JOURNAL OF VIROLOGY, Jan. 1991, p. 511–513
0022-538X/91/010511-03$02.00/0

Vol. 65, No. 1

A Tyrosine Residue in the Small Nuclear Inclusion Protein of Tobacco Vein Mottling Virus Links the VPg to the Viral RNA

JOHN F. MURPHY,[1] WOJCIECH RYCHLIK,[2] ROBERT E. RHOADS,[2] ARTHUR G. HUNT,[3] AND JOHN G. SHAW[1]*

Departments of Plant Pathology,[1] Biochemistry,[2] and Agronomy,[3] University of Kentucky, Lexington, Kentucky 40546-0091

Received 16 June 1990/Accepted 11 October 1990

The identity of the amino acid residue that links the VPg of the potyvirus tobacco vein mottling virus (TVMV) to the viral RNA was determined. ^{32}P-labeled TVMV RNA was digested with RNase A and micrococcal nuclease. The resulting ^{32}P-labeled VPg was isolated and partially hydrolyzed with 6 N HCl at 110°C for 2 h. Analysis by thin-layer electrophoresis revealed the presence of [^{32}P]phosphotyrosine but not [^{32}P]phosphoserine or [^{32}P]phosphothreonine. Another preparation of TVMV RNA was treated with endoproteinase Lys-C, and the resulting peptide-RNA was purified by sodium dodecyl sulfate-sucrose gradient centrifugation. The sequence of the N-terminal 15 amino acid residues of the peptide, when compared with the RNA-derived amino acid sequence of the TVMV polyprotein, demonstrated that the peptide occurs in the small nuclear inclusion protein. These data suggest that Tyr-1860 of the polyprotein is the amino acid residue that links the TVMV VPg to the viral RNA.

Plant potyviruses are flexuous rod-shaped particles whose monopartite genomes consist of a 9- to 10-kb, single-stranded, polyadenylated, positive-sense RNA molecule (8). The RNAs of several potyviruses have been shown to have a protein (VPg) covalently linked to the 5′ terminus. The VPg of tobacco etch virus (TEV) was originally reported to be a protein of 6 kDa (10), but other investigations indicated that the VPgs of tobacco vein mottling virus (TVMV) and plum pox virus (PPV) may be 22- to 24-kDa proteins (20, 25) and that in TVMV, the VPg is the N-terminal 24-kDa portion of NIa, the small nuclear inclusion protein (24). In a more recent report, Murphy et al. (17) have suggested that the 49-kDa proteinase (3) of TEV, the homolog of the NIa of TVMV, is itself the VPg and that the 24-kDa protein associated with some TEV RNA molecules is a breakdown product thereof.

With some members of the picornavirus and comovirus groups, viruses which, like potyviruses, express their constituent genes by viral protease-mediated processing of viral proteins (9, 18), the amino acid residue in the VPg that links the VPg to the genomic RNA has been identified. Such information concerning potyviruses has not been reported, although it is essential for studies of the role of the VPg in replication and pathogenicity. We therefore undertook an analysis of the VPg of TVMV and identified a tyrosine residue in NIa as the amino acid residue to which the RNA is linked.

We used two approaches to investigate the TVMV VPg-RNA linkage. (i) ^{32}P-labeled TVMV RNA was digested with nucleases, and the resulting [^{32}P]VPg, after partial hydrolysis, was analyzed to identify a ^{32}P-labeled amino acid. (ii) Viral RNA was treated with a protease, and the peptide that remained associated with the RNA was subjected to Edman degradation sequence analysis.

TVMV RNA was labeled with ^{32}P in infected tobacco mesophyll protoplasts. *Nicotiana tabacum* L. cv. Xanthi mesophyll protoplasts were isolated as described by Takebe et al. (26), with modifications (16). Approximately 4×10^7 protoplasts were inoculated by electroporation (15) with 750 μg of TVMV RNA which had been isolated from purified TVMV as described by Murphy et al. (17). The electroporated protoplasts were incubated in 40 ml of incubation medium (16, 26) containing 15 mCi of $H_3{}^{32}PO_4$ at 25°C under 35 to 55 microeinsteins of fluorescent illumination m^{-2} s^{-1}. After 50 h, the protoplasts were disrupted by addition of 20 ml of cold chloroform. The mixture was shaken vigorously, and the phases were separated by centrifugation. Nonradioactive TVMV (500 μg), NaCl (to 0.25 M), and polyethylene glycol 8000 (to 4%, wt/vol) were added to the aqueous phase, which was stirred overnight at 4°C and then centrifuged at 12,000 × *g* for 10 min. The resulting pellet was suspended in 10 mM Tris hydrochloride (pH 7.4)–1 mM EDTA and subjected to density gradient centrifugation in Cs_2SO_4 (19). The virus was collected, and the RNA was purified by incubation in guanidinium isothiocyanate and centrifugation through CsCl (4). The labeled TVMV RNA was analyzed by electrophoresis in a formaldehyde-formamide agarose gel. A band which comigrated with ethidium bromide-stained, nonlabeled TVMV RNA was detected by autoradiography (data not shown).

The ^{32}P-labeled TVMV RNA was treated with RNase A and micrococcal nuclease, and the resulting [^{32}P]VPg was isolated and partially hydrolyzed in 6 N HCl. Analysis by thin-layer electrophoresis revealed three ^{32}P-labeled species (Fig. 1). The fastest-migrating species was P_i; the slowest was not identified but may have been incompletely hydrolyzed RNA. The intermediate species migrated to the same position as the ninhydrin-stained phosphotyrosine standard, suggesting that a tyrosine residue links the VPg to the viral RNA.

With the second approach used to identify the amino acid that links the VPg to TVMV RNA, the intention was to cleave the VPg into peptides while retaining the integrity of the RNA. One of the resulting peptides should remain covalently linked to the RNA and would be isolated by sodium dodecyl sulfate (SDS)-sucrose gradient purification of the RNA. This peptide would then be subjected to Edman degradation sequence analysis.

TVMV RNA (1 mg) in 25 mM Tris hydrochloride (pH

* Corresponding author.

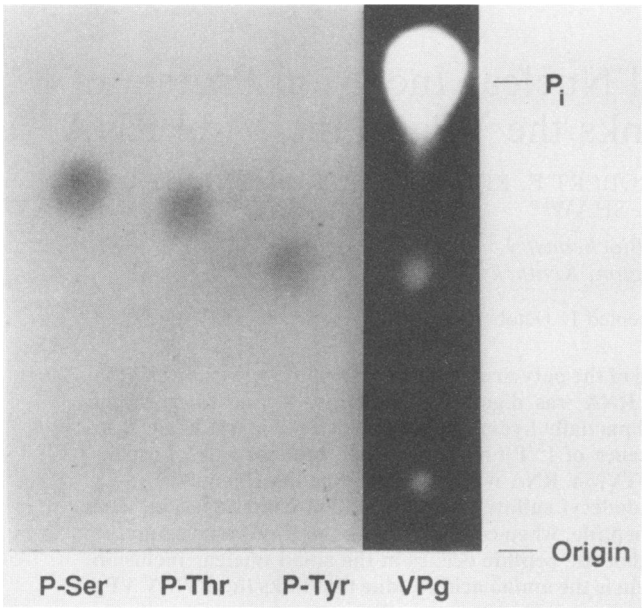

FIG. 1. Thin-layer electrophoretic analysis of ^{32}P-labeled amino acids from TVMV VPg. ^{32}P-labeled TVMV RNA was treated with 1 μg of RNase A for 2.5 h at 37°C, followed by addition of $CaCl_2$ to 2 mM and treatment with 15 U of micrococcal nuclease for 2 h at 37°C in a total volume of 100 μl. The resultant [^{32}P]VPg was purified by Sephadex G25 chromatography (11) and precipitated after addition of NaCl to a final concentration of 10 mM and 3 volumes of methanol-acetic acid (100:1) at −20°C. The [^{32}P]VPg was hydrolyzed for 2 h at 110°C in 6 N HCl (23), and the product was analyzed by thin-layer electrophoresis on cellulose plates in 0.05% acetic acid and 0.005% pyridine. Phosphoserine (P-Ser), phosphothreonine (P-Thr), and phosphotyrosine (P-Tyr) standards were visualized with ninhydrin, and P_i was identified relative to a $^{32}P_i$ standard.

7.7)–1 mM EDTA containing 80 U of RNasin (Promega Biotec, Madison, Wis.) was treated with 0.05 U of endoproteinase Lys-C (Promega) in a total volume of 0.5 ml for 24 h at 37°C. The protease-treated RNA was then subjected to SDS-sucrose gradient centrifugation (17), collected, and precipitated with ethanol after addition of sodium acetate to 0.1 M. The RNA was pelleted by centrifugation at 12,000 × *g* for 40 min, suspended in water, and subjected to a second ethanol-acetate precipitation. In preliminary experiments, these treatments resulted in cleavage of the VPg into peptides of 2 to 10 kDa. When the RNA was subjected to additional cycles of protease digestion and SDS-sucrose gradient purification, electrophoretic analysis indicated that a single peptide species of approximately 6 kDa remained linked to the viral RNA (data not shown). Thus, for complete cleavage of the VPg and rigorous purification of the RNA-linked peptide, four successive treatments consisting of protease digestion, SDS-sucrose gradient centrifugation, and two ethanol-acetate precipitations were carried out. After the fourth cycle, the RNA was subjected to an additional cycle of SDS-sucrose gradient purification and two ethanol-acetate precipitations. The peptide that resulted from these treatments of TVMV RNA was subjected to Edman degradation sequence analysis with an Applied Biosystems 477A protein sequencer.

The analysis yielded the sequence of the N-terminal 15 amino acid residues of the peptide that remained linked to the TVMV RNA. Interpretation of the data generated during separation of phenylthiohydantoin (PTH) amino acids by high-performance liquid chromatography was made by the staff of the sequencing facility (Macromolecular Structure Analysis Facility, University of Kentucky, Lexington) without reference to the known sequence of the TVMV polyprotein (6) (Table 1). The resulting sequence was then compared with the predicted amino acid sequence of the TVMV polyprotein and found to correspond to residues 1856 to 1870. Thirteen of the amino acid assignments matched the predicted sequence of the polyprotein (Fig. 2). A proline residue was suspected in cycle 9 but was not unambiguously identified because of the small amount of the amino acid that was recovered. No amino acid residue was recovered from the fifth cycle, which, according to the predicted amino acid sequence of the polyprotein, is a tyrosine (residue 1860). However, other tyrosine residues, Tyr-1867 and Tyr-1869, were obtained during subsequent cycles. We suggest that Tyr-1860 was not identified during sequence analysis because it was linked, via a phosphodiester bond, to the RNA.

TABLE 1. Edman degradation sequence analysis of an RNA-linked peptide in the VPg of TVMV[a]

Cycle	Amino acid identified	Concn (pmol)
1	Phe	64.07
2	Val	63.88
3	Asn	37.05
4	Met	32.28
5	—	—
6	Gly	52.90
7	Val	42.75
8	Ser	27.92
9	Pro?	4.86
10	Asp	13.64
11	Glu	14.41
12	Tyr	13.36
13	Ser	17.71
14	Tyr	20.53
15	Val	24.74

[a] The endoproteinase Lys-C-digested peptide (approximately 200 pmol) which remained attached to TVMV RNA was subjected to Edman degradation sequence analysis. In each cycle, the most likely amino acid residue released was identified without reference to the nucleotide sequence of TVMV RNA (6). The identity of the predominant amino acid residue and the concentration of that residue recovered after each cycle are shown. —, No PTH amino acid recovered. Cleavage of proline residues during Edman degradation is typically low; thus, identification of proline in cycle 9 was not certain but suspected.

TVMV (peptide)		F	V	N	M	-	G	V	S	P?	D	E	Y	S	Y	V	
TVMV	1855 K	**F**	**V**	**N**	**M**	**Y**	**G**	**V**	**S**	**P**	**D**	**E**	**Y**	**S**	**Y**	1870 **V**	(ref. 6)
TEV	1906 K	**F**	I	**N**	**M**	**Y**	**G**	F	D	**P**	T	D	F	S	Y	1921 I	(ref. 1)
PPV	1914 K	**F**	**V**	**N**	**M**	**Y**	**G**	V	D	**P**	T	D	Y	N	F	1929 **V**	(ref. 14)
PVY	1902 R	**F**	I	**N**	**M**	**Y**	**G**	F	D	**P**	T	E	Y	S	F	1917 I	(ref. 21)

FIG. 2. Alignment of the amino acid sequence of the RNA-associated peptide of TVMV VPg with amino acid sequences in the polyproteins of TVMV, TEV, PPV, and PVY. The TVMV VPg peptide sequence was determined by Edman degradation sequence analysis, and those of TVMV, TEV, PPV, and PVY were predicted on the basis of RNA sequences (references are noted at the right). Boldface lettering is used for conserved residues, and the proposed tyrosine residue linking the VPg to the RNA is shaded.

Thus, on the basis of the association of phosphotyrosine with the RNA (Fig. 1) and the results of sequence analysis of the RNA-bound peptide (Table 1), we propose that Tyr-1860 is the amino acid residue that links the TVMV VPg to the viral RNA.

Extensive sequence homology was found between the amino acid residues surrounding Tyr-1860 of the polyprotein of TVMV and the corresponding residues of the polyproteins of TEV (1), PPV (14), and potato virus Y (PVY; 21) (Fig. 2). Within the consensus sequence B-F-Z-N-M-Y-G-Z-(S,D)-P-X-Z-(S,N)-Z-Z, perfect identity occurred at positions −4, −2, −1, +1, and +4 relative to Tyr-1860 of TVMV and similar amino acid residues were found at the other positions (i.e., basic [B] at −5, acidic [X] at +5, and hydrophobic [Z] at −3, +2, +6, +8, and +9). It should be noted that the TVMV RNA-linked peptide contained a fourth tyrosine residue (Tyr-1872) that was not present in the sequenced N-terminal 15 amino acids described above. However, in the regions of the four potyviral polyproteins corresponding to the TVMV RNA-linked peptide, the only conserved tyrosine residue is that corresponding to Tyr-1860 in the TVMV polyprotein. This provides further evidence that Tyr-1860 is the tyrosine residue that is linked to TVMV RNA. It also suggests that the VPgs of TEV, PPV, PVY, and perhaps other potyviruses are linked to the viral RNAs by these conserved tyrosine residues.

The organization of the potyviral genome is similar to that of the picornaviruses and comoviruses (7). The poliovirus VPg was shown to be linked to the viral RNA by a tyrosine residue (2, 22), and the cowpea mosaic virus VPg is linked to the viral RNA by a serine residue (12). Studies on the function of the poliovirus VPg suggest that it is involved in virus replication (13), acting as a primer (27), or is involved in cleavage of the replicative form of the RNA (28). Identification of Tyr-1860 as the amino acid that links the TVMV VPg to the viral RNA allows us to begin studies on the function of this VPg by analysis of in vitro-generated mutant TVMV RNAs (5).

We thank Carol Beach for interpretation of the sequence data and Richard Spalding for excellent assistance.

This work was supported by grant 4E021 from the University of Kentucky Tobacco and Health Research Institute and grant 85-CRCR-1-1536 from the U.S. Department of Agriculture Competitive Grant Program.

REFERENCES

1. **Allison, R., R. E. Johnston, and W. G. Dougherty.** 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology **154**:9–20.
2. **Ambros, V., and D. Baltimore.** 1978. Protein is linked to the 5′-end of poliovirus RNA by a phosphodiester linkage to tyrosine. J. Biol. Chem. **60**:5263–5266.
3. **Carrington, J. C., and W. G. Dougherty.** 1987. Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. J. Virol. **61**:2540–2548.
4. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
5. **Domier, L. L., K. M. Franklin, A. G. Hunt, R. E. Rhoads, and J. G. Shaw.** 1989. Infectious *in vitro* transcripts from cloned cDNA of a potyvirus, tobacco vein mottling virus. Proc. Natl. Acad. Sci. USA **86**:3509–3513.
6. **Domier, L. L., K. M. Franklin, M. Shahabuddin, G. M. Hellmann, J. H. Overmeyer, S. T. Hiremath, M. F. E. Siaw, G. P. Lomonossoff, J. G. Shaw, and R. E. Rhoads.** 1986. The nucleotide sequence of tobacco vein mottling virus RNA. Nucleic Acids Res. **14**:5417–5430.
7. **Domier, L. L., J. G. Shaw, and R. E. Rhoads.** 1987. Polyviral proteins share amino acid sequence homology with picorna-, como-, and caulimoviral proteins. Virology **158**:20–27.
8. **Dougherty, W. G., and J. C. Carrington.** 1988. Expression and function of potyviral gene products. Annu. Rev. Phytopathol. **26**:123–143.
9. **Goldbach, R., and A. van Kammen.** 1985. Structure, replication and expression of the bipartite genome of cowpea mosaic virus, p. 83–120. *In* J. W. Davies (ed.), Molecular plant virology, vol. 2. CRC Press, Inc., Boca Raton, Fla.
10. **Hari, V.** 1981. The RNA of tobacco etch virus: further characterization and detection of protein linked to RNA. Virology **112**:391–399.
11. **Harris, T. J. R., J. J. Dunn, and E. Wimmer.** 1978. Identification of specific fragments containing the 5′ end of poliovirus RNA after ribonuclease III digestion. Nucleic Acids Res. **5**:4039–4054.
12. **Jaegle, M., J. Wellink, and R. Goldbach.** 1987. The genome-linked protein of cowpea mosaic virus is bound to the 5′ terminus of virus RNA by a phosphodiester linkage to serine. J. Gen. Virol. **68**:627–632.
13. **Kuhn, R. J., H. Tada, M. F. Ypma-Wong, J. J. Dunn, B. L. Semler, and E. Wimmer.** 1988. Construction of a "mutagenesis cartridge" for poliovirus genome-linked viral protein: isolation and characterization of viable and nonviable mutants. Proc. Natl. Acad. Sci. USA **85**:519–523.
14. **Lain, S., J. L. Riechmann, and J. A. Garcia.** 1989. The complete nucleotide sequence of plum pox potyvirus RNA. Virus Res. **13**:157–172.
15. **Luciano, C. S., R. E. Rhoads, and J. G. Shaw.** 1987. Synthesis of potyviral RNA and proteins in tobacco mesophyll protoplasts inoculated by electroporation. Plant Sci. **51**:295–303.
16. **Murphy, J. F., U. Jarlfors, and J. G. Shaw.** Phytopathology, in press.
17. **Murphy, J. F., R. E. Rhoads, A. G. Hunt, and J. G. Shaw.** 1990. The VPg of tobacco etch virus RNA is the 49 kDa proteinase or the N-terminal 24 kDa part of the proteinase. Virology **178**:285–288.
18. **Nicklin, M. J. H., H. Toyoda, M. G. Murray, and E. Wimmer.** 1986. Proteolytic processing in the replication of polio and related viruses. Biotechnology **4**:33–42.
19. **Reddick, B. B., and O. W. Barnett.** 1983. A comparison of three potyviruses by direct hybridization analysis. Phytopathology **73**:1506–1510.
20. **Riechmann, J. L., S. Lain, and J. A. Garcia.** 1989. The genome-linked protein and 5′ end RNA sequence of plum pox potyvirus. J. Gen. Virol. **70**:2785–2789.
21. **Robaglia, C., M. Durand-Tardif, M. Tronchet, G. Boudazin, S. Aster-Manifacier, and F. Casse-Delbart.** 1989. Nucleotide sequence of potato virus Y (N strain) genomic RNA. J. Gen. Virol. **70**:935–947.
22. **Rothberg, P. G., T. J. R. Harris, A. Nomoto, and E. Wimmer.** 1978. O^4-(5′-uridylyl)tyrosine is the bond between the genome-linked protein and the RNA of poliovirus. Proc. Natl. Acad. Sci. USA **75**:4868–4872.
23. **Rychlik, W., P. R. Gardner, T. C. Vanaman, and R. E. Rhoads.** 1986. Structural analysis of the messenger RNA cap-binding protein. J. Biol. Chem. **261**:71–75.
24. **Shahabuddin, M., J. G. Shaw, and R. E. Rhoads.** 1988. Mapping the tobacco vein mottling virus VPg cistron. Virology **163**:635–637.
25. **Siaw, M. F. E., M. Shahabuddin, S. Ballard, J. G. Shaw, and R. E. Rhoads.** 1985. Identification of a protein covalently linked to the 5′ terminus of tobacco vein mottling virus RNA. Virology **142**:134–143.
26. **Takebe, I., Y. Otsuki, and S. Aoki.** 1968. Isolation of tobacco mesophyll cells in intact and active state. Plant Cell Physiol. **9**:115–124.
27. **Takegami, T., R. J. Kuhn, C. W. Anderson, and E. Wimmer.** 1983. Membrane-dependent uridylylation of the genome-linked protein VPg of poliovirus. Proc. Natl. Acad. Sci. USA **80**:7447–7451.
28. **Tobin, G. J., D. C. Young, and J. B. Flanegan.** 1989. Self-catalyzed linkage of poliovirus terminal protein VPg to poliovirus RNA. Cell **59**:511–519.