Genomic organization of RNA2 of *Tomato ringspot virus*: processing at a third cleavage site in the N-terminal region of the polyprotein in *vitro*

Karma Carrier,† Yu Xiang and Hélène Sanfaçon

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

The protease of *Tomato ringspot virus* (genus *Nepovirus*) is responsible for proteolytic cleavage of the RNA2-encoded polyprotein (P2) at two cleavage sites, allowing definition of the domains for the movement protein (MP) and coat protein. In this study, we have characterized a third cleavage site in the N-terminal region of P2 using an *in vitro* processing assay and partial cDNA clones. Results from site-directed mutagenesis of putative cleavage sites suggest that cleavage occurs at dipeptide Q301/G. Cleavage at this site is predicted to result in the release of two proteins from the N-terminal region of P2: a 34 kDa protein located at the N terminus of P2 (assuming translation initiation at the first AUG codon) and a 71 kDa protein located immediately upstream of the MP domain. In contrast, only one protein domain is present in the equivalent region of the P2 polyprotein of other characterized nepoviruses.

Nepoviruses are bipartite positive-strand RNA viruses belonging to the family *Comoviridae* (Sanfaçon, 1995; Mayo & Robinson, 1996). Each RNA encodes a large polyprotein which is proteolytically processed by a viral serine-like proteinase (Pro; encoded by RNA1) to release the mature proteins. Nepoviruses have been subdivided into three subgroups based on the size of the RNA2 (Mayo & Robinson, 1996). Members of subgroup C (including *Tomato ringspot virus*; ToRSV) are characterized by having a larger RNA2 that includes substantial regions of sequence identity shared with RNA1 (Rott et al., 1991b). The proteolytic processing of ToRSV polyproteins has been studied. The conserved amino acid sequence at cleavage sites recognized by the ToRSV proteinase is (C or V)Q/(G or S) (Wang et al., 1999; Carrier et al., 1999). Site-directed mutagenesis of two cleavage sites has confirmed the importance of these amino acids for efficient processing at the cleavage sites (Carrier et al., 1999). Five cleavage sites have been identified on the RNA1-encoded polyprotein (P1) and these define the domains for the putative NTP-binding protein, VPg, Pro and RNA-dependent RNA polymerase and for two additional proteins (X1 and X2) located at the N terminus of the polyprotein (Wang et al., 1999; Wang & Sanfaçon, 2000). Two cleavage sites have been identified in the C-terminal region of the RNA2-encoded polyprotein (P2), allowing definition of the movement protein (MP) and coat protein (CP) domains (Hans & Sanfaçon, 1995; Carrier et al., 1999). However, a large region at the N terminus of P2 (coding capacity for a protein or a precursor of approximately 102 kDa) has not been characterized. In this study, we report the identification of a new cleavage site in the N-terminal region of P2.

To test for the presence of possible cleavage sites in the N-terminal region of P2, an *in vitro* processing assay was used. Plasmid pT7-N-termFL, containing the entire N-terminal region of P2 under the control of the T7 polymerase promoter (Fig. 1a), was constructed by amplifying a cDNA fragment from plasmid pMR14 (Rott et al., 1991a) using Pfu polymerase (Stratagene) and primers PX6 [5′ TTTCCATGGGCCAAAATCTGGTGATATTCC 3′], containing an engineered EcoRI site (underlined) and corresponding to ToRSV RNA2 nucleotides 4–25 (numbering according to Rott et al., 1991a) and PX4 [5′ CAAAGATTTCTCCCTATCGACAAACCCTC 3′], containing an engineered EcoRI site (underlined) and complementary to ToRSV RNA2 nucleotides 2831–2813. The cDNA fragment was digested with EcoRI and inserted into the corresponding sites of plasmid pCITE (4a+) (Novagen). This resulted in an in-frame fusion of the viral sequence with an ATG present in the vector and with some non-viral sequences (Fig. 1a). The labelled polyprotein precursor was synthesized using a coupled transcription/translation rabbit reticulocyte system (TNT, Promega) in the

Author for correspondence: Hélène Sanfaçon.
Fax +1 250 494 0755. e-mail SanfaconH@em.agr.ca
† Present address: Department of Biochemistry, School of Medicine, Emory University, Atlanta, GA 30322, USA.
Fig. 1. Proteolytic processing at a new cleavage site on P2 in vitro. (a) Schematic representation of the processing of three polyprotein precursors. The predicted genomic organization of RNA2 is shown on the top of the figure. Previously determined X4–MP and MP–CP cleavage sites (Carrier et al., 1999; Hans et al., 1995) are shown by continuous vertical lines. Two putative cleavage sites delineating the putative X3 and X4 domains are shown by the discontinuous vertical lines. Precursors
presence of $[^35]S$-methionine and separated by SDS–PAGE as described (Wang et al., 1999). A predominant protein of approximately 120 kDa (calculated molecular mass for the precursor is 112.3 kDa) was observed (Fig. 1b, lane 1). Additional minor proteins were also produced, presumably through internal translation initiation. Purified recombinant ToRSV protease was added to the translation products and incubated at 15 °C (as described in Wang & Sanfaçon, 2000). Two additional proteins with relative molecular masses of approximately 75 and 40 kDa were observed (Fig. 1b, lanes 2 and 3). These proteins did not appear when the translation products were incubated at 15 °C in the absence of exogenous proteinase (Fig. 1b, lanes 4 and 5), demonstrating that they were produced by proteolytic cleavage of the N-termFL precursor at a new cleavage site by the ToRSV protease. Processing at this site was relatively inefficient in vitro with only partial cleavage observed after overnight incubation under conditions that allow complete cleavage of precursors containing the X–MP and MP–CP cleavage sites in 3 h (Carrier et al., 1999; Wang & Sanfaçon, 2000). Examination of the deduced amino acid sequence of the N-terminal region of the P2 polyprotein revealed the presence of only two putative cleavage sites (diadenylates Q301/G and Q319/G, Fig. 1a). A valine was present at the –2 position of the putative sites, which is in agreement with the criteria established for ToRSV cleavage sites (Carrier et al., 1999). The sizes of the 40 and 75 kDa proteins were consistent with the predicted size of cleaved products obtained after proteolytic processing at either one of the putative sites (38.5 and 73.8 kDa for Q301/G; 40.4 and 71.9 kDa for Q319/G).

To verify the hypothesis that cleavage occurred at one (or both) of the aforementioned sites, two additional precursors (N-termA1 and N-termA2) were studied that contained deletion of sequences at either the N or the C terminus of the N-termFL precursor (Fig. 1a). Plasmid pT7-N-termA1 was constructed by amplification of a cDNA fragment using primers PX4 and W034 (Wang & Sanfaçon, 2000). The cDNA fragment was digested with NcoI and EcoRI and inserted in the corresponding sites of plasmid pCITE (4a+). Plasmid pT7-N-termA2 was constructed by first ligating an 811 nt Smal–HindIII cDNA fragment (corresponding to ToRSV RNA2 nucleotides 363–1234) into the EcoRI–HindIII sites of pET21B (Novagen). The cleaved EcoRI site of plasmid pET21B had been filled-in using the Klenow fragment of DNA polymerase I to create a blunt end prior to the ligation. The NrdI–NotI 649 nt fragment from this intermediate plasmid was then ligated with the large NrdI–NotI fragment of plasmid pT7-N-TermFL to create plasmid pT7-N-termA2. Coupled in vitro transcription/translation of plasmids pT7-N-termA1 and pT7-N-termA2 resulted in the production of predominant proteins of approximately 90 and 55 kDa, respectively (Fig. 1b, lanes 10 and 11), corresponding to the calculated molecular masses for N-termA1 (84.5 kDa) and N-termA2 (49.1 kDa). As above, some minor proteins were also produced, presumably through internal translation initiation. The 75 kDa protein was produced upon incubation of the N-termA1 precursor in the presence of ToRSV protease (Fig. 1b, compare lane 7 to lane 10) while the 40 kDa protein was produced upon maturation of the N-termA2 precursor (Fig. 1b, compare lanes 8 to lane 11). These results confirmed that the 40 kDa and the 75 kDa proteins were the N-terminal and C-terminal processing products of the N-termFL precursor, respectively.

To further study the proteolytic processing at the new cleavage site(s), point mutations at the two putative sites were introduced into the N-termFL and N-termA2 precursors. We have previously shown that mutation of the conserved Q at the –1 position of the cleavage site to an A results in a complete inhibition of processing (Carrier et al., 1999). Site-directed mutagenesis was done on plasmids pT7-N-termFL and pT7-N-termA2 to introduce an A at the –1 position of the Q301/G and Q319/G diadenylates using the Quickchange site-directed mutagenesis kit (PDI bioscience). Oligonucleotides KC-55 (5′ GGTCTCAGCCTGTTGCAGGGGTTCTCCCA 3′; mutated nucleotides shown in bold) and KC-56 (5′ GGGAGAACCCCCCTGCACAGCTGAGACC 3′) were used for the A301 mutation while oligonucleotides KC-57...
(5′-GTCGCTCCACCGTTGCGGTTGGTGCACGC-3′) and KC-58 (5′-GCGGCAACCCACCCGGCAAGGTGGG-AGCGAC-3′) were used for the A319 mutation. The results from in vitro processing experiments of the wild-type and mutated precursors were similar regardless of the precursor (N-termFL and N-term∆2) into which the mutations were introduced. The precursors containing the A319 mutation were cleaved by the exogenous proteinase in vitro, resulting in cleavage patterns similar to those observed with the corresponding wild-type precursors (Fig. 1c, compare lane 3 to lane 1, and lane 7 to lane 5). In contrast, cleavage was not detected in precursors containing the A301 mutation (Fig. 1c, lanes 2 and 6) or a double mutation (A301 + A319; Fig. 1c, lanes 4 and 8). These results suggest that cleavage occurred at the Q301/G dipeptide in vitro. Confirmation of this cleavage site by direct N-terminal amino acid sequencing of the 75 kDa product released from the N-termFL or N-term∆1 precursors was not successful as it was produced in low amounts and ran in close proximity to other labelled proteins on SDS–PAGE. Cleavage at the new cleavage site (presumably at dipeptide Q319/G) is predicted to allow the release of two proteins from the N-terminal region of P2: a 34 kDa protein (assuming translation initiation at the first AUG codon), arbitrarily called X3, and a 71 kDa protein (X4; see Fig. 2). Although Q301/G is suggested as the new cleavage site based on in vitro processing experiments, we cannot exclude the possibility that both the Q301/G and the Q319/G dipeptides are recognized in vivo or that the Q319/G dipeptide may act as an alternative cleavage site.

Although two closely spaced putative cleavage sites are present at the N terminus (VQ301/G and VQ319/G; this study) and the C terminus (VQ907/S and VQ934/S; Carrier et al., 1999) of the X4 domain, results from in vitro processing experiments indicate that cleavage occurs predominantly at
dipeptides Q^{101}/G and Q^{98}/S (this study; Carrier et al., 1999). All four putative cleavage sites satisfied the criteria established for ToRSV cleavage sites (C or V/Q/G or S; Carrier et al., 1999). This suggests that features other than the primary sequence are implicated in the recognition of the cleavage sites by the protease. An analysis of secondary structure prediction around the P2 cleavage sites was done using a combination of prediction methods as described in Guermeur et al. (1999). In this model, the MP–CP (Q^{132}/G), X4–MP (Q^{98}/S) and proposed X3–X4 (Q^{101}/G) cleavage sites were located in areas predicted to adopt a random coil conformation while the other dipeptides (Q^{319}/G and Q^{96}/S) were located in areas predicted to adopt an α-helix conformation. The three-dimensional structure of the precursor may therefore play a role in the recognition of the cleavage sites as suggested for poliovirus and comovirus cleavage sites (Ypma-Wong et al., 1988; Clark et al., 1999).

Our results show that the ToRSV protease recognizes a third cleavage site on P2 in vitro. This observation contrasts with results obtained with other nepoviruses (Fig. 2). Indeed, analysis of the proteolytic processing of the P2 polyproteins of *Grapevine fanleaf virus* (GFLV; a nepovirus from subgroup A) and of *Tomato black ring virus* and *Grapevine chrome mosaic virus* (two nepoviruses from subgroup B) revealed the presence of only two cleavage sites (Margis et al., 1993; Demangeat et al., 1991; Hibrand et al., 1992). *Blackcurrant reversion virus* (BRV) is the only other member of nepovirus subgroup C for which the entire sequence of the RNA2 is available (Latvala-Kilby & Lehto, 1999). As with ToRSV, BRV has a large P2, which raises the possibility that a third cleavage site may be present. Since the only BRV cleavage site identified so far is rather unusual (D/S; Latvala et al., 1998), it is difficult to predict additional cleavage sites based on the deduced amino acid sequence. The presence of two protein domains in the N-terminal region of the ToRSV P2 may therefore either represent a unique feature of ToRSV or a feature common to other nepoviruses of subgroup C.

The possible function of the X3 and X4 proteins in the ToRSV replication cycle is not known. A protein sequence homology search using BLAST (Altschul et al., 1990) did not reveal any similarities of the X4 protein to other known proteins. The X3 protein contains extensive regions of homology to the X1 protein domain present at the N terminus of P1 (Wang & Sanfaçon, 2000; Fig. 2), including an alanine-rich motif also present in the N-terminal region of the BRV P2 (Latvala-Kilby & Lehto, 1999) and in the N-terminal region of other nepovirus P1 polyproteins (Mayo & Robinson, 1996). The X3 protein also contains several proline motifs which are also present in the RNA2-encoded 28 kDa protein of GFLV. The GFLV 28 kDa protein and the corresponding 58 kDa protein of *Coupea mosaic virus* (genus Comovirus) are critical for the replication of RNA2 (Van Bokhoven et al., 1993; Gaire et al., 1999). Further studies will be aimed at detecting the putative X3, X4 and/or X3–X4 proteins in infected plants and at examining the role of these proteins in the ToRSV replication cycle.

We wish to thank Drs D. Rochon and A. Wang and Mrs Joan Chisholm for critical reading of this manuscript. This work was supported in part by an NSERC research grant awarded to H.S. Karma Carrier and Yu Xiang contributed equally to this work.

References


Received 2 February 2001; Accepted 22 March 2001