Nucleotide sequence of carnation ringspot dianthovirus RNA-2

T. L. Kendall and S. A. Lommel*

Department of Plant Pathology, Box 7616, North Carolina State University, Raleigh, North Carolina 27695-7616, U.S.A.

RNA-2 of carnation ringspot virus (CRSV), the type member of the dianthovirus group, has been cDNA cloned and sequenced. CRSV RNA-2 is 1394 nucleotides in length and contains a single open reading frame encoding a 304 amino acid polypeptide of 33.8K. Amino acid sequence alignment of this polypeptide with the cell-to-cell movement proteins encoded by RNA-2 of red clover necrotic mosaic virus (RCNMV) Australian (Aus) and Czechoslovakian (TpM-34) isolates indicates 59.6% and 55.7% sequence identity, respectively. The N-terminal 230 amino acids are more highly conserved, with 64.3% and 62.6% sequence identity, respectively. The cell-to-cell movement proteins of the two RCNMV isolates are themselves 82.5% and 91.7% identical when the amino-terminal 230 amino acids are compared. Structural prediction comparison of the RCNMV-Aus, RCNMV-TpM-34 and tobacco mosaic virus cell-to-cell movement proteins to the putative CRSV RNA-2-encoded movement protein suggests that even though no primary amino acid sequence similarity exists, the movement protein polypeptides are possibly similar in structure and function.

CRSV was maintained in Nicotiana clevelandii under greenhouse conditions and purified as described by Lommel et al. (1982). Virion RNA was isolated and cDNA clones generated as described by Lommel et al. (1988). A nearly full-length CRSV RNA-2 cDNA was identified by Southern blot analysis (pCRS132) and subcloned into the phagemid vector pBS(+) (Stratagene). The nucleotide sequence of pCRS132 was determined by the dideoxynucleotide chain termination method (Sanger et al., 1977) from either full-length DNA or exonuclease III-generated nested deletions (Henikoff, 1987). A portion of the sequence unaccessible to sequencing with the T7 primer was derived by priming with the oligonucleotide 5' CAAGTCGACCATGAA-TGT 3' corresponding to nucleotides 679 to 696. The 5'-terminal 28 nucleotides not represented in pCRS132 were determined directly from RNA-2 by primer extension using the oligonucleotide 5' TACATGAACT-GCCATCTC 3', complementary to nucleotides 44 to 61.

The complete and compiled nucleotide sequence of CRSV RNA-2 is depicted in Fig. 1. Computer analysis and translation of the deduced sequence (Rhoads & Roufa, 1985) revealed the presence of a single large open reading frame (ORF) beginning at the first initiation codon at nucleotide 47. The presence of a purine at position -3 and a G in position +4 in the sequence GAGAUGGCAG relative to the adenine residue in the initiation codon agrees with Kozak's consensus sequence (Kozak, 1986) and the consensus sequence of Lutcke et al. (1987) required for the initiation of translation. The
domain but rather a region necessary for correct protein passing residues 65 to 86) identified originally by movement protein nucleic acid-binding domain (encom-

The 'Plotstructure' profiles of the movement proteins (1978) and Hopp & Woods (1983), respectively. The described by Chou & Fasman (1978), Gamier

desirability and hydrophilicity predictions based on the methods The 'Plotstructure' profiles of the movement proteins were determined by the formula described by Emini (Fig. 2) of this polypeptide with the RCNMV Australian (Fig. 2) and Czechoslovakian (TpM-34) RNA-2-encoded proteins like the G protein-coupled receptors (Caron et al., 1990) and, therefore, suggests this region is critical for maintaining overall protein structure.

TMV and dianthovirus movement proteins have similar unstructured basic hydrophilic C termini, which are not necessary for function (Fig. 3, domain 4). Xiong et al. (1992) and Osman et al. (1991 b) have shown that the C-terminal portion of the RCNMV cell-to-cell movement protein is dispensable for cell-to-cell movement. Likewise, Berna et al. (1991) and Okada et al. (1990) have shown that the C-terminal 55 and 31 amino acids can be deleted from the TMV and tomato mosaic virus movement proteins, respectively, while maintaining limited cell-to-cell movement activity. Interestingly, domain 4 of all four movement proteins is predicted to be on the surface of the protein (Fig. 4).

Citovsky et al. (1992) identified two independently active single-strand nucleic acid-binding domains in the TMV movement protein. The first domain (A) between amino acid residues 112 and 185 is predicted to be internal whereas domain B, between residues 185 and 268, is predicted to be on the surface of the movement

Fig. 1. Complete nucleotide sequence of CRSV RNA-2 and the amino acid sequence of the major 34K ORF. Amino acids are presented in the one-letter code below the first base of each codon within the ORF. Numbers to the right of each line identify the last residue of that line. The asterisk identifies the first nucleotide of the stop codon terminating the ORF.

ORF ends with an opal termination codon at nucleotide 959.

The single major ORF has the capacity to encode a 304 amino acid polypeptide. Based on studies with RCNMV (Paje-Manalo & Lommel, 1989; Osman & Buck, 1987), this polypeptide is predicted to be the cell-to-cell movement protein. Amino acid sequence alignment (Fig. 2) of this polypeptide with the RCNMV Australian (Aus) and Czechoslovakian (TpM-34) RNA-2-encoded cell-to-cell movement proteins indicates 59.6% and 55.7% sequence identity, respectively. The N-terminal 230 amino acids are even more highly conserved, showing 64.3% and 62.6% sequence identity, respectively. The cell-to-cell movement proteins of the two RCNMV isolates are 82.5% and 91.7% identical when the N-terminal 230 amino acids are compared.

We predicted the secondary structures and surface probabilities for the three dianthovirus proteins as well as the tobacco mosaic virus (TMV) movement protein. The 'Plotstructure' profiles of the movement proteins depicted in Fig. 3 incorporate α-helix, β-sheet, flexibility and hydrophilicity predictions based on the methods described by Chou & Fasman (1978), Garnier et al. (1978) and Hopp & Woods (1983), respectively. The surface probability profiles for the movement proteins were determined by the formula described by Emini et al. (1985) (Fig. 4).

Citovsky et al. (1992) suggested that the TMV movement protein nucleic acid-binding domain (encompassing residues 65 to 86) identified originally by Citovsky et al. (1990) is not the actual active binding domain but rather a region necessary for correct protein folding. The greatest degree of structural similarity between the movement proteins of the three dianthoviruses and TMV exists within the four blocked and marked domains (Fig. 3). The first three clustered domains include a hydrophilic turn domain (Fig. 3, domain 2) flanked by strong hydrophobic regions (Fig. 3, domains 1 and 3). Domain 2 of the TMV movement protein includes the region identified by Citovsky et al. (1992) necessary for proper folding. The organization of these three domains is consistent with membrane-bound proteins like the G protein-coupled receptors (Caron et al., 1990) and, therefore, suggests this region is critical for maintaining overall protein structure.

Fig. 2. Amino acid sequence alignment of the cell-to-cell movement proteins of RCNMV-Aus (Lommel et al., 1988) and RCNMV-TpM-34 (Osman et al., 1991 a) with the putative CRSV movement protein. The alignment was compiled with FASTP (Lipman & Pearson, 1985). Gaps introduced to effect the best alignment are indicated by dashes. Identical amino acids in two or more sequences are boxed.
protein. Domain 4 of the dianthovirus movement protein may be functionally analogous to the TMV single-stranded nucleic-acid binding domain B. If dianthovirus domain 4 and TMV domain B are functionally analogous, this would suggest that the dianthoviruses would also have a second independent nucleic acid-binding domain, based on the dispensability of the C-terminal portion of the protein for minimal activity (Osman et al., 1991b). The similarity in the predicted protein structures coupled with the genetic data suggest that the dianthovirus and TMV movement proteins may be similar in structure and function.

We are grateful to Long V. Nguyen for help with the protein structure analysis. This work was supported in part by a grant from the North Carolina Biotechnology Center.

References


(Received 9 March 1992; Accepted 3 June 1992)