Volume 284, number 1, 117–119 FEBS 09819 June 1991
 00145793/91/$3.50
ADONIS 0014579391005274

Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus

M. Pinck[1], J. Reinbolt[2], A.M. Loudes[1], M. Le Ret[2] and L. Pinck[1]

[1]*Institut de Biologie Moléculaire des Plantes du C.N.R.S. et Université Louis Pasteur, Laboratoire de Virologie, 12 rue Gl Zimmer, 67084 Strasbourg, France* and [2]*Institut de Biologie Moléculaire et Cellulaire du C.N.R.S., 15 rue Descartes, 67084 Strasbourg, France*

Received 2 April 1991

The genome linked protein VPg covalently linked to the RNAs of grapevine fanleaf nepovirus has been sequenced. The VPg ($M_r = 2931$) composed of 24 residues is linked by its N-terminal Ser β-OH group to the viral RNAs. The VPg mapped from residues 1218 to 1241 of the 253K polyprotein encoded by GFLV RNA1.

Nepovirus; Grapevine fanleaf virus; VPg protein; Primary structure

1. INTRODUCTION

The genome of grapevine fanleaf nepovirus (GFLV) consists of two messenger-sense RNA [1] of 7342 nucleotides for RNA1 (Ritzenthaler, C., personal communication) and 3774 nucleotides for RNA2 [2], translated into large primary translation products (polyproteins) of 253K and 122K, respectively, which are processed in functional proteins by proteolytic cleavage. The two genomic RNA and the satellite RNA in the F13 isolate of GFLV are characterized by a 3′-terminal poly(A) tail and a small viral protein, denoted VPg, is assumed to be covalently attached to the 5′ terminal nucleotide of each RNA species [3]. The viral RNA of nepoviruses, comoviruses and other viruses of the superfamily of the picorna-like viruses share similar properties [4]. For nepoviruses the sequences of VPg and its mapping in the genome are unknown. For comoviruses which have a genomic organisation similar to the nepoviruses the sequence of the VPg of CPMV and its linkage by a phosphodiester bond between a serine and the 5′ terminal uridylate of CPMV RNA have been extensively studied [5,6].

The VPg and the coat protein are the two structural proteins which can be purified from the virus and for which amino acid sequence data can be obtained from microsequencing. This information allows the identification of the proteolytic cleavage sites in the polyprotein and their mapping on the viral genome. We have so far identified the proteolytic cleavage at R/G for the coat protein gene of GFLV and located the CP gene on RNA2 [2]. Here we report the complete amino acid sequence of the VPg from GFLV and the mapping of VPg on the polyprotein encoded by RNA1.

Correspondence address: L. Pinck, Laboratoire de Virologie, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France. Fax: (33) (88) 614442

2. MATERIALS AND METHODS

2.1. *Virus extraction and RNA purification*

The F13 isolate of GFLV was propagated on *Chenopodium quinoa* and virus was extracted from systemically infected leaves 15 days after infection. Virus particles were purified as described previously [3] except that after two ultracentrifugation steps the virus was purified through a 5–50% linear sucrose gradient in 10 mM sodium phosphate, 1 mM EDTA, pH 7.0 buffer for 5 h at 27000 rpm in a SW27 rotor Beckman instead of sedimentation through a 20% sucrose cushion. The viral RNA was prepared by the phenol-SDS procedure from the gradient fractions corresponding to the bottom component, collected and sedimented by ultracentrifugation. The purification of RNA was achieved by precipitation of the RNA from the aqueous layer by 2 M LiCl at 0°C. The RNA was pelleted, dissolved in sterile water, again precipitated with 2 M LiCl and finally ethanol precipitated in the presence of 0.1 M NaCl.

2.2. *VPg preparation for microsequencing*

Two methods were used to hydrolyse the RNA linked to VPg, (1) GFLV RNA (370 μg) was hydrolysed in 100 μl of 20% trifluoroacetic acid (TFA) for 48 h at room temperature, (2) the same amount of RNA was digested with RNase T1 (0.1 U/μg) for 30 min incubation at room temperature according to [7]. The hydrolysed material was directly subjected to sequencing.

2.3. *Sequence analysis of VPg protein*

The VPg was entirely sequenced by automated Edman's degradation using an Applied Biosystems 470A protein sequencer equipped with a PTH 120A analyser [8]. Carboxypeptidase A from Boehringer (Mannheim, Germany) was used to confirm the C-terminal sequence of VPg. The digestion of VPg, obtained from 350 μg RNA and estimated to 0.6 μg of protein, was performed at 37°C for 5 h in NH_4HCO_3 50 mM, diisopropylfluorophosphate (DIFP) 10^{-4} M, pH

8 buffer. The enzyme/protein ratio was 1:30. Analysis of the amino acids released by this treatment was performed with a 420A-130A derivatisation and analysis system (Applied Biosystems, USA). The possible presence of phosphoserine was checked by the method using ethanethiol to convert phosphoserine into *S*-ethylcysteine [9].

3. RESULTS AND DISCUSSION

As determined for the other picorna-like viruses [4] it is expected that VPg is linked to GFLV RNAs by a phosphodiester bond. Two different approaches were used to cleave the RNA from VPg which yield either VPg linked to the ultimate 5′-end nucleotide after TFA hydrolysis or to a short oligonucleotide after RNase T1 digestion of the viral RNAs (VPg-AUG for RNA1 and 2 and VPg-UAUG for the satellite RNA as deduced from their 5′ terminal sequence [1,2,10]). In a first attempt the complex VPg-viral RNA was subjected to the hydrolysis in TFA 20% as described above. Under these conditions no cleavage occurred within the polypeptide chain of VPg, the RNA however was almost cut off in its nucleotides or oligonucleotides as observed on a 20% polyacrylamide sequencing gel upon analysis of labelled RNA. As to the VPg we could establish unambiguously its sequence:

```
1                  10                  20    24
S E P R L E E R Y S P R N R F V S R I S K I R G
```

The C-terminal residue Gly was confirmed by digestion with carboxypeptidase A which liberated only a Gly residue. Furthermore the targets of a possible phosphodiester linkage with the RNA, namely one Tyr and 4 Ser residues, are characterized during the sequencing as free of any additional nucleotide or phosphate. Thus so far the 3 GFLV RNAs could be iodinated [3], indicating, since the Tyr residues are preferentially labelled with the method used [11], that the only Tyr residue of VPg is not involved in the linkage to the viral RNA. As to the Ser residues, the cleavage of the RNA with 20% TFA could have generated a phosphoserine-containing VPg, which obviously yields the dithiothreitol (DTT) adduct of PTH-Ser upon Edman degradation as mentioned in [9]. Thus no identifiable PTH derivatives of phosphoserine can be obtained by sequencing. To overcome this difficulty Meyer et al. [9] described a method converting quantitatively phosphoserine into *S*-ethylcysteine by β-elimination and subsequent addition of ethanethiol just before sequencing. This derivative is detectable on an Applied Biosystems gas-phase sequencer. As to VPg, there was no phosphoserine within the sequence at all, it seems that 20% TFA would cut off the phosphodiester bond as well.

Therefore, in a second experiment the complex VPg-RNA was digested with RNase T1. Indeed RNase T1 should generate VPg linked to AUG and/or UAUG and the conditions used for this cleavage should be milder than those of a 20% TFA cleavage. This might allow the identification of the residue involved in the phosphodiester bond between VPg and viral RNA. Again we were able to determine the whole sequence of VPg from residue Glu at position 2 until the C-terminal residue Gly at position 24. However the first N-terminal residue Ser could not be identified: indeed there is a blank at the first sequencing step which strongly suggests that Ser at position 1 is still modified after digestion of the VPg-RNA complex with RNase T1. Thus this result favourably supports the fact that viral RNA is bound to N-terminal Ser of VPg, all other possible targets being recovered unmodified in the automated sequential Edman's degradation.

All these data seem to indicate that VPg protein is likely to be linked to the viral RNA through its N-terminal Ser by a phosphodiester bond. We also could establish its primary structure by microsequencing of about 20 pmol protein. This sequence corresponds to the protein between residues 1218 and 1241 of the 253 kDa polyprotein encoded by RNA1 (Fig. 1). Its appears that VPg (M_r = 2931) is matured from the polyprotein by the cut of a Cys-Ser bond (1217-1218) at its N-terminal part and by the cleavage of a Gly-Glu bond (1241-1242) at its C-terminal end. The proteins upstream and downstream of the VPg correspond to viral proteins displaying the characteristic signatures of NTP-binding protein and of viral protease found in comoviruses [1,12].

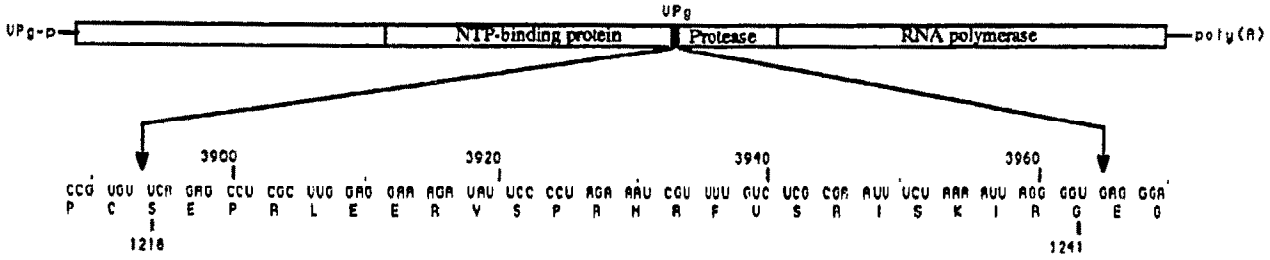

Fig. 1. Amino acid sequence of GFLV VPg deduced from the sequence of the cDNA copy of RNA1. The extent of the VPg sequence is indicated by arrows pointing the cleavage sites in the 253K polyprotein. The 253K polyprotein of 2284 residues is represented in the upper part by a box on RNA1, in which the positions of different viral proteins and VPg are indicated drawn to scale.

REFERENCES

[1] Pinck, L., Pinck, M., Serghini, M.A., Ritzenthaler, C., Fuchs, B. and Walter, B. (1990) Proceedings of the 10th ICVG Meeting in Volos (Greece), in press.
[2] Serghini, M.A., Fuchs, M., Pinck, M., Reinbolt, J., Walter, B. and Pinck, L. (1990) J. Gen. Virol. 71, 1433–1441.
[3] Pinck, L., Fuchs, M., Pinck, M., Ravelonandro, M. and Walter, B. (1988) J. Gen. Virol. 69, 233–239.
[4] Eggen, R. and Van Kammen, A. (1988) in: RNA genetics (Ahlquist, P., Holland, J. and Domingo, E. eds) vol. I, pp. 49–69. CRC Press, Boca Raton.
[5] Zabel, P., Moerman, M., Lomonossoff, G., Shanks, M. and Beyreuther, K. (1984) EMBO J. 3, 1629–1634.
[6] Jaegle, M., Wellink, J. and Goldbach, R. (1987) J. Gen. Virol. 68, 627–632.
[7] Wikman, F.P., Romby, P., Metz, M.H., Reinbolt, J., Clark, B.F.C., Ebel, J.P., Ehresmann, C. and Ehresmann, B. (1987) Nucleic Acids Res. 15, 5787–5801.
[8] Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981) J. Biol. Chem. 256, 7990–7997.
[9] Meyer, H.E., Hoffmann-Posorske, E., Korte, H. and Heilmeyer Jr, L.M.G. (1986) FEBS Lett. 204, 61–66.
[10] Fuchs, M., Pinck, M., Serghini, M.A., Ravelonandro, M., Walter, B. and Pinck, L. (1989) J. Gen. Virol. 70, 955–962.
[11] Koenig, I. and Fritsch, C. (1982) J. Gen. Virol. 60, 343–353.
[12] Wellink, J. and Van Kammen, A. (1988) Arch. Virol. 98, 1–26.