The tobacco necrosis virus p7a protein is a nucleic acid-binding protein

Samuel K. Offei, Robert S. Coffin and Robert H. A. Coutts*

Biology Department, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BB, UK

The two centrally located open reading frames (ORFs) of necroviruses may, by analogy with the similarly located and related ORFs of carmoviruses, be expected to have a function in virus movement in plants. In the case of tobacco necrosis virus (TNV) strain D these proteins both have a molecular mass of approximately 7 kDa and are thus known as p7a and p7b. We over-expressed p7a in Escherichia coli, separated it from bacterial proteins and renatured it on gels, and showed that p7a was able to bind single-stranded RNA and single-stranded DNA, but was unable to bind double-stranded DNA. These protein–nucleic acid complexes were stable at moderately high salt concentrations. Protein p7b could not be expressed in a number of bacterial systems. We speculate that in TNV, unlike some other viruses which encode a single movement protein with separate functional domains for RNA binding and plasmodesmatal targeting, pVa and pVb may respectively provide these functions on separate proteins.

Tobacco necrosis virus (TNV) is a small icosahedral plant virus, belonging to the necrovirus group, with a single, positive-sense RNA of about 4 kb. The complete nucleotide sequence of the genome of TNV strain D (TNV-D; Coutts et al., 1991) reveals four open reading frames (ORFs); a 5'-proximal ORF encoding a protein with a molecular mass of 22 kDa which may be read through to produce a protein with a molecular mass of 82 kDa with the characteristics of a putative RNA-dependent RNA polymerase, two centrally located ORFs encoding two out-of-frame proteins which both have a molecular mass of 7 kDa (p7a and p7b) and the 3'-proximal coat protein with a molecular mass of 29 kDa. The genome organization of TNV-D is similar to both TNV-A (Meulewaeter et al., 1990) and the carmovirus turnip crinkle virus (TCV; Carrington et al., 1989). The TNV-D p7a and p7b proteins show sequence similarity with the corresponding ORFs of both TNV-A (with molecular masses of 8 kDa and 6 kDa, respectively) and TCV (with molecular masses of 8 kDa and 9 kDa, respectively) (Courts et al., 1991). For both TNV strains these ORFs are probably expressed from a single subgenomic RNA of about 1.5 kb (Meulewaeter et al., 1990, 1992; Coutts et al., 1991; S. K. Offei & R. S. Coutts, unpublished) either by an internal initiation or frame shifting mechanism, or a mixture of the two (see later). In TCV the 8K and 9K proteins are both required for the cell-to-cell movement of the virus (Hacker et al., 1992), and thus it is likely that the corresponding TNV proteins play a similar role.

The cell-to-cell movement of viruses is not a passive process but requires a virus-encoded movement protein (MP), which in the case of tobacco mosaic virus (TMV) in vivo is located in the plasmodesmata (Tomenius et al., 1987) and also binds single-stranded (ss) RNA to produce an unfolded structure suitable for transport between cells (Citovsky et al., 1990, 1992). The protein is thought to increase the size exclusion limit for plasmodesmal movement and thus facilitate transport of the viral RNA–MP complex (Wolf et al., 1989). The TMV MP binds ss nucleic acids in a cooperative manner (Citovsky et al., 1990) and several other plant virus proteins thought to be involved in movement also bind ss nucleic acids in a similar fashion. With these observations in mind we have investigated the nucleic acid binding capabilities of the basic TNV-D p7a protein as a start in the investigation of the possible role of p7a and p7b in virus movement.

The p7a gene was cloned using RT–PCR and purified TNV-D RNA (Coutts et al., 1988) which was denatured for 5 min at 65 °C and quenched on ice immediately before cDNA synthesis. An oligonucleotide (5' GAAATGGATCCCGTTGATGAGCTCTAATGTA 3') complementary to the 3' end of the gene but including downstream mis-matches in the sequence to introduce a unique BamHI site (underlined) was used for first-strand synthesis and together with a second oligonucleotide (5' AGATCGATCCATATGGAAAATA 3'), comple-
incillin (Sambrook et al., 1989) to give pETp7a. The structure of the construct and the complete inserted sequence was confirmed by dideoxy sequencing (Sanger et al., 1977).

Proteins were expressed using E. coli strain BL21(DE3) pLysE (Studier & Moffat, 1986) and induced with isopropyl-β-D-thiogalactopyranoside (IPTG). Briefly, a 20 ml culture of the transformed bacteria was grown overnight at 37 °C in 2TY containing 150 μg/ml ampicillin (Sambrook et al., 1989). The culture was diluted twofold with fresh medium (including ampicillin) and IPTG added to 0.4 mM to induce expression of the inserted gene. The cells were grown for a further 3 h at 37 °C when 1 ml aliquots of the culture were harvested by centrifugation, lysed in 200 μl of sample buffer (20% glycerol, 2% SDS, 0.7 M-2-mercaptoethanol, 12.5 mM-Tris-HCl pH 6.8, 0.1 mg/ml bromophenol blue) and the proteins separated by SDS denaturing electrophoresis on 18% polyacrylamide gels. A band of the expected size for the p7a protein was present in extracts of E. coli cells transformed with pETp7a which was absent from uninduced cells and cells transformed with pET-3a (Fig. 1).

Unfortunately, unlike many other proteins expressed using the pET system, p7a did not form insoluble aggregates after induction and thus could not be easily purified, and all attempts to purify it using a variety of procedures based on those described by Citovsky et al. (1990) were unsuccessful. Therefore, all nucleic acid binding experiments were conducted using the crude extracts described and separated as above. Following electrophoresis, proteins were renatured by soaking the gel in 192 mM-glycine and 20% methanol in 25 mM-Tris-HCl pH 8.3 for 1 h and then transferred to nitrocellulose by standard procedures. The concentrations of p7a in the extracts were roughly estimated by excising the p7a band from Coomassie Blue stained gels, electroeluting the stained band and assaying by the Bradford method (Bio-Rad).

Radioactively labelled RNA probes were produced by in vitro transcription of a 400 bp pBluescript clone corresponding to nucleotides (nt) 1631–2030 in the TNV-D sequence (Coutts et al., 1991) in the presence of [α-32P]UTP (Amersham) as described previously (Osman et al., 1992). Radioactive double-stranded (ds) DNA probes, corresponding to nt 1–348 of TNV-D RNA, were produced by RT–PCR as above using suitable primers, but including 10 μCi [α-32P]dCTP (Amersham). These were purified using Magic PCR Preps (Promega) according to the manufacturer’s instructions. Radioactively labelled ssDNA was prepared by reverse transcription of TNV-D RNA from nt 1–348 using Superscript reverse transcriptase (Gibco-BRL) and a complementary primer, and including 10 μCi [α-32P]dCTP in the reaction mixture. The cDNA was treated with 300 mM-NaOH to remove RNA, neutralized with 300 mM-HCl, phenol-extracted and ethanol-precipitated.

In initial experiments about 50 ng of labelled ss- or dsDNA was incubated with nitrocellulose filters onto which various concentrations of protein had been transferred, as described by Mellor et al. (1985) but with minor modifications. In these experiments both ss- and dsDNA bound efficiently to p7a. However, as many of the bacterial proteins also showed a similar binding pattern, and although binding would be expected for some bacterial proteins, these interactions could have
have been due to a non-specific process. Thus in subsequent experiments probe concentration was decreased to 1 ng allowing a clear discrimination between the binding capabilities of the bacterial proteins and of p7a to be shown, with efficient binding of ssRNA and ssDNA (Fig. 2) but not now of dsDNA (result not shown). In further experiments the specificity of the p7a–RNA interaction was explored using a range of salt concentrations by including different concentrations of NaCl in the binding and washing buffers. Strong signals were obtained with 20–200 mM-NaCl, with a weaker signal at 300 mM and no signal at 400 mM or higher NaCl concentrations (Fig. 3).

These observations indicate that p7a stably interacts with RNA up to salt concentrations of between 200 and 300 mM in vitro and imply that a stable p7a–RNA complex could form at physiological salt concentrations in vivo.

The exact biological role of RNA binding by p7a cannot be inferred from these results, but as discussed above it is likely that the ORF has a role in virus cell-to-cell movement. The differing affinity of p7a for ss and ds nucleic acids is similar to the binding capabilities of the TMV MP and of some other MPs (Citovsky et al., 1990; Osman et al., 1992; Brantley & Hunt 1993; Soumounou & Laliberté, 1994). Unfortunately, although it would have been of interest to compare the properties of p7a and p7b, when p7b was cloned into pET-3a in a similar fashion it failed to express the protein upon induction. Expression was also unsuccessful using a glutathione S-transferase fusion system, a His-tag vector and the pMAL system (New England Biolabs). The reasons for this lack of success are unknown, although it seems likely that the expressed protein is toxic to E. coli.

Although a comparison of MPs from several different groups of plant viruses has revealed the presence of conserved amino acid motifs (Koonin et al., 1991), these cannot be found in the TNV-D 7K proteins (Coutts et al., 1991) or in the related proteins of TNV-A or TCV, implying that if indeed these proteins do fulfill the movement function for these viruses then they may represent a separate evolutionary lineage. However, it would be unsurprising if differing solutions to the problem of movement in plants had not evolved, with, in this case, a possible separation of the necessary functions of RNA binding and plasmodesmatal targeting into separate proteins, allowing differential regulation of the levels of expression of these functions. Thus it is of interest that in all the viruses discussed here and in several other viruses with similar centrally located ORFs (which are as yet, in many cases, of unknown function), these genes appear to be expressed from a single RNA transcript, with no adequate method of expression of the internal ORFs having been demonstrated. Recent work with the carlaviruses potato virus M has shown that the two 3'-terminal ORFs, probably translated from a single sub-genomic RNA, are separated by a sequence responsible for producing a transframe fusion of the two proteins (Gramstat et al., 1994). The second protein is...
also produced by an internal initiation mechanism. These ‘shifty’ sequences, consisting only of AAAA followed by a stop codon, were mutated to either UUUU, GGGG or CCCC before the stop codon, with UUUU increasing the efficiency of the slippage but GGGG or CCCC decreasing the efficiency dramatically (Grammat et al., 1994). An examination of the sequences at the borders between p7a and the out of frame p7b ORF and at the borders of the corresponding genes of TNV-A, TNV-NE, TCV (Carrington et al., 1989) and melon necrotic spot carmovirus (MNSV) shows in each case sequences identical or nearly identical to the sequence giving maximum transframe fusion above (i.e. TNV-D, UUUUUAA; TNV-A, CUUUUAA; TNV-NE, CUUUUAG; TCV, CUUCUAA; MNSV, UUUU-UAG - in each case the stop codon is underlined). Thus a possible mechanism of expression can be proposed where, in the case of TNV-D, p7a is expressed both singly and as a translational fusion with p7b, and p7b as well as being expressed as a fusion with p7a is expressed separately by internal ribosome entry. This would allow the levels of the three protein types (p7a, p7b and p7a/b) to be differentially regulated, thus allowing expression of higher levels of p7a for RNA binding and unfolding and lower levels of p7b for possible plasmodesmal enlargement and of p7a/b for potential targeting of unfolded RNA-protein complexes to the plasmodesmata. These hypotheses, together with attempts to express p7b in eukaryotic cells, are currently under investigation.

S.K.O. would like to thank the Association of Commonwealth Universities for a grant to complete his doctorate at the University of London. Sequence analysis was performed using the SEQNET computer facility of the BBSRC Daresbury Laboratory. R.S.C. thanks MAFF for support. We also acknowledge an equipment grant from the Computer facility of the BBSRC Daresbury Laboratory. R. S. C. thanks MAFF for support. We also acknowledge an equipment grant from the University of London Central Research Fund.

References


