VIROLOGY 234, 300–303 (1997)
ARTICLE NO. VY978654

The Genome-Linked Protein of Potato Leafroll Virus Is Located Downstream

F. van der Wilk,[1] M. Verbeek, A. M. Dullemans, and J. F. J. M. van den Heuvel

Department of Virology, DLO Research Institute for Plant Protection (IPO-DLO), P.O. Box 9060, 6700 GW Wageningen, The Netherlands

Received April 22, 1997; returned to author for revision May 8, 1997; accepted June 4, 1997

The sequence of the 32 N-terminal amino acids of the protein (VPg) which is covalently linked to the RNA of potato leafroll virus has been determined. The obtained VPg sequence mapped to position 400 to 431 of the PLRV ORF1 product, downstream of the putative protease domain and in front of the RNA-dependent RNA polymerase. Comparison with other viral sequences revealed significant similarities with the ORF1 products of beet western yellows virus, cucurbit aphid-borne yellows virus, and beet mild yellowing virus.

INTRODUCTION

The genus *Luteovirus* is classified into the subgroups I and II, which differ in their genomic organization and type of RNA-dependent RNA polymerase. The polymerases of subgroup II luteoviruses are similar to those of the sobemoviruses, while the subgroup I polymerases are related to those of the diantho- and carmoviruses. The genome of potato leafroll virus (PLRV), like all subgroup II luteoviruses, consists of a single-stranded (ss) messenger-sense RNA molecule which is covalently linked to a small protein (VPg) (Mayo *et al.,* 1982). The genomic sequence of PLRV contains six open reading frames (ORFs) (Van der Wilk *et al.,* 1989; Mayo *et al.,* 1989; Keese *et al.,* 1990), which are separated into two clusters of three genes by an intergenic region (Fig. 1). ORF3 and ORF5 which are present in the 3′-half of the genome encode the structural proteins. The ORF4 underlying the ORF3 encodes the putative movement protein (Tacke *et al.,* 1993). The function of the 5′-terminal ORF0 product is unknown but expression of this ORF in transgenic potato plants has been shown to induce viral disease-like symptoms (Van der Wilk *et al.,* 1997). The ORF1 product (P1) contains motifs characteristic of serine-like proteinases (Gorbalenya *et al.,* 1989; Bazan and Fletterick, 1989). ORF2 overlaps ORF1 and is expressed through a −1 translational frameshift (Prüfer *et al.,* 1992; Kujawa *et al.,* 1993) and codes for the putative RNA-dependent RNA polymerase (Van der Wilk *et al.,* 1989, Mayo *et al.,* 1989; Habili and Symons, 1989).

The gene encoding the VPg of luteoviruses has not been identified yet. The molecular masses of the VPg's of two luteoviruses, PLRV and barley yellow dwarf virus-RPV (BYDV-RPV), have been determined, and are 7 and 17 kDa, respectively (Mayo *et al.,* 1982; Murphy *et al.,* 1989). Analogous to the picornaviral polyprotein arrangement: VPg-protease-polymerase, it has been suggested that the P1–P2 fusion protein of luteoviruses comprises the VPg (Miller *et al.,* 1995). To identify which ORF encodes the VPg, we have isolated the PLRV VPg and determined its N-terminal sequence.

MATERIALS AND METHODS

Virus purification

The Wageningen isolate of PLRV was maintained on *Physalis floridana* by repeated aphid transfers. Inoculated plants were kept in a glasshouse at 22 ± 2°C for symptom development. Four weeks after inoculation, infected plants were harvested and frozen. Batches of infected leaf material were purified using an enzyme-assisted purification procedure essentially as previously described (Van den Heuvel *et al.,* 1991). The final step was a 20–50% sucrose gradient in 0.1 *M* sodium citrate buffer, pH 6.0. The virus-containing fractions were then collected and the virus was sedimented at 40,000 *g.*

RNA extraction

RNA was extracted from the virus pellet using the RNeasy Total RNA kit (Qiagen, Hilden, Germany) according to the manufacturers instructions. Briefly, the virus particles were disrupted in a highly denaturing lysis buffer, containing guanidinium isothiocyanate and β-mercaptoethanol. Subsequently, the RNA was separated from non-covalently-bound proteins by binding the RNA to silica-gel in a spin column and stringent washing with a buffer containing guanidinium isothiocyanate. The RNA was eluted from the column with DEPC-treated water, ethanol precipitated, and resuspended in 40 μl of water.

[1] To whom correspondence and reprint requests should be addressed. Fax: ++31 317 410113. E-mail: F.vanderWilk@IPO.DLO.NL.

0042-6822/97 $25.00

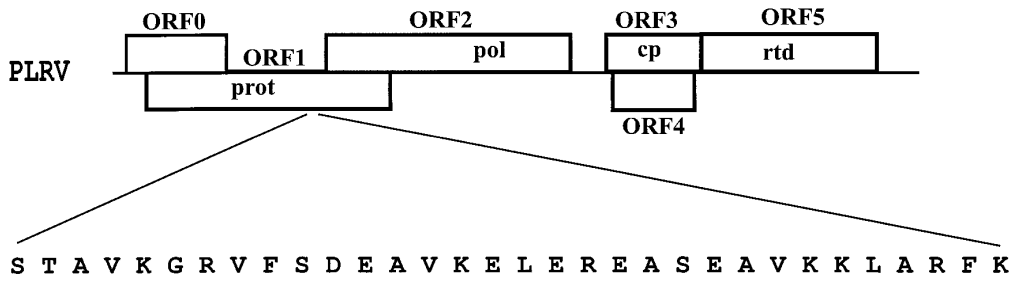

FIG. 1. Genomic organization of PLRV and location of the N-terminal amino acid sequence of the VPg. prot, putative protease domain; pol, putative RNA-dependent RNA polymerase; cp, coat protein; rtd, readthrough domain.

Protein microsequencing

To prepare the VPg for microsequencing, the viral RNA (approximately 70 μg) linked to the protein was hydrolysed in 50 μl of 20% trifluoroacetic acid (TFA) for 48 hr at room temperature (Pinck *et al.,* 1991). The concentration of TFA was then reduced to less than 5% by evaporation, and the hydrolyzed material was blotted onto PVDF membrane using a ProSorb cartridge (Perkin–Elmer, Foster City, CA). The PVDF-immobilized protein was directly subjected to N-terminal amino acid sequencing by automated Edman degradation (Ariad Pharmaceuticals Inc., Cambridge, MA).

Computer analysis

The sequences were compared and compiled using the computer programs Bestfit, Gap, and Pileup of the Genetics Computer Group of the University of Wisconsin (Devereux *et al.,* 1984) and the Blast suite (Altschul *et al.,* 1990).

RESULTS AND DISCUSSION

The N-terminal sequence (Fig. 1) of the PLRV VPg has been determined. A sequence of 32 amino acid residues was unambiguously established. Background signals were not observed, except in the cycles 18, 23, 27, and 29 in which an additional minor signal was present, thus confirming that only one protein was present in the sample and that no degradation of this protein had occurred. Comparison of the obtained amino acid sequence with PLRV sequences revealed that it lined up with residues 400 to 431 of PLRV P1. Hence, it was concluded that the ORF1 encodes the VPg and that the VPg is located downstream of the putative protease domain and upstream of the polymerase.

A search for similarity between the obtained N-terminal sequence and sequences present in peptide databases, using the computer program BLAST, revealed significant similarities (53–56%) with internal P1 sequences of beet western yellows luteovirus (BWYV) (Veidt *et al.,* 1988), cucurbit aphid-borne yellows luteovirus (CABYV) (Guilley *et al.,* 1994), and beet mild yellowing luteovirus (BMYV)(Guilley *et al.,* 1995) (Fig. 2). These sequences probably represent the VPg N-terminal part of each virus. Similarity between the PLRV VPg sequence and BYDV-RPV P1 (Vincent *et al.,* 1991) was not detected. However, alignment of the VPg sequence with the BYDV-RPV P1, using Bestfit, disclosed significant similarity (50%) (Fig. 2).

Unexpectedly, the position of the VPg on the luteoviral genome differs from the picornaviral genomic arrangement: VPg-protease-polymerase, which prevails in all other ssRNA viruses with a VPg.

No similarity could be detected between the obtained PLRV VPg sequence and sequences derived of subgroup I luteoviruses. Previously, it has been suggested that the subgroup I luteoviruses lack a VPg (Miller *et al.,* 1995). The ORF1-products of subgroup I and II luteoviruses share no homology. Moreover, the subgroup I ORF1 products are markedly smaller, which supports the suggestion that a VPg is absent in subgroup I luteoviruses.

The VPg of *Sobemoviridae* has been predicted to be composed of the N-terminal 134 amino acids of the ORF2 product (Gorbalenya *et al.,* 1988; Mäkinen *et al.,* 1995; Ngon A Yassi *et al.,* 1994). The RNA-dependent RNA polymerases of the sobemoviruses are related to those of the subgroup II luteoviruses. Moreover, the polymerase of cocksfoot mottle virus (CfMV) is expressed by a −1 translational frameshift (Mäkinen *et al.,* 1995). It has been proposed that the subgroup II luteoviruses have emerged by a recombination event between a sobemovirus and a subgroup I luteovirus (Miller *et al.,* 1995), the 3′-terminal structural genes originating from the luteovirus, and the 5′-terminal nonstructural genes from the sobemovirus. Therefore, the PLRV VPg sequence was compared to those of the sobemoviral ORF2 products using the computer program Bestfit. Although no significant similarity was observed with the proposed sobemovirus VPg sequences, the PLRV VPg sequence showed significant similarities (65–53%) with residues 347–379 of CfMV, residues 771–814 of Southern bean mosaic virus (SBMV) (Wu *et al.,* 1987) and residues 498–537 of rice yellow mottle virus (RYMV) (Ngon A Yassi *et al.,* 1994). Possibly, the sobemovirus VPg domain is located between the putative protease and polymerase domain, analogous to the luteoviruses.

```
PLRV        401   STAVKGRVFSDEAVKELEREAS.EAVKKLARFK 432
BWYV        401   TTAPQGRVFAQEDIAEIEGLYA.QVMKRVQQAE 432
BMYV        407   TTALAGKFFSQEEVEELMEDFSLQEIYSIATAR 439
CABYV       396   TTAPQGRLFTQEEIEELIEEFSFSEITSIMGHR 428
BYDV-RPV    403   SDPPQGLVFPEEVTESIE.....AAIKEATMYN 430
consensus         ttapqGrvFsqEe.eEleeefs.qeiksia.ar
```

FIG. 2. Alignment of the PLRV VPg and P1 sequences of BWYV (Veidt *et al.,* 1988), CABYV (Guilley *et al.,* 1994), BMYV (Guilley *et al.,* 1995) and BYDV-RPV (Vincent *et al.,* 1991). Numbering refers to amino acid positions in each P1.

Maturation of the VPg requires both proteolytic cleavage at the N- and C-terminus. It is most likely that the PLRV VPg is released from P1 by proteolytic activity of its putative protease domain. In all subgroup II luteoviral P1 sequences the VPg N-terminal residue (S/T) is preceded by a glutamic acid residue, indicating that the N-terminal proteolytic processing site consists of the residues E-S/T. The reported size of 7 kDa of the PLRV VPg maps its C-terminus approximately at residue 465 of the ORF1 product. However, a sequence similar to the putative N-terminal processing site is absent both in the immediate vicinity of the putative VPg C-terminus as well as in the downstream P1 sequence, making it impossible to predict the exact position of the C-terminal residue.

Previously, Habili and Symons (1989) identified two putative helicase domains (IV and VI) in the PLRV and BWYV P1 and P2. The putative helicase motif IV (NYVFESTAV) believed to be present in P1, overlaps with the N-terminal residues and putative cleavage site of the VPg. Therefore, it seems highly unlikely that this sequence represents a conserved helicase domain. This finding corroborates that of Koonin and Dolja (1993) who stated that helicase motifs are absent in luteoviruses.

The polymerases of luteoviruses are expressed by a −1 translational frameshift occurring at a frequency of ~1% (Prüfer *et al.,* 1992). The PLRV VPg domain is located upstream of the translational frameshift site and consequently is expressed at a 100-fold excess to the polymerase. The BYDV-RPV VPg, with a reported size of 17 kDa, might extend beyond the frameshift site and is possibly produced by translational frameshifting. However, it is more likely to assume that analogous to PLRV the BYDV-RPV VPg is encoded entirely by the ORF1. In all other VPg RNA viruses the VPg and polymerase are encoded by the same gene, and expression of the gene products is regulated by sequential proteolytic events. Both the low expression level of the luteo/sobemoviral polymerase and the unusual excess of the VPg to the polymerase suggest that the luteoviral polymerase possesses a high activity.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

Bazan, J. F., and Fletterick, R. J. (1989). Comparative analysis of viral cysteine protease structural models. *FEBS Lett.* 249, 5–7.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.

Gorbalenya, A. E., Koonin, V. E., Blinov, V. M., and Donchenko, A. P. (1988). Sobemovirus genome appears to encode a serine protease related to cysteine proteases of picornaviruses. *FEBS Lett.* 236, 287–290.

Gorbalenya, A. E., Donchenko, A. P., Blinov, V. M., and Koonin, V. E. (1989). Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. *FEBS Lett.* 243, 103–114.

Guilley, H., Richards, K. E., and Jonard, G. (1995). Nucleotide sequence of beet mild yellowing virus RNA. *Arch. Virol.* 140, 1109–1118.

Guilley, H., Wipf-Scheibel, C., Richards, K., Lecoq, H., and Jonard, G. (1994). Nucleotide sequence of cucurbit aphid-borne yellows luteovirus. *Virology* 202, 1012–1017.

Habili, N., and Symons, R. H. (1989). Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicases. *Nucleic Acids Res.* 17, 9543–9555.

Keese, P., Martin, R. R., Kawchuk, L. M., Waterhouse, P. M., and Gerlach, W. L. (1990). Nucleotide sequences of an Australian and a Canadian isolate of potato leafroll luteovirus and their relationships with two European isolates. *J. Gen. Virol.* 71, 719–724.

Koonin, E. V., and Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* 28, 375–430.

Kujawa, A. B., Drugeon, G., Hulanicka, D., and Haenni, A.-L. (1993). Structural requirements for efficient translational frameshifting in the synthesis of the putative viral RNA-dependent RNA polymerase of potato leafroll virus. *Nucleic Acids Res.* 21, 261–2171.

Mäkinen, K., Tamm, T., Naess, V., Truve, E., Puurand, Ü., Munthe, T., and Saarma, M. (1995). Characterization of cocksfoot mottle sobemovirus genomic RNA and sequence comparison with related viruses. *J. Gen. Virol.* 76, 2817–2825.

Mayo, M. A., Barker, H., Robinson, D. J., Tamada, T., and Harrison, B. D. (1982). Evidence that potato leafroll virus is positive-stranded, is linked to a small protein and does not contain polyadenylate. *J. Gen. Virol.* 59, 163–167.

Mayo, M. A., Robinson, D. J., Jolly, C. A., and Hymann, L. (1989). Nucleotide sequence of potato leafroll virus RNA. *J. Gen. Virol.* 70, 1037–1051.

Miller, W. A., Dinesh-Kumar, S. P., and Paul, C. P. (1995). Luteovirus gene expression. *Crit. Rev. Plant Sci.* 14, 179–211.

Murphy, J. F., D'Arcy, C. J., and Clark, J. M. J. (1989). Barley yellow dwarf virus RNA has a 5′-terminal genome-linked protein. *J. Gen. Virol.* 70, 2253–2256.

Ngon A Yassi, M., Ritzenthaler, C., Brugidou, C., Fauquet, C., and Beachy, R. N. (1994). Nucleotide sequence and genome characterization of rice yellow mottle virus RNA. *J. Gen. Virol.* 75, 249–257.

Pinck, M., Reinbolt, J., Loudes, A. M., Le Ret, M., and Pinck, L. (1991). Primary structure and location of the genome-linked protein (VPG) of grapevine fanleaf nepovirus. *FEBS Lett.* 284, 117–119.

Tacke, E., Schmitz, J., Prufer, D., and Rohde, W. (1993). Mutational analysis of the nucleic acid-binding 17 kDa phosphoprotein of potato

leafroll luteovirus identifies an amphiphatic α-helix as the domain for protein/protein interactions. *Virology* 197, 274–282.

Prüfer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A., and Rohde, W. (1992). Ribosomal frameshifting in plants: a novel signal directs the −1 frameshift in the expression of the putative viral replicase of potato leafroll luteovirus. *EMBO J.* 11, 1111–1117.

Veidt, I., Lot, H., Leiser, M., Scheidecker, D., Guilley, H., Richards, K., and Jonard, G. (1988). Nucleotide sequence of beet western yellows RNA. *Nucleic Acids Res.* 16, 9917–9932.

Van den Heuvel, J. F. J. M., Boerma, T. M., and Peters, D. (1991). Transmission of potato leafroll virus from plants and artificial diets by *Myzus persicae*. *Phytopathology* 81, 150–154.

Van der Wilk, F., Huisman, M. J., Cornelissen, B. J. C., Huttinga, H., and Goldbach, R. (1989). Nucleotide sequence and organization of potato leafroll virus genomic RNA. *FEBS Lett.* 245, 51–56.

Van der Wilk, F., Houterman, P., Molthoff, J., Hans, F., Dekker, B., Van den Heuvel, J., Huttinga, H., and Goldbach. R. (1997). Expression of the potato leafroll virus ORF0 induces viral disease-like symptoms in transgenic potato plants. *Mol. Plant-Microbe Interact.* 10, 153–159.

Vincent, J. R., Lister, R. M., and Larkins, A. (1991). Nucleotide sequence analysis and genomic organization of the NY-RPV isolate of barley yellow dwarf virus. *J. Gen. Virol.* 72, 2347–2355.

Wu, S., Rinehart, C., and Kaesberg, P. (1987). Sequence and organization of southern bean mosaic virus genomic RNA. *Virology* 161, 73–80.