Cellular Proteins Bind to Multiple Sites of the Leader Region of Cauliflower Mosaic Virus 35S RNA

DIANA I. DOMINGUEZ, THOMAS HOHN,1 and WALTRAUD SCHMIDT-PUCHTA

Friedrich Miescher-Institut, P.O. Box 2543, CH-4002 Basel, Switzerland

Received August 15, 1996; accepted September 24, 1996

UV crosslinking assays were performed to characterize interactions between the leader of the cauliflower mosaic virus (CaMV) 35S RNA and cellular, cytoplasmic proteins. From turnip, a host for CaMV, three different proteins, p35, p49, and p100, interacted with multiple binding sites within the leader. p35 binds to RNA nonspecifically and p49 binds with low specificity, whereas p100 interacts specifically with viral sequences. The expression of the proteins is not induced upon virus infection, as there is no difference in the protein pattern between healthy and infected cell extracts. In a cellular fractionation, p35 and p100 remain in the high-speed supernatant whereas p49 cosediments with polysomes. A possible involvement of these proteins in the translation of the CaMV 35S RNA is discussed.

INTRODUCTION

Cauliflower mosaic virus (CaMV) is the type member of caulimoviruses, a group of plant pararetroviruses [for a review see Rothnie et al. (1994)]. In CaMV-infected cells, mRNAs are transcribed from two different promoters (Fig. 1): the polycistronic 35S RNA serves as the template for reverse transcription [for a review see Bonneville and Hohn (1993)], as mRNA for viral proteins (Dixon and Hohn, 1984), and also as substrate for various splicing events (Kiss-László et al., 1995); and the subgenomic, monocistronic 19S RNA encodes the viral inclusion body protein (Covey and Hull, 1981), which acts also as translational transactivator (Bonneville et al., 1989; Fütterer and Hohn, 1992; Sha et al., 1995; Zijlstra and Hohn, 1992) and probably has additional functions (Himmelbach et al., 1996).

The terminally redundant 35S RNA contains a 600-nucleotide-long leader (Fig. 2A) that includes important signals controlling different functions of the viral life cycle: it comprises all the signals required for polyadenylation (Rothnie et al., 1994; Sanfaçon et al., 1991), the primer binding site for reverse transcription (Pfeiffer and Hohn, 1983), a splice donor (Kiss-László et al., 1995), and translation control signals (Fütterer et al., 1990; Fütterer et al., 1993). It has also been proposed to contain the packaging signal(s) (Fütterer et al., 1988; Richins et al., 1987).

Translation of most eukaryotic mRNAs is preceded by scanning of the 43S preinitiation complex along the RNA and initiation at the first AUG in a favorable context [scanning model of translation (Kozak, 1989)]. However, about 5–10% of mRNAs carry upstream small open reading frames (sORFs) that per se hinder translation (Kozak, 1987). Compared to cellular mRNAs, the CaMV 35S RNA leader is exceptional in several respects: it is unusually long and folds into a complex stem–loop structure (Fütterer et al., 1988). Furthermore, it contains, depending on the strain, 6 to 7 sORFs, 1 to 35 codons in length (Fütterer et al., 1988; Hohn et al., 1989). Upstream sORFs and long structured leader sequences are features of mRNAs whose expression is highly regulated, such as mRNAs of transcription factors [e.g., Lc (Damiani and Wessler, 1993)], proteins involved in signal transduction [e.g., opaque-2 (Lohmer et al., 1993); GCN4 (Müller and Hinnebusch, 1986); ornithine decarboxylase (Manzella and Fütterer and Hohn, 1992; Sha et al., 1995; Zijlstra and Hohn, 1992) and probably has additional functions (Himmelbach et al., 1996).

The negative effect of the leader sORFs on CaMV translation is partially compensated for by cis-acting signals present in the leader. It was proposed that these act through a shunt mechanism allowing the scanning complex to bypass the central leader region (Fütterer et al., 1990, 1993). Ribosome shunting has also been observed in a badnavirus, in rice tungro bacilliform virus (RTBV) (Fütterer et al., 1996), and in adenovirus late mRNAs, directed by a complex group of secondary structures in the tripartite leader (Yueh and Schneider, 1996). The molecular mechanism of the shunt is still unclear. However, since shunting can occur in the absence of any viral gene product, it is possible that (one or more) cellular factors could participate in the process.

Protein–RNA interactions between leader or trailer regions and cellular proteins were studied in a number of animal viruses [e.g., Leopardi et al., 1993; Pogue et al.,...
interacts strongly with two defined regions within the Protein extracts (5–10% leader of the CaMV 35S RNA).

UV crosslinking of RNA–protein complexes

Three proteins bind to multiple sites in this region. p100 interacts specifically with RNA. One of the p100 binding sites was further characterized by mutational analysis.

Preparation of cell extracts

Plasmid constructions

In vitro transcription

Uniformly labeled RNAs were prepared by in vitro transcription of the linearized plasmids, using SP6 (Boehringer Mannheim), T3 (Boehringer Mannheim), or T7 (Bio-finx) RNA polymerase and [α-32P]UTP (800 Ci/mmol; NEN) according to the manufacturer’s instructions. Transcripts were purified on 6% denaturing polyacrylamide gels. In competition experiments transcripts were prepared in the presence of tracer amounts of [α-32P]UTP in order to allow precise quantification of the amount of RNA added.

UV crosslinking of RNA–protein complexes

Protein extracts (5–10 μg protein) were preincubated at room temperature for 10 min with 100 μg/ml tRNA (Boehringer Mannheim) as a nonspecific competitor in a buffer containing 10 mM HEPES (pH 7.6), 8 mM MgCl2, 1 mM DTT, 4 mM spermidine, 1 mg/ml heparin, 5% glycerol, 20 U RNase inhibitor (Boehringer Mannheim) in a final volume of 10 μl. In competition experiments, com-
FIG. 2. UV crosslinking of different parts of the CaMV 35S RNA leader to cytoplasmic proteins from noninfected turnip leaves (S12). (A) 35S RNA leader with the small ORFs (A–F) and the first large ORF, ORF VII. (B) The entire leader and the various subfragments used to identify the binding sites of the proteins. The numbering of the leader sequence is taken from Fütterer et al. (1990). As a control, a nonspecific plasmid-derived RNA, a 169-nt-long transcript of the vector pT3T7 linearized with PvuII, was used. The binding of the proteins is indicated on the right. (C) Autoradiograph of a 12.5% SDS–polyacrylamide gel. In each sample, equimolar amounts (30 fmol) of the RNA were incubated with 5 μg of a protein extract. The size in kDa of protein markers is indicated on the left. Positions of p35, p49, and p100 are indicated on the right.

RESULTS

Identification of cellular proteins that bind to the CaMV leader

RNA–protein interactions were analyzed in UV crosslinking experiments after incubating the full-length leader of the CaMV 35S RNA (Fig. 2B) with a cytoplasmic protein extract (S12) from noninfected turnip leaves. Three proteins with apparent molecular masses of 35, 49, and 100 kDa were detected by autoradiography after separation by SDS–PAGE (Fig. 2C, lane 1). p49 usually appears as a doublet (see Discussion).

The protein character of the labeled bands was tested in control reactions that included proteinase K digestion prior to gel electrophoresis. After this treatment, the labeled bands disappeared, indicating that they resulted from protein–RNA interactions (data not shown). In addition, when the UV crosslinking experiment was performed in the absence of protein extracts, no band was visible (Fig. 3B, lane 1). This indicates that the RNA digestion was complete, and that the labeled bands resulted indeed from RNA–protein interactions and not from partial RNA digestion.

To test whether these complexes were formed with specific sequences within the leader, subclones were prepared to produce four RNA fragments (I to IV, Fig. 2B) that span the entire leader. In addition, subfragments of I and IV (Ia and Ib, and IVa and IVb, respectively) were tested, with Ib being in fact a partially overlapping fragment of I, containing 17 additional nucleotides at its 3′-end. All of these leader-specific fragments, as well as a nonspecific plasmid-derived RNA, could be crosslinked to p35 (Fig. 2C, lanes 1 to 10), indicating that p35 binds RNA nonspecifically. p49 was strongly crosslinked to fragments I, II, III, and IV (lanes 2, 5, 6, and 7). Within fragment I, p49 did not bind to Ia (lane 3), but it interacted with Ib (lane 4), whereas within fragment IV, p49 bound to both subfragments, IVa and IVb (lanes 8 and 9). The fact that neither the first 60 nucleotides of the leader (lane 3) nor the vector-derived RNA (lane 10) could be crosslinked to this protein suggests that p49 interacts with CaMV leader with a certain specificity, although low. A third protein, p100, was bound to fragment IV (lane 7) and a weaker interaction was also observed with fragment I (lane 2). When subfragments were tested, p100 interacted with Ib and IVa (lanes 4 and 8, respectively), but not with Ia and IVb (lanes 3 and 9, respectively), indicating a specific binding of p100 to certain leader sequences.
The experiments discussed above were performed with extracts from healthy turnip leaves. When extracts from CaMV-infected turnip leaves were used, the same proteins interacted with the leader and no additional complex could be observed (data not shown). This indicated that the proteins interacting with the leader were neither induced nor degraded upon virus infection.

Specificity of binding

The specificity of the complexes formed between the cellular proteins and CaMV leader fragments was further tested in competition experiments where increasing amounts of competitor RNA were preincubated with the protein extract. Two different types of competitors were used: (i) in vitro transcribed RNA fragments and (ii) homopolymers.

In a first set of experiments, competition with unlabeled in vitro transcribed RNA fragments was used to investigate the specificity of binding of these proteins to fragment Ib (Fig. 3). Even a 100-fold excess of fragment Ia, which binds only p35, did not drastically affect cross-linking of p49 and p100 to fragment Ib (Fig. 3A, lane 5). A 200-fold excess of competitor was necessary to observe an effect on p49 and p100 binding (lane 6). However, competition was visible for p35 binding at 50-fold excess of competitor (lane 4). The same competition pattern was obtained when the vector derived RNA fragment was used as competitor (not shown). In contrast, competition with a homologous unlabeled probe inhibited the binding of all three proteins at a lower molar excess (Fig. 3B). A reduction of complex formation between the labeled Ib fragment and p49 and p100 was already visible at a 25-fold excess of competitor (Fig. 3B, lane 4). These experiments confirmed the specific binding of p100 to fragment Ib of the CaMV leader and shows that p49 binding is not completely nonspecific.

To examine whether the proteins interact preferentially with certain bases, poly(A), poly(U), poly(C), and poly(G) were tested as competitors. These single-stranded homopolymers did not compete well for the binding of any of the proteins including p35 (Fig. 4) at concentrations up to 100-fold excess. At 1000-fold excess, poly(A) still did not compete for binding (lane 4), whereas poly(U) showed competition for all three proteins (lane 7). poly(C) and poly(G) could compete to some extent for p35 binding, to a lesser extent for p49 binding, and almost not at all for p100 binding (lanes 10 and 13). To analyze whether this low competition capacity was due to the single-stranded nature of the homopolymers, double-stranded poly(A)·poly(U) was also tested. This polymer had no effect on the binding of the three proteins to fragment Ib, even at 1000-fold excess (Fig. 4, lane 17).

To demonstrate that the same proteins and not only proteins of similar molecular weights bind to the different leader regions, cross-competition experiments were carried out. Using fragment Ib as labeled RNA, fragment IVa was a very strong competitor for all three proteins, indicating that the same proteins of 35, 49, and 100 kDa bind to the two fragments (Fig. 5, lanes 1–4). This result was supported by an experiment in which fragment IVa was used as the labeled probe and fragment Ib was used as the competitor. Again, Ib was a strong competitor for the binding of all three proteins (not shown). These results show that the 100-kDa protein binding to fragment...
FIG. 4. Competition for binding of p35, p49, and p100 by homopolymers. Turnip S12 cytoplasmic extracts were crosslinked to fragment Ib in the presence of no competitor (lanes 1 and 14), and a 10, 100, and 1000 times molar excess of poly(A) (lanes 2 to 4), poly(U) (lanes 5 to 7), poly(C) (lanes 8 to 10), poly(G) (lanes 11 to 13) and poly(A)-poly(U) (lanes 15 to 17). Proteins were separated on SDS-12.5% polyacrylamide gels. Positions of p35, p49, and p100 are indicated.

Ib is the same as the one that binds to fragment IVa. The same is true for p49 and p35. Fragment III, on the other hand, turned out to be a weak competitor for binding of p35 and p49. A 100-fold excess of fragment III was necessary to visualize competition with p49 (Fig. 5, lane 7). As expected, less competition was observed for p100 binding, since this protein does not interact with fragment III (Fig. 2C, lane 6).

In a different approach, the effect of increasing ionic strength on the binding of p49 and p100 to various leader fragments was studied (Fig. 6). The standard reactions were carried out at a concentration of 50 mM KCl. Binding of p35 to fragments Ib and IVa was drastically reduced at 600 mM KCl (lanes 4 and 7). On the other hand, p35 binding to fragment III was strongly reduced already at 300 mM KCl (lane 12). The binding of p49 was notably reduced at 200 mM KCl where measured (lanes 2 and 11). However, at 400 mM KCl p49 still binds to fragment Ib and IVa (lanes 3 and 6, respectively), whereas binding to fragment III was nearly undetectable at 300 mM KCl (lane 12). These findings confirm the results obtained in the competition experiments, suggesting a weaker interaction of fragment III with these proteins. On the other hand, complex formation between p100 and fragments Ib and IVa was still visible at 800 mM KCl (lanes 5 and 8), indicating a high binding affinity of p100 for these RNA fragments.

Effect of point mutations in fragment IVa on protein binding

Fragments Ib and IVa, which interacted strongly with both p100 and p49, have no obvious sequence similarity besides their AU-richness. To investigate the cis-acting elements important for protein binding, fragment IVa was further analyzed. This fragment was chosen because it contains three repeats of an UAUAC sequence motif that had a substantial effect on translational efficiency in transfected protoplasts of a reporter gene located downstream of the leader (Fütterer et al., 1993). Point mutations were introduced into this fragment (Fig. 7A) and the interaction of the resulting RNA with the different proteins...
was tested (Fig. 7B). In fragment IVp, four bases were changed and one was deleted compared to the wild-type sequence. The UAUAC motifs were not touched. These changes did not alter binding of the proteins (Fig. 7B, lane 1). However, the additional mutation of the UAUAC motifs to CAUGC (mutant IVc, Fig. 7A) dramatically reduced the affinity for p100. Binding of p49 was also affected whereas binding to p35 was not altered (Fig. 7B, lane 3).

The mutation analysis suggests that the UAUAC motifs in fragment IVa form part of the important cis-elements required for p49 and especially p100 binding. According to computer predictions, the IVp and IVc RNAs can form a secondary structure similar to the wild-type IVa RNA, indicating that it is unlikely that the reduction in binding is due to major conformational changes of the RNAs.

**Protein p49 cosediments with polysomes**

The proteins p35, p49, and p100 were all found in the crude cytoplasmic extract (S12) of turnip leaves. A different protein pattern was observed in UV crosslinking experiments under the conditions described above, when nuclear extracts from turnip or Orychophragmus violacesus, another CaMV host plant, were used; no bands corresponding to molecular masses of 35, 49, and 100 kDa were observed (data not shown). In order to get a first insight into their cytoplasmic distribution, the S12

![Diagram of RNA binding complexes](image)

**FIG. 6.** Effect of ionic strength on complex formation. UV crosslinking with turnip S12 extract and (a) fragment Ib, (b) fragment IVa, and (c) fragment III in the presence of different KCl concentrations as indicated. The KCl concentration in standard reactions was 50 mM.

![Diagram of RNA sequences](image)

**FIG. 7.** Mutation analysis of fragment IVa. (A) Primary sequence of the different RNA transcripts used for the analysis. Vector sequences are indicated by lowercase letters; the start site of leader sequences is indicated by an arrowhead. Arrows show the position of the three UAUAC motifs. Fragment IVa was used as a reference. Deviations from this sequence are shown in the mutants. Hyphens, identical sequences; dots, deleted bases. (B) UV crosslinking experiment using S12 cytoplasmic extract from turnip and the different mutants shown in A. The positions of the proteins are indicated.
leader and to the pseudoknot domain in the 3'-UTR of TMV RNA and has been proposed to mediate the positive effect that these regions have on translation (Tanguay and Gallie, 1996).

The leader sequences of CaMV and RTBV pregenomic RNAs and adenovirus late RNAs have been proposed to direct translation by a mechanism known as ribosome shunting or jumping. (Fütteter et al., 1993, 1996; Yueh and Schneider, 1996). The fact that shunting can occur in the absence of viral gene products prompted us to search for cellular proteins binding to CaMV leader as possible candidates to participate in the process. The UV crosslinking experiments reported here revealed three cytoplasmic proteins that interact with the 5' leader of CaMV 35S RNA (Fig. 2). The relative amount of each of the three proteins that was crosslinked to RNA was variable between experiments. The reason for this is unknown, but it could reflect variations in the amount of each of the proteins in the different batches of cellular extracts. p35 was shown to bind rather nonspecifically to all RNAs tested. p49 did not bind to a nonviral plasmid-derived RNA; however, it interacted with most leader subfragments (Fig. 2C), with the exception of the first 60 nucleotides of the leader (fragment la). In many experiments p49 appeared as a double band. Whether this second band corresponds to a posttranslational modification of p49, a degradation product, or, alternatively, to an additional protein is not yet known. However, since this second band shows the same specificity as p49 (see competition experiments, Figs. 3 and 4) and the same cellular localization (Fig. 8), it is most probably related to p49 and not an additional protein. A more specific interaction with only two defined regions of the leader, subfragments Ib and IVa, could be demonstrated for p100. This interaction seems to be strong, since it requires more than 800 mM KCl to dissociate the complex (Fig. 6). This is comparable to the binding affinity of the TMV p30 protein to RNA (Citovsky et al., 1990). Besides these interactions, p100 was also bound to the larger fragments I and IV (Fig. 2C). The fact that p100 interacts more weakly with fragments I and IV than with the subfragments Ib and IVa could have different interpretations. One possibility is that the structure of the RNA is important for binding; a more favorable structure might be formed by fragment Ib and IVa, whereas fragments I and IV might adopt a conformation for which p100 would have less affinity. For fragment I, another possibility exists: subfragment Ib contains 17 additional nucleotides (Fig. 2B) which could be important for p100 recognition.

By cross-competition experiments it was shown that the same p35, p49, and p100 protein species bind to the different leader fragments and not different proteins having similar mobility in SDS–PAGE. These experiments also indicate that within each protein the same

![Figure 8. Localization of the proteins. Different protein fractions (see Materials and Methods) were crosslinked to fragment IVa. S12, crude cytoplasmic extract (lane 1); P, polysomal pellet (lane 2) and S120, high-speed supernatant (lane 3). The positions of the proteins are indicated.](image-url)
The entire CaMV 35S RNA leader is AU-rich. A comparison of the different binding sites for p100, however, did not reveal any apparent sequence similarity. The specific binding of a protein to different RNAs without any apparent sequence similarity was also described for the binding of p57 to different parts of the internal ribosome entry site of poliovirus RNA (Hellen et al., 1994) and foot and mouth disease virus RNA (Luz and Beck, 1991) and for the binding of the wheat germ protein p102 to the 5'- and 3'-UTRs of TMV RNA (Tanguay and Gallie, 1996).

To determine more precisely cis-acting elements required for binding, fragment IVa was further investigated. In transient expression experiments this region was shown to be important for translation of a reporter gene downstream of the leader. It contains three copies of a cis-acting RNA-binding element that have been defined as important for the shunt mechanism in translation (Fütterer et al., 1990, 1993). In particular, mutations introduced into the UAUAC motifs of fragment IVa reduced translation efficiency in protoplasts and drastically reduced binding of p100 and also p49.

In picornaviruses, the RNA binding proteins are thought to be part of a larger complex (Jang and Wimmer, 1990). This could also be true for the CaMV leader binding proteins. The two binding sites of p100 are about 200 nt apart in the primary sequence but are very close to each other in the proposed secondary structure (Fütterer et al., 1998). An interaction between different proteins and different parts of the leader might stabilize a specific tertiary conformation. Such a complex might allow the shunt mechanism to take place.

The 35S RNA is terminally redundant and the first 180 nt of the leader, including most of fragment Ib, are duplicated at the 3'-end of the RNA. Therefore an even more complex interaction between the 5'- and 3'-nontranslated ends of the CaMV 35S RNA could be considered. Studies are underway to purify the leader binding proteins and to investigate further their function in the viral life cycle, and as they are also present in noninfected cells, their function in the host.

ACKNOWLEDGMENTS

We thank Karl Gordon for providing the fragment Ib clone, his communication concerning the pseudoknot structure, and helpful discussions. We very much appreciate the technical assistance of Matthias Müller. We thank Helen Rothnie, Johannes Fütterer, Witold Filipowicz, and Lyubov Ryabova for critical reading of the manuscript. W.S.-P. was supported by an EMBO fellowship.

REFERENCES

Covey, S. N., and Hull, R. (1981). Transcription of cauliflower mosaic virus DNA. Detection of transcripts, properties, and location of the

binding site is used for interaction with the different RNAs.

ever this is not a prerequisite for proteins involved in translation, since some initiation and elongation factors are isolated from the supernatants of cytoplasmic extracts after a high-speed centrifugation to pelleting the poly-somal fraction (Lax et al., 1986). In addition, both binding sites of p100 overlap with regions of the CaMV leader that have been defined as important for the shunt mechanism of translation in transient expression experiments (Fütterer et al., 1990, 1993). In particular, mutations introduced into the UAUAC motifs of fragment IVa reduced translation efficiency in protoplasts and drastically reduced binding of p100 and also p49.


Pfeiffer, P., and Hohn, T. (1983). Involvement of reverse transcription in...