agronomie (1995) 15, 447-452

Sequence of the 3'-part of the RNA of ryegrass mosaic virus, a potyvirus

J Schubert [1]*, F Rabenstein [2], E Proll [2]

[1] *Federal Centre of Breeding Research on Cultivated Plants, Institute for Resistance Research;*
[2] *Federal Centre of Breeding Research on Cultivated Plants, Institute for Pathogen Diagnostics, PO Box 1505, D-06435 Aschersleben, Germany*

(Received 20 May 1995; accepted 4 July 1995)

Summary — The sequence of the first 2 772 nucleotides upstream from the 3'-end was identified a Danish isolate of ryegrass mosaic (RgMV) considered till now to belong to *Rymovirus* genus. Comparisons of the deduced amino acid sequences for the coat protein (cp) and nuclear inclusion protein (NIb) with those of other potyviruses revealed a high degree of homology with viruses of the genus *Potyvirus* and only weak homologies with another rymovirus, wheat streak mosaic potyvirus. The putative proteolytic cleavage site NIb/cp exhibits a previously undescribed amino acid motif. The classification of RgMV should be reviewed.

ryegrass mosaic virus / potyvirus / sequence / coat protein

Résumé — Séquence de la partie 3' du RNA du virus de la mosaïque du ray-grass, un potyvirus. *Nous avons déterminé la séquence de 2 772 nucléotides de la portion terminale 3' d'un isolat danois du virus de la mosaïque du ray-grass (RgMV) considéré jusque là comme appartenant aux rymovirus. Les comparaisons des séquences des acides aminés de la protéine capside (cp) et de la protéine d'inclusion nucléaire (NIb) avec des potyvirus révèlent des homologies fortes avec les virus de genre* Potyvirus *et seulement une faible homologie avec le virus de la mosaïque striée du blé, un autre rymovirus. Le site putatif de clivage protéolytique de la NIb/cp montre la présence de motifs acides aminés non décrits jusque là. La classification du RgMV doit être modifiée.*

***virus de la mosaïque du ray-grass / potyvirus / séquence* / protein capside**

INTRODUCTION

Ryegrass mosaic potyvirus (RgMV) is an important viral pathogen of several species of the Gramineae. The virus was found to occur naturally in Italian ryegrass *(Lolium multiflorum)*, perennial ryegrass *(L perenne)* and cocksfoot *(Dactylis glomerata)* (Slykhuis and Paliwal, 1972). On the basis of its transmission, it is grouped together with wheat streak mosaic virus (WSMV), agropyron mosaic virus (AgMV), hordeum mosaic virus (HorMV), and oat necrotic mottle virus (ONMV) into the newly proposed genus *Rymovirus* within the family Potyviridae (Zagula, 1992). Brome streak mosaic virus (BrSMV) is a putative member of the genus. The eriophyid gall mites *Abacarus hystrix* and *Aceria*

* Correspondence and reprints

tulipae act as vectors. The virus particles are flexible rods, approximately 705 nm long and contain a single-stranded RNA (2.7 x 10^6 Da). The molecular weight of the coat protein (cp) of this virus is 29.2 kDa (Paliwal and Tremaine, 1976; Salm *et al,* 1993).

MATERIALS AND METHODS

Virus isolate, purification of the virus and isolation of RNA

For cloning, we used a Danish isolate of RgMV (RgMV-DK) obtained from Dr Albrechtsen (Lyngby, Denmark). In DAS-ELISA its serological reactions were identical to isolates from Australia, Bulgaria, The Czech Republic, Germany, Great Britain, and The Netherlands (results not shown). The virus was propagated on mechanically infected *L multiflorum* plants. Virus purification and RNA extraction followed a method published by Schubert and Rabenstein (1995). The purity of the RNA was checked electrophoretically. It was transferred to positively charged nylon membranes (Hybond N^+, Amersham) by capillary blot.

cDNA synthesis and cloning

cDNA was synthesised using 5 µg RNA and oligo dT_{15} as a primer according to Gubler and Hoffman (1989). For the first strand synthesis we applied AMV reverse transcriptase (Amersham). The ends of the ds cDNA were blunted with T_4-DNA-polymerase (USB). The success of the reactions was monitored by the addition of ^{32}P-dATP (Amersham; Sambrook *et al,* 1989). The ds cDNA was fractionated on Sephadex G50. The larger fragments eluted from the column first and were precipitated with ethanol and ligated into pUC 18/Sma I (Pharmacia). The plasmid was transferred into highly competent cells of *Escherichia coli* XL2-blue (Stratagene). Positive, white clones were proofed on nitrocellulose filters with ^{32}P-dATP labelled cDNA after alkali lysis of bacteria (Sambrook *et al,* 1989).

Sequencing

Sequencing was performed with plamids as templates by means of cycle (Vent polymerase, Biolabs) or conventional dideoxy-sequencing (Sequenase, USB) using ^{35}S-dATP (Amersham) for labelling.

Serology and protein expression

For size estimation of the cp infected plant material was homogenized with water (w/v = 1:2). The homogenate was immediately centrifuged, or after 6 h, and the supernatant mixed 3:1 with 4^x loading buffer and boiled for 1 min; 5 µl were used for electrophoresis. For *in vitro* expression of the cp the cDNA of clone Rg 12 (1737 bp) was recloned in the expression vector pGEX4-T1 (Pharmacia). It was in frame with the sequence of glutathione *S*-transferase (GST) of the vector. For induction of the protein synthesis we added IPTG to 0.1 mM to freshly grown culture of transformed *E coli* JM 109 (in LB with 2% glucose). Three hours after induction, cells were collected by centrifugation and disrupted with an ultrasonifier (Hilscher). The fusion protein was purified by glutathione *S*-sepharose (Pharmacia) according to the manufacturer's instructions. For electrophoresis 10 µl of the fractions were generally used. After electrophoresis in 10% SDS-polyacrylamide gels (Sambrook *et al,* 1989), the recombinant as well as natural cp were blotted semi-dry on nitrocellulose (Bio-Rad). The cp of RgMV were detected on the Western blots with a polyclonal antiserum or monoclonal antibodies against native virus particles (Rabenstein, unpublished results) and alkaline phosphatase conjugated goat-anti rabbit/mouse IgG. As substrate we used BCIP/NBT.

RESULTS

Western blots of the viral cp for all of the 10 investigated RgMV isolates gave sizes of approximately 46 kDa. In addition different bands of degradation products could be identified (figs 1 and 2).

Five positively reacting RgMV clones were chosen for sequencing. All hybridized with the blotted viral RNA. The largest of them, Rg 5, contains an insertion of 2 772 nt excluding the polyA tract. Non-nucleic acid (na) sequence differences were observed in the overlapping regions of the investigated clones. The deduced amino acid (aa) sequence of Rg 5 is shown in figure 3. It contains one open reading frame. The non-translated region amounts to 165 nt. The cloned sequence of Rg 12 (1 727 nt, fig 3), containing the cp gene and a part of Nlb, was successfully expressed in *E coli* (fig 2). The fusion protein showed the expected size of approximately 80 kDa and reacted in Western blots with the corresponding antisera demonstrating that the correct virus was cloned. Compared with similar constructs of other potyviruses (Schubert, unpublished data) the expression level was very low.

DISCUSSION

One large open reading frame, which is typical for potyviruses, was detected for the clone Rg 5.

The na and the deduced aa sequences showed homologies to those of other potyviruses especially of the genus *Potyvirus* (fig 3). Using the CLUSTAL 4 option of the DNASIS software (Hitachi), the 3 mite-transmitted Potyviridae WSMV (Niblett *et al,* 1991), BrSMV (Schubert and Rabenstein, 1995) and RgMV as well as the

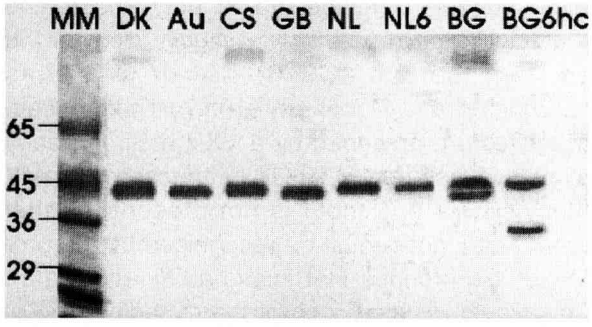

Fig 1. Estimation of molecular weights of cp of different RgMV isolates from *L multiflorum* by Western blot (12% PAAG). For detection a polyclonal antiserum was used. Aliquots of the homogenates of the BG and NL isolates were further processed after 6 h to study the action of plant proteases. On the left side, the Coomassie-stained gel with molecular weight markers (MW in kDa). MM–MW marker mixture; designation of isolates: DK: Danish; Au: Australian; CS: Czech; GB: British; NL: Dutch; BG: Bulgarian; NL6 and BG6: protease exposed probes; hc: healthy control.

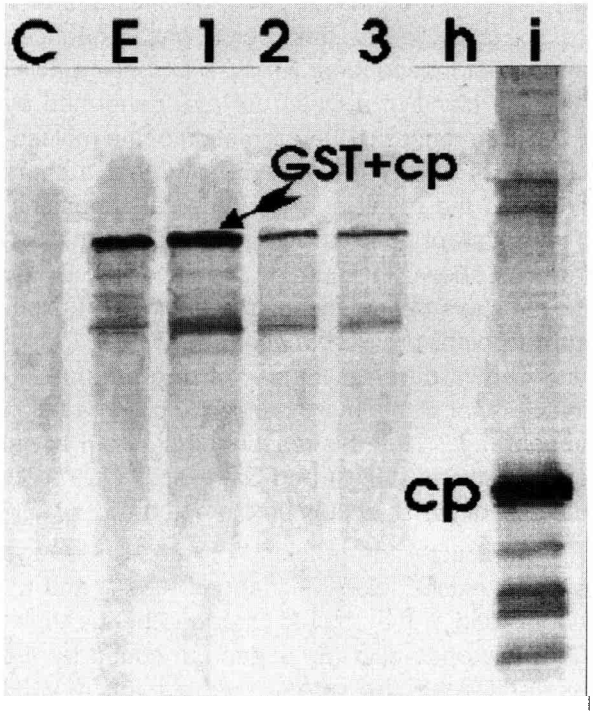

Fig 2. Western blot of recombinant and natural RgMV cp with polyclonal antiserum. C-negative control (proteins of *E coli*/pGEX4-T), E-proteins from lysed *E coli*/pGEX4-T with insertion. 1–3 - fractions from gluthatione *S*-sepharose column with recombinant fusion proteins, proteins of (h) healthy and (i) RgMV-DK-infected plants of *A sativa* L. GST+cp: band of undegraded fusion protein; cp: band of undegraded RgMV-DK cp (molecular weight 46 kDa).

Fig 3. Comparison of the amino acid sequences of WSMV, RgMV and PVY. Common aa are marked as follows: RgMV/PVY – bold letters; WSMV/RgMV – underlined letters; for all 3 viruses – bold italic letters; asterisks mark conserved aa of the genus *Potyvirus*, lower case letters show the conserved aa, if it is missing for RgMV; ⋆ motifs, thought to be important for formation of virus particles (Jagadish *et al,* 1991); ➡..◆ region of na homology between RgMV cDNA and gliadin DNA; ➡Rg 12: sequence of clone Rg 12, which was expressed in *E coli.*

aphid transmitted PVY (Robaglia *et al*, 1989) were grouped on the basis of the na sequences of their cp (fig 4). It is clearly visible that both definitive members of the genus *Rymovirus*, WSMV and RgMV, do not group together. Though being the type member of this genus, RgMV is more closely related to PVY, the type member of the genus *Potyvirus*, than to WSMV.

Shukla and Ward (1988) defined criteria for the classification of viruses on the basis of homologies of aa of the viral cp. This criterion has only a limited value for RgMV, as it is much larger than most known potyviruses of the genus *Potyvirus*. The problem can be avoided if only the homologous regions of the proteins are taken into account for comparison. These results are given in table I and figure 5. It is clearly visible that both the cp and the NIb protein of RgMV have limited homologies with those of WSMV but a high degree of homology with those of PVY. Even most of the conserved aa of viruses of the genus *Potyvirus* appear in the genome of RgMV but not in that of WSMV.

The sizes and structures of the Cp of RgMV, WSMV and BrSMV are similar. It seems that during evolution the C-terminal parts of the polymerases were integrated into the cp of all 3 viruses. This assumption is supported by the sequence homologies of the corresponding regions (fig 3). Furthermore, some additional aa appear in their cp (fig 3). For RgMV possibly, 60 aa were integrated in the *N*-terminal part of the core region of cp, 61 additional aa seem to be integrated in the C-terminal part of the cp of WSMV (depending on the aa alignment).

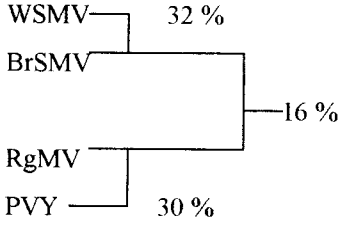

Fig 4. CLUSTAL-analysis for the na sequence relationships of 3 rymoviruses and PVY.

Table I. Degree of amino acid homology (%) between NIb and cp of RgMV, PVY and WSMV.

	NIb	*cp*
RgMV/PVY	50/70[a]	48
RgMV/WSMV	36	13
PVY/WSMV	37	10

[a] For full size of NIb and /size corresponding to the (shorter) published part of WSMV sequence. As standard for alignment, the sizes of the RgMV proteins were taken.

The protease cleavage site between NIb and cp of RgMV is different from all other described corresponding cleavage sites of potyviruses. The proteolytic active part of NIa usually cleaves the cp from the NIb at Q/S, Q/A or Q/G sites (Carrington and Dougherty, 1988). These cleavage sites are preceded by a VXH motif in nearly all potyviruses (Maiß, 1993). In the case of RgMV the VXHQ/S,A,G motif is not present, so that other sequences must be recognised by the protease. We propose that this is a Q/L site (fig 3). This would result in a cp of the expected size of 45.4 kDa. Nevertheless it cannot be excluded that the Q/F site, located 2 aa upstream, is recognised resulting in a cp of 45.7 kDa. For the corresponding cleavage sites of tobacco vein mottling and tobacco etch viruses a phenylalanin (F) was identified, preceding the glutamine (Q). An F precedes the putative Q/L cleavage site of RgMV also supporting the assumption that Q/L is the active site. A mutation which changed the aa S, A or G of the putative cleavage sites to L or F can be excluded as this region was sequenced for 2 other independent RgMV-DK clones as well as for a clone of a Dutch isolate, generated by PCR. Differences in the estimation of the molecular weight of the cp by PAGE by other authors could be the result of environmental conditions, different plant species and influences of the isolate. As shown in figure 1, the cp of the Dutch isolate stays nearly unchanged after 6 h of exposure to plant proteases while the Bulgarian isolate shows different bands of degradation. In infected oat plants more bands of cp degradation appear (fig 2). For this reason the data of Paliwal and Tremaine (1976) and Salm *et al* (1993) for the size of the cp should be corrected.

Interestingly the 'ancient' protease cleavage site still exists nearly unchanged (figs 3 and 6), but instead of the original Q, coded by the triplet CAA, glutamic acid (E) appeared, coded by the triplet GAA. The conserved aa motif VXH upstream from the cleavage site is still present; 21 aa downstream for the 'ancient' cleavage site the above-mentioned additional aa appeared to be integrated. The hypothetical integration and mutation possibly led to the appearance of the altered structure of the cp. The results demonstrate that the C-terminal part of the NIb is not

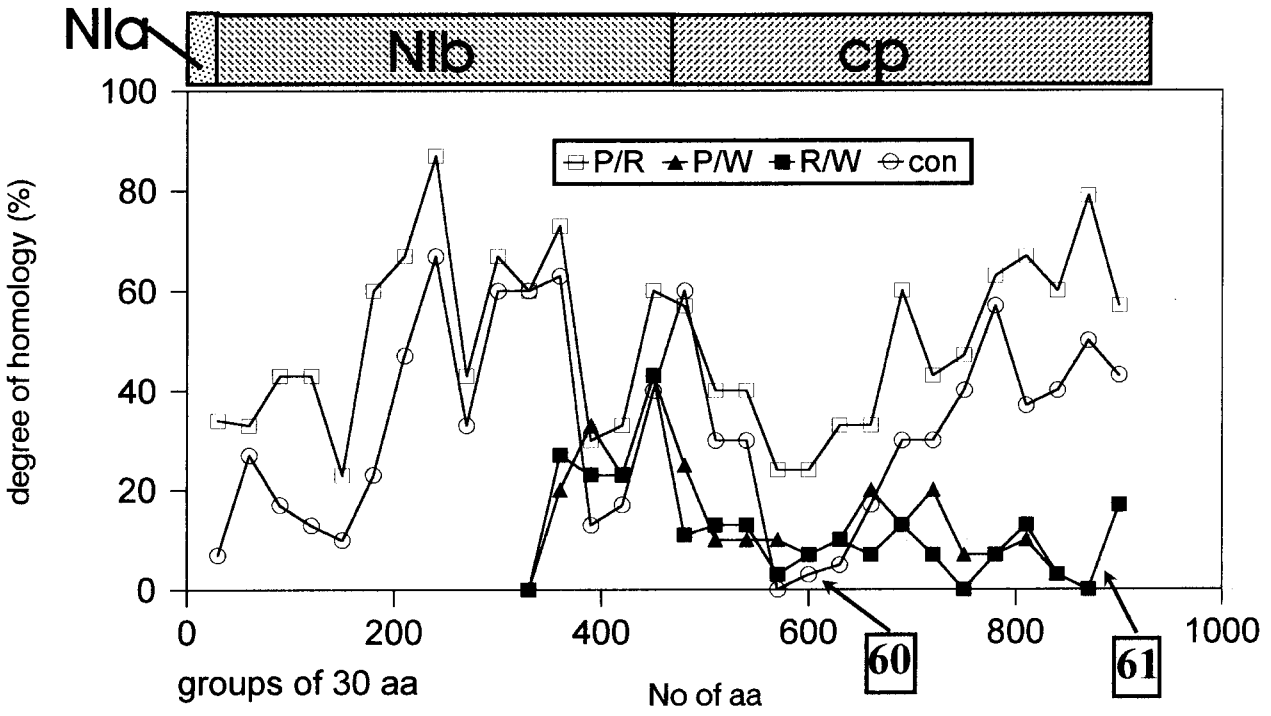

Fig 5. Degree of aa homologies in the different regions of the viral polyprotein. W – WSMV; P – PVY; R – RgMV; con: conserved aa of aphid transmitted potyviruses (data for conserved aa from Maiß, 1993). For calculation the hypothetically inserted regions of RgMV (60 aa) and WSMV (61 aa), having no corresponding regions in the other compared viruses, were omitted.

```
K   V   V   H   E(/)A   N   (aa)
AAA GTG GTG CAC GAA GCT AAC (na)
                ↓
                CAA = Q
```

Fig 6. Possible mutation in the RgMV genome leading to a larger size of the coat protein. (/)- 'lost' cleavage site; V, H: conserved amino acids (motif V x H) ↓ possible mutation; (aa) amino acid sequence; (na) nucleic acid sequence.

necessary for the enzymatic activity of this RNA-dependent RNA polymerase.

The cp of RgMV is lacking the DAG motif thought to be involved in aphid transmissibility (Atreya *et al,* 1991). A sequence responsible for mite transmissibility cannot be proposed. It was interesting to find out where the integrated sequence of 60 aa of RgMV could originate from. Even for plum pox virus (PPV), pea seed-borne mosaic, turnip mosaic and papaya ringspot viruses such additional aa can be observed for the corresponding regions of the cp (Maiß, 1993). Comparison of the na sequence of RgMV with sequence data of the EMBL library revealed some homology of a part of the insertion with a part of a gliadin sequence (tagliaba). The same was true for a part of the corresponding sequence of PPV. The homologous regions of the sequences of gliadin and RgMV are shown in figure 7. The multiple CAA motif is typical for a plant genome microsatellite. It would be interesting to test if gliadin or similar genes of other plants have an even higher degree of homology with this sequence. This could be an evidence that parts of the host genome can be integrated in the viral genome. $(CAA)_n$ motifs were recorded for the Ω' region of TMV and the 5' NTR of potyviruses too. They are thought to be responsible for translation enhancement (Tordo *et al,* 1995).

```
R  TCTGCAACAACGACACCATCAACAGGAACCACTGCGCCAACAACAC
G  TCTGCATCAACAACACCACCACCACCAACAACAACAACAACAACAA
   ****** **** ****** ** **  *** **  **********

R  CAACAGTGAGAGCCGCAAACTTGGACATCGCAGGACACAGAA
G  CAACAACAACCGTTGAGCCAAGTGTCCTTCCAACAGCCTCAG
   ******  *  *  *          * *  **  *  *  *
```

Fig 7. Homology (na) between the inserted regions of the RgMV cp and gliadin. R: RgMV; G: gliadin; asterisk: common nucleotide; bold: CAA-microsatellite.

Summarizing our results about the characteristics of RgMV we propose to review the grouping of WSMV and RgMV into the genus *Rymovirus.* They should be grouped into different genera.

ACKNOWLEDGMENTS

We wish to thank M Nielitz, M Wesemann, H Mühlheim and E Zimmermann for excellent technical assistance.

REFERENCES

Atreya PL, Atreya CD, Pirone TD (1991) Amino acid substitutions in the coat protein results in a loss of insect transmissibility of a plant virus. *Proc Natl Acad Sci USA* 88, 7887-7891

Carrington JC, Dougherty WG (1988) A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc Natl Acad Sci USA* 85, 3391-3395

Jagadish MN, Ward CW, Gough KH, Tulloch PA, Whittaker LA, Shukla DD (1991) Expression of potyvirus coat protein in *Escherichia coli* and yeast and its assembly into virus-like particles. *J Gen Virol* 72, 1543-1550

Maiß E (1993) Das Scharka-Virus der Pflaume (PPV). Habilitationsschrift. Universität Hannover, Germany

Milne RG (1988) Taxonomy of the rod-shaped filamentous viruses. *In: The Plant Viruses. Vol 4. The Filamentous Plant Viruses* (RG Milne, ed), Plenum press, NY, USA, 3-50

Niblett CL, Zagula KR, Calvert LA *et al* (1991) cDNA cloning and nucleotide sequence of the wheat streak mosaic virus capsid protein gene. *J Gen Virol* 72, 499-504

Paliwal YC, Tremaine JH (1976) Multiplication, purification, and properties of ryegrass mosaic virus. *Phytopathology* 66, 406-414

Robaglia C, Durand-Tarif M, Trouchet, M, Boudazin G, Astier-Manifacier S, Casse-Delbart F (1989) Nucleic acid sequence of potato virus Y (N strain) genomic RNA. *J Gen Virol* 70, 935-947

Salm SN, Rey MEC, Wolfson MMW (1993) A South African isolate of ryegrass mosaic virus. *Plant Pathol* 43, 708-712

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn, Cold Spring Harbour Laboratory, NY, USA, 743-745

Schubert J, Rabenstein F (1995) Sequence of the 3'-terminal region of the RNA of a mite transmitted potyvirus from *Hordeum murinum* L. *Eur J Plant Pathol* 101, 123-132

Shukla DD, Ward CW (1988) Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *J Gen Virol* 69, 2703-2710

Slykhuis JT, Paliwal YC (1972) Ryegrass mosaic virus. CMI/AAB Descript Plant Virus. No 86 Commonw Mycol Inst, Assoc Appl Biol, Kew, UK, 4 p

Tordo VM-J, Chachulska AM, Fakhafakh H, Le Romancer M, Robaglia C, Astier-Manifacier S (1995) Sequence polymorphism in the 5'NTR and in the P1 coding region of potato virus Y genomic RNA. *J Gen Virol* 76, 939-949

Zagula KR, Niblett CI, Robertson NL, French R, Lommel SA (1992) Potyviridae: genus *Rymovirus*. *Arch Virol* Suppl 5, 269-276