Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12541–12546, October 1996
Microbiology

Organization and expression of the double-stranded RNA genome of *Helminthosporium victoriae* 190S virus, a totivirus infecting a plant pathogenic filamentous fungus

(internal initiation/RNA-dependent RNA polymerase)

SHAOHUA HUANG AND SAID A. GHABRIAL*

Department of Plant Pathology, University of Kentucky, Lexington, KY 40546-0091

Communicated by Robert J. Shepherd, University of Kentucky, Lexington, KY, July 18, 1996 (received for review May 8, 1996)

ABSTRACT **The complete nucleotide sequence, 5178 bp, of the totivirus *Helminthosporium victoriae* 190S virus (Hv190SV) double-stranded RNA, was determined. Computer-assisted sequence analysis revealed the presence of two large overlapping ORFs; the 5′-proximal large ORF (ORF1) codes for the coat protein (CP) with a predicted molecular mass of 81 kDa, and the 3′-proximal ORF (ORF2), which is in the −1 frame relative to ORF1, codes for an RNA-dependent RNA polymerase (RDRP). Unlike many other totiviruses, the overlap region between ORF1 and ORF2 lacks known structural information required for translational frameshifting. Using an antiserum to a C-terminal fragment of the RDRP, the product of ORF2 was identified as a minor virion-associated polypeptide of estimated molecular mass of 92 kDa. No CP–RDRP fusion protein with calculated molecular mass of 165 kDa was detected. The predicted start codon of the RDRP ORF (2605-AUG-2607) overlaps with the stop codon (2606-UGA-2608) of the CP ORF, suggesting RDRP is expressed by an internal initiation mechanism. Hv190SV is associated with a debilitating disease of its phytopathogenic fungal host. Knowledge of its genome organization and expression will be valuable for understanding its role in pathogenesis and for potential exploitation in the development of biocontrol measures.**

Although a large number of the viruses that infect plant pathogenic fungi have been reported to be avirulent, it is becoming increasingly clear that phenotypic consequences of harboring specific mycoviruses or certain double-stranded RNA (dsRNA) molecules can range from symptomless to severely debilitating, and from hypovirulence to hypervirulence (1, 2). Virus-induced diseases and virus-mediated attenuation of virulence in phytopathogenic fungi provide excellent opportunities for fundamental studies aimed at developing novel biological control measures. In this regard, the dsRNA viruses infecting *Helminthosporium victoriae* (teleomorph: *Cochliobolus victoriae*), the causal agent of Victoria blight of oats, are of special interest because they are associated with a lytic disease of their fungal host (3). Two isometric dsRNA viruses, designated according to their sedimentation values as the *Helminthosporium victoriae* 190S virus (Hv190SV) and 145S virus (Hv145SV) have been identified in diseased fungal isolates (4). Recent studies suggest that the dsRNAs associated with the Hv145S component are satellites dependent on Hv190SV for encapsidation and replication (ref. 1 and unpublished work). This paper focuses on the molecular characterization of Hv190SV.

The Hv190SV is a member of the genus *Totivirus* in the family *Totiviridae* (5). The *Saccharomyces cerevisiae* L-A (ScV-L-A) virus is the only member of the genus that has been well-characterized at the molecular level (5). Complete sequence of the ScV-La virus (synonymous ScV-L-BC virus), and partial sequence of the Ustilago maydis virus H1 have been reported (6, 7). The complete nucleotide sequence of several of the totiviruses that infect the parasitic protozoa *Giardia lamblia* (Giardia lambia virus, GLV) and *Leishmania* species [two strains of Leishmania RNA virus 1 (LRV1-1 and LRV1-4) and LRV2-1] have been determined (8–11). These protozoal viruses belong to the genera *Giardiavirus* and *Leishmaniavirus* in the family Totiviridae (5).

Although the capsid of Hv190SV, like other totiviruses, is encoded by a single gene, it contains three closely related polypeptides, p78, p83, and p88. Whereas p83 and p88 are phosphoproteins, p78 is nonphosphorylated (12). Purified Hv190S virion preparations contain two types of virions that differ slightly in sedimentation rates, 190S-1 and 190S-2 virions (12). These two virion types, which also differ in transcriptional activity, are believed to represent different stages in the virus life cycle (1). The 190S-1 and 190S-2 virions can be further differentiated based on capsid composition; the 190S-1 capsids contain two coat proteins (CPs), p88 and p83, occurring in approximately equimolar amounts, and the 190S-2 capsids are comprised of similar amounts of p88 and p78 (12). Thus the capsids of Hv190SV contain two related major CPs, whereas all other totiviruses so far characterized appear to contain only a single major CP.

Because none of the totiviruses (or tentative totiviruses) that infect filamentous fungi have been characterized at the molecular level and because of the apparent unique properties of Hv190SV (virion-associated protein kinase activity and capsid protein heterogeneity), it was of interest to study the organization and expression of its dsRNA genome. In this study, we report the complete nucleotide sequence of Hv190SV dsRNA and elucidate its genome organization. We also show that the Hv190SV RNA-dependent RNA polymerase (RDRP) is a separate virion-associated polypeptide and propose that it is expressed by an internal initiation mechanism. Hv190SV thus differs from other totiviruses that express their RDRP as a CP–RDRP fusion protein via a translational frameshift.

MATERIALS AND METHODS

Virion Purification and dsRNA Extraction. Virion purification was by the procedure of Ghabrial and Havens (13). dsRNA was isolated from purified virions by SDS/phenol extraction and purified by linear-log sucrose density gradient

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Hv190SV, *Helminthosporium victoriae* 190S virus; RDRP, RNA-dependent RNA polymerase; CP, coat protein; UTR, untranslated region; ScV-L-A, *Saccharomyces cerevisiae* L-A virus; GLV, Giardia lambia virus; LRV1, Leishmania RNA virus 1; dsRNA, double-stranded RNA.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U41345).

*To whom reprint requests should be addressed.

centrifugation (4). Labeling of the 5′ terminus of dsRNA was made as previously described (14).

Molecular Cloning and Sequence Analysis. A cDNA library representing Hv190SV dsRNA was prepared from gradient-purified dsRNA. First strand cDNA was synthesized using BRL-Superscript RT (RNase H$^-$ reverse transcriptase) according to the manufacturer's instructions (GIBCO/BRL). Hv190SV dsRNA, denatured by treatment with 90% dimethyl sulfoxide for 20 min at 60°C, was used as a template, and mixed random hexamers were used as primers. Additionally, poly(A) tails were added enzymatically to the two strands of dsRNA by treatment with poly(A) polymerase (15), and cDNA synthesis was primed with oligo(dT)$_{12-18}$ primers. Furthermore, synthetic primers based on known sequences were used to initiate cDNA synthesis from denatured dsRNA (Fig. 1). Second strand synthesis was made by the procedure of Gubler and Hoffman (16) using the Amersham cDNA synthesis kit. The ds-cDNA was blunt-ended with T4 DNA polymerase, ligated into the phagemid vector pUC119 at the *Sma*I site, and transformed into *Escherichia coli* strain DH5α. Selection of ampicillin-resistant cDNA clones containing inserts was made by blue/white colony screening on 5-bromo-4-chloro-3-indolyl β-D-galactoside/isopropylthioβ-D-galactoside medium and the veracity of the inserts was tested by colony hybridization using ^{32}P-end-labeled denatured dsRNA (17). A full-length cDNA clone was constructed from a series of overlapping clones. cDNA sequencing was done by the dideoxy-nucleotide chain termination method (18) using α-^{35}S and Sequenase (United States Biochemical). The full-length clone and its subclones were sequenced completely in both directions using a combination of subcloning and synthetic primers designed from internal sequences (Fig. 1). Sequence analyses and a homology search were performed using the GCG Sequence Analysis software package (Genetics Computer Group, Madison, WI). Sequence alignments were made with the GAP program (GCG package).

Recombinant DNA Constructions. *Expression of 5′ large proximal ORF (ORF1) in* E. coli. A near full-length cDNA clone (nucleotides 4–5178) was inserted into the *Sac*I and *Xho*I sites of pBluescript II KS+ to generate the plasmid pHV-4. The orientation of the insert in pHV-4 is such that the 5′ end of the Hv190SV plus strand is proximal to the *Sac*I site, and the 3′ end is flanked by two unique restriction sites, *Bam*HI and *Not*I. pHV-4 was digested with *Sac*I–*Sal*I, and the *Sac*I–*Sal*I fragment (nucleotides 4–941) was isolated and subcloned into *Sac*I–*Sal*I-digested pBluescript II KS+. The subclone was designated as pBS1. A pair of primers, an upstream primer CP1, 5′-CCCTAAATTTC**CATATG**TCTC3-′, containing Hv190SV sequence nucleotides 278–296 in which two nucleotides (boldface type) were changed to generate a recognition site for *Nde*I (underlined), and a downstream T3 promoter primer were used to direct the amplification of the DNA fragment (nucleotides 278–941) with plasmid pBS1 as a template. The PCR-amplified product was digested with *Nde*I and *Sal*I and isolated by low melting point agarose gel electrophoresis. The cDNA fragments *Sal*I–*Pst*I and *Pst*I–*Bam*HI, derived from pHV-4 in separate reactions with the respective enzymes, were isolated by low melting point agarose gel electrophoresis. Three isolated fragments [*Nde*I–*Sal*I (nucleotides 290–941), *Sal*I–*Pst*I (941–3328), and *Pst*I–*Bam*HI (3328–5178)] were ligated into *Nde*I–*Bam*HI-digested pET-22b(+) to produce recombinant plasmid pET-HV1. Expression of CP in bacteria followed the manufacturer's instructions (Novagen).

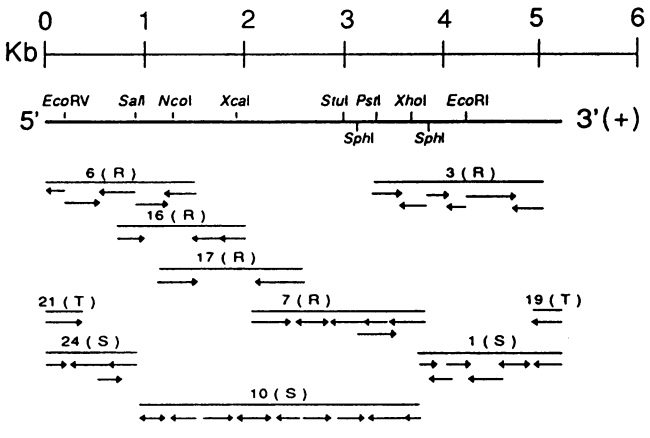

FIG. 1. Representative overlapping cDNA clones (horizontal lines) spanning the entire length of the Hv190SV dsRNA genome. The arrows indicate the size of cDNA fragment (subclone) sequenced and the direction of sequencing. Clones 3, 6, 7, 16, and 17 were synthesized using random hexamer primers. Clones 19 and 21 were synthesized using poly(A)-tailed dsRNA template and oligo(dT) primers. Clones 1, 10, and 24 were synthesized using synthetic oligonucleotide primers based on known sequences. R, random primers; S, synthetic primers; T, oligo(dT) primers.

Expression of 3′-proximal ORF (ORF2) in E. coli. A PCR-based approach was used to amplify a cDNA fragment containing the gene encoding RDRP (nucleotides 2605–4917) as follows. Plasmid pHV-4 was digested by *Spe*I and *Pst*I, and the *Spe*I–*Pst*I fragment (nuceotides 2204–3328) was subcloned in pBluescript KS(+) to generate plasmid pBS2 and amplified by PCR using a T7 promoter primer and a primer (5′-GGGA**CCATGG**TGATCCTCAGGAA-3′) identical to nucleotides 2600–2622 except for two nucleotide changes (bold) made to create an *Nco*I site (underlined). The PCR product was digested with *Nco*I–*Pst*I and isolated by low melting point agarose gel electrophoresis. The *Nco*I–*Pst*I fragment (nucleotides 2605–3328) and another cDNA fragment (*Pst*I–*Bam*HI; nucleotides 3328–5178) were ligated into *Nco*I–*Bam*HI-digested pET21d(+), and the recombinant bacterial expression vector was designated pET-HV2.

Two in-frame deletions of the RDRP gene were produced as follows. Plasmid pBS2 (nucleotides 2204–3328) was digested with *Nde*I–*Pst*I, and the isolated *Nde*I–*Pst*I fragment (nucleotides 2683–3328) and the *Pst*I–*Bam*HI (3328–5178) were ligated into *Nde*I–*Bam*HI-digested pET22b(+) to produce plasmid pET-HV3. The second in-frame deletion that contained nucleotides 2990–5178 was constructed by digesting pHV-4 with *Stu*I–*Eco*RI and the isolated fragment (nucleotides 2990–4276) was subcloned into *Sma*I–*Eco*RI-digested pBluescript II KS+. The subclone was then digested with *Bam*HI–*Eco*RI, and the isolated *Bam*HI–*Eco*RI fragment and the *Eco*RI–*Not*I fragment (nucleotides 4276–5178) were ligated into *Bam*HI–*Not*I-digested pET22b(+) to produce pET-HV4. Expression of the RDRP constructs in bacteria followed the manufacturer's instructions (Novagen). A C-terminal fragment of RDRP, expressed in bacteria transformed with pET-HV4, was purified from SDS/polyacrylamide gels and used to prepare a polyclonal antiserum to RDRP in rabbits.

Mutational Analysis of CP Stop Codon. In addition to the UGA at position 2606–2608, the predicted termination codon of ORF1, there are two other out of phase UGAs in close proximity. It was therefore important to verify that the UGA at position 2606–2608 is indeed the stop codon for ORF1. For this purpose, mutational analysis using a PCR-based approach was used to individually mutate these UGAs to sense codons. The first UGA at position 2587–2589 was mutated to AGA, and the second UGA (nucleotides 2606–2608) was mutated to UGG. Isolated cDNA fragments containing the mutated UGAs were separately exchanged with the corresponding wild-type fragments in the bacterial expression vector pET-HV1, and the expression of the mutated CPs followed the manufacturer's instructions (Novagen).

RESULTS AND DISCUSSION

Nucleotide Sequence of the Hv190SV dsRNA. The complete sequence, 5178 bp, of Hv190SV dsRNA was determined from a series of overlapping cDNA clones spanning the entire length of dsRNA (Fig. 2). The similarity of Hv190S dsRNA 5′-terminal sequence (5′-GAAAUUUA) to that of the ScV-L-A (5′-GAAAAAUUUUU) (20) was useful in identifying the 5′-terminal clones synthesized using poly(A)-tailed dsRNA templates and oligo(dT)$_{12-18}$ as primers. Likewise, the similarity of the 3′-terminal sequence of Hv190SV (3′-CGAAAAACC) to that of L-A (3′-CGUAUACCC) was also helpful in this regard. The sequences of at least three 5′- and 3′-terminal cDNA clones generated from poly(A)-tailed dsRNA were reproducible, verifying that the 5178-bp sequence (Fig. 2) assembled from the overlapping cDNA clones represents the full-length viral genome.

Computer-assisted analysis of the sequence of the full-length cDNA clone revealed the presence of two large overlapping ORFs. The large ORF1 initiates at the AUG at nucleotide

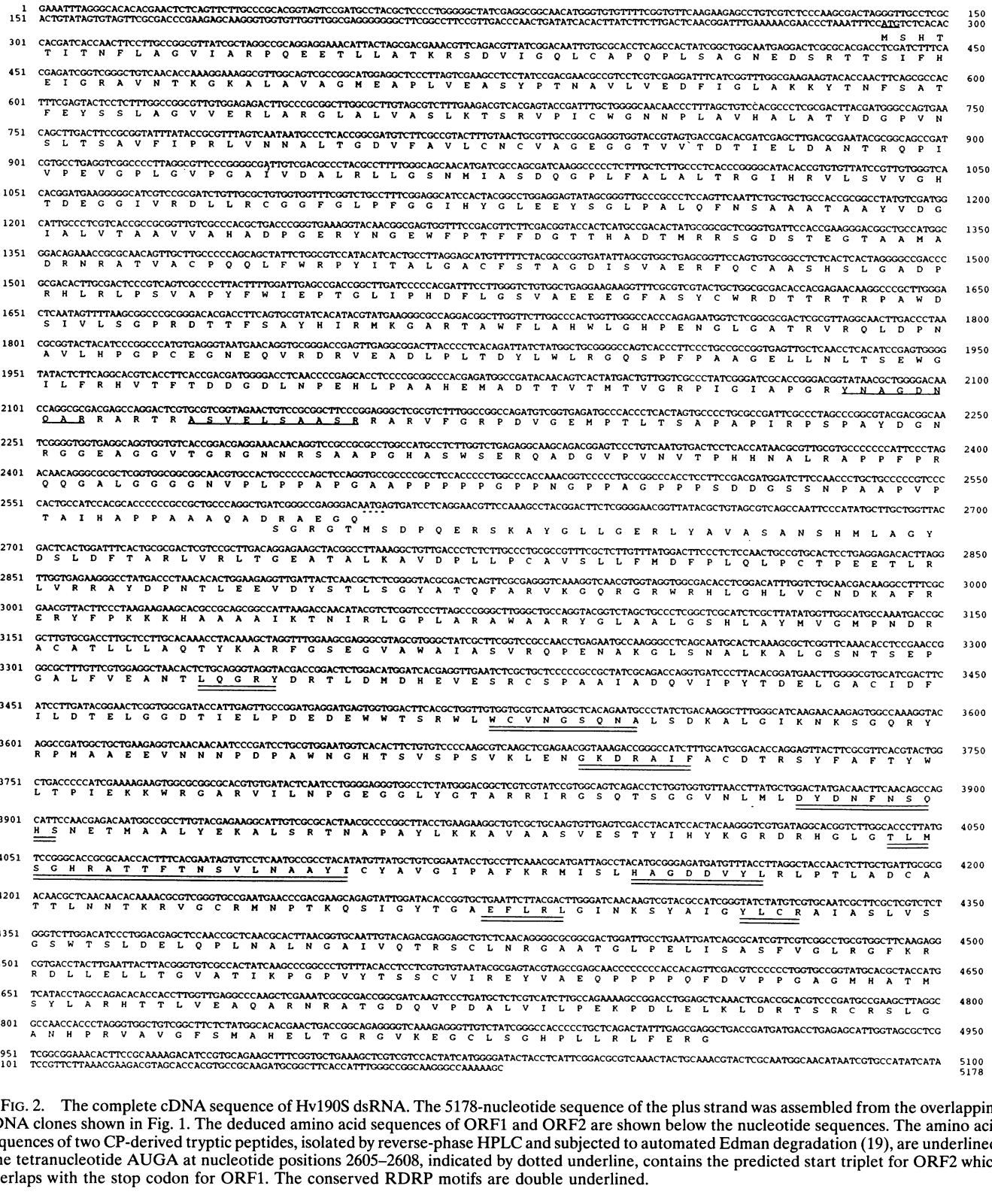

FIG. 2. The complete cDNA sequence of Hv190S dsRNA. The 5178-nucleotide sequence of the plus strand was assembled from the overlapping cDNA clones shown in Fig. 1. The deduced amino acid sequences of ORF1 and ORF2 are shown below the nucleotide sequences. The amino acid sequences of two CP-derived tryptic peptides, isolated by reverse-phase HPLC and subjected to automated Edman degradation (19), are underlined. The tetranucleotide AUGA at nucleotide positions 2605–2608, indicated by dotted underline, contains the predicted start triplet for ORF2 which overlaps with the stop codon for ORF1. The conserved RDRP motifs are double underlined.

position 290 (underlined) and terminates at the UGA at nucleotide position 2608. The first initiation codon in the ORF2 is the AUG at position 2605, and ORF2 terminates at the UGA at position 4917. The overlap region (16 nucleotides) between the two ORFs is considerably smaller than that in ScV-L-A (130 nucleotides), LRV1-1 and LRV1-4 (71 nucleotides), and GLV (122 nucleotides). The overlap region of these totiviruses contain sufficient information (structures necessary for ribosomal frameshifting including a slippery site and a pseudoknot structure involving a predicted stem-loop structure) to promote fusion of ORF1 and ORF2 *in vivo*. The overlap region in Hv190SV dsRNA genome, on the other hand, lacks a heptamer slippery site and a potential pseudoknot structure cannot be predicted from the secondary structure of the sequences flanking the 3′-terminal region of the CP gene (data not shown). These observations suggest that expression of RDRP occurs by a mechanism different from translational frameshifting.

The 5′ Region Upstream of ORF1. There are two short ORFs upstream of ORF1 (in different reading frames) that can potentially code for polypeptides of 36 and 23 amino acid residues. Secondary structure analysis of this region, however, indicates that it is highly structured (Fig. 3) and suggest that these "minicistrons" may not be translated. It is of considerable interest to note that the 11-nucleotide sequence extending from nucleotides 1–11 is complementary to the sequence starting at nucleotides 278–288, and that the resultant secondary structure (Fig. 3) renders the AUG at nucleotide position 290 (the putative start codon for the CP gene) closest to the 5′ end of the plus strand. It will be of interest to determine whether all or any of the 11 5′-terminal nucleotide are required for efficient translation of ORF1.

Like other totiviruses, the 5′ end of the plus strand is not capped since treatment of dsRNA with alkaline phosphatase followed by polynucleotide kinase and [γ-^{32}P]ATP produced molecules with labeled 5′ termini (data not shown). Since the plus strand of Hv190SV dsRNA contains a relatively long 5′ untranslated region (UTR) (289 nucleotide) with two noninitiator AUGs, it must then utilize a cap-independent mechanism to express its genome. It is not known whether the 5′ UTR of Hv190SV RNA may serve as an internal ribosome entry site, comparable to the 5′ UTR of the picornaviral mRNAs (22), and thus initiates translation specifically at internal AUG codons in a cap-independent manner. That the 5′ UTR of a totivirus may function as an internal ribosome entry site element was supported in a recent study with the LRV1 system (23).

ORF1 Encodes the Viral Capsid Protein p88. Previous results indicated that the capsid protein p88 is the major *in vitro* translation product of either the denatured Hv190SV dsRNA (24) or the full-length *in vitro* transcript of genomic dsRNA (13), suggesting that the ORF1 codes for the CP. Direct evidence for this was provided by amino acid sequencing of two internal tryptic peptides derived from p88; the sequence of one peptide (YNAGDNQAR) matched perfectly the amino acid sequence deduced from the cDNA sequence at nucleotides 2084–2110 (Fig. 2, underlined) and the sequence of the second peptide (ASVELSAAAR) matched perfectly that deduced from the sequence at nucleotides 2126–2155 (Fig. 2, underlined).

The nucleotide sequencing data predict that ORF1 of Hv190SV dsRNA (2319 bp) can encode a CP of 772 amino acid residues with molecular mass of 81.2 kDa and pI of 5.85. This value (81 kDa) is closer to that of p78 and p83 than to p88, suggesting that either p78 or p83 may be the primary translation product of ORF1. We have previously proposed that p78 is the primary translational product of ORF1 and that p83 and p88 represent postranslational phosphorylation products of p78 (1, 12). We reasoned that the separation of these three closely related proteins (p78, p83, and p88) in SDS/polyacrylamide gels was due to their differential binding to SDS as determined by their state of phosphorylation (1).

Bacterially expressed CP from cDNA constructs that initiate translation from the AUG at nucleotide position 290 comigrated with p88 (Fig. 4*B*, lanes 2 and 5). Since bacterially expressed eukaryotic phosphoproteins are not expected to be phosphorylated (assuming autophosphorylation does not occur), these results suggest that p88 is the primary translation

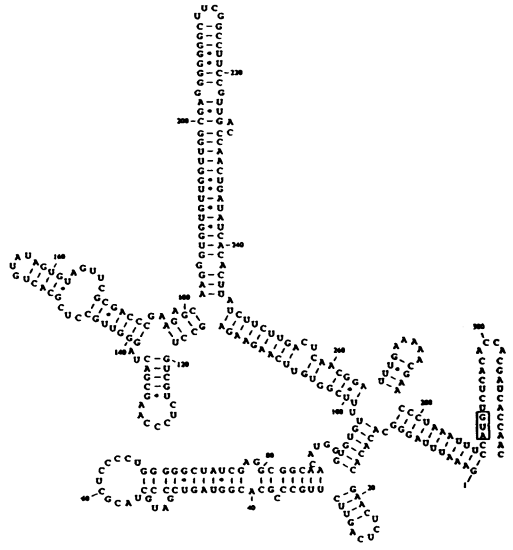

FIG. 3. Predicted secondary structure in the 5′-UTR (289 nucleotide) region of the positive strand of Hv190SV dsRNA, as determined by the GCG fold program (21). The CP start codon (290-AUG) is enclosed in a rectangle. Note the sequence complementarity between nucleotide 1–11 and nucleotides 278 to 288 and the ensuing secondary structure that makes the AUG initiating at nucleotide 290 the closest to the 5′ end of the plus strand.

A

	-1							0	+1	
2586	CT	GAT	CGG	GCC	GAG	GGA	CAA	TGA	GTG	AT 2613

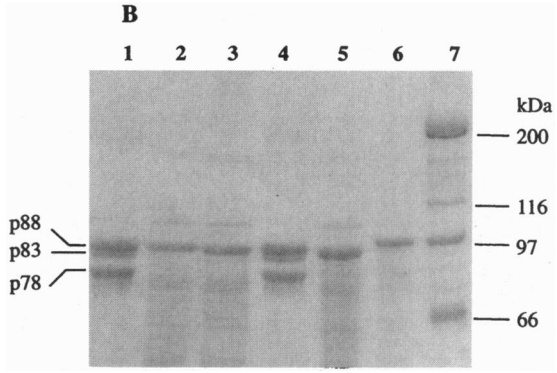

FIG. 4. Expression of wild type and mutatagenized Hv190SV CP gene (ORF1) in *E. coli*. The nucleotide sequence (nucleotides 2586–2613) flanking the predicted stop codon (TGA, 0) of ORF1 is shown in *A*; the presence of two additional TGAs (in the -1 and +1 frames relative to ORF1) is indicated by horizontal lines over the respective triplets. A Coomassie-stained SDS polyacrylamide gel of bacterially expressed wild type and mutated CP is shown in *B*. Lanes: 1 and 4, purified Hv190S virions; 2 and 5, bacterially expressed products from wild-type CP construct; 3 and 6, expressed products from constructs in which the TGA triplet in the -1 frame or the 0 frame, respectively, was separately mutated to a sense codon; 7, molecular mass protein standards. The positions of virions p88, p83, and p78 are indicated to the left.

product of ORF1. This conclusion is consistent with recent results indicating that p78 represents a posttranslational proteolytic processing product of p88 at its C-terminal region (unpublished work). The reasons why an estimate of 88 kDa was obtained by SDS/PAGE for the CP with the predicted 81-kDa size remain unknown. Interestingly, most estimates by SDS/PAGE of the CP size of ScV-L-A is about 88 kDa even though its predicted molecular mass is 76 kDa (25).

Although present evidence indicates that translation of ORF1 starts at the AUG at position 290, this initiation codon resides in an unfavorable sequence context (UCCAUGU). The next in frame AUG is at position 503 and is in a strong context (GGCAUGG). However, the bacterially expressed protein from AUG at 503 (predicted molecular mass of 74 kDa) migrated significantly faster than p88 in SDS polyacrylamide gels (data not shown). Furthermore, we have recently shown that all three CPs (p78, p83, and p88) have common N-terminal sequence since they reacted equally strong with antibodies raised against the peptide encoded by the sequence starting at nucleotides 290 to 502 (unpublished work). Because the N termini of all three CPs were found to be blocked, we could not resort to direct amino acid sequencing to provide unequivocal evidence for the start site of ORF1. It is of interest in this regard that the CP initiator codons for at least two other totiviruses [ScV-L-A (CCCAUGC) and GLV (C]CGAUGG)], have also weak or relatively weak sequence contexts.

In addition to the UGA at position 2606–2608, the predicted termination codon of ORF1, there are two other out-of-frame UGAs in close proximity (Fig. 4*A*). These are the UGA at position 2587–2589 (-1 frame relative to ORF1) and the UGA at position 2610–2612 (+1 frame relative to ORF1). It was therefore important to verify that the UGA at position 2606–2608 is the stop codon for ORF1. For this purpose, site-directed mutagenesis of the predicted CP stop codon as well as the upstream UGA was carried out. Bacterial expression of the construct in which the upstream UGA was mutated yielded a product indistinguishable in size from wild-type p88 (Fig. 4*B*, compare lanes 2 and 3). On the other hand, mutating the putative CP stop codon yielded a fusion product of estimated molecular mass of 97 kDa (Fig. 4*B*, lane 6). These results verified that the stop codon at position 2606–2608 is indeed the termination codon for the CP ORF. Because mutating the CP stop codon did not result in the production of a CP–RDRP fusion protein with a calculated molecular mass of 165 kDa, expression of RDRP as a CP–RDRP fusion protein by a readthrough mechanism can be ruled out.

ORF2 Encodes RDRP. The predicted amino acid sequence of ORF2 contains all eight conserved motifs (Fig. 2, double underlined) characteristic of RDRPs from dsRNA viruses of simple eukaryotes (7). Of the 70 amino acid positions contained in the 8 conserved motifs, Hv190SV RDRP is identical in 48, 46, 40, 38, and 21 positions, respectively, to the RDRPs of LRV1, ScV-L-A, Ustilago maydis virus H1, ScV-La, and GLV. These results are in agreement with the previous conclusion (7) that the RDRPs of dsRNA viruses of simple eukaryotes (except for GLV) are very closely related.

In Western blots of denatured Hv190S virions, RDRP was detected as a minor virion-associated polypeptide using a polyclonal antiserum to a bacterially expressed C-terminal fragment of RDRP (Fig. 5*B*, lane 1). The virion-associated RDRP comigrated with a protein expressed in bacteria from a cDNA construct that initiates translation from the AUG at position 2605 and terminates at the UGA at position 4917 (Fig. 5*B*, lane 2). An M_r value of 92×10^3 was estimated for RDRP by SDS/PAGE. This value is higher than that calculated from the deduced amino acid sequence of RDRP ORF (molecular mass of 84.3 kDa and pI of 9.01). No proteins with molecular mass higher than 92 kDa were detected in Hv190S virions by the RDRP-specific antiserum suggesting that fusion proteins (CP–RDRP) comparable to those associated with the virions

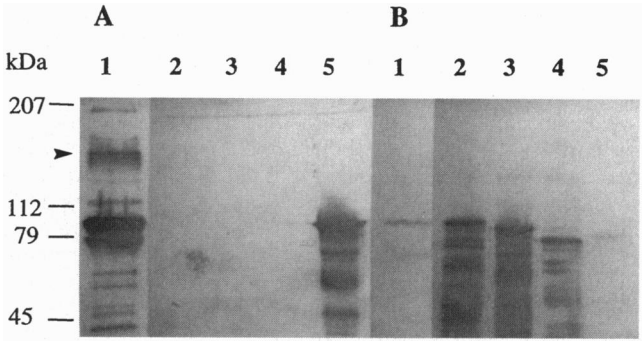

FIG. 5. Western blot analysis of bacterially expressed full-length and in-frame deletions of Hv190SV ORF2 (RDRP gene). Transferred proteins were detected by using antisera to virion p88 (*A*) and to a C-terminal fragment of RDRP (*B*). Lanes: 1, purified Hv190S virions; 2, bacterially expressed full-length RDRP (construct pET-HV2); 3 and 4, truncated RDRP expressed from in-frame deletions of ORF2 (constructs pET-HV3 and -HV4, respectively); 5, bacterially expressed full-length CP. Sizes of the protein standards are indicated to the left.

of the totiviruses ScV-L-A and GLV were not present (11, 20). Stable dimers of each of the capsid proteins p78, p83, and p88 were detected by the antiserum to p88 (Fig. 5*A*, arrowhead), and this was initially a cause for confusion because of the similarity in size to a hypothetical CP–RDRP fusion protein.

Although the possibility that RDRP is expressed as a CP–RDRP fusion protein that is rapidly processed cannot be dismissed, it appears unlikely since generation of a fusion protein by a readthrough mechanism was ruled out (this study) and fusion via translational frameshifting is difficult to justify. The overlap region between the CP and RDRP ORFs lacks a heptamer slippery site and an RNA pseudoknot structure cannot be predicted from the secondary structure flanking the 3′ terminus region of ORF1. The small size of the overlap region between the CP and RDRP ORFs in Hv190SV does not rule out the possibility of translational frameshifting. The overlap region between the 39 kDa protein and the 60 kDa putative RNA polymerase of the luteovirus barley yellow dwarf virus is 13 nucleotides, yet the ribosomes slip into the −1 frame in the overlap region that includes a consensus slippery heptanucleotide (26). The recent finding that ORF1 and ORF2 of the totivirus LRV2-1 do not overlap (9) is of interest in this regard. Because a potential pseudoknot structure was found near the 3′-terminal region of the viral CP gene, the authors proposed that fusing of the two ORFs may occur by ribosomal hopping since it cannot be produced by frameshifting.

The finding that the termination codon of ORF1 (the UGA at nucleotide positions 2606–2608) overlaps with the predicted start codon for the RDRP ORF (AUG at nucleotide positions 2605–2607) suggests that RDRP is translated by an internal initiation mechanism. The overlapping of these two ORFs may serve as a means for regulating the level of the internally encoded product (RDRP). By analogy to the yeast L-A virus (27), only 1–2 RDRP molecules are assembled with 120 CP molecules to form the mature virions. It is well known that the upstream reading frame in dicistronic eukaryotic mRNAs can profoundly influence the level of internal translation initiation depending on the length of the intercistronic region (28–30); the longer the intercistronic region the more efficient the level of internal translational initiation. Therefore, the initiation of translation at a downstream AUG codon would be expected to be markedly suppressed when it is overlapped by an upstream cistron.

It has been previously shown that internal ribosome entry site elements can direct internal initiation of a second cistron in a dicistronic mRNA (31). Experiments involving the generation of dicistronic constructs with the Hv190SV 5′-UTR sequences, or the 3′-terminal region of the CP ORF, inserted

between the two cistrons are underway to verify whether these leader sequences possess internal ribosome entry site elements and whether they can mediate translation of the downstream cistrons.

In conclusion, our study presents the first reported sequence for a totivirus infecting a filamentous fungus. Like other members of the genus *Totivirus*, Hv190SV contains two large overlapping ORFs with the 5′-proximal ORF encoding a CP and the 3′-ORF encoding an RDRP. Whereas the RDRPs of other totiviruses are expressed as CP–RDRP fusion proteins, the Hv190SV-RDRP, as shown in this study, is expressed as a separate nonfused virion-associated polypeptide. We also describe the structural features of the 5′ UTR and the overlap region between the two ORFs and suggest that both ORFs are expressed by internal initiation of translation. Future experiments will be designed to elucidate the nature of the sequence elements regulating the expression of Hv190SV CP and RDRP. Studies on the expression of viral genes via internal initiation mechanisms are presently attracting considerable general interest (22). Because Hv190SV is associated with a debilitating disease of its fungal host, knowledge of its genome structure and expression strategy will be valuable for understanding its role in pathogenesis and potential use as a biocontrol agent.

1. Ghabrial, S. A. (1994) *Adv. Virus Res.* **43,** 303–388.
2. Nuss, D. L. & Koltin, Y. (1990) *Annu. Rev. Phytopathol.* **28,** 37–58.
3. Ghabrial, S. A. (1986) in *Fungal Virology*, ed. Buck, K. W. (CRC, Boca Raton, FL), pp. 163–176.
4. Sanderlin, R. S. & Ghabrial, S. A. (1978) *Virology* **87,** 142–151.
5. Ghabrial, S. A., Bruenn, J. A., Buck, K. W., Wickner, R. B., Patterson, J. L., Stuart, K. D., Wang, A. L. & Wang, C. C. (1995) in *Virus Taxonomy: Sixth Report of the International Committee on Taxonomy of Viruses*, eds. Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A. & Summers, M. D. (Springer, New York), pp. 245–252.
6. Park, C.-M., Lopinski, J. D., Masuda, J., Tzeng, T.-H. & Bruenn, J. A. (1996) *Virology* **216,** 451–454.
7. Bruenn, J. A. (1993) *Nucleic Acids Res.* **21,** 5667–5669.
8. Scheffter, S., Widmer, G. & Patterson, J. L. (1995) *Virology* **212,** 82–90.
9. Scheffter, S., Ro, Y. T., Chung, I. K. & Patterson, J. L. (1995) *Virology* **199,** 479–483.
10. Stuart, K. D., Weeks, R., Guilbride, L. & Myler, P. J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8596–8600.
11. Wang, A. L., Yang, H. -M., Shen, K. A. & Wang, C. C. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8595–8599.
12. Ghabrial, S. A. & Havens, W. M. (1992) *Virology* **188,** 657–665.
13. Ghabrial, S. A. & Havens, W. M. (1989) *J. Gen. Virol.* **70,** 1025–1035.
14. Hiremath, S., L'Hostis, B., Ghabrial, S. A. & Rhoads, R. E. (1986) *Nucleic Acids Res.* **14,** 9877–9896.
15. Asamizu, T., Summers, D., Motika, M. B., Anzola, J. V. & Nuss, D. E. (1985) *Virology* **144,** 398–409.
16. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
17. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
19. Matsudaira, P. (1993) in *A Practical Guide to Protein and Peptide Purification for Microsequencing*, ed. Matsudaira, P. (Academic, New York), 2nd Ed., pp. 3–13.
20. Icho, T. & Wickner, R. B. (1989) *J. Biol. Chem.* **264,** 6716–6723.
21. Zucker, M. (1989) *Methods Enzymol.* **180,** 262–288.
22. Meerovitch, K., & Sonenberg, H. (1993) *Semin. Virol.* **4,** 217–227.
23. Maga, J. A., Widmer, G. & LeBowitz, J. H. (1995) *Mol. Cell. Biol.* **15,** 4884–4889.
24. Ghabrial, S. A., Bibb, J. A., Price, K. H., Havens, W. M. & Lesnaw, J. A. (1987) *J. Gen. Virol.* **68,** 1791–1800.
25. Tipper, D. J. & Bostian, K. A. (1984) *Microbiol. Rev.* **48,** 125–156.
26. Di, R., Dinesh-Kumar, S. P. & Miller, W. A. (1993) *Mol. Plant–Microbe Interact.* **6,** 444–452.
27. Cheng, R. H., Caston, J. R., Wang, G.-J., Gu, R., Smith, T. J., Baker, T. S., Bozarth, R. F., Trus, B. L., Cheng, N., Wickner, R. B. & Steven, A. C. (1994) *J. Mol. Biol.* **244,** 255–258.
28. Peabody, D. S., Suresh, S. & Berg, P. (1986) *Mol. Cell. Biol.* **6,** 2704–2711.
29. Kozak, M. (1987) *Mol. Cell. Biol.* **7,** 3438–3445.
30. Iizuka, N., Najita, L., Franzusoff, A. & Sarnow, P. (1994) *Mol. Cell. Biol.* **14,** 7322–7330.
31. Jackson, R. J., Howell, M. T. & Kaminski, A. (1990) *Trends Biochem. Sci.* **15,** 477–483.