RNA 2 of *Citrus psorosis virus* is of negative polarity and has a single open reading frame in its complementary strand

María Eugenia Sánchez de la Torre¹, Carmelo López², Oscar Grau¹ and María Laura García¹

¹Instituto de Bioquímica y Biología Molecular, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Calles 47 y 115–1900 La Plata, Argentina
²Instituto de Biología Molecular y Celular de Plantas (UPV-CSIC), Universidad Politécnica de Valencia, Spain

*Citrus psorosis virus* (CPsV) causes a citrus disease occurring worldwide. Isolate CPV 4 has a genome with three single-stranded RNAs. The complete sequence of RNA 2 (1643 nucleotides) is reported here. Northern blot hybridization with strand-specific probes showed that most of the encapsidated RNA 2 is of negative polarity, although a small amount of the complementary strand may also be present in particles. The RNA 2 complementary strand contained a single open reading frame encoding a protein of 476 amino acids, which includes a motif resembling a nuclear localization signal. The sequence of this putative protein shows no significant similarity to any other in the databases. In the 3'-terminal untranslated region there is a putative polyadenylation signal. No subgenomic RNAs derived from RNA 2 were detected.

Psorosis is a serious disease affecting citrus in many countries (Roistacher, 1993). In Argentina, it seems to be spread by an unknown vector (Beñatena & Portillo, 1984) and it causes important losses (Danos, 1990). The presumed causal agent, *Citrus psorosis virus* (CPsV), is the type member of the genus *Ophiovirus* (Milne et al., 2000) but the low concentration of virus particles in citrus tissues and their instability has so far impeded characterization of the genome. All species of the genus, *Ranunculus white mottle virus* (RWMV), *Tulip mild mottle mosaic virus* (TMMMV) and *Mirafiori lettuce virus* (MiLV), have similar morphology: circular, filamentous naked nucleocapside about 3 nm in diameter of different lengths. Limited information is available about their genomic structure: RWMV has at least three RNAs of approximately 7.5, 1.8 and 1.5 kb and a coat protein of 43 kDa (Vaira et al., 1997); TMMMV presents a 47 kDa coat protein (Morikawa et al., 1995) and MiLV also contains three RNAs of 8.5, 1.9 and 1.7 kb and a coat protein of 48 kDa (Roggero et al., 2000).

Although the morphology of ophioviruses resembles that of the nucleocapsids of bunyaviruses and tenuiviruses, their RNAs 3 and 2 (as shown in this paper) have no sequence homology with them or with any other negative or ambisense viruses and their coat proteins have no serological relationship (Vaira et al., 1997; Milne et al., 2000). CPsV particles can be separated in a sucrose gradient into two components, ‘top’ (T) and ‘bottom’ (B), which individually are not infectious (Derrick et al., 1988; García et al., 1991). Particles from both components share a single coat protein and they have a similar highly sinuous morphology, but B particles have contour lengths about five times larger than T particles (Derrick et al., 1988; García et al., 1991, 1994). The CPV 4 isolate of CPsV has been reported to contain one large RNA in the B component and a small single-stranded RNA in the T component (Derrick et al., 1991). More recent results have confirmed that RNA 1 is in the B component, but Northern blot hybridization indicated that in the T component there are two RNAs (RNA 2 and RNA 3) of approximately 1650 and 1500 nucleotides, respectively (Sánchez de la Torre et al., 1998). RNA 3 codes for the coat protein (Sánchez de la Torre et al., 1998; Barthe et al., 1998), but the sequences and possible functions of the other RNAs are unknown. Here we report the molecular characterization of CPsV RNA 2, which is of negative polarity. The complementary strand of this RNA has a single open reading frame (ORF) potentially encoding a protein of unknown function with a putative nuclear localization signal (NLS).

The CPV 4 isolate (Garney & Timmer, 1980) was mechanically transmitted from *Citrus sinensis* to *Chenopodium quinoa* leaves, which developed local lesions. Virus particles were partially purified from those lesions by differential and sucrose gradient centrifugation as described by Garcia et al. (1991). RNA extracted from the T component (T RNA) was used to prepare a cDNA library with the Librarian II cloning kit (Invitrogen). The sequences of 55 clones, determined in both directions with an automatic ABI Prism apparatus, were assembled in two contigs using the GCG software package.
Fig. 1. Nucleotide and predicted amino acid sequences of the complementary strand (positive polarity) of CPsV RNA 2. The nucleotides corresponding to the bipartite nuclear targeting sequence are underlined, and the conserved amino acids of that sequence are in bold. A 23 nucleotide sequence including a putative polyadenylation signal in bold is boxed.

Once it was established that the encapsidated viral RNA 2 was of negative polarity (see below), two strategies were used to search for its two termini. The 3' end of RNA 2 was determined after polyadenylation of RNA from the T component with yeast poly(A) polymerase (Amersham). First-strand cDNA was synthesized with primer PM1 (5' CCAGGATCCCTCTAGAAGCCGCCTGAT7'V-3'; where V is A, C or G) (kindly provided by Pedro Moreno, IVIA, Spain) and SuperScript II reverse transcriptase (Gibco BRL), and then PCR amplifications with Taq DNA polymerase a touchdown cycling was carried out in a Perkin Elmer 2400 Thermocycler:
five cycles for 10 s at 94 °C, 5 s at 70 °C and 45 s at 72 °C, followed by 30 cycles of 10 s at 94 °C, 5 s at 68 °C and 45 s at 72 °C, with a final extension for 5 min at 72 °C. PCR products were cloned in plasmid pGEM-T (Promega) and sequenced as described above. Ten clones obtained by this procedure extended the sequence up to 1643 nucleotides.

The 5′ end of RNA 2 was determined using two strategies. The 5′ end of positive RNA 2 was polyadenylated and the first strand of cDNA was synthesised with primer PM1 and PCR amplified with primers Pr260 (5′ GGACAGCGGATGAAGGAA 3′), homologous to positions 1350–1369 (Fig. 1), and PM1 under the following conditions: 35 cycles for 45 s at 94 °C, 45 s at 55 °C and 2 min at 72 °C, with a final extension for 10 min at 72 °C. PCR products were cloned in plasmid pGEM-T (Promega) and sequenced as described above. Six clones were obtained, four inserts ended at position 1643, one had an extra C, and another a T at position 1644.

2. The 5′ end of RNA 2 was studied using the Smart PCR cDNA Synthesis Kit (Clontech) and primers Smart (Clontech) and Stop (5′ GAGTATTTGATAGATCGAAGGAGTGAG 3′), homologous to positions 1432–1461 in the positive strand of RNA 2 (Fig. 1). Cloning and sequencing were performed as above. Four clones had the same sequence up to position 1643; however, an extra C cannot be ruled out because of the use of the Smart strategy Cs are added at the end of the first cDNA strand.

The 5′ end of the RNA 2 was further confirmed by primer extension using primer Stop. A single band of the predicted size was obtained (Fig. 2). The sequence of the insert used for comparison contains Cs at positions 1643/1644 but this could be a consequence of the Smart strategy.

The consensus sequence of RNA 2 was therefore derived in each region from 9 to 27 overlapping sequences.

No complementarity between the 5′- and 3′-terminal sequences was observed, in spite of the circular structures observed by electron microscopy of CPsV particles (García et al., 1994), which suggested a possible ‘panhandle’ structure as in other viral RNAs. In the case of CPsV, the 5′ ends of the complementary strand of RNA 1 (unpublished results), RNA 2 and RNA 3 are identical, GATAC(T) 7, thus reinforcing the probability that this sequence is the 5′ terminus. The same results have been also obtained with RNAs 1, 2 and 3 from a Spanish CPsV isolate, giving further credence to this view (Pedro Moreno, personal communication). The circular structures observed must be maintained by some other type of interaction.

Northern blot hybridizations with riboprobes derived from cloned RNA 2 cDNA were used to investigate the polarity of the encapsidated form of this RNA and to explore the possible existence of subgenomic RNAs in purified virions (T RNA) and crude extracts (total RNA), as previously described (Sánchez de la Torre et al., 1998). One intense band of the expected size was observed in the total RNA fraction from infected tissue when hybridized with a riboprobe containing a region of the coding sequence of RNA 2 (Fig. 3A). This band was not
present in the healthy control thus confirming its viral origin. Using a second riboprobe of the complementary polarity, one single band of the same size but significantly less intense was observed (Fig. 3B). A weaker band of larger size that was detected in extracts from both infected and healthy tissue was considered non-specific. A similar hybridization pattern was found using T RNA instead of total RNA (Fig. 3C, D). An intense signal was generated with the riboprobe detecting negative strands whereas the complementary riboprobe produced only a faint signal. Therefore, we conclude that CPsV RNA 2 is of negative polarity. No subgenomic RNAs were found in either T or total RNA preparations indicating that the single ORF (see below) is translated from the positive strand of the genomic RNA.

Using the Translate program, a single ORF was found in the complementary strand of RNA 2, potentially encoding a 476 amino acid protein with an $M_0$ of 53694 (Fig. 1). The first AUG codon was found at positions 46–48 and a second one, in-frame, at positions 52–54. Based on the consensus sequences flanking the translation start codon of eukaryotic mRNAs [an A and a G at positions −3 and +4, respectively; Kozak (1989) and Lutcke et al. (1987)], it is not possible to decide which is the actual initiation codon. An A at position −3 precedes the first AUG which is followed by a U, instead of a G, at position +4. The second AUG has a G at position +4 but it is preceded by a U, instead of an A, at position −3. Both AUGs have a C residue at position −2, which is also present in Kozak’s and Lutcke’s consensus.

Comparison of the sequences of all clones indicated very limited and dispersed variability. Eight changes were located in the ORF, none of them producing a stop codon, and five causing changes in the amino acid sequence. The remaining two variations were located in the 3′ UTR.

The hydropathy profile of the putative protein coded by RNA 2, obtained according to Kyte & Doolittle (1982), showed that both terminal regions are hydrophilic, as well as six internal regions. No significant hydrophobic regions were observed, suggesting that this is not a transmembrane protein. Since this protein has no significant similarity to any other in databases (GenBank/EMBL, SWISSPROT and PIR, using FASTA, BLAST and PROSITE), we cannot speculate on its function. However, most likely it does not have a structural role because only a 48 kDa protein is found in purified particles (Garcia et al., 1991).

In the amino acid sequence, two motifs located between positions 86–102 and 255–271 bear similarity to a putative ‘bipartite nuclear targeting sequence’, which is considered to be a nuclear localization signal (NLS) (Nigg, 1997). These motifs contain two adjacent basic amino acids, Arg or Lys, a spacer region of 10 residues, and at least three basic residues in the next five positions after the spacer region (Fig. 1, underlined). These karyophilic signals are present in the N protein of the negative-stranded non-segmented viruses that replicate in the host nucleus, like Sonchus yellow net virus (Martins et al., 1998) and Borna disease virus (Pyper & Gartner, 1997). In the case of influenza virus, with a negative-stranded segmented genome, a nonconventional NLS has been found in the viral nucleoprotein (Wang et al., 1997).

The positive strand of RNA 2 has 5′ and 3′ UTRs of 45 and 167 nucleotides, respectively, flanking the ORF (Fig. 1). The 3′ UTR has a 23 nucleotide sequence located 50 nucleotides downstream of the stop codon (Fig. 1, boxed) that is also found essentially unchanged in the 3′ UTR of the positive strand of RNA 3, at 29 residues downstream of the stop codon (Sánchez de la Torre et al., 1998). The presence of this 23 nucleotide stretch in the 3′ UTR of the positive strand of RNAs 2 and 3 suggests its involvement in regulating replication or stabilization. Within these conserved 23 nucleotides there is an AAUAAA sequence, which is a highly conserved polyadenylation signal in eukaryotes (Zhao et al., 1999) that is also present in several negative-stranded viruses (Matthews, 1991), indicating that it might function as a polyadenylation signal in both RNA 2 and 3.

In summary CPsV, the type species of the Ophiovirus genus, is a multipartite RNA virus with three RNAs. At least two of them are of negative polarity, encoding in their complementary strands a single protein: the coat protein (RNA 3) and a 54 kDa protein with a putative NLS (RNA 2).

We thank L. W. Timmer and K. S. Derrick for the CPV 4 isolate; P. Moreno for the PM1 primer and for critical reading and providing unpublished information; N. Costa for citrus seedlings and G. Chiarrone for greenhouse work. The authors are indebted to R. Flores and R. G. Milne for critically reading the manuscript. M. L. G. belongs to the staff of Facultad de Ciencias Exactas, UNLP; O. G. is a recipient of the research career award from CICBA, and M. L. G. from CONICET. M. E. S. belongs to the staff of Facultad de Ciencias Agrarias y Forestales, UNLP. This work was supported by grants BID802 OC-AR PID 332, SECyT-CONICET, UE contract ERBCIICT960044 and CICBA 1454/97.

References


Citrus psorosis virus RNA 2


Received 26 February 2002; Accepted 25 March 2002