Genome organization of grapevine fanleaf nepovirus RNA2 deduced from the 122K polyprotein P2 in vitro cleavage products

Rogério Margis,1,2 Christophe Ritzenthaler,1 Joseph Reinbolt,3 Monique Pinck1 and Lothaire Pinck1*

1 Institut de Biologie Moléculaire des Plantes du C.N.R.S. et Université Louis Pasteur, Laboratoire de Virologie, 12 rue du Général Zimmer, 67084 Strasbourg, France, 2 Escola Técnica Federal de Química do Rio de Janeiro, Rua Senador Furtado No. 121, 20270 Rio de Janeiro, Brazil and 3 Institut de Biologie Moléculaire et Cellulaire du C.N.R.S., U.P.R. 9002, 15 rue Descartes, 67084 Strasbourg, France

The full-length transcript of grapevine fanleaf virus (GFLV) RNA2 produces a primary product of 122K when translated in the rabbit reticulocyte system. This 122K polyprotein is completely processed in vitro by the RNA1-encoded 24K proteinase. The positions of the cleavage sites within the polyprotein have been mapped and the genome organization of GFLV-F13 RNA2 has been established. The order of mature proteins in the 122K polyprotein is the amino-terminal 28K protein, the 38K protein followed by the 56K coat protein at the carboxy terminus. These proteins represent the final cleavage products of the 122K polyprotein. A 66K protein which yields 28K and 38K proteins constitutes the major maturation intermediate. Microsequencing of the amino extremity of radioactively labelled 38K protein allowed identification of the Cys257/Ala258 site as the cleavage site recognized by the GFLV proteinase in addition to the Arg605/Gly606 site between the 28K and the 38K proteins in the 66K protein in addition to the Arg605/Gly606 site between the 38K protein and the coat protein.

Introduction

Grapevine fanleaf virus (GFLV) belongs to the plant nepovirus group. Members of this group are characterized by nematode transmissibility, polyhedral viral particles, a plus-sense bipartite RNA genome, the presence of a small covalently genome-linked protein (VPg) at the 5′ extremity and a poly(A) sequence at the 3′ OH extremity of virus RNAs (Martelli, 1991). The complete nucleotide sequence of RNA1 of the F13 isolate of GFLV, 7342 nucleotides (nt), contains a major open reading frame (ORF) encoding a product of Mr 252K corresponding to the polyprotein P1 in which conserved regions characteristic of NTP-binding protein, viral proteinase and RNA polymerase have been identified (Ritzenthaler et al., 1991). The primary sequence of the VPg was determined by microsequencing and was located on the polyprotein P1 sequence (Pinck et al., 1991). The nucleotide sequence of GFLV RNA2 (3774 nt) codes for a polyprotein P2 of Mr 122K which contains the coat protein (CP) in its carboxy-terminal part. The CP of Mr 55-6K is produced by a proteolytic cleavage at the Arg605/Gly606 site of P2 as deduced from the amino-terminal sequence of the CP (Sergolini et al., 1990). Biologically active full-length transcripts of RNA2 (clone pACN) and RNA1 (clone pMV) have been obtained (Viry et al., 1993).

A cDNA clone (pVP7) encoding the proteinase activity of GFLV has been previously described (Margis et al., 1991). Its insert contains the sequence of the proteinase flanked by the entire VPg sequence (24 residues) at its amino extremity and 57 amino acids corresponding to the N-terminal portion of the viral RNA polymerase at its carboxy extremity. This protein displays a proteinase activity which cleaves autocatalytically (by an in cis mechanism) between the proteinase itself and the polymerase amino extremity, but also processes the polyprotein P2 in trans to produce the CP and a 66K protein (Margis et al., 1991).

Polyprotein synthesis is the strategy of genome expression used by nepoviruses as well as by different groups of animal and plant positive-strand RNA viruses (Hellen et al., 1989). Several reports have shown that nepovirus polyproteins are processed by a proteinase activity present in polyprotein P1 (Morris-Krsinich et al., 1983; Brault et al., 1989; Demangeat et al., 1990, 1991; Margis et al., 1991). Morris-Krsinich et al. (1983) reported that GFLV RNA1 translation products synthesized in a rabbit reticulocyte lysate system can cleave polyprotein P2 into two major proteins, a 68K protein and the CP. The polyproteins encoded by RNA2 of tobacco ringspot virus, tomato black ring virus and grapevine chrome mosaic virus, also members of the nepovirus group, have been shown to be processed into
three final cleavage products (Forster & Morris-Krinsinich, 1985; Demangeat et al., 1991).

In this paper we provide evidence that the 66K protein is an intermediate in GFLV polypolyprotein P2 processing. The site of the proteolytic cleavage in the 66K protein was identified by microsequencing of the labelled cleavage product. Different intermediate polypeptides occurring during the processing of P2 polypolyprotein have been identified by analysis of the proteins produced from transcripts of mutagenized RNA2 in reticulocyte lysates. We present the genome organization of GFLV RNA2 deduced from these data.

**Methods**

**Virus, bacteria and plasmids.** GFLV strain F13 was routinely multiplied in *Chenopodium quinoa* and purified as previously described (Pinck et al., 1988). *Escherichia coli* strain NM22 was used to amplify plasmid constructs. Competent cells were transformed by electrophoration (Pfu & Youderian, 1990) or by chemical methods (Chung & Miller, 1988). The constructs were made from the phagemids BlueScript m13 plus (BS') (Stratagene) or BlueScript II KS' (KS') (Stratagene). Plasmid pVP7 was used to produce, by *in vitro* translation and autacotytic cleavage, an active proteinase, named Pro, which includes Vpg and proteinase sequences (Margis et al., 1991). Clone pACN was used to produce full-length transcripts of RNA2 (Viry et al., 1993).

**PCR mutagenesis and cloning.** Plasmid N1, a cDNA clone of GFLV RNA2 from nt 198 to 3754 (Viry et al., 1993), was used as template for PCR mutagenesis of the 66K protein-coding sequence. PCR conditions for production of clones p38, p38RV, p28 and p94 were as follows. The reaction mixture contained 1 µL of double-stranded full-length cDNA of RNA2, 50 pmol of each primer, 2 µL of 10 mM-dNTPs, 10 µL of 10 × PCR buffer (100 mM-Tris-HCl pH 8.3, 0.1% gelatin, 15 mM-MgCl2, 500 mM-KCl) in a final volume of 100 µL. After 5 min boiling, 5 units of Taq DNA polymerase (Beckman) was added and 12 cycles of PCR were performed in a programmable 480 thermal cycler (Perkin-Elmer Cetus). The first three cycles were performed with a denaturation step (94 °C, 1 min), hybridization step (42 °C, 1 min) and elongation step (72 °C, 1 to 3 min, depending on the size of the fragment to be amplified) followed by eight cycles under the same conditions except that the hybridization temperature was raised to 60 °C. The PCR product was precipitated and digested with EcoRV for 2 h. After agarose gel electrophoresis, the band corresponding to the digested PCR product was purified and ligated to *EcoRV*-cut 5′ dephosphorylated KS′ plasmid. The clones were analysed for their size and integrity were checked by electrophoresis under denaturing conditions (1% agarose-formaldehyde gel (Gustafson et al., 1982). Transcripts at concentrations ranging from 0.25 to 1 µg/µL, estimated by their A260, were translated in a nuclease-treated rabbit reticulocyte lysate system (NT-RRLS) obtained from Promega. Translations in the reticulocyte system were performed with an incubation mixture composed of 35 µL of NT-RRLS (Promega), 1 µL of a 1 mm-amino acid mix (isoleucine-, leucine-, lysine- or methionine-free) and either 4 µL of [35S]methionine (15 µCi/µL, 1000 Ci/mmol, Amersham), 5 µL of l-[4,5-3H]lysine (5 µCi/µL, 120 Ci/mmol, Amersham), l-[4,5-3H]leucine (5 µCi/µL, 120 to 190 Ci/mmol, Amersham) or l-[4,5-3H]isoleucine (5 µCi/µL, 85 Ci/mmol, Amersham). Four µL of the above translation mixture was added to 0.5 to 1 µL of transcripts (i.e. 0.12 to 1 µg) and the final volume was adjusted to 5 µL with sterile water. Translations were performed at 30 °C for different times. The samples were then incubated for 10 min at 30 °C with 100 ng/µL RNase A (final concentration) before their use in maturation studies. Translations and maturation reactions were stopped by addition of an equivalent volume of electrophoresis loading buffer (10% SDS, 25% 2-mercaptoethanol, 160 mM-Tris-HCl pH 6.8) (Demangeat et al., 1990). Translation with non-radioactive amino acids was performed in the presence of 10 µL of 1 mM-methionine instead of labelled methionine. Samples were analysed on polyacrylamide-SDS gels according to the procedure of Laemmli (1970). Autoradiograms were scanned at 560 nm with a Shimadzu CS9000 densitometer (Shimadzu Corporation).

**Protein purification and microsequencing.** Transcript of clone p66 linearized by EcoRI (166) and an equivalent amount of the pVP7-Pmol linearized transcript were translated in 400 µL NT-RRLS incubation mixture in the presence of [35S]methionine or [3H]lysine for 5 h at 30 °C to allow maximal cleavage to occur. The reaction was stopped by addition of an equivalent volume of electrophoresis loading buffer. The sample was stored at −80 °C or immediately loaded in several slots of a 10% polyacrylamide-SDS gel. After electrophoresis, bands corresponding to the 38K protein were excised from the gel and the protein was electroeluted. The eluted protein was concentrated with a Centricon-10 (Amicon) and subjected to automated Edman degradation using an Applied Biosystems 470A protein sequencer (Hewick et al., 1981). Samples from each sequencing cycle were mixed with 5 µL of Ready Gel Cocktail (Beckman) and radioactivity was counted in a Beckman LS9000 scintillator.

**Results**

**Synthesis of GFLV-F13 polypolyprotein P2 in vitro**

We chose the rabbit reticulocyte system for the maturation studies since, as shown earlier (Margis et al., 1991), proteolytic processing is inhibited in the wheat germ system probably by inhibitors of viral proteinase (Shih et al., 1987). The sequence of the RNA2 transcribed from clone pACN differs from that previously reported (Sergini et al., 1990) by a uridine in position 7 instead of adenosine, a change which eliminates the reported first
Maturation of GFLV polyprotein P2

Fig. 1. Detection of the in vitro translation and maturation products of GFLV-F13 virion RNAs and the full-length RNA2 transcript. (a) In vitro translation of GFLV-F13 total RNAs (lane 1) and the full-length RNA2 transcript (tr2) (lane 2) in NT-RRLS. (b) Kinetics of maturation in trans of polyprotein P2 by the GFLV proteinase; autoradiography of the time course of maturation of [35S]methionine-labelled polyprotein P2 during 15, 45, 60, 90, 120 and 240 min (lanes 2 to 8 respectively) analysed in a 10% polyacrylamide gel containing SDS. Lanes 1 and 9 show tr2 translated for 45 and 285 min, respectively. (c) Autoradiography of the time course of maturation during 30, 60 and 120 min of [3H]leucine-labelled tr2 translated polyprotein P2 analysed in a 12.5% polyacrylamide gel containing SDS (lanes 2 to 4 respectively) [lane 1 as in (a) lane 2]. Positions of M_r markers are indicated to the right. The major translation product of tr2, the polyprotein P2 and its maturation products, the 66K, 38K, 28K proteins and the CP are indicated. P1, P2 and P3 indicate the position of polyproteins produced by RNA1, RNA2 and RNA3 respectively.

AUG initiation codon of the major ORF. The translation of the transcripts obtained from clone pACN (tr2) or from clone pN1 yields the same polyprotein P2 of M_r 122K (data not shown). These polyproteins comigrate with the P2 translated from GFLV-F3 total RNA in vitro (Fig. 1a).

Polyprotein P2 cleavage by the GFLV proteinase

Maximal polyprotein P2 synthesis was obtained after 45 min of tr2 translation in NT-RRLS (Fig. 1b, lane 1). A time course study was performed, in which [35S]-labelled tr2 translation products (45 min of translation) were incubated with an equal volume of non-labelled proteinase. Incubation times varied from 15 to 240 min (Fig. 1b, lanes 2 to 8). The major protein bands appearing during the incubation corresponded to polyprotein P2 partial or final maturation products and will be referred to as 66K, CP, 38K and 28K according to their M_r calculated from the P2 sequence and the various cleavage sites discussed below (Fig. 1b). Some of these M_r values differ from those estimated from SDS–polyacrylamide gel analysis because some of the products have abnormal mobilities when compared with the M_r markers used in the different experiments. Such abnormal mobility had already been described for other proteins (Hames, 1987). Densitometric analysis showed that 58% of polyprotein P2 is cleaved after 15 min and 94% after 240 min of incubation (Fig. 1b, lanes 2 and 8). The CP and the 66K cleavage products were assumed to correspond to the viral coat protein and the 68K protein previously described by Morris-Krsinich et al. (1983). Labelling of polyprotein P2 with [3H]leucine produced a difference in intensity of cleavage products (Fig. 1c) compared with the [35S]methionine-labelled protein pattern (Fig. 1b). The major difference observed for the 28K cleavage product can be related to the uniform distribution of leucine residues in the P2 polyprotein, in contrast to the methionine residues, which are poorly represented in the amino extremity. If RNA2 transcript (tr2) translation was continued for 240 min the polyprotein P2 band and some other bands became less intense (Fig. 1b, lanes 1 and 9). These bands probably resulted from internal initiation, premature termination and to some extent degradation. Similarly some degradation may also occur for CP since its level, maximal after 90 min incubation, decreases with longer incubation time (Fig. 1b, lanes 6 to 8).

To study the complete maturation of polyprotein P2, partial constructs which allowed the synthesis of transcripts corresponding to the region coding for the 66K protein and the CP were produced. Based upon the Arg605/Gly606 cleavage site (Serghini et al., 1990), two clones were constructed: a CP clone and a p66 clone corresponding to the P2 polyprotein upstream of the Arg/Gly site (residues 1 to 605).
In vitro expression of the coat protein and the 66K protein

Transcripts trCPM of the CP-coding clone pCPM, translated in the wheat germ system, produced the 56K protein which comigrates with the purified GFLV-F13 CP and two shorter proteins of 52K and 50.5K produced by internal initiation of translation (Serghini et al., 1991). In the NT-RRLS system, the trCPM yields the same three proteins (Fig. 2a, lane 2) and the tr2 produced the CP and the 66K protein from polyprotein P2 maturation (Fig. 2a, lane 1). The CP is not further processed when given longer incubation with viral proteinase (data not shown).

The 66K protein was translated from transcripts of a clone in which the Gly^{608} codon of polyprotein P2 was replaced by a TAG stop codon. The mutagenesis was performed by PCR amplification with clone pN1 as template, pBS {sup} {sup} reverse primer (5' AACAGCTATGAC- CATG) and primer P1656 (5' CCTCTACCAAGTCTTCTAT- ATCCAAAGCTTTCTATCCACAGTTAGCTC) which introduced a stop codon (underlined) and a HindIII cloning site (italics). The resulting PCR amplification product of 1863 bp, digested by HindIII and cloned in the HindIII site of the BS{sub}5, yielded p66, corresponding to polyprotein P2 but with the CP-coding region deleted. The clone p66 was linearized by AatII digestion and the tr66 transcript was produced by in vitro transcription with T3 RNA polymerase. Another partial transcript (tr2-PstI), corresponding to the 66K coding region plus 53 amino acids of the CP amino terminus, was produced after digestion of pACN by PstI (Fig. 2c).

Translation of transcript tr66 in the NT-RRLS system produced a major 66K protein (Fig. 2b, lane 6) which comigrated with the 66K cleavage product obtained after the proteinase processing of polyprotein P2 (Fig. 2b, lane 1). In the presence of proteinase, the tr66 translation product (66K) was processed to two proteins which comigrated with the 38K and the 28K cleavage products of the polyprotein P2 (Fig. 2b, lanes 5 and 1). The low amount of 28K protein in our maturation assays is due to its instability in the reticulocyte lysate as indicated by the following time course experiment. If the transcript encoding the 28K protein is translated in the NT-RRLS, a continuous degradation of the 28K protein is detected, resulting in its almost complete disappearance after a 4 h incubation under the conditions used for maturation reported here (data not shown). Measurement of radioactivity in the 28K protein band after 1 h and 2 h of incubation indicated a decrease of 75%. The degradation of the 28K protein and to a lesser extent of CP, as previously mentioned, may result from the presence of endogenous proteinases in the rabbit reticulocyte lysate system (Mumford et al., 1981; Oberst et al., 1993).
In vitro translation of the GFLV tr2-PstI produced a protein (66') with an \( M_r \) greater than that of the 66K protein (Fig. 2b, lane 4), as expected from the 53 additional amino acids of the CP cistron present at its carboxy terminus. If proteinase is added to tr2-PstI translation product, four new bands corresponding to the species indicated as 66K, 38' (a protein with \( M_r \) greater than 38K), 38K and 28K were generated (Fig. 2b, lane 3). The presence of a 66K and a 38K protein indicated that the proteinase was able to recognize and cleave at the Arg/6°Gly/6° site present in the 66K protein and release the 53 residues of the CP amino terminus (Fig. 2c). The presence of a 28K protein among the polyprotein P2 and the 66K protein maturation products (Fig. 2b, lanes 1 and 3) indicates that a cleavage occurs within the 66K protein. Protein 38', assumed to correspond to the 38K protein plus 53 amino acids from the CP amino terminus, is produced if the first cleavage of the 66' product occurs between the 28K and 38K proteins in the absence of cleavage at the Arg/Gly site at the end of the 66' protein (Fig. 2c). These results indicate that the 38K protein is located at the carboxy terminus of the 66K protein. In addition, analysis of the cleavage products of the polyproteins translated from various truncated transcripts of p66 confirmed that the 28K protein could only arise from the amino-terminal part of the 66K protein (data not shown).

**Determination of the cleavage site within the 66K protein**

The analysis of the carboxy extremity of the CP (Serghini et al., 1990) and the complete VPg sequence (Pinck et al., 1991) indicated that proteolytic cleavage occurred respectively at Arg/Gly, Cys/Ser and Gly/Glu sites. On the basis of these results, several clones carrying inserts corresponding to a protein of about 38K were obtained by PCR mutagenesis of the central region of GFLV RNA2. In each insert the first amino acid downstream of a potential cleavage site was substituted by an ATG codon and the ultimate Arg/6° residue was followed by a termination codon. The initiation and termination codons were built into the PCR primers used to construct each cDNA (see Methods). Clone p38, initiated after Cys/6°, produced a protein comigrating with the 38K maturation product of the polyprotein P2 after transcription and in vitro translation in NT-RRLS (Fig. 3a, lane 1). Clone RV, which contained a methionine in front of Val/18, produced proteins larger than the expected 38K protein (Fig. 3a, lane 3). This protein is processed in the presence of the viral proteinase, and the expected 38K maturation product (Fig. 3b, lane 2) is produced.

To identify the cleavage site within the 66K protein, microsequencing of the labelled maturation products...
Fig. 4. (a) Diagram showing the action of the GFLV viral proteinase (Pro) on the 66K and 94K proteins. The position and nature (CA, Cys/Ala or RG, Arg/Gly) of each cleavage site and the final cleavage products (28K, 38K, and CP) are indicated. (b) Kinetic analysis of 66K and 94K protein cleavage by the GFLV proteinase. The abscissa corresponds to the time of incubation with the viral proteinase and the ordinate the percentage of uncleaved 66K or 94K protein.

Discussion

In this paper we demonstrate that the polyprotein P2 encoded by the GFLV-F13 RNA2 is sequentially processed in vitro by the viral proteinase to yield an N-terminal 28K protein, a 38K protein and the C-terminal 56K CP as final cleavage products. A major 66K intermediate species is also produced which represents with the CP the first products from in vitro maturation of polyprotein P2 by the viral proteinase after cleavage in trans at the Arg<sup>605</sup>/Gly<sup>606</sup> site. This 66K protein is further processed by the same viral proteinase at the Cys<sup>257</sup>/Ala<sup>258</sup> site to yield the 28K and 38K proteins. The GFLV-F13 proteinase is thus able to recognize and cleave in trans two distinct and specific sites in polyprotein P2 with different amino acid residues at the P1 position (Arg<sup>605</sup> or Cys<sup>257</sup>). A greater cleavage efficiency of the proteinase at the Arg/Gly site is observed in the maturation kinetics (Fig. 1b and 2b). In addition the kinetics of 94K and 66K protein maturation show that the 94K protein is almost completely cleaved but the 66K intermediate is only partially processed. This result suggests that the polyprotein P2 is sequentially cleaved by the viral proteinase. A first cleavage occurs at the Arg<sup>605</sup>/Gly<sup>606</sup> site which produces the 66K protein and the CP and, later, a second cleavage at the Cys<sup>257</sup>/Ala<sup>258</sup> site produces the 28K and 38K proteins. Nevertheless, even if the first cleavage of polyprotein P2 were to occur
at the Cys/Ala site, the transient 94K putative intermediate would be rapidly cleaved, and the 94K species would be hardly detectable.

Despite several reports indicating that polyproteins encoded by nepovirus RNA2 produce three proteins, up to now only the cleavage sites at the CP amino terminus have been established (Serghini et al., 1990; Demangeat et al., 1991; Hibrand et al., 1992; Blok et al., 1992). Recently, Blok et al. (1992) reported a Cys/Ala cleavage site which produced the CP from the polyprotein of raspberry ringspot nepovirus (RRV) and speculated that the 66K protein of the RRV polyprotein could also contain a Cys/Ala cleavage site between a 46K and a 20K protein. In contrast to what is suggested for the other nepoviruses reported above, two distinct cleavage sites were involved in the processing of GFLV polyprotein P2 by the same viral proteinase.

The difference in cleavage efficiency observed between the Arg/Gly and Cys/Ala cleavage sites in the GFLV polyprotein P2 may be attributed either to the primary sequence at the cleavage site or to steric accessibility to the active site of the proteinase. However, it cannot be excluded that the incomplete cleavage of the 66K protein may result from our in vitro approach. Moreover, several possible phosphorylation, glycosylation and myristylation sites are present in the viral proteinase, and modifications in planta could alter or regulate the enzyme activity.

In comoviruses, a proteinase cofactor present at the amino terminus of the 200K polyprotein coded by the B-RNA of cowpea mosaic virus, is thought to be responsible for specific recognition of the Gln/Met amino terminus of the 200K polyprotein coded by the RNA-2 of raspberry ringspot nepovirus (RRV). Since our maturation studies were conducted only with the 28K viral proteinase, the existence of a cofactor that could increase the cleavage of the 66K intermediate cannot be ruled out. The incomplete cleavage of the 66K protein could also reflect a mechanism by which the virus regulates the level of each cleavage product during its life cycle.

No biological function has so far been attributed to the 28K and 38K proteins of nepoviruses. It will now be of interest to undertake the detection in vivo of these two non-structural proteins and to determine their biological role.

The authors are grateful to Dr. K. Richards for improving the manuscript. Rogério Margis received a grant from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil).

References


(Received 16 February 1993; Accepted 4 May 1993)