Sequence analysis and product assignment of segment 7 of the rice dwarf virus genome

Kazuo Nakashima,1 Tetsuji Kakutani2 and Yuzo Minobe2*

1Tropical Agriculture Research Center and 2National Institute of Agrobiological Resources, Tsukuba, Ibaraki 305, Japan

The complete nucleotide sequence of segment 7 of the rice dwarf virus (RDV) genome was determined. The segment was 1696 bp long and its plus-strand terminal sequence, 5' GGCAAA---UGAU 3', was in agreement with the consensus sequence previously found in other segments of RDV. A 10 bp inverted repeat was found adjacent to the termini. A single long open reading frame extended for 1518 bp from the first AUG triplet (positions 26 to 28), and encoded a polypeptide of 506 amino acids (M, 55 339). This protein had 32% identity in the amino acid sequence to the 57K protein encoded by segment 7 of the wound tumour virus genome. The translation product of transcript RNA made from 'tailored' cDNA of RDV segment 7 comigrated with the 60K core protein of RDV in 10% polyacrylamide gel and reacted with antiserum against the 60K core protein of RDV. Segment 7 of the RDV genome therefore codes for the 60K core protein.

Rice dwarf virus (RDV) (Phytoreovirus subgroup) consists of 12 genome segments of dsRNA (Fuji-Kawata et al., 1970; Kimura et al., 1987) and seven structural proteins (Nakata et al., 1978; Kimura et al., 1987). These proteins are encoded by the viral genomic RNAs and include those with polymerase activity (Kodama & Suzuki, 1973). The protein product of segment 7 was previously predicted to be the 60 K core protein of RDV (Nakata et al., 1978) because of its size (about 1600 bp). Here, we report on the results of sequence analysis and expression in vitro of segment 7.

A cDNA library of RDV genome segments (Minobe, 1987) was constructed as described by Cashdollar et al. (1982). Twenty-seven transformants containing cDNA of segment 7 were selected by colony hybridization using the purified 32p-labelled segment 7 dsRNA as a probe. Seven recombinant plasmids had inserts of approximately 1800 bp, which correspond in size to full-length copies of segment 7 with the tail attached during the process of cloning. One of these cDNAs was subcloned into M13 mp18 in both orientations. This insert was digested from one end by exonuclease III to various extents, so that a series of unidirectionally deleted cDNAs were formed (Henikoff, 1984). Deleted cDNAs thus obtained were sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977), using 2’-deoxy-7-deazaguanosine triphosphate instead of dGTP (Mizusawa et al., 1986). The entire sequence of the cDNA was analysed in both orientations. The 5'- and 3'-terminal nucleotide sequences were confirmed by primer extension of synthetic oligonucleotides, 5’ AATTTATAACAATTAGAAA 3’ at nucleotides 96 to 114 and 5’ AACATTATATGCTCCCGGTTT 3’ at nucleotides 1583 to 1604, with four independent cDNAs of segment 7, respectively. Since the 5’-terminal sequence showed a compression, we added 0.1 μg of T4 gene 32 protein (Boehringer) for 20 min at 65 °C, after termination of the reaction.

The nucleotide sequence of segment 7 is shown in Fig. 1(a). The segment was 1696 bp long and its plus-strand terminal sequence, 5’ GGCAAA---UGAU 3’, was in agreement with the consensus sequence plus-strand 5’ GG[AAA---]GAAA 3’, at nucleotides 96 to 114 and 5’ AACATTATATGCTCCCGGTTT 3’ at nucleotides 1583 to 1604, with four independent cDNAs of segment 7, respectively. Since the 5’-terminal sequence showed a compression, we added 0.1 μg of Proteinase K (Merck) for 20 min at 65 °C, after termination of the reaction.

The nuclidean sequence of segment 7 is shown in Fig. 1(a). The segment was 1696 bp long and its plus-strand terminal sequence, 5’ GGCAAA---UGAU 3’, was in agreement with the consensus sequence plus-strand 5’ GG[AAA---]GAAA 3’, previously found in six other RDV genome segments (Fukumoto et al., 1989; N. Hayashi, T. Omura & Y. Minobe, unpublished results; Minobe, 1987; Omura et al., 1988, 1989). A 10 bp inverted repeat was found adjacent to the termini, as shown in Fig. 1(b).

The terminal consensus sequence is similar to that of wound tumour virus (WTV), also a member of the Phytoreovirus subgroup, namely plus-strand 5’ GGUAUU---UGAU 3’. Inverted repeats were also observed in the segments of WTV (Anzola et al., 1987). The consensus terminal sequences and the inverted repeats of RDV may play roles in the sorting and assembly of RNA genome segments, as previously suggested for WTV by Anzola et al. (1987).

Segment 7 of RDV had one long open reading frame, which started from the first AUG triplet (positions 26 to
and extended for 1518 bp. The sequence context of the initiation codon is GNNAUGU, which is in agreement with the proposed consensus initiation sequence (Kozak, 1987). The open reading frame encoded a polypeptide of 506 amino acids (Mr less than 8526). Reading frames in the minus-sense strand RNA only) and predicted amino acid sequence of the translation product of the large open reading frame. The consensus terminal 5' hexanucleotide and 3' tetranucleotide sequences found in several other RDV genome segments are underlined. (b) The inverted repeat of 28) and extended for 1518 bp. The sequence context of the initiation codon is GNNAUGU, which is in agreement with the proposed consensus initiation sequence (Kozak, 1987). The open reading frame encoded a polypeptide of 506 amino acids (Mr, 55339; hereafter referred to as the 55K protein). The predicted amino acid sequence was compared with a protein sequence library (NBRF database) but no similar protein was found. Other reading frames code for small polypeptides of Mr less than 8526. Reading frames in the minus-sense also code for small polypeptides of Mr, less than 11361.

During the preparation of this manuscript the complete nucleotide sequence of segment 7 of WTV and the predicted amino acid sequence of the 57K encoded by
this segment were reported (Anzola et al., 1989). The two proteins are similar in sequence (Fig. 2a), with 32% of residues being identical when aligned by a computer program (DNASIS; Hitachi). The greatest similarity was in the regions between amino acids 61 and 141, in which 57% were identical (Fig. 2b). This suggests that the 55K protein of RDV and the 57K protein of WTV are evolutionarily related and that their functions are similar.

Xu et al. (1987) reported that the presence of complimentary homopolymer regions 5' and 3' to the coding RNA of WTV causes inefficient in vitro translation. 'Tailored' cDNA of segment 7 of RDV lacking homopolymer regions was prepared by the polymerase chain reaction method (Saiki et al., 1988). The reaction mixtures (100 μl) contained 0.1 ng segment 7 cDNA, 1 μg each of 5' GGTACCGGCAAAAAGCTCATTGC 3' and 5' GAGCTCATATTAAAGCTCATTCC 3' (underlined sequences are KpnI and SacI cutting sites added to the terminal sequences of segment 7 of RDV). The mixture was subjected to 25 cycles of 1 min at 92 °C and 2 min at 42 °C, followed by 5 min at 72 °C. The 'tailored' cDNA of segment 7 was subcloned between bacterial T7 and T3 promoters of a transcription vector (Bluescript; Stratagene). After linearization by digestion with KpnI or SacI and treatment with 200 μg/ml proteinase K (Merck), DNA (1 μg) was incubated with 40 mM-Tris–HCl pH 8.0, 8 mM-MgCl₂, 2 mM-spermidine, 50 mM-NaCl, each rNTP at 0-4 mM, 30 mM-DTT, 40 units of RNase inhibitor and 10 units of either T3 or T7 polymerase for 1 h at 37 °C.

Transcribed RNAs (10 to 20 μg/μg DNA) of each strand were about 1800 nucleotides in length. Translation in vitro of the transcripts (1 μg) was in 22 μl of rabbit reticulocyte lysate (Amersham) and 4 μl of L-[³⁵S]methionine (37 TBq/mmol) for 1 h at 37 °C, as described by Pelham & Jackson (1976). The major translation product of the plus-strand RNA comigrated with the 60K core protein of RDV during electrophoresis in 10% polyacrylamide gel, prepared and used according to Laemmli (1970) (Fig. 3b, lane 3). When the transcribed RNA with the homopolymer tail attached was used as a template the major in vitro translation product was smaller than the 60K core protein (about Mr 50K, data not shown) as previously reported for non-tailed WTV cDNA transcripts by Xu et al. (1987). There were no in vitro translation products specific to the minus-strand RNA transcript (Fig. 3b, lane 4).

In order to prepare antibodies against the 60K core protein, the structural proteins of RDV were separated by SDS–PAGE. The 60K core protein was eluted from gels by shaking overnight with distilled water, mixed with Freund’s complete adjuvant and injected into a mouse. The antiserum thus obtained reacted only with 60K core protein of RDV by the immunoblotting method (Towbin et al., 1979) (Fig. 3a). When 10 μl of translation products of plus-strand RNA were immunoprecipitated (Olliver & Boyd, 1984) with 0.1 μl of the antiserum and analysed by SDS–PAGE, the 60K protein was found (Fig. 3b, lane 6). Several other proteins smaller than the 60K core protein were also observed, which we presume are the result of degradation during immunoprecipitation. No proteins produced by the minus-strand RNA transcript were immunoprecipitated with the antiserum against the viral 60K protein (Fig. 3b, lane 7). Therefore, segment 7 of RDV appears to code for the 60K core protein.

Xu et al. (1989) recently reported that the polypeptide synthesized in vitro by a transcript generated in vitro from
the tailored cDNA of segment 7 of the WTV genome comigrated in polyacrylamide gel with the viral 57K core protein of WTV. They concluded that the 57K core protein of WTV was the product of segment 7. Therefore, each of the viral segments named segment 7 from the two related viruses, RDV and WTV, codes for their respective viral core proteins. Although these proteins are believed to play important roles in viral RNA transcription, the function of the 60K protein of RDV remains to be seen.

We thank Drs Miyoji Sugiura and Yuriko Hayano for advice and encouragement and Ken-ichi Higo for critical reading of the manuscript.

References


(Received 11 July 1989; Accepted 7 November 1989)