Evidence for the occurrence of two distinct subgroups of peanut stunt cucumovirus strains: molecular characterization of RNA3

C.-C. Hu, A. E. Aboul-Ata,† R. A. Naidu‡ and S. A. Ghabrial

Department of Plant Pathology, University of Kentucky, Lexington, Kentucky 40546, USA

Strains of peanut stunt cucumovirus (PSV) were classified into two distinct subgroups, I and II, based on Western and Northern blot analyses using antisera and cloned cDNA probes to strains PSV-ER and PSV-W. These results were corroborated by nucleotide sequence analyses of full-length cDNA clones of RNA3 from representative strains of the two subgroups. Whereas the percentage nucleotide sequence identity between PSV-ER (or PSV-J) and PSV-W RNA3s was determined to be 80%, the corresponding value between strains ER and J was 91%, confirming that strains ER and J belong to the same subgroup (subgroup I) whereas strain W belongs to a separate subgroup (subgroup II). PSV-W and PSV-ER RNA3s are 2173 and 2188 nucleotides long, respectively. Each is dicistronic, encoding a putative movement protein (3a protein) and a coat protein (CP). The intercistronic and 5' untranslated region (UTR) sequences of PSV strains, unlike those of cucumber mosaic cucumovirus (CMV) strains, are highly conserved and thus not useful for distinguishing the two subgroups. However, the 3' UTR sequences of PSV strains, like those of CMV strains, can discriminate between the two subgroups since strains within the same subgroup are 95% identical in their 3' UTRs whereas those in different subgroups are only 74–78% identical. PSV-W and PSV-ER RNA4s were determined to be 994 and 1006 nucleotides long, respectively. PSV 3a and CP genes have higher percentage nucleotide sequence identities to those of tomato aspermy cucumovirus than to those of CMV.

Introduction

Peanut stunt virus (PSV) is a member of the genus Cucumovirus in the family Bromoviridae (Rybicki, 1995). Other members of the genus are cucumber mosaic virus (CMV), the type member, and tomato aspermy virus (TAV). Cucumoviruses have tripartite genomes of positive strand RNAs, designated RNA1, 2 and 3 in order of decreasing size, which are packaged in isometric particles of about 28 nm diameter. In addition to the genomic RNAs, the virions also encapsidate a fourth RNA (RNA4), a subgenomic RNA 3'-co-terminal with RNA 3, which functions as mRNA for the viral coat protein (for a review, see Palukaitis et al., 1992). A small overlapping gene (2b), encoded by RNA2, was recently discovered in all cucumoviruses sequenced to date and is most likely expressed through the subgenomic mRNA, RNA4A (Ding et al., 1994). Naturally occurring virions of CMV and PSV, but not TAV, may also package a fifth RNA, designated satellite RNA (satRNA), along with their genomic and subgenomic RNAs (Naidu et al., 1991a, 1992; Roossinck et al., 1992). CMV and TAV, but not PSV, support the replication of CMV satRNAs. Only PSV supports the replication of PSV satRNAs (Kaper et al., 1978; Naidu et al., 1995).

PSV is an economically important pathogen of legumes world-wide (Mink, 1972; Xu et al., 1986). On the basis of gel double diffusion immunassays and nucleic acid competition hybridization, isolates of PSV have been differentiated into two major strains, eastern and western (Mink et al., 1969; Mink, 1972; Diaz-Ruiz & Kaper, 1983; Xu et al., 1986). Several
variants of the eastern strain have been described based on host reactions and/or gel immunodiffusion patterns (Xu et al., 1986). Recently, Naidu et al. (1995) differentiated eastern and western strains based on satRNA support and sequence homology.

The complete nucleotide sequences of the genomic RNAs of several strains of CMV have been reported and CMV strains have been classified into two subgroups, I and II, based on sequence homology (Palukaitis et al., 1992). Partial sequences of some TAV strains are also available (Moriones et al., 1991; O’Reilly et al., 1991). Only the complete sequence of the Japanese strain of PSV (Karasawa et al., 1991, 1992) has been published. In this paper we report the complete nucleotide sequence of RNA3 from two strains of PSV and present serological and molecular evidence that strains of PSV, like those of CMV, fall into two distinct subgroups.

**Methods**

- **Virus strains and purification.** The sources of PSV strains ER, W, BV-15 and CMV strain Fny were as described previously (Naidu et al., 1995). A subculture of the Japanese strain (PSV-J) was provided by K. Hanada (Kyushu Agricultural Experiment Station, Japan) and the Spanish strain PSV-B (Diaz-Ruiz et al., 1979) was supplied, as purified viral RNAs, by P. Garcia-Arenal (Universitat de Valencia, Spain). All PSV strains were maintained in cowpea (Vigna unguiculata) or in Burley tobacco (Nicotiana tabacum, cv. Ky 14) or in cowpea. Purification of PSV virions as described by Ghabrial et al. (1997). CMV was increased in Burley tobacco and purified according to the procedure of Lot et al. (1972).

- **Antisera preparation.** Antisera to glutaraldehyde-stabilized virions from two strains, PSV-ER and PSV-W, were produced in rabbits by a series of subcutaneous injections administered at biweekly intervals.

- **Western blotting.** Following SDS–PAGE (Hu & Ghabrial, 1995), the proteins were transferred to Immobilon-P transfer membrane (Millipore) with a Mini Trans-Blot electrophoretic transfer apparatus (Bio-Rad) at 90 V for 1.5 h. The membranes were then incubated overnight at 4°C in PBS (10 mM sodium phosphate buffer containing 150 mM NaCl, pH 7.4) containing 0.5% non-fat milk, washed three times with 1× TBS (20 mM Tris–HCl buffer containing 150 mM NaCl, pH 7.4) and incubated at room temperature for 2 h with an antiserum to PSV virions diluted 1:6000. The membranes were washed as before, then reacted with the secondary antibody, goat anti-rabbit IgG conjugated with alkaline phosphatase (Sigma; diluted 1:1500 in 1× TBS containing 4% non-fat milk) for 1 h at room temperature. The bound antibody was detected using 5-bromo-4-chloro-3-indolyl phosphate – nitro blue tetrazolium (Promega) as substrates.

- **In vitro translation.** Purified total viral RNA or full-length RNA3 transcripts of strains PSV-W and PSV-ER were translated in rabbit reticulocyte lysate (Promega). After 1 h of translation, the products were analysed in 12.5% SDS–polyacrylamide gels (Laemmli, 1970). The gels were fixed, dried and exposed to a Phosphor screen (Molecular Dynamics) for 5–16 h. The images were visualized by PhosphorImager 445 SI system (Molecular Dynamics) and analysed with the ImageQuant 4.1 program. The images were then transferred to a Powerpoint program (Microsoft) for adding labels.

- **Northern hybridization analysis.** Viral RNAs were extracted from purified virions as described by Naidu et al. (1995), denatured with 6% formaldehyde and separated by electrophoresis in 1.5% agarose gels made in TBE (89 mM Tris–borate buffer, pH 8.3, containing 25 mM disodium EDTA). The RNAs were then transferred and fixed onto Hybond-N membrane (Amersham) according to the manufacturer’s instructions. The membranes were pre-hybridized for 2 h at 42°C with buffer containing 5 × SSC (1× SSC: 150 mM NaCl, 15 mM sodium citrate), 50% formamide, 0.5% SDS, 100 μg/ml salmon sperm DNA, then hybridized with 32P-labelled probes prepared by nick-translation of full-length cDNA clones of RNA3 of PSV-ER (pER3), or PSV-W (pW3–1) for 16–18 h at 42°C, followed by washing four times with 2× SSC, 0.1% SDS at 25°C for 15 min and 2 times with 0.1% SSC, 0.1% SDS at 68°C for 15 min. The membranes were then air-dried and exposed for 5–12 h to Bio-Max film (Kodak) with intensifying screens at −80°C.

- **cDNA synthesis and cloning.** PSV RNA3 was purified from low-melting-point agarose following electrophoresis of unfractionated virion RNAs. cDNA libraries representing RNA3 from either PSW-V or PSV-ER were synthesized using components of the Amersham cDNA synthesis kit. Poly(A) tails were added enzymically to PSV RNA3 by treatment with poly(A) polymerase (Sippel, 1973) and cDNA synthesis was primed with oligo(dT)12–18 primers. Alternatively, a 21-mer synthetic oligonucleotide complementary to the 3′-terminal sequence of PSV-ER RNA4 (Naidu et al., 1991b) was used to prime the synthesis of first strand cDNA; the second strand was synthesized according to the procedure of Gubler & Hoffman (1983). Double-stranded (ds) cDNA was blunt-ended with T4 DNA polymerase and ligated into Smal-linearized pUC119 or pUC118. Selection of ampicillin-resistant cDNA clones containing inserts was made by blue/white colony screening on X-Gal/IPTG medium (Sambrook et al., 1989).

- **Sequence analysis.** Multiple independent cDNA clones of appropriate size were used for sequencing. For DNA sequencing analysis, ds cDNA as well as single-stranded (ss) DNA of both strands of selected cDNA clones (subcloned in the phagemids pUC118/119 and generated with the aid of helper phage) (Vieira & Messing, 1987) were used. Furthermore, nested sets of unidirectional deletion mutants of selected cDNA clones (subcloned in M13EcoK vector) were generated according to the procedure of Shen & Waye (1988), and using the M13 single-step nested deletion kit (Bio101). Nucleotide sequences of selected cDNA clones were obtained with the dideoxy chain-termination method (Sanger et al., 1977) using T7 DNA polymerase (Sequenase version 2, USB) and ss and ds cDNA templates. As sequences were generated using universal sequencing primers, a directed sequencing protocol using progressive synthetic oligonucleotide primers was followed. At least two independent clones were sequenced from both orientations for each region of the cDNAs. The 5′-terminal sequences were determined by the dideoxy chain-termination method using a reverse transcriptase RNA sequencing kit method (USB). Internal primers complementary to nucleotide positions 118–131 and 121–134 of RNA3 of PSV-W and PSV-ER, respectively, were used to prime reverse transcription of the template RNAs. For the 5′-terminal sequence of PSV-W RNA4, an internal primer corresponding to positions 1318–1332 of RNA3 was used. The sequences were assembled using the GEL program in the IG-Suite (IntelliGenetics) and analysed using both IG-Suite and the University of Wisconsin GCG package (Devereux et al., 1984). The GAP, PILEUP and PRETTY programs in the UWCGC package (version 8.0) were used to align and compare nucleotide and deduced amino acid sequence identities (and amino acid sequence similarities) between RNA3s of the following cucumoviruses (GenBank accession number in parentheses): PSV-W
Table 1. Lengths of the coding and noncoding regions in the RNA3s of the cucumoviruses included in this study.

<table>
<thead>
<tr>
<th>Virus</th>
<th>5' UTR</th>
<th>3a-encoding</th>
<th>IR</th>
<th>3' UTR</th>
<th>Full length</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSV-W</td>
<td>116</td>
<td>867</td>
<td>262</td>
<td>645</td>
<td>2173</td>
<td>This study</td>
</tr>
<tr>
<td>PSV-ER</td>
<td>119</td>
<td>870</td>
<td>259</td>
<td>678</td>
<td>2188</td>
<td>This study</td>
</tr>
<tr>
<td>PSV-J</td>
<td>118</td>
<td>867</td>
<td>257</td>
<td>654</td>
<td>2186</td>
<td>Karasawa et al. (1991)</td>
</tr>
<tr>
<td>CMV-Fny</td>
<td>119</td>
<td>840</td>
<td>297</td>
<td>657</td>
<td>2216</td>
<td>Owen et al. (1990)</td>
</tr>
<tr>
<td>CMV-Y</td>
<td>120</td>
<td>840</td>
<td>299</td>
<td>657</td>
<td>2217</td>
<td>Nitta et al. (1988)</td>
</tr>
<tr>
<td>CMV-Trk7</td>
<td>96</td>
<td>840</td>
<td>295</td>
<td>657</td>
<td>2209</td>
<td>Salanki et al. (1994b)</td>
</tr>
<tr>
<td>CMV-Q</td>
<td>95</td>
<td>840</td>
<td>284</td>
<td>657</td>
<td>2197</td>
<td>Davies &amp; Symons (1988)</td>
</tr>
<tr>
<td>TAV-C</td>
<td>89</td>
<td>834</td>
<td>295</td>
<td>690</td>
<td>2214</td>
<td>O’Reilly et al. (1991)</td>
</tr>
<tr>
<td>TAV-P</td>
<td>89</td>
<td>843</td>
<td>295</td>
<td>657</td>
<td>2222</td>
<td>Salanki et al. (1994a)</td>
</tr>
</tbody>
</table>

Results and Discussion

Differentiation of PSV strains into two subgroups based on Western and Northern blot analyses

In Western blots, the antiserum to PSV-ER virions reacted strongly with its homologous antigen (Fig. 1a, lane 5) as well as with strain BV-15 (lane 7). Weak or very weak reactions were detected with PSV-W and CMV-Fny (Fig. 1a, lanes 6 and 8, respectively). The antiserum to PSV-W virions, however, reacted strongly with its homologous antigen (Fig. 1b, lane 6) and weakly with PSV strains ER and BV-15 (lanes 5 and 7). No reaction was detected with CMV-Fny (Fig. 1b, lane 8). Results of Western blot analysis of PSV-J and PSV-B were similar to those obtained with PSV-ER and PSV-W, respectively (data not shown). Coomassie-stained gels of the cucumovirus coat proteins analysed by Western blotting are shown in Figs 1(a) and 1(b), lanes 1–4. The results with the reassortant strain BV-15 verify that RNA3 (or more specifically the CP gene) of this strain is derived from an eastern strain (Naidu et al., 1995) since it is serologically indistinguishable from other eastern strains.

Northern blots of total RNAs from purified PSV virions strains ER, W, J and B were probed with either PSV-W or PSV-ER RNA3-specific cloned probes. Whereas the PSV-W RNA3-specific probe hybridized strongly with RNA3 from strains W and B, it did not react with RNA3 from either PSV-ER or PSV-J (Fig. 2a). However, the PSV-ER RNA3-specific probe hybridized strongly with its homologous RNA (Fig. 2b) and also with RNA3 from PSV-J (data not shown). No hybridization signals, however, were detected with RNA3 from either PSV-W or PSV-B (Fig. 2b). No cross-hybridization was observed between the RNA3 probes and the homologous RNAs 1 and 2 (Fig. 2) even though the sequences of the 3’ untranslated regions (UTRs) of the three genomic RNAs are 83–88% identical (C.-C. Hu and others, unpublished). Very weak signals, however, were detected with RNAs 1 and 2 in some autoradiograms exposed for 72 h or more (data not shown). The lack or weakness of cross-hybridization may be explained on the basis that the full-length RNAs share only 45% sequence similarity (C.-C. Hu and others, unpublished) and that efficient detection of hybridization, under the high stringency conditions we used, requires that sequences have at least 90% identity to the probe (Memelink et al., 1994).

The results of Western and Northern blot analyses provide further evidence that PSV strains can be differentiated into two distinct subgroups (subgroups I and II instead of eastern and western). Thus, it was confirmed that strain J and strain B belong to subgroup I (Naidu et al., 1995) and subgroup II (Diaz-Ruis & Kaper, 1983), respectively.

Comparative nucleotide sequence analysis of RNA3 from PSV strains W and ER

The nucleotide sequences of full-length cDNA clones of RNA3 from PSV strains W and ER were determined (Figs 3 and 4). PSV-W and PSV-ER RNA3s are 2173 and 2188 nucleotides
The 5’ UTR of PSV RNA3, like those of other cucumoviruses, contains the conserved UG tract (Figs 3 and 4, boxed) known to be required for efficient accumulation of CMV RNA3 (Boccard & Baulcombe, 1993). The UG tract of PSV-W occurs between position 85 and 100, those for strains ER and J occur between positions 88 and 103, and 87 and 102, respectively (Figs 3 and 4; Karasawa et al., 1991). Interestingly, TAV has a significantly longer UG tract (between position 38 and 80; O’Reilly et al., 1991) than either PSV or CMV. The IR of both PSV-W and PSV-ER contains the conserved internal control region (ICR2)-like conserved motif (Figs 3 and 4, boxed) shown to be necessary for efficient synthesis of RNA3 (Boccard & Baulcombe, 1993). Like other cucumoviruses, the 3’ UTRs of PSV-W and PSV-ER RNA3s contain the 40 nt sequence domain conserved in all cucumovirus RNAs so far examined (Figs 3 and 4, underlined). This highly conserved sequence, which is important for RNA3 accumulation (Boccard & Baulcombe, 1993; McGarvey et al., 1995), may be exploited in designing universal probes for the detection and diagnosis of cucumovirus infections and also in designing universal primers for cDNA synthesis cloning of cucumovirus RNAs. It is of interest to note that the only known variant nucleotide (arrowhead above sequence in the 3’ UTR, Figs 3 and 4) in this highly conserved 40 nt sequence (McGarvey et al., 1995) is different for each of the three PSV strains: G for strain W, U for strain ER and A for strain J.

The percentage nucleotide sequence identity between PSV-W and PSV-ER RNA3s, or between those of PSV-W and PSV-J, are 80-4% and 80-5%, respectively. The corresponding value between strains ER and J, however, is 91% (Fig. 5a, above diagonal) confirming the earlier proposal that these two strains belong to the same subgroup whereas strain W belongs to a separate subgroup (Naidu et al., 1995). The percentage nucleotide sequence identity between CMV strains within the same subgroup is 97-98% and between strains in different subgroups is 72-74% (Fig. 5a). CMV strains have been classified into two subgroups based on their sequence homology (Palukaitis et al., 1992); for example, strains Fny and Y belong to subgroup I whereas CMV strains Trk7 and Q belong to subgroup II. The percentage nucleotide sequence identity between PSV and CMV RNA3s is 55-60%, depending on the strain, and between PSV and TAV is about 65%, suggesting that overall PSV RNA3 is more closely related to TAV RNA3 than to that of CMV (Fig. 5a).

The conclusion that PSV strains can be classified into two subgroups based on sequence analysis of RNA3 is in agreement with that deduced from sequence comparisons of RNA1 and RNA2 from three PSV strains. The values for percentage nucleotide sequence identity between strains in the same subgroup (between strains J and ER) are 91 and 94%, respectively, for RNA1 and RNA2. The corresponding values for strains in different subgroups (between strains W and ER/J) are 79 and 76%, respectively, for RNA1 and RNA2 (C.-C. Hu and others, unpublished; Karasawa et al., 1992).
Comparative sequence analyses of IRs and 5′ and 3′ UTRs of cucumoviruses

The percentage nucleotide sequence identity between the IRs of PSV strains were similar to each other (84–88%; Fig. 5a, below diagonal) regardless of whether the strains are in the same or in different subgroups. This was also true for percentage nucleotide sequence identity of the 5′ UTR (94–97%; Fig. 5b, above diagonal). The IR and 5′ UTR sequences of CMV strains, unlike those of PSV strains, are not as highly conserved. Whereas the values for percentage sequence identity of IR between strains in the same subgroup varied from 93.3 to 99.6%, those between strains in different subgroups were found to be 66–70% (Fig. 5a, below diagonal). For the 5′ UTR, the corresponding values are 90–99% and 42–48%, respectively (Fig. 5b). Based on the RNA3 IR sequence, higher values for percentage nucleotide sequence identity were determined between PSV and TAV than between PSV and CMV (Fig. 5a). Interestingly, the 5′ UTR sequence of TAV strains has higher nucleotide sequence identity to CMV than to PSV.

Contrary to PSV IR and 5′ UTR, the percentage sequence identity of RNA3 3′ UTR between strains depended on whether they are in the same subgroup (95%) or in different subgroups (74–78%; Fig. 5b, below diagonal). This is also true for the 3′ UTR of RNA3 of CMV strains (Fig. 5b, below diagonal). This subgroup-distinguishing property of the 3′ UTR sequence of cucumovirus RNA3s is shared by RNA1 and RNA2 (Palukaitis et al., 1992; C.-C. Hu and others, unpublished), and therefore may be useful in diagnosis and classification of PSV and CMV strains. Comparable values in percentage sequence identity of the 3′ UTRs were obtained between PSV and either CMV or TAV strains (Fig. 5b).

Characterization of the PSV 3a protein and comparison with other cucumoviruses

For PSV-W, ORF1 starts at the codon AUG at positions 117–119 and terminates at the codon UAA at positions 981–983, thus encoding a protein of 288 aa with a predicted molecular mass of 31,504 Da (Fig. 3). ORF1 of PSV-ER RNA3 starts at the AUG at position 120–122 and terminates at the UAA at position 987–989, thus encoding a protein of 289 aa with calculated molecular mass of 31,379 Da (Fig. 4). In both strains, the initiation codon of ORF1 (which is the first AUG in the sequence) has an optimal context, according to Kozak (1989): an A at position −3 and a G at position +4. Unlike strains W and ER, PSV-J RNA3 has two AUG codons in frame proximal to the 5′ terminus (Karasawa et al., 1991), at positions 113–115 and 119–121 (underlined and double-underlined, respectively, in alignment below). Karasawa et al. (1991) were not certain as to which one of the AUGs is the initiator codon. With the availability of the complete sequences of RNA3 from two other strains (this study) and examining alignments of the three strains, it is most likely that ORF1 of strain J initiates translation from the AUG at position 119–121. An alignment of the 10 nt flanking the initiator AUG in the three strains is shown below:

PSV-ER 111-CAUUUAGAGAUGG-123
PSV-J 110-CGUAGAGAGUUG-122
PSV-W 108-CUUUUAGAGUGG-120
Fig. 3. Nucleotide sequence of PSV-W RNA3 cDNA. The conserved UG tract in the 5′ UTR is boxed (TG tract in the cDNA sequence). The predicted amino acid sequences of the 3a protein (the 3a gene is predicted from nt 117 to 980) and CP (the CP gene is predicted from nt 1246 to 1887) are indicated in the one-letter code below the nucleotide sequence. The principal conserved motif typical of the 30K movement protein superfamily is double-underlined. The conserved ICR2-like motif in the IR is boxed and the first nucleotide of RNA4, determined to correspond to the G at position 1180 (see text) is indicated by an arrowhead below the sequence. The amino acid sequences of two CP-derived tryptic peptides, isolated by reverse-phase HPLC and subjected to automated Edman degradation, are double-underlined. The 40 nt domain conserved in the 3′ UTR of all cucumovirus RNA sequences is underlined and the position of the only known variant nucleotide in this highly conserved sequence is indicated with an arrowhead.
Fig. 4. Nucleotide sequence of PSV-ER RNA3 cDNA. The conserved UG tract in the 5′ UTR is boxed (TG tract in the cDNA sequence). The predicted amino acid sequences of the 3a protein (the 3a gene is predicted from nt 120 to 986) and CP (the CP gene is predicted from nt 1249 to 1923) are indicated in the one-letter code below the nucleotide sequence. The principal conserved motif typical of the 30K movement protein superfamily is double-underlined. The conserved ICR2-like motif in the IR is boxed and the first nucleotide of RNA4, predicted to correspond to the G at position 1183 (see text) is indicated by an arrowhead below the sequence. The 40 nt domain conserved in the 3′ UTR of all cucumovirus RNA sequences is underlined and the position of the only known variant nucleotide in this highly conserved sequence is indicated with an arrowhead.
Fig. 5. Percentage nucleotide sequence identity of RNA3 between PSV strains and other cucumoviruses. (a) Full-length RNA3 above diagonal, IR below diagonal. (b) 5' UTR above diagonal, 3' UTR below diagonal. (c) 3a gene above diagonal, CP gene below diagonal. Values are the identity scores as calculated by the GAP program in the UWGCG package.

In vitro translation in a rabbit reticulocyte lysate system of RNA3 from either PSV strain or of transcripts from full-length RNA3 cDNA clones yielded a protein (3a protein) with an estimated molecular mass of 34 kDa as the major product (Fig. 6a). This value is slightly higher than the 31–32 kDa value predicted from the deduced amino acid sequence of ORF1. The amino acid sequence of ORF1 from the two strains of PSV and that from PSV-J contains the principal conserved motif (Figs 3 and 4, underlined) typical of the 30K movement protein superfamily (Mushegian & Koonin, 1993). An align-
generated by the PILEUP program and the consensus sequence was derived from the alignment (Mushegian & Koonin, 1993), present in the movement proteins from nine cucumoviruses. The conserved motif was typical of the 30K movement protein superfamily.

Values for percentage amino acid sequence similarity between PSV 3a proteins were found to be 95% and 87–88%, respectively, between strains in the same subgroup and in different subgroups (Fig. 7a; above diagonal). The corresponding values for percentage amino acid sequence identity were 92% and 77–80%, respectively (Fig. 7a, below diagonal). The data assembled on the percentage nucleotide (or amino acid) sequence identity between 3a genes (or 3a proteins) of cucumoviruses were 92% and 87–88%, respectively (Fig. 5b). An alignment of the principal conserved sequence motif typical of the 30K movement protein superfamily (Mushegian & Koonin, 1993), present in the movement proteins from nine cucumoviruses. The conserved motif was derived from the alignment generated by the PILEUP program and the consensus sequence was produced using the PRETTY program.

Figure 6. (a) Autoradiogram of polypeptides synthesized in rabbit reticulocyte lysates. Lanes: 1, total virion RNA, PSV-ER; 2, RNA3, PSV-ER; 3, total virion RNA, PSV-W; 4, transcript from cloned full-length cDNA of RNA3, PSV-ER; 5, no RNA added; 6, transcript from cloned full-length cDNA of RNA3, PSV-W. (b) An alignment of the principal conserved sequence motif typical of the 30K movement protein superfamily (Mushegian & Koonin, 1993), present in the movement proteins from nine cucumoviruses.

Characterization of the PSV CP gene and comparison with that of other cucumoviruses

The 5’ terminus of RNA4 of PSV-W was determined by direct RNA sequencing to correspond to the G at position 1180 of RNA3, and that of PSV-ER RNA4 was predicted to correspond to the G at position 1183 of RNA3 (Figs 3 and 4, bold letter and arrowhead). Therefore, RNA4 of PSV-W and PSV-ER are 994 and 1006 nt long, respectively. The CP amino acid sequence, deduced from cDNA sequencing, was verified by direct amino acid sequencing of two CP internal tryptic peptides derived from purified PSV-W virions; the sequence of one peptide (GHSFGQR) matched perfectly the amino acid sequence deduced from the cDNA sequence at nt 1485–1506 (Fig. 3, double-underlined) and the sequence of the second peptide (IPSASK) matched perfectly that deduced from the sequence at nt 1861–1878 (Fig. 3, double-underlined).

Although the calculated values for CP molecular mass of all cucumoviruses so far sequenced are approximately 24 kDa, the experimentally determined values using the Laemmli SDS–PAGE system (Laemmli, 1970) were found to be 30–31 kDa (Hu & Ghabrial, 1995). We have recently reported that cucumovirus CPs show anomalously electrophoretic mobilities when subjected to SDS–PAGE (Hu & Ghabrial, 1995). This anomalous behaviour of CPs from purified virions was also observed with bacterially expressed cloned cucumovirus CP genes as well as with in vitro translation products of RNA4s and of transcripts from full-length cloned CP genes. Evidence was presented that the highly conserved, hydrophilic and arginine-rich N-terminal domain of these CPs is a major contributing factor to their anomalous behaviour in SDS–polyacrylamide gels (Hu & Ghabrial, 1995).

The percentage nucleotide sequence identity in the CP gene between PSV strains within the same subgroup and in different subgroups were calculated to be 87% and 75–78%, respectively (Fig. 5c, below diagonal). The value of 87% nucleotide identity (and 84% amino acid sequence identity; Fig. 7b) for strains in the same subgroup appears to be relatively low in comparison with the corresponding values for CMV strains (98% and 97%, respectively). However, the values for percentage amino acid identity for PSV strains in different subgroups (65–70%) were proportionately lower than the corresponding values for CMV strains (81–83%, Fig. 7b). The finding that the percentage amino acid sequence similarity between PSV strains in the same subgroup is 90%
Fig. 7. Percentage deduced amino acid sequence identity (similarity) of 3a protein (a) and CP (b) between PSV strains and other cucumoviruses. (a) Amino acid sequence similarity of 3a protein above diagonal, amino acid sequence identity below diagonal. (b) Amino acid sequence similarity of CP above diagonal, amino acid identity below diagonal. Values are the identity (similarity) scores as calculated by the GAP program in the UWGCG package.

We would like to thank Margaret Sanger and M. R. Hajimorad for their comments. This work was supported by a grant from the University of Kentucky Tobacco and Health Research Institute (5-41124) and is published as Journal Series Paper 96-12-164 of the Kentucky Agricultural Experiment Station.


Received 9 September 1996; Accepted 14 November 1996