

Mapping of the RNA-binding domain of the cucumber mosaic virus movement protein

C. Vaquero, Y.-C. Liao, J. Nähring and R. Fischer

Institut für Biologie I (Molekulargenetik/Botanik), RWTH Aachen, Worringerweg 1, D-52074 Aachen, Germany

A series of in-frame deletion mutants was used to identify a domain within the 3a protein of cucumber mosaic virus (CMV) that is required for RNA-binding activity. Deletions in the 3a gene were generated by PCR and restriction digestion, and the resulting mutated 3a sequences were cloned either in pT7-7 or in pGEX-5X3 expression vectors. The mutated 3a proteins or fusions with glutathione S-transferase (GST) were expressed in *E. coli*, purified, and their nucleic acid-binding activities analysed by photochemical UV cross-linking assays using digoxigenin-UTP-labelled RNA probes. Comparative analyses of seven mutated 3a proteins obtained from inclusion bodies and eight GST fusion proteins revealed that there is an RNA-binding domain located between amino acids 174 and 233. This RNA-binding domain is able to bind single-stranded RNA out of the context of the complete 3a movement protein and is highly conserved within both subgroups of CMV.

The cucumber mosaic virus (CMV) 3a protein has been generally considered to be the movement protein of the virus. Analysis of 3a gene deletion mutants revealed that the 3a protein is not required for replication, but is essential for the spread of infection *in planta* (Suzuki *et al.*, 1991; Bocard & Baulcombe, 1993). Transgenic tobacco plants expressing the 3a gene have been shown to complement movement-deficient mutants (Kaplan *et al.*, 1995). Furthermore, the CMV 3a protein shares several properties with other well-characterized viral movement proteins, such as cell-wall localization in transgenic and infected tobacco plants and the ability to modify the size-exclusion limit of plasmodesmata (Vaquero *et al.*, 1994, 1996) and to traffic RNA from cell-to-cell (Ding *et al.*, 1995).

The ability to bind nucleic acids *in vitro* is a common feature of plant virus movement proteins (Citovsky *et al.*, 1990, 1991;

Osman *et al.*, 1992; Schoumacher *et al.*, 1992*a, b*; Tacke *et al.*, 1991; Rouleau *et al.*, 1994; Ivanov *et al.*, 1994; Bleykasten *et al.*, 1996). Distinct domains essential for this binding activity have been identified in several movement proteins, such as the 30 kDa ('30K') protein of tobacco mosaic virus (TMV) (Citovsky *et al.*, 1992), the 35 kDa protein of red clover necrotic mosaic virus (RCNMV) (Osman *et al.*, 1993; Giesman-Cookmayer & Lommel, 1993), the P3 protein of alfalfa mosaic virus (AIMV) (Schoumacher *et al.*, 1994) and the P1 protein of cauliflower mosaic virus (CaMV) (Thomas & Maule, 1995). The 30K protein of TMV comprises two domains: domain 'A', which interacts with nucleic acids only in the context of the complete movement protein, and has a low surface probability, and the independent domain 'B', which has a high surface probability (Citovsky *et al.*, 1992). The CMV 3a protein also binds single-stranded (ss) nucleic acids *in vitro* (Li & Palukaitis, 1996). To define the domain of the 3a protein required to bind ssRNA, we employed deletion mutagenesis using mutated 3a proteins from inclusion bodies as well as soluble glutathione S-transferase (GST) fusion proteins.

The recombinant plasmid pT7-7C3a containing the cDNA sequence of the RNA 3 corresponding to nucleotides (nt) 99–1022 of a Spanish isolate of CMV (CMV-24) has been described previously (Vaquero *et al.*, 1996). Deletions within the 3a gene (nt 99–938) were created by PCR amplification using plasmid pT7-7C3a as template and specific oligonucleotides containing appropriate restriction endonuclease sites as primers, or by exploiting the existing *StyI* or *KpnI* restriction sites in the 3a gene. Mutated 3a sequences were cloned either into *NdeI*- and *SalI*-digested pT7-7 expression vector (Tabor, 1990) to create the DNA constructs M1–M7, or into pGEX-5X3 (Pharmacia) between the *BamHI* and *XhoI* sites for constructs M8–M15 (Fig. 1). The full-length 3a gene was cloned in pGEX-5X3 to create plasmid pGEX-5X3/C3a. All recombinant plasmids were confirmed by sequence analysis.

The 3a protein and deletion proteins M1–M7 were expressed in *E. coli* BL21(DE3) (Novagen). The recovery of inclusion bodies was carried out as described by Vaquero *et al.* (1996), with the exception that bacteria were disrupted by sonication. The proteins were solubilized by resuspending inclusion bodies with 8 M urea in binding buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 50 mM NaCl, 1 mM PEFA-

Author for correspondence: Rainer Fischer.

Fax +49 241 8888 182. e-mail Rainer.Fischer@post.rwth-aachen.de

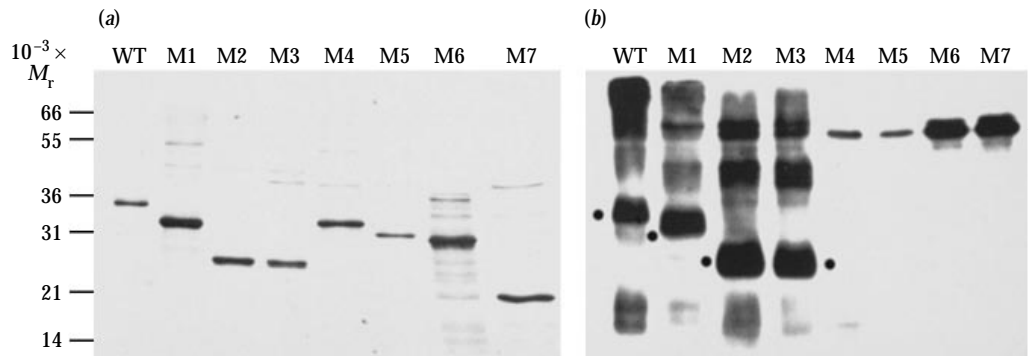


Fig. 2. RNA-binding analysis of the CMV 3a protein (WT) and deletion mutant proteins (M1–M7) expressed in *E. coli* BL21 and purified from inclusion bodies by UV cross-linking. Five ng of DIG-labelled ssRNA probe was incubated with approximately 100 ng of WT and M5, 200 ng of M2, M3 and M4, and 500 ng of M1, M6 and M7. Reaction mixtures were subjected to UV treatment and digested with RNase A as described in the text. Samples were analysed in duplicate by (a) SDS-PAGE followed by Coomassie blue staining, and (b) Western blot hybridization using anti-DIG Fab labelled with alkaline phosphatase. ●, Positions of 3a and deletion protein hybridization signals. M_r values of marker proteins are shown on the left side of the gel. *, Signal at $M_r \sim 70000$ corresponding to contaminating *E. coli* protein binding the ssRNA probe.

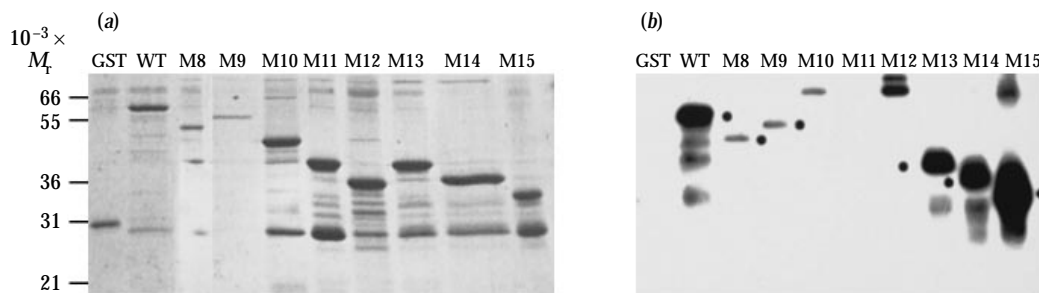


Fig. 3. RNA-binding analysis of the GST-3a and GST-deletion mutant proteins. Approximately 200 ng of GST, M8 and M9, 500 ng of WT and M15, and 1000 ng of M10, M11, M12, M13 and M14 were used for a UV cross-linking assay. (a) SDS-PAGE analysis; (b) Western blot hybridization analysis. ●, Positions of GST-3a and GST-deletion protein hybridization signals. Marker proteins and (*), as in Fig. 2.

(Fig. 3b), although it was not clearly visible on Coomassie-stained gels. These bands were shown to be of bacterial origin by Western blotting using a polyclonal antiserum raised against BL21(DE3) bacterial proteins (data not shown). The presence of this co-purified bacterial protein in nucleic acid-binding assays has been reported for other virus movement proteins (Rouleau *et al.*, 1994; Thomas & Maule, 1995).

To identify the region of the 3a protein involved in interaction with RNA, we tested the ability of deletion proteins M1–M7, isolated from inclusion bodies, to bind ssRNA (Figs 1 and 2). As shown in Fig. 2, proteins M1 and M2, carrying deletions at the N terminus ($\Delta 4$ –39 and $\Delta 1$ –85, respectively), and M3, with a deletion in the central region ($\Delta 86$ –170), were fully competent for ssRNA binding. In contrast, all deletions at the C terminus located between amino acids (aa) 176 and 279 (M4–M7) completely abolished RNA binding. These results indicated that the region between aa 176 and 279 is involved in the protein–RNA interaction.

Previous studies on the nucleic acid-binding properties of virus movement proteins have shown that some deletion proteins were unable to bind RNA due to incorrect refolding (Citovsky *et al.*, 1992; Osman *et al.*, 1993). To exclude this

possibility and to confirm the presence of an RNA-binding domain within the C-terminal region of the 3a protein, we tested 3a deletion proteins fused to GST (M8–M15) (Figs 1 and 3). Analyses of mutants M10 and M8, retaining two-thirds of the N terminus (aa 1–175) or C terminus (aa 86–279) of the protein, respectively, showed that only M8 retained the ability to interact with RNA, indicating that a domain located between aa 175–279 is responsible for RNA interaction. To find out whether any cryptic binding sites were hidden in M8, three deletions fused to GST, namely M11, M12 and M13, which retained the N-terminal (aa 1–107), central (aa 86–175) and C-terminal (aa 174–279) regions of the 3a protein, respectively, were assayed for their RNA-binding activity. Neither M11 nor M12, which both lack the C-terminal region, was able to interact with the ssRNA probe, while M13 strongly bound RNA. This result confirmed that the C-terminal RNA-binding site is active out of the context of the 3a protein. To further map the RNA-binding domain of the 3a protein, we tested the GST fusion proteins M9, M14 and M15, which contain aa 1–233, 1–6 and 174–256, and 1–6 and 174–233, respectively. As shown in Fig. 3, the removal of 23 (M14) or up to 46 aa (M9 and M15) from the most C-terminal region of the

protein did not affect the RNA-binding activity of the fusion proteins. Therefore, we concluded that there is an RNA-binding site located between aa 174 and 233 of the CMV 3a protein which is able to interact with ssRNA both in the context of the 3a protein and when fused to GST (M15).

Proteins M4 and M5 were purified in several independent experiments using different buffers for solubilization and refolding from inclusion bodies. In contrast to the results obtained with soluble M9, M14 and M15 GST fusion proteins, none of the purified M4 or M5 proteins bound RNA, suggesting that proteins carrying deletions within the C terminus were unable to adopt the correct conformation after denaturation and refolding.

The amino acid sequence of the 3a protein deduced from the nucleotide sequence of RNA 3 from CMV-24 (Vaquero *et al.*, 1996) is identical to that of the 3a protein of CMV-Fny (Owen *et al.*, 1990). Complementation studies using the Fny strain showed that deletion of 43 aa from the C terminus of the 3a protein did not affect the ability of the virus to infect non-transformed tobacco plants (Kaplan *et al.*, 1995). Here, we show that proteins M9 and M15, which lack the C-terminal 46 aa, were fully competent for RNA binding, suggesting that this region is also not involved in the interaction with RNA. The function of this C-terminal portion of the 3a protein during virus infection remains unclear. Kaplan *et al.* (1995) also reported that a deletion of up to 70 aa (Δ 209–279) at the C terminus resulted in the virus losing its ability to infect wild-type tobacco plants, indicating the importance of aa 209–236 of the CMV 3a protein for establishment of infection. Our results demonstrate that the region between aa 174 and 233 is an essential domain of the CMV 3a protein for protein–RNA interaction *in vitro*. The inability of the Δ 209–279 CMV mutant to infect host plants was probably due to the inability of the mutated 3a protein to interact with the viral RNA, as has been proposed for other virus movement proteins (Citovsky *et al.*, 1990, 1991).

The biological significance of the identified RNA-binding domain (aa 174–233) is also supported by its high degree of sequence conservation among the 3a proteins of both subgroups of CMV (Kaplan *et al.*, 1995). The sequence between aa 174–233 of the CMV-24 3a protein shows 100% sequence identity with four strains of subgroup I (CMV-Fny, CMV-O, CMV-Y and CMV-M), and differs at only nine positions from three strains belonging to subgroup II (CMV-Q, CMV-Kin and CMV-Trk7). Furthermore, this stretch of 60 aa includes a large number of positively charged residues which could form ionic bonds with negatively charged phosphate groups, as well as many residues with the potential to interact with nucleic acids (Fig. 1*b*). Similar to RNA-binding domain 'B' of TMV (Citovsky *et al.*, 1992), the RNA-binding domain of CMV 3a protein has a high surface probability.

In summary, the CMV-24 3a protein has an ssRNA-binding domain located between aa 174 and 233 which is able to interact with ssRNA of CMV as well as TMV and PRLV (data

not shown) out of the context of the remainder of the movement protein. As this RNA-binding domain is highly conserved in both subgroups of CMV, ongoing experiments will further elucidate the role of this domain in virus movement.

We would like to thank Dr I. García-Luque for providing plasmid pT7-7C3a, and Drs D. Baulcombe and R. Willmott for critical reading of the manuscript. C.V. was supported by a postdoctoral fellowship from the Ministerio de Educación y Ciencia, Spain.

References

- Bleykasten, C., Gilmer, D., Guilley, H., Richards, K. E. & Jonard, G. (1996). Beet necrotic yellow vein virus 42 kDa triple gene block protein binds nucleic acid *in vitro*. *Journal of General Virology* **77**, 889–897.
- Boccard, F. & Baulcombe, D. C. (1993). Mutational analysis of cis-acting sequences and gene function in RNA 3 of cucumber mosaic virus. *Virology* **193**, 563–578.
- Citovsky, V., Knorr, D., Schuster, G. & Zambryski, P. (1990). The P30 movement protein of tobacco mosaic virus is a single-strand nucleic acid protein. *Cell* **60**, 637–647.
- Citovsky, V., Knorr, D. & Zambryski, P. (1991). Gene I, a potential cell-to-cell movement locus of cauliflower mosaic virus, encodes an RNA-binding protein. *Proceedings of the National Academy of Sciences, USA* **88**, 2476–2480.
- Citovsky, V., Wong, M. L., Shaw, A. L., Venkataram, P. & Zambryski, P. (1992). Visualization and characterization of tobacco mosaic virus movement protein binding to single-stranded nucleic acids. *Plant Cell* **4**, 397–411.
- Ding, B., Li, Q., Nguyen, L., Palukaitis, P. & Lucas, W. J. (1995). Cucumber mosaic virus 3a protein potentiates cell-to-cell trafficking of CMV RNA in tobacco plants. *Virology* **207**, 345–353.
- Giesman-Cookmayer, D. & Lommel, S. A. (1993). Alanine scanning mutagenesis of a plant virus movement protein identifies three functional domains. *Plant Cell* **5**, 973–982.
- Ivanov, K. I., Ivanov, P. A., Timofeeva, E. K., Dorokhov, Y. L. & Atabekov, J. G. (1994). The immobilized movement proteins of two tobamoviruses form stable ribonucleoprotein complexes with full-length viral genomic RNA. *FEBS Letters* **346**, 217–220.
- Kaplan, I. B., Shintaku, M. H., Li, Q., Zhang, L., Marsh, L. E. & Palukaitis, P. (1995). Complementation of virus movement in transgenic tobacco expressing the cucumber mosaic virus 3a gene. *Virology* **209**, 188–199.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Li, Q. & Palukaitis, P. (1996). Comparison of the nucleic acid- and NT-binding properties of the movement protein of cucumber mosaic virus and tobacco mosaic tobamovirus. *Virology* **216**, 71–79.
- Osman, T. A. M., Hayes, R. J. & Buck, K. W. (1992). Cooperative binding of the red clover necrotic mosaic virus movement protein to single-stranded nucleic acids. *Journal of General Virology* **73**, 223–227.
- Osman, T. A. M., Thömmes, P. & Buck, K. W. (1993). Localization of a single-stranded RNA-binding domain in the movement protein of red clover necrotic mosaic dianthovirus. *Journal of General Virology* **74**, 2453–2457.
- Owen, J., Shintaaku, M., Aeschleman, P., Ben Tahar, S. & Palukaitis, P. (1990). Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 3. *Journal of General Virology* **71**, 2243–2249.
- Rouleau, M., Smith, R. J., Bancroft, J. B. & Mackie, G. A. (1994).

Purification, properties, and subcellular localization of foxtail mosaic potyvirus 26-kDa protein. *Virology* **204**, 254–265.

Schoumacher, F., Erny, C., Berna, A., Godefroy-Colburn, T. & Stussi-Garaud, C. (1992a). Nucleic acid-binding properties of the alfalfa mosaic virus movement protein produced in yeast. *Virology* **188**, 654–667.

Schoumacher, F., Gagey, M.-J., Maira, M., Stussi-Garaud, C. & Godefroy-Colburn, T. (1992b). Binding of RNA by alfalfa mosaic virus movement protein is biphasic. *FEBS Letters* **308**, 231–234.

Schoumacher, F., Giovane, C., Maira, M., Poirson, A., Godefroy-Colburn, T. & Berna, A. (1994). Mapping of the RNA-binding domain of the alfalfa mosaic virus movement protein. *Journal of General Virology* **75**, 3199–3202.

Smith, D. B. & Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* **67**, 31–40.

Suzuki, M., Kuwata, S., Kataoka, J., Masuta, C., Nitta, N. & Takanami, Y. (1991). Functional analysis of deletion mutants of cucumber mosaic virus RNA 3 using an *in vitro* transcription system. *Virology* **183**, 106–113.

Tabor, S. (1990). Expression using the T7 RNA polymerase/promotor system. *Current Protocols in Molecular Biology* 16.2.1–16.2.11.

Tacke, E., Prüfer, D., Schmitz, J. & Rohde, W. (1991). The potato leafroll luteovirus 17K protein is a single-stranded nucleic acid-binding protein. *Journal General Virology* **72**, 2035–2038.

Thomas, C. L. & Maule, A. J. (1995). Identification of the cauliflower mosaic virus movement protein RNA-binding domain. *Virology* **206**, 1145–1149.

Vaquero, C., Turner, A. P., Demangeat, G., Sanz, A., Serra, M. T., Roberts, K. & García-Luque, I. (1994). The 3a protein from cucumber mosaic virus increases the gating capacity of plasmodesmata in transgenic tobacco plants. *Journal of General Virology* **75**, 3193–3197.

Vaquero, C., Sanz, A. I., Serra, M. T. & García-Luque, I. (1996). Accumulation kinetics of CMV RNA 3-encoded proteins and subcellular localization of the 3a protein in infected and transgenic tobacco plants. *Archives of Virology* **141**, 987–999.

Received 15 November 1996; Accepted 21 March 1997