Molecular characterization of *Penicillium chrysogenum virus*: reconsideration of the taxonomy of the genus *Chrysovirus*

Daohong Jiang† and Said A. Ghabrial

Correspondence
Said A. Ghabrial
saghab00@uky.edu

Department of Plant Pathology, University of Kentucky, 201F Plant Science Building, 1405 Veterans Drive, University of Kentucky, Lexington, KY 40546-0312, USA

Molecular cloning and complete nucleotide sequencing of *Penicillium chrysogenum virus* (PcV) dsRNAs indicated that PcV virions contained four dsRNA segments with sizes of 3562, 3200, 2976 and 2902 bp. Each dsRNA segment had unique sequences and contained a single large open reading frame (ORF). *In vitro* translation of transcripts derived from full-length cDNA clones of PcV dsRNAs yielded single products of sizes similar to those predicted from the deduced amino acid sequences of the individual ORFs. Sequence similarity searches revealed that dsRNA1 encodes a putative RNA-dependent RNA polymerase. In this study, it was determined that dsRNA2 encodes the major capsid protein and that p4, encoded by dsRNA4, is virion-associated as a minor component. All four dsRNAs of PcV, like the genomic segments of viruses with multipartite genomes, were found to have extended regions of highly conserved terminal sequences at both ends. In addition to the strictly conserved 5’-terminal 10 nt, a second region consisting of reiteration of the sequence CAA was found immediately upstream of the AUG initiator codon. These (CAA)n repeats are reminiscent of the translational enhancer elements of tobamoviruses. The 3’-terminal 14 nt were also strictly conserved. As PcV and related viruses with four dsRNA segments (genus *Chrysovirus*) have not been previously characterized at the molecular level, they were provisionally classified in the family *Partitiviridae*, comprising viruses with bipartite genomes. This study represents the first report on molecular characterization of a chrysovirus and the results suggest the creation of a new family of mycoviruses with multipartite dsRNA genomes to accommodate PcV and related viruses.

INTRODUCTION

*Penicillium chrysogenum virus* (PcV) was one of the first fungal viruses to be isolated and subjected to extensive biochemical, biophysical and ultrastructural studies (Lemke & Ness, 1970; Wood & Bozarth, 1972; Yamashita et al., 1973; Buck & Girvan, 1977; Edmondson et al., 1984). The discovery of PcV particles in many *P. chrysogenum* industrial strains used for penicillin production generated considerable interest in the study of *Penicillium* viruses, as it was believed that virus infection might be responsible for the instability of some of these strains (Banks et al., 1969; Lemke et al., 1973). Although the results of earlier studies on the biochemical and biophysical properties of PcV were mostly in agreement, they differed in their explanation of the nature of genome complexity.

PcV and the related *Penicillium* viruses *Penicillium brevicompactum virus* (PbV) and *Penicillium cyaneo-fulvum virus* (Pc-fV) have isometric particles, 35–40 nm in diameter, and contain four separately encapsidated dsRNA segments (Wood et al., 1971; Wood & Bozarth, 1972; Buck & Girvan, 1977; Edmondson et al., 1984; Castón et al., 2003). As they have not been previously characterized at the molecular level, these three viruses (grouped under the genus *Chrysovirus*) were provisionally placed in the family *Partitiviridae* (Ghabrial et al., 2000) with the assumption that their genomes are bipartite, with dsRNA1 encoding the RNA-dependent RNA polymerase (RDRP) and dsRNA2 encoding the major capsid protein (CP). The additional dsRNAs (dsRNAs 3 and 4), like those of some partitiviruses, were presumed to be defective or satellite dsRNAs (Ghabrial et al., 2000; Ghabrial, 2002). In the present study, we report the complete nucleotide sequence and genome organization of each of the four monocistronic dsRNA segments associated with PcV virions and discuss similarities to viruses with multipartite and multicomponent RNA genomes. Based on the consistent co-presence of their four dsRNA segments, the existence of extended regions of highly conserved terminal sequences at both ends of all four segments, sequence comparisons and
phylegenetic analysis, PcV and related viruses should not be classified with the family Partitiviridae. Our results support the creation of a new family of fungal viruses with multipartite genomes.

**METHODS**

**Fungal culture and virion purification.** *Penicillium chrysogenum* strain ATCC 9480 was used as a source of PcV. Mycelium was harvested from 10 day stationary cultures grown on potato dextrose broth containing 0·5 % yeast extract and homogenized in a blender with 0·1 M sodium phosphate buffer, pH 7·4 containing 0·2 M KCl and 0·5 % β-mercaptoethanol. The homogenate was mixed with an equal volume of chloroform and the emulsion was separated by centrifugation at 8000 g for 20 min. The aqueous layer was then subjected to two cycles of differential centrifugation (27 000 r.p.m. for 150 min in a Beckman type 30 rotor and 10 000 r.p.m. for 10 min in a Beckman JA-20 rotor). Pellets were resuspended in 50 mM Tris/HCl buffer, pH 7·6. The final step in purification was made by rate zonal centrifugation in sucrose density gradients (100–400 mg ml⁻¹). The gradients were made in 50 mM Tris/HCl buffer (pH 7·6) and centrifuged at 24 000 r.p.m. in a Beckman SW28 rotor for 150 min. The broad virus band was withdrawn with a syringe from the side of the tube and diluted with buffer and the virions were concentrated by overnight centrifugation at 40 000 r.p.m. in a Beckman 50Ti rotor. The pellets were resuspended in 50 mM Tris/HCl buffer (pH 7·6).

**Extraction of viral dsRNA and Northern hybridization analysis.** Viral dsRNA was isolated from purified virions (100 μg virions resuspended in 100 μl DEPC-treated water) by SDS/phenol extraction and ethanol precipitation. The pellet was washed twice with 70 % ethanol and resuspended in TE buffer, pH 8·0. For Northern hybridization analysis, dsRNAs were separated on a 1·5 % agarose gel for 6–7 h at 70 V. The gel was then soaked for 20 min in 0·1 M NaOH, followed by soaking in 0·1 M Tris/HCl, pH 8·0, for 20 min, and the RNA was transferred by capillary action to a Hybond-N+ nylon membrane (Amersham) in 10 × SSC buffer. Prehybridization and hybridization were performed under high stringency conditions in a hybridization solution (5 × Denhardt’s reagent, 6× SSC, 20 mM Tris/HCl, pH 7·5, 0·5 % SDS) containing 50 % formamide for 14–16 h at 42 °C. The RNA blots were probed with 32P-labelled probes prepared by random-primer labelling of cloned cDNA to the PcV dsRNAs.

**cDNA synthesis and molecular cloning.** cDNA clones of the viral dsRNA were synthesized using the T7 panucrRNA cDNA synthesis kit (Amerham Pharmacia Biotech). Virion dsRNA (2 μg) was mixed with 0·74 μg random hexamers and incubated in 90 % DMSO at 65 °C for 30 min, ethanol-precipitated and resuspended in 20 μl DEPC-treated water. First and second strand cDNA synthesis and addition of EcoRI–NotI adaptors were performed according to the manufacturer’s instructions. The cDNAs were inserted into the EcoRI and NotI sites of a Bluescript vector using T4 DNA ligase and recombinant plasmids were transformed into E. coli strain DH5α. Clones corresponding to the different dsRNA segments were identified by Northern hybridization analysis of viral dsRNA using the cloned cDNAs as probes.

Gaps between non-overlapping cDNA clones were connected using RT-PCR and sequence-specific primers. Viral dsRNA (5–10 μg) was denatured by incubation in 90 % DMSO at 90 °C for 5 min and then chilled on ice for 5 min. The denatured dsRNA was reverse transcribed using Superscript II reverse transcriptase (BRL) and a sequence-specific reverse primer. The reaction mixture was incubated for 60 min at 45 °C followed by treatment with RNase H (2 U, at 37 °C for 20 min) and 5 % of the reaction volume (1 μl) was used for PCR amplification by using the pertinent sequence-specific forward and reverse primers. PCR was carried out by using Platinum Pfu DNA polymerase (BRL) and cycling parameters for ‘touch-down’ PCR were 94 °C for 4 min; 94 °C for 1 min, 65–50 °C (–0·5 °C per cycle) for 2 min, 72 °C for 3 min, 30 cycles; 94 °C for 1 min, 50 °C for 2 min, 72 °C for 3 min, 10 cycles; 72 °C for 12 min. PCR products were purified with QIAquick gel extraction kit (Qiagen), A-tailed with Taq DNA polymerase at 70 °C for 30–45 min and cloned into the pGEM-T Easy cloning vector (Promega).

In addition, clones for terminal sequences of the four dsRNAs were generated by T4 RNA ligase oligo-mediated amplification. T4 poly-nucleotide kinase-phosphorylated oligo-x-lig [5′-CCATATGCGGCC- GCTAGCAGATCTCTAAAGGCGAC-(NH2)-3′] was ligated to the 3′ ends of each strand of the viral dsRNAs using T4 RNA ligase (NEB) in 1 × RNA ligase buffer, at 4 °C for 18 h. The unligated oligonucleotide was removed from the RNA ligation reaction by membrane filtration using a Microcon YM-100 (Amicon; Millipore). The oligo-ligated dsRNA was denatured by incubation in 90 % DMSO at 90 °C for 5 min and then quickly chilled on ice for 5 min. An aliquot equivalent to 5 % of the reaction mixture was used for the reverse transcription reaction with Superscript II reverse transcriptase (BRL) in 1 × first strand buffer, 10 mM DTT, 0·5 mM dNTPs in a total reaction volume of 20 μl at 48 °C for 30 min and 5 pmol of a primer with sequence complementary to the oligonucleotide used for RNA ligation (oligo-REV, 5′-TGCGCTTTAGCTCAGATT-3′). The reaction mixture was then treated with RNase H (2 U, at 37 °C for 20 min) and a sample equivalent to 5 % of the reaction mixture was used for PCR amplification with the oligo-REV primer and a sequence-specific primer corresponding to dsRNA1, 2, 3 or 4. PCR conditions were as described above by using Platinum Pfu DNA polymerase (BRL) and cycling parameters for ‘touch-down’ PCR. The PCR products were gel-purified using a gel extraction kit (Qiagen), A-tailed with Taq DNA polymerase (5 U, in 1 × reaction buffer, 1·5 mM MgCl2, BRL) and dATP (0·2 mM) and then cloned into pGEM-T Easy.

**Generation of full-length cDNA clones and in vitro translation.** Full-length cDNA clones of the four dsRNAs were produced using sequence-specific primers and the RT-PCR protocol described above. Forward and reverse primers of the four dsRNAs corresponding to the 5′- and 3′-terminal nucleotide sequences were designed and used for RT-PCR. The PCR products were gel-purified using a gel extraction purification kit (Qiagen) and cloned into pGEM-T Easy.

**In vitro translation.** In vitro translation was performed using the TNT T7 Quick coupled transcription–translation kit (rabbit reticulocyte lysates; Promega) in the presence of [35S]methionine (RediMix; Amersham). Recombinant plasmids were linearized with the pertinent restriction enzymes (SpeI for dsRNA1; SalI for dsRNAs 2 and 3; NdeI for dsRNA4) and purified with the Concert Rapid PCR purification system (Gibco-BRL, Life Technologies) and approximately 1 μg template DNA was used for the *in vitro* translation reaction. The *in vitro* translation products were separated by SDS-PAGE on an 8 % polyacrylamide gel and analysed by autoradiography using a PhosphorImager (Molecular Dynamics).

**Protein sequence analysis.** For sequencing internal peptides of PcV CP, gradient-purified PcV virions were digested with sequencing-grade trypsin (Promega) at 37 °C for 18 h and the digestion products were separated by reverse-phase HPLC on an analytical C18 column (Vydak). Two highly resolved peptides were selected for amino acid sequencing by automated Edman degradation.

**Nucleotide sequencing and sequence analysis.** All sequencing was performed by dideoxy chain termination sequencing using the
Big Dye terminator sequencing kit (ABI) and an ABI 310 system for automated sequencing. Sequencing was carried out bidirectionally on both strands using either M13 universal primers or sequence-specific ‘walking’ primers. At least two independent clones, usually three to five clones, were analysed for sequence determination at all nucleotide positions. Paired and multiple sequence alignments were performed with the programs BESTFIT, GAP and PILEUP (University of Wisconsin GCG software package version 10). Sequence similarity searches of GenBank, Swissprot and EMBL databases were conducted using the BLAST program (Altschul et al., 1997). Sequence alignments and phylogenetic analysis were performed using the programs CLUSTAL X and PAUP* (Thompson et al., 1997; Swofford, 2000). Searches for amino acid signatures and protein motifs were conducted using the programs included in the ExPASy proteomics tools (http://www.expasy.org/tools/).

**RESULTS**

**Number and size of dsRNA segments associated with PcV virions**

Although the dsRNA extracted from purified PcV virions was resolved into three bands by agarose gel electrophoresis (Fig. 1, lane EB), cDNA cloning and nucleotide sequencing analysis revealed the presence of four distinct dsRNA segments in association with purified virions (Table 1). Northern hybridization analysis using cloned cDNA probes representing the four dsRNA segments showed that each segment has unique sequences (Fig. 1). As dsRNA3 and 4 differ in size by only 74 bp, they co-migrated when separated by agarose gel electrophoresis. These results present conclusive evidence for the presence of four dsRNA segments in association with PcV virions and explain why some earlier studies reported that PcV virions packaged only three dsRNA segments (Wood & Bozarth, 1972; Buck & Girvan, 1977).

**Nucleotide sequencing analysis and in vitro translation of genomic dsRNAs**

The complete nucleotide sequences of each of the four dsRNAs were determined from a series of cDNA clones spanning the entire length of each dsRNA (see Methods). The sequences were deposited in GenBank and the accession numbers are listed in Table 1. Sequence analysis and in vitro coupled transcription–translation assays showed that each of the four dsRNAs is monocistronic, as each dsRNA was found to contain a single major open reading frame (ORF) and each was translated into a single major product of the size predicted from its deduced amino acid sequence (Fig. 2). The origin of the polypeptide of <47 kDa (Fig. 2, lane pG-RNA1) is not known. It might have resulted from premature termination of translation. No such smaller polypeptide was observed when the corresponding dsRNA1 from the chrysovirus Hv145SV was translated (A. Soldevila and S. A. Ghabrial, unpublished).

*Table 1. Size of the dsRNA genomic segments of PcV and size and function of their encoded proteins*

<table>
<thead>
<tr>
<th>Segment (accession no.)</th>
<th>Size (bp)*</th>
<th>Size of encoded protein (Da)†</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>dsRNA1 (AF296439)</td>
<td>3562</td>
<td>128 548 (1117 aa)</td>
<td>RDRP</td>
</tr>
<tr>
<td>dsRNA2 (AF296440)</td>
<td>3200</td>
<td>108 806 (982 aa)</td>
<td>Capsid protein</td>
</tr>
<tr>
<td>dsRNA3 (AF296441)</td>
<td>2976</td>
<td>101 458 (912 aa)</td>
<td>Unknown</td>
</tr>
<tr>
<td>dsRNA4 (AF296442)</td>
<td>2902</td>
<td>94 900 (847 aa)</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

* Determined by nucleotide sequencing of cDNA clones spanning the entire length of the individual dsRNA segments.
† Determined from the predicted amino acid sequence derived from full-length cDNA clones of the individual genome segments. Numbers of amino acid residues of the encoded viral proteins are given in parentheses.
positions of size marker (M) are indicated.

amino acid sequence. Lane PcV contains purified PcV virions.

major product similar in size to that predicted from its deduced

 autoradiography. Each of the dsRNAs was translated into a

analysed by SDS-PAGE on an 8 % polyacrylamide gel and

the T N T T7 coupled transcription–translation system in the pre-

nated pG-RNA1 to pG-RNA4) were translated

in vitro (Table 1). This is consistent with the size estimated by

SDS-PAGE of PcV CP, derived from purified virions, as well as

of the in vitro translation product of full-length transcript

dsRNA2 cDNA (Fig. 2). Direct evidence that dsRNA2 encodes

PcV CP was provided by amino acid sequencing of a tryptic peptide derived from a gradient-purified PcV capsid; the sequence of the peptide (KMPTGAFTTRW-VAAK) matched perfectly to the amino acid sequence deduced from the cDNA sequence at nt positions 1712–1756 (Fig. 5, boxed). Interestingly, the sequence of a second tryptic peptide obtained from the trypsin-digested PcV capsid indicated that it was derived from the putative protein encoded by dsRNA4 (Fig. 6; see below). BLAST searches of the deduced amino acid sequence of dsRNA2 ORF showed only a single significant high similarity hit (3e−62) to the chrysovirus Hv145SV CP (AF297177; 29 % identity and 50 % aa sequence similarity). It is interesting that the region of high sequence similarity between the CPs of these two related viruses was limited to the N-terminal half of the proteins (aa 19–560 of PcV CP; data not shown). This finding may have implications when considering the structural organization of chrysovirus capsids (Castón et al., 2003).

dsRNA2

Sequence analysis of full-length dsRNA2 cDNA revealed that PcV dsRNA2 is 3200 bp in size and contains a single large ORF (dsRNA2 ORF) at nt positions 158 to 3106 (Fig. 5). A molecular mass of 109 kDa was calculated for the protein (982 aa residues) encoded by dsRNA2 ORF. This predicted size is similar to that estimated by SDS-PAGE of PcV CP, derived from purified virions, as well as of the in vitro translation product of full-length transcript of dsRNA2 cDNA (Fig. 2). Direct evidence that dsRNA2 encodes PcV CP was provided by amino acid sequencing of a tryptic peptide derived from a gradient-purified PcV capsid; the sequence of the peptide (KMPTGAFTTRW-VAAK) matched perfectly to the amino acid sequence deduced from the cDNA sequence at nt positions 1712–1756 (Fig. 5, boxed). Interestingly, the sequence of a second tryptic peptide obtained from the trypsin-digested PcV capsid indicated that it was derived from the putative protein encoded by dsRNA4 (Fig. 6; see below). BLAST searches of the deduced amino acid sequence of dsRNA2 ORF showed only a single significant high similarity hit (3e−62) to the chrysovirus Hv145SV CP (AF297177; 29 % identity and 50 % aa sequence similarity). It is interesting that the region of high sequence similarity between the CPs of these two related viruses was limited to the N-terminal half of the proteins (aa 19–560 of PcV CP; data not shown). This finding may have implications when considering the structural organization of chrysovirus capsids (Castón et al., 2003).

dsRNA1

Sequence analysis of full-length dsRNA1 cDNA indicated that it is 3562 bp in size and that it contains a single large ORF (dsRNA1 ORF) from nt positions 145 to 3498. A protein with the molecular mass of 129 kDa was calculated from the 1117 aa residues encoded by dsRNA1 ORF (Table 1). This is consistent with the size estimated by SDS-PAGE of the in vitro translation product of full-length transcripts of dsRNA1 cDNA (Fig. 2). Examination of the deduced amino acid sequence of dsRNA1 ORF revealed the presence of eight conserved motifs (Fig. 3) characteristic of RDRPs of dsRNA viruses of simple eukaryotes (Bruenn, 1993; Ghabrial, 1998). BLAST searches of the deduced amino acid sequence of dsRNA1 ORF showed that it has significantly high sequence similarity (40 % identity and 57 % aa sequence similarity) to the RDRP encoded by the chrysovirus Helminthosporium victoriae 145S virus Hv145SV (GenBank accession no. AF297176; A. Soldevila and S. A. Ghabrial, unpublished). High similarities (BLAST hits of e−16 or lower) were also found to the RDRPs of the unclassified Agaricus bisporus virus 1 (AbV1; Van der Lende et al., 1996), the totiviruses Ustilago maydis virus H1 (UmV-H1; Kang et al., 2001) and Trichomonas vaginalis virus (TVV; Su & Tai, 1996). Still high similarity hits (e−5 or lower) were obtained with the RDRPs of several members of the family Totivi-

ridae. Interestingly, no significant hits were evident with viruses in the family Partitiviridae. A comparison of the conserved motifs of PcV RDRP with those of totiviruses and partitiviruses revealed that the RDRPs of the chrysovirus PcV and Hv145SV are more closely related to those of the totiviruses than to those of the partitiviruses (Fig. 3). This conclusion was also supported by the results of phylogenetic analysis of RDRP conserved motifs and flanking sequences of the two chrysoviruses and viruses in the families Totiviridae and Partitiviridae (Fig. 4).

dsRNA2

Sequence analysis of full-length dsRNA2 cDNA revealed that PcV dsRNA2 is 3200 bp in size and contains a single large ORF (dsRNA2 ORF) at nt positions 158 to 3106 (Fig. 5). A molecular mass of 109 kDa was calculated for the protein (982 aa residues) encoded by dsRNA2 ORF. This predicted size is similar to that estimated by SDS-PAGE of PcV CP, derived from purified virions, as well as of the in vitro translation product of full-length transcript of dsRNA2 cDNA (Fig. 2). Direct evidence that dsRNA2 encodes PcV CP was provided by amino acid sequencing of a tryptic peptide derived from a gradient-purified PcV capsid; the sequence of the peptide (KMPTGAFTTRW-VAAK) matched perfectly to the amino acid sequence deduced from the cDNA sequence at nt positions 1712–1756 (Fig. 5, boxed). Interestingly, the sequence of a second tryptic peptide obtained from the trypsin-digested PcV capsid indicated that it was derived from the putative protein encoded by dsRNA4 (Fig. 6; see below). BLAST searches of the deduced amino acid sequence of dsRNA2 ORF showed only a single significant high similarity hit (3e−62) to the chrysovirus Hv145SV CP (AF297177; 29 % identity and 50 % aa sequence similarity). It is interesting that the region of high sequence similarity between the CPs of these two related viruses was limited to the N-terminal half of the proteins (aa 19–560 of PcV CP; data not shown). This finding may have implications when considering the structural organization of chrysovirus capsids (Castón et al., 2003).

dsRNA3 and 4

Sequence analysis of PcV dsRNAs 3 and 4 cDNA clones showed that they are 2976 and 2902 bp in size, respectively. Each dsRNA contains a single large ORF encoding proteins of calculated molecular mass of 101 and 95 kDa for dsRNAs 3 and 4. These values are similar to those estimated by SDS-PAGE of in vitro translation products of full-length transcripts derived from cDNAs of dsRNAs 3 and 4 (Fig. 2). PcV dsRNA3 ORF was predicted to initiate at the AUG at nt positions 162 to 164 and to terminate at the UAA at nt positions 2898 to 2900 (data not shown). dsRNA4 ORF was predicted to initiate at the AUG at nt positions 163 to 165 and to terminate at the UAG at nt positions 2704 to 2706 (Fig. 6).

BLAST searches of the deduced amino acid sequences of the proteins encoded by PcV dsRNAs 3 and 4 (designated hereafter as p3 and p4, respectively) showed high similarity
Fig. 3. Comparison of the conserved motifs of RDRPs of selected isometric dsRNA mycoviruses including the putative RDRP encoded by PcV dsRNA1. Numbers 1–8 refer to the eight conserved motifs characteristic of RDRPs of RNA viruses. The amino acid positions corresponding to conserved motifs 1 and 2 for the RDRPs of viruses in the family Partitiviridae are not well-defined and, therefore, they are not presented. Multiple sequence alignments were obtained using the CLUSTAL X program with RDRP amino acid sequences of the following viruses. Upper set: viruses in the families Totiviridae and Chrysoviridae: SsRV1, SsRV2 (Sphaeropsis sapinea RNA virus 1 and 2, respectively; Preisig et al., 1998), Hv190SV (Helminthosporium victoriae 190S virus; Huang & Ghabrial, 1996), UmV-H1 (Kang et al., 2001), Hv145SV (A. Soldevila and S. A. Ghabrial, unpublished) and PcV (this study). Lower set: viruses in the family Partitiviridae: FpV1 (Fusarium poae virus 1; Compel et al., 1999), RaV (Rhizoctonia solani virus; Strauss et al., 2000), AhV (Atkinsonella hypoxylon virus; Oh & Hillman, 1995), HaV (Heterobasidion annosum virus, Ihmark et al., 2001), BCV3 (Beet cryptic virus 3; Xie et al., 1993) and FsV1 (Fusarium solani virus 1; Nogawa et al., 1996). Asterisks signify identical residues (shaded) at the indicated position; colons signify at least four identical residues within a column; single dots signify 50% identical residues at the indicated position. Numbers in parentheses correspond to the number of amino acid residues separating the motifs.
confirmed when a p4-specific antiserum becomes available. Furthermore, p4 contains a consensus signature of the glycosyl hydrolases family 10 (Fig. 6; underlined). This conserved region is centred on a conserved glutamic acid (aa position 709; Fig. 6, boxed), which was demonstrated for one member of the glycosyl hydrolases family to be directly involved in glycosidic bond cleavage by acting as a nucleophile (Tull et al., 1991). The significance of a putative glycosyl hydrolase activity of p4 to the PcV lifecycle is not known at present.

Fig. 5. Nucleotide sequence of PcV dsRNA2 cDNA. The deduced amino acid sequence of PcV CP is indicated in the one letter code below the nucleotide sequence. The translation initiation codon ATG at nt positions 158 to 160 is underlined, the termination codon TAA at nt positions 3104 to 3106 is indicated with an asterisk. The amino acid sequence of a CP-derived tryptic peptide, isolated by reverse-phase HPLC and subjected to automated Edman degradation, is boxed. The complete nucleotide and deduced amino acid sequences of PcV dsRNA2 are deposited in GenBank under the accession no. AF296440; the shortened presentation shown in Fig. 5 points out some features described in Results.

Fig. 4. Phylogenetic analysis of the RDRP conserved motifs and flanking sequences derived from aligned deduced amino acid sequences of PcV and selected totiviruses and partitiviruses. Phylogeny estimates were done using PAUP* 4.0b10, maximum-parsimony setting, and a 100-replicate bootstrap search with the heuristic search option. The resulting consensus tree is shown; numbers above each node indicate the percentage of bootstrap replicates in which that node was recovered. The tree was outgroup-rooted to the totivirus UmV-H1. See Fig. 3 for abbreviations of virus names; the following viruses were not included in Fig. 3: HmV (Helicobasidium mompa virus; Osaki et al., 2002), DdV1 and DdV2 (Discula destructiva virus 1 and 2; Rong et al., 2002), GaRVMS1 (Gremmeniella abietina RNA virus-MS1; Tuomivirta et al., 2002), PPV (Pyrus pyrifolia virus; Osaki et al., 1998) and AbV1 (Agaricus bisporus virus 1; Van der Lende et al., 1996).
The 5' UTRs of PcV dsRNAs are relatively long, between 144 and 162 nt in length. Direct comparison of the nucleotide sequences of the 5' UTR of the four dsRNA segments revealed regions of high sequence similarity (Fig. 7a). In addition to the absolutely conserved 5' termini (positions 1 to 10), a 50 nt region with high sequence identity is present within the 5' UTR of all four dsRNAs. Fifty-two of the 5' terminal 60 nt are identical in at least three of the four dsRNAs. A second region of strong sequence similarity is positioned immediately upstream of the AUG initiator codon (indicated in bold, Fig. 7a). This region consists of a stretch of 50–75 nt containing reiteration of the sequence CAA (underlined, Fig. 6a). The (CAA)n repeats are similar to the translational enhancer elements present at the 5' UTRs of tobamoviruses (Gallie & Walbot, 1992). A highly conserved region (a stretch of 36 nt) with sequence identity above 80% is present within the 3' UTR of all four PcV dsRNAs (Fig. 7b). The 14 nt at the 3' end are strictly conserved among the four genomic dsRNAs (shaded area, Fig. 7b) and 44 out of the 3' terminal 46 nt are identical in at least three of the four dsRNAs.

DISCUSSION

We have determined the complete genomic sequence and genomic organization of PcV, the first chryssovirus to be characterized at the molecular level. Molecular cloning, cDNA sequencing and Northern hybridization analysis of PcV dsRNAs demonstrated quite clearly that each of the PcV dsRNA segments has unique sequences and that none is defective in nature. Each of the dsRNA segments was found to contain a single large ORF, which can be translated in vitro to generate a product of size similar to that above 80%.

Fig. 6. Nucleotide sequence of PcV dsRNA4 cDNA. The deduced amino acid sequence of the protein encoded by PcV dsRNA4 (p4) is indicated. The translation initiation codon, ATG, at nt positions 163–165 is underlined, the termination codon, TAG, at nt positions 2704–2706 is indicated with an asterisk. The amino acid sequence of a capsid-derived tryptic peptide, isolated by reverse-phase HPLC and subjected to automated Edman degradation, is boxed. The peptide sequence matched perfectly the amino acid sequence deduced from dsRNA4 cDNA sequence at nt positions 357–380. The consensus signature of the glycosyl hyrolases family 10 is underlined; the conserved glutamic acid (aa position 709) is boxed. The complete nucleotide and deduced amino acid sequences of PcV dsRNA4 are deposited in GenBank under accession no. AF296442; the shortened presentation shown in Fig. 6 points out some features described in Results.
predicted from its deduced amino acid sequence. Evidence was presented that dsRNA1 encodes a putative RDRP and that dsRNA2 encodes the major CP. Although the functions of the proteins encoded by dsRNAs 3 and 4 are presently unknown, their genomic sequences revealed features consistent with those of genomic dsRNAs of multipartite viruses. Sequence comparisons of the 5’ and 3’ UTRs of all four dsRNAs showed common structural features, indicating the presence of strictly conserved terminal sequences at both ends, long stretches of highly conserved sequences adjacent to both termini and tobamovirus-like sequences at both ends, long stretches of highly conserved sequences at their 5’ ends and 3’ ends that share extended (50–270 nt) regions of highly conserved terminal sequences in all viral segments of extended regions of high sequence similarity at dispensable for virus infection, they are subject to loss during repeated subculturing and single spore isolation. This is well exemplified by the different dsRNA patterns in various strains of Ustilago maydis infected with the totivirus UmV-H and associated defective and satellite dsRNAs (Koltin, 1988). To provide unequivocal evidence that all four dsRNA segments of PcV are necessary for infection, however, would require performing infectivity assays. Unfortunately, such infectivity assays are not presently available for dsRNA mycoviruses (Ghabrial, 1998).

Results from molecular cloning, sequencing and Northern hybridization analysis of PcV dsRNAs (this study) provided conclusive evidence that none of the dsRNA segments is defective in nature. The possibility that dsRNAs 3 and 4 of PcV and related viruses are satellites can be dismissed, considering their consistent and stable co-presence with all four dsRNA segments of extended regions of high sequence similarity at 5’ and 3’ UTRs. These properties are typical of RNA viruses with multipartite and multicomponent genomes, as discussed above, and are unlike those of satellite RNAs. Satellite RNAs, which are generally significantly smaller in size than the genomic RNAs of the helper virus, are known for ssRNA plant viruses with multipartite and multicomponent genomes (e.g. bromoviruses, cucumoviruses and comoviruses; Palukaitis et al., 1992; Shanks & Lomonossoff, 1992; Ahlquist, 1999; Mertens & Sanger, 1985; Palukaitis et al., 1982; Imai et al., 1983; Kuchino et al., 1982; Kudo et al., 1991; Mertens & Sanger, 1985; Palukaitis et al., 1992; Shanks & Lomonossoff, 1992; Ahlquist, 1999; Di et al., 1999). These conserved terminal sequences have been suggested to play an important role in transcription, replication and packaging of viral RNA (Attoui et al., 1997; Mertens et al., 2000; Wei et al., 2003). It is of interest to note that the genome segments of ssRNA plant viruses with multipartite and multicomponent genomes (e.g. bromoviruses, cucumoviruses and comoviruses; Palukaitis et al., 1992; Shanks & Lomonossoff, 1992; Ahlquist, 1999; Di et al., 1999), like chrysoviruses (Ghabrial et al., 2002; A. Soldevila and S. A. Ghabrial, unpublished), are separately encapsidated and share extended (50–270 nt) regions of highly conserved sequences at their 5’ and 3’ UTRs.

The dsRNA pattern of PcV virions isolated from different strains of Penicillium chrysogenum [e.g. strain Wis. Q176 (ATCC 10002) and a number of strains derived from strain NRRL 1951 (ATCC 9480)] has remained unchanged throughout the years since PcV was first isolated (Banks et al., 1969; Buck et al., 1971; Border et al., 1972; Wood & Bozarth, 1972; Nash et al., 1973; Buck & Girvan, 1977; Edmondson et al., 1984; this study). This is also true for other chrysoviruses isolated from different Penicillium species (Buck & Girvan, 1977; Wood et al., 1971) and from various strains of Helminthosporium victoriae (Sanderlin & Ghabrial, 1978; Soldevila et al., 2000; Ghabrial et al., 2002). The co-presence of all four segments in different fungal species and strains harbouring chrysoviruses and the stability of the dsRNA patterns support the contention that all four segments are essential for infection and that none is defective or satellite in nature (see below for additional discussion). As defective and satellite dsRNAs are dispensable for virus infection, they are subject to loss
to share very little or no sequence similarity with their helper viruses (Mayo et al., 2000; Bruening, 2002). PcV dsRNAs 3 and 4 (as well as those of the chrysovirus Hv145SV; Ghabrial et al., 2002; A. Soldevila and S. A. Ghabrial, unpublished) are of comparable size to dsRNAs 1 and 2 and all four dsRNAs share extended regions (50–75 nt) of high sequence similarity (>80%) at both of their termini. Furthermore, all four segments of PcV and Hv145SV contain tobamovirus-like enhancer elements, (CAA)n repeats, upstream of their AUG initiator codons. We are not aware of any satellite RNAs (ssRNA or dsRNA; messenger or non-messenger type) that share extended sequence similarity at the 5′ and 3′ UTR with their helper viruses (Bruening, 2002). Like other satellite RNAs, the known satellite dsRNAs that are associated with totiviruses and partitiviruses share only 2–5 nt at their termini with the helper genomic dsRNAs (Esteban et al., 1989; Wickner, 1996; Rong et al., 2002). Although the terminal sequence at the 3′ end of the well-characterized M1-dsRNA, a satellite dsRNA associated with the totivirus Saccharomyces cerevisiae virus L-A (Esteban et al., 1989; Wickner, 1996), is substantially different from its helper virus at the level of primary nucleotide sequence, it shares similar RNA secondary structures (stem–loop structures) at the 3′ terminus and at an internal site with its helper virus. These structures have been shown to serve as cis-acting signals for replication and packaging (Wickner, 1996). Moreover, the extreme 5′ and 3′ ends of the sense strand of all four dsRNAs from PcV and Hv145SV have the same sequence, 5′-GAUAAAAA ... UAAGUGU-3′ (Ghabrial et al., 2002; this study). The property of having regions of terminal sequences conserved within a genus is characteristic of viruses with multipartite dsRNA genomes, as exemplified by members of the family Reoviridae that include the orthoreoviruses, phytoreoviruses, orbiviruses, rotaviruses and cypoviruses (Mertens et al., 2000). Furthermore, direct evidence based on amino acid sequencing of a tryptic peptide derived from purified PcV capsid showed that p4, encoded by dsRNA4, is virion-associated as a minor component, a characteristic of some gene products of viruses with multipartite genomes (Mertens et al., 2000), but unknown for satellite RNA-encoded proteins (Bruening, 2002). Collectively, these findings are consistent with the conclusion that each of the four dsRNAs of PcV and related viruses represents a component of a multipartite (multicomponent) viral genome and that none is satellite or defective dsRNA.

Amino acid sequence analysis of the protein encoded by PcV dsRNA1 revealed the presence of the eight conserved motifs characteristic of RDRPs of dsRNA viruses of simple eukaryotes (Bruenn, 1993; Ghabrial, 1998). Comparison of the conserved motifs of PcV RDRP with those of viruses in the families Partitiviridae and Totiviridae indicated that it is most closely related to that of another chrysovirus, Hv145SV (Ghabrial et al., 2002), and is more closely related to the totiviruses than to the partitiviruses. This conclusion is also supported by phylogenetic analysis of the conserved motifs and flanking sequences of viruses in the families Totiviridae and Partitiviridae in comparison with the chrysoviruses (Fig. 3). Furthermore, whereas the amino acid positions corresponding to motifs 1 and 2 are clear for PcV and Hv145SV, those for partivirus RDRPs are not well defined (Fig. 3). Interestingly, phylogenetic analysis of the RDRP conserved motifs also revealed that the previously unclassified Agaricus bisporus virus 1 (AbV1; Van der Lende et al., 1996) forms a sister clade to the chrysoviruses PcV and Hv145SV (this study), suggesting its placement as a tentative member of the genus Chrysovirus. Additionally, purified preparations of PcV as well as of other members of the genus Chrysovirus consistently contained four dsRNA segments, as discussed above. In contrast, members of the genus Partivirus, known to support satellite dsRNA, may contain none, one or two satellite dsRNAs besides their two genomic segments (Oh & Hillman, 1995; Rong et al., 2002). Thus, the number of dsRNA segments associated with partivirus infection may vary from two to four segments among members within the same genus. Taken together, these attributes of PcV and similar viruses suggest that they should not be classified with the family Partitiviridae, which includes viruses with bipartite genomes. A proposal, which was based on the properties of the chrysovirus so far characterized at the molecular level (PcV and Hv145SV; Ghabrial et al., 2002; this study), to establish a new family of isometric mycoviruses with multipartite dsRNA genomes was recently accepted by the International Committee on Taxonomy of Viruses (ICTV) and the new family was named Chrysoviridae. The genus Chrysovirus, which was removed from the family Partitiviridae, was designated the type genus of the new family and PcV as the type species of the genus (see http://www.danforthcenter.org/iltab/ictvnet under taxonomic proposals accepted by ICTV).

The three-dimensional structure of PcV was recently determined and it showed that PcV has an authentic T = 1 capsid with 60 equivalent protein subunits. This is in contrast to the ubiquitous T = 2 capsid, with 120 copies of the structural protein, typical of dsRNA viruses including viruses in the family Totiviridae (Castón et al., 2003). This structural study, which utilized cryo-electron microscopy combined with three-dimensional reconstruction techniques, described major conformational differences between full and empty PcV particles and suggested that a number of interactions between the inner surface of the protein shell and the genomic RNA take place. Furthermore, a mechanism for transcript release from transcribing particles was proposed based on the structure of empty particles. The availability of the complete nucleotide sequences of the PcV genome and full-length cDNA clones of viral dsRNAs and the elucidation of the common terminal structural features shared by all PcV dsRNA segments (this study) should prove valuable in future structural studies on protein–RNA interactions, mechanisms of RNA packaging and release and the role of dsRNA in the conformation of the structural subunits.
ACKNOWLEDGEMENTS

This work was supported in part by Grant Agreement 2001-35319-10010 from the United States Department of Agriculture National Research Initiative Competitive Research Program (to S.A.G.) and is published with the approval of the Director of the Kentucky Agricultural Experiment Station as Journal Article 03-12-111. We also thank Ana Soldevila for valuable advice and assistance during the course of this study and Wendy Havens for technical assistance.

REFERENCES


Molecular characterization of PcV


