenase kit (US Biochemical) as recommended by the manufacturer.

The 35S RNA leader sequence used in this study was derived from CaMV strain CM4-184 (Figure 1A; Dixon *et al.*, 1986). The numbering refers to the transcription start site as +1 which is located at genome position 7014 [by analogy with the transcription start site of the Cabb-S strain (Guilley *et al.*, 1982)].

Suspension cultures, protoplast preparation and transient expression assays

Culture conditions for the *Orychophragmus violaceus* suspension culture, protoplast preparation and the transfection assay have been described previously (Fütterer *et al.*, 1989, 1990a). Routinely, 2×10^6 protoplasts were transfected by electroporation (Fromm *et al.*, 1985) with 10–15 μg of circular plasmid DNA as described (Fütterer *et al.*, 1988, 1989). Transactivator was provided by cotransfecting 5 μg plasmid pHELP7 (Bonneville *et al.*, 1989). After overnight incubation (12–24 h) a protein extract was prepared and the CAT or GUS activity was determined (Gorman *et al.*, 1982; Jefferson *et al.*, 1986; Bonneville *et al.*, 1989). For CAT quantification, assay conditions were chosen such that not more than 30% of the radioactive chloramphenicol was converted to the acetylated forms.

The response of plant protoplasts to transfection with a given construct varied by about ±20% of the respective induced reporter enzyme activity levels in the absence of transactivator. In its presence, variation was ±50% when values obtained over several years were compared. Especially the response of the GUS ORF to transactivation varied in a way that seems to be linked to the season. Most constructs were tested more than ten times and none less than four times. Since the absolute expression levels varied with the batch of protoplasts used, whereas the relative levels remained constant (within the range mentioned), we present the GUS and CAT expression data relative to the reference plasmid pGC1 in the absence (GUS) or the presence (CAT) of the transactivator.

Acknowledgements

We are grateful to W.Zürcher and F.Fischer for synthesis of oligonucleotides and to Drs H.Rothnie, B.Hohn and S.-J.Morley for critically reading the manuscript. We especially thank Hanny Schmid-Grob, Gundula Pehling and Mathias Müller for expert technical assistance.

References

Abastado,J.-P., Miller,P.F., Jackson,B.M. and Hinnebusch,A.G. (1991) *Mol. Cell. Biol.*, **11**, 486–496.

Angenon,G., Uotila,J., Kurkela,S.A., Teeri,T.H., Botterman,J., van Montagu,M. and Depicker,A. (1989) *Mol. Cell. Biol.*, **9**, 5676–5684.

Azad,A.A. and Deacon,N.J. (1980) *Nucleic Acids Res.*, **8**, 4365–4376.

Baim,S. and Sherman,F. (1988) *Mol. Cell. Biol.*, **8**, 1591–1601.

Barkan,A. (1988) *EMBO J.*, **7**, 2637–2644.

Baughman,G. and Howell,S.H. (1988) *Virology*, **167**, 125–135.

Baughman,G., Jacobs,J.D. and Howell,S.H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 733–737.

Bonneville,J.M., Hohn,T. and Pfeiffer,P. (1988) In Domingo,E., Holland,J.J. and Ahlquist,P. (eds), *RNA Genetics*. CRC Press. Vol. 2, pp. 23–42.

Bonneville,J.M., Sanfaçon,H., Fütterer,J. and Hohn,T. (1989) *Cell*, **59**, 1135–1143.

Chevrier,D., Vézina,C., Bastille,J., Linhard,C., Sonenberg,N. and Boileau,G. (1988) *J. Biol. Chem.*, **263**, 902–910.

Cone,K.C. and Steege,D.A. (1985) *J. Mol. Biol.*, **186**, 733–742.

Curran,J. and Kolakofsky,D. (1988) *EMBO J.*, **7**, 2869–2874.

Darlix,J.-L., Zuker,M. and Spahr,P.-F. (1982) *Nucleic Acids Res.*, **10**, 5183–5196.

Dixon,L. and Hohn,T. (1984) *EMBO J.*, **3**, 2731–2736.

Dixon,L., Nyffenegger,T., Delley,D., Martinez-Izquierdo,J.A. and Hohn,T. (1986) *Virology*, **150**, 463–468.

Edery,I.R., Petryshyn,R. and Sonenberg,N. (1989) *Cell*, **56**, 303–312.

Fajardo,J.E. and Shatkin,A.J. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 328–332.

Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Nelson,T. and Turner,D.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9373–9377.

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

Fütterer,J., Gordon,K., Bonneville,J.M., Sanfaçon,H., Pisan,B., Penswick,J. and Hohn,T. (1988) *Nucleic Acids Res.*, **16**, 8377–8390.

Fütterer,J., Gordon,K., Pfeiffer,P., Sanfaçon,H., Pisan,B., Bonneville,J.M. and Hohn,T. (1989) *Virus Genes*, **3**, 45–55.

Fütterer,J., Gordon,K., Sanfaçon,H., Bonneville,J.M. and Hohn,T. (1990a) *EMBO J.*, **9**, 1697–1707.

Fütterer,J., Bonneville,J.M., Gordon,K., DeTapia,M., Karlsson,S. and Hohn,T. (1990b) In McCarthy,J.E.G. and Tuite,M.F. (eds), *Post-transcriptional Control of Gene Expression*. NATO ASI series, volume H49, Springer, Heidelberg, pp. 349–357.

Fütterer,J., Bonneville,J.M. and Hohn,T. (1990c) *Physiol. Planta.*, **79**, 154–157.

Goldberg,K.-B., Kiernan,J.M. and Shepherd,R.J. (1991) *Mol. Plant–Microbe Interact.*, in press.

Gordon,K., Pfeiffer,P., Fütterer,J. and Hohn,T. (1988) *EMBO J.*, **7**, 309–317.

Gordon,K., Fütterer,J. and Hohn,T. (1991) *Plant J.*, in press.

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

Gowda,S., Wu,F.S., Scholthof,H.B. and Shepherd,R.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9203–9207.

Guilley,H., Dudley,R.K., Jonard,G., Balazs,E. and Richards,K.E. (1982) *Cell*, **30**, 763–773.

Herman,R.C. (1989) *Trends Biochem. Sci.*, **14**, 219–222.

Hinnebusch,A.G. (1988) *Trends Genet.*, **4**, 169–174.

Horvath,C.M., Williams,M.A. and Lamb,R.A. (1990) *EMBO J.*, **9**, 2639–2647.

Hull,R. (1984) *Plant Mol. Biol.*, **3**, 121–125.

Hull,R., Sadler,J. and Longstaff,M. (1986) *EMBO J.*, **5**, 3083–3090.

Jackson,R.J., Howell,M.T. and Kaminski,A. (1990) *Trends Biochem. Sci.*, **15**, 477–483.

Jang,S.K., Kräusslich,H.-G., Nicklin,M.J., Duke,G.M., Palmenberg,A.C. and Wimmer,E. (1988) *J. Virol.*, **62**, 2636–2643.

Jang,S.K., Pestova,T., Hellen,C.U.T., Witherell,G.W. and Wimmer,E. (1991) In Thach,E. (ed.), *Translationally Regulated Genes in Higher Eukaryotes*. Karger, Basel, in press.

Jefferson,R.A. (1987) *Plant Mol. Biol. Rep.*, **5**, 387–405.

Jefferson,R.A., Burgess,S.M. and Hirsch,D. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8447–8451.

Kaufman,R.J., Murtha,P. and Davies,M.V. (1987) *EMBO J.*, **6**, 187–193.

Kozak,M. (1986a) *Cell*, **44**, 283–292.

Kozak,M. (1986b) *Cell*, **47**, 481–483.

Kozak,M. (1987) *Mol. Cell. Biol.*, **7**, 3438–3445.

Kozak,M. (1989a) *J. Cell. Biol.*, **108**, 229–241.

Kozak,M. (1989b) *Mol. Cell. Biol.*, **9**, 5134–5142.

Liu,C.-C., Simonsen,C.C. and Levinson,A.D. (1984) *Nature*, **309**, 82–85.

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

Mason,W.S., Taylor,J.M. and Hull,R. (1987) *Adv. Virus. Res.*, **32**, 35–96.

Miller,P.F. and Hinnebusch,A.G. (1989) *Genes Dev.*, **3**, 1217–1225.

Pain,V.M. (1986) *Biochem. J.*, **235**, 625–637.

Peabody,D.S. and Berg,P. (1986) *Mol. Cell. Biol.*, **6**, 2695–2703.

Peabody,D.S., Subramani,S. and Berg,P. (1986) *Mol. Cell. Biol.*, **6**, 2704–2711.

Pelletier,J. and Sonenberg,N. (1985) *Cell*, **40**, 515–526.

Pelletier,J. and Sonenberg,N. (1988) *Nature*, **334**, 320–325.

Pietrzak,M., Shillito,M., Hohn,T. and Potrykus,I. (1986) *Nucleic Acids Res.*, **14**, 5857–5868.

Rhoads,R.E. (1988) *Trends Biochem. Sci.*, **13**, 52–56.

Sanfaçon,H. and Hohn,T. (1990) *Nature*, **346**, 81–84.

Schwarz,S., Felber,B.K., Fenyö,E.-M. and Pavlakis,G. (1990) *J. Virol.*, **64**, 5448–5465.

Sedman,S.A. and Mertz,J. (1988) *J. Virol.*, **62**, 954–961.

Sedman,S.A., Gelembiuk,G.W. and Mertz,J.E. (1989a) *J. Virol.*, **64**, 453–457.

Sedman,S.A., Good,P.J. and Mertz,J.E. (1989b) *J. Virol.*, **63**, 3884–3893.

Spanjaard,R.A. and van Duin,J. (1989) *Nucleic Acids Res.*, **17**, 5501–5507.

Steege,D.A. and Ivey-Hoyle,M. (1990) In McCarthy,J.E.G. and Tuite,M.F. (eds), *Post-transcriptional Control of Gene Expression*. NATO ASI series, volume H49, Springer, Heidelberg, pp. 197–206.

Takahashi,H., Shimamoto,K. and Ehara,Y. (1989) *Mol. Gen. Genet.*, **216**, 188–194.

Thomas,K.R. and Capecchi,M.R. (1986) *Nature*, **324**, 34–38.

Wang,F., Petti,L., Braun,D., Seung,S. and Kieff,E. (1987) *J. Virol.*, **61**, 945–954.

Williams,N.P., Mueller,P. and Hinnebusch,A.G. (1988) *Mol. Cell. Biol.*, **8**, 3827–3836.

Zoller,M.J. and Smith,M. (1983) *Methods Enzymol.*, **100**, 468–500.

Received on July 2, 1991; revised on August 6, 1991