Rice ragged stunt Oryzavirus genome segment 9 encodes a 38600 \( M_r \) structural protein

Ichiro Uyeda,* Haruhisa Suga, Sang Yong Lee,† Jin Yan, Tatsuji Hataya, Ikuo Kimura and Eishiro Shikata

Department of Agrobiology and Bioresources, Faculty of Agriculture, Hokkaido University, Kita 9, Nishi 9, Kita-ku, Sapporo 060, Japan

The complete nucleotide sequence of rice ragged stunt virus genome segment 9 (S9) was determined. The S9 segment is 1132 nucleotides long and has a long open reading frame starting from the first AUG codon at nucleotide position 14–16 and terminating at a UAG codon located at 1028–1030, which could encode a polypeptide with an \( M_r \) of 38600 (P9). The encoded polypeptide has no sequence homology to polypeptides of any other plant reoviruses published previously. An immunological study demonstrated that P9 was the smallest of the structural proteins. The P9 polypeptide was expressed as a fusion protein with maltose binding protein in Escherichia coli. Antisera to purified virions and to the fusion protein reacted with both the bacterially expressed polypeptide and the smallest polypeptide of the purified virus in immunoblotting analyses.

Introduction

Rice ragged stunt virus (RRSV) is reported to occur in several south-east and east Asian countries and is an economically important disease of rice (Milne et al., 1982). The virus is transmitted by the planthopper Nilaparvata lugens and belongs to the new genus Oryzavirus of the family Reoviridae (Francki et al., 1991; Mayo & Martelli, 1993). The genome consists of 10 segmented double-stranded RNAs (dsRNAs) with estimated \( M_r \) values ranging from \( 0.78 \times 10^6 - 2.58 \times 10^6 \) (Kawano et al., 1984; Uyeda et al., 1990). Five major structural polypeptides were detected when the purified virus particles were subjected to SDS–PAGE (Hagiwara et al., 1990; Chen et al., 1989). Hagiwara et al. (1990) reported that a high concentration of MgCl\(_2\) removed \( M_r \) 47000 and 37000 structural proteins, suggesting that these two proteins are components of the spike. Neither of these has been assigned to a particular genomic segment. Molecular characterization of the genome has not been reported except for analyses of the nucleotide sequences of the terminal portions (Yan et al., 1992). In this paper, we have determined the complete nucleotide sequence of the genome segment 9 (S9) and demonstrated that it encodes an \( M_r \) 38000 structural protein (P9).

The viral transcripts were synthesized from purified virions (Lee et al., 1987a) and a poly(A) tail was added by poly(A) polymerase and used for cDNA synthesis by the method of Gubler & Hoffman (1983) as previously described (Lee et al., 1987b). cDNA clones that reacted with the electrophoretically separated S9 probe in a dot-blot hybridization (Lee et al., 1987b) were selected for nucleotide sequencing. A cDNA clone that covered a complete 3′-terminal portion was selected by a colony hybridization using a synthetic 17-mer oligonucleotide that was complementary to the 3′ terminus as a probe. Four cDNA clones were sequenced by the dideoxy chain termination method of Sanger et al. (1977) and were found to cover the entire genomic sequence, except for the 5′-terminal 19 nucleotides. One clone had a sequence identical to that of the 3′-terminal oligonucleotides determined directly from the genome dsRNA followed by a poly(d/A) tail (Yan et al., 1992). More than 96% of the segment was sequenced in both directions. Since the cDNA clones did not cover the entire 5′-terminal region, its sequence was determined by reverse transcription of the viral transcript using the dideoxy termination method. A synthetic 16-mer that was complementary to nucleotide positions 49–64 from the 5′ terminus was used as a primer (Yan et al., 1992). The entire nucleotide sequence of S9 is presented in Fig. 1. It is 1132 nucleotides long and has one long open reading frame (ORF).
GATAAATTCGGCCATGAGACTGCTTTTGCACAGATCTCTTTTACAGCTCTGCTACAGG
M K T A F A R D P F T A PATG

TACGTTACCGAACATATGCGCTCCAGATCTATTCCCAGCTCTCTATCTCTAGGTTCTTT
T Y G T I Y A S R L P R L S I S K F L

GAAGAGTGCAAATCCAGAAATTTAGAACATACGAGATGAGAGCGCTCAGAACCAACACAG
E D A N P E I V E L S R Y E A L G T N R

ACCCAGCTCAGGAAGCAGCATACTGGAACACCCGAGCTGATGAAACCAAGCTTTACTGGAAAC
P S S G K R A M Q P A V S K P A L L E T

CGTATTCACACTGATATATGGCAAGAGACAAACACATCGCAGAATTAAAGCAAGCAAGCA
V F T L D I W Y R R T N N Q N I G N L R

AGATTCGGTATCCTCGATTTCTCTATCTGATGCTGATAAGAAGACGCGTATGCGACT
D S V S R F L S D D R V R E A V M V R L

AGATTATACATTGTTTGTCCGCTAAAGAAATACGTGGTGATGATGAAACGAAAAAGGCA
D L D I V T K L E Y W L I V K D K E A

CCGACTTTTCGCGGAGCATTGGCCCTTTGACTCGATCTCTTTTTGCTCAATCGCGCAGAAA
Q T F A D R L A P D S H L F V N R G E N

CGGGAATTATATGATCTGGTAACTCAGACGTTACCGAGTTGATTGTCTATTAAGGATAC
A N Y D L V T Q T F I P S D A F L K D N

CTTTAAGACGGAACAGCATAAGGAATGGTTGCTGAAGAGGCTGCAAACCCATACGGGACTAGA
F K T E A L K K L L L S V Q N H T G L D

CGCTGATTGCGAAGGTAGTATGCTGACAGACCGACTTACAACTGCGCAGACATATCT
A G L Q G D S S K A T Y N I G L Q Y L

CGAGGATAGAGGCCTCTCTGACTGGCAAGGTTGCTCTCAACACAAATGCGCTTTTGCGGA
E D E A L L Y R Q G V A L Q Q M A F A E

GTTGGAATATAGCTAGAGAGCAAGGAAAGAGCGCTTTCCCATCAACCTTCGATCTAACGAA
L E L A R G A E K A F P S T F D L S N

TAGACCGCTTGACACTTTGGATCCTTAAGCGGAACGTGTAAGTGATGATCACAACAAACGTTTAA
R P A C N L I L K R T C K W Y Q Q T F K

GGATGAGGAGAAAGGAAATTTTGTCAAGAGGTGTTGTGATGATCGCTTTTGCGAGGCTAA
D E R E K F A K S L W V D D F A E A N

TTGGAAACACAGGTAAATTTTGCCTTTTGCGCTTTTCAACCACTTAAATGTAATTGAGAGATG
W N T G N L S F G F S T T L N V I E R W

GGCTTTTGACAGAGTTATGTCTCATATGTCACATCTCTGTCATATATCTCGCAAGAGCC
R L T R F Y V H M Y S S S V H I Y S Q K A

CTCGGGTTAGGAAACGCTGCGGATCAAGTGTAATTGAAATAGGACTTCTCACCCCTTGAT
G S G *

GACATTTCGGCCTCTTACCAGGGTGCGACACCCGAAAACGGCGAGACCATGGTGCC
1132

Fig. 1. Nucleotide sequence of cDNA to rice ragged stunt virus genome segment 9 (upper) and amino acid sequence (lower) of the encoded polypeptide.

encoding 338 amino acids (P9). The predicted polypeptide has no significant sequence homologies to any of those polypeptides encoded by plant reoviruses sequenced so far (Uyeda et al., 1994; Nuss & Dale, 1990).

In order to assign a polypeptide encoded by 9, the longest ORF was expressed in E. coli as a fusion with the maltose binding protein as described in Masuta et al. (1992). A cDNA clone was made by reverse transcription coupled with PCR of genomic dsRNAs (Murao et al., 1994) using a set of primers complementary to the 5’ translational start site and the 3’-terminal 17 nucleotides. EcoRI and PstI recognition sites were added to their 5’ termini for subsequent cloning into bacterial plasmid pUC119. The longest ORF was then cloned in the EcoRI and PstI sites of pMAL C2 (New England Biolabs). The E. coli strain JM83 was transformed with the plasmid and
the fusion protein was expressed as described by Masuta et al. (1992). The bacterial cells expressing the fusion protein produced a band with an $M_r$ of 81000 in SDS–PAGE (Laemmli, 1970). Cleavage of the fusion protein by the proteinase factor Xa yielded a band with an apparent $M_r$ of 36200, which is close to the $M_r$ (38600) of P9 expected from the nucleotide sequence. This band reacted in immunoblotting analyses with antisera raised against purified virions, suggesting that P9 was the smallest structural protein (Fig. 2a), previously reported as the $M_r$ 37000 structural protein by Hagiwara et al. (1990). The mobility of the cleavage product was slightly slower than the smallest structural protein, probably because four amino acid residues derived from the plasmid vector sequence had been added to the N terminus of the protein. Antisera were raised against the fusion protein in rabbits. When the purified virus was electrophoresed, blotted to an Immobilon-P membrane (Millipore), and tested with the antisera, only the smallest structural protein specifically reacted (Fig. 2b). These experiments prove that P9 is the smallest structural protein with an $M_r$ of 38600.

This study was supported in part by Grant-in-Aid for Scientific Research no. 62480042 and 01560045, and Grant-in-Aid for Cooperative Research no. 6304013 from The Ministry of Education, Science and Culture, and a grant from Suhara Foundation. We thank Dr C. Masuta, Japan Tobacco Inc., for suggestions on handling pMAL expression vectors.

References


(Received 25 July 1994; Accepted 16 November 1994)