

Structural elements of the 3'-terminal coat protein binding site in alfalfa mosaic virus RNAs

Chantal B. E. M. Reusken and John F. Bol*

Institute of Molecular Plant Sciences, Gorlaeus Laboratories, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands

Received April 16, 1996; Revised and Accepted May 30, 1996

ABSTRACT

The 3'-termini of the three genomic RNAs of alfalfa mosaic virus (AIMV) and ilarviruses contain a number of AUGC-motifs separated by hairpin structures. Binding of coat protein (CP) to such elements in the RNAs is required to initiate infection of these viruses. Determinants for CP binding in the 3'-terminal 39 nucleotides (nt) of AIMV RNA 3 were analyzed by band-shift assays. From the 5'- to 3'-end this 39 nt sequence contains AUGC-motif 3, stem-loop structure 2 (STLP2), AUGC-motif 2, stem-loop structure 1 (STLP1) and AUGC-motif 1. A mutational analysis showed that all three AUGC-motifs were involved in CP binding. Mutation of the A- and U-residues of motifs 1 or 3 had no effect on CP binding but similar mutations in motif 2 abolished CP binding. A mutational analysis of the stem of STLP1 and STLP2 confirmed the importance of these hairpins for CP binding. Randomization of the sequence of the stems and loops of STLP1 and STLP2 had no effect on CP binding as long as the secondary structure was maintained. This indicates that the two hairpins are not involved in sequence-specific interactions with CP. They may function in a secondary structure-specific interaction with CP and/or in the assembly of the AUGC-motifs in a configuration required for CP binding.

INTRODUCTION

The family *Bromoviridae* consists of four genera of plant viruses with a similar tripartite genome. These are the bromoviruses, the cucumoviruses, the ilarviruses and the alfamoviruses. There are two distinctive features that set the alfamo- and ilarviruses apart from other genera in this family. First, the RNAs of alfamo- and ilarviruses lack the 3'-terminal tRNA-like structure that is present in the RNAs of bromo- and cucumoviruses. Second, the coat protein (CP) is required for initiation of alfamo- and ilarvirus infection while a mixture of the three genomic RNAs of bromo- and cucumoviruses is infectious as such. A mixture of the three genomic RNAs of alfalfa mosaic virus (AIMV), type member of the alfamoviruses, is infectious only when CP or its subgenomic messenger RNA 4 is present. This early function of CP has been termed 'genome activation'. Evidence has been obtained that

each of the genomic RNAs in an AIMV inoculum has to be complexed with a few CP molecules to permit initiation of infection. RNAs 1 and 2 of AIMV encode the replicase subunits P1 and P2 respectively, whereas RNA 3 encodes both the movement protein P3 and CP (for a review see 1). The four RNAs are separately encapsidated into bacilliform particles that are predominantly stabilized by RNA-CP interactions (2).

In addition to genome activation, interactions between AIMV CP and RNAs have been implicated in RNA stability (3), RNA replication (4,5), a switch from minus-strand RNA synthesis to plus-strand RNA synthesis (6-8) and cell-to-cell spread (9). The evidence of the importance of RNA-CP interactions in the life cycle of AIMV has led to several studies on the identification of CP binding determinants in the AIMV RNAs. The observation that AIMV RNAs are able to withdraw CP subunits from virus particles *in vitro* indicated that the RNAs contain specific binding sites with a high affinity for CP (10). In subgenomic RNA 4, which is identical in sequence to the 3'-terminal sequence of 881 nucleotides (nt) of RNA 3, these high-affinity binding sites were predominantly localized in the 3'-terminal untranslated region (UTR) (11). The 3'-terminal 145 nt of RNAs 1, 2 and 3 show a sequence similarity of 80% and despite the nucleotide differences in these sequences, all termini can be folded into a similar secondary structure (12). The proposed secondary structure, based both on phylogenetic comparisons (12,13) and enzymatic structure mapping data (14), consists of several stem-loop structures interspersed with single-stranded AUGC-motifs. This is illustrated in Figure 2A for the 3'-UTR of RNA 3. This pattern is conserved in the 3'-terminal structures of ilarviruses (Fig. 1). Previously, we demonstrated by band-shift assays the presence of a minimum of two specific binding sites for CP in the 3'-UTR of RNA 3 (23). Site 1 consists of the region between nt 11 and 127 and site 2 is located between nt 133 and 208 from the 3'-end. Both sites contain several hairpins and AUGC-motifs (Fig. 2A). Deletion studies revealed that the two sites could bind CP independently of each other and by site-directed mutagenesis it was shown that the AUGC-motifs are important for CP binding to both sites. The importance of the AUGC-motifs for CP binding was further illustrated by Houser-Scott *et al.* (13) who demonstrated that mutation of the second AUGC-motif in the 3'-terminal 39 nt fragment of RNA 3 resulted in a loss of CP binding. Moreover, these authors showed that poly(AUGC)₁₀ was unable to bind CP, indicating that the AIMV CP recognizes invariant AUGC motifs in the context of conserved structural elements.

* To whom correspondence should be addressed

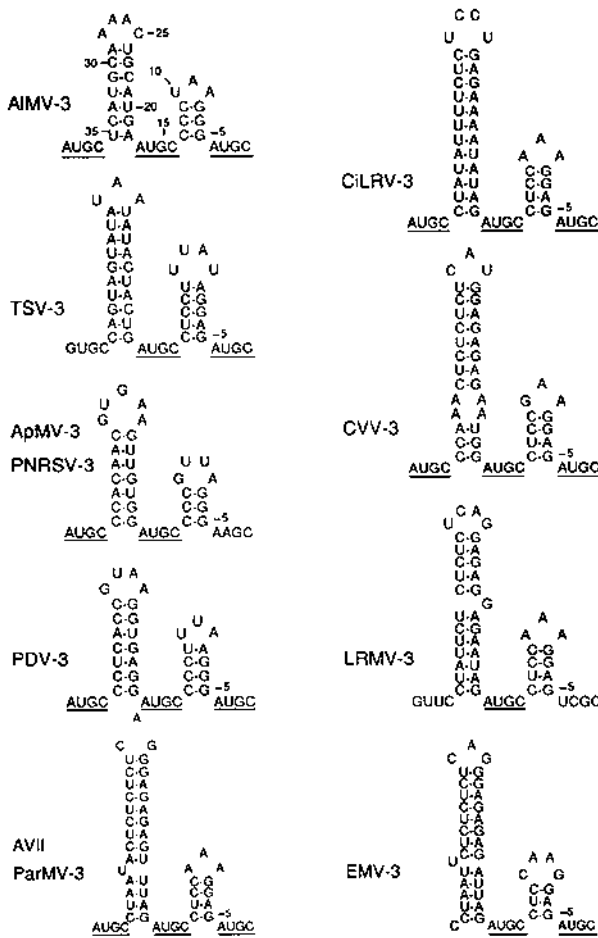


Figure 1. Predicted secondary structure at the 3'-terminus of RNA 3 of alfalfa mosaic virus (AIMV-3) (12), tobacco streak virus (TSV-3) (15,16), apple mosaic virus (ApMV-3) (17), prunus necrotic ringspot virus (PNRSV-3) (18), prune dwarf ilarvirus (PDV-3) (19), asparagus virus II (AVII) (20), parietaria mottle virus (ParMV-3) (13), citrus leaf rugose virus (CiLRV-3) (21), citrus variegation virus (CVV-3) (21), lilac ring mottle virus (LRMV-3) (22) and elm mottle virus (EMV-3) (13). AUGC-boxes are underlined.

In this study we investigated determinants for CP binding in the extreme 3'-terminal 39 nt of RNA 3. This sequence contains two hairpin structures flanked by three AUGC-motifs (Fig. 2B) and has been shown to bind CP specifically (13). The secondary structure pattern at the extreme 3'-end of the AIMV RNAs is extremely well conserved among AIMV and ilarviruses, as is illustrated in Figure 1. AIMV CP is able to bind to the 3'-termini of ilarvirus RNAs and vice versa and the CPs of AIMV and ilarviruses can be exchanged in genome activation (1). Thus, data on the structure of the AIMV CP binding site are probably also relevant for ilarviruses.

MATERIALS AND METHODS

RNA synthesis and purification

RNAs corresponding to the 3'-terminal 39 nt of AIMV RNA 3 were transcribed *in vitro* by T7 RNA polymerase from synthetic DNA templates containing a 24-base double-stranded T7 promoter and single-stranded sequences corresponding to the 3'-terminal

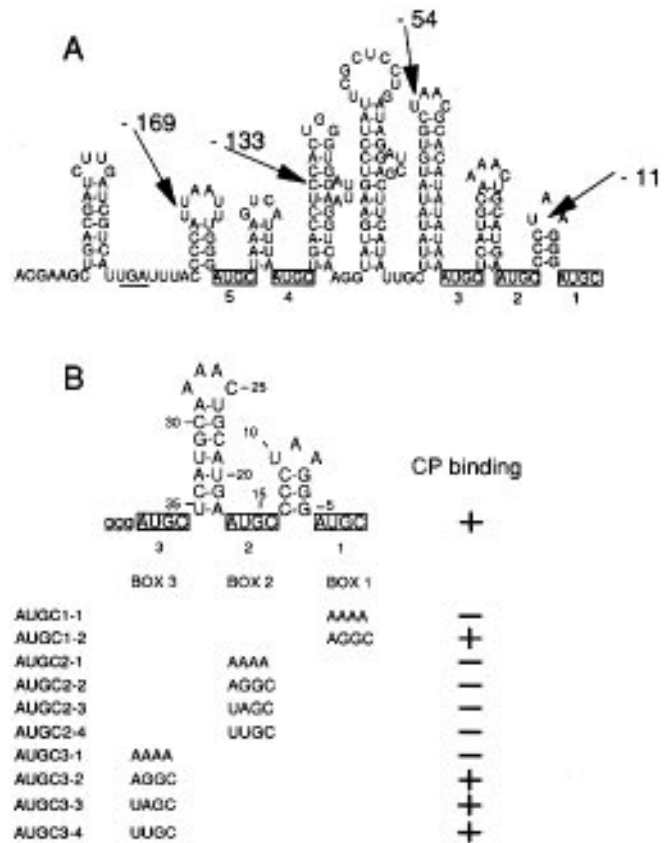


Figure 2. Predicted secondary structure at the 3'-terminus of AIMV RNA 3 and mutations introduced in AUGC-boxes 1, 2 and 3. (A) Predicted secondary structure of the 3'-UTR of RNA 3. The stop codon of the CP gene is underlined. AUGC-motifs 1-5 are boxed. The arrows indicate deletion end points discussed in the text. Numbering of nucleotides is from the 3'-end of RNA 3. (B) Predicted secondary structure of the T7 RNA polymerase transcript corresponding to the 3'-terminal 39 nt of RNA 3. Non-viral nucleotides are represented by lower case letters. AUGC-motifs 1, 2 and 3 are boxed and mutations made in these motifs are listed. The results from band-shift assays with the RNA transcripts are summarized by a plus sign (wt level of CP binding) or a minus sign (no detectable CP binding).

region of AIMV RNA 3. The partially double-stranded DNA templates were obtained by annealing a standard T7-oligodeoxyribonucleotide fragment (5'-AATTAAATACGACTCACTATAGG-3') to a second oligodeoxyribonucleotide fragment which is in part complementary to this T7-oligodeoxyribonucleotide fragment and in part coding for the RNA molecule of choice. Five μM of each oligodeoxyribonucleotide fragment was incubated in annealing buffer (10 mM Tris-HCl pH 7.8, 5 mM MgCl₂, 1 mM EDTA) for 15 min at 65°C and 20 min at room temperature prior to transcription. All RNAs contained an addition of GGG at their 5'-end which promotes efficient transcription (24) and were radiolabelled by incorporation of [³²P]UTP during the transcription reaction (reaction conditions: 200 mM Tris-HCl pH 8.2, 50 mM MgCl₂, 25 mM DTT, 5 mM spermidine, 40% PEG-6000, 0.05% Triton-X-100, 500 nM DNA, 1 mM NTPs, 0.05 mCi [³²P]UTP, 28 U RNasin and 17.5 U T7 RNA polymerase in total volume of 100 μl). All RNAs were purified on 15% polyacrylamide-7 M urea gels, eluted from the gels with 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1% SDS by shaking overnight at 37°C and ethanol

precipitated. Purified RNA was resuspended in deionized water and concentrations were determined on the basis of the specific activity of [32 P]UTP incorporated into the transcripts. Mutant RNA molecules were obtained by introducing the desired mutations in the oligodeoxyribonucleotide fragment containing the RNA 3 sequence.

Isolation of AIMV CP

CP was purified from AIMV particles in the form of a stable dimer as described previously (23). Briefly, virions were dissociated into protein and RNA by mixing one volume of virus suspension (~20 mg/ml) with one volume of 1 M MgCl₂ at 4°C. The RNA precipitate was pelleted by centrifugation and the protein in the supernatant was dialyzed extensively against a buffer containing 0.05 M sodium acetate, pH 5.5.

Gel mobility shift assays

Gel mobility shift assay conditions were adapted from Reusken *et al.* (23). One pmol of radiolabelled, gel-purified RNA and 3.75 pmol of purified AIMV CP were incubated in a total volume of 10 μ l of 10 mM phosphate, 1 mM EDTA, pH 7.0. The molar concentration of CP was calculated on the basis of its presence in the form of dimers in the preparation (23,25). Prior to use all RNAs were heated for 10 min at 65°C and chilled on ice. The resulting mixture was incubated for 30 min at room temperature. When indicated, unlabeled competitor RNAs were added at a 10-fold mass excess prior to addition of CP. Purified AIMV RNA 4 was used as the homologous competitor and RNA of turnip yellow mosaic virus (TYMV) was used as non-homologous competitor (23). The incubation mixtures were analyzed by electrophoresis into a 7.5% polyacrylamide gel (acrylamide–bisacrylamide ratio of 40:1) with 1 \times Tris-borate–EDTA as the running buffer. After electrophoresis for 2 h at 375 V, the gel was dried and exposed to X-ray film.

RESULTS

Specific binding of CP to 3'-terminal fragments of RNA 3

Zuidema *et al.* (26) reported that CP was able to rebind a 36 nt 3'-terminal fragment of RNA 3 that was protected by CP against digestion with ribonuclease T1. Subsequently, Houser-Scott *et al.* (13) showed that the 3'-terminal 39 nt of RNA 3 did bind CP with high affinity. The possibility to transcribe this 3'-terminal RNA fragment with T7 RNA polymerase from synthetic DNA templates greatly facilitated a mutational analysis of this sequence. The wild-type (wt) transcript of 39 nt with three nonviral G-residues at the 5'-end is shown in Figure 2B. It consists of two hairpin structures flanked by AUGC-boxes 1, 2 and 3. Binding of CP to wt and mutant fragments was tested by band-shift assays. An example of the technique is shown in Figure 3A. Lane 1 of Figure 3A shows the free radiolabelled wt transcript. Addition of CP results in a shift of part of the fragment (Fig. 3A, lane 2) which is completely abolished by addition of a 10-fold excess of the non-labelled homologous competitor (Fig. 3A, lane 4). The observation that addition of a 10-fold excess of a non-homologous competitor did not affect the band-shift (Fig. 3A, lane 3) demonstrates the specificity of the complex formation. Experiments with wt and mutant fragments were repeated at least four times with reproducible results.

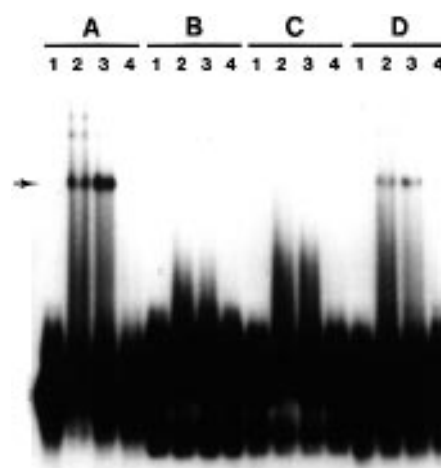


Figure 3. Band-shift analysis of the binding of CP to wt and mutant transcripts corresponding to the 3'-terminal 39 nt of RNA 3. The assays were done with wt RNA (A), mutant STLP2-1 (B), mutant STLP2-2 (C) and mutant STLP2-3 (D). The mutants are schematically shown in Figure 4. In lanes 1 no CP was added to the transcript. Binding of CP to the transcripts was studied in the absence of unlabelled competitor RNA (lanes 2) and in the presence of a 10-fold mass excess of TYMV RNA (lanes 3) or AIMV RNA 4 (lanes 4). The major RNA–protein complex is indicated by an arrow.

Role of AUGC-boxes in CP binding

To analyze the role of AUGC-boxes 1, 2 and 3 in the binding of CP to the 3'-terminal 39 nt of RNA 3, the mutations listed in Figure 2B were made. The results of band-shift assays with the mutant transcripts are summarized in this figure by a plus sign (binding similar to wt) or a minus sign (no detectable binding). When the sequence of boxes 1, 2 or 3 was changed into AAAA, no binding of CP was observed. This indicates that all three boxes play a role in the interaction of CP with the 39 nt fragment. Mutation of the sequence of box 2 to AGGC abolished CP binding but similar mutations of boxes 1 or 3 did not affect CP binding under our assay conditions (Fig. 2B). Apparently, sequence requirements for box 2 are more stringent than those for boxes 1 and 3. This conclusion is further substantiated by the observation that mutation of box 2 to UAGC or UUGC abolished CP binding whereas these mutations in box 3 had no effect (Fig. 2B).

Role of stem-loop structure 2 in CP binding

The 5'-proximal hairpin in the 39 nt fragment of RNA 3 is referred to as stem-loop structure 2 (STLP2) and the 3'-proximal hairpin is designated stem-loop structure 1 (STLP1). The stem structure of STLP2 contains an AUGC-motif at the position of nt 33 to 30 from the 3'-end of RNA 3. To investigate whether the stem structure of STLP2 or the primary sequence of this fourth AUGC-motif in the 39 nt fragment were important for CP binding, the mutations shown in Figure 4 were made (mutated nucleotides are indicated by a dot). Mutagenesis was targeted to the adjacent U–A and A–U basepairs (nt 32–21 and 33–20) highlighted in the upper right panel of Figure 4. Mutation of nt 33–32 from A–U into U–A in mutant STLP2-1 changed the AUGC-motif into UAGC and disrupted basepairing. Figure 3B shows that mutant STLP2-1 was unable to bind CP; the data are

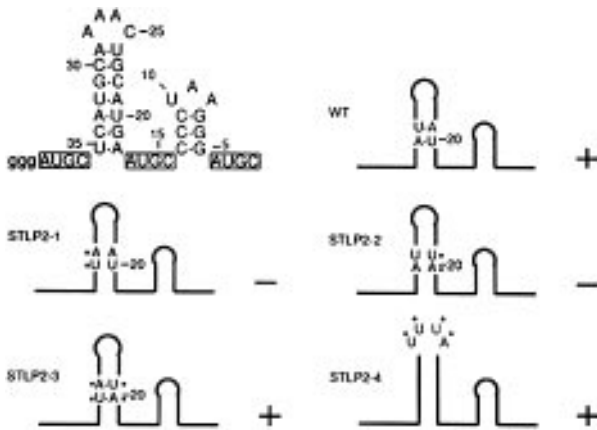


Figure 4. Mutations made in the 5'-proximal stem-loop structure (STLP2) of transcripts corresponding to the 3'-terminal 39 nt of RNA 3. Mutated nucleotides are indicated by dots. The two upper panels represent the wt sequence. The results of band-shift assays with the transcripts are summarized by a plus sign (wt level of CP binding) or a minus sign (no detectable CP binding).

summarized by a minus sign in Figure 4. In mutant STLP2-2 the A-U sequence of nt 21-20 was changed into U-A. This mutation disrupts basepairing but does not affect the AUGC-motif. As shown in Figure 3C and summarized by a minus sign in Figure 4, mutant STLP2-2 was unable to bind CP. In mutant STLP2-3 the mutations of STLP2-1 and STLP2-2 were combined, resulting in a change of the wt basepairs U-A and A-U into A-U and U-A respectively. Thus, basepairing is restored but mutant STLP2-3 contains the same modification of the AUGC-motif at position 33-30 as mutant STLP2-1. The observation that mutant STLP2-3 did bind CP (Fig. 3D; summarized by a plus sign in Fig. 4) demonstrates that the secondary structure of the stem of STLP2 is essential for CP binding rather than the AUGC-motif contained in this stem structure.

In mutant STLP2-4 the loop sequence AAAC of STLP2 (nt 28-25) was changed into UUUA. As this change did not affect CP binding (summarized by a plus sign in Fig. 4), the loop of STLP2 is probably not involved in the interaction of CP with the 39 nt fragment.

Role of stem-loop structure 1 in CP binding

Mutations that were introduced in the 3'-proximal stem-loop structure in RNA 3 (STLP1) are shown in Figure 5. In the literature it has been proposed that the G-residue at position 5 basepairs either with C-residue 13 (13) (conformation STLP1a shown in the upper left panel of Fig. 5) or with the C-residue at position 14 (12,29) (conformation STLP1b). In this latter conformation, G-7 would basepair with C-12 instead of C-11 and STLP1 has a tetranucleotide loop. Mutation of G-7 into a C-residue in mutant STLP1-1 would disrupt a basepair in both STLP1a and STLP1b. This mutant was unable to bind CP as summarized by the minus sign in Figure 5. A compensatory mutation of C-11 into G in mutant STLP1-3 would restore basepairing in the STLP1a conformation and the observation that this mutant is able to bind CP (plus sign in Fig. 5) supports the existence of this conformation. However, when only C-11 is mutated into a G-residue in mutant STLP1-5, a basepair is disrupted in the STLP1a conformation but the mutant is still able

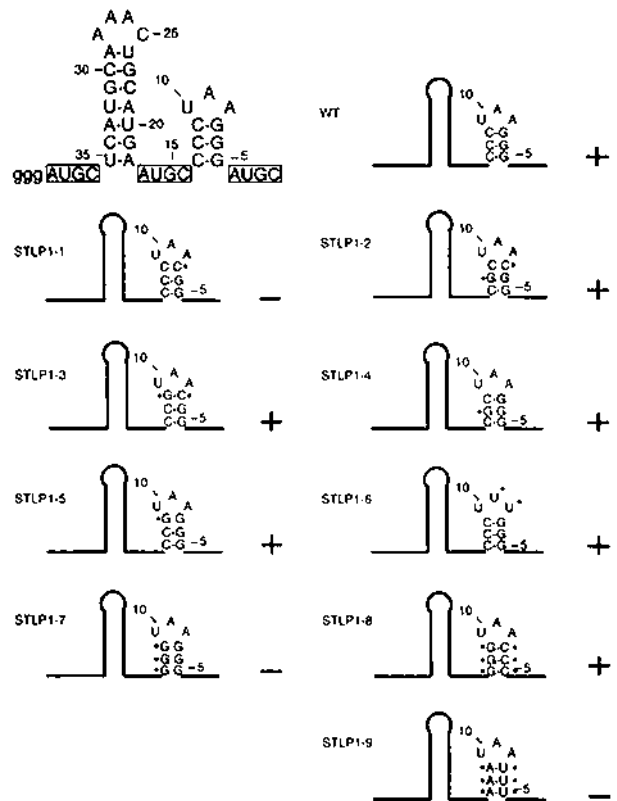


Figure 5. Mutations made in the 3'-proximal stem-loop structure (STLP1) of transcripts corresponding to the 3'-terminal 39 nt of RNA 3. Mutated nucleotides are indicated by dots. The two upper panels represent the wt sequence. The results of band-shift assays are summarized by a plus sign (wt level of CP binding) or a minus sign (no detectable CP binding).

to bind CP (Fig. 5). This could be explained by the assumption that a three basepair stem is restored by a conformational change from STLP1a to STLP1b. Similarly, the mutation of G-7 into C and C-12 into G in mutant STLP1-2 would disrupt two basepairs in the STLP1a conformation but in the STLP1b conformation of this mutant three G-C basepairs can be formed. The observation that this mutant is able to bind CP (Fig. 5) supports the existence of the STLP1b conformation. A role of the stem structure of STLP1 in CP binding is most convincingly illustrated by mutants STLP1-7 and STLP1-8. Mutation of C-residues 11, 12 and 13 into G-residues in mutant STLP1-7 would disrupt all three basepairs in the STLP1a conformation and two basepairs in STLP1b. This mutant is unable to bind CP (Fig. 5). The compensatory mutation of G-residues 5, 6 and 7 into C-residues in mutant STLP1-8 restores basepairing and CP binding (Fig. 5). When the three G-C basepairs of the STLP1 stem are changed into A-U basepairs in mutant STLP1-9, no CP binding was observed (Fig. 5). However, the computer program M-FOLD did not predict the formation of STLP1 in mutant STLP1-9, indicating that the structure of this mutant shown in Figure 5 is not stable. Only the results with mutant STLP1-4 were difficult to explain. The mutation of C-12 into G in this mutant disrupts a basepair in both the STLP1a and STLP1b conformation but yet the mutant was able to bind CP in four independent experiments (Fig. 5). Possibly, the interaction with CP stabilizes a stem structure of two G-C basepairs in this mutant.

The sequence of the loop in STLP1 was altered by changing A-residues 8 and 9 into U-residues in mutant STLP1-6. The observation that this mutant was able to bind CP (Fig. 5) indicates that the sequence of the loop is not critical for CP binding.

Structural elements of CP binding sites in ilarvirus RNAs

The 3'-termini of AIMV and ilarvirus RNAs contain several AUGC-boxes brought together by two hairpin structures of variable lengths (Fig. 1). To see if our assay system permits the analysis of binding of AIMV CP to the 3'-termini of ilarvirus RNAs, we transcribed a synthetic oligonucleotide *in vitro* into a RNA molecule corresponding to the 3'-terminal 49 nt of tobacco streak virus (TSV) RNA 3. This transcript did bind AIMV CP in the band-shift assay at the same level as obtained with the wt AIMV 39 nt transcript (Fig. 6, TSV-3). Also, a chimeric molecule containing STLP1 of TSV and STLP2 of AIMV did bind CP (Fig. 5, AIMV-TSV). The notion that the major function of the two hairpins is to arrange the AUGC boxes in the correct position was confirmed by randomizing the stem and loop sequences of STLP1 and STLP2 in the AIMV 39 nt fragment in such a way that the folding pattern and overall ΔG value was maintained. The resulting non-viral RNA did bind AIMV CP as efficiently as the wt 39 nt fragment in the band-shift assay (Fig. 6, random).

DISCUSSION

Although AIMV and ilarvirus CP's show little sequence similarity, they are able to bind the 3'-termini of each others RNAs and are interchangeable in genome activation (30–35). One of the 3'-structures recognized by these proteins in the viral RNAs probably consist of three AUGC-boxes or related motifs, interspersed by two hairpins. We have used short *in vitro* synthesized transcripts to study the determinants in this structure required for CP binding. Under the conditions of our band-shift assay, CP did bind to mutant transcripts either at a wt level or at a non-detectable level. We cannot rule out the possibility that some mutant transcripts had a reduced affinity for CP that would have been detectable at other conditions. Despite this limitation of the technique, the results presented in Figure 4 indicated that a stable stem structure of the 5'-hairpin (STLP2) in the 39 nt fragment is required for binding of AIMV CP *in vitro*. Because randomization of the sequence of the stem and loop of STLP2 had little effect on CP binding (Fig. 6), this hairpin is probably not involved in sequence-specific interactions with CP. The AUGC-motif in the stem of STLP2 is not conserved in the 3'-terminal structures of ilarviruses (Fig. 1) and the results from Figures 4 and 6 indicate that this motif has no role in CP binding other than participating in the stem structure.

The results with mutations in the 3'-hairpin (STLP1) of AIMV RNA 3 were more difficult to explain and suggested that the effect of some mutations was compensated by a conformational change which involved a slippage of the four C-residues at position 11–14 along the three C-residues at positions 5, 6 and 7 from the 3'-end. Such a slippage, resulting in basepairing of the C-residue of AUGC-box 2 with G-5, would considerably destabilize the stem of STLP1 of TSV RNA and several other ilarviruses shown in Figure 1. Possibly, the TSV fragment shown in Figure 6 would be a better candidate to study structural requirements of STLP1. However, the results with mutants STLP1-7 and STLP1-8 (Fig. 5) clearly support a role of the stem of STLP1 in CP binding.

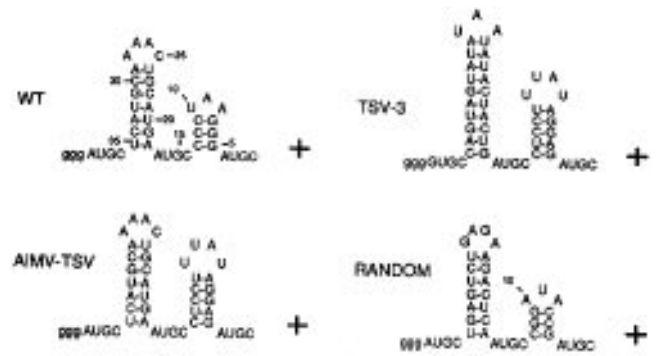


Figure 6. Predicted secondary structure of AIMV, TSV and mutant transcripts. The transcripts correspond to the 3'-terminal 39 nt of AIMV RNA 3 (wt), the 3'-terminal 49 nt of TSV RNA 3 (TSV-3), a RNA molecule with an AIMV derived 5'-proximal stem-loop structure and a TSV-derived 3'-proximal stem-loop structure (AIMV-TSV) and a RNA molecule with stem-loop structures of which the primary sequence was randomized (RANDOM). The results of band-shift assays are summarized by a plus sign (wt level of CP binding) or a minus sign (no detectable CP binding).

The results with mutant STLP1-6 indicate that the sequence of the loop of this hairpin is not important to CP binding. This is in agreement with the observation that binding of CP to the 3'-end of AIMV RNA 3 does not protect the loop of STLP1 against RNase A digestion (28) or in hydroxyl radical footprinting assays (13). In other RNA structures involved in protein binding, loop sequences have been implied in specific recognitions since they expose the RNA backbone and bases to interaction with protein groups (36).

A comparison of the 3'-termini of ilarvirus RNAs shown in Figure 1 reveals that only AUGC-box 2 is strictly conserved. Previously, it was shown that mutation of this box to AAAA abolished binding of CP to a 3'-terminal fragment of AIMV RNA 3 (13). We showed that similar mutations of AUGC-boxes 1 or 3 also inhibited CP binding (Fig. 2). However, mutation of the A- or U-residues of boxes 1 or 3 had no effect on CP binding whereas similar mutations in box 2 blocked the binding of CP (Fig. 2). Possibly, AUGC-box 2 and the G- and/or C-residues of boxes 1 and 3 are the minimal requirements of a structure with CP binding activity. Hairpins 1 and 2 would serve to bring these elements in the correct orientation and proximity. This type of protein binding element does not resemble any of the protein binding RNA structures that have been recently reviewed in the literature (37–39).

In addition to the minimal binding site in the 3'-terminal 39 nt, other CP binding sites are present in the 3'-UTR of AIMV RNA 3. The arrows in Figure 2A indicate the end points of 3'-terminal deletions of RNA 3 that were analyzed in a previous study (23). Deletion of the 3'-169 nt abolished CP binding whereas deletion of the 3'-133 nt permitted the binding of CP to site II which included AUGC-boxes 4 and 5. When this site II was deleted, binding of CP could be observed to a 3'-terminal sequence of 127 nt of RNA 3 (23). Deletion of the 3'-11 nt did not affect binding of CP to this sequence but when the 3'-54 nt were deleted no binding of CP was observed (23). Possibly, when the 3'-11 nt are deleted from the 127 nt sequence, AUGC-boxes 2 and 3 fulfil the role of boxes 1 and 2, respectively, and the function of box 3 is exerted by the UUGC-sequence flanking the 12 basepair hairpin upstream of box 3. The results with mutant AUGC3-4 in Figure 2 demonstrate that a UUGC-sequence at the position of

box 3 is functional in CP binding. Possibly, the seven hairpin structures in the 3'-UTR of AIMV RNA 3 arrange the five AUGC-boxes and UUGC-sequence in such a way that tandem binding sites are created that permit cooperative binding of CP dimers. Previously, we showed that mutation of AUGC-box 3 to AGGC abolished binding of CP to the 3'-terminal 127 nt sequence of RNA 3 whereas such mutations of boxes 1 and 2 did not affect CP binding (23). In the present study, we observed that mutation of AUGC-box 2 to AGGC in the 39 nt fragment abolished CP binding (Fig. 2B). This suggests that in the 127 nt sequence the function of box 3 is equivalent to that of box 2 in the 39 nt fragment whether or not box 1 is deleted from the 127 nt sequence. When AUGC-boxes were changed into AGGC in an infectious clone of RNA 3 and the accumulation of mutant RNA in tobacco plants was monitored, mutation of box 1 had no effect, mutation of box 2 reduced RNA accumulation and mutation of box 3 abolished RNA accumulation (9). Recently, we observed that these mutations have no effect on the recognition of RNA 3 by the purified AIMV replicase *in vitro* (manuscript in preparation). These data indicate that in the full-length RNA 3 box 3 has a major role in primary or cooperative binding of CP.

CP is not required for the recognition of plus-strand AIMV RNA 3 by the viral replicase *in vivo* (9) or *in vitro* (40). Our recent experiments showed that the purified AIMV replicase did not accept a RNA transcript corresponding to TSV RNA 4 as a template, indicating that the 3'-termini of AIMV and ilarvirus RNAs are equivalent in CP binding but not in replicase recognition (manuscript in preparation). Mutations analyzed in the present study are currently being introduced in an infectious clone of RNA 3 to investigate their effects on template activity towards the AIMV replicase *in vitro* and RNA 3 replication *in vivo*.

ACKNOWLEDGEMENTS

We thank Dr G. W. Witherell for his helpful advice concerning the T7 transcription procedure and C. M. A. van Rossum for the stimulating discussions. TYMV RNA was a gift from Mrs C. J. Houwing. This work was supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific research (NWO).

REFERENCES

- Jaspars, E.M.J. (1985) In J.W. Davies (ed.) *Molecular Plant Virology*. CRC Press, Boca Raton, Vol I, pp 155–221.
- Kaper, J.M. (1973) *Virology* **55**, 299–304.
- Neeleman, L., Van der Vossen, E.A.G. and Bol, J.F. (1993) *Virology* **196**, 883–887.
- Smit, C.H., Roosien, J., Van Vloten-Doting, L. and Jaspars, E.M.J. (1981) *Virology* **112**, 169–173.
- Quadt, R., Rosdorff, H.J.M., Hunt, T.W. and Jaspars, E.M.J. (1991) *Virology* **182**, 309–315.
- Nassuth, A. and Bol, J.F. (1983) *Virology* **124**, 75–85.
- Van der Kuyl, A.C., Neeleman, L. and Bol, J.F. (1991) *Virology* **183**, 687–694.
- Van der Kuyl, A.C. Neeleman, L. and Bol, J.F. (1991) *Virology* **185**, 496–499.
- Van der Vossen, E.A.G., Neeleman, L. and Bol, J.F. (1994) *Virology* **202**, 891–903.
- Van Vloten-Doting, L. and Jaspars, E.M.J. (1972) *Virology* **48**, 299–304.
- Houwing, C.J. and Jaspars, E.M.J. (1978) *Biochemistry* **17**, 2927–2933.
- Koper-Zwarthoff, E.C., Brederode, F.Th., Walstra, P. and Bol, J.F. (1979) *Nucleic Acids Res* **7**, 1887–1900.
- Houser-Scott, F., Baer, M.L., Liem, K.F., Cai, J.-M. and Gehrke, L. (1994) *J. Virol.* **68**, 2194–2205.
- Quigley, G.J., Gehrke, L., Roth, D.A. and Auron, P.E. (1984) *Nucleic Acids Res.* **12**, 347–366.
- Koper-Zwarthoff, E.C. and Bol, J.F. (1980) *Nucleic Acids Res.* **8**, 3307–3318.
- Cornelissen, B.J.C., Janssen, H., Zuidema, D. and Bol, J.F. (1984) *Nucleic Acids Res.* **12**, 2427–2437.
- Alrefai, R.H., Shiel, P.J., Domier, L.L., D'Arcy, C.J., Berger, P.H. and Korban, S.S. (1994) *J. Gen. Virol.* **75**, 2847–2850.
- Sanchez-Navarro, J.A. and Pallas, V. (1994) *J. Gen. Virol.* **75**, 1441–1445.
- Bachman, E.J., Scott, S.W., Xin, G. and Bowman Vance, V. (1994) *Virology* **201**, 127–131.
- EMBL/GenBank/DDBJ databases. Accession no. X86352.
- Scott, S.W. and Ge, X. (1995) *J. Gen. Virol.* **76**, 957–963.
- Scott, S.W. and Ge, X. (1995) *J. Gen. Virol.* **76**, 1801–1806.
- Reusken, C.B.E.M., Neeleman, L. and Bol, J.F. (1994) *Nucleic Acids Res.* **22**, 1346–1353.
- Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
- Kruseman, J., Kraal, B., Jaspars, E.M.J., Bol, J.F., Brederode, F.Th. and Veldstra, H. (1971) *Biochemistry* **10**, 447–455.
- Zuidema, D., Cool, R.H. and Jaspars, E.M.J. (1984) *Virology* **136**, 282–292.
- Zuidema, D. and Jaspars, E.M.J. (1985) *Virology* **140**, 342–350.
- Zuidema, D., Bierhuizen, M.F.A. and Jaspars, E.M.J. (1983) *Virology* **129**, 255–260.
- Houwing, C.J. and Jaspars, E.M.J. (1982) *Biochemistry* **21**, 3408–3414.
- Van Vloten-Doting, L. (1975) *Virology* **65**, 215–225.
- Reusken, C.B.E.M., Neeleman, L. and Bol, J.F. (1995) *J. Virol.* **69**, 4552–4555.
- Gonsalves, D. and Garnsey, S.M. (1974) *Virology* **64**, 343–353.
- Gonsalves, D. and Garnsey, S.M. (1975) *Virology* **66**, 25–31.
- Gonsalves, D. and Garnsey, S.M. (1975) *Virology* **67**, 311–318.
- Gonsalves, D. and Fulton, R.W. (1977) *Virology* **81**, 398–407.
- Chastain, M. and Tinoco, I. (1991) *Prog. Nucleic Acid Res.* **41**, 131–177.
- Burd, C.G. and Dreyfuss, G. (1994) *Science* **265**, 615–621.
- Draper, D.E. (1995) *Annu. Rev. Biochem.* **64**, 593–620.
- Nagai, K. and Mattaj, I.W. (1994) *RNA-Protein Interactions*. IRL Press, Oxford.
- De Graaff, M., Man in't Veld, M.R. and Jaspars, E.M.J. (1995) *Virology* **208**, 583–589.