Characterization of proteinase cleavage sites in the N-terminal region of the RNA1-encoded polyprotein from *Arabis mosaic virus* (subgroup A nepovirus)

Thierry Wetzela, Joan Chisholmb, Alexandra Basslera, Hélène Sanfaçonb,⁎

a RLP Agroscience, AllPlanta Institute for Plant Research, Breitenweg 71, 67435, Neustadt an der Weinstrasse, Germany
b Pacific Agri-Food Research Centre, Agriculture and Agri-Food Canada, P.O. Box 5000, 4200 Highway 97, Summerland, B.C., Canada V0H 1Z0

Received 19 November 2007; returned to author for revision 20 December 2007; accepted 30 January 2008

Available online 4 March 2008

Abstract

*Arabis mosaic virus* is a subgroup A nepovirus. The RNA1-encoded polyprotein (P1) contains the domains for the NTP-binding protein (NTB), VPg, proteinase (Pro) and polymerase at its C-terminus. Putative cleavage sites delineating these domains have been proposed. However, the number and location of cleavage sites upstream of the NTB domain are not known. Using *in vitro* processing assays, we have confirmed proteolytic cleavage at the NTB-VPg and VPg-Pro sites. In addition, we have identified two cleavage sites in the N-terminal region of P1. Site-directed mutagenesis and immunoprecipitation experiments using inserted peptide tags confirmed that the position of these cleavage sites correspond to that of cleavage sites delineating the X1 and X2 domains in *Tomato ringspot virus* (subgroup C nepovirus). Amino acid alignments implied the presence of similar cleavage sites in the P1 polyprotein of other nepoviruses. Our results suggest that the presence of two protein domains upstream of NTB is a common feature of nepoviruses.

© 2008 Elsevier Inc. All rights reserved.

Keywords: Positive-strand RNA virus; Picorna-like virus; Nepovirus; Comoviridae; Protease; Cleavage site; Genomic organization

Introduction

The genus *Nepovirus* is a group of plant-infecting picorna-like viruses. They have a bipartite genome and share sequence identity, common genome expression strategies and biological properties (Le Gall et al., 2005, 2007). Nepoviruses belong to the family *Comoviridae* along with members of the genera *Comovirus* and *Fabavirus*. Among picorna-like viruses, nepoviruses have the unique property of encapsidating their genome using a single large coat protein (CP) subunit (Le Gall et al., 2005, 2007). In spite of common characteristics, nepoviruses are diverse and the genus has been divided into three subgroups based on the length of RNA2, immunological properties of the virions, sequence identities in the CP and cleavage site specificity of the viral proteinase (Le Gall et al., 2005; Mayo and Robinson, 1996).

Each nepovirus genomic RNA encodes one large polyprotein that is cleaved into mature proteins and intermediate precursors by the viral proteinase (Pro). Nepovirus proteinases are related to the 3C Pro of picornaviruses (Gorbalenya et al., 1989b). A conserved histidine in the substrate-binding pocket of the 3C Pro of picornaviruses is responsible for the specific recognition of a glutamine or glutamate at the −1 position of the cleavage sites (Gorbalenya et al., 1989a; Seipelt et al., 1999). This histidine is present in the substrate-binding pocket of proteinases of subgroup C nepoviruses and polyprotein cleavage was shown to occur after a glutamine, asparagine or aspartate (Bacher et al., 1994; Carrier et al., 1999; Latvala et al., 1998). The proteinases of subgroup A and B nepoviruses have a leucine instead of a histidine in their substrate-binding pockets (Margis and Pinck, 1992). Subgroup B nepovirus cleavage sites have a lysine or arginine at the −1 position (Demangeat et al., 1991; Hemmer et al., 1995). Subgroup A nepovirus proteinases cleave after a cysteine, arginine or glycine (Blokh et al., 1992; Buckley et al., 1993; Margis et al., 1993; Pinck et al., 1991; Serghini et al., 1990).
Nepoviruses of different subgroups have similar but distinct genomic organizations. The RNA2-encoded polyprotein (P2) of all nepoviruses contains the domains for the movement protein (MP) and CP at its C-terminus. However, while the N-terminal region of subgroup A and B nepovirus P2 polyproteins contains a single protein domain upstream of the MP (Demangeat et al., 1991; Margis et al., 1993), the P2 polyprotein of *Tomato ringspot virus* (ToRSV, a subgroup C nepovirus), contains two protein domains in addition to the MP and CP (Carrier et al., 2001). The RNA1-encoded polyprotein (P1) of all nepoviruses includes in its C-terminal region the domains for the putative nucleotide triphosphate-binding protein (NTB), the small viral protein linked to the genome (VPg), the Pro and the RNA-dependent RNA polymerase (Pol) (see Fig. 1). Furthermore, the arrangement of these protein domains within the polyprotein is conserved in all picorna-like viruses (Le Gall et al., 2007). The N-terminal region of nepovirus P1 polyproteins is less characterized. In the case of ToRSV, *in vitro* processing studies have identified two cleavage sites upstream of the 66 kDa NTB domain (Wang and Sanfacon, 2000). These cleavage sites define the 45 kDa X1 protein and the 22 kDa X2 protein. The ToRSV X2 protein shares many similarities with the comovirus 32 kDa protein. Both proteins are highly hydrophobic and possess a conserved amino acid motif [F-x(28)-W-x(11)-L-x(23)-E] (Zhang and Sanfacon, 2006). The ToRSV X1 protein does not have a functional equivalent in the comovirus genome. Amino acid sequences highly similar to the ToRSV X2 protein and comovirus 32 kDa protein are also found in the N-terminal region of the P1 polyproteins of subgroup A and B nepoviruses (Zhang and Sanfacon, 2006) (see Fig. 1). A cleavage site was identified upstream of the conserved NTB motif in the P1 polyproteins from *Beet ringspot virus* (BRSV, a subgroup B virus previously referred to as *Tomato black ring virus*) (Demangeat et al., 1990; Le Gall et al., 2005) and *Grapevine fanleaf virus* (GFLV, a subgroup A virus) (Margis et al., 1994). However, the position of these cleavages relative to the conserved NTB and X2 motifs differs (Fig. 1). Cleavage in the BRSV P1 polyprotein was detected between the X2 and NTB motifs at a position corresponding to the ToRSV X2-NTB site (Demangeat et al., 1990). In the case of GFLV, the observed cleavage was upstream of the X2 motif, at a position corresponding to the ToRSV X1–X2 site (Margis et al., 1994). The possibility that additional cleavage sites exist in the N-terminal region of subgroup A and B nepovirus P1 polyproteins has not been rigorously examined.

In this study, we have examined the proteolytic processing of the P1 polyprotein of *Arabis mosaic virus* (ArMV), a subgroup A nepovirus closely related to GFLV (Wetzel et al., 2004, 2001). Using *in vitro* processing assays, we have identified two cleavage sites upstream of the NTB domain at positions equivalent to the ToRSV X1–X2 and X2-NTB cleavage sites. Amino acid sequence alignments also implied similar cleavage sites in the polyproteins of other nepoviruses, suggesting that the presence of two protein domains upstream of NTB is a common feature of nepoviruses.

**Results**

Detection of processing events at the NTB-VPg and VPg-Pro cleavage sites

To study processing events in the ArMV P1 polyprotein, we used previously described *in vitro* processing assays (Wang et al., 1999; Wang and Sanfacon, 2000). We first tested a small truncated P1 polyprotein to confirm processing at the NTB-VPg and VPg-Pro sites. The precursor was designed to include the putative 66 kDa NTB domain, as well as the VPg and Pro domains (NTB-Pro precursor, see Materials and methods and Fig. 2A). As a negative control, a mutation was introduced in the protease domain of the NTB-Pro precursor to replace the conserved histidine (H\(^{128}\)) of the catalytic triad for a serine (HS mutation, numbering from the N-terminus of the deduced amino acid sequence for the P1 polyprotein according to Wetzel et al., 2004). Previous studies with the GFLV and ToRSV proteases have shown that mutation of this conserved histidine eliminates...
their proteolytic activity (Hans and Sanfacon, 1995; Margis and Pinck, 1992).

After in vitro translation, the NTB-Pro<sup>HS</sup> mutant polyprotein appeared as a protein of approximately 95 kDa in SDS-polyacrylamide gels (SDS-PAGE, see Fig. 2B, lane 1). The apparent molecular mass of this protein is close to that expected for the full-length NTB-Pro polyprotein (100 kDa including an S-tag at the N-terminus and a His-tag at the C-terminus). Minor proteins of lower apparent molecular masses were also observed. These proteins were likely produced by translation initiation at an internal AUG or premature termination of the translation. The unprocessed wild-type (WT) NTB-Pro precursor was also 100 kDa, but after overnight incubation at 16 °C, additional proteins appeared with apparent molecular masses of 69 kDa, 66 kDa, 32 kDa and 31 kDa that likely corresponded to the NTB-VPg, NTB, VPg-Pro and Pro cleavage products, respectively (Fig. 2B, lane 2 and summary table in C). These proteins were not produced after overnight incubation of the NTB-Pro<sup>HS</sup> mutant polyprotein (Fig. 2B, lane 1). To verify the nature of the cleavage products, we took advantage of the S-tag, which was fused in frame with the N-terminus of the NTB-Pro polyprotein, and conducted pull-down assays (Fig. 2B). As expected, the pull-down fraction included the full-length polyprotein precursor. Several background proteins were also concentrated in this fraction suggesting that they included the S-tag and were probably produced by premature termination of translation. The 69 kDa and 66 kDa cleavage products were enriched in the pull-down fraction while the 32 kDa and 31 kDa cleavage products remained in the unbound fraction (Fig. 2B, lanes 5 and 8). Because the S-tag was fused to the N-terminus of the NTB domain, this result confirmed that the 69 kDa and 66 kDa proteins contain this domain and provided experimental support for the suggestion that they correspond to the NTB-VPg intermediate and to the mature NTB protein, respectively.

The above results suggested that the ArMV proteinase recognized both the NTB-VPg and the VPg-Pro cleavage sites. By analogy with characterized GFLV cleavage sites, putative cleavage sites between the NTB, VPg and Pro domains were previously proposed to correspond to dipeptide C<sup>1216</sup>S (for NTB-VPg) and either G<sup>1240</sup>D or R<sup>1239</sup>G (for VPg-Pro) (Wetzel et al., 2004). After processing of the WT NTB-Pro precursor, the intensity of the bands corresponding to the 66 kDa (NTB) and 32 kDa (putative VPg-Pro) cleavage products was two to three times stronger than that of bands corresponding to the 69 kDa (NTB-VPg) and 31 kDa (putative Pro) products as judged by quantification of the radioactivity contained in each band, taking into account the number of radiolabeled methionine incorporated in each protein (Fig. 2C and Materials and methods). This suggested that cleavage occurred predominantly between the NTB and VPg domains. To verify this, we deleted the C/S

---

**Fig. 2.** In vitro processing of the NTB-Pro polyprotein and identification of the NTB-VPg cleavage site. (A) Schematic representation of the processing of the ArMV NTB-Pro polyprotein. The polyprotein is represented at the top. An S-tag epitope (grey box) and a histidine tail (black circle) fused in frame with the N- and C-termini of the polyprotein are shown. Vertical lines represent cleavage sites present in the precursor with the predicted C/S dipeptide indicated above the NTB-VPg cleavage site. (B) In vitro processing assays. The wild-type (WT) NTB-Pro protein and mutant derivatives NTB-Pro<sup>HS</sup> (mutation in the catalytic triad of the proteinase) and NTB-Pro<sup>ΔCS</sup> (deletion of the C/S dipeptide at the junction between the NTB and VPg domains) were synthesized in vitro and incubated at 16 °C overnight to allow for proteolytic processing of the polyprotein. After processing, S-tag pull-down assays were conducted as described in Materials and methods and proteins present in the starting reactions (in vitro processing), the immunoprecipitated fraction (S-tag pull-down) or the supernatant after immunoprecipitation (unbound) were separated by SDS-PAGE as indicated above each lane. The precursor polyprotein and cleavage products are shown. These molecular masses were calculated taking into account the non-viral sequences fused in frame with the proteins. (C) Summary of the results and deduced identity of the cleavage products. Presence (+) or absence (−) of each cleavage product in the wild-type (WT) or mutated (ΔCS) NTB-Pro processing reactions is indicated. The ability of each cleavage product to be immunoprecipitated by the S-tag antibody (S-Tag column) is also shown. The number of methionines (N<sub>met</sub>) present in each cleavage product is indicated. Relative intensity of each cleavage product was estimated by calculating the number of digital light units (DLU) present in each band, corrected for the number of methionines (see Materials and methods) and expressed as follows: 7–10 (+++++), 3–4 (+++), 0.5–1 (+), and less than 0.05 (+/-).
dipeptide at the predicted NTB-VPg cleavage site. In vitro processing of the NTB-ProΔCS mutant resulted in the release of the 69 kDa and 31 kDa cleavage products (Fig. 2B, lane 3). The 66 kDa and 32 kDa products were not detected. This result suggested that cleavage at the junction between the NTB and VPg domains occurs at or near the predicted C1216/S dipeptide. However, our experiments do not exclude the possibility that deletion of dipeptide C1216/S induces conformational changes in the polyprotein that could affect cleavage at a different dipeptide.

Identification of a sub-optimal cleavage event at the junction between the X2 and NTB domains

To identify cleavage events upstream of the NTB domain, we designed additional truncated P1 polyproteins of increasing length. The X2-Pro precursor was designed to include sequences corresponding to the ToRSV X2 and NTB domains (Fig. 3A). Although a cleavage site was previously identified in the GFLV P1 polyprotein at a position immediately upstream of the putative X2 domain, cleavage at the junction between the putative X2 and NTB domains had not been tested (Margis et al., 1994).

As above, the H1283 to S mutation was introduced in the proteinase domain of the X2-Pro polyprotein as a negative control. The X2-ProHS full-length polyprotein migrated as a protein with an apparent molecular mass of 115 kDa in SDS-PAGE (expected molecular mass is 120 kDa, Fig. 3B, lane 1). As with the NTB-ProHS polyprotein, several background bands were also observed in addition to the full-length polyprotein. After overnight processing of the WT X2-Pro precursor at 16 °C, several additional bands appeared that were not present after incubation of the X2-ProHS mutant polyprotein under the same conditions (compare lane 2 to lane 1). Two major proteins of approximately 85 kDa and 32 kDa and three minor proteins of approximately 95 kDa, 62 kDa and 23 kDa were observed. The 95 kDa and 66 kDa proteins were difficult to resolve due to the presence of many background bands of similar size but were more easily visualized when the cleavage products were separated by SDS-PAGE using a lower concentration of acrylamide (8%, Fig. 3B lower panel). Based on their apparent molecular masses, we speculated that the predominant 85 kDa and 32 kDa proteins corresponded to the X2-NTB and VPg-Pro cleavage products (see Fig. 2A for the calculated molecular mass of each cleavage product). The 95 kDa, 62 kDa and 23 kDa proteins may correspond to the NTB-Pro intermediate, the mature NTB protein and the mature X2 protein, respectively.

S-tag pull-down experiments resulted in the concentration of the 23 kDa and 85 kDa cleavage products in the pull-down fraction while the 95 kDa, 66 kDa and 32 kDa cleavage products remained in the unbound fraction (Fig. 2B, lanes 7 and 12). Because the S-tag is fused to the N-terminus of the X2 domain, this result provided experimental support for the suggestion that the 85 kDa and 23 kDa proteins correspond to the X2-NTB intermediate and to the mature X2, respectively.

Fig. 3. In vitro processing of the X2-Pro polyprotein and identification of the X2-NTB cleavage site. (A) Schematic representation of the processing of the ArMV X2-Pro polyprotein. The polyprotein is shown at the top. As in Fig. 2, an S-tag epitope (grey box) and a histidine tail (black circle) are indicated. Vertical lines represent cleavage sites. The C/S dipeptide predicted as the NTB-VPg cleavage site is shown. Two possible cleavage sites (dipeptides G/V and R/V) are indicated at the junction between the X2 and NTB domains. Cleavage products observed after processing of the polyprotein are represented below the arrow. The calculated molecular masses of the precursor polyprotein and cleavage products are shown. (B) In vitro processing assays. The wild-type (WT) X2-Pro protein and mutant derivatives X2-ProΔCS (mutation in the catalytic triad of the proteinase), X2-ProΔCS (deletion of the C/S dipeptide), X2-ProΔRV and X2-ProΔGV (deletion of the R/V or G/V dipeptides at the junction between the X2 and NTB domains) were tested using in vitro processing assays as described in the legend of Fig. 2. After processing, S-tag pull-down assays were conducted as described in Materials and methods and proteins present in the starting reactions (in vitro processing), the immunoprecipitated fraction (S-tag pull-down) or the supernatant after immunoprecipitation (unbound) were separated by SDS-PAGE as indicated above each lane. Smaller cleavage products were visualized after separation on a 12% polyacrylamide gel (top panel). Larger intermediate polyproteins were also separated on an 8% polyacrylamide gel (bottom panel). Bands corresponding to cleavage products are identified by the symbols on the left side of each lane. The migration of molecular mass markers (in kDa) is indicated at the left of the gel. (C) Summary of the results and deduced identity of the cleavage products are as explained in the legend of Fig. 2.
To further confirm the identity of the cleavage products, we conducted site-directed mutagenesis of predicted cleavage sites as above. Deletion of the C1216/S dipeptide between the NTB and VPg domains prevented the release of the 85 kDa and 32 kDa products (see X2-ProΔCS mutant, Fig. 3B, lane 3, and C). This result confirmed that the 85 kDa and 32 kDa proteins correspond to the X2-NTB and VPg-Pro cleavage products after cleavage at the NTB-VPg site. The 62 kDa cleavage product was also not observed in the cleavage products of the mutated X2-ProΔCS polyprotein (this band appears as a faint band below one of the background proteins; see Fig. 3B, lower panel). This result was in agreement with the suggestion that this protein corresponds to the mature NTB domain, which could not be released after mutation of the NTB-VPg cleavage site.

Detection of the X2, putative NTB and putative NTB-Pro cleavage products (23 kDa, 62 kDa and 95 kDa proteins, respectively) suggested that a cleavage occurred between the predicted X2 and NTB domains. Examination of the deduced amino acid sequence did not reveal the presence of dipeptides previously identified or suggested for ArMV or other subgroup A nepovirus cleavage sites (R/G, C/A, C/S, G/E). However, two alternative dipeptides were identified that may function as possible cleavage sites (G604/V and R618/V). Deletion of R618/V did not alter the processing of the X2-Pro precursor (Fig. 3B, lanes 4 and 9). In contrast, deletion of G604/V prevented the release of the 23 kDa cleavage product (X2, lanes 5 and 10). The 62 kDa (putative NTB) and 95 kDa (putative NTB-VPg-Pro) cleavage products were not detected, although it should be noted that the presence of many background bands obscured somewhat the resolution of these proteins. As expected, release of the 85 kDa (X2-NTB) and 32 kDa (putative Vpg-Pro) proteins was not affected by this mutation. This result confirms that a proteolytic cleavage occurs between the X2 and NTB domains. The amount of the 23 kDa (X2) and 62 kDa (putative NTB) proteins in the processing products was ten times less than that of the 85 kDa (X2-NTB) and 32 kDa (putative Vpg-Pro) proteins as determined by quantification of the radioactivity in each band, see Fig. 2C. This suggested that cleavage at the X2-NTB site is inefficient. Based on the result from the site-directed mutagenesis experiments, we tentatively propose dipeptide G604/V as the X2-NTB cleavage site.

Identification of a cleavage site between the X1 and X2 domains

A cleavage site was previously identified immediately upstream of the putative X2 domain in the N-terminal region of the GFLV P1 polyprotein (Margis et al., 1994). Dipeptide C416/A was proposed as a putative cleavage site (Margis et al., 1994). In ArMV, the corresponding dipeptide C414/G was suggested as a putative cleavage site (Wetzel et al., 2004). To determine whether the ArMV proteinase could recognize this proposed cleavage site,

---

**Fig. 4. In vitro processing of the X1-Pro polyprotein and identification of the X1–X2 cleavage site.** (A) Schematic representation of the processing of the ArMV X1-Pro polyprotein. The polyprotein is shown at the top. As in Fig. 2, an S-tag epitope (grey box) and a histidine tail (black circle) are indicated. Vertical lines represent cleavage sites. The C/S and C/G dipeptides predicted as the NTB-VPg and X1–X2 cleavage sites are shown. Cleavage products observed after processing of the polyprotein are shown below the arrow with their calculated molecular masses. (B) In vitro processing assays. The wild-type (WT) X1-Pro polyprotein and mutant derivatives X1-ProHS (mutation in the catalytic triad of the proteinase), X1-ProΔCS (deletion of the C/S dipeptide), X1-ProΔCG (deletion of the C/G dipeptide) and X1-ProΔCS+CG (deletion of both the C/S and C/G dipeptides) polyproteins were tested using in vitro processing assays as described in the legend of Fig. 2. After processing, S-tag pull-down assays were conducted as described in Materials and methods and proteins present in the starting reactions (in vitro processing), the immunoprecipitated fraction (S-tag pull-down) or the supernatant after immunoprecipitation (unbound) were separated by SDS-PAGE as indicated above each lane. Cleavage products were visualized after separation on a 12% polyacrylamide gel (top panel). The 20 kDa cleavage product was also separated using a 15% polyacrylamide gel (bottom panel). Bands corresponding to cleavage products are identified by the symbols on the left side of each lane. The migration of molecular mass markers (in kDa) is indicated at the left of the gel. (C) Summary of the results and deduced identity of the cleavage products are as explained in the legend to Fig. 2. NT: not tested.
we designed a new precursor that included the entire N-terminal region of the ArMV P1 polyprotein (precursor X1-Pro, Fig. 4A). As above, the HS mutation was introduced in the proteinase domain to create mutant X1-Pro<sup>HS</sup>. The results are shown in Fig. 4B and summarized in Fig. 4C.

Processing of the WT X1-Pro precursor resulted in the release of multiple cleavage products with apparent molecular masses of 116 kDa, 95 kDa, 85 kDa, 64 kDa, 55 kDa, 32 kDa and 20 kDa (Fig. 4B, lane 2). The apparent molecular masses of these proteins were close to those calculated for the X2-Pro (116 kDa), NTB-Pro (94 kDa), X2-NTB (88 kDa), NTB (66 kDa), X1 (50 kDa) and VPg-Pro (28 kDa) cleavage products. S-tag pull-down assays resulted in the concentration of the 55 kDa protein, confirming that it corresponds to the N-terminal X1 cleavage product (Fig. 4B, lane 9).

The 20 kDa protein (putative X2) was difficult to detect in the cleavage products from the WT X1-Pro polyprotein. To improve the detection of this protein and confirm its identity, an HA tag was inserted at the N-terminus of the predicted X2 domain; i.e., three amino acids downstream of the proposed X1–X2 cleavage site (mutant X1-Pro<sup>N3-X2-HA</sup>, Fig. 5A). After processing of the HA-tagged polyprotein, immunoprecipitation experiments were conducted using commercially available anti-HA antibodies (see Materials and methods). The 20 kDa protein (which migrates as a 21 kDa protein after fusion to the HA tag) was concentrated in the HA pull-down fraction (Fig. 5B, lane 5), confirming that it is the mature X2. The 116 kDa and 85 kDa proteins were also present in the HA pull-down fraction. This result was in agreement with the suggestion that these proteins correspond to the X2-Pro and X2-NTB cleavage products, respectively. We also inserted an HA tag between the G<sup>604</sup>/V and R<sup>618</sup>/V dipeptides at the junction between the X2 and NTB domains (mutant X1-Pro<sup>N4-NTB-HA</sup>, Fig. 5A). Based on the suggestion that G<sup>604</sup>/V corresponds to the X2-NTB cleavage site (see above, Fig. 3), the tag was predicted to be inserted at the N-terminus of the NTB domain. Indeed, the 64 kDa protein (putative NTB) was concentrated in the pull-down fraction after immunoprecipitation of the cleavage products of X1-Pro<sup>N4-NTB-HA</sup> (Fig. 5B, lane 6). The 116 kDa, 95 kDa and 85 kDa proteins were also found in the HA pull-down fraction, validating their deduced identity as the X2-Pro, NTB-Pro and X2-NTB intermediates, respectively. This result also confirms that cleavage between the X2 and NTB domains occurs at or near the predicted G<sup>604</sup>/V dipeptide.

To further analyze the processing of the X1-Pro precursor, specific deletions of dipeptides C<sup>1216</sup>/S (proposed NTB-VPg cleavage site) or C<sup>414</sup>/G (putative X1–X2 cleavage site) were introduced either individually (mutants ΔCS and ΔCG) or in combination (mutant ΔCS+CG). Deletion of C<sup>1216</sup>/S prevented the release of the 20 kDa (X2), 32 kDa (putative VPg-Pro), 64 kDa (NTB) and 82 kDa (X2-NTB) cleavage products, while deletion of C<sup>414</sup>/G eliminated the production of the 116 kDa (X2-Pro), 85 kDa (X2-NTB), 55 kDa (X1) and 20 kDa (X2) proteins (Fig. 4B, lanes 1–2). After simultaneous deletion of both proposed cleavage sites the 95 kDa protein (NTB-Pro) was the only cleavage product detected (Fig. 4B, lane 3). These results were consistent with the deduced identity of each cleavage product (Fig. 4C) and supported the suggestion that cleavage between the X1 and X2 domains occurs at or near dipeptide C<sup>414</sup>/G.

Taken together, our results suggest that the X1-Pro polyprotein is predominantly cleaved at the X1–X2 and NTB-VPg
sites. The accumulation of large amounts of the X2-NTB intermediate polyprotein and the low concentration of the X2 and NTB cleavage products as determined by quantification of the radioactivity present in each cleavage product (Fig. 4C) also confirmed that the X2-NTB cleavage site is a sub-optimal cleavage site.

**Ability of various P1 cleavage sites to be processed intermolecularly**

The X1–X2 cleavage site of GFLV was previously shown to be recognized *in trans* by the VPg-Pro precursor form of the proteinase (Margis et al., 1994). In contrast, the GFLV Pro-Pol site is cleaved only *in cis* (Margis et al., 1991). In ToRSV, none of the P1 truncated precursors tested (including precursors that contained the X1–X2 site) were recognized *in trans* by the ToRSV Pro or VPg-Pro (Chisholm et al., 2001; Wang and Sanfacon, 2000).

To determine if ArMV P1 cleavage sites could be recognized *in trans* by the proteinase, the radiolabeled mutated version of the X1-Pro polypeptide precursor containing an inactive proteinase domain (HS mutant) was supplemented with processing products derived after processing of *in vitro* translated non-labeled NTB-Pro polyprotein and incubated at 16 °C overnight to allow for processing to occur. In addition to the proteinase domain present in NTB-Pro, two other forms of the proteinase (VPg-Pro and Pro) are released from this polyprotein (see Fig. 2). Thus, this assay would allow us to simultaneously test if the proteinase domain contained in NTB-Pro, VPg-Pro or Pro could recognize *in trans* the cleavage sites present in the X1-ProHS polyprotein. As a negative control, we used a mutated version of NTB-Pro containing the inactive proteinase domain (HS mutant). As expected, proteolytic processing of the X1-ProHS polyprotein was not observed when it was incubated alone or in combination with the NTB-ProHS polyprotein (Fig. 6, lanes 2 and 5). In contrast, the X1 and X2-Pro cleavage products were released from the X1-ProHS polyprotein after incubation with the processing products derived from the wild-type NTB-Pro (lane 3). Other possible cleavage products of the X1-ProHS polyprotein (e.g. X2-NTB or VPg-Pro) were not detected in the *trans*-processing assays. The X1-ProHS polyprotein was also incubated with the processing products of the NTB-ProΔCS mutant polyprotein. The efficiency of release of the X1 and X2-Pro cleavage products was similar after overnight incubation of X1-ProHS with the processing products derived from the WT or mutated (ΔCS) versions of the NTB-Pro polyprotein. We have shown above that VPg-Pro is not released from NTB-ProΔCS. These results suggest that either the mature proteinase (released in small amounts from the NTB-ProΔCS and NTB-Pro polyproteins) and/or the proteinase domain present in the unprocessed NTB-Pro are responsible for *trans*-cleavage of X1-ProHS at the X1–X2 cleavage site.

**Prediction of proteinase cleavage sites in the N-terminal region of the P1 polyprotein from other nepoviruses**

The results presented above suggest that the ArMV proteinase recognizes two cleavage sites upstream of the NTB domain, at least *in vitro*. The position of these cleavage sites was similar to those previously reported for the ToRSV P1 polyprotein (Wang and Sanfacon, 2000) and delineates two protein domains upstream of NTB. Using the ClustalW program, we aligned the deduced amino acid sequences of the entire P1 polyprotein of nepoviruses currently available in the database. The quality of the alignment was verified by ensuring that previously identified conserved amino acids in the X2, NTB, Pro and Pol domains were properly aligned (Fig. 7 and data not shown). The portions of the alignment corresponding to regions flanking the putative X1–X2 and X2-NTB cleavage sites are shown in Fig. 7. Cleavage sites were deduced based on the known specificity of the proteinases from viruses of the different subgroups. The previously identified or suggested ToRSV X1–X2 and X2-NTB sites, GFLV X1–X2 site, BRSV X2-NTB site and ArMV X1–X2 and X2-NTB sites (Demanget et al., 1990; Margis et al., 1994; Wang and Sanfacon, 2000 and this study) were also used as a guide to predict other cleavage sites.

Dipeptides R/A or R/S were predicted as putative cleavage sites in the P1 polyprotein of BRSV, *Tomato black ring virus,*
Grapevine chrome mosaic virus or Cycas necrosis stunt virus (subgroup B nepoviruses, see Fig. 7), which is in agreement with previously identified subgroup B nepovirus cleavage sites that have a lysine or arginine at the \(-1\) position (Demangeat et al., 1991; Hemmer et al., 1995). In the case of Blackcurrant reversion virus and Peach rosette mosaic virus, two subgroup C nepoviruses, dipeptides N/G, D/G, N/A or D/L were predicted as possible cleavage sites. Processing after a glutamine, asparagine or aspartate was previously reported for this subgroup of nepoviruses (Bacher et al., 1994; Carrier et al., 1999; Latvala et al., 1998). Finally, for GFLV, Tobacco ringspot virus and Raspberry ringspot virus, three subgroup A nepoviruses, dipeptides C/G, C/A, C/S, G/V, G/A or G/L were predicted as possible cleavage sites based on the previous identification of subgroup A nepovirus cleavage sites with a cysteine, arginine or glycine at the \(-1\) position (Blok et al., 1992; Buckley et al., 1993; Margis et al., 1993; Pinck et al., 1991; Serghini et al., 1990). In all cases, putative cleavage sites consistent with the specificity of the corresponding protease were found at positions equivalent to that of the X1–X2 and X2-NTB cleavage sites from the ToRSV and ArMV P1 polyproteins, suggesting that the presence of two cleavage sites upstream of the NTB domain is conserved among nepoviruses.

Discussion

In this study, we have analyzed the processing of the N-terminal region of the ArMV P1 polyprotein. Four cleavage sites were shown to be recognized by the ArMV protease. Based on site-directed mutagenesis of the X2 region of the P1 polyprotein of nepoviruses and prediction of X1–X2 and X2-NTB cleavage sites. Amino acid sequences from the P1 polyprotein of nepoviruses were aligned using the ClustalW program as described in Materials and methods. Two regions of the alignment are shown. The upper portion of the figure includes the N-terminal region of X2 and the putative X1–X2 cleavage site. The lower portion of the figure includes the N-terminal region of NTB and the putative X2-NTB cleavage site. A boxed phenylalanine (F) within the X2 domain corresponds to the first consensus amino acid of the motif conserved in nepoviruses and comoviruses (Zhang and Sanfacon, 2006). Predicted cleavage sites are underlined. Cleavage sites supported by experimental evidence (mutagenesis) are also shown in bold italics. In the case of the TRSV X2-NTB cleavage site, two possible dipeptides are underlined.
plants, the mature NTB protein and NTB-VPg intermediate polyprotein are found in association with endoplasmic reticulum-derived membranes, confirming that cleavage between the X2 and NTB domains occurs in vivo (Han and Sanfacon, 2003). However, small amounts of an intermediate precursor of the expected molecular mass for the X2-NTB-VPg protein are also consistently detected in association with ER membranes in infected cells, suggesting that cleavage at the junction between the X2 and NTB domains is incomplete (Chisholm et al., 2007; Han and Sanfacon, 2003). Further experiments will be necessary to study the efficiency of cleavage at the ArMV X2-NTB site in vivo and to compare the relative accumulation of the X2-NTB, X2 and NTB proteins in infected plants.

Nepovirus X2 proteins share many common properties with the 32 kDa protein of Cowpea mosaic virus (CPMV, a comovirus) including having a very hydrophobic nature and a signature amino acid motif (Zhang and Sanfacon, 2006). The ToRSV X2 protein and the CPMV 32 kDa protein both have the ability to interact directly with endoplasmic reticulum (ER) membranes and have been suggested to play a role in anchoring the viral replication complex into the membranes (Carette et al., 2002; Zhang and Sanfacon, 2006). The NTB proteins of ToRSV and CPMV are also associated with ER-derived membranes and are probably also involved in the assembly of the replication complex (Carette et al., 2002; Han and Sanfacon, 2003; Wang et al., 2004; Zhang et al., 2005). The accumulation of intermediate polyproteins that contain both the X2 and NTB domains may allow coordination of the function of these proteins in viral replication complex assembly. It is interesting to note that in the case of CPMV, although cleavage between the 32 kDa protein and the NTB domain is very efficient in vivo and in vitro, the two proteins interact with each other and remain associated after proteolytic cleavage (Peters et al., 1992).

The CPMV 32 kDa protein has been shown to modulate the activity of the proteinase (Peters et al., 1992). It has been proposed that an interaction between the 32 kDa protein and the NTB domain present within the NTB-VPg-Pro-Pol polyprotein alters the conformation of this polyprotein and renders cleavage sites less accessible to recognition by the proteinase. However, there is no evidence that the X2 protein of nepoviruses play a similar role. Indeed, the presence or absence of the X2 domain on truncated P1 polyproteins from ArMV or ToRSV does not significantly influence the efficiency of processing at the NTB-VPg site (Wang and Sanfacon, 2000 and this study).

Our results indicate that the ArMV X1–X2 cleavage site is recognized very efficiently by the proteinase, at least in vitro. Indeed, precursors that include both the X1 and X2 domains are not detected in significant amounts in the cleavage products of the X1-Pro polyprotein. In contrast, ToRSV intermediate polyproteins containing the X1 and X2 domains were consistently detected in the cleavage products of a truncated P1 precursor that included the C-terminal portion of X1 (cX1-Pro precursor, Wang and Sanfacon, 2000). Another notable difference is that while the ArMV and GFLV X1–X2 cleavage sites can be recognized in trans by one or several forms of the proteinase (Margis et al., 1994 and this study), trans-cleavage at the ToRSV X1–X2 cleavage site by Pro or VPg-Pro was not detected (Chisholm et al., 2001; Wang and Sanfacon, 2000). One possible explanation for these differences is that the absence of the N-terminal portion of X1 on the truncated ToRSV cX1-Pro precursor influences the folding of the polyprotein and prevents optimal presentation of the X1–X2 cleavage site for recognition by the proteinase. Alternatively, it is also possible that regulation of cleavage at the X1–X2 site differs among nepoviruses.

While X2 is related to the comovirus 32 kDa protein, the X1 protein domain is unique to nepoviruses within the family Comoviridae. This protein domain is not conserved among nepoviruses and amino acid alignments of this region of the polyprotein are difficult to validate because of the lack of conserved amino acid sequence motifs (data not shown). Further studies will be necessary to decipher the biological function of this protein in the replication cycle of nepoviruses and determine why it is absent from the genome of the related comoviruses and fabaviruses.

In conclusion, our results suggest that the presence of two protein domains upstream of the NTB domain is a distinguishing common feature of nepoviruses, although the specific regulation of proteolytic cleavage at these sites may differ from one nepovirus to another. The genus Nepovirus constitutes an interesting group of viruses that share many similar properties but are at the same time very diverse (Le Gall et al., 2005). Nepoviruses have been separated into three subgroups based on the length and genomic organization of RNA 2, the cleavage site specificity of the proteinase and phylogenetic analysis using the coat protein sequence (Le Gall et al., 2005, 2007). Based on these distinct characteristics, it could be argued that nepoviruses should be separated into three distinct genera. However, the three nepovirus subgroups do not separate in distinct clades in phylogenetic studies using the Pro-Pol region (Le Gall et al., 2007). A distinct and unifying property of members of the genus Nepovirus is the presence of a single large CP subunit. The results presented here suggest that the genomic organization of RNA 1 is another common property of nepoviruses which is shared by members of all three subgroups.

Materials and methods

Construction of plasmids

Plasmids pCITE-NTB-Pro, pCITE-X2-Pro and pCITE-X1-Pro were constructed by amplifying the corresponding region of the ArMV genome from a full-length cDNA clone of a grapevine isolate of ArMV (Wetzel et al., 2004 and unpublished) using the Phusion High-Fidelity PCR kit (Finnzymes, distributed by New England Biolabs). The following primer pairs were used for the amplification: for plasmid pCITE-NTB-Pro, primer Pro-as [5′ acgc gatctctcacaat gaggagctataac 3′], complementary to nts 4064–3982 of ArMV RNA1 and including a SalI restriction site (underlined) and primer NTB-s [5′ acgc gatctctagcatgtgcaattttctgaggg 3′], corresponding to nts 2082–2102 of ArMV RNA1 and including an EcoRV restriction site; for plasmid pCITE-X2-Pro, primers Pro-as and X2-s [5′ acgc gatctctagcatgtgcaattttctgaggg 3′], corresponding to nts 1470–1492 of ArMV RNA1 and including an EcoRV restriction site; and for plasmid pCITE-X1-Pro, primers Pro-as and X1-s [5′ acgc gatctctagcatgtgcaattttctgaggg 3′], corresponding to nts 1470–1492 of ArMV RNA1 and including an EcoRV restriction site.
3’, corresponding to nts 228–247 of ArMV RNA1 and including an EcoRV restriction site]. The amplified fragments were digested by EcoRV and Sall and inserted into the corresponding sites of plasmid pCITE-4a(+) (Novagen). Mutation, deletion or insertion of amino acids was produced by site-directed mutagenesis using specific primers, as described (Urban et al., 1997). All plasmids were sequenced prior to analysis to verify that the desired mutations were correctly introduced.

**In vitro processing assays and immunoprecipitations**

Coupled in vitro transcription and translation reactions and in vitro processing assays were conducted as described previously for the processing of ToRSV polyproteins (Wang et al., 1999). Briefly, in vitro translations were conducted for 90 min at 30 °C in the presence of [35S]-methionine using the TNT coupled rabbit reticulocyte system (Promega). After translation, RNase A’ and cold methionine were added and the samples were incubated at room temperature for 10 min. The translation products were then diluted 1:2 in 2× processing buffer (200 mM Tris–HCl pH 8.0, 2 mM DTT, 20% (v/v) glycerol) and incubated at 16 °C overnight to allow for proteolytic processing. For trans-processing assays, wild-type or mutated versions of the NTB-Pro polyprotein were first synthesized in vitro and processed as above but in the presence of cold methionine rather than radiolabeled methionine. The cold NTB-Pro processing products were then added to the radiolabeled X1-Pro[35S] polyprotein produced by in vitro translation in the presence of [35S]-methionine. The mixture was then incubated at 16 °C overnight to allow for processing of X1-Pro[35S]. Labeled proteins were separated by SDS-PAGE (Laemmli, 1970) and visualized by autoradiography. For quantification of the relative concentration of each cleavage product, gels were exposed to a phosphorimager screen and the radioactivity present was posed to a phosphorimager (Cyclone Plus, Perkin Elmer). The values (digital light units) corresponding to each band were readjusted to the relative concentration of each cleavage product, gels were separated by SDS-PAGE and visualized by autoradiography. For quantitative measurements, the phosphorimager signal was corrected for the area under the curve of each peak and the area of the control peak was set to 1.

**Amino acid alignments**

Amino acid sequences from the entire P1 polyprotein of nepoviruses were aligned using the ClustalW program (Chenna et al., 2003) and were manually readjusted to ensure alignments of previously identified conserved amino acids. The following accession numbers were used for the alignment: Peach rosette mosaic virus (PRMV, AAB69867), Blackcurrant reversion virus (BRV, AF3682772), ToRSV (DQ469829 for the X2 domain and L19655 for the rest of the P1 polyprotein), Tomato black ring virus (TBRV, AY157993), Grapevine chrome mosaic virus (GCMV, X15346), Cucurbit necrotic stunt virus (CNSV, AB073147), BRSV (D00322), Tobacco ringspot virus (TRSV, U50869), Raspberry ringspot virus (RpRSV, AY303787), ArMV (AY303786), and GFLV (D00915).

**Acknowledgments**

This research was supported by a grant from Agriculture and Agri-Food Canada and by a grant from the Klaus Tschira Stiftung in Germany, and was facilitated by a collaborative grant from the German International Bureau of the federal Ministry of Education and Research (Internationales Büro des Bundesministerium für Bildung und Forschung, CAN06/A06).

**References**


Han, S., Sanfacon, H., 2003. Tomato ringspot virus proteins containing the nucleoside triphosphate binding domain are transmembrane proteins that associate with the endoplasmic reticulum and cofractionate with replication complexes. J. Virol. 77, 523–534.


