Proteolytic processing at a novel cleavage site in the N-terminal region of the tomato ringspot nepovirus RNA-1-encoded polyprotein in vitro

Aiming Wang† and Hélène Sanfaçon²

¹ Department of Botany, The University of British Columbia, Vancouver, BC, Canada, V6T 1Z4
² Pacific Agri-Food Research Centre, Summerland, BC, Canada, V0H 1Z0

Tomato ringspot nepovirus RNA-1-encoded polyprotein (P1) contains the domains for the putative NTP-binding protein, VPg, 3C-like protease and a putative RNA-dependent RNA polymerase in its C-terminal region. The N-terminal region of P1, with a coding capacity for a protein (or a precursor) of 67 kDa, has not been characterized. Using partial cDNA clones, it is shown that the 3C-like protease can process the N-terminal region of P1 at a novel cleavage site in vitro, allowing the release of two proteins, X1 (located at the N terminus of P1) and X2 (located immediately upstream of the NTB domain). P1 precursors in which the protease was inactive or absent were not cleaved by exogenously added protease, suggesting that P1 processing was predominantly in cis. Results from site-directed mutagenesis of putative cleavage sites suggest that dipeptides Q423/G and Q620/G are the X1-X2 and X2-NTB cleavage sites, respectively. The putative X1 protein contains a previously identified alanine-rich sequence which is present in nepoviruses but not in the related comoviruses. The putative X2 protein contains a region with similarity to the comovirus 32 kDa protease co-factor (the only mature protein released from the N terminus of comovirus P1 polyproteins) and to the corresponding region of other nepovirus P1 polyproteins. These results raise the possibility that the presence of two distinct protein domains in the N-terminal part of the P1 polyprotein may be a common feature of nepoviruses.

Introduction

Tomato ringspot virus (ToRSV) is a member of the genus Nepivirus, family Comoviridae. Nepoviruses have been divided into three subgroups (Mayo & Robinson, 1996). ToRSV is the only member of subgroup C whose genome has been completely sequenced (Rott et al., 1991a, 1995). The ToRSV genome consists of two mRNAs which are polyadenylated at their 3′ ends and are covalently attached to a small virus-encoded protein (VPg) at their 5′ ends (Sanfaçon, 1995; Mayo & Robinson, 1996). RNA-1 is 8214 nt in length, excluding the 3′ poly(A) tail, and contains a single long open reading frame (ORF) of 6591 nt, beginning at the first AUG codon at nucleotide position 78 (Rott et al., 1995). Translation of this ORF would produce a polyprotein (P1) with a predicted molecular mass of 244 kDa, which is cleaved by a virus-encoded protease to release the mature proteins. The C-terminal portion of the ToRSV P1 polyprotein contains the domains for a putative NTP-binding protein (NTB), the VPg, the protease (Pro) and a putative RNA-dependent RNA polymerase (Pol; Rott et al., 1995), which are also present on the polyproteins encoded by related comoviruses (Argos et al., 1984; Goldbach, 1987). Protease cleavage sites at the junction of the NTB-VPg, VPg-Pro and Pro-Pol domains have been identified, thereby allowing the definition of the VPg, Pro and Pol domains (Wang et al., 1999).

The ToRSV protease is a serine-like protease, related to the 3C protease of picornaviruses and the 3C-like proteases of comoviruses and potyviruses (Hans & Sanfaçon, 1995). ToRSV cleavage sites characterized so far consist of Q/G or Q/S dipeptides (Hans & Sanfaçon, 1995; Wang et al., 1999; Carrier et al., 1999). Site-directed mutagenesis of two ToRSV cleavage
sites has confirmed that efficient processing by the protease requires the presence of a Q at position —1 and a S or a G at position +1 of the cleavage sites (Carrier et al., 1999).

The N-terminal domain of the ToRSV P1 polyprotein, with a total coding capacity for a protein or a precursor of 67 kDa, has not been characterized. This domain includes a region that is identical to the amino acid sequence of P2 and a region that is unique to P1 (Rott et al., 1991b). In this study, we present evidence for the presence of a novel cleavage site in the N terminus of the polyprotein that is recognized in vitro by the ToRSV protease. This would result in the release of two mature proteins from the N-terminal region of the ToRSV P1 (i.e., upstream of NTB). Sequence analysis revealed that the presence of two distinct protein domains in the N terminus of the P1 polyprotein may be a common feature of nepoviruses. In contrast, only one mature protein is released from the N-terminal region of the P1 polyprotein from comoviruses (Goldbach & Wellink, 1996).

Methods

Plasmid construction. Plasmid pMR10, a full-length cDNA clone of ToRSV RNA-1, has been described (Rott et al., 1995). To construct plasmid pTrxFus-X1, a 285 bp fragment of pMR10 (RNA-1, nt 1366–4534) was amplified using plasmid pTrxFus (Invitrogen). To create plasmid pT7-X1-Pro, a 3168 nt fragment of pMR10 (RNA-1, nt 1058–1343) was amplified using oligonucleotides W018 and W019 (Table 1). The amplified fragment was digested with KpnI and XhoI and inserted into the corresponding sites of plasmid pCITE-4a (Novagen). For the construction of plasmid pT7-X2-Pro\(^{11283\text{D}}\), the following three fragments were ligated: a large Ncol–BamHI fragment from plasmid pT7-X2-Pro, a 2050 nt Ncol–XhoI fragment from plasmid pT7-X2-Pro and a 1040 nt XhoI–BamHI fragment from plasmid pT7-NTB-VPg-Pro\(^{11283\text{D}}\) (Wang et al., 1999). To create plasmid pT7-X2-Pro\(^{\Delta\text{NS1}}\), a large DNA fragment was amplified using oligonucleotides W036 and W037 and plasmid pT7-X2-Pro served as a template. The amplified fragment was self-ligated. To construct plasmid pT7-X1-Pro (Fig. 1), a 3748 nt fragment of pMR10 (RNA-1, nt 786–4534) was amplified with oligonucleotides W034 and W023, digested with Ncol and Sall and ligated into the corresponding sites of plasmid pCITE-4a(+) . The three-piece ligation strategy described above for the construction of plasmid pT7-X2-Pro\(^{11283\text{D}}\) was used to produce plasmid pT7-X1-Pro\(^{\Delta\text{NS1}}\). Plasmid pT7-X1-Pro\(^{\Delta\text{NS1}}\) was constructed by ligation of the large Sall–Sall fragment of pT7-X1-Pro with the 2700 nt Sall–Sall fragment of pT7-X2-Pro\(^{\Delta\text{NS1}}\). Plasmids pT7-X1-Pro\(^{\Delta\text{NS1}}\) and pT7-X1-Pro\(^{\Delta\text{NS1}}\) were created by ligation of the 1900 nt EcoRI–Sall fragment of pT7-NTB-VPg-Pro\(^{\Delta\text{NS1}}\) (Wang et al., 1999) with the large EcoRI–Sall fragment of pT7-X1-Pro and pT7-X1-Pro\(^{\Delta\text{NS1}}\), respectively. Plasmid pT7-X1-Pro\(^{\Delta\text{NS1}}\) was obtained by self-ligation of a fragment amplified using oligonucleotides W074 and W075 with plasmid pT7-X1-Pro serving as a template. Plasmid pT7-X1-Pro\(^{\Delta\text{NS1}+\text{NS2}+\text{NS3}}\) was created in a similar manner using plasmid pT7-X1-Pro\(^{\Delta\text{NS1}+\text{NS2}+\text{NS3}}\) as a template for amplification. To construct plasmid pT7-X1-X2, a 1496 bp fragment of plasmid pMR10 (RNA-1, nt 441–1937) was amplified with oligonucleotides W027 and W029, digested with Ncol and Sall, and inserted into the corresponding sites of pCITE-4a(+) (Novagen). For the construction of plasmid pT7-NTB-Pro and pT7-NTB-VPg-Pro\(^{11283\text{D}}\) (Wang et al., 1999) to give

<table>
<thead>
<tr>
<th>Name</th>
<th>Polarity</th>
<th>Sequence (5’ to 3’)</th>
<th>Corresponding sequence of RNA-1</th>
<th>Incorporated restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>W018</td>
<td>+</td>
<td>GGCCTACAGCATGATCTCTGCGC</td>
<td>1058–1076</td>
<td>KpnI</td>
</tr>
<tr>
<td>W019</td>
<td>–</td>
<td>AGGCTCTAGCCTCTGAGGAGAAATGAGG</td>
<td>1343–1322</td>
<td>Xhol</td>
</tr>
<tr>
<td>W023</td>
<td>–</td>
<td>ACGCTGACCTGAAAGGACAAAAAG</td>
<td>4534–4518</td>
<td>Sall</td>
</tr>
<tr>
<td>W025</td>
<td>–</td>
<td>ACCGGTACATGATCTCTGAGGAGAAATGAGG</td>
<td>3794–3778</td>
<td>Sall</td>
</tr>
<tr>
<td>W026</td>
<td>–</td>
<td>ACCGGTACATGATCTCTGAGGAGAAATGAGG</td>
<td>1366–1383</td>
<td>Ncol</td>
</tr>
<tr>
<td>W027</td>
<td>–</td>
<td>ACCGGTACATGATCTCTGAGGAGAAATGAGG</td>
<td>1957–1922</td>
<td>Sall</td>
</tr>
<tr>
<td>W028</td>
<td>+</td>
<td>ACCGGATGAAATAGTCTGAGTGTGT</td>
<td>2948–2967</td>
<td>Ncol</td>
</tr>
<tr>
<td>W029</td>
<td>+</td>
<td>ACCGGATGAAATAGTCTGAGTGTGT</td>
<td>441–459</td>
<td>Ncol</td>
</tr>
<tr>
<td>W034</td>
<td>+</td>
<td>ACCGGATGAAATAGTCTGAGTGTGT</td>
<td>786–801</td>
<td>Ncol</td>
</tr>
<tr>
<td>W035</td>
<td>–</td>
<td>ACCGGTACAGCGAGCTACTCCGCT</td>
<td>6696–6682</td>
<td>Sall</td>
</tr>
<tr>
<td>W036</td>
<td>–</td>
<td>ACCGGTACAGCGAGCTACTCCGCT</td>
<td>1938–1956</td>
<td>Ncol</td>
</tr>
<tr>
<td>W037</td>
<td>–</td>
<td>ACCGGTACAGCGAGCTACTCCGCT</td>
<td>1934–1917</td>
<td>Ncol</td>
</tr>
<tr>
<td>W074</td>
<td>+</td>
<td>ACCGGTACAGCGAGCTACTCCGCT</td>
<td>1347–1369</td>
<td>Ncol</td>
</tr>
<tr>
<td>W075</td>
<td>–</td>
<td>ACCGGTACAGCGAGCTACTCCGCT</td>
<td>1343–1325</td>
<td>Ncol</td>
</tr>
</tbody>
</table>

* The sequences of engineered restriction sites are underlined.

NA. Not applicable.
plasmids pT7-NTB-Pol and pT7-NTB-Pol%H%+%+%. Plasmid pET-MPCAT was constructed by inserting the SacI–SalI fragment of pT7-MPCAT-AATG (see below) into the SacI–XhoI sites of pET21b (Novagen). Plasmid pT7-MPCAT-AATG was obtained by site-directed mutagenesis (Kunkel, 1985) using uridyl-single-stranded DNA from pT7-MPCAT (Hans & Sanfaçon, 1995) and oligonucleotide 34. The nucleotide sequence of all plasmids was verified by sequencing using the ABI PRISM Dye Terminator cycle sequencing reaction kit (Perkin Elmer) and an ABI PRISM 310 Genetic Analyser (Perkin Elmer).

Production of polyclonal antibodies against X1. Polyclonal antibodies against NTB and VPg and monoclonal antibodies against Pro were previously described (Wang et al., 1999). Polyclonal antibodies against the putative X1 domain were produced against the Trx-Fus-X1 fusion protein. Plasmid pTrxFus-X1 was transformed into E. coli strain GI 724 (Invitrogen) and expression of the recombinant protein was induced as described by the supplier. Purification of the inclusion bodies containing the X1 fusion protein, solubilization of the fusion protein with 7 M urea and renaturation of the protein by dialysis in PBS containing a decreasing concentration of urea were as described (Wang et al., 1999). The purified fusion protein was injected into mice and the polyclonal antibodies were harvested as described (Wang et al., 1999).

In vitro transcription and translation, immunoprecipitations of translation products and trans-processing assays using purified recombinant active protease. In vitro transcription and translation of cDNA clones were carried out using a TNT coupled transcription/translation system (Promega) in the presence of [35S]methionine at 30 °C for 2 h as described (Wang et al., 1999). To allow optimal proteolytic processing and to arrest the translation reaction, the samples were diluted 1:3 in processing buffer (100 mM Tris–HCl, pH 8.0; 10 µg/ml RNase A; 1 mM DTT; 10% (v/v) glycerol) and incubated at 16 °C overnight. Translation products were separated by SDS–PAGE (Laemmli, 1970) and visualized by autoradiography. Immunoprecipitations of in vitro translation products were conducted as previously described (Hans & Sanfaçon, 1995).

Recombinant active protease was purified from the expression products of plasmid pET15BPro-N-Pol as described (Wang et al., 1999). Upon expression, the protease was insoluble. Solubilization of the protease from the purified inclusion bodies with urea and renaturation of the protease by gradual dialysis was as described (Wang et al., 1999). Aliquots of purified protease were stored in 50 mM Tris–HCl, pH 8.0, 1 mM DTT and 10% (v/v) glycerol at −70 °C. Trans-processing assays were conducted by adding the purified protease diluted in the processing buffer (described above) to the translation products and incubating the samples at 16 °C overnight.

Results

Production of antibodies against the N-terminal part of the RNA-1-encoded polyprotein

To facilitate the analysis of proteolytic processing in the N-terminal region of the RNA-1-encoded polyprotein, antibodies were produced against a fusion protein containing viral sequences derived from this region. Examination of the deduced amino acid sequence of the N-terminal region of the polyprotein encoded by RNA-1 revealed the presence of only two potential cleavage sites corresponding to the criteria for ToRSV cleavage sites: dipeptide Q^230/G, immediately upstream of the NTB domain (predicted by Rott et al., 1995), and Q^223/G (Fig. 1a). Evidence is presented that cleavage at these sites occurs in vitro. Two potential protein domains were delineated by these putative cleavage sites. These potential proteins were arbitrarily called X1 and X2, starting at the N terminus of P1. The putative X1 protein contains a region of similarity to the RNA-2-encoded polyprotein at its N terminus and a region that is unique to P1 at its C terminus. Mouse polyclonal antibodies were raised against a fusion protein containing the C-terminal region of the putative X1 protein (aa E^327 to Q^423) fused to the thioredoxin protein (Fig. 1a). As expected, these anti-X1 antibodies reacted with the corresponding fusion protein in immunoblot experiments (Fig. 1b) and immunoprecipitated in vitro translation products con-

Fig. 1. Production of polyclonal antibodies against the N-terminal region of P1. (a) Schematic representation of the P1 domain contained in the fusion protein. The predicted genomic organization of ToRSV RNA-1 is shown. The VPg protein is indicated by a circle at the 5’-end and the predicted ORF is indicated by the initiation (AUG) and termination (UAG) codons. A region of similarity between the RNA-1- and RNA-2-encoded polyproteins is shown by the speckled area. The darker speckled area represents 100% amino acid identity while the lighter speckled area represents 50% amino acid identity. The putative functions of the mature products and their predicted molecular masses are shown inside and under the corresponding boxes, respectively. The determined cleavage sites are indicated in bold and by solid vertical lines. The predicted cleavage sites are shown by dashed vertical lines. The cDNA clone containing a portion of the N-terminal region of the RNA-1 ORF fused in-frame with the thioredoxin protein (Thio) is shown with the predicted molecular mass of the fusion protein. (b) Expression of the X1 fusion protein (pTrxFus-X1) in E. coli and specificity of antibodies raised against the X1 fusion protein. Proteins were separated by 12% SDS–PAGE and stained with Coomassie-blue R-250 (lanes 1 and 2) or transferred to PVDF membrane for immunodetection using antibodies raised against the fusion protein (lanes 3 and 4). Lanes 1–4 represent uninduced (–) and induced (+) extracts of E. coli transformed with plasmid pTrxFus-X1 (X1). Positions of molecular mass markers are indicated on the left of the gel and the expected position for the Fus-X1 protein is shown on the right.
taining this domain of P1 (see below). We have also attempted to produce antibodies against the putative X2 protein using this method. Unfortunately, we were unsuccessful in obtaining antibodies that could immunoprecipitate in vitro translation products containing the X2 domain.

Processing at the cleavage site immediately upstream of the NTB domain in vitro

To characterize the N-terminal region of P1, in vitro translations were conducted using polyprotein precursors (X1-Pro and X2-Pro) that included various regions of P1. The coding region for these precursors was inserted into vector pCITE-4a(+), as described in Methods, and includes the T7 promoter and translational enhancing sequences to optimize the efficiency of in vitro translation. The X2-Pro precursor, encoded by plasmid pT7-X2-Pro, contained the domains for NTB, VPg and Pro and a region N-terminal of NTB with a coding capacity of 21 kDa, corresponding to the putative X2 protein (Fig. 2a). This precursor was predicted to contain one potential cleavage site located immediately upstream of the NTB domain (X2-NTB cleavage site) in addition to two previously identified cleavage sites (NTB-VPg and VPg-Pro; Wang et al., 1999).

In vitro transcription/translation of plasmid pT7-X2-Pro resulted in the synthesis of a protein with an apparent molecular mass of approximately 120 kDa (10% SDS–PAGE), close to the predicted size for the entire precursor polyprotein.
Fig. 3. Processing of the X1-Pro precursor polypeptide in vitro. (a) Schematic representation of processing of the X1-Pro polypeptide. The calculated molecular masses of the precursors and cleaved proteins are shown. The predicted and identified cleavage sites (including the cleavage site between X1 and X2) are indicated by short dashed lines and solid vertical lines. VS, vector sequence. (b, c) Proteolytic processing of the X1-Pro precursor. Lanes 1–7: translation products of plasmid pT7-X1-Pro (WT) and of mutant derivatives [pT7-X1-Pro\textsuperscript{HD} (HD), pT7-X1-Pro\textsuperscript{\DeltaX1X2} \textsuperscript{\DeltaX2N} (\DeltaX1X2), pT7-X1-Pro\textsuperscript{\DeltaNV} \textsuperscript{\DeltaX2N} \textsuperscript{\DeltaNV} (\DeltaX2N) and pT7-X1-Pro\textsuperscript{\DeltaX1X2\DeltaX2N\DeltaNV} (\DeltaX2N\DeltaNV)] were incubated for 20 h at 16 °C and separated by 11% SDS–PAGE (b) and 7% SDS–PAGE (c). Lane 8: translation products of plasmid pT7-X2-Pro. Arrowheads indicate the position of the cleavage products on the gel. The relative molecular mass of these proteins, which in some cases differs slightly from the calculated molecular mass shown in (a), is indicated.

(Fig. 2b, lane 1). Following overnight incubation of this precursor at 16 °C, smaller proteins with apparent molecular masses of 88 kDa, 66 kDa, 33 kDa and 22 kDa accumulated (Fig. 2b, lanes 2–4). These proteins did not appear following incubation of an equivalent precursor containing a mutation in the putative catalytic triad of the protease (HD mutation, histidine 1283 replaced with asparagine, numbering from the putative first amino acid of P1; Fig. 2c, lanes 5 and 6). This mutation was previously found to abolish protease activity (Hans & Sanfaçon, 1995; Wang et al., 1999). These results indicated that the small proteins were produced by proteolytic cleavage of the wild-type precursor polypeptide by the ToRSV protease contained in the precursor. The 88 kDa, 66 kDa and 33 kDa proteins correspond approximately to the calculated molecular masses for the X2-NTB intermediate (90 kDa), the mature NTB protein (66 kDa) and the VPg-Pro intermediate (32 kDa). The nature of these proteins was confirmed by immunoprecipitation experiments (Fig. 2b, lanes 5–8) using antibodies raised against the NTB, VPg and Pro proteins. The 22 kDa protein probably corresponded to the region upstream of NTB (calculated molecular mass of 24 kDa; 21 kDa of viral sequence and 3 kDa of vector sequence). The nature of this protein could not be confirmed, as antibodies against the X2 domain were not available for immunoprecipitation experiments. This protein accumulated at low levels, possibly due to its intrinsic instability. Taken together, these results suggested that proteolytic processing occurred at the NTB-VPg and X2-NTB cleavage sites. Processing at the VPg-Pro cleavage site was very inefficient (or absent) in this precursor. This was similar to previous results using a precursor containing the NTB-VPg-Pro domain (Wang et al., 1999).

As mentioned above, dipeptide Q\textsuperscript{G} located immediately upstream of NTB (Rott et al., 1995) satisfied the criteria for ToRSV cleavage sites. Site-directed mutagenesis of this potential cleavage site was conducted. In plasmid pT7-X2-Pro\textsuperscript{\DeltaX2N}, the codon for amino acid Q\textsuperscript{G} was precisely deleted. Maturation of this mutated precursor resulted in the production of the 88 kDa (X2-NTB) and 33 kDa (VPg-Pro) intermediates, whereas the 22 kDa (X2) and 66 kDa (NTB) proteins did not accumulate (Fig. 2c, lanes 3 and 4). Based on these results, we tentatively propose dipeptide sequence Q\textsuperscript{G} as the X2-NTB cleavage site. To confirm the potential nature of the cleavage site, we attempted direct N-terminal sequencing of the 66 kDa protein (NTB). However, the results were not
conclusive as a result of cross-contamination of the NTB protein with another labelled protein that ran in close proximity in SDS gels (data not shown).

**Processing at a novel cleavage site in the N-terminal domain of the P1 polyprotein**

To test if additional cleavage site(s) upstream of the NTB domain could be detected in vitro, plasmid pT7-X1-Pro was constructed using the pCITE-4a(+) vector as previously described. The X1-Pro precursor contained the coding region for NTB, VPg and Pro and for a region upstream of NTB with a coding capacity of 45 kDa (Fig. 3a). Coupled transcription/translation reactions were conducted in vitro as described above and the translation products were separated by 11% or 7% SDS–PAGE to allow visualization of small proteins and large precursors. As expected, a polyprotein with an apparent molecular mass of 146 kDa was observed following in vitro transcription/translation of this clone. After incubation of this polyprotein for 20 h at 16 °C, smaller proteins with apparent molecular masses of 26 kDa, 33 kDa, 42 kDa, 66 kDa and 105 kDa accumulated (Fig. 3b and c, lane 2). These proteins were not released following incubation of a polyprotein containing an inactive protease (plasmid pT7X1-ProH128A; Fig. 3b and 3c, lane 1). To determine the nature of these proteins, two sets of experiments were conducted. First, three derivatives of the X1-Pro precursor were created: precursor X1-ProΔNV (containing a precise deletion of residue Q^{1212} at position —1 of the NTB-VPg cleavage site), precursor X1-ProAX2N (containing a deletion of residue Q^{820} at position —1 of the putative X2-NTB cleavage site) and precursor X1-ProAX2NΔNV (containing deletions of residues Q^{1212} and Q^{820}). Second, immunoprecipitation of the processing products was conducted using the anti-NTB, anti-VPg, anti-Pro and anti-X1 antibodies. The results of the mutagenesis experiments are shown in Fig. 3(b and c) and the results of the immunoprecipitation experiments are shown in Fig. 4. These results are summarized in Table 2. Using these approaches, the following processing products were identified: VPg-Pro (33 kDa protein, calculated molecular mass 32 kDa), X1-X2 (42 kDa protein, calculated molecular mass 45–5 kDa), NTB (66 kDa protein) and X1-X2-NTB (105 kDa protein, calculated molecular mass of 111 kDa).

The 26 kDa protein was immunoprecipitated by the anti-X1 antibodies, but not by other available antibodies, suggesting that it corresponded to a mature protein containing part of the region upstream of the NTB domain. Mutation of the X2-NTB and/or the NTB-VPg cleavage sites did not prevent the production of this protein. Therefore, the 26 kDa protein was probably produced by proteolytic processing at a new cleavage site (X1-X2 cleavage site) and was released from the immediate N terminus of the precursor. An additional protein with an apparent molecular mass of 110 kDa was likely to have been produced by cleavage at the putative X1-X2
cleavage site. This protein was produced upon maturation of the X1-ProANV and X1-ProAX2ANV precursors (and also in smaller amounts upon maturation of the X1-Pro and X1-ProAX2 precursors) and may correspond to the X2-NTB-VPg-Pro intermediate.

Taken together these results indicate that proteolytic processing occurred at the X2-NTB and NTB-VPg cleavage sites and at a new cleavage site in the N-terminal region of the polyprotein. Cleavage at the VPg-Pro site was not detected in the X1-Pro precursor. As mentioned above, examination of the deduced amino acid sequence of the RNA-1-encoded polyprotein at the putative location for the novel X1-X2 cleavage site revealed the presence of only one potential cleavage site that satisfied the criteria for ToRSV cleavage sites (dipeptide QG). The calculated molecular mass of the putative X2-NTB-VPg-Pro precursor produced by the cleavage of this dipeptide is 119 kDa; this is slightly larger than the apparent mass of the 110 kDa protein observed in the cleavage products of the X1-Pro precursor. These results support the hypothesis that the 110 kDa protein is the X2-NTB-VPg-Pro intermediate produced by cleavage at the putative QG cleavage site.

Site-directed mutagenesis of the possible QG cleavage site between X1 and X2 was conducted. In construct pT7-X1-ProAX2X2, amino acid Q was precisely deleted. Following proteolytic processing of this precursor, the 26 kDa (X1) protein was not produced (Fig. 3b, lane 3). The 26 kDa and 110 kDa (X2-NTB-VPg-Pro) proteins were also not produced in a triple mutant containing mutations of the X1-X2, X2-NTB and NTB-VPg cleavage sites (compare construct pT7-X1-ProAX1AX2ANV with construct pT7-X1-ProAX1AX2ANV; Fig. 3b and c, lanes 6 and 8). The presence of X2, therefore, causes all proteins to run anomalously faster on SDS–PAGE, as observed for X1-X2-NTB, X1-X2 and X2-NTB-VPg-Pro cleaved products and the X2-Pro precursor. These results support the hypothesis that the 110 kDa protein is the X2-NTB-VPg-Pro intermediate produced by cleavage at the putative QG cleavage site.

Table 2. Summary of results from in vitro processing assays and immunoprecipitation of wild-type and mutated X1-Pro constructs

<table>
<thead>
<tr>
<th>Apparent molecular mass of</th>
<th>X1-Pro (cleavage site mutated)*</th>
<th>Immunoprecipitation (Abs)‡</th>
<th>Deduced nature of protein</th>
<th>Calculated molecular mass of processed product (kDa)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>processed product (SDS–PAGE) (kDa)</td>
<td>WT</td>
<td>ΔX1X2</td>
<td>ΔX2N</td>
<td>ΔN</td>
</tr>
<tr>
<td>26</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>33</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>42</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>66</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>97</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>105</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

* Presence (or absence) of each processed product upon maturation of the wild-type (WT) X1-Pro precursor or of precursors containing mutations in the putative X1-X2 cleavage site, putative X2-NTB cleavage site (X2N) or NTB-VPg cleavage site (NV) is indicated.
† The ability of antibodies raised against the putative X1 protein (X1), putative NTP-binding protein (NTB), VPg and protease (Pro) to immunoprecipitate each cleavage product is indicated. For the 42 kDa and 110 kDa proteins, the amounts of protein produced were too low to allow their identification by immunoprecipitation. nd, Not determined.
‡ The calculated molecular mass of each processed product is based on the position of the identified (NTB-VPg) and proposed (X1-X2 and X2-NTB) cleavage sites.
Intramolecular processing of P1 cleavage sites in vitro

To investigate if P1 cleavage sites could be processed in *trans* by the protease in *vitro*, precursors containing various regions of P1 were incubated in the presence of purified recombinant ToRSV protease (Fig. 5). Plasmid pET-MPCAT was used as a positive control for protease activity. This plasmid allowed the expression of a precursor containing the cleavage site between the RNA-2-encoded movement protein (MP) and coat protein (CP) that is cleaved in *trans* by the protease (Hans & Sanfaçon, 1995). The MP-CAT precursor with an apparent molecular mass of 65 kDa (calculated...
molecular mass of 59 kDa) was processed into two proteins with apparent molecular masses of 36 kDa and 27 kDa. These corresponded approximately to the predicted sizes for the C-terminal half of MP (32 kDa) and a fusion protein (CP-CAT) containing the N-terminal region of CP and CAT (27 kDa). The nature of these proteins was confirmed by immunoprecipitation experiments using antibodies against MP and CAT (Hans & Sanfaçon, 1995; data not shown). Cleavage by exogenously added protease was not detected in two large P1-derived polyprotein precursors containing an inactive protease. Precursor X1-Pro$^{H128SD}$ included all the potential cleavage sites upstream of Pro and precursor NTB-Pol$^{H128SD}$ contained the domains for NTB, VPg, Pro and Pol and therefore all the potential cleavage sites downstream of NTB. Trans-cleavage was also not detected upon addition of the exogenous protease to three smaller precursors that did not include the protease domain: a precursor containing the C-terminal half of NTB followed by VPg (plasmid pT7-C-NTB-VPg; Fig. 5), a precursor containing the entire NTB and VPg domains (data not shown) and a precursor containing the C-terminal region of X1 followed by X2. These results suggest that processing at P1 polyprotein cleavage sites is predominantly an intramolecular event.

Discussion

The results presented in this study provide evidence for the presence of a novel cleavage site in the N-terminal region of the ToRSV P1 polyprotein (X1-X2 cleavage site). Cleavage at this site by the ToRSV protease was demonstrated in vitro using partial cDNA clones and confirmed by mutagenesis and immunoprecipitation experiments. Interestingly, although proteolytic cleavage was also demonstrated at the X2-NTB and NTB-VPg cleavage sites, processing at the VPg-Pro cleavage site was very inefficient in vitro in all the P1-derived precursors studied here (X1-Pro, X2-Pro) or previously (NTB-VPg-Pro, VPg-Pro-N-Pol; Wang et al., 1999). Cleavage at this site is also inefficient in E. coli using either a small precursor (VPg-Pro-N-Pol; Wang et al., 1999) or a larger precursor (NTB-VPg-Pro-Pol; unpublished data). It will be interesting to determine if the VPg-Pro precursor also accumulates in infected plants.

The results presented in this study and earlier (Carrier et al., 1999) suggest that P1 processing occurs predominantly in cis. Indeed, precursors in which the protease domain was mutated or absent were not cleaved in vitro by purified recombinant ToRSV protease supplied exogenously. The accumulation of several intermediates in vitro (X1-X2-NTB, X1-X2) supports the notion that they are not further cleaved by the protease. Interestingly, although trans-cleavage was not detected in the P1 polyprotein of tomato black ring virus (TBRV, a nepovirus of subgroup B; Hemmer et al., 1995), trans-cleavage was observed at the X-NTB cleavage site of grapevine fanleaf virus (GFLV, a nepovirus of subgroup A; Margis et al., 1994). Our results do not exclude the possibility that some trans-cleavage of the ToRSV P1 cleavage sites may occur on the full-length P1 polyprotein.

Based on the results of site-directed mutagenesis, dipeptides Q423/G and Q629/G are proposed to be the putative X1-X2 and X2-NTB cleavage sites, respectively. In a previous study (Carrier et al., 1999), we showed that the identity of the amino acid at position −2 of the cleavage site played a role in the efficiency of cleavage at two ToRSV cleavage sites. While previously characterized ToRSV cleavage sites contain a C or a V at position −2, introduction of an A at position −2 of the Pro-Pol and X-MP cleavage sites resulted in efficient cleavage at these sites. In contrast, introduction of other amino acids (R, G, F) resulted in inefficient cleavage. The presence of an A at position −2 of the putative X1-X2 cleavage site is consistent with these results. The presence of a T at position −2 of the presumed X2-NTB cleavage site suggests that this is also an acceptable amino acid at this position. Efficiency of processing at cleavage sites including a T at position −2 was not directly tested in our previous mutagenesis study (Carrier et al., 1999).

The results presented here suggest that in ToRSV, the region upstream from NTB contains two distinct protein domains delineated by the putative X1-X2 and X2-NTB cleavage sites. The predicted molecular mass of these proteins is 46 kDa for the X1 protein (assuming that translation is initiated at the first AUG codon) and 21 kDa for the X2 protein. Comoviruses and nepoviruses have very similar genomic organizations in that the NTB, VPg, Pro and Pol domains are located in the C-terminal region of the P1 polyprotein. However, the size of the region N-terminal of NTB varies with different viruses (Fig. 6). In cowpea mosaic virus (CPMV, a comovirus), this region contains a single protein domain with a molecular mass of 32 kDa (Co-Pro protein; Goldbach & Wellink, 1996). In nepoviruses, this region has a coding capacity for a protein (or a precursor) with a predicted molecular mass of 67 kDa for ToRSV, 71 kDa for peach rosette mosaic virus (PRMV, a nepovirus of subgroup C; Lammers et al., 1999), 60 kDa for grapevine chrome mosaic virus and TBRV (two nepoviruses of subgroup B; Greif et al., 1988; Le Gall et al., 1989) and 45 kDa for GFLV (a nepovirus of subgroup A; Ritzenhaller et al., 1991). Indirect evidence based on comparisons of the amino acid sequence of the N-terminal region of the P1 polyprotein of nepoviruses and comoviruses suggests that the presence of two distinct protein domains may be a common feature of nepoviruses. Indeed, two regions of sequence similarity were identified (Fig. 6). The first region of similarity is an alanine-rich sequence which has been identified in the X1 domain of the ToRSV P1 polyprotein and in the equivalent regions of the P1 polyproteins of all other nepoviruses for which sequence information is available (Rott et al., 1995; Mayo & Robinson, 1996; Lammers et al., 1999). This sequence is not present in the comovirus genome. The second region of similarity is present in the X2 domain of the ToRSV P1 polyprotein, in the Co-Pro domain of the comovirus polyproteins and in the corresponding regions of all other
Nepoviruses

ToRSV (subgroup C)

PRMV (subgroup C)

TBRV (subgroup B)

GFLV (subgroup A)

Comoviruses

CPMV

Fig. 6. Genomic organization of ToRSV RNA-1 compared to that of related nepoviruses and comoviruses. Only the coding region is represented. The rectangle represents the polyproteins encoded by RNA-1. Continuous vertical lines represent cleavage sites identified by in vitro processing experiments (in ToRSV the putative sequence for the X1-X2 and X2-NTB cleavage sites is indicated above the corresponding vertical lines). Discontinuous vertical lines represent cleavage sites proposed after examination of the amino acid sequence (in PRMV, dipeptide Q647/S; Discontinuous vertical lines represent cleavage sites proposed after experiments (in ToRSV the putative sequence for the X1-X2 and X2-NTB rectangle represents the polyproteins encoded by RNA-1). Continuous lines represent cleavage sites for which the amino acid sequence is available (Lammers et al., 1999). The stars indicate a region of similarity to the comovirus Co-Pro domain (Ritzenthaler et al., 1991; Rott et al., 1995). ToRSV, tomato ringspot virus; PRMV, peach rosette mosaic virus; TBRV, tomato black ring virus; GFLV, grapevine fanleaf virus; CPMV, cowpea mosaic virus.

The potential functions of the ToRSV X1 and X2 domains are not known. A protein sequence identity search with BLAST (Altschul et al., 1990) using the amino acid sequence of the X2 protein of ToRSV did not reveal any significant similarities to known proteins, other than the ones already mentioned above. The X1 protein contains a large region of sequence identity with the N-terminal portion of the P2 polyprotein (Rott et al., 1991). Duplication of this domain appears to be unique to nepoviruses of subgroup C. The X2 protein has similarities with the 32 kDa protein (Co-Pro) of CPMV that acts as a protease co-factor by decreasing cleavage efficiency of the P1 polyprotein, and increasing cleavage efficiency of the P2 polyprotein (Peters et al., 1992). However, these and previously presented data do not support the idea that the X2 protein regulates the activity of the ToRSV protease in a similar manner. Indeed, the presence of the X2 domain in P1 precursors did not seem to decrease cleavage efficiency at the P1 cleavage sites, i.e. release of NTB and VPg-Pro was similarly efficient in precursor X1-Pro (containing X2; this study) and in precursor NTB-VPg-Pro (Wang et al., 1999). The protease can also efficiently process P2 cleavage sites in the absence of the X2 domain (Hans & Sanfaçon, 1995; Carrier et al., 1999). Similar results were obtained with other nepoviruses (Margis et al., 1994; Hemmer et al., 1995). The conserved sequences among the X2 protein of ToRSV, the 32 kDa Co-Pro protein of CPMV and the equivalent region of the other nepovirus polyproteins may therefore possess an alternative, unknown function. It is possible that either the X1 or X2 protein of P1 (and perhaps the N-terminal protein of P2) plays an accessory role in RNA replication, similar to that suggested for the N-terminal protein of P2 in CPMV (Van Bokhoven et al., 1993) and GFLV (Gaire et al., 1999). Future work will be aimed at investigating the presence of the X1 and X2 proteins in infected plants and the potential function of the proteins encoded by the N-terminal region of P1.

The authors would like to thank Mr Andrew Wieczorek for the initial purification of active recombinant protease and for the construction of clone PET-MPCAT, Mrs Joan Chisholm for expert technical assistance and Mr Mike Weis for photographs. Drs Guus Bakkeren and Paul Wiersma and Mrs Joan Chisholm are acknowledged for critical reading of the manuscript. This study was supported in part by a grant from NSERC (Canada) to H.S. A.W. was a recipient of a University Graduate Fellowship from the University of British Columbia.

References


Received 2 June 2000; Accepted 21 July 2000