

Characterization of VPg and the polyprotein processing of Cocksfoot mottle virus (genus Sobemovirus)

Kristiina Mäkinen,¹ Katri Mäkeläinen,¹ Natalya Arshava,² Tiina Tamm,^{3,4} Andres Merits,¹ Erkki Truve,^{3,4} Sergei Zavriev² and Mart Saarma¹

¹Institute of Biotechnology, Program for Plant Molecular Biology, Viikki Biocenter, PO Box 56, FIN-00014 University of Helsinki, Finland

²Institute of Agricultural Biotechnology, Timiryazevskaya st. 42, Moscow 127550, Russia

^{3,4}National Institute of Chemical Physics and Biophysics³ and Gene Technology Center, Tallinn Technical University⁴, Akadeemia tee 23, EE12618 Tallinn, Estonia

The polyprotein of Cocksfoot mottle virus (CfMV; genus *Sobemovirus*) is translated from two overlapping open reading frames (ORFs) 2a and 2b by a -1 ribosomal frameshifting mechanism. In this study, a 12 kDa protein was purified from viral RNA-derived samples that appears to correspond to the CfMV genome-linked protein (VPg). According to the determined N-terminal amino acid sequence, the VPg domain is located between the serine proteinase and replicase motifs and the N terminus of VPg is cleaved from the polyprotein between glutamic acid and asparagine residues. Western blot analysis of infected plant material showed that the polyprotein is processed at several additional sites. An antiserum against the ORF 2a product recognized six distinct proteins, whereas, of these, the VPg antiserum clearly recognized only a 24 kDa protein. This indicates that the fully processed 12 kDa VPg detected in viral RNA-derived samples is a minor product in infected plants. An antiserum against the ORF 2b product recognized a 58 kDa protein, which indicates that the fully processed replicase is entirely or almost entirely encoded by ORF 2b. The origin of the detected cleavage products and a proposed polyprotein processing model are discussed.

Introduction

Cocksfoot mottle virus (CfMV) has a monopartite, single-stranded, 4082 nucleotide, positive-sense RNA genome (Mäkinen *et al.*, 1995a; Ryabov *et al.*, 1996). The translational strategy of sobemoviruses is based on production of a large polyprotein from the genomic RNA and synthesis of the coat protein (CP) from a subgenomic RNA (Matthews, 1991). Unlike other sobemoviruses with known genomic RNA sequences, CfMV lacks the continuous open reading frame (ORF) encoding the entire polyprotein (Mäkinen *et al.*, 1995a; Ryabov *et al.*, 1996). The polyprotein of CfMV is translated from two overlapping ORFs, 2a and 2b, by a -1 ribosomal frameshifting mechanism (Mäkinen *et al.*, 1995b).

Several viruses from various virus groups have a protein (VPg) linked covalently at the 5'-end of the viral RNA. A model for the initiation of poliovirus RNA synthesis was

proposed by Andino *et al.* (1993). According to this model, uridynylated VPg (VPg-pU-pU) serves as a primer for 3D replicase. Both genomic and subgenomic RNAs of *Southern bean mosaic virus* (SBMV; genus *Sobemovirus*) contain a VPg of about 12 kDa covalently linked to their 5'-ends (Ghosh *et al.*, 1979, 1981; van der Wilk *et al.*, 1998). The VPg of *Southern cowpea mosaic virus* (SCPMV; genus *Sobemovirus*) is about 10 kDa (Mang *et al.*, 1982). In the case of most of the single-stranded RNA viruses of which the VPg has been identified, the domains are arranged in the polyprotein in the following order: VPg-proteinase-polymerase (Koonin & Dolja, 1993). An increasing number of viruses, including SBMV and viruses related to sobemoviruses, have been shown to have a different polyprotein arrangement: proteinase-VPg-polymerase (van der Wilk *et al.*, 1997, 1998; Revill *et al.*, 1998; Wobus *et al.*, 1998). Although the region encoding VPg is similarly located in the genomes of *Potato leafroll virus* (PLRV), SBMV, *Pea enation mosaic virus* (PEMV-1) and *Mushroom bacilliform virus* (MBV), no significant similarity has been reported in the amino acid sequence or in the size of the VPgs. Each of these viruses

Author for correspondence: Kristiina Mäkinen.

Fax +358 9 19159571. e-mail kristiina.makinen@helsinki.fi

has been either shown or proposed to express its replicase by a -1 ribosomal frameshifting mechanism (Demler & Zoeten, 1991; Prüfer *et al.*, 1992; Kujawa *et al.*, 1993; Revill *et al.*, 1994). In SBMV, the N-terminal amino acid of VPg is threonine, which is preceded by a glutamic acid, indicating that an E/T processing site is used for VPg maturation (van der Wilk *et al.*, 1998). It is also known that no polyprotein processing of CfMV occurs in an *in vitro* translation reaction (Tamm *et al.*, 1999). No further information about the polyprotein processing of sobemoviruses exists.

In this work, a protein that appears to correspond to the CfMV VPg was isolated and its size and N-terminal amino acid sequence were determined. The proteolytic cleavage site at the VPg N terminus was identified, providing the first indication of how the CfMV polyprotein is processed *in vivo*. Immunoblotting of infected plant material showed that the polyprotein is processed at several additional sites; one of them was predicted by sequence analysis. Based on these data, a proteolytic cleavage scheme is proposed for the CfMV polyprotein.

Methods

■ **Virus preparation and RNA isolation.** Leaves of *Hordeum vulgare* infected with CfMV Russian isolate were homogenized in 100 mM phosphate buffer, pH 7.5, containing 0.5% Na₂SO₃ and 5 mM EDTA. The virus was pelleted by ultracentrifugation and then purified by centrifugation into a 20% sucrose cushion. In order to identify and characterize the VPg of CfMV, the genomic RNA was isolated either by NaClO₄ extraction (Atabekov, 1981) or with the plant RNeasy Total RNA kit (Qiagen) according to the manufacturer's instructions.

■ **Inoculation.** For Western blot analysis, 12-day-old seedlings of *Hordeum vulgare* cv. Kustaa were inoculated mechanically by grinding 1 g of leaves of CfMV-infected barley cv. Lise with a mortar and pestle per 10 ml of distilled water and rubbing the sap onto the lowest leaf of each barley plant (each plant had a total of two to three leaves) dusted with carborundum. Western blot analysis was used for virus detection.

■ **Preparation of total soluble protein samples from infected leaves.** Approximately 50 mg samples of CfMV-infected barley cv. Kustaa and the corresponding mock-inoculated leaves were collected 21 days post-inoculation and immediately frozen in liquid nitrogen. Frozen leaf tissues were ground with a pestle in a microcentrifuge tube and the homogenized material was mixed with 300 µl Laemmli sample buffer (Laemmli, 1970). After heating the samples for 5 min at 100 °C, the insoluble material was removed by centrifugation (14 000 g) and the supernatant was loaded on the gel.

■ **Preparation of VPg antiserum.** The VPg-encoding region in the CfMV polyprotein gene was amplified by PCR from plasmid pORF2a/2b (Tamm *et al.*, 1999) by using the oligonucleotides 5' CGCGGATCCA-ACAGTGAGTTATATCCC (contains a *Bam*HI site and CfMV nucleotides 1385–1403) and 5' ACGCGTCTGACTTATTCCTTCGTCACG-CCAG (contains a *Sal*I site, a UAA termination codon and sequence complementary to CfMV nucleotides 1708–1724). Base numbering refers to the CfMV genome as described in Mäkinen *et al.* (1995a). The resulting fragment was cloned into pGEM-T (Promega) and sequenced. The correct fragment was subcloned into the *Bam*HI and *Sal*I sites of pQE30 as a 6-His fusion (Qiagen). The resulting expression plasmid was

used to transform *E. coli* strain M15 (Qiagen). Expression and purification of the 6-His-VPg fusion were carried out according to the manufacturer's protocol. Rabbit polyclonal antiserum was produced against the expressed 6-His fusion protein by using a denatured protein solution (in 4 M urea) for immunization.

■ **Electrophoresis and Western blot analysis.** SDS-PAGE was carried out on an RNase A-treated CfMV RNA-derived sample and the resulting gel was silver-stained. Western blot analysis was carried out on the same sample and on virus particles. Virus particles were prepared by disrupting them with 1% SDS prior to RNase A treatment. Proteins were separated in a 16.5% Tricine-SDS-PAGE system (Schagger & von Jagow, 1987) and then electroblotted onto a nitrocellulose filter. The filter was blocked with 3% BSA. A specific antiserum produced in rabbits against ORF 2a-encoded recombinant protein of CfMV (P2a antiserum; Tamm *et al.*, 1999) was used at a 1:2000 dilution (Fig. 1, lane 3). Anti-rabbit antibody conjugated with alkaline phosphatase (Sigma) was used as the second antibody.

Western blot analysis from the leaf samples was carried out as follows. SDS-PAGE-separated proteins were electroblotted onto a PVDF membrane (Millipore) and the membrane was blocked with 5% non-fat milk powder in PBS. Rabbit anti-CfMV polyclonal antibodies against purified CfMV particles (Truve *et al.*, 1997) were used at 1:5000 dilution for detection of virus from infected plants. The P2a, P2b and VPg antisera were also used at 1:5000 dilution. Specific P2b antiserum was produced against the replicase encoded by the -1 frame in the polyprotein-encoding region (Tamm *et al.*, 1999). A 1:5000 dilution of anti-rabbit antibody conjugated with horseradish peroxidase (Sigma) was used as the second antibody. The TMB (3,3',5,5'-tetramethyl benzidine) stabilized substrate for horseradish peroxidase (Promega) was used for detection.

■ **Iodination.** CfMV RNA (13–40 µg depending on the experiment) was iodinated by the chloramine T method as described by Sainio *et al.* (1997). The reaction contained 29 µl 0.5 M phosphate buffer, pH 7.5, 20 µl sample, 2.5 µl Na¹²⁵I (250 µCi; Amersham) and 4 µl chloramine T (Serva; 2 mg in 1 ml 0.5 M phosphate buffer, pH 7.5). The chloramine T treatment lasted for 20 s and then 50 µl Na₂S₂O₅ (2.4 mg in 1 ml 0.5 M phosphate buffer, pH 7.5) was added. After 30 s, 15 µl 0.1 M NaI and 50 µl 1% BSA in PBS were added to the reaction. After iodination, the free label was removed from the sample by using a Sephadex G-25 column (Pharmacia) in the presence of 1% BSA. The radioactivity of 1 µl samples from each fraction was counted by using a beta counter (Wallac). To eliminate excess BSA, the RNA sample was ethanol-precipitated prior to RNase A treatment (10 µg/ml, 1 h, 37 °C). Part of the sample was treated with proteinase K (50 µg/ml, 1 h, 37 °C) in addition to RNase A. Samples were analysed in a 15% SDS-PAGE gel.

■ **Immunoprecipitation.** Two hundred µl of the iodinated products (approximately 2.5×10^3 c.p.m. per sample) was immunoprecipitated (2 h, 0 °C) by using 5 µl antiserum raised against CfMV particles or ORF 2a- and ORF 2b-encoded proteins expressed in *E. coli* and then diluted with 200 µl lysis buffer (Tris-borate buffer, pH 7.5, 2 mM EDTA, 10% glycerol, 1% Nonidet P-40, 1% Triton X-100). Complexes were collected by adding 30 µl of a 50% suspension of Protein A-Sepharose granules (Pharmacia) in lysis buffer. Granules were previously treated with 0.2% BSA. The suspension was incubated for 1 h on a rotating wheel in a cold room (4 °C). Granules were washed four times with PBS/2% Triton X-100, 30 µl Laemmli sample buffer was added and the samples were boiled prior to separation by 15% SDS-PAGE.

■ **Amino acid sequencing.** A CfMV RNA sample was isolated from the purified virus stock (13 mg/ml) by using the plant RNeasy Total RNA kit (Qiagen) according to the manufacturer's instructions. Seventy

μg RNA was hydrolysed in 10% trifluoroacetic acid (TFA) for 48 h at room temperature. The concentration of TFA was reduced by evaporation. The hydrolysed material was blotted onto a PVDF membrane by using a ProSorb cartridge. The PVDF-immobilized protein was sequenced directly by automated Edman degradation.

Results

Samples of CfMV RNA were purified by NaClO_4 extraction, treated with RNase A and separated on a 16.5% Tricine-SDS-PAGE gel which was then silver-stained (Fig. 1, lane 2). The RNA sample contained a 32 kDa protein that comigrated with the CP (data not shown) and a protein of approximately 16 kDa, which is the RNase used to degrade the CfMV RNA. Also, a 12 kDa protein co-purified together with the viral RNA (Fig. 1, lane 2). This 12 kDa protein was recognized specifically by the P2a antiserum (Fig. 1, lane 3). When virus particles were analysed with the P2a antiserum, the 12 kDa protein was not detected (data not shown). We propose that either the large excess of CP or incomplete degradation of RNA under the conditions used prevented migration of the 12 kDa protein in the gel as a clearly pronounced, single band.

For further characterization of the 12 kDa protein, the RNA isolation method was changed to the plant RNeasy Total RNA kit (Qiagen). This was done in order to reduce the amount of contaminating non-covalently bound CP (Fig. 1, lane 1) in viral RNA samples. RNA obtained with the kit was labelled with ^{125}I by the chloramine T method. The labelled samples were analysed by radiography of the polyacrylamide gel and immunoprecipitation (Fig. 2). Two labelled products were observed in the ^{125}I -labelled CfMV RNA sample treated with RNase A. The sizes of these proteins were 12 and 32 kDa (Fig. 2A, lane 2). To confirm that the signals detected following SDS-PAGE indeed originated from iodinated proteins, the RNase A-treated samples were treated with proteinase K. In this case, the ^{125}I -labelled products disappeared, confirming that the iodinated targets were proteins (Fig. 2A, lane 1). Much less of the 32 kDa CP was observed compared with the amount of the 12 kDa protein in the purified RNA sample isolated with the kit than in the corresponding sample isolated by NaClO_4 extraction, showing that the kit yielded purer RNA.

The origin of the iodinated proteins was studied by immunoprecipitation. Approximately 2.5×10^5 c.p.m. per sample of iodinated CfMV RNA was treated with RNase A prior to immunoprecipitation with P2a, P2b or CP (P3) antiserum. P2a antiserum immunoprecipitated the 12 kDa protein and the P3 antiserum immunoprecipitated the 32 kDa protein (Fig. 2A, lanes 3 and 5). The P2b antiserum did not immunoprecipitate either of the proteins (Fig. 2A, lane 4). From this experiment, we conclude that a 12 kDa protein, attached to the CfMV RNA, is recognized by the antiserum raised against the ORF 2a-encoded protein. The 12 kDa iodinated protein from the sample treated with RNase A

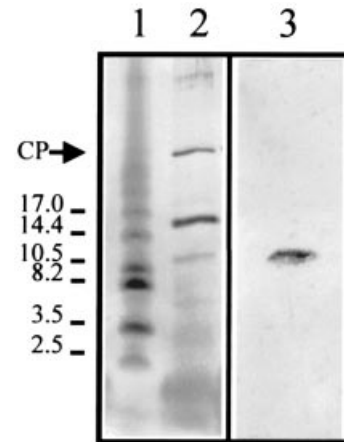


Fig. 1. Analysis of viral particles. Silver-stained SDS-PAGE gel (lanes 1–2) and Western blot analysis (lane 3) of a viral-RNA-derived sample. The separation of samples was done by electrophoresis in the gel, which was cut into two halves prior to staining and immunoblotting. P2a antiserum was used as the specific antiserum and alkaline phosphatase-conjugated anti-rabbit antibody as the second antibody in Western blot analysis. Lanes: 1, molecular mass markers (sizes in kDa); 2–3, CfMV RNA-derived sample (NaClO_4 extraction).

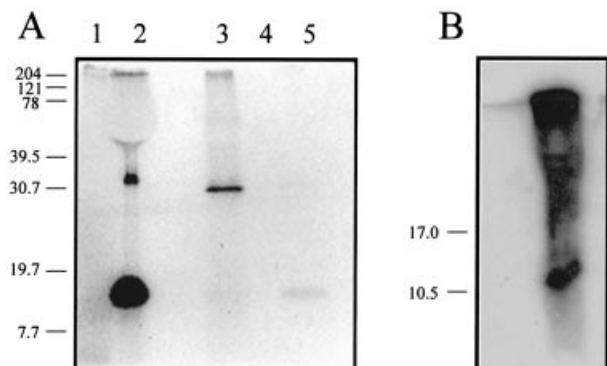


Fig. 2. Analysis of ^{125}I -labelled VPg linked to the CfMV RNA. (A) Lanes: 1, ^{125}I -labelled CfMV-derived RNA sample treated with RNase A and proteinase K; 2, ^{125}I -labelled CfMV RNA-derived sample treated with RNase A; 3–5, iodinated, RNase A-treated, CfMV-RNA-derived sample immunoprecipitated with P3 (lane 3), P2b (4) or P2a (5) antiserum. (B) ^{125}I -labelled CfMV-derived RNA sample not treated with RNase A. Sizes of molecular mass markers are indicated on the left (in kDa). Viral RNA for the iodination experiments was isolated with the plant RNeasy Total RNA kit (Qiagen).

migrated into the resolving portion of the SDS-PAGE gel, but a large part of the radioactivity from the non-treated sample remained at the top of the resolving gel, or its migration was retarded. A minor part of the 12 kDa protein from the non-treated sample migrated normally (Fig. 2B).

The sequence of the N-terminal 17 amino acids of the 12 kDa protein associated with the CfMV RNA was determined to be N*EL*PDQSSGPARELD. Amino acids represented by asterisks gave a signal that did not correspond to any residue used as a standard. According to the nucleotide sequence, they should be serine and tyrosine, respectively, and

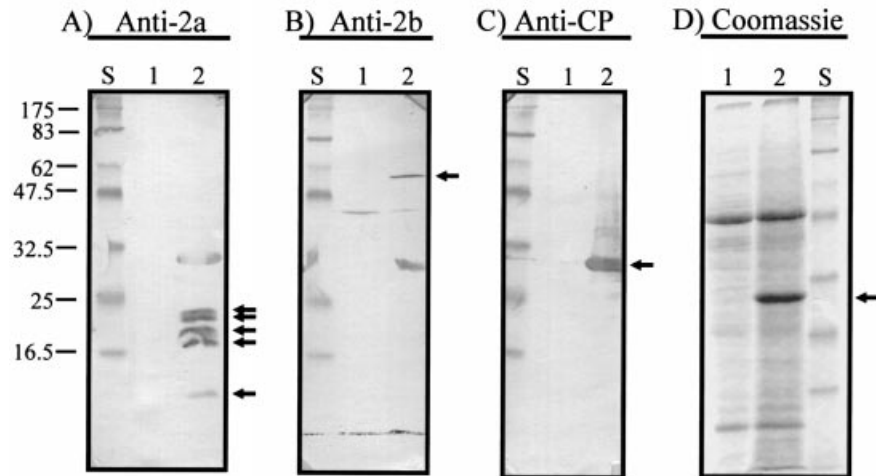


Fig. 3. Western blot analysis of total protein derived from CfMV-infected leaves with P2a (A), P2b (B) and P3 (CP) antisera. Protein samples in each blot as well as in the Coomassie blue-stained gel (D) are as follows. Lanes: 1, mock-inoculated barley cv. Kustaa; 2, infected barley cv. Kustaa. The immunoblots presented originate from one gel that was cut into three prior to immunodetection. Specific bands observed in each blot are indicated by arrows. The Coomassie blue-stained gel indicates the amount of protein loaded. Lanes containing molecular mass markers are indicated by S; the masses of the markers are given on the left (in kDa).

we suggest that these amino acids are probably modified in the mature 12 kDa protein. The other signals were unambiguous and the presence of trace amounts of the CP in the sample did not disturb the sequence analysis. The identified residues represent amino acids 320–337 of the polyprotein (Mäkinen *et al.*, 1995a; Ryabov *et al.*, 1996). The N-terminal amino acid sequences of VPg molecules of other related viruses have been identified and studied by similar techniques (van der Wilk *et al.*, 1997, 1998; Revill *et al.*, 1998; Wobus *et al.*, 1998). The data obtained taken together lead to the suggestion that the 12 kDa protein associated with the CfMV RNA represents the VPg molecule.

In order to analyse the size of the *in vivo* processing products of the CfMV polyprotein, Western blot analysis of CfMV-infected leaf material was carried out with P2a, P2b and P3 (CP) antisera (Fig. 3). The P2a antiserum recognized proteolytically processed proteins of 12, 18, 19, 20, 23, 24 and 30 kDa (Fig. 3A). Occasionally, a protein of 8 kDa was also detected (data not shown). The major products recognized specifically by the P2b antiserum were 30 and 58 kDa (Fig. 3B). The CP of CfMV, detected with the P3 antibody, is 30 kDa (Fig. 3C, D). Both the P2a and P2b antisera detected a protein that has the same mobility as the CP. Either a polypeptide that has the same mobility as the CP can be recognized by these two antibodies or a cross-reaction has occurred. The 18–24 kDa polypeptides recognized by the P2a antiserum are not related to the CfMV CP (Fig. 3A, C).

A recombinant protein, identical to VPg in its N terminus and ending at a glutamic acid residue (435) of P2a (calculated molecular mass 12.3 kDa), was produced in *E. coli*. This protein was purified and used for production of polyclonal antibody in

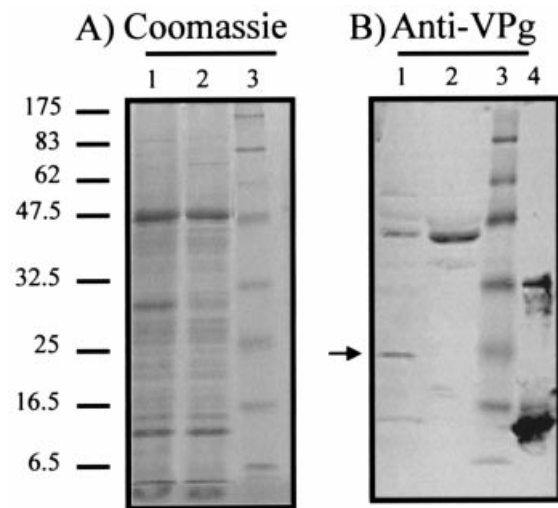


Fig. 4. Western blot analysis with VPg antiserum of a total protein sample derived from CfMV-infected leaves. (A) Coomassie blue-stained gel indicating the amount of protein loaded. Lanes are numbered 1–3 as in the blot. (B) Western blot analysis with VPg antiserum. Lanes: 1, infected barley cv. Kustaa; 2, mock-inoculated barley cv. Kustaa; 3, molecular mass markers (mass of each protein marker is given on the left in kDa); 4, purified recombinant VPg produced in *E. coli*. The specific band recognized by the VPg antiserum in infected leaves is indicated.

rabbits. The resulting antiserum recognized the recombinant VPg produced in *E. coli* (Fig. 4B, lane 4). In infected leaves, the VPg antiserum specifically recognized a 24 kDa protein (Fig. 4B). The faintly detectable 12 kDa protein may represent the VPg, but we occasionally detected a similar band in the mock control. Occasionally, a protein of 62 kDa, probably repre-

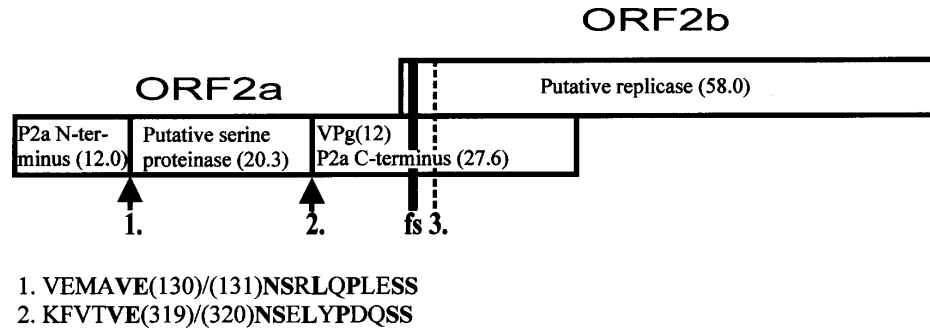


Fig. 5. Model for proteolytic processing of CfMV polyprotein. Based on the CfMV VPg N-terminal amino acid sequence, the CfMV polyprotein is cleaved between glutamic acid (E) and asparagine (N) (site 2). An additional site (site 1), resembling the determined cleavage site, was found in the polyprotein of CfMV. Amino acid sequences at the cleavage sites are given. Amino acid numbering refers to the CfMV polyprotein according to the sequence in Mäkinen *et al.* (1995a). The calculated molecular masses of the putative cleavage products are given in parentheses (in kDa). The location of the signal for -1 ribosomal frameshifting (fs) is represented by a thick line. The C terminus of the mature VPg may originate from either frame; therefore, the estimated location of the VPg C terminus (site 3), based on the determined sizes of VPg and replicase, is indicated by the dashed line crossing both ORFs 2a and 2b.

senting the unprocessed form of P2a, as well as partially processed forms of P2a of 58 and 32.5 kDa were detected with VPg antiserum (data not shown). The lack of the clearly detectable, expected 12 kDa VPg cleavage product in this blot may indicate that the amount of it in leaf extracts is very small.

Based on the N-terminal amino acid sequence of VPg, the first asparagine (N) residue is preceded by a glutamic acid residue (E), indicating that the proteolytic processing site is between E and N. The amino acid sequence surrounding this cleavage site is **VE/NSELYPDQSS**. Another putative cleavage site, **VE/NSRLQPLESS**, resembling that found at the N terminus of VPg, can be found in the P2a sequence. The putative serine proteinase of CfMV is located between these two cleavage sites. The VPg molecule is located within the polyprotein, after the putative serine proteinase. A proteolytic processing model for the polyprotein of CfMV is proposed in Fig. 5. The model is based on the N-terminal sequence of VPg, immunological data and the predicted processing sites.

Discussion

In this work, it has been shown that a protein of approximately 12 kDa, recognized by the P2a antiserum, is attached to the CfMV RNA. This protein largely remains attached to viral RNA when it is subjected to SDS-PAGE (Fig. 2B). This result suggests that the 12 kDa protein is the VPg of CfMV. The N-terminal amino acid sequence of the CfMV VPg places it between the serine proteinase and replicase motifs in the CfMV polyprotein. Computer comparisons of the amino acid sequence surrounding the cleavage site of the N terminus of CfMV VPg do not reveal any similar sites in SBMV (Othman & Hull, 1995) or *Rice yellow mottle virus* (RYMV; Ngon A Yassi *et al.*, 1994). The determined N-terminal sequence of the SBMV VPg is RSQE/TLPELSVIE (van der Wilk *et al.*, 1998), which differs completely from that of CfMV

(VTVE/NSELYPDQSS). Previously, we reported that an amino acid motif W(A,G)D followed by an aspartic acid- and glutamic acid-rich region can be found upstream of the established or putative -1 ribosomal frameshift sites in several viruses related to sobemoviruses (Mäkinen *et al.*, 1995a). This sequence was found in the sobemoviruses CfMV, RYMV, SBMV and SCPMV, in the poleroviruses PLRV and *Beet western yellows virus* (BWYV), in the enamovirus PEMV, in the barnavirus MBV and in a human astrovirus, H-Ast2 (Mäkinen *et al.*, 1995a). The electrophoretically determined sizes of sobemovirus (CfMV, SBMV and SCPMV), polerovirus (PLRV) and barnavirus (MBV) VPgs suggest that the motif W(A,G)D followed by a D- and E-rich region is a part of the VPgs of these viruses. This is the only sequence motif conserved between VPgs of these related viruses.

The putative serine proteinase region in the ORF 2a-encoded part of the CfMV polyprotein resembles the picornavirus 3C^{Pro} proteinase (Mäkinen *et al.*, 1995a). It has been predicted that the luteovirus and sobemovirus polyproteins are cleaved at (Q,E)/(G,S,A) sites (Gorbalenya *et al.*, 1988). The finding that the N-terminal cleavage site of the CfMV VPg is between glutamic acid (E) and asparagine (N) was thus unexpected. The fact that another similar cleavage site was found in the polyprotein sequence at the beginning of the serine proteinase region gives some hints as to how the CfMV polyprotein may be processed. The calculated molecular masses of polypeptides cleaved from P2a using the E/N sites presented in Fig. 5 are 12.0 kDa for the protein preceding the putative proteinase region, 20.3 kDa for the putative serine proteinase and 26.7 kDa for the rest of P2a, including the VPg region. P2a antibody detected a 12 kDa protein in infected leaf material, which was either not detected or was detected faintly with the anti-VPg antibody from the same samples. Most of the 12 kDa protein detected by the P2a antibody probably corresponds to the polypeptide preceding the putative serine

proteinase in the CfMV polyprotein. The polypeptide starting at asparagine-131 and ending at glutamic acid-319 would be 189 amino acids long. The number of amino acid residues in different picornavirus 3C proteins, which are closely related to sobemovirus serine proteinases, is close to 200 (Rueckert, 1996). The 18–23 kDa polypeptides recognized in infected leaf extracts by the P2a antiserum may represent differently modified, processed or degraded forms of the putative serine proteinase. Both the P2a and VPg antisera recognized a 24 kDa protein, which represents the part of P2a from the beginning of VPg to the C terminus of P2a. The 12 kDa protein, which appears to correspond to VPg, is a minor product in infected plants and can only be clearly detected by Western blot analysis of viral RNA-derived samples.

In PLRV-infected plants, a protein of 25 kDa is detected by P1 (ORF1) monoclonal and polyclonal antibodies (Prüfer *et al.*, 1999). The N terminus of this protein is either identical to or located adjacent to the PLRV VPg. In a model suggested by Prüfer *et al.* (1999), it was proposed that the hydrophobic N-terminal region of PLRV polyprotein P1 targets the protein to cellular membranes and that the basic nucleic acid-binding domain at its C terminus interacts with the PLRV RNA. This membrane-bound complex can then serve as a proteolytic processing site for VPg maturation. Several aspects of this model could be applicable to the polyprotein processing of CfMV, since, because of similarities in the polyprotein sequence and arrangement, sobemoviruses and poleroviruses may share similar proteolytic processing strategies. The 60 N-terminal amino acids of P2a contain hydrophobic residues and have been proposed to form a transmembrane domain in the CfMV polyprotein (Ryabov *et al.*, 1996). The C-terminal part of P2a contains a strong basic region (amino acids 539–552) in the P2a sequence according to Mäkinen *et al.* (1995a) and may determine the RNA-binding property of P2a (Tamm & Truve, 2000). No proteolytic processing of the CfMV polyprotein was observed in a cell-free translation system (Tamm *et al.*, 1999). Therefore, proteolytic processing of the CfMV polyprotein may require association with the replication complex and/or a proper cellular environment.

Interestingly, since the size of the CfMV VPg appears to be approximately 12 kDa, the coding region of the VPg probably overlaps the site in the CfMV polyprotein of the –1 frameshifting signal (Fig. 5). The antiserum against the ORF 2b product recognized a 58 kDa protein, which indicates that the fully processed replicase is either entirely or almost entirely encoded by ORF 2b. Two possibilities exist: the VPg attached to the CfMV RNA originates either from P2a or from the transframe protein. No suitably positioned E/N site, similar to the N-terminal VPg maturation site, is present downstream of the N terminus of the VPg sequence in either frame. Also, it is not possible to find a clearly conserved putative cleavage site by comparing the downstream sequences of sobemoviruses and poleroviruses. Therefore the C-terminal site used for maturation of CfMV VPg cannot be predicted. It will be

interesting to determine the biological significance of the fact that VPg or a molecule N-terminally similar to VPg is produced in fusion with the replicase. In particular, this may be important for establishing the replication complex. Furthermore, –1 ribosomal frameshifting regulates the abundance of the various forms of the molecules related to the VPg as well as that of the replicase, which may contribute to the regulation of CfMV replication.

The authors wish to acknowledge Dr Viktor Novikov for CfMV preparations, Dr Nisse Kalkkinen for sequencing the N terminus of VPg and Dr Alan Schulman for critical reading of the manuscript. This work was supported by EU INCO Copernicus grant IC15-CT96-0907, the Technology Development Center (TEKES; grant number 40673/98), Biocentrum Helsinki and the Estonian Science Foundation grant #2308.

References

- Andino, R., Rieckhof, G. E., Achacoso, P. L. & Baltimore, D. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5' end of viral RNA. *EMBO Journal* **12**, 3587–3598.
- Atabekov, J. G. (editor) (1981). Extraction of TMV RNA using SDS and NaClO₄. In *Practice in General Virology*, pp. 104–106. Moscow: Moscow State University.
- Demler, S. A. & de Zoeten, G. A. (1991). The nucleotide sequence and luteovirus-like nature of RNA 1 of an aphid non-transmissible strain of pea enation mosaic virus. *Journal of General Virology* **72**, 1819–1834.
- Ghosh, A., Dasgupta, R., Salerno-Rife, T., Rutgers, T. & Kaesberg, P. (1979). Southern bean mosaic viral RNA has a 5'-linked protein but lacks 3' terminal poly(A). *Nucleic Acids Research* **7**, 2137–2146.
- Ghosh, A., Rutgers, T., Ke-Qiang, M. & Kaesberg, P. (1981). Characterization of the coat protein mRNA of southern bean mosaic virus and its relationship to the genomic RNA. *Journal of Virology* **39**, 87–92.
- Gorbalenya, A. E., Koonin, E. V., Blinov, V. M. & Donchenko, A. P. (1988). Sobemovirus genome appears to encode a serine protease related to cysteine proteases of picornaviruses. *FEBS Letters* **236**, 287–290.
- Koonin, E. V. & Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Critical Reviews in Biochemistry & Molecular Biology* **28**, 375–430.
- Kujawa, A. B., Dugeon, G., Hulanicka, D. & 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 Research* **21**, 2165–2171.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Mäkinen, K., Tamm, T., Næss, V., Truve, E., Puurand, Ü., Munthe, T. & Saarma, M. (1995a). Characterization of cocksfoot mottle sobemovirus genomic RNA and sequence comparison with related viruses. *Journal of General Virology* **76**, 2817–2825.
- Mäkinen, K., Næss, V., Tamm, T., Truve, E., Aaspõllu, A. & Saarma, M. (1995b). The putative replicase of cocksfoot mottle sobemovirus is translated as a part of the polyprotein by –1 ribosomal frameshift. *Virology* **207**, 566–571.
- Mang, K.-Q., Ghosh, A. & Kaesberg, P. (1982). A comparative study of the cowpea and bean strains of southern bean mosaic virus. *Virology* **116**, 264–274.

- Matthews, E. F. (1991).** Sobemovirus group. In *Plant Virology*, 3rd edn, pp. 239–241. San Diego: Academic Press.
- Ngon A Yassi, M., Ritzenthaler, C., Brugidou, C., Fauquet, C. & Beachy, R. N. (1994).** Nucleotide sequence and genome characterization of rice yellow mottle virus RNA. *Journal of General Virology* **75**, 249–257.
- Othman, Y. & Hull, R. (1995).** Nucleotide sequence of the bean strain of southern bean mosaic virus. *Virology* **206**, 287–297.
- Prüfer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A. & Rohde, W. (1992).** Ribosomal frameshifting in plants: a novel signal directs the –1 frameshift in the synthesis of the putative viral replicase of potato leafroll luteovirus. *EMBO Journal* **11**, 1111–1117.
- Prüfer, D., Kawchuk, L., Monecke, M., Nowok, S., Fischer, R. & Rohde, W. (1999).** Immunological analysis of potato leafroll luteovirus (PLRV) P1 expression identifies a 25 kDa RNA-binding protein derived via P1 processing. *Nucleic Acids Research* **27**, 421–425.
- Revell, P. A., Davidson, A. D. & Wright, P. J. (1994).** The nucleotide sequence and genome organization of mushroom bacilliform virus: a single-stranded RNA virus of *Agaricus bisporus* (Lange) Imbach. *Virology* **202**, 904–911.
- Revell, P. A., Davidson, A. D. & Wright, P. J. (1998).** Mushroom bacilliform virus RNA: the initiation of translation at the 5' end of the genome and identification of the VPg. *Virology* **249**, 231–237.
- Rueckert, R. R. (1996).** *Picornaviridae: the viruses and their replication* In *Fields Virology*, 3rd edn, pp. 609–654. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott–Raven.
- Ryabov, E. V., Krutov, A. A., Novikov, V. K., Zhelenznikova, O. V., Morozov, S. Yu. & Zavriev, S. K. (1996).** Nucleotide sequence of RNA from the sobemovirus found in infected cocksfoot shows a luteovirus-like arrangement of the putative replicase and protease genes. *Phytopathology* **86**, 391–397.
- Sainio, K., Suvanto, P., Davies, J., Wartiovaara, J., Wartiovaara, K., Saarma, M., Arumäe, U., Meng, X., Lindahl, M., Pachnis, V. & Sariola, H. (1997).** Glial-cell-line-derived neurotrophic factor is required for bud initiation from ureteric epithelium. *Development* **124**, 4077–4087.
- Schagger, H. & von Jagow, G. (1987).** Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* **166**, 368–379.
- Tamm, T. & Truve, E. (2000).** RNA-binding activities of cocksfoot mottle sobemovirus proteins. *Virus Research* **66**, 197–207.
- Tamm, T., Mäkinen, K. & Truve, E. (1999).** Identification of genes encoding for the cocksfoot mottle virus proteins. *Archives of Virology* **144**, 1557–1567.
- Truve, E., Næss, V., Blystad, D. R., Järvekülg, L., Mäkinen, K., Tamm, T. & Munthe, T. (1997).** Detection of cocksfoot mottle virus particles and RNA in oat plants by immunological, biotechnical and electron-microscopical techniques. *Archiv für Phytopathologie und Pflanzenschutz* **30**, 473–485.
- van der Wilk, F., Verbeek, M., Dullemans, A. M. & van den Heuvel, J. F. J. M. (1997).** The genome-linked protein of potato leafroll virus is located downstream of the putative protease domain of the ORF1 product. *Virology* **234**, 300–303.
- van der Wilk, F., Verbeek, M., Dullemans, A. M. & van den Heuvel, J. F. J. M. (1998).** The genome-linked protein (VPg) of southern bean mosaic virus is encoded by the ORF2. *Virus Genes* **17**, 21–24.
- Wobus, C. E., Skaf, J. S., Schultz, M. H. & de Zoeten, G. A. (1998).** Sequencing, genomic localization and initial characterization of the VPg of pea enation mosaic enamovirus. *Journal of General Virology* **79**, 2023–2025.

Received 5 June 2000; Accepted 21 July 2000