Several genes in Chlorella virus strain CVG-1 encode putative virion components

Barbara Plugge,1 Barbara Becker2 and Andreas H. Wolf1

1 Albrecht-von Haller-Institut für Pflanzenwissenschaften, Universität Göttingen, Untere Karspule 2, 37073 Göttingen, Germany
2 Abt. Landwirtschaftliche und Lebensmittel-Mikrobiologie, Universität Bonn, Meckенheimer Allee 168, 53115 Bonn, Germany

We have started to characterize the capsid components of European Chlorella virus isolate CVG-1, a member of the Pbi subgroup of the Phycodnaviridae. The major coat protein, Vp49, was biochemically characterized and the amino acid sequence of the N terminus was determined. Subsequently, the corresponding gene was isolated from CVG-1 genomic DNA. Sequence data were compared to those available from PBCV-1 and other Chlorella virus isolates representing the NC64A subgroup of the Phycodnaviridae. The major coat proteins of all strains are homologous and similar in size, but apparently differ in their degree of glycosylation. Like PBCV-1, the major coat protein of CVG-1 is part of a gene family, as two open reading frames with high similarity to Vp49 were also isolated and characterized in this study. The predicted amino acid sequences of the CVG-1 and PBCV-1 virus genes examined show, with one exception, a divergence of about 25%. Taking into account that corresponding genes of NC64A viruses are almost identical, this divergence supports the original placement of the NC64A and Pbi viruses into separate subgroups of the Phycodnaviridae.

Introduction

Chlorella virus Göttingen (CVG-1) belongs to a family of large dsDNA viruses, the Phycodnaviridae (reviewed by Van Etten et al., 1991). Chlorella viruses are ubiquitous in freshwater. Their hosts, symbiotic ‘zoochlorella’ species, are associated with Paramecium bursaria or Hydra viridis. Only algae separated from their symbiosis partner are infected by the viruses.

Chlorella virus isolates show strong specificity towards certain host strains and have been grouped accordingly: Pbi viruses infect Chlorella strain Pbi isolated from Paramecium bursaria collected in Göttingen, Germany; NC64A viruses infect Chlorella strain NC64A isolated from Paramecium bursaria collected in the United States; HVCV viruses infect Chlorella strains isolated from Hydra viridis.

Chlorella strains Pbi and NC64A differ in chloroplast morphology and cell wall composition. Neither can be assigned to any of the free-living Chlorella species, but both are believed to belong to the Chlorella vulgaris/sorokiniana group (Reisser & Widowski, 1992).

Of the NC64A viruses, PBCV-1 has been studied extensively, while little is known about the European Pbi isolates. To date, the only available genomic sequence of a Pbi virus is a portion of a DNA polymerase gene. Phylogenetic analysis of DNA polymerase sequences of several algal viruses places the Pbi and NC64A viruses in separate, though closely related groups (Chen & Suttle, 1996). The viruses cannot infect host strains of the other group, because they do not attach to the algal cells (Reisser et al., 1988). This may be related to cell wall characteristics of the host strains, but must also depend on variation in virion proteins. Of the virion’s more than 50 structural proteins, several are glycosylated and located in the outer layer of the capsid (Skrdla et al., 1984; Songsri et al., 1997). The surface structure of the virion is dominated by the major capsid protein, a glycoprotein accounting for approximately 40% of the total virion protein (Skrdla et al., 1984). A role of the major capsid protein in the attachment/host-recognition process is disputable though, since several PBCV-1 antigenic variants with incomplete oligosaccharide chains are still able to infect the original host alga NC64A (Wang et al., 1989).

Author for correspondence: Andreas Wolf.
Fax +49 551 397838. e-mail awolf@gwdg.de
The GenBank accession numbers of the sequences reported in this paper are AF076921–AF076923.
1993). The key element for host recognition might rather be one of the minor virion proteins, probably located at the virion's vertices where attachment occurs.

To investigate the relationship between NC64A and Pbi viruses, we have started to characterize structural components of Pbi virus CVG-1 at a molecular level. We were interested in isolating the major capsid protein gene as well as related virion components, and in comparing our data to corresponding genes isolated from NC64A viruses. The amino acid sequences inferred from the major coat protein gene and several related open reading frames (ORFs) show that individual genes are closely related between the two virus groups.

**Methods**

**Growth conditions.** Growth of *Chlorella* strain Pbi (*Paramecium bursaria* isolate) and production of virus have been described elsewhere (Reisser et al., 1986). CVG-1, isolated from a pond in Göttingen (Reisser et al., 1988), was treated accordingly.

**Virus purification, DNA preparation and PCR.** To obtain a suspension of genetically identical viruses, a single plaque isolate was suspended in 50 mM Tris–HCl and further purified by successive plaque assays. The final single plaque isolate was repeatedly centrifuged on sucrose density-gradients (Van Etten et al., 1983). Virus DNA was extracted by protease K digestion and phenol–chloroform extraction, followed by precipitation of DNA.

PCR was carried out on a Perkin Elmer DNA thermal cycler 480. One degenerate oligonucleotide (GTNGCNYGGNCNCARGAYGTNTA) corresponded to aa 9–18 of the CVG-1 major protein, Vp49. The second oligonucleotide (AGRTAYTTNGTNGGRTGARTARAAR) corresponded to aa 254–262 of the PBCV-1 major capsid protein, Vp54 (see Fig. 2). The ends of the PCR products were made blunt and ligated to pUC18 vector (Sure Clone Ligation kit, Pharmacia). The constructs were transformed into *E. coli* strain DH5αF′ and prepared for sequencing.

**Protein extraction and sequencing.** Purified virus particles were diluted in 1 vol. of sample buffer and heated at either 60 or 100 °C for 5 min. Proteins were fractionated on 12.5% polyacrylamide gels in the presence of SDS. Protein bands were visualized by silver staining or incubation with 100 U of the respective enzymes at 37 °C for at least 12 h before SDS–PAGE. Proteins were visualized by fluorography.

**Detection of glycoproteins.** Glycoproteins were stained with Schiff’s reagent (Kapitany & Zebrowski, 1973). For *in vivo* labelling, 25 ml of virus-infected Pbi suspension cultures and an uninfected control were incubated with 750 μCi [3H]glucosamine, 250 μCi [3H]mannose and 250 μCi [3H]galactose. After 48 h incubation, algal proteins and virus particles were isolated and fractionated by SDS–PAGE. Labelled proteins were visualized by fluorography.

Infected cultures (40 ml) were incubated with 200 μCi [35S]methionine and 200 μCi [35S]cysteine before treatment with endoglycosidase F (endo-β-N-acetylglucosaminidase), N-glycosidase F (peptide-N-β-glucosaminyl)-asparagine amidase), endoglycosidase H (endo-β-N-acetylglucosaminidase H) and O-glycosidase (O-glycan peptide hydrolase; all enzymes obtained from Boehringer Mannheim). Viruses were harvested after 28 h. The labelled virus preparations were mixed with 2 vols of denaturing buffer (0.2% SDS, 1% N-octylglucoside, 0.1% mercaptoethanol) and sodium acetate buffer (0.1 M sodium acetate, 0.01 M EDTA) respectively and boiled for 10 min. The assays were incubated with 100 U of the respective enzymes at 37 °C for at least 12 h before SDS–PAGE. Proteins were visualized by fluorography.

**Construction and screening of a genomic library.** A genomic library was prepared by partially digesting CVG-1 DNA with *SalIII*, enriching for 20 kbp DNA fragments on sucrose density-gradients, and ligating the purified DNA fragments into the BamHI site of EMBL3 vector (Stratagene). The constructs were packaged in *vitro* (GigapackGold, Stratagene) and used to infect *E. coli* strain XL-1 Blue MRA.

Phage plaque.lifts were screened with [α-32P]dCTP-labelled probes. All hybridizations were performed overnight in sodium phosphate buffer (Church & Gilbert, 1984). Post-hybridization washes at room temperature were carried out twice in 0.1 SSC (1 X SSC = 0.15 M NaCl, 0.015 M sodium citrate), twice in 0.2 SSC and twice in 0.2 X SSC for 15 min. Final stringency conditions were 0.1 X SSC, 1% SDS at 65 °C for 10 min. Plaque.lifts were autoradiographed at −70 °C for 24 h with intensifying screens on Kodak X-OMAT AR film. Phages and phage DNA were prepared according to Sambrook et al. (1989).

**Southern blotting.** For DNA blot analysis, 10 μg of phage DNA was digested for 6 h with the appropriate restriction enzymes. Fragments were separated by electrophoresis on 0.5% agarose gels and transferred to a nylon membrane (Nytran N, Schleicher & Schuell). For hybridization, PCR products were labelled with biotin-14-dCTP (Gibco BRL). Membranes were incubated in prehybridization solution (4 X SSC, 1% SDS) for 1 h and then transferred to hybridization solution (4 X SSC, 5 X Denhardt’s solution, 0.1% SDS) for a further hour. After adding the denatured probe, membranes were incubated overnight. All incubation steps were performed at 60 °C. Post-hybridization washes were carried out twice in 0.2 X SSC, 0.1% SDS and twice in 0.2 X SSC, 0.1% SDS. Final stringency conditions were 0.1 X SSC, 0.1% SDS at 50°C for 30 min. Incubation with calf intestinal alkaline phosphatase and colour development were performed as described in the Bio-Rad Blot Detection Manual.

Restriction fragments hybridizing to the probes were cloned into pBluescript vector (Stratagene), transformed into *E. coli* strain DH5αF′ and prepared for sequencing.

**DNA sequencing.** Sequencing was done on a Macrophor sequencing system (Pharmacia), using the 35Sequencing kit (Pharmacia) and [α-32P]dATP for labelling. DNA sequences were analysed and amino acid sequences predicted with MacVector software (Eastman Kodak). Multiple alignments were generated with the Clustal V program (Higgins & Sharp, 1988). Sequences were compared to the NCBI databases with the BLAST computer algorithm (Altschul et al., 1990). Amino acid motifs were identified using the Prosite database (Bairoch et al., 1997).

**Results and Discussion**

Based upon SDS–PAGE of purified virus, CVG-1 contains about 50 proteins ranging in size from 10–200 kDa (Fig. 1A).

The most abundant CVG-1 protein, Vp49, is dimeric (Fig. 1A) and is identical to the PBCV-1 major coat protein, Vp54, in the amino acid sequence of the N terminus (Table 1). Protein sequencing also revealed post-translational modification of Vp49, as the N-terminal methionine residue is missing. This is also observed in the major capsid protein of PBCV-1 (M. V. Graves, personal communication) and outer envelope proteins of the Japanese isolate CVK2 (Songsri et al., 1997; see also Table 1).

To detect the presence of glycoproteins, virus preparations were subjected to several glycosylation detection methods. Schiff’s reagent labelled three proteins of molecular mass > 120 kDa and the major protein Vp49 (data not shown).
Table 1. N termini of Chlorella virus capsid proteins

<table>
<thead>
<tr>
<th>Virus</th>
<th>Protein</th>
<th>Gene/clone</th>
<th>SDS–gel</th>
<th>Calculated</th>
<th>N-terminal sequencing</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBCV-1</td>
<td>Vp54</td>
<td>a430l</td>
<td>54*</td>
<td>48</td>
<td>Yes</td>
<td>Acc. no. PBU42580</td>
</tr>
<tr>
<td>CVG-1</td>
<td>Vp49</td>
<td>vcr7</td>
<td>49*</td>
<td>48</td>
<td>Yes</td>
<td>Acc. no. AR076921</td>
</tr>
<tr>
<td>CVT2</td>
<td>Vp54</td>
<td>up54</td>
<td>54</td>
<td>48</td>
<td>NA</td>
<td>Acc. no. AR032579</td>
</tr>
<tr>
<td>CVK2</td>
<td>Vp52</td>
<td>mcp1</td>
<td>52</td>
<td>48</td>
<td>Yes</td>
<td>Acc. no. AR032579</td>
</tr>
<tr>
<td>PBCV-1</td>
<td>Vp45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Songsri et al. (1997)</td>
</tr>
<tr>
<td>CVG-1</td>
<td>Vp45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Songsri et al. (1997)</td>
</tr>
<tr>
<td>CVK2</td>
<td>Vp52</td>
<td>mcp1</td>
<td>52</td>
<td>48</td>
<td>Yes</td>
<td>Songsri et al. (1997)</td>
</tr>
<tr>
<td>PBCV-1</td>
<td>Vp45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Songsri et al. (1997)</td>
</tr>
<tr>
<td>CVG-1</td>
<td>Vp45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Songsri et al. (1997)</td>
</tr>
</tbody>
</table>

*The indicated molecular masses apply to the monomeric forms of Vp54 and Vp49 (the native proteins are dimers). The higher molecular mass of Vp54 as determined by SDS–PAGE is due to glycosylation. The protein runs at 48 kDa when deglycosylated.

†Protein band with corresponding size of 58 kDa in both PBVC-1 and CVG-1 protein preparations.

NA, Not applicable.
Fig. 2. For legend see facing page.
Coat protein genes of a Chlorella Pbi virus

Further sequence analysis

Two complete ORFs were isolated from the CVG-1 genomic library: clone pcr8 was part of a 1566 bp ORF (viral coding region vcr8); clone pcr7 belonged to a 1299 bp ORF, termed vcr7. For clone pcr3, no matching phage subclones were identified.

The A/T content of the 50 bp preceding the start codons of vcr7 and vcr8 (76 and 78% respectively) was significantly higher than in the coding regions (42 and 51%). This is typical for major ORFs in PBCV-1 (Schuster et al., 1990; Lu et al., 1995). Also, a string of 10 and 7 A residues respectively is found at position −30 upstream from the start codons. This motif plays an important role in promoter regions of Shope fibroma virus early genes (Macaulay & McFadden, 1989) and is present 5' to a number of PBCV-1 genes (Graves & Meints, 1992).

Sequence similarities to proteins from other Chlorella viruses

We compared our data mainly to nucleotide sequences from PBCV-1, the prototype of the NC64A group. When available, genes from other NC64A viruses were used for a group-specific evaluation of sequence similarities.

vcr7. We believe that vcr7 encodes the CVG-1 major capsid protein for the following reasons. (1) The predicted amino acid sequence of vcr7 exactly matched the N-terminal sequence of the isolated protein. (2) The calculated molecular mass of 48 kDa for the vcr7 peptide agrees with peptides encoded by the major coat protein genes from NC64A isolates PBCV-1 (Graves & Meints, 1992) and CVT2 (Nishida et al., 1998; see also Table 1). Differences in the apparent molecular masses, as determined by SDS–PAGE, seem to result from differences in glycosylation (Graves & Meints, 1992). (3) The vcr7 peptide shares 78% amino acid identity with the major coat proteins of PBCV-1 and CVT2 (Table 1). Strikingly, the latter two proteins are almost identical in sequence (>98% amino acid identity), despite being from virus strains isolated in countries remote from each other (USA/Japan).

The similarity between the NC64A virus protein sequences and the divergence of the Pbi virus protein sequence respectively support the grouping of the NC64A and Pbi viruses according to host specificity. This observation is further substantiated by a comparison of amino acid sequences inferred from PCR fragments of DNA polymerase genes (Chen & Suttle, 1996). Again, the peptides from NC64A viruses (PBCV-1 and NY-2A) are more than 90% identical, while the sequences of PBCV-1 and Pbi virus CVA-1 share 78% amino acid identity.

The major capsid proteins of Pbi virus CVG-1 and NC64A virus PBCV-1 both seem to be part of a gene family. In our study we identified two more ORFs which share several conserved amino acid motifs with the major capsid protein gene vcr7 (overall amino acid identity about 40%). Similarly, the genome of PBCV-1 contains four such ORFs (Li et al., 1997). In fact, a homologue for each CVG-1 ORF is present in the PBCV-1 gene family.

vcr8. This ORF encodes a protein of 521 amino acids with a calculated molecular mass of 58 kDa. The sequence shares 72% amino acid identity with a622l of PBCV-1, which is expressed as a late gene (M. V. Graves, personal communication), and is also similar to the major coat protein gene of that virus. Protein bands of approximately 58 kDa appear in both virus protein patterns (Fig. 1A and Skrdla et al., 1984), but are not labelled as glycosylated or surface proteins. The vcr8 peptide is longer than the major coat protein, with stretches of similarity in the N-terminal and C-terminal regions and a domain of about 100 amino acids that cannot be aligned to the other ORFs, except a622l (Fig. 2).

pcr3. This fragment contains an incomplete ORF which is moderately similar to a558l from PBCV-1 (Fig. 2). With a calculated amino acid identity of 36%, the predicted amino acid sequences of pcr3 and a558l are clearly more divergent than the other matching sets of ORFs.

Conclusions

Our results establish several common features between Pbi virus CVG-1 and NC64A viruses. (1) The major capsid protein Vp49 is similar in size and dimeric form to major capsid proteins from members of the NC64A group. (2) The corresponding major coat protein gene is part of a gene family which includes at least two more ORFs. (3) The amino acid sequences predicted from the three ORFs identified in this study each have homologues in the coat protein gene family of NC64A virus PBCV-1. In this context, it is interesting that the predicted amino acid sequences of the Pbi and NC64A viruses examined to date show, with one exception, a consistent

Fig. 2. Multiple sequence alignment of CVG-1 ORFs with corresponding PBCV-1 genes. Note that certain domains of vcr7 and a430l differ decidedly from each other (aa 310–330, 413–441 and 519–536). A stretch of about 100 amino acids (aa 92–196) in vcr8 cannot be aligned to the other ORFs except a622l. Bold letters indicate amino acid residues used to design oligonucleotides. Inverted (white-on-black) boxes denote amino acid motifs occurring in all ORFs. PBCV-1 ORFs were taken from the complete genome sequence (accession no. PBU42580).
divergence of ca. 25%. Taking into account that NC64A virus sequences are almost identical, this divergence supports the original virus grouping according to host specificity. The virus strains clearly have evolved separately for some time, probably parallel to their respective host strains, endosymbiotic Chlorella strains Pbi and NC64A.

The authors thank Professor Dietrich Gradmann for accommodating this project in his group. We also thank Dr Heinz Feldmann, Institut für Virologie, Philipps-Universität Marburg, for the opportunity to perform the glycosylation experiments in his laboratory. Special thanks to Dr Michael Graves for critically reading the manuscript.

References


Received 17 July 1998; Accepted 25 November 1998