Molecular cloning, sequencing and expression in *Escherichia coli* of the odontoglossum ringspot virus coat protein gene

Yoshinori Isomura,1,2 Yoshinori Matumoto,2 Akiko Murayama,1 Masaaki Chatani,2 Narinobu Inouye3 and Masato Ikegami1*

1NODAI Research Institute, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156, 2Nippon Oil Company Ltd., Yamaguchi-ken 744 and 3Research Institute for Bioresources, Okayama University, Kurashiki 710, Japan

The sequence of the 3'-terminal 1865 nucleotides of the genome of the tobamovirus odontoglossum ringspot virus (ORSV) was determined. This sequence contained two open reading frames (ORFs), 912 and 477 nucleotides long. The 912 nucleotide ORF has been identified as the cell-to-cell transport protein gene. The 477 nucleotide ORF was expressed in *Escherichia coli*, and the product was detected by antibodies specific for the coat protein of ORSV. The amino acid sequence of protein encoded by this ORF shares 84% similarity with the tobacco mosaic virus (TMV) (vulgare) coat protein. The 3'-terminal untranslated region of ORSV comprises 414 nucleotides, 210 nucleotides more than that of TMV (vulgare) RNA.

Odontoglossum ringspot virus (ORSV) was first isolated from *Odontoglossum grande* (Jensen & Gold, 1951). This virus causes ringspot symptoms in *O. grande*, and diamond mottle in *Cymbidium*. The particles of ORSV are rod-shaped and contain an ssRNA molecule, approximately 6 kb in length (Paul, 1975). ORSV is a member of the tobamoviruses (Paul, 1975). We have reported the nucleotide sequence of the cell-to-cell transport protein gene of ORSV (Isomura et al., 1990). In this paper we present the nucleotide sequence of the 3'-terminal region of ORSV RNA and identification of the coat protein gene.

ORSV, originally from the laboratory of Professor N. Inouye, was purified from *Cymbidium* plants as described by Inouye (1966), and RNA isolation from purified virions was carried out as described by Ikegami et al. (1987). ORSV RNA was polyadenylated using poly(A) polymerase (Takara Shuzo) as described by Sippel (1973). DNA complementary to the RNA was synthesized using a cDNA kit (purchased from Pharmacia) according to the manufacturer's specifications. Synthesized double-stranded cDNA was inserted into the EcoRI site of pUC119 using EcoRI linkers. One clone containing an insert of approximately 1800 nucleotides was designated pORE1033. The cloned DNA was digested into fragments using EcoRI, Accl, PstI, NsiI, ApaI and XbaI. The fragments were then subcloned into pUC119 and/or pUC119, as well as into M13mp18 and/or M13mp19. Nucleotide sequence analysis was carried out by the enzymic dideoxynucleotide chain termination sequencing method with a DNA sequencer (model 370A, Applied Biosystems). The 3' 40 nucleotide residues of ORSV end-labelled genomic RNA were sequenced by direct RNA enzymic sequencing (Haseloff & Symons, 1981). Nucleotide sequence data were analysed using the program DNASIS (Takara Shuzo). An ORSV open reading frame (ORF) expression vector was constructed by digesting pORE1033 with XbaI and NsiI and subcloning the 694 bp fragment (717 to 1411 nucleotides in Fig. 1) containing ORF 2 (477 bp) into pUC119 at the XbaI/PstI site. This cloned DNA was digested with KpnI and SmaI, then deleted with exonuclease III leaving four nucleotides before the ATG codon where the coat protein mRNA translation starts. After digesting with S1 nuclease, this fragment was ligated with EcoRI linkers and digested again with HindIII. The fragment containing ORF 2 was ligated to the EcoRI/HindIII site of the pKK223-3 expression vector (Pharmacia), followed by transfection of *E. coli* JM109. One recombinant, designated pBORCPD-83, was used for the expression of the ORSV ORF 2 protein. The latter was done in *E. coli* JM109 cells as described before (Murayama et al., 1991) and the lysate of *E. coli* cells harbouring pBORCPD-83 was analysed in 12-5% polyacrylamide gels containing 0-1% SDS according to Laemmli (1970). After electrophoresis, the proteins in the gel were transferred electrophoretically, and bound to nitrocellulose membranes (0-45 μm, Toyo Roshi Kaisha). The membrane was incubated with ORSV antiserum (our own laboratory stock), followed by incubation with antibodies conjugated with horseradish peroxidase (Immun-blot assay kit, Bio-Rad) and immer-
turer's directions (Murayama et al., 1991). For comparison, some gels were stained directly with Coomassie blue R250 without immunoblotting. The nucleotide sequence of the 3'-terminal region of ORSV RNA is shown in Fig. 1. The plus strand (virion polarity) was examined for potential protein coding regions (protein of M, > 10000). There are two possible regions (ORFs 1 and 2) including in-phase AUG codons terminating with an UGA or UAA codon, as shown in Fig. 1. ORFs 1 and 2 consisted of 912 and 477 nucleotides respectively. ORF 1 represents the cell-to-cell transport protein gene (Isomura et al., 1990). We examined whether ORF 2 is the coat protein gene. The amino acid sequence of the coat protein of tobacco mosaic virus (TMV) (vulgare). Paul (1975) has published the amino acid sequence for the coat protein of the German isolate of ORSV, which shares only 71% identity with the sequence encoded by ORF 2 of the Japanese isolate; this may be due to differentiation between ORSV strains. The ORSV ORF 2 of the Japanese isolate; this may be due to differentiation between ORSV strains. The ORSV ORF 2 was three nucleotides shorter than the TMV (vulgare) coat protein gene. The amino acid sequence of the protein encoded by ORF 2 of this ORSV strain shared 84% similarity with respect to the coat protein of tobacco mosaic virus (TMV) (vulgare). In ORSV, a 3' untranslated region of 414 nucleotides was seen, 210 nucleotides longer than the 3' untranslated region of TMV RNA. Analysis of the untranslated region of ORSV revealed three domains, each of 57 nucleotides, which were very similar (Fig. 1). Such repeated sequences are not present in the TMV 3' terminal untranslated region. These repeated sequences are not present in the TMV 3' terminal untranslated region. These repeated sequences contained two pseudoknotted structures (Fig. 3) similar to two of three pseudoknotted structures formed in the 3' terminal region of TMV RNA, as reported by Van Belkum et al. (1985). The 39 nucleotides at the 3' terminus of ORSV RNA were also highly similar to the 38 nucleotides at the 3' terminus of TMV RNA.
Fig. 2. SDS-PAGE of lysate of E. coli cells harbouring pBORCPD-83 (lanes 1 to 3) and detection of ORSV capsid polypeptide by Western blotting (lanes 4 to 6). Lanes 1 to 3: proteins separated by SDS-PAGE in a 12.5% polyacrylamide gel and stained with Coomassie blue R250. Lane 1, proteins from control E. coli cells; lane 2, proteins from E. coli cells harbouring pBORCPD-83; lane 3, coat proteins from purified ORSV particles. A band of M, about 17000 is seen in lanes 2 and 3. Lanes 4 to 6 correspond to lanes 1 to 3 after Western blotting onto nitrocellulose and probing with antiserum to ORSV virions. M, markers are cytchrome c tetramer, 49.6K; cytchrome c trimer, 37.2K; cytchrome c dimer, 24.8K; and cytochrome c monomer, 12.4K.

Fig. 3. Pseudoknotted structures in two repeated sequences of 3′-terminal untranslated region of ORSV RNA. These structures are similar to two of three pseudoknotted structures formed in the 3′-terminal region of TMV RNA as reported by Van Belkum et al. (1985).

Although the amino acid sequences of the putative cell-to-cell transport protein (Isomura et al., 1990) and coat protein of ORSV shared 70% and 84% identity with the TMV 30K protein and coat protein respectively, the difference between the sequences of 3′ non-coding regions of ORSV (Japanese isolate) and TMV RNAs supported the previous conclusion that the tobamovirus isolated from Cymbidium by Inouye (1966) was ORSV and not a strain of TMV. Jensen & Gold (1951) isolated ORSV from O. grande, and Kado et al. (1968), Corbett (1967) and Lawson (1970) isolated TMV orchid strain (TMV-O) from orchids. The distinction between ORSV and TMV-O is not clear. It would be interesting to compare the nucleotide sequences of ORSV and TMV-O.

References


(Received 2 January 1991; Accepted 14 May 1991)