Localization of a single-stranded RNA-binding domain in the movement protein of red clover necrotic mosaic dianthovirus

T. A. M. Osman, P. Thömmes and K. W. Buck*

Department of Biology, Imperial College of Science, Technology and Medicine, London SW7 2BB, U.K.

Mutant movement proteins of red clover necrotic mosaic dianthovirus (RCNMV), consisting of in-frame deletions or fusions with a maltose-binding protein, were produced in *Escherichia coli* using expression vectors. The ability of the mutant proteins to bind to ssRNA was tested by photochemical cross-linking and gel retardation. The results showed that the region between amino acids 181 and 225 of the RCNMV movement protein contains an ssRNA-binding domain.

There is a growing body of evidence that movement of plant viruses from cell to cell occurs via the intercellular channels known as plasmodesmata, and is mediated by virus-encoded movement proteins (for review, see Hull, 1989; Atabekov & Talianksy, 1990; Robards & Lucas, 1990; Maule, 1991; Citovsky & Zambryski, 1991; Deom *et al.*, 1992). It has been shown that the movement protein of tobacco mosaic tobamovirus (TMV) interacts with plasmodesmata, thereby increasing the size exclusion limit for movement of molecules between cells (Tomienius *et al.*, 1987; Wolf *et al.*, 1989; Deom *et al.*, 1990; Ding *et al.*, 1992), and that it binds cooperatively in *vitro* to ssRNA, producing long, thin unfolded complexes (Citovsky *et al.*, 1990, 1992). These properties led to the hypothesis that TMV movement protein–ssRNA complexes are targeted to, and translocated through, plasmodesmata (Citovsky *et al.*, 1990). The finding that movement proteins of viruses in other groups interact with plasmodesmata, or are found in cell wall fractions, and bind to ssRNA, e.g. cauliflower mosaic virus (Albrecht *et al.*, 1988; Linstead *et al.*, 1988; Citovsky *et al.*, 1991), alfalfa mosaic virus (Godefroy-Colbourn *et al.*, 1986; Stussi-Garaud *et al.*, 1987; Schoumacher *et al.*, 1992; Erny *et al.*, 1992), red clover necrotic mosaic dianthovirus (RCNMV) (Osman & Buck, 1991; Osman *et al.*, 1992), suggests that transport of movement protein–ssRNA complexes through modified plasmodesmata may be a common mechanism of virus cell-to-cell movement (Citovsky & Zambryski, 1991).

The movement protein of TMV has been shown to have two ssRNA-binding domains (Citovsky *et al.*, 1992). Domain A is located in a predominantly hydrophobic region with low surface probability (amino acids 112 to 185) and only binds to ssRNA in the context of the native protein. Domain B is located in the C-terminal 83 amino acids (residues 186 to 268), a predominantly hydrophilic region with high surface probability, and its RNA-binding activity is expressed both in the context of the native protein and when fused to the phage T7 gene 10 protein. Here we report the localization of an ssRNA-binding domain in the movement protein of RCNMV and compare its sequence with the two RNA-binding domains of the TMV movement protein.

Movement proteins with deletions (Table 1) were created by *in vitro* mutagenesis (Kunkel *et al.*, 1987) of a full-length cDNA clone of RNA 2 (Osman *et al.*, 1991b). An *Nde*I site was introduced at the translational initiation codon using the oligonucleotide AGGTAGGTTCATATGGCTATTC (Osman *et al.*, 1992). In-frame deletions (deletion numbers 1 to 6) were created using oligonucleotides containing 20 nucleotides on each side of the deletion. To create a movement protein lacking the 86 C-terminal amino acids (deletion no. 7), a stop codon was introduced after the codon for amino acid 240 using the oligonucleotide AGCCAATGACTGTAGTCTGAGG (Ingles, 1993). DNA fragments containing the mutant movement protein–coding regions and the 3′ untranslated regions were then excised using *Nde*I and *Bam*HI and cloned into the corresponding sites of the expression vector pET3a (Rosenberg *et al.*, 1987). Expression of the mutated movement proteins in *Escherichia coli* strain BL21(DE3)pLysE (Studier & Moffatt, 1986) and their purification were as described by Osman *et al.* (1992). When analysed by SDS–PAGE, each protein gave a single band of the expected size (Fig. 1a; Table 1).

The ability of each protein to bind to ssRNA was tested by photochemical cross-linking as described previously (Osman *et al.*, 1992). Each protein was incubated with a *32*P-labelled 387 nucleotide ssRNA, produced by *in vitro* transcription of a cDNA clone of...
Table 1. Deletion and fusion mutants of the RCNMV movement protein and their ability to bind to ssRNA

<table>
<thead>
<tr>
<th>Deletion no.</th>
<th>Amino acids deleted</th>
<th>Calculated Mr (K)</th>
<th>RNA binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type*</td>
<td></td>
<td>36</td>
<td>+</td>
</tr>
<tr>
<td>1</td>
<td>3-40</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>41-77</td>
<td>31</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>78-124</td>
<td>31</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>125-165</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>166-203</td>
<td>31</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>204-238</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>241-326</td>
<td>26</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fusion no.</th>
<th>Amino acids retained†</th>
<th>RNA binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2-326</td>
<td>79</td>
</tr>
<tr>
<td>2</td>
<td>1-180</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>181-326</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>181-240</td>
<td>49</td>
</tr>
<tr>
<td>5</td>
<td>181-225</td>
<td>48</td>
</tr>
</tbody>
</table>

* Wild-type is TpM-34.
† Amino acids linked to the maltose-binding protein.

RCNMV RNA 2, and irradiated with u.v. light. The complexes were then incubated with ribonuclease A and the products analysed by SDS-PAGE and autoradiography. The results showed that four of the deletion mutants, like the wild-type protein, had the ability to bind to the ssRNA probe (Table 1). The binding was not due to an undetected *E. coli* protein in the movement protein preparations because (i) when an equivalent protein preparation from *E. coli* transformed with pET3a lacking the RCNMV sequence was made, no bands could be detected in SDS-PAGE after Coomassie blue staining of the free protein or after carrying out the u.v. cross-linking assay, (ii) no binding was detected by preparations of deletion numbers 1, 4 and 6 and (iii) the electrophoretic mobilities of the covalently cross-linked protein–nucleotide complexes (Fig. 1b) correlated with those of the free wild-type and deletion mutants. Similar results were obtained with ssRNA probes of different lengths and with ssDNA probes (not shown), as reported previously for the wild-type protein (Osman et al., 1992). The three mutated proteins that did not bind were again expressed in *E. coli* and further protein preparations were made in the presence of 0.4% NP40 or Triton X-100. However, tests failed to reveal any ssRNA- or ssDNA-binding activity.

RNA-binding activity was also tested by gel retardation analysis. Each protein was incubated with a $^{32}$P-labelled ssRNA probe as above and the products were analysed by non-denaturing PAGE and autoradiography as described previously (Osman et al., 1992). The results (Fig. 2) showed that the four mutant proteins that had RNA-binding activity in the photochemical cross-linking assay also retarded the ssRNA probe in PAGE. The electrophoretic mobilities of the RNA–protein complexes formed by the mutant proteins and by the wild-type protein were similar, suggesting that similar amounts of protein were bound. Similar results were obtained with ssDNA probes (not shown). No retardation was observed with the three mutant proteins that did not bind ssRNA or ssDNA in the photochemical cross-linking assay.

The lack of *in vitro* ssRNA-binding activity of movement proteins with deletions of amino acids 3 to 40, 125 to 165 or 204 to 238 (Table 1) could indicate that amino acids from all of these regions participate in a single binding site. Alternatively it is possible that there is a binding site contained completely within one of the

Fig. 1. Analysis of the RNA binding of the deletion mutants of the RCNMV movement protein by photochemical cross-linking. (a) SDS–PAGE of purified wild-type and mutant proteins. After electrophoresis, the gel was stained with Coomassie blue. (b) The wild-type and mutant proteins (25 ng) were incubated with a $^{32}$P-labelled 387 nucleotide ssRNA probe (1.0 ng) in 0.2 M-NaCl and the complexes were analysed by photochemical cross-linking, SDS–PAGE and autoradiography as described by Osman et al. (1992). Lane 1, marker proteins ($M_r$ values are indicated on the side of the gel); lane 2, wild-type protein; lane 3, deletion no. 1; lane 4, deletion no. 2; lane 5, deletion no. 3; lane 6, deletion no. 4; lane 7, deletion no. 5; lane 8, deletion no. 6; lane 9, deletion no. 7.
regions, in which case the other two regions would not contain an RNA-binding site and the absence of either would therefore affect the binding indirectly.

To obtain more direct evidence on the location of one or more RNA-binding domains in the RCNMV movement protein, fusions of the whole or part of the protein with the E. coli maltose-binding protein were made using the expression vector pMAL-c2 (New England Biolabs), a derivative of pMAL-c (Maina et al., 1988). Fusion number 1, containing the whole of the movement protein coding region except the first methionine, was constructed as described by Osman & Buck (1991). For the other fusions, the appropriate regions were amplified using PCR (Mullis & Faloona, 1987) from either a cDNA clone of wild-type RNA 2 (Osman et al., 1991b) (fusion numbers 2 and 3) or cDNA clones of RNA 2 in which a stop codon had been introduced after the codon for amino acid 240 (fusion number 4) or amino acid 225 (fusion number 5) by in vitro mutagenesis using the oligonucleotides AGCCAATGACTGATGTCTGAGG and GGAGACACACTAATCTGAGGTATCGCC respectively (Ingles, 1993). PCR was carried out using the following oligonucleotide primers: fusion number 2, forward primer, GGGACTGAATTCATGGGCTATTCACGTGG, reverse primer, CTAGGCCGTCAAGTCAGTGTC; fusion number 4, forward primer, GGGACTGAATTCGGACGTTACTGTCGAGC, reverse primer, CTAGGCTGAGTCAGGTCACGTGTC. The forward primers contained an EcoRI site upstream of the coding region and the reverse primers contained a PstI site downstream of the coding region. The PCR products were purified by gel electrophoresis and cloned between the EcoRI and PstI sites of pMAL-c2 in E. coli DH5α. Expression of the proteins and purification by affinity chromatography on an amylose resin were as described by Osman & Buck (1991). In SDS-PAGE each protein gave a band of the expected size (Fig. 3a; Table 1).

The ssRNA-binding activity of each fusion protein was tested by photochemical cross-linking and SDS-PAGE as described for the deletion mutant proteins. The results are shown in Fig. 3 (b) and Table 1. A positive result was obtained with fusion number 1 which contained the whole of the RCNMV movement protein sequence. This established that the movement protein retained its ssRNA-binding activity when covalently linked to the carboxyl terminus of the maltose-binding protein, since no RNA-binding activity could be detected with the maltose-binding protein alone. The ssRNA-binding activity of fusion number 3 shows that there is a binding domain in the C-terminal half of the movement protein (amino acids 181 to 326).

Deletion numbers 1 and 4 which retain the C-terminal half of the movement protein did not show binding activity. Eukaryotic proteins produced in E. coli are often inactive; this can be because they do not adopt their native conformation and tend to aggregate (Kohno et al., 1990). The native movement proteins and deletion numbers 1 to 7 were all insoluble when produced in E. coli and were solubilized by denaturation in 4 M-urea at an elevated temperature. Subsequent removal of the urea allowed the protein to refold. It is possible that deletion of amino acids 3 to 40 or 125 to 165 prevented correct refolding of the protein. Although addition of a non-ionic detergent, such as NP40, can aid in refolding by preventing intermolecular interactions between hydrophobic regions and was successfully used in the refolding of some deletion mutants of the TMV movement protein (Citovsky et al., 1992), such a detergent would have little effect on other types of interaction, such as electrostatic interactions or hydrogen bonding. In contrast to the native movement protein and its deletions, the maltose-binding protein–movement protein fusions were all soluble and did not require to be denatured and refolded.

It was shown previously that a truncated movement protein of an RCNMV mutant (TpM-341) which lacked the C-terminal 88 amino acids of the wild-type sequence, a region poorly conserved between different dianthoviruses (Osman et al., 1991a; Kendall & Lommell, 1992; Ge et al., 1992), had ssRNA-binding activity (Osman et al., 1991b, 1992). TpM-341 movement protein contained an additional 34 unique C-terminal amino acids which
might have affected its binding activity. However, deletion number 7 (Table 1) which lacked the C-terminal 86 amino acids and contained no additions also had ssRNA-binding activity. We therefore decided to delete the C-terminal 86 amino acids from fusion number 3 to create fusion number 4, which was also found to have RNA-binding activity (Fig. 3). A deletion of a further 15 amino acids created fusion number 5, which was also found to bind to RNA (Fig. 3). An RNA-binding site is therefore contained within the region between amino acids 181 and 225.

Within the sequence of amino acids 181 to 225 (Fig. 4), there are many amino acids with the potential for interacting with nucleic acids i.e. basic residues (H, K, R) which could form ionic bonds with negatively charged phosphate groups, amides (N, Q), acids (D, E) or hydroxylated amino acids (S, T, Y) which could form hydrogen bonds with nucleic acid bases, and aromatic amino acids (F, Y) which could form stacking interactions with nucleic acid bases. Amino acids 181 to 199 are predominantly hydrophobic with a low surface probability (Kendall & Lommell, 1992). However, the RNA-binding activity of deletion number 5 which lacks amino acids 166 to 203 and the absence of RNA-binding activity of deletion number 6 which lacks amino acids 204 to 238 suggest that the RNA-binding site may lie within the sequence of amino acids 204 to 225. This is a predominantly hydrophilic region with relatively high surface probability (Kendall & Lommell, 1992).

Sequence comparisons using FASTA and LFASTA programs (Pearson & Lipman, 1988) failed to reveal any significant similarity between sequences of amino acids 181 to 225 of the RCNMV movement protein and proteins in the SWISSPROT and NBRF/PIR protein databases or with RNA-binding domains A and B of the TMV movement protein (Citovsky et al., 1992). Based on predicted secondary structure and surface probability, Kendall & Lommell (1992) noted a similarity between the TMV movement protein RNA-binding domain B and amino acids 270 to 326 of the RCNMV movement protein i.e. distinct from the binding region identified here. A recent alignment of sequences of plant virus movement proteins placed TMV and RCNMV movement proteins in different families, although the RCNMV RNA-binding domain identified here was not part of the aligned sequence (Koonin et al., 1991). In summary there is no obvious similarity between the RNA-binding region identified here and either of the RNA-binding regions of the TMV movement protein. However, determination of the three-dimensional structures, and further in vitro mutagenesis and binding studies, will be needed to identify those amino acids in the movement proteins of both viruses which make contact with ssRNA and to evaluate the possibility of an evolutionary relationship between their RNA-binding domains.
We thank the AFRC for a grant. Sequence analysis was carried out using the SEQUENET computer facility of the SERC Daresbury Laboratory. Work with RCNMV was carried out under MAFF licence no. PHF 1250B/13 (22). P.T. thanks the European Molecular Biology Organisation for the award of a long-term fellowship.

References


(Received 22 March 1993; Accepted 8 June 1993)