Nucleotide sequences of genome segments S8, encoding a capsid protein, and S10, encoding a 36K protein, of rice gall dwarf virus

Hiroaki Noda, Koichi Ishikawa, Hiroyuki Hibino, Hajime Kato and Toshihiro Omura

1National Institute of Sericultural and Entomological Science and 2National Agriculture Research Center, Tsukuba, Ibaraki 305, Japan

The nucleotide sequences of DNAs complementary to the eighth (S8) and the tenth (S10) largest of the 12 genome segments of rice gall dwarf virus (RGDV) were determined. The S8 and S10 segments consist of 1578 and 1198 nucleotides, each with a single open reading frame extending for 1278 nucleotides from nucleotide 21, and 960 nucleotides from nucleotide 22, respectively. S8 encodes a polypeptide of 426 amino acids with an Mr of 47419. The amino acid sequences of several peptide fragments of the major outer capsid protein reported as 45K were contained in the predicted polypeptide. This protein, renamed the 47K protein, showed high homology with the outer capsid proteins of rice dwarf virus (RDV) and wound tumour virus (WTV); there was 56, 52 and 48% amino acid sequence identity between RGDV and WTV, RGDV and RDV, and RDV and WTV, respectively. S10 had the coding potential for a polypeptide of 320 amino acids with an Mr of 36095 (36K protein), which exhibits 32% and 35% amino acid sequence identity with the predicted translation product of RDV S9 and the P9 capsid protein encoded by WTV S11, respectively. The conserved terminal sequences 5' GG...GAU 3' which are present in all genome segments of WTV and RDV so far analysed, and in S9 of RGDV, were also found in RGDV S8 and S10. This conserved sequence together with the segment-specific inverted repeats found in the terminal sequence of RGDV S8 and S10 are thus characteristic structures common to all three phytoreoviruses. The nucleotide sequence of the region surrounding the inverted repeats was more similar between RGDV and WTV than between RGDV and RDV.

The phytoreoviruses, wound tumour virus (WTV), rice dwarf virus (RDV) and rice gall dwarf virus (RGDV), have icosahedral double-shelled particles approximately 65 to 70 nm in diameter, containing 12 segments of dsRNA and several proteins (Nuss & Dall, 1990). Of the proteins, the outer capsid protein is the major constituent of the virus particles. Therefore, information on the primary structure of the capsid protein would be useful for understanding a major part of the organization of the particle. The nucleotide sequence of the eighth largest genome segment (S8) encoding the outer capsid protein has been analysed for WTV (Xu et al., 1989a) and RDV (Omura et al., 1989). In WTV, the primary structure of the P9 protein, which was reported to be another constituent of the capsid, was studied by nucleotide sequence analysis of genome segment S11 (Dall et al., 1989). This paper describes the nucleotide sequence of genome segment S8 of RGDV which encodes the outer capsid protein and that of S10 which encodes a protein with an amino acid sequence highly homologous to the P9 protein of WTV.

†Present address: Kobe University, Nada, Kobe 657, Japan.
Short communication

(a) 

(b)
Among the RGDV proteins, the 45K protein, the major constituent of the outer capsid (Omura et al., 1985), was the closest in size to the predicted polypeptide. Hence, partial amino acid sequences of the outer capsid protein were analysed and compared with those of the predicted 47K polypeptide. RGDV was purified as reported previously (Omura et al., 1982). Amino acid sequencing of the major capsid was carried out according to the method of Omura et al. (1989). After electrophoresis of the dissociated proteins of purified RGDV, the band corresponding to the outer capsid protein reported as the 45K protein was electroeluted and digested with trypsin. The peptide fragments were isolated by HPLC with a reverse-phase column and were subjected to amino acid sequence analysis using automatic protein sequencers (Applied Biosystems, 477A and Shimadzu PSQ-1). As shown in Fig. 1(a), amino acid sequences of the polypeptide fragments obtained by digestion with trypsin correlated with the amino acid residues 85 to 95, 143 to 148, 160 to 166, 363 to 375 and 404 to 407 of the polypeptide predicted from the nucleotide sequence. These results demonstrate that genome segment S8 of RGDV encodes the major outer capsid protein which has previously been called the 45K protein. The protein is to be renamed the 47K protein.

The phytoreoviruses have almost identical morphologies, based on electron microscopic observation of RDV (Omura et al., 1989), RGDV (Omura & Inoue, 1985) and WTV (Streissle & Granados, 1968). They all have icosahedral double-shelled spherical particles of 65 to 70 nm. The morphological similarity of these viruses is considered to depend on the spatial conformation of structural proteins; i.e. the subunit proteins of the capsomere and the core proteins. The outer capsid of RDV is composed of 180 capsomeres (Kimura & Shikata, 1968; Uyeda & Shikata, 1982) which are trimers of 46K protein subunits (Omura et al., 1989) and the core of RDV consists of 114K proteins (Kano et al., 1990). The major outer capsid proteins of the three viruses are similar in size, 47K in RGDV (Fig.1a), 46K in RDV (Omura et al., 1989) and 48K (one of the outer capsid proteins, P8, encoded by S8) in WTV (Xu et al., 1989a). As shown in Fig. 1(b), the primary structures of these proteins show homology: 56, 52 and 48% amino acid sequence identity between RGDV and WTV, RGDV and RDV, and RDV and WTV, respectively. Furthermore, approximately 38% of the sequences are common among the three viruses. The amino- and carboxy-terminal domains of the proteins are especially highly conserved among the three viruses. i.e. 14 of the first 19 amino-terminal amino acids and 18 of the last 25 carboxy-terminal amino acids are identical in the three viruses. Some regions with stretches of 10 amino acids identical were detected between residues 381 and 390 of RGDV and 377 to 386 of RDV, and between residues 414 to 423 of RGDV and 415 to 424 of WTV. This similarity in primary structure may result in subunits being folded to give capsomeres of the same dimensions, which would make the three viruses indistinguishable in electron microscopy. The high scores (70 to 75%) obtained for chemically similar amino acids (Dayhoff et al., 1972) would also support the supposition that the spatial conformations of the subunit proteins are identical.

The similar structures, with homologous arrangement of the capsid protein amino acids as described above, nevertheless seem to be serologically distinguishable. No cross-reaction was observed between RGDV and RDV (Omura et al., 1985), or between RDV and WTV (Liu & Black, 1978) when intact virus particles were used as antigens. The domains with long identical amino acid sequences are therefore not thought to be on the surface of the virus particles. This assumption correlates with the fact that RGDV reacts with antiserum against dissociated RDV particles (Matsuoka et al., 1986). The spatial disposition of amino acids in capsomeres, and the interactions between capsomeres and the core should be discernible by X-ray diffraction studies using crystals of RDV (Mizuno et al., 1990).

The nucleotide sequence of S10 is shown in Fig. 2(a). The segment contains 1198 bp of dsRNA with a calculated $M_r$ of $0.77 \times 10^6$ and a GC content of 45.2%. It has one long open reading frame, which starts from residue 22 and extends for 960 nucleotides, followed by a 3' non-coding region of 217 nucleotides. None of the other open reading frames exceeds 61 amino acids.

The open reading frame has the coding potential for a 320 amino acid polypeptide with a calculated $M_r$ of 36095 (36K). High homology was detected between the 36K protein of RGDV and the 38-9K protein of RDV, encoded by genome segment S9 (Fukumoto et al., 1989), and the P9 protein of WTV, encoded by genome segment S11 (Dall et al., 1989) (Fig. 2b). There is 35%, 32% and 37% amino acid sequence identity between RGDV and WTV, RGDV and RDV, and RDV and WTV, respectively. There are a series of identical six-amino acid sequences at residues 3 to 8 in the 36K protein of RGDV and the corresponding region in the 38-9K protein of RDV and WTV.

**Fig. 1.** (a) Nucleotide sequence of the plus-sense strand of segment S8 of RGDV and the amino acid sequence of its predicted translation product. The in-phase termination codon is indicated with an asterisk. The amino acid sequences of several peptide fragments which have been determined are underlined. (b) Alignment of the predicted amino acid sequences of the major outer capsid proteins of RGDV, RDV and WTV, which are all encoded by genome segment S8. Identical amino acids are boxed. Gaps (-) were inserted to maximize the alignment. Numbers are amino acid positions from the N terminus.
Fig. 2. (a) Nucleotide sequence of the plus-sense strand of segment S10 of RGDV and the amino acid sequence of its predicted translation product. The in-phase termination codon is indicated with an asterisk. (b) Alignment of the predicted amino acid sequences encoded by genome segment S10 of RGDV, S9 of RDV and S11 of WTV. Identical amino acids are boxed. Gaps (-) were inserted to maximize the alignment. Numbers are amino acid positions from the N terminus.
Fig. 3. Comparison of the terminal sequence domains of the inverted repeats of the corresponding segments (see text) of RGDV, RDV and WTV. No segment corresponding to S9 of RGDV and WTV could be detected in segments S3 to S10 of RDV. Data cited were reported by Koganezawa et al. (1990) for RGDV S9, Xu et al. (1989a) for WTV S8, Anzola et al. (1989) for WTV S9, Dall et al. (1989) for WTV S11, Omura et al. (1989) for RDV S8 and Fukumoto et al. (1989) for RDV S9.

protein of RDV, and at residues 6 to 11 and 80 to 85 in the RGDV protein and corresponding regions in the P9 protein of WTV, respectively.

The P9 protein of WTV (Xu et al., 1989b) has been proposed as one of the capsid proteins (Reddy & MacLeod, 1976). As mentioned above, the predicted amino acid sequence of the P9 protein of WTV (Dall et al., 1989) is homologous to the 36K protein of RGDV which does not correspond to any protein released from purified RGDV (Omura et al., 1985). There is no evidence indicating that the 38.9K protein of RDV is also a capsid protein.

By including the sequence information for genome segments S8 and S10 with S9 of RGDV (Koganezawa et al., 1990), striking similarities were found in the terminal structures of the genome segments of all three phyto-reoviruses. The terminal sequences, 5' GG...GAU 3', conserved in all the genome segments of WTV (Anzola et al., 1987), RDV S3 to S10 (Uyeda et al., 1987, 1989, 1990; Omura et al., 1988, 1989; Fukumoto et al., 1989; Nakashima et al., 1990; Suzuki et al., 1990a, b) and RGDV S9 (Koganezawa et al., 1990), were found in RGDV S8 and S10. Segment-specific inverted repeats were also found in residues 5 to 21 and 1557 to 1573 for S8, and 4 to 16 and 1282 to 1195 for S10 of RGDV (Fig. 3), as reported in all the genome segments of WTV, RDV and RGDV described above. Thus, the terminal sequence, 5' GG...GAU 3', and the inverted repeat are structures characteristic of genome segments of phyto-reoviruses and may be associated with common functions.

Terminal nucleotide sequences associated with the inverted repeat were compared among the three phyto-reoviruses (Fig. 3). All nine nucleotides at the 5' and five nucleotides at the 3' termini were identical among RGDV S8, S9 and S10 and WTV S8, S9 and S11. However, homology was low between RGDV and RDV, except for the conserved 5' (GG) and 3' (GAU) termini. The molecular structures of the genome termini, assumed to regulate their own expression in WTV (Xu et al., 1989a), were closer between RGDV and WTV than between RGDV and RDV, despite the fact that the plant host and vectors are different for RGDV and WTV and similar for RGDV and RDV.

The authors are grateful to Dr H. Kano, Dr H. Hirano and Dr T. Watanabe for their continued interest and valuable suggestions.

References


(Received 2 April 1991; Accepted 10 July 1991)