Identification of the RNA-binding sites of the triple gene block protein 1 of bamboo mosaic potexvirus

Chiung-Hua Wung,1 Yau-Heiu Hsu,2 Dann-Ying Liou,2 Wei-Cheng Huang,3 Na-Sheng Lin4 and Ban-Yang Chang3

1, 2, 3 Agricultural Biotechnology Laboratories1, Institute of Agricultural Biotechnology2, and Institute of Biochemistry3, National Chung-Hsing University, Taichung, Taiwan 402, Republic of China
4 Institute of Botany, Academia Sinica, Taipei, Taiwan 115, Republic of China

The triple gene block protein 1 (TGBp1) encoded by open reading frame 2 of bamboo mosaic potexvirus (BaMV) was overexpressed in Escherichia coli and purified in order to test its RNA-binding activity. UV crosslinking assays revealed that the RNA-binding activity was present mainly in the soluble fraction of the refolded TGBp1. The binding activity was nonspecific and salt concentration-dependent: activity was present at 0–50 mM NaCl but was almost abolished at 200 mM. The RNA-binding domain was located by deletion mutagenesis to the N-terminal 3–24 amino acids of TGBp1. Sequence alignment analysis of the N-terminal 25 amino acids of the TGBp1 homologues of potexviruses identified three arginine residues. Arg-to-Ala substitution at any one of the three arginines eliminated most of the RNA-binding activity, indicating that they were all critical to the RNA-binding activity of the TGBp1 of BaMV.

Introduction

The flexuous rod-shaped particles of potexviruses have single-strand positive-sense RNA genomes that encode five well-conserved open reading frames (ORFs) (Bancroft et al., 1991; Zuidema et al., 1989). Mutational analysis of the genome of white clover mosaic potexvirus (WCIMV) has revealed that the central three ORFs (ORFs 2, 3 and 4), the ‘triple gene block’ (TGB), are required for the systemic spread of virus in host plants (Beck et al., 1991). The TGB counterparts are also found in the genomes of other viruses such as carla-, beny-, pomo-, plena- and hordeiviruses, the products of which are also essential for virus movement (Gilmer et al., 1992; Petty & Jackson, 1990).

The movement of plant viruses from cell to cell is assumed to occur by passage through plasmodesmata. Virus-encoded movement proteins have been shown to be involved in this process (Maule, 1991; Deom et al., 1992). The P30 protein of tobacco mosaic virus (TMV) induces an increase in the size-exclusion limit of plasmodesmata (Wolf et al., 1989) and possesses nonspecific single-stranded RNA-binding activity to shape viral nucleic acids into a transferable form (Citovsky et al., 1990, 1992). The movement proteins of comoviruses (Shanks et al., 1989; van Lent et al., 1990), caulimoviruses (Linstead et al., 1988) and nepoviruses (Wieczorek & Sanfaçon, 1993) participate in the formation of tubular structures extended from plasmodesmata in which virus-like particles are detected. Different from the above-mentioned, the TGBp1 homologues of potexviruses are predominantly associated with cytoplasmic inclusions (Davies et al., 1993; Rouleau et al., 1994; Chang et al., 1997). None of them was detected in plasmodesmata or in tubular structures. Recently, the TGBp1 homologue of white clove mosaic virus (WCIMV) was shown to facilitate the transport of the infectious transcripts in company with capsid protein, and movement of this ribonucleoprotein complex absolutely requires the presence of the other two TGB proteins (Lough et al., 1998).

The TGBp1 encoded by ORF2 of foxtail mosaic potexvirus (FMV) is predominantly fractionated in the soluble fraction (S30) of infected tissue homogenate (Rouleau et al., 1994); however, the TGBp1 homologue of potato virus X (PVX) is mainly in the insoluble P1 and P30 fractions (Davies et al., 1993), and the TGBp1 homologue of bamboo mosaic potexvirus (BaMV) is associated mainly with the cell wall and P30 fractions (Chang et al., 1997). Despite the difference in

Author for correspondence: Ban-Yang Chang.
Fax +886 4 286 1905, e-mail bychang@mail.nchu.edu.tw
Table 1. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>5'-AACTAACCATGGATAACCC-3'</td>
<td>All mutations</td>
</tr>
<tr>
<td>Reverse</td>
<td>5'-CTCTGGTAGCTTCATGCGG-3'</td>
<td>All mutations</td>
</tr>
<tr>
<td>DI-1</td>
<td>5'-CTAACCATGGAT&lt;CAAGAATCGTCGGTGTGCAC-3'</td>
<td>Deletion mutant 1</td>
</tr>
<tr>
<td>DII-1</td>
<td>3'-CTTGGTTTCCTCCCGGCCC&gt;&lt;GAGTGGTAGA-5'</td>
<td>Deletion mutant 2</td>
</tr>
<tr>
<td>DII-2</td>
<td>5'-GGCGGGCGG&gt;&lt;CTTCACTACTACATC-3'</td>
<td>Deletion mutant 3</td>
</tr>
<tr>
<td>DIII-1</td>
<td>5'-CCACCAAAAC&gt;&lt;TGGGAATGTTGATGCTG-3'</td>
<td>Deletion mutant 3</td>
</tr>
<tr>
<td>DIII-2</td>
<td>5'-GATGGGCTGTTGATGCTG-3'</td>
<td>Deletion mutant 4</td>
</tr>
<tr>
<td>DIV-1</td>
<td>5'-GGCGGGCGG&gt;&lt;GTCAAGACTTCGACAC-3'</td>
<td>Deletion mutant 4</td>
</tr>
<tr>
<td>DIV-2</td>
<td>5'-CCCACAAC&gt;&lt;CAAGACATCAGAC-3'</td>
<td>Deletion mutant 5</td>
</tr>
<tr>
<td>DVI-1</td>
<td>5'-GCTGCGGCTAAAGGCTG-3'</td>
<td>Deletion mutant 6</td>
</tr>
<tr>
<td>DVI-2</td>
<td>5'-ATTTGACGA&gt;&lt;ATCTCTAGCCTCTCTG-3'</td>
<td>Deletion mutant 7</td>
</tr>
<tr>
<td>DVI-2</td>
<td>5'-GATGGGCTGTTGATGCTG-3'</td>
<td>Arg-11-Ala substitution</td>
</tr>
<tr>
<td>R11A-1</td>
<td>5'-GGATGAGCTGCTAGACCAGATG-5'</td>
<td>Arg-11-Ala substitution</td>
</tr>
<tr>
<td>R11A-2</td>
<td>5'-CTTACACCACTCACTGACATACATC-3'</td>
<td>Arg-11-Ala substitution</td>
</tr>
<tr>
<td>R16A-1</td>
<td>5'-GATTCTCTGTTATTCTGCAGTTGCAAGAC-5'</td>
<td>Arg-16-Ala substitution</td>
</tr>
<tr>
<td>R16A-2</td>
<td>5'-CCAGCCGGCGGGCGG-3'</td>
<td>Arg-16-Ala substitution</td>
</tr>
<tr>
<td>R16A-2</td>
<td>5'-CCAGCCGGCGGGCGG-3'</td>
<td>Arg-21-Ala substitution</td>
</tr>
<tr>
<td>R16A1-1</td>
<td>5'-CCAGCCGGCGGGCGG-3'</td>
<td>Arg-16,21-Ala substitution</td>
</tr>
<tr>
<td>R16A1-2</td>
<td>5'-CCAGCCGGCGGGCGG-3'</td>
<td>Arg-16,21-Ala substitution</td>
</tr>
</tbody>
</table>

< >, DNA sequences being deleted.
[ ] sites of Arg-to-Ala substitution.

solubility, the TGBp1 homologues of the three potexviruses share high amino acid sequence similarity, suggesting that they have similar biochemical properties. An NTP-binding helicase motif has been found for the three homologues (Bancroft et al., 1991; Skryabin et al., 1988; Lin et al., 1994), and RNA-binding and ATPase activities are also reported for the TGBp1 homologues of FMV and PVX (Kalinina et al., 1996; Rouleau et al., 1994).

A common feature of plant virus movement proteins is that they bind nucleic acids in vitro (Citovsky et al., 1990; Osman et al., 1992; Schoumacher et al., 1992; Tacke et al., 1991; Rouleau et al., 1994; Ivanov et al., 1994; Bleykasten et al., 1996). Distinct domains essential for this binding have been identified in several movement proteins, such as the P30 protein of TMV (Citovsky et al., 1992), the 35 kDa protein of red clover necrotic mosaic virus (RCNMV) (Giesman-Cookmayer & Lommel, 1993), the P3 protein of alfalfa mosaic virus (Schoumacher et al., 1992), the P1 protein of cauliflower mosaic virus (CaMV) (Thomas & Maule, 1995), the 3a protein of cucumber mosaic virus (Vaquero et al., 1997) and the 3a protein of brome mosaic bromovirus (Fujita et al., 1998). However, no RNA-binding domain has been identified for the TGBp1 homologues of potexviruses. Here, we report the RNA-binding domain, the RNA-binding properties and the possible mechanisms leading to RNA binding of the TGBp1 homologue of BaMV.

Methods

Bacterial strains and plasmids. Escherichia coli HMS 174 was used as host for cloning the mutated ORF2 of BaMV-O, while E. coli BL21(DE3) was used as host for overexpression of the mutant proteins. Plasmid pJP1 containing the cDNA of ORF2 (Chang et al., 1997) served as template for construction of the ORF2 derivatives by PCR. Plasmids pBaHB (Lin et al., 1993), pJP1 (Chang et al., 1997), pCT14 and pWC1 (B. Y. Chang, unpublished data), which contain cDNA fragments of the whole genome, ORF2, ORF3 and ORF4 of BaMV, respectively, were linearized and used as templates for in vitro transcription.

Construction of deletion and substitution mutants of TGBp1. Mutations resulting in either amino acid deletions or substitutions in TGBp1 of BaMV were constructed by PCR (Ho et al., 1989; Horton et al., 1990). The linker primer sequences used are shown in Table 1. The forward and reverse primers contain a BamHI and an NcoI site (underlined bases), respectively. After PCR synthesis, the mutated DNA fragments were cut with both BamHI and NcoI and used to replace the wild-type counterpart on pJP1 plasmid. The positions of amino acids
deleted in the mutant proteins (M1–M7) are shown in Fig. 1. The substitution mutants including R11A, R16A, R21A and R16,21A were created by the same PCR method. Oligonucleotides used are also shown in Table 1. Each of the mutant derivatives of both types was shown to have the predicted structure by DNA sequencing and restriction enzyme digestion.

**Overexpression and purification of the wild-type and mutant TGBp1.** *E. coli* BL21(DE3) harbouring a plasmid containing a mutated ORF2 was grown at 37 °C in 2 × YT medium (10 g tryptone, 10 g yeast extract, 5 g NaCl per litre) and induced with 0.4 mM IPTG when the culture had reached an optical density at 550 nm of 0–0.8; 50 min after IPTG induction, rifampicin was added to the culture at a final concentration of 100 µg/ml. The cells were harvested 4 h later by centrifugation at 5000 r.p.m. for 15 min and washed twice with buffer L (10 mM Tris–HCl, pH 8.0, 0.200 mM NaCl, 1 mM EDTA, 10% glycerol, 1 mM PMSF, 1 mM DTT). The washed cells were then homogenized with a French press at 18000 p.s.i. (ca. 124 MPa) followed by centrifugation at 12000 r.p.m. for 10 min. The cell debris, which contained inclusions of the target protein, was washed several times with buffer L and suspended in the same buffer. The target protein was present to about 70% homogeneity at this step. To further purify the protein, an equal volume of sample buffer (0.002% bromophenol blue) was added to the protein suspension. The protein sample was then run in SDS–polyacrylamide gels and stained with 0.25 M KCl (4°C). Gel fragments containing the target protein were excised and chopped into small pieces. The target protein was electroeluted with a Blue tank (Isco) and precipitated with 80% (50 mM Tris–HCl, pH 8.0, 0.200 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF, 10% glycerol) by centrifugation at 12000 r.p.m. for 10 min. The cell debris, which was pelleted and soluble states, the method of photochemical crosslinking between protein and RNA was adopted. As described in the Experimental section, RNA-binding activity has been observed for the insoluble aggregates of CaMV P1 protein (Thomas & Maule, 1995) and the 66 kDa cytoplasmic inclusion of tamarillo mosaic potyvirus (TamMV) (Eagles et al., 1994). In order to determine the RNA-binding activity of TGBp1 in both the pelleted and soluble states, the method of photochemical crosslinking between protein and RNA was adopted. As shown in Fig. 2(a), both monomeric (M) and dimeric (D) forms of the TGBp1 crosslinked to the 32P-labelled RNA transcript of the 3′ noncoding region of BaMV were observed for the suspension of refolded protein (Sup1). Moreover, the soluble fraction (Sup1) possessed most of the RNA-binding activity of the UV crosslinking method. Basically, 10 ng of 32P-labelled RNA transcript (1 × 107 c.p.m./µg) and an appropriate amount of renatured TGBp1 were mixed in buffer containing 10 mM Tris–HCl, 50 mM NaCl, 5 mM MgCl2, 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF, 10% glycerol. The final volume of binding mixture was 12 µl. The mixture was incubated on ice for 15 min before irradiation for about 8 min in a Stratalinker (Stratagene) at 8 cm from the light source (0.78 J/cm²). After UV crosslinking, 10 µg of RNase A was added to the mixture and the sample was further incubated at 37 °C for 30 min to digest uncrosslinked RNA. The resulting sample was then boiled for 5 min with an equal volume of sample buffer (0.125 M Tris–HCl, pH 6.8, 4% SDS, 20% glycerol, 0.002% bromophenol blue) and electrophoresed through a 12.5% SDS–polyacrylamide gel. The gel was stained with Coomassie blue, dried, and autoradiographed after electrophoresis.

**Results**

**RNA-binding properties of TGBp1**

Unlike the TGBp1 homologue of FMV (Rouleau et al., 1994), the TGBp1 encoded by ORF2 of BaMV (Fig. 1a) was fairly low in solubility when overexpressed in *E. coli* (Chang et al., 1997). The solubility of the refolded TGBp1 was also low, about 0.06–0.1 µg/µl depending on buffer pH and salt concentration. RNA-binding activity has been observed for the insoluble aggregates of CaMV P1 protein (Thomas & Maule, 1995) and the 66 kDa cytoplasmic inclusion of tamarillo mosaic potyvirus (TamMV) (Eagles et al., 1994). In order to determine the RNA-binding activity of TGBp1 in both the pelleted and soluble states, the method of photochemical crosslinking between protein and RNA was adopted. As shown in Fig. 2(a), both monomeric (M) and dimeric (D) forms of the TGBp1 crosslinked to the 32P-labelled RNA transcript of the 3′ noncoding region of BaMV were observed for the suspension of refolded protein (Sup1). Moreover, the soluble fraction (Sup1) possessed most of the RNA-binding activity of the UV crosslinking method. Basically, 10 ng of 32P-labelled RNA transcript (1 × 107 c.p.m./µg) and an appropriate amount of renatured TGBp1 were mixed in buffer containing 10 mM Tris–HCl, 50 mM NaCl, 5 mM MgCl2, 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF, 10% glycerol. The final volume of binding mixture was 12 µl. The mixture was incubated on ice for 15 min before irradiation for about 8 min in a Stratalinker (Stratagene) at 8 cm from the light source (0.78 J/cm²). After UV crosslinking, 10 µg of RNase A was added to the mixture and the sample was further incubated at 37 °C for 30 min to digest uncrosslinked RNA. The resulting sample was then boiled for 5 min with an equal volume of sample buffer (0.125 M Tris–HCl, pH 6.8, 4% SDS, 20% glycerol, 0.002% bromophenol blue) and electrophoresed through a 12.5% SDS–polyacrylamide gel. The gel was stained with Coomassie blue, dried, and autoradiographed after electrophoresis.
the refolded protein (Sus) compared with the pellet (P1). Resuspension of P1 did not further solubilize the refolded protein (data not shown); only trace RNA-binding activity was detected in both supernatant (Sup2 and Sup3) and pellet (P3) fractions of the resuspended sample, indicating that only a certain portion of the refolded TGBp1 was soluble and active.

The stability of TGBp1 and RNA complexes was analysed at different salt concentrations (Fig. 2b). The TGBp1 bound RNA maximally at 0 mM NaCl; the binding stability decreased as salt concentration was increased. At 200 mM NaCl, the protein–RNA complexes were almost abolished. We also investigated whether TGBp1 bound preferentially to particular nucleic acid sequences. To accomplish this, four more RNA probes including ORF2, ORF3, ORF4 and the complementary sequence of the 3' noncoding region were prepared. The similarity in binding strength among different RNA transcripts (Fig. 2c) clearly indicated that TGBp1 had no sequence specificity in interaction with BaMV RNA.

Mapping of the RNA-binding domain(s) of TGBp1

To map the RNA-binding domain(s) of TGBp1, we expressed and purified a series of mutant proteins with deletions spanning different regions from the N- to the C-terminal end of the protein (Fig. 1b). All of the mutant proteins overexpressed in E. coli BL21(DE3) had observed molecular masses ranging from 22 to 25 kDa, comparable to those calculated from their amino acid sequences. Moreover, all of them were able to react with antiserum raised against the wild-type TGBp1 (data not shown), suggesting that they were all in the correct reading frame.

Since the wild-type TGBp1 overexpressed in E. coli in the form of inclusion bodies was capable of binding to RNA after urea denaturation and subsequent refolding processes (see Methods), the same strategy was adopted to prepare mutant proteins for RNA-binding assays. Furthermore, the decrease in the RNA-binding activity of TGBp1 at a relatively high salt condition (Fig. 2b) suggested that ionic interaction was one of the major forces involved in TGBp1 and RNA binding. Since the surface charge of a protein might be influenced by buffer pH (the lower the buffer pH, the higher the net surface positive charge), we located the RNA-binding site by analysing the effects of buffer pH and deletion sites on binding of the mutant proteins to RNA (Figs 1 and 3). Clearly, all the mutant proteins were able to bind RNA molecules when the buffer pH was as low as 5.0 or 6.0 (Fig. 3a, b). However, M1 failed to bind RNA as the pH was raised to 7.4 (Fig. 3c); M2, M3 and M7 also lost their RNA-binding activities as the pH was further increased to 8.0. Only M5 and M6, which had relatively higher pI values (9.14 and 9.11, respectively), and M4, with a pI of 8.01, retained their RNA-binding activity at the highest pH (Fig. 3d). The loss of RNA-binding activity of M1, M2, M3 and M7 at pH 8.0 suggested that multiple positively charged regions were responsible for the RNA binding of TGBp1 at elevated pH. However, the exclusive loss of RNA-binding activity of M1 at pH 7.4 indicated that amino acids 3–24 in the N-terminal region of TGBp1 were more critical than those in other regions for RNA binding at physiological pH.
RNA-binding sites of BaMV movement protein

Involvement of three arginine residues in RNA binding

To identify amino acid residues which were responsible for binding of TGBp1 to RNA at physiological pH, the N-terminal 25 amino acids of six TGBp1 homologues of potexviruses were aligned. Besides certain conserved hydrophobic and uncharged amino acids (such as isoleucine, leucine, threonine and glycine), the basic amino acid at position 16 was also conserved among the homologues; it was either arginine or lysine (Fig. 4). We suspected that Arg-16 and possibly also Arg-11 and Arg-21 in the neighbourhood were involved in RNA binding. To test this idea, four mutant proteins with either single (residues 11, 16 or 21) or double (both residues 16 and 21) Arg-to-Ala substitutions were constructed. These
These results indicate that all three arginine residues are essential for RNA binding activity of the M1 protein at pH 7.4, and only trace activity was retained by R11A and R16,21A. Similar to the loss of RNA-binding activity of the wild-type TGBp1, RNA-binding activities of R11A, R16A, R21A and R16, 21A were almost abolished by the four mutant proteins with Arg-to-Ala substitutions shown in Fig. 4 (b). M indicates the monomeric form of TGBp1. The numbers in the right margin indicate the molecular masses (kDa) of protein markers.

Proteins were overexpressed and purified (data not shown). Effects of Arg-to-Ala substitutions on RNA-binding activity of the mutant proteins were analysed. Similar to the loss of RNA-binding activity of the M1 protein at pH 7.4, the RNA-binding activities of R16A and R21A were almost abolished (Fig. 5), and only trace activity was retained by R11A and R16,21A. These results indicate that all three arginine residues are essential for RNA-binding activity of the TGBp1 of BaMV.

**Discussion**

We have been able to test the RNA-binding activity of the BaMV TGBp1 overexpressed in *E. coli* in the form of inclusion bodies. UV crosslinking assays show that most of the RNA-binding activity resides in the soluble fraction of the refolded TGBp1. Studies on the RNA-binding activities of the deletion and substitution mutant proteins have enabled us to locate the RNA-binding domain to the N-terminal region between amino acids 3–24 of TGBp1, within which three arginine residues were shown to be important for RNA binding.

Several approaches such as photochemical crosslinking, gel mobility shifting, filter binding and Northern blotting have been adopted to detect the interaction between protein and RNA, but preliminary experiments revealed that not all of them were suitable for our purpose. However, we found that the photochemical crosslinking method, which trapped the binding complexes, worked well in our case. The stability of protein–RNA complexes with respect to salt concentration is often a criterion to evaluate the binding strength of the complexes (Li & Palukaitis, 1996). The stability of the RNA–BaMV TGBp1 complex is lower than that formed with the more water-soluble TGBp1 homologue of FMV, which remains strongly bound to RNA at 0.2 M NaCl (Rouleau et al., 1994). It is also not as stable as complexes formed with the movement proteins of RCNMV (Osman et al., 1992) and TMV (Cítovsky et al., 1990), which are stable even at 0.4 and 0.6 M NaCl, respectively. However, it is similar to those reported for the TGBp1 homologue of PVX (Kalinina et al., 1996; Karpova et al., 1997). The physiological significance of the solubility of movement proteins or the stability of RNA–movement protein complexes remains unknown.

The proximity of virus aggregates to the cytoplasmic inclusion bodies of TGBp1 homologues in PVX, FMV- and BaMV-infected tissues (Davies et al., 1993; Rouleau et al., 1994; Chang et al., 1997) raises the possibility that these inclusions are the source of active TGBp1. Possibly, active TGBp1 might be continuously released from protein inclusions in soluble form due to chemical equilibration under certain circumstances. Our finding that most of the RNA-binding activity resides in the soluble fraction of the refolded TGBp1 is consistent with this idea (Fig. 2a). However, the detection of only a very low amount of soluble protein and a very low level of RNA-binding activity after resuspension of the pelleted TGBp1 is not consistent with this hypothesis (Fig. 2a). Possibly the protein inclusions in BaMV-infected tissues do not function in the same manner as reported for the insoluble P1 protein of CaMV (Thomas & Maule, 1995) or the 66 kDa cytoplasmic inclusions of TamMV (Eagles et al., 1994). Alternatively, the pelleted TGBp1 may not possess the same structural properties as the protein inclusions in infected tissues due to the improper folding during renaturation.

This is the first report on mapping the RNA-binding site in the TGBp1 homologues of potexviruses. Many different structure motifs responsible for the interaction between protein and RNA have been reported (Burd & Dreyfuss, 1994). They include the RNP motif, the arginine-rich motif (ARM), the KH homology (KH) motif, the RGG box and the double-stranded RNA-binding motif (DSRM). The RNA-binding domain of TGBp1 was mapped to the N-terminal 24 amino acids, a region within which no such typical structure motif was observed. However, the involvement of three surrounding arginine residues in RNA binding (Fig. 5) suggests that the TGBp1 of BaMV possesses an ARM-like structure. There are probably nonspecific ionic interactions between the positively charged arginines of TGBp1 and the negatively charged phosphate groups on RNA, as well as specific hydrogen bonding networks between arginines and the RNA sugar-phosphate backbone and bases (Burd & Dreyfuss, 1994). These interactions appear to require a certain structure(s), rather than just particular amino acids, since each of the four Arg-to-Ala substitutions tested destroys most of the RNA-binding activity of TGBp1 (Fig. 5). The predicted absence of α-helix or β-structure in the N-terminal 24 amino acids (Chou & Fasman, 1974) suggests that a loop or bulge structure(s) is present in this region and is important to RNA binding.
This research was supported by National Science Council of the Republic of China grant NSC 86-2311-B-005-025-B11.

References


Received 19 November 1998; Accepted 26 January 1999