

Proteolytic activity of plum pox virus–tobacco etch virus chimeric NI_a proteases

Juan Antonio García and Sonia Lain

Centro de Biología Molecular (CSIC-UAM), Universidad Autónoma de Madrid, Canto Blanco, 28049 Madrid, Spain

Received 7 December 1990; revised version, received 5 February 1991

Plasmids encoding chimeric NI_a-type proteases made of sequences from the potyviruses plum pox virus (PPV) and tobacco etch virus (TEV) have been constructed. Their proteolytic activity on the large nuclear inclusion protein (NI_b)-capsid protein (CP) junction of each virus was assayed in *Escherichia coli* cells. The amino half of the protease seemed to be involved neither in the enzymatic catalysis nor in substrate recognition. In spite of the large homology among the PPV and TEV NI_a-type proteases, the exchange of fragments from the carboxyl halves of the molecules usually caused a drastic decrease in the enzymatic activity. Inactive chimeric proteases did not interfere with cleavage by PPV wild type protease expressed from a second plasmid. The results suggest that the recognition and catalytic sites of the NI_a proteases are closely interlinked and, although residues relevant for the correct interaction with the substrate could be present in other parts of the protein, a main determinant for substrate specificity should lie in a region situated, approximately, between positions 30 and 90 from the carboxyl end. This region includes the conserved His at position 360 of PPV or 355 of TEV, which has been postulated to interact with the Gln at position -1 of the cleavage sites.

Viral protease; Potyvirus; Chimeric proteins

1. INTRODUCTION

Plum pox virus (PPV) is a member of a potyvirus group and expresses its genetic information through a polyprotein of about 355 kDa that is proteolytically processed, probably following a regulated pathway [1,2]. The virus-encoded NI_a protein is responsible for the processing at the central and carboxy-terminal regions of the PPV polyprotein [1,3]. The cleavage sites are characterized by conserved heptapeptides [3]; the one corresponding to the NI_b-CP junction is still recognized when introduced in a non-specific position of the polyprotein [4].

The most extensively studied potyviral protease is the NI_a or 49 kDa protein of tobacco etch virus (TEV) [5]. Abundant information has been obtained on the nature and specificity of the cleavage sites [6-9] as well as on the mechanism of catalysis and the amino acid residues involved in it [5,10-11]. The potyviral NI_a-type protease has been defined as a trypsin-like protease but with a Cys as the active site nucleophile [11]. His-234, Asp-269 and Cys-339 and His-239, Asp-274 and Cys-344 have been postulated as the catalytic residues of the TEV 49 kDa [11] and PPV NI_a [1] proteases, respectively. However, no experimental data are available on the residues of the NI_a protease involved in the specific recognition of the substrate. Although TEV and PPV

NI_a proteases are highly homologous [1] and the cleavage sites of both viruses share some features [12], the seven amino acid cassette of the TEV NI_b-CP junction was not efficiently processed by the PPV protease [4]. In this paper we describe the construction of PPV-TEV chimeric proteases and their proteolytic activity on PPV and TEV substrates in order to study the compatibility of functional domains from different proteins as well as to gain insight into the protein regions responsible for the specificity of the enzymatic reaction.

2. MATERIALS AND METHODS

2.1. Bacterial strains and plasmids

Cloning of plasmids was done in *E. coli* JM109 [13]. *E. coli* strain CJ236 [14] was used in site directed mutagenesis experiments.

PPV and TEV cDNA fragments were obtained from pPPV12 [13], pTL0059 and pTL5495 (purchased from the American Type Culture collection). Several restriction sites were introduced in the PPV and TEV sequences by site directed mutagenesis, employing the method described by Kunkel et al. [14], to facilitate the correct exchange of fragments. A *Sall* site was generated on a PPV *Bam*HI-*Hind*III fragment (nt 6933-8019) at position 7210 using the oligodeoxynucleotide pCAAAAGAGTCGACTAAATAAG; pPPV12s was obtained by substituting the *Bam*HI-*Nco*I fragment (nt 6933-7677) with the new *Sall* site, for the corresponding one of pPPV12 [1]. A *Sma*I site was generated on a TEV *Hinc*II-*Bam*HI fragment (nt 4045-5915) at position 4217 using the oligodeoxynucleotide pCATGAATTCCTGGGAAAGTC. A *Stu*I site was introduced on a TEV *Bam*HI-*Sall* fragment (nt 5915-7167) at position 6275 using the oligodeoxynucleotide pGTTTAAAGGCCTACGTGA. *Pvu*I, *Xho*I and *Bam*HI restriction sites were independently generated at positions 6429, 6730 and 6891 on the TEV *Bam*HI-*Sall* fragment (nt 5915-7167) carrying

Correspondence address: J. A. García, Centro de Biología Molecular (CSIC-UAM), Universidad Autónoma de Madrid, Canto Blanco, 28049 Madrid, Spain

the previously introduced *Sma*I mutation, using the oligodeoxynucleotides pTGGTCCGATCCGCTACATG, pTATCAACTCCGAG-ATGG and pTGTGGGGATCCCATAAAG, respectively.

In all cases, viral cDNA beginning at the *Sma*I site of PPV (nt 4238) or of TEV (nt 4217, introduced by mutation) was ligated to the *Bam*HI site of pUC8 filled in with Klenow enzyme, in phase with the β -galactosidase gene of the vector. Chimeric plasmids contained different combinations of TEV and PPV restriction fragments (see Fig. 1). The names of the clones specify the number of amino acids from the amino end of the PPV NI₁ protein (P number T) or of the TEV 49 kDa protein (T number P) included in the chimeric protease encoded by the plasmid. Two series of plasmids differing in the 3' region of the viral insert were constructed: p and t series (noted at the end of the name of the plasmids) contained the *Sall*-*Pst*I fragment of pPPV12s (nt 7209 to the 3' end) and of pTL5495 (nt 7168-9180), respectively (see Fig. 1). The correct ligation of fragments was checked by restriction analysis and sequencing across the ligation points. The numbering of the nucleotide sequences of PPV and TEV cDNA was taken from Lain et al. [12] and Allison et al. [16], respectively, and that of the amino acid sequences of the NI₁ proteases from Garcia et al. [1].

The medium copy number plasmid pPPVS20 has been previously described [15].

2.2. Immunological detection of viral polypeptides

The preparation of the extracts, separation on SDS-polyacrylamide gels and immunodetection of proteins transferred to nitrocellulose membranes (0.2 μ m) were as previously described [1]. The anti-PPV CP serum was obtained from rabbits immunized with CP from disrupted virions. The anti-TEV serum was purchased from the American Type Culture Collection.

3. RESULTS

3.1. Chimeric plasmids

Fig. 1 shows the characteristics of the chimeric plasmids employed in this study. Plasmids of p and t series differ in the 3' terminal region of the cloned cDNA, which codes for NI₁ and CP sequences. The polyprotein encoded by the p plasmids contains most of the PPV NI₁ protein and the complete CP and, thus, the PPV NI₁-CP junction. Plasmids of the t series encode TEV NI₁ and CP sequences, including the NI₁-CP cleavage site, but CP is truncated at amino acid 221 and fused to 13 amino acids encoded by vector sequences.

The amino region of the polyprotein encoded by pP plasmids consists of a 50.4 kDa amino-truncated CI protein and the 6 kDa peptide from PPV and different chimeric proteases formed by amino-terminal sequences of PPV and carboxy-terminal ones of TEV. On the contrary, pT plasmids code for a 49.6 kDa truncated CI and the 6 kDa TEV proteins and chimeric protease, with the amino region from TEV and the carboxyl one from PPV. A Pro $\rightarrow$ Leu mutation at position 196 of the TEV protease, originated by the introduction of a *Stu*I site (see section 2), was present in all pT plasmids.

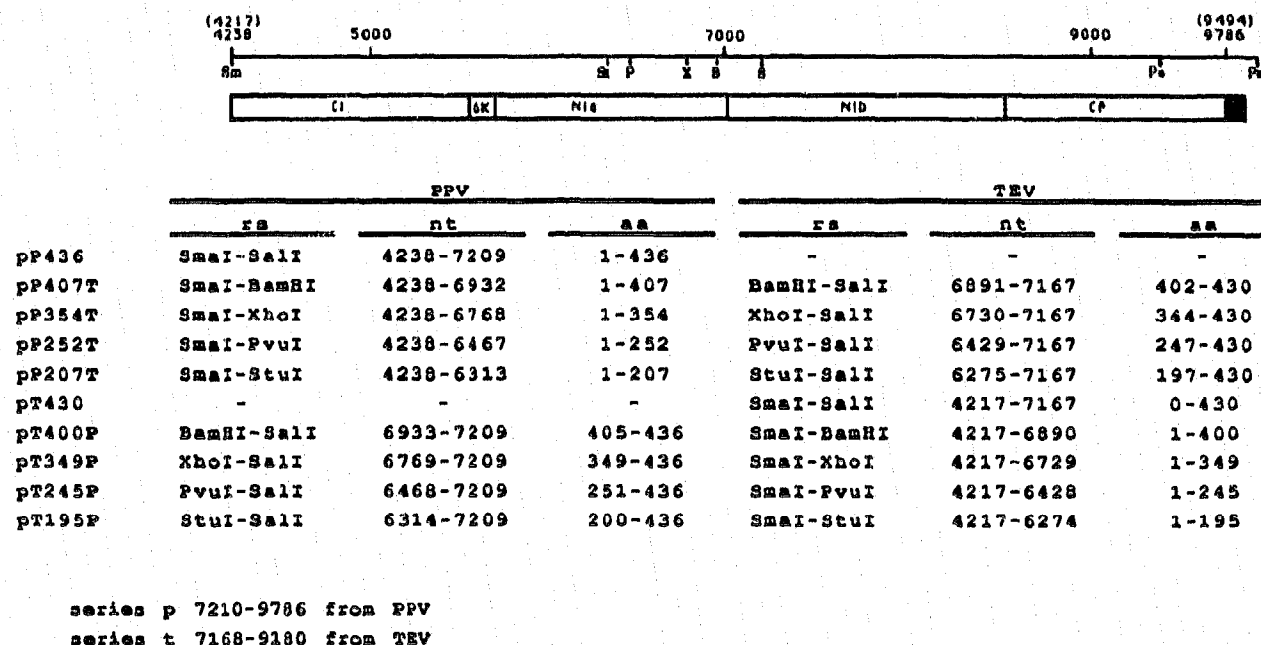


Fig. 1. Characteristics of the PPV-TEV chimeric cDNA clones used in this study. A cistron map is shown on the top of the Figure; length scale indicates the PPV nucleotides and, in parentheses, the ends of cloned TEV cDNA. Restriction endonuclease sites relevant for the constructions are also indicated: Sm = *Sma*I; St = *Stu*I; P = *Pvu*I; X = *Xho*I; B = *Bam*HI; S = *Sall*; Ps = *Pst*I. *Pst*I site at position 9180 of TEV cDNA was not present in PPV cDNA. The Figure shows the restriction fragments (rs) derived from each viral cDNA which make up the 5' region of the plasmid inserts, and the nucleotides that they comprise (nt). The specific amino acids of the regions of the chimeric proteases corresponding to each virus are presented in the column aa. Because of the sequence homology between PPV and TEV proteases, some amino acids at the junction should be considered as included in both regions of the chimeric proteins although no duplications or extra residues were produced in any of the constructions. The origin of the 3' region of the cDNA sequence of each series of plasmids is shown at the bottom of the Figure. Details of the construction of the plasmids are given in section 2.

3.2. Cleavage at PPV NI_b-CP junction

Cleavage at the PPV NI_b-CP junction was evaluated by immunodetection of the free CP released from the polyprotein, in Western blots of extracts of *E. coli* cells harboring the plasmids encoding the chimeric protease. Essentially, complete processing at the NI_b-CP site was carried out by the proteases encoded by pP436p, pT195Pp and pT245Pp which contain the complete sequence, amino acids 200-436 and amino acids 251-436 of the PPV NI_a protease, respectively (Fig. 2A, lanes a, j and l, Fig. 4). The PPV region of the T195P protease includes His-239, Asp-274 and Cys-344 which presumably form the catalytic triad [2], as well as His-360 which has been implicated in substrate binding [17]. However, the catalytic His is TEV-encoded (residue 234) in T245P protease, the remaining residues mentioned above being PPV-encoded. The protease encoded by pP407Tp, which consists of the PPV NI_a protein with the 29 carboxy-terminal amino acids replaced by the corresponding ones of the TEV protease, was on-

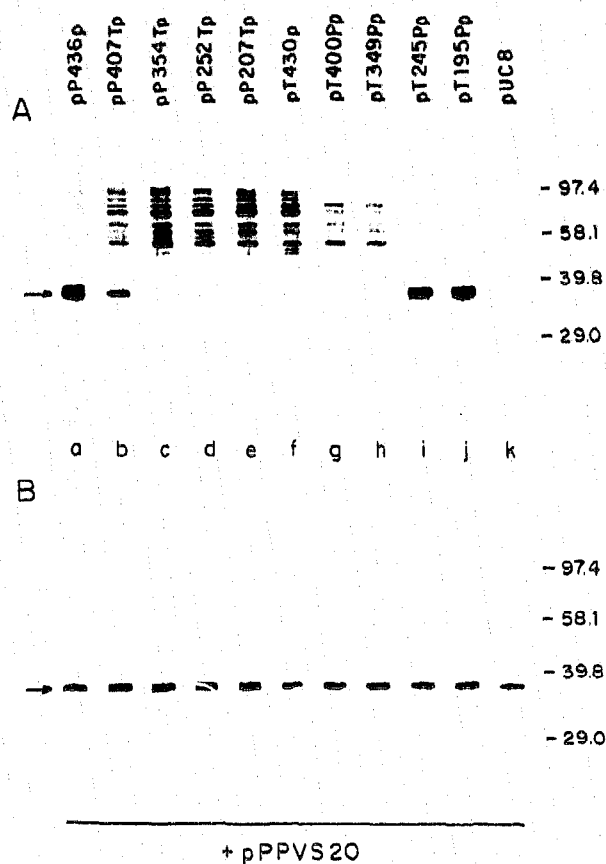


Fig. 2. Cleavage at PPV NI_b-CP junction by PPV and TEV NI_a proteases. Extracts were prepared from cells harboring the indicated plasmids either alone (A) or in addition to pPPVS20 (B), transferred to nitrocellulose membranes after electrophoresis and probed with anti-PPV CP serum. The molecular masses of biotinylated protein markers are indicated beside the panels. The arrow shows the position of PPV CP.

ly partially active (Fig. 2A, lane b, Fig. 4), although it contained the catalytic triad and His 360 from PPV. The confusing pattern of uncleaved products immunoreacting with anti-CP serum seems to mainly originate from internal initiations of translation along the PPV ORF [1] and cleavages at other sites of the polyprotein (not shown). The TEV protease and the rest of the chimeric proteins were unable to cleave the PPV NI_b-CP junction (Fig. 2A, lanes c-h, Fig. 4).

Cleavage at the NI_b-CP site of the polyproteins encoded by the different pP and pT plasmids of the p series by exogenous wild type PPV protease encoded by the medium copy number pPPVS20 was assayed in *E. coli* cells harboring two plasmids. Fig. 2B shows that proteolytic processing took place in all cases; the efficiency of the cleavage performed by the wild type protease was not affected by the expression of any of the chimeric ones.

3.3. Cleavage at TEV NI_b-CP junction

The immunodetection with anti-TEV serum of the truncated TEV CP released from the viral polyproteins encoded by the different plasmids of the t series allowed the evaluation of the proteolytic activity of the chimeric proteases on the TEV NI_b-CP site. When processing at NI_b-CP junction did not take place, a complex pattern, similar to that revealed with anti-PPV CP serum for the p series (see above), was detected with anti-TEV serum.

The TEV 49 kDa protease encoded by pT430t plasmid had a Pro→Leu mutation at position 196 which was also present in the rest of the pT plasmids (see section 3.1). This mutation did not affect the activity of the TEV 49 kDa protein on the TEV NI_b-CP cleavage site (Fig. 3, lane f, Fig. 4). The protease encoded by

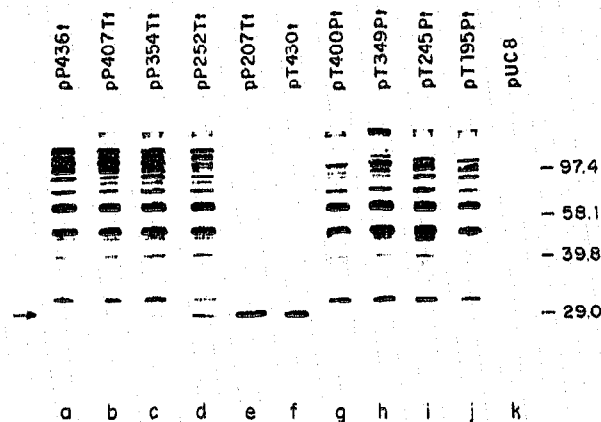


Fig. 3. Cleavage at TEV NI_b-CP junction by PPV and TEV proteases. Extracts were prepared from cells harboring the indicated plasmids, transferred to a nitrocellulose membrane after electrophoresis and probed with anti-TEV serum. Half of the amount of the rest of the samples was loaded on lanes e and f. The molecular masses of biotinylated protein markers are indicated beside the panels. The arrow shows the position of the truncated TEV CP.

pP207Tt, which contains amino acids 197-430 of the TEV 49 kDa protein, including the catalytic triad and His-355, was as active as the complete TEV protease (Fig. 3, lanes e,f, Fig. 4). The replacement of a TEV region containing the catalytic His-234 by the corresponding one of PPV (P252T protease) produced a drastic decrease of cleavage at TEV NI_b-CP site (Fig. 3, lane d, Fig. 4), in contrast with the normal activity on the PPV NI_b-CP site of the converse exchange (T245P, Fig. 2A, lane i, Fig. 4). Also contrasting with the PPV activity, a TEV protease with the 30 carboxy-terminal amino acids replaced by the PPV ones (T400P protease) was essentially inactive for processing at the TEV NI_b-CP junction (Fig. 3, lane g, Fig. 4).

Plasmid pP354T codes for a chimeric protease with the complete catalytic triad of PPV but with the 87 carboxy-terminal residues, including His-355, specific of TEV. We could not detect proteolytic activity of P354T protease on the PPV NI_b-CP site (Fig. 2A, lane c). However, it processed, although with very low efficiency, the TEV cleavage site (Fig. 3, lane c, Fig. 4). This low activity disappeared with further replacement of TEV amino acids by PPV ones; no cleavage at the TEV NI_b-CP junction by P407T protease, with only the 29 carboxy-terminal amino acids of TEV origin, was detected (Fig. 3, lane b, Fig. 4). No activity on the TEV NI_b-CP site could either be attributed to T349P protease, the converse chimera to P354T, with the catalytic triad of TEV and 89 carboxy-terminal amino acids of PPV, or to T245P and T195P proteases (Fig. 3, lanes

h-j, Fig. 4). Surprisingly, the wild type PPV NI_a protease seemed to be able to recognize the TEV NI_b-CP site, but with even less efficiency than P354T protease, since a very faint band at the position of the free truncated TEV CP could be detected when a large amount of bacterial extract was applied to the gel (the band is difficult to see in Fig. 3, lane a).

4. DISCUSSION

The results of sequence analysis and biochemical experimentation suggest that the potyviral NI_a proteases are a subgroup of trypsin-like proteases [17,18]. The NI_a proteins of TEV and PPV share 50.6% of identical amino acids [1] and recognize cleavage sites defined by conserved heptapeptides. Although their consensus sequences are different, several features, such as the presence of polar residues, mainly Glu, at position -6, hydrophobic amino acids at position -4, Gln at position -1 and small amino acids at position +1, are common to both viruses [12]. In spite of these similarities, the proteolytic processing is virus specific: the inability of PPV NI_a protease to cleave an artificial TEV site has been previously reported [4] and only extremely poor activity could be detected on the natural TEV NI_b-CP junction (Fig. 3, lane a, Fig. 4). No cleavage by the TEV protease of the PPV NI_b-CP site could be observed (Fig. 2A, lane f, Fig. 4), even when the electrophoresis gels were overloaded.

Chimeric proteins consisting of PPV and TEV protease sequences have been constructed and the results of the analysis of their proteolytic activities on the PPV and TEV NI_b-CP junctions are summarized in Fig. 4. Since the chimeric proteases are flanked by one cleavage site of each virus, the release from the polyprotein of the protease depended on its proteolytic activity at each cleavage site. However this should not be relevant for the analysis of the proteolytic activity on the NI_b-CP junction since normal trans cleavage at this site has been reported for potyviral NI_a proteases which form part of a polyprotein [1,10]. Chimeric proteins consisting of about half of each protease (P207T, T195P) were fully active on the NI_b-CP cleavage site of the virus supplying the carboxyl half of the chimera, which contains the catalytic triad, and inactive on that of the virus providing the amino half (Fig. 2A, Fig. 3, lanes e and j, Fig. 4). This result is in agreement with the activity in a cell-free extract of a mutant TEV protease lacking the 202 amino terminal residues [5]. Although a role in the *in vivo* processing cannot be ruled out, the *in vitro* and in *E. coli* results, together with the recent mapping of the potyviral VPg at the amino terminus of the NI_a protein [19,20], suggest that the amino half of the NI_a protease is not relevant for the proteolytic activity.

The efficient cleavage on the PPV NI_b-CP site by the T245P protease, which contains the catalytic Asp and

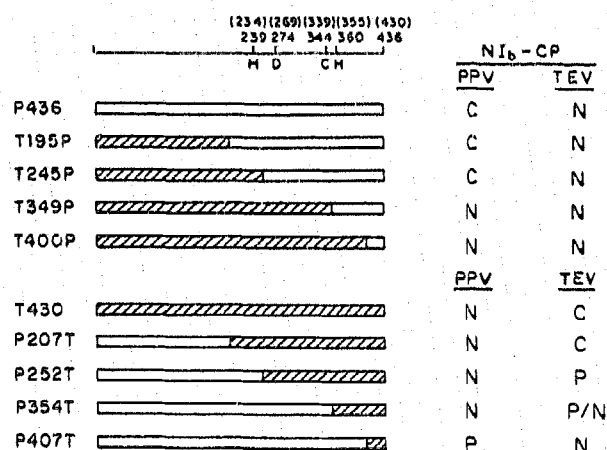


Fig. 4. Summary of the proteolytic activity of several PPV-TEV chimeric proteases. The location of the catalytic triad and the His postulated to interact with Gln at position -1 of the cleavage site is indicated in a map of the NI_a-type protease, at the top of the Figure. The number of these residues in the PPV and TEV (in parentheses) proteases is also shown. The PPV and TEV regions of the chimeric proteins are represented by open and hatched bars, respectively. P436 and T430 are the wild type PPV NI_a protease and the 49 kDa TEV protease with a P→L substitution at position 196, respectively. The level of proteolytic activity of the different chimeric proteases on the PPV or TEV NI_b-CP junctions is symbolized by C (complete), P (partial), and N (no or very low activity).

Cys from PPV and the catalytic His from TEV (Fig. 2A, lane i, Fig. 4) indicates that structural similarity among PPV and TEV NI_a proteases is enough to allow the coupling of fragments from different viruses in an active catalytic center. However, the decrease in proteolytic activity of the converse chimera, P252T, on the TEV NI_b-CP site (Fig. 3, lane d, Fig. 4), shows the existence of restrictions to the compatibility. On the other hand, both chimeric proteases maintain substrate specificity, only cutting when the carboxyl region of the protein and the cleavage site belong to the same virus (Fig. 4). Thus, these proteases do not require their own catalytic His (239 in PPV; 234 in TEV) for the specific recognition of the substrate. However, some involvement of the catalytic His in substrate recognition cannot be ruled out since preliminary results seemed to indicate that T245P protease (with TEV His-234) cleaved, with very low efficiency, at the TEV 6 kDa-NI_a site (but not at the TEC CI-6 kDa site), while cleavage at this junction could not be detected in cells producing T195P protease (with PPV His-239) (unpublished results).

A drastic drop in activity was observed in chimeric proteases consisting of a large amino-terminal fragment, including the complete catalytic triad, from one virus and a small carboxyl one, including His-360 (PPV) or -355 (TEV), from the other (T349P and P354T), indicating that a precise interaction between the catalytic site and the carboxyl region of the protein is required for the enzymatic activity. However, the P354T protease retained a very low, but reproducibly observed, activity at the NI_b-CP cleavage site of TEV, the virus that codes for its 86 carboxy-terminal amino acids, while it was unable to cut at the NI_b-CP junction of PPV, the virus from which its catalytic center derives (Fig. 2A, Fig. 3, lane c, Fig. 4). On the other hand, sequences preceding the 29 carboxy-terminal amino acids confer specificity of substrate recognition since the P407T chimera, the PPV protease with the last 29 amino acids replaced by the corresponding TEV ones, was partially active on the PPV NI_b-CP site, but inactive on the TEV site (Fig. 2A, Fig. 3, lane b, Fig. 4). Some deletion/substitution mutations have been previously described at this region of the PPV protease [1,15]; it is noteworthy that Δ B7 protease, with 16 non-specific amino acids instead of the 30 carboxy-terminal ones, was almost as active as the wild type NI_a protein [1,15], being clearly more efficient than P407T protein, in which the carboxy-terminal amino acids were replaced by homologous ones. Thus, structural rather than amino acid sequence features are required on the carboxyl end of the molecule. In this sense, the lack of activity of the T400P protease on both substrates (Fig. 2A, Fig. 3, lane g, Fig. 4) seems to indicate more strict structural requirements at the carboxy-terminal region of the TEV protease. A deletion mutant of the TEV 49 kDa protease, lacking the last 30 amino acids, was

reported to be unable to cleave in trans at the NI_b-CP site, but conserved some activity for cis processing at its amino border [5]; we could not detect autoprocessing activity of the T400P chimeric protease (unpublished results).

PPV NI_a protease expressed from an independent medium copy number plasmid was able to cleave at the PPV NI_b-CP sites of all the chimeric constructions (Fig. 2B), indicating that the inactive hybrid proteases did not interfere with the recognition and processing of the substrate by the wild type protease.

In summary, although the potyviral NI_a proteases, particularly the PPV ones, are sufficiently malleable to allow some homologous exchanges of fragments, that may be of concern to the catalytic region, they seem to have stringent structural requirements which are easily disturbed in chimeric constructions, even when they are built with proteins so homologous as PPV NI_a and TEV 49 kDa proteases. The results described in this paper suggest that the sequences involved in substrate recognition lie in the carboxyl region of the protease. In particular, the activity of the P354T protease on the TEV NI_b-CP site (but not on the PPV one) and of P407T protease on the PPV NI_b-CP site (but not on the TEV one) seem to indicate that there are important determinants for substrate specificity in the region located between amino acids 354 and 407 of PPV protease, including His-360 (355 in TEV) which has been postulated to be involved in the interaction with Gln at position -1 of the cleavage sites [17]. However, the scarcity of the activity of P354T and the ability of PPV NI_a protease to recognize the TEV NI_b-CP site, although it took place at an almost undetectable level, makes our conclusion only tentative. In any case, the close and precise interaction which seems to connect the regions involved in the recognition of the substrate and in catalysis makes the possibility very unlikely that they exist as two different functional domains.

Acknowledgements: We thank Dr. J. Ortín, M. Serrano and J.L. Riechmann for critical reading of the manuscript. This investigation was aided by grant B1088-0257 from CICYT and an institutional grant from Fundación Ramón Areces. S.L. was recipient of a fellowship from Fondo de Investigaciones Sanitarias.

REFERENCES

- [1] García, J.A., Riechmann, J.L. and Laín, S. (1989) *Virology* 170, 362-369.
- [2] García, J.A., Laín, S., Cervera, M.R., Riechmann, J.L. and Martín, M.T. (1990) *J. Gen. Virol.* 71, 2773-2779.
- [3] Martín, M.T., López-Ortín, C., Laín, S. and García, J.A. (1990) *Virus Res.* 15, 97-106.
- [4] García, J.A., Riechmann, J.L. and Laín, S. (1989) *J. Virol.* 63, 2457-2460.
- [5] Carrington, J.C. and Dougherty, W.G. (1987) *J. Virol.* 61, 2540-2548.
- [6] Carrington, J.C. and Dougherty, W.G. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3391-3395.

- [7] Dougherty, W.G., Carrington, J.C., Cary, S.M. and Parks, T.D. (1988) *EMBO J.* 7, 1281-1287.
- [8] Dougherty, W.G., Cary, S.M. and Parks, T.D. (1989) *Virology* 171, 356-364.
- [9] Dougherty, W.G. and Parks, T.D. (1989) *Virology* 172, 145-155.
- [10] Carrington, J.C., Cary, S.M. and Dougherty, W.G. (1988) *J. Virol.* 62, 2313-2320.
- [11] Dougherty, W.G., Parks, T.D., Cary, S.M., Bazan, J.F. and Fletterick, R.J. (1989) *Virology* 172, 302-310.
- [12] Lain, S., Riechmann, J.L. and Garcia, J.A. (1989) *Virus Res.* 13, 157-172.
- [13] Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* 33, 103-119.
- [14] Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) In: *Methods in Enzymology*, vol. 154, (Wu, R. and Grossman, L., eds) Academic Press, New York, pp. 367-382.
- [15] Garcia, J.A., Riechmann, J.L., Martin, M.T. and Lain, S. (1989) *FEBS Lett.* 257, 269-273.
- [16] Allison, R., Johnston, R.E. and Dougherty, W.G. (1986) *Virology* 154, 9-20.
- [17] Bazan, J.F. and Fletterick, R.J. (1988) *Proc. Natl. Acad. Sci. USA* 85, 7872-7876.
- [18] Gorbatenya, A.E., Donchenko, A.P., Blinov, V.M. and Koonin, E.V. (1989) *FEBS Lett.* 243, 103-114.
- [19] Shahabuddin, M., Shaw, J.G. and Rhoads, R.E. (1988) *Virology* 163, 635-637.
- [20] Murphy, J.F., Rhoads, R.E., Hunt, A.G. and Shaw, J.G. (1990) *Virology* 178, 285-288.