Cleavage profiles of tobacco etch virus (TEV)-derived substrates mediated by precursor and processed forms of the TEV NIa proteinase

T. Dawn Parks,1 Holly A. Smith1 and William G. Dougherty1,2*

1Department of Microbiology and 2The Center for Gene Research and Biotechnology, 220 Nash Hall, Oregon State University, Corvallis, Oregon 97331-3804, U.S.A.

Nucleotide sequences coding for proteins containing the tobacco etch virus (TEV) NIa proteinase were generated by polymerase chain reaction amplification and/or site-directed mutagenesis. These coding regions contained sequences for the proteinase alone or as part of higher Mr precursors. Following transcription and translation of these sequences in a cell-free system, the various polyproteins, all containing an active small nuclear inclusion protein (NIa) proteinase, were used to process a TEV substrate series. Most substrates were processed in a similar fashion by all proteolytic forms. However, one substrate which contained the TEV 50K/71K protein junction was differently processed by several of the polyproteins containing NIa proteinase. Substrates which previously had no identified TEV NIa proteinase cleavage sites also were tested and were not cleaved by any of the proteinase-containing polyprotein forms.

Introduction

Differential processing of particular cleavage sites by viral polyproteins containing proteinase has been proposed to be important in the replication cycle of a number of animal viruses. In the alphavirus family, there is a clear pathway for polyprotein processing by precursors containing the nsP2 proteinase. It has been hypothesized that this differential processing scheme is responsible for regulation of RNA synthesis (for review, see Strauss & Strauss, 1990). The involvement of proteinase precursors in the differential cleavage of polyproteins has also been demonstrated with picornaviruses. During poliovirus genome expression, the 3CD precursor form of the 3C proteinase is required for processing of the P1 polyprotein into the VP0, VP1 and VP3 capsid proteins. Other members of the picornavirus group may require 3C proteinase alone or as part of the 3ABC polyprotein for maturation of the structural P1 polyprotein into capsid proteins (reviewed by Harris et al., 1990; Palmenberg, 1990).

The ability of a plant viral proteinase to process different cleavage sites, with varying cleavage profiles, has been demonstrated with the proteolytic activity associated with the small nuclear inclusion protein (NIa) of tobacco etch virus (TEV) (Dougherty & Parks, 1989). TEV differential processing is determined by those particular amino acids found in the cleavage site heptapeptide consensus sequence, and does not appear to be influenced by amino acid sequences which flank the cleavage site.

This study was conducted to determine whether the TEV NIa proteinase domain or polyproteins containing this proteolytic activity differentially processed substrates derived from the TEV genome. In the TEV proteinase study cited above, the source of proteolytic activity was the 49K NIa derived from partially purified nuclear inclusion bodies isolated from TEV-infected plants. This 49K protein has recently been shown to be a polyprotein containing the domain for the VPg (viral protein, genome-linked) moiety in the N-terminal portion, in addition to a proteinase function which resides in the C-terminal portion of the polyprotein (Murphy et al., 1990; Dougherty & Parks, 1991). In the study described here, TEV NIa-associated proteolytic activity was generated in several contextual forms and tested in a cell-free system. These proteolytic activities were used to process various TEV-derived substrates. These studies indicated that, for most TEV substrates, there were no detectable differences in cleavage by various TEV polyprotein proteinases. However, one substrate did exhibit differential cleavage profiles for several proteolytic forms. Based on observations made in the cell-free studies, we would propose that, unlike the polioviral 3CD protease system, the TEV system would not have a prerequisite for particular proteinase precursors in the processing of its genome-derived polyprotein.

Methods

Construction of plasmids containing proteinase coding sequences. All proteinase coding sequences used in this study were generated by polymerase chain reaction (PCR) amplification (Erlich, 1989) of TEV...
Fig. 1. A schematic diagram of the protease forms used in this study. The TEV genome is shown at the top in bold. The VPg is indicated by a circle at the 5' end and a polyadenylated region is shown at the 3' end. The \( M_r \) values of individual gene products are shown above, and the gene products present as inclusion bodies in infected cells are indicated: CI, NIa and NIb, and AI, the amorphous inclusion protein. Putative and demonstrated activities are listed below the genome map and include (hel) helicase, (pro) protease, (rep) replicase and (cap) capsid (Dougherty & Carrington, 1988; Carrington et al., 1989; Lain et al., 1990; Murphy et al., 1990). Cleavage sites are indicated by vertical lines, with an as yet unidentified proteolytic site indicated by a (?) and the HC-Pro cleavage site indicated by an arrowhead. The vertical dotted line within the 49K polypeptide indicates an internal cleavage site which delineates VPg and protease domains. An expanded view of the TEV genome used for protease constructs is shown underneath. Plasmid constructs which coded for proteinase forms are listed on the left with predicted product \( M_r \) values indicated. Each plasmid contained the 5' TEV untranslated region (nucleotides 1 to 144), with the coding sequence of the polypeptides inserted at an ATG translation start site which provided the initiating Met codon. The filled-in boxes represent the polypeptides which contain the proteinase; open boxes show peptides which, although contained in the coding sequence of the plasmid, are cleaved by autocatalysis of the TEV NIa proteolytic activity contained within the translation product. The picornaviral counterparts are indicated to the right of each construct.

The TEV encoding sequence for the 5' end was used as a template for construction of plasmids pN49 and pN27, with pTL-5473\(^\text{a/b}\) (Carrington & Dougherty, 1987b) used in generating pN49\(^\text{a/c}\). The TEV coding region contained in pTL5473-11N (Carrington et al., 1988), which had a mutation inactivating the cleavage site which delineates 6K from 49K amino acid sequence, was used in creating plasmid p6 + 49. Two other plasmids with altered cleavage site coding sequences were used as sources of TEV cDNA. Plasmids pTL5473-12C (with an altered C-terminal 49K cleavage site) and pTL5473-11N/12C (containing altered N- and C-terminal 49K cleavage sites (Carrington et al., 1988)) were cleaved with the restriction enzymes Sall (TEV nucleotide 7166) and PstI (vector sequences). A cDNA fragment corresponding to TEV nucleotides 7166 to 9176 was then inserted. The resulting plasmids were used as templates in PCR amplification to create plasmids p27 + 58, p49 + 58 and p6 + 49 + 58. The vector sequences previously described, and are shown in Fig. 1. The proteinase coding sequence found in plasmid pTL-5473 (Carrington & Dougherty, 1987a), was used as a template for construction of plasmids pN27 and pN27 + 58, which created an \( M_r \) restriction site corresponding to nucleotides 6251 to 6257 of the TEV genome. Two 3' oligonucleotide primers were used. One corresponded to vector sequences and was used in generating pN49, pN49\(^\text{a/c}\), p6 + 49 and pN27. A second oligonucleotide created a PstI restriction site at the coding sequence corresponding to the 58K/30K protein junction and was used in the synthesis of p27 + 58, p49 + 58 and p6 + 49 + 58. Standard recombinant DNA techniques (Sambrook et al., 1989) were used to insert the PCR-generated products described above into the transcription vector pTL27N. This plasmid vector was similar to pTL17 described previously (Carrington et al., 1987), except that an origin of replication for single-stranded DNA production was inserted and a site-directed mutation creating an \( NcoI \) restriction site at the ATG translation start codon (TEV nucleotides 43 to 145) was introduced (Taylor et al., 1985a, b).

Construction of plasmids used in substrate production. Six plasmids were used for the production of substrate molecules in this study and are shown in Fig. 3(a). Three of these plasmids (pAS, pAS-58/30 and pTL37-8595) have been described previously and contain cleavage sites recognized by the NIa protease (Dougherty & Parks, 1989). Briefly, pAS contained sequences corresponding to the region of the TEV genome surrounding the 50K/71K cytoplasmic cylindrical inclusion (CI) protein junction. The plasmid pAS-58/30 was similar, except that it contained a cleavage site sequence which had been mutated such that it coded for the seven amino acid cleavage site sequence found at the junction of the 58K and capsid (30K) proteins. This cleavage sequence was found in its normal polyprotein context in pTL37-8595. Two substrates derived from pTL0036 and pN71 did not contain any previously identified TEV NIa proteinase cleavage sites. The plasmid pTL0036 contained coding sequences from the 5' region of the TEV genome, which included an 87K polypeptide [31K and 56K helper component-protease (HC-Pro), a 50K protein, with a TGA stop codon at the 3' end] of the 50K coding sequence. The coding sequence of the 71K CI protein (nucleotides 3634 to 5532) found in pN71 was generated by PCR amplification using primers which created an NcoI restriction site, an AUG start codon at the 5' end and a stop codon at the 3' end. The inactive proteinase molecule described above (pN49\(^\text{a/c}\)) also served as a substrate source. Diagrams of these molecules are shown in Fig. 3(a).

Transcription, translation and processing. Following linearization with restriction enzymes, phenol extraction and ethanol precipitation, all plasmid DNAs were transcribed with T7 RNA polymerase as described (Carrington & Dougherty, 1987a). After ethanol precipitation, aliquots of all transcripts were translated in the presence of \([\text{35S}]\) methionine in a rabbit reticulocyte lysate as described by Dougherty & Hiebert (1980), and protein production was quantified by TCA precipitation. Proteinase transcripts were then translated in the presence of unlabelled amino acids, and substrate molecules were translated in the presence of labelled methionine. Based on TCA-precipitable counts, equimolar amounts of each proteinase were incubated with various substrates in a 5:1 ratio of proteinase to substrate.

Aliquots were taken at 5, 10, 15, 30, 60, 90 and 120 min after the addition of proteinase for reactions using substrates derived from pAS-58/30, pAS and pTL37-8595. Additionally, substrates were incubated for 120 min with TEV NIa proteinase in the form of a partially purified nuclear inclusion body preparation. Substrates and products were separated on 12.5% polyacrylamide gels containing SDS using the buffer system of Laemmli (1970). The gels were impregnated with ENHANCE (NEN), dried and exposed to X-ray film. Autoradiographs of these gels were scanned using a Zeineh soft laser scanning densitometer, and the amount of substrate present as product was...
were incubated together at 30 °C for 2 h and 24 h. For samples which
terminated by the addition of RNase A. Proteinases and substrates
were incubated for 24 h, PMSF was added to a final concentration of
0.5 mM. Substrates and products were separated using SDS-PAGE and
gels were visualized by fluorography.

Results
The six proteinase sources used in this study presented
the NIa proteolytic activity in different contextual forms
(Fig. 1). These forms included the 27K NIa proteinase
domain expressed alone and five additional sources
which presented the proteolytic activity in TEV polyprotein
precursor forms. These five polyproteinmes were
generated from plasmids containing site-directed mutations (single amino acid changes) which eliminated
cleavage sites and consisted of the following: (i) a 49K
NIa polyprotein [(21K VPg and 27K NIa proteinase activity), p49], (ii) 49K NIa proteinase with the ad-
tional amino acids of the 6K protein product at the N
terminus (p6 + 49), (iii) 49K polyprotein with the 58K
putative replicase sequences attached at the C terminus
(p49 + 58), (iv) a 6K–49K–58K polyprotein (p6 +
49 + 58) and (v) the 27K NIa proteolytic domain with
the 58K protein at the C terminus (p27 + 58). After in
vitro transcription and translation in a rabbit reticulocyte
lysate containing [35S]methionine, aliquots of each
translation were analysed by SDS-PAGE. Each protein-
ase form containing the TEV NIa proteolytic activity
migrated to a position in the gel corresponding to the
predicted Mr, (data not shown). Additional aliquots of
each translation were precipitated with TCA and
radioactivity was counted. Based on the level of
incorporation and the number of methionine residues,
relative equivalent molar amounts of each proteinase-
containing polyprotein were used in processing
reactions.

We investigated whether TEV cleavage sites, demon-
strated to be processed in trans by the proteolytic activity
associated with the TEV NIa polypeptide (49K), would
be processed differently by the various artificial polypro-
tein sources described above. Processing reactions using
substrates derived from pAS, pAS-58/30 and pTL37-8595
were conducted. Substrate precursor and processed
products were separated by PAGE and visualized by
fluorography, with examples shown in Fig. 2(a). For pAS
and pAS-58/30, the 32K substrate precursor was convert-
ed over time into 20K and 12K (not detected in our gel
system) products. These autoradiographs were subjected
to densitometric analysis. The percentage of substrate
converted to product was determined for each time point
sample and is presented for each proteinase in Fig. 2(b). These graphs represent the mean percentage values of
processed product at each time point, based on four or
five independent processing reactions. For processing
reactions in which pAS-58/30 was used as a substrate
source, each proteinase appeared to cleave the substrate
to generate a similar cleavage profile. This also was the
case with processing of pTL37-8595 (data not shown).

In contrast, there was differential processing of the
cleavage site contained in the pAS substrate source when
the various proteinase forms were analysed. In these
processing reactions, the p49 + 58-derived proteinase
appeared to process the pAS substrate at an enhanced
rate, while the pN27-derived proteinase processed
somewhat slower. The other four proteinase forms were
similar and intermediate in processing efficiency. These
results are presented in Fig. 2. The cleavage profiles
generated by proteinases derived from p49 + 58 and
pN27 were statistically faster and slower, respectively,
when compared to the other proteinases (P < 0:001).

Another substrate molecule, encoded by sequences
contained in the plasmid pN49C/A, consisted of amino
acid sequences corresponding to the 6K protein linked to
a proteolytically inactive NIa 49K polyprotein and 62
amino acids of the amino terminus of the large nuclear
inclusion protein (NIb) (58K). This molecule contained
two cleavage sites (6K/49K and 49K/NIb junctions)
normally processed in an autocatalytic or cis reaction, as
well as an inefficiently processed internal cleavage site
which delineates the VPg and NIa proteinase domains.
Processing of this precursor molecule by nuclear inclu-
sion bodies proceeded slowly, with limited processing
evident after 2 h. Some processing was evident after
24 h, with approximately half of the sample remaining in
the 62K unprocessed form, as shown in Fig. 3(b).
Polyacrylamide gel electrophoretic analysis of the cleavage
products of this reaction indicated that limited
cleavage occurred at both sites normally processed in cis,
as well as at the internal site in the 49K proteinase form.
The predicted Mr's of the bands shown in Fig. 3(b)
correspond to 62K for the uncleaved substrate, 56K
representing the product formed by cleavage at the N
terminus of the 49K polyprotein, 49K representing
cleavage at both the N and C termini of the 49K
polyprotein, 34K representing cleavage at the internal
site only (C-terminal fragment), 27K for a doublet
representing the product formed by cleavage at only
the internal site (N-terminal fragment) and the 27K protein-
ase fragment, and 21K representing the product formed
by cleavage at the N terminus and internal site. Cleavage
profiles similar to that observed for nuclear inclusion
bodies were exhibited when the pN49C/A substrate was
incubated with the six proteinase forms (data not
shown).
We subsequently investigated whether there were existing cryptic cleavage sites in the TEV polyprotein which could be processed by only one of the different Nia proteinase sources. Several additional substrates were tested for the ability to be processed by the six proteolytic forms. Two of these, derived from plasmids pTL0036 and pN71, contained sequences corresponding to the 31K–HC-Pro–50K TEV polyprotein and the 71K TEV cytoplasmic inclusion protein, respectively. The cDNA contained in these two particular plasmids did not code for a previously experimentally identified Nia proteinase cleavage site. Plasmid pTL0036 did contain sequences which coded for an active HC-Pro. Translation products derived from this plasmid were autocatalytically processed to generate 87K and 50K products, and the pN71 translation product migrated at a position corresponding to 71K. The substrates generated from these two plasmids were incubated for 2 h with the six proteinase forms and polyacrylamide gels were used to separate the 32K substrate (S) from the 22K product (P). The top three panels show processing assays using pN27-derived, p6 + 49-derived and p49 + 58-derived proteinases as their sources. The first lane in each corresponds to substrate incubated with nuclear inclusion bodies as a proteinase source (+) and the second lane shows substrate incubated without exogenously added proteinase (−). Numbers above the remaining lanes indicate time points (min) when aliquots were taken and processing was stopped by the addition of Laemmli loading buffer. (b) Processing profiles of each of the six proteinase forms (□, pN27; ○, pN49; ●, p6 + 49; ◇, p27 + 58; ■, p49 + 58; □, p6 + 49 + 58) are shown. For each time point, the amount of substrate present as product was determined by densitometric measurement, and a reaction profile was determined. The graphs presented represent the mean reaction profile for each proteinase, based on four or five repetitions of each assay.
proteinase forms, and with the proteolytic activity associated with TEV nuclear inclusion bodies. There was no evidence that processing by any of the NIa proteolytic forms had occurred. Even after 24 h incubation with the six proteinase forms or with TEV nuclear inclusion bodies, there was no apparent cleavage of either substrate. Incubation of these substrates for 24 h with five times the normal amount of proteinase also did not result in any detectable processing (Fig. 3b).

Discussion

We have used a cell-free processing system to examine the possible role of processing by TEV NIa proteinase precursors. This experimental system has been used previously to determine that specific amino acids within TEV NIa cleavage sites affect the rate at which cleavage occurs (Dougherty & Parks, 1989). It has been proposed that these varying rates of cleavage may be important in the production of various viral proteins, hence providing a mechanism for regulation of gene expression. This study was undertaken to examine the possibility that an additional level of control over protein production might exist, whereby differential processing by protease precursor forms could enhance or inhibit the production of specific viral proteins.

The results of this study with various TEV NIa proteinase forms indicated that, for most substrates, all tested proteinase forms cleaved with essentially equivalent profiles. The substrates derived from pAS-58/30 and pTL37-8595 were cleaved by all six proteinase activities to generate similar processing profiles. Both of these substrates contain the seven amino acid cleavage sequence E-N-L-Y-F-Q/S, which has been identified previously as a rapidly processed TEV cleavage site sequence (Dougherty & Parks, 1989; Dougherty et al., 1989). However, different reaction profiles were consistently observed when substrate derived from pAS was tested. The substrates derived from pAS and pAS-30 are identical except for three amino acid differences localized within the heptapeptide cleavage sequence. The cleavage sequence (E-I-I-Y-T-Q/S) encoded in pAS has been demonstrated to be a slowly processed site. In cleavage assays utilizing pAS-derived substrate, two of the proteolytic forms showed somewhat different efficiencies. Proteinase derived from p49 + 58 cleaved the pAS substrate more rapidly and to a greater extent than had been observed when either TEV nuclear inclusion bodies or other polyprotein forms were used as a proteinase source. Essentially, p49 + 58 proteinase cleaved the slowly processed site of pAS with the same profile as the rapidly processed site sequence (with averages of 81% and 84% substrate converted to product after 2 h, respectively). The proteinase derived from pN27 cleaved the pAS substrate more slowly and to a lesser extent than any other proteinase tested over the 2 h assay period. This was considerably slower than the processing exhibited when this same proteinase was used in cleavage assays with pAS-58/30-derived substrate (average cleavages of substrate to product of 26% and 82%, respectively).

It can only be speculated as to why the different proteinase forms vary in their ability to cleave the pAS substrate and yet cleave the pAS-58/30 translation product with similar profiles. This suggests that the three amino acid differences which are proximal to the cleavage site are important in the processing phenotypes observed. One possibility is that the pAS-58/30 cleavage site is so efficiently recognized and processed that differences in proteinase form are not important and cleavage proceeds regardless of the specific proteinase.
We would suggest that the pAS-58/30 cleavage site takes on a conformation which maximizes proteinase–substrate interaction at the heptapeptide cleavage sequence, relegating the rest of the substrate molecule to an indifferent state, unimportant in defining cleavage.

In the case of the pAS substrate, a suboptimal configuration may be assumed by the heptapeptide sequence, and secondary interactions, not important in pAS-58/30 cleavage, come into play. Additional sequences contained in the proteinase polyprotein might interact with the substrate, or with other regions of the polyprotein, to enhance or inhibit proteinase–substrate contact. Thus, various proteinase polyproteins might have different affinities for a given substrate. The final cleavage profile obtained would reflect the summation of these various interactions.

Two substrates, derived from pTL0036 and pN71, which had no known NIa proteinase cleavage sites, were incubated with the polyprotein proteinase forms to determine whether cryptic cleavage existed which could be cleaved by only one of these precursor forms. A possible cleavage site in the carboxyl-terminal portion of the TEV 50K protein (TEV amino acids 1110 to 1111) has been proposed by Lain et al. (1989). We found no cleavage by any proteinase form, even when NIa proteinase was added in a large excess to the substrate synthesized from pTL0036, as shown in Fig. 3(b). This putative cleavage sequence is present in substrates derived from pAS-58/30 and pAS, but we were also unable to detect a second cleavage event in either of these substrates (data not shown). It is unlikely that a functional NIa cleavage site is located in this region of the TEV genome-derived polyprotein. However, we cannot rule out the possibility that substrate molecules larger than those used in this study are necessary to direct cleavage, as has been shown in processing studies of the poliovirus P1 protein (Ypma-Wong & Semler, 1987). Additionally, this region of the 50K protein may be located on the interior of the protein and be inaccessible to the proteinase in the absence of an accessory factor. This factor, not present in our cell-free processing system, may be required to facilitate cleavage, as has been demonstrated for cowpea mosaic virus (Vos et al., 1988).

In the poliovirus system, proteolytic processing by proteinase precursor forms has been shown to be important in cleavage of viral polyproteins in the generation of specific viral proteins. In some picornaviruses (poliovirus), the 3CD polyprotein is required for efficient cleavage during the generation of capsid protein subunits VP0, VP3 and VP1 (Jore et al., 1988; Ypma-Wong et al., 1988), while other members of this virus family (genera aphtho- and cardiovirus) appear to produce capsid proteins quite efficiently via cleavage by 3C proteinase alone, or as a component of a larger polyprotein such as 3ABC (reviewed by Harris et al., 1990; Palmenberg, 1990).

In comparing the results presented here with the picornaviral studies, several similarities and differences are apparent. Both virus groups make extensive use of a proteinase in a polyprotein form. In the case of potyviruses, the 49K NIa protein is a VPG-proteinase or 3BC polyprotein; in picornaviruses it is often a proteinase-replicase (3CD) polyprotein. In the poliovirus studies, precursor proteinase forms are necessary to mediate the cleavage required to generate capsid protein subunits, while the single cleavage necessary to generate the TEV capsid appears to proceed efficiently with a variety of proteinase forms, although the potyvirus 49K polyprotein is probably involved in vivo. The cleavage of the pAS substrate by the various proteinase forms reflects differences in the rate and extent of processing, not an absolute requirement for a particular proteinase form. This has also been exhibited in in vitro studies of poliovirus, where the 3C proteinase, when added in large excess, was able to cleave the P1 precursor (Nicklin et al., 1988). The differential processing of the 50K/71K cleavage site sequence found in pAS-derived substrate would be analogous to a faster cleavage by a 3BCD precursor (p49 + 58), with slower cleavage by the 3C proteinase (pN27).

It must be emphasized that the cell-free studies reported here are based on artificial proteinase forms, which may or may not reflect proteins actually present in significant amounts in infected cells. With this caveat in mind, we would suggest that during TEV genome expression and replication, autocatalysis of the 49K NIa proteinase from the genome-derived polyprotein is essentially complete and likely to be cotranslational, with larger precursor molecules usually undetectable. The cleavage of the 49K protein to the N-terminal VPg and C-terminal 27K proteinase proceeds much more slowly. Therefore, it is probable that in TEV-infected cells, cleavage of the genome-derived polyprotein is directed by the abundant 49K polyprotein or 27K NIa proteinase and that larger polyprotein precursors play a minimal role in gene expression.

We gratefully acknowledge Stephen S. Whitehead for his invaluable assistance with the statistical analysis. This work was supported by a grant from the Department of Energy to W.G.D. This is Oregon Agricultural Experiment Station Paper number 9637.

References


(Received 8 July 1991; Accepted 23 September 1991)