Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Proteinases Involved in Plant Virus Genome Expression

JUAN ANTONIO GARCÍA, MARÍA ROSARIO FERNÁNDEZ-FERNÁNDEZ, AND JUAN JOSÉ LÓPEZ-MOYA

Centro Nacional de Biotecnología (CSIC), Campus de la Universidad Autónoma de Madrid, 28049 Madrid, Spain

I. INTRODUCTION

Viruses that infect plants have genomes of rather small size that are expressed following very different strategies, which in many cases involve proteolytic processing of protein precursors (Zaccomer *et al.*, 1995; Maia *et al.*, 1996b). Host proteinases have been shown to participate in the maturation of animal virus proteins that are targeted to the viral envelope (see chapter by Kido). Although these events should also take place in plants, they have not yet been studied in detail in viruses infecting plants, probably due to the scarcity of enveloped plant viruses. For this reason, this chapter focuses only on virus-encoded proteinases.

The genome of most plant viruses consists of one or various single-stranded (ss) RNA molecules of positive (+) polarity. Despite their large diversity of genome organization and virion morphology, nucleotide sequence data have revealed that most eukaryotic (+)RNA viruses (irrespective of whether they infect plants or animals) can be classified into two large "supergroups" (picornalike and alphalike) and a limited number of less well-defined minor

Proteases of Infectious Agents

"supergroups" (as carmolike and sobemolike) (Goldbach *et al.*, 1991). Table I lists plant virus groups in which virus-encoded proteinases have been identified, or its existence has been suggested on the basis of sequence analysis. Cysteine, serine, and serinelike (with a cysteine in the active center) proteinases, but not aspartic proteinases nor metalloproteinases, have been shown to be involved in the proteolytic processing of protein precursors encoded by plant (+)RNA viruses (Maia *et al.*, 1996b).

The genomic RNAs of picornalike viruses encode large polyproteins that are processed by one or more virus-encoded proteinases. Proteolytic processing of polyproteins might not only determine the time of appearance of each gene

TABLE I Genera of Plant Viruses That Code for Demonstrated or Putative Proteinases

Virus genus	Proteinase class[a]
Picornalike	
Potyviridae	
Potyvirus	Ser, Cys, Ser-like
Rymovirus	Ser, Cys, Ser-like
Bymovirus	Cys, Ser-like
Comoviridae	
Comovirus	Ser-like
Nepovirus	Ser-like
Sequiviridae	
Sequivirus	Ser-like
Waikavirus	Ser-like
Alphalike	
Tymovirus	Cys
Marafiviruses	Cys
Carlavirus	Cys
Capillovirus	Cys
Trichovirus	Cys
Benyvirus	Cys
Closterovirus	Cys
Sobemolike	
Luteovirus (subgroup II)	Ser
Sobemovirus	Ser
Pararetrovirus	
Caulimoviridae	Asp

[a]The class of proteinase(s) that the viral protein(s) is associated (or proposed to be associated) with is presented. Abbreviations: Ser, serine; Cys, cysteine; Ser-like, serine-like with cysteine in the active center; Asp, aspartic.

product but also regulate its activity or subcellular localization by removal of functional domains.

Genes of alphalike, carmolike, and sobemolike viruses are expressed through different combinations of strategies acting at the level of transcription (synthesis of subgenomic (sg) RNAs) and/or translation (alternative translation initiation sites, frameshifting, and readthrough at suppressible termination codons). In some cases, gene expression is completed posttranslationally by proteolysis mediated by a virus-encoded proteinase.

Proteolytic processing of protein precursors has also been observed in plant pararetroviruses. These viruses have double-stranded (ds) DNA genomes that replicate via RNA intermediates (Hohn and Fütterer, 1997). In this case, the virus-encoded proteinase belongs to the aspartic family and a host cysteine proteinase seems to be involved in the maturation of a viral protein.

Here, we have attempted to give an overall view of the use of proteolytic processing by different plant virus groups for the expression of their genomes.

II. PICORNALIKE SUPERGROUP

A. POTYVIRIDAE

The family Potyviridae includes three definite genera, the monopartite poty- and rymoviruses and the bipartite bymoviruses and some still unclassified viruses such as those tentatively designated ipomoviruses and macluraviruses (Murphy *et al.*, 1995). All potyviruses have similar genome organizations (Fig. 1). Each potyviral genomic RNA encodes all its proteins in single open reading frames (Shukla *et al.*, 1994).

The proteolytic processing of the polyprotein into which potyvirus RNA is translated has been extensively studied both *in vitro* (using cell-free translation systems) and in heterologous (*Escherichia coli*) or homologous (infected and transgenic plants) *in vivo* systems (reviewed by Riechmann *et al.*, 1992). Two potyviral proteinases located at the N-terminus of the polyprotein, P1 (Carrington *et al.*, 1990; Verchot *et al.*, 1991) and HC (Carrington *et al.*, 1989a), autocatalytically cleave at their respective C-termini, whereas the NIa proteinase processes intra- and intermolecularly the rest of the polyprotein sites (Carrington and Dougherty, 1987; Hellmann *et al.*, 1988; García *et al.*, 1989b; Ghabrial *et al.*, 1990). The three potyviral proteinases have RNA binding activity that, in contrast with the RNA affinity of the picornaviral 3C proteinase, seems to be nonspecific (Brantley and Hunt, 1993; Maia and Bernardi, 1996; Daros and Carrington, 1997). Most potyviral proteins have been shown to be multifunctional. Thus, RNA binding and proteolysis might be independent activities of the potyviral proteinases.

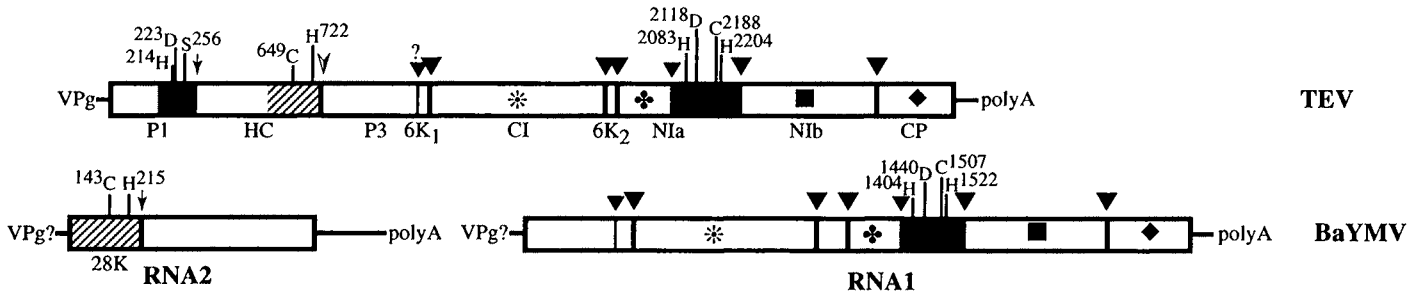

FIGURE 1 Genomic maps of a potyvirus (TEV) and a bymovirus (BaYMV). Nontranslated regions and ORFs are represented by horizontal lines and open boxes, respectively. Arrowheads indicate the cleavage sites recognized by serine-type (P1) (➛), papain-like (HC) (➢) and 3C-like (NIa) (▼) proteinases. Vertical lines separate the products of the polyprotein proteolytic processing. A thin line between P3 and $6K_1$ indicates that cleavage at this site is only inferred by sequence analysis and that, in the case of TEV, it does not take place *in vitro*. The proteinase domains are represented by black (NIa 3C-like), hatched (HC papain-like), and gray (P1 serine type) boxes. The positions of predicted catalytic and substrate-binding residues are shown above the map. Different symbols indicate other proteins with known or predicted functions: helicase-like (✳), VPg (✤), replicase-like (■), and coat protein (◆). Gene products mentioned in the text are written below the maps.

1. P1 Proteinase

The P1 proteinase is derived from the amino-terminal region of the polyprotein (Fig. 1). The proteolytic activity of the P1 protein resides in its C-terminal half. In spite of the high variability found among the potyviral P1 proteins, amino acids characteristic of serine proteinases have been found to be very well conserved, showing the signature $Hx_8D/Ex_{29\text{-}32}GxSG$ (D predominates over E) (Riechmann *et al.*, 1992; Ryan and Flint, 1997). The same sequence, although with slight differences in spacing, can be identified in brome streak mosaic rymovirus (BrSMV) and sweetpotato mild mottle ipomovirus (SPMMV), but not in the bipartite bymoviruses. Site-directed mutagenesis of His214, Asp223, and Ser256 of P1 from tobacco etch potyvirus (TEV) supported the assumption that this conserved motif corresponds to the active center of the enzyme (Verchot *et al.*, 1992). The sequence GxSG is identical to the consensus motif around the Ser in the active site of trypsin- and chymotrypsinlike serine proteinases (Barrett, 1986). However, the distances between His, Asp, and Ser of the proposed catalytic triad of P1 proteinases are considerably shorter than those separating the active site residues in the cellular enzymes, indicating a quite broad evolutionary distance between them.

In vitro processing experiments in a wheat germ system have shown that P1 cleaves at its carboxyl end and is unable to act in trans (Verchot *et al.*, 1992). Sequence comparisons among potyviral polyproteins show a nonstrictly conserved consensus sequence H/Q-Y/F↓S for the P1 cleavage site, which resembles those of the cellular cathepsin C and chymotrypsin serine proteinases. An unusual feature of the P1 proteinase is its inability to function in an *in vitro* rabbit reticulocyte system (Mavankal and Rhoads, 1991; Verchot *et al.*, 1991). Addition of relatively small amounts of wheat germ extract to the reticulocyte lysate promoted P1 proteinase activity, suggesting that inactivity in the latter is probably due to the lack of a host factor, rather than to the presence of an inhibitor (Verchot *et al.*, 1992). This factor should also be present in insect cells, since P1 proteinase is active when expressed using a baculovirus vector (Thornbury *et al.*, 1993). However, the fact that a chimerical plum pox virus (PPV)-TEV P1 proteinase is active in a reticulocyte lysate system (P. Sáenz and J. A. G., unpublished results) seems to indicate that the host factor is not strictly required for P1 proteolytic activity.

It has been shown that in TEV P1 acts in trans as an accessory factor for genome amplification (Verchot and Carrington, 1995b). The infectivity of P1 proteinase-debilitated TEV mutants was restored by second-site mutations that inserted a cleavage site recognized by the NIa proteinase (Verchot and Carrington, 1995a). This result indicates that proteolytic separation of P1 from the next gene product, HC, but not P1 proteolytic activity per se, is essential for virus viability in plants.

2. HC Proteinase

The HC protein is adjacent to the C-terminus of P1 proteinase (Fig. 1). This protein, which often aggregates into amorphous inclusion bodies (De Mejia *et al.*, 1985), was first identified as a factor required for aphid transmission of the virus (helper component) (Berger *et al.*, 1989), but later on it has been shown to be involved in efficient genome amplification (Atreya *et al.*, 1992) and cell-to-cell (Rojas *et al.*, 1997) and long-distance (Cronin *et al.*, 1995; Kasschau *et al.*, 1997) movement (reviewed by Maia *et al.*, 1996a). Recent reports indicate that HC behaves as a broad-range pathogenicity enhancer (Pruss *et al.*, 1997; Shi *et al.*, 1997).

A papainlike proteinase domain has been localized to the C-terminal half of HC protein (Carrington *et al.*, 1989a). In TEV, the catalytic dyad was shown by sequence comparisons and mutagenesis analysis to be composed of Cys649 and His722 (Oh and Carrington, 1989). The HC proteinase domain is well conserved among potyviruses and may be easily aligned, specially at the sequences around the catalytic residues and at a LGxWP motif, with homologous polyprotein regions in the mite-transmitted BrSMV rymovirus and the whitefly transmitted SPMMV ipomovirus (Fig. 7). A considerable similarity was also detected between the HC proteinase domain and the N-terminal 28K protein encoded by the RNA-2 of fungus-transmitted bymoviruses. Other cysteine proteinases that bear striking resemblance to HC are encoded by closteroviruses (Agranovsky *et al.*, 1994; Karasev *et al.*, 1995) (see Section III,A) and by ORFA and ORFB of the hypovirulence-associated dsRNA virus (HyAV) of chestnut blight fungus (Choi *et al.*, 1991; Shapira and Nuss, 1991) (Fig. 7). Viral cysteine proteinases more distantly related to HC have been identified as products of different plant alphalike viruses (see Section III,B), aphtoviruses, alphaviruses, rubiviruses, arteriviruses, and coronaviruses (Gorbalenya *et al.*, 1991; Rozanov *et al.*, 1995).

The HC proteinase is responsible for processing at its own C-terminus via an autocatalytic mechanism. When analyzed *in vitro*, it exhibits little or no proteolytic activity in trans (Carrington *et al.*, 1989a,b). The TEV mutant genomes modified at the HC proteinase active site were amplification defective in protoplasts and plants. Introduction of a heterologous cleavage site recognized by the NIa proteinase at the HC C-terminus was not sufficient to restore genome amplification. In addition, an active-site mutant was not complemented by wild-type HC protein supplied in trans in transgenic plants. These results suggest that an active HC proteinase is required in cis for virus amplification (Kasschau and Carrington, 1995).

Tobacco etch potyvirus HC proteinase cleavage *in vitro* occurs at a Gly-Gly dipeptide (aa 763–764). There is a good consensus, YxVG↓G, at the presumed cleavage sites of different potyviruses, SPMMV ipomovirus and BrSMV rymovirus. The marked preference for specific amino acids at this site has been dem-

onstrated by site-directed mutagenesis (Carrington and Herndon, 1992). With the exceptions of Tyr to Phe (P4) and Val to Leu (P2) changes, which were partially tolerated, even very conservative substitutions at P4, P2, P1, and P1′ positions were found to eliminate or nearly eliminate proteolysis. Substitutions at the P5, P3, and P2′ positions permitted processing to occur, although in some cases at reduced rates. This level of specificity, that it is also shown by the potyviral NIa proteinase (see below), is not usual in viral proteinases. Interestingly, the sequence around the putative cleavage site at the C-end of the 28K proteinase of bymoviruses differs from the potyvirus HC consensus sequence, reflecting the divergence between the potyvirus HC and the bymovirus 28K proteinases.

3. NIa Proteinase

The NIa (nuclear inclusion protein a) forms, together with the RNA replicase NIb, crystalline inclusions within the nucleus of cells infected with some, but not all, potyviruses (Hiebert *et al.*, 1984). Nuclear inclusion protein a consists of an amino domain that constitutes the genome-linked protein (VPg) and a carboxyl domain that is associated with the inter- and intramolecular proteolytic activity responsible for most of the potyvirus polyprotein processing (Carrington and Dougherty, 1987; Hellmann *et al.*, 1988; Dougherty and Parks, 1991; García and Laín, 1991).

Nuclear inclusion protein a belongs to a family of viral proteinases whose archetypal member is the picornavirus 3C protein. These proteins are related to the trypsinlike family of cellular serine proteinases, but Cys replaces Ser in their active center (Bazan and Fletterick, 1988, 1990; Gorbalenya *et al.*, 1989). Sequence comparison analysis and site-directed mutagenesis enabled localization of the probable catalytic triad, composed of His, Asp, and Cys, and the His residue in the substrate-binding pocket characteristic of a Gln-x substrate specificity (Fig. 4) (Carrington *et al.*, 1988; Dougherty *et al.*, 1989b; Gorbalenya *et al.*, 1989; Bazan and Fletterick, 1990; García *et al.*, 1990; Ghabrial *et al.*, 1990). Similar catalytic triads formed by His, an acidic residue (Asp or Glu) and Cys have been suggested for other viral 3C-like proteinases (Ryan and Flint, 1997). However, crystal structure of hepatitis A virus (HAV) 3C proteinase poses some doubts on the catalytic relevance of the acidic residue and suggests that viral 3C proteinases may have a catalytic dyad rather than a triad (Bergmann *et al.*, 1997; see chapter by Bergmann and James). Interestingly, a substitution of Glu for the proposed catalytic Asp in the TEV and PPV NIa proteinases had different effects on proteolytic activity depending on the cleavage site analyzed (Dougherty *et al.*, 1989b; García *et al.*, 1990). These results suggest that there may be different structural requirements in the active center

of the proteinase for processing at different cleavage sites. The proposed catalytic triad is well conserved among genera of the family Potyviridae, although sequences around them have diverged considerably (Fig. 4). It is remarkable that the place of the His involved in recognition of Gln at position P1 of the cleavage site is occupied in SPMMV ipomovirus by an Asn residue; in agreement with this fact, cleavage sites of SPMMV NIa proteinase seem to differ from those of the rest of NIa proteinases (see below).

The NIa cleavage sites are defined by conserved heptapeptide sequences (Carrington and Dougherty, 1987, 1988; García *et al.*, 1989a; Martín *et al.*, 1990). The requirement of such an extended sequence motif is a peculiarity of the potyvirus NIa proteinases, and it is not shared by the 3C-like proteinases of other picornalike viruses. Although there is not an extended conserved cleavage motif used by all potyviruses, the NIa cleavage sites share enough features to be easily identified by sequence analysis. Position P1 is always occupied by Gln or, in much fewer sites, by Glu. The Val residue is the preferred one for position P4, which in rare occasions is occupied by another hydrophobic residue. The other positions of the heptapeptide are less conserved, although P1′ residue is most of the time Gly, Ala, or Ser, while His or an aromatic residue predominate at the P2 position and acidic or amide residues (specially Glu) are quite common at the P6 position. Site-directed mutagenesis studies of two TEV cleavage sites demonstrated that the presence of particular amino acids at positions P6, P4, P3, P1, and P1′ are essential for the cleavage site functionality (Dougherty *et al.*, 1988), whereas the amino acid at positions P5 and P2 influence the cleavage reaction profile (Dougherty *et al.*, 1989a; Dougherty and Parks, 1989). It is important to note that P6 and P3 positions are more conserved in TEV than P4 and P2 ones, whereas the opposite situation is observed in other potyviruses; thus, it is possible that the relevance of each heptapeptide sequence residue can vary depending on the particular potyvirus. Artificial NIa cleavage sites have been constructed by inserting the appropriate heptapeptide sequences in nonspecific protein regions (Carrington and Dougherty, 1988; García *et al.*, 1989a). However, sequences and/or conformational context outside the conserved heptapeptide have been shown to modulate the cleavage reaction efficiency (García *et al.*, 1992). Sequences similar to those of the potyvirus NIa cleavage sites are present at the expected positions of the polyproteins encoded by BrSMV rymovirus RNA and barley yellow mosaic virus (BaYMV) and barley mild mosaic (BaMMV) bymovirus RNA-1 molecules. However, we have not been able to identify putative cleavage sites for the NIa-like protein of SPMMV ipomovirus, suggesting that their sequences have diverged from those recognized by the potyvirus NIa. This supposition is in agreement with the lack in SPMMV of the typical His of the 3C-like proteinase substrate-binding pocket (Fig. 4).

Although both NIa proteinases and their recognition sequences have consid-

erable similarity among different potyviruses, NIa cleavage sites of each potyvirus seem to be efficiently recognized only by their own proteinases (García *et al.*, 1989a; García and Laín, 1991; Parks and Dougherty, 1991). Results from chimeric NIa proteinases (TEV-TVMV and TEV-PPV) suggest that NIa proteinase recognition and catalytic sites are closely interlinked. Several protein domains, one of them including the substrate-binding pocket His residue, appear to be important in determining substrate specificity (García and Laín, 1991; Parks and Dougherty, 1991).

The main task of NIa proteinase is to obtain the final processing products of the polyprotein. Nevertheless, regulation of the processing pathway is probably essential to: (1) synthesize the required product at the right time and place, (2) maintain functional partially processed products, and (3) control protein activity by cleavage of functional domains. Little is known about this level of regulation in potyviruses. However, some processing steps whose objective goes beyond producing the final protein products have been proposed. Cleavage between the PPV P3 protein and a putative $6K_1$ peptide has been shown to occur *in vitro* (García *et al.*, 1992). However, *in vitro* cleavage at the equivalent TEV site (discernible by sequence alignments) has not been detected (Parks *et al.*, 1992), and processing of the PPV polyprotein P3-$6K_1$ junction seems not to be essential for virus viability, although it affected virus infectivity and symptom induction. Thus, it has been suggested that the role of cleavage at the P3-$6K_1$ site, rather than producing two proteins, P3 and $6K_1$, is to regulate the activity of a single functional protein, P3-$6K_1$ (Riechmann *et al.*, 1995).

Another small peptide, $6K_2$, is placed upstream of NIa in the potyviral polyprotein. It has been shown that, whereas TEV NIa is transported to the nucleus, the $6K_2$-NIa precursor is directed to membranous structures, where potyvirus replication takes place (Restrepo-Hartwig and Carrington, 1992, 1994; Schaad *et al.*, 1997); thus, $6K_2$ probably lacks activity by itself, playing its role in the context of the $6K_2$-NIa unprocessed product. In this scenario, cleavage at the $6K_2$-NIa junction would be involved in the control of NIa activity.

Another clear example of a regulated proteolytic event is the internal cleavage that splits the VPg and proteinase domains of the NIa protein (Dougherty and Parks, 1991; Laliberté *et al.*, 1992). Although the unprocessed NIa product can readily function as a proteinase, and also it has been found covalently attached to a fraction of the TEV genomic RNAs examined by Murphy *et al.* (1990), a mutation that inhibited internal cleavage of NIa abolished TEV infectivity, indicating that proteolytic separation of the VPg and Pro NIa domains is essential for viral viability (Carrington *et al.*, 1993). The NIa internal cleavage site is processed incompletely in infected cells and inefficiently *in vitro*. Thus, whereas TEV NIa is a very abundant protein that accumulates in nuclear inclusions in infected cells, the 21K VPg is only found linked to the viral RNA and the 27K proteinase domain is found neither in nuclear inclusions nor in total

protein extracts from infected tissues (Dougherty and Parks, 1991; Carrington *et al.*, 1993). While most sites recognized by the NIa proteinases have a Gln at P1 position, a Glu residue is present at this position of all proposed VPg–Pro cleavage sites. The residue at position P3 of the TEV VPg–Pro cleavage site also deviated from the heptapeptide consensus sequence. These changes might be involved in slowing cleavage. Mutations that accommodated the VPg–Pro junction to a consensus cleavage site accelerated internal processing *in vitro* very much. Genome amplification was drastically disturbed by these substitutions, suggesting that the slow-processing feature may accomplish an important regulatory function (Schaad *et al.*, 1996).

The proteolytic processing strategy of gene expression provides the opportunity to use partially processed forms of viral proteinases to play alternative roles (Dessens and Lomonossoff, 1992; Hellen and Wimmer, 1992; Margis *et al.*, 1994). However, when the cleavage profiles of precursor and processed forms of the TEV NIa proteinase were analyzed, most substrates were processed in a similar fashion by all proteolytic forms. Only at the $6K_1$-CI site, slight processing differences could be observed (Parks *et al.*, 1992). Further autoprocessing at specific positions of the NIa C-terminal region has been described in turnip mosaic potyvirus (TuMV) (Kim *et al.*, 1995; Ménard *et al.*, 1995; Kim *et al.*, 1996) and TEV (Parks *et al.*, 1995). The sequences around these cleavage sites were not similar to the typical heptapeptide recognition signals. Whereas a truncated TuMV 25K product (lacking the last 20 aa) was as active as the complete 27K proteinase for the cleavage at the $6K_1$-CI site (Kim *et al.*, 1995), a TuMV 24K protein (lacking the last 30 aa) did not cleave at this site (Kim *et al.*, 1996), and a TEV 25K protein (lacking the last 24 aa) was approximately one-twentieth as efficient in proteolysis of the NIb-CP site as the full-length form (Parks *et al.*, 1995). The functional relevance of sequences at the C-terminal region of the NIa proteinase (that may depend on the substrate analyzed), has also been shown by deletion and site-directed mutagenesis (García *et al.*, 1989b; Kim *et al.*, 1996). Thus, although it has not been demonstrated that trimming at the NIa C-end takes place *in vivo*, the possibility of regulation of activity by removal of C-terminal sequences of the NIa proteinase must be carefully considered.

B. Comoviridae

The family Comoviridae includes the genera *Comovirus*, *Nepovirus*, and *Fabavirus* (Murphy *et al.*, 1995). Since at present data on the molecular biology of fabaviruses are not available, only como- and nepovirus proteinases are discussed. In both cases, their genomes are split in two RNA molecules, which encode large polyproteins that are proteolytically cleaved by viral proteinases.

Whereas comovirus RNA-B and nepovirus RNA-1 code for all the proteins required for RNA replication, including the viral proteinases, comovirus RNA-M and nepovirus RNA-2 encode the capsid proteins (CPs) and proteins involved in virus movement. Comovirus RNA-M differs from nepovirus RNA-2 in being translated from two alternative in frame AUG codons giving rise to two coterminal polyproteins (Fig. 2).

The best studied member of the Comoviridae family is cowpea mosaic comovirus (CPMV). Its RNA-B encodes a 24K proteinase homologous to the 3C-like proteinases. The genomic location of the proteinase gene, as part of a VPg–proteinase–replicase segment, is consistent with that of the rest of picornalike viruses. Sequence alignments (Gorbalenya *et al.*, 1989; Bazan and Fletterick, 1990; Shanks and Lomonossof, 1990) and site-directed mutagenesis analysis (Dessens and Lomonossoff, 1991) have identified the probable catalytic triad, composed of His987, Glu1023, and Cys1113. Sequence analysis also predicts that His1131 is the residue of the substrate-binding pocket that interacts with Gln at cleavage site P1 position.

The CPMV 24K proteinase recognize Q–G, Q–S, and Q–M cleavage sites (Wellink and van Kammen, 1988). Similar dipeptides are present at the putative cleavage sites of other sequenced comoviruses (Chen and Bruening, 1992a,b; Shanks and Lomonossoff, 1992). A weak consensus for Ala at positions P2 and P4 is the only other feature observed at the comovirus cleavage sites.

Comovirus polyprotein processing differs from those of other picornalike viruses in being regulated by a viral product without proteolytic activity. After *in vitro* translation, an N-terminal 32K protein is released from the CPMV 200K RNA-B-encoded polyprotein by an intramolecular cleavage and remains associated with the remaining 170K protein, probably by interaction with its 58K domain (Peters *et al.*, 1992b). The same complex can be formed if 32K and 170K proteins are translated simultaneously from different RNA molecules, but the proteins cannot associate if they are translated separately and mixed later (Peters *et al.*, 1992b). When the 170K polyprotein, which contains the 24K proteinase, is associated with the 32K cofactor, further 170K self-cleavage is very slow (Peters *et al.*, 1992b). Also, trans cleavage between the two RNA-M-encoded CPs mediated by the 32K–170K complex is very inefficient (Vos *et al.*, 1988). In contrast, this complex (or noncleaved 200K protein) accomplishes the processing at the Gln–Met site that separates the 48- and 58-kDa proteins from the 60K CP precursor (Vos *et al.*, 1988; Peters *et al.*, 1992b). In the absence of the 32K factor, the 170K protein was efficiently processed, essentially by cis cleavages, following three different pathways that start with the synthesis of 60K+110K, 80K+87K, and 58K+112K proteins (Peters *et al.*, 1992a). The puzzling final pattern includes not only the fully processed products, but also stable intermediates that are not efficiently processed due to the cis preference

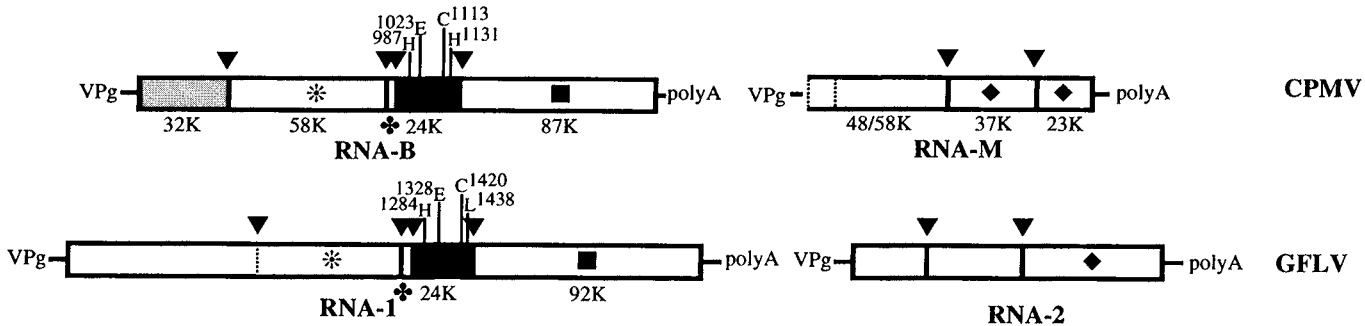

FIGURE 2 Genome maps of a comovirus (CPMV) and a nepovirus (GFLV) depicted as explained in the legend to Fig. 1. A light gray box represents the 32K proteinase cofactor. Vertical dashed lines indicate the alternative translation initiation sites of the CPMV M polyprotein.

of the 24K proteinase (Peters *et al.*, 1992a) and the requirement of upstream sequences for efficient cleavage of the 24K proteinase at its C-end (Dessens and Lomonossoff, 1992). This complex control of the proteolytic processing seems to be essential even for the replication of CPMV RNA-B alone, since a mutated RNA-B that lacks the sequences coding for the 32K protein did not replicate in cowpea protoplasts (Peters *et al.*, 1992b). It has been suggested that the 32K protein might act as a molecular chaperon blocking a certain folding pathway for the 170K protein that could lead to the formation of abortive structures as a result of premature self-cleavage (Peters *et al.*, 1992b).

Sequence alignments of the 24K proteinases from different comoviruses show large sequence conservation, even at the residues predicted to be involved in substrate binding by Bazan and Fletterick (1988). In spite of this fact, the 24K proteinase of a comovirus is not able to process cleavage sites of another comovirus neither in trans (at the 58/48K–60K junction) (Goldbach and Krijt, 1982) nor in cis (at the 32K–170K junction) (Shanks *et al.*, 1996). The fact that changing the Gln–Gly dipeptide at the 37K–23K junction of the M polyprotein into either Gln–Ser or Gln–Met resulted in a dramatic decrease of proteolysis efficiency indicates that cleavage specificity of the 24K proteinases is determined in part by the amino acid sequence of the junction site (Vos *et al.*, 1988). However, cleavage by the CPMV 24K proteinase at the CPMV 32K–170K junction is not prevented when the Gln–Ser site is changed into a His–Met site (Peters *et al.*, 1992b), whereas CPMV 24K proteinase is not able to cleave at Gln–Gly site of the red clover mottle comovirus 32K–170K junction (Shanks *et al.*, 1996), suggesting that in this case the specificity is determined by tertiary structure interactions between the substrate and the substrate-binding pocket of the proteinase.

The proteinase involved in the proteolytic processing of the nepovirus polyproteins occupies the same genomic place and has similar size as the comovirus 24K proteinase (Figs. 2 and 4). The catalytic triad of the grapevine fanleaf nepovirus (GFLV) 24K proteinase, predicted on sequence alignments and studied by site-directed mutagenesis, is formed by His1284, Glu1328, and Cys1420 (Margis and Pinck, 1992). The GFLV 24K proteinase differs from the 3C-like proteinases of poty- and comoviruses in two main aspects. First, its active site Cys can be mutated to Ser without loss of activity and, second, Leu1438 substitutes for the typical His of the 3C proteinase substrate-binding pocket (Margis and Pinck, 1992). Similar Leu residues are found in the 24K proteinases of tomato black ring nepovirus (TBRV) and grapevine chrome mosaic nepovirus (GCMV). All these data suggest that the substrate specificity of nepovirus enzymes is more similar to that of cellular serine proteinases than to that of their viral counterparts. In agreement with this assumption, some nepovirus proteinase cleavages take place at Arg–Ala, Arg–Gly, and Arg–Ala dipeptides, although cleavages at Cys–Ala, Cys–Ser, and Gly–Glu dipeptides have also been

described (Brault *et al.*, 1989; Pinck *et al.*, 1991; Margis *et al.*, 1993). The 24K proteinase of tomato ringspot nepovirus (TomRSV), which probably form part of a distinct subgroup, seems to be different from those of GFLV and other related nepoviruses, being more similar to the CPMV 24K proteinase since it has a His in the putative substrate-binding pocket and probably cleaves Gln-x sites (Rott *et al.*, 1995).

Although the proteolytic processing of nepovirus polyproteins should be also tightly regulated, the control mechanisms seem to be quite different from those of comoviruses. A virus-encoded cofactor is not required for *in vitro* trans processing of the RNA-2-derived polyprotein (Margis *et al.*, 1993). However, the identification at the N-terminal region of some nepovirus polyproteins of sequence motifs also present in the comovirus 32K protein (Ritzenhaler *et al.*, 1991; Rott *et al.*, 1995) might suggest that, at least in some cases, this protein could collaborate with the 24K protein in the nepovirus proteolytic processing. Similarly to the comovirus proteinase, the activity of the nepovirus 24K proteinase is modulated by sequences surrounding it. However, while sequences upstream of the comovirus proteinase enhance *in vitro* cleavage at its C-terminus (Dessens and Lomonossoff, 1992), the GFLV 24K–92K precursor is better cleaved than the VPg–24K–92K intermediates (Margis *et al.*, 1994).

C. Sequiviridae

The family Sequiviridae consists of the genera *Sequivirus* and *Waikavirus* (Murphy *et al.*, 1995). The monopartite genome of sequi- and waikaviruses differs from that of the potyviruses in encoding three capsid proteins (like animal picornaviruses), located internally near the N-terminus of the large polyprotein (Fig. 3). Particular features of the waikaviruses are long AUG-containing sequences upstream of the large genomic ORF and small 3′ ORFs that might be expressed by subgenomic RNAs (Shen *et al.*, 1993; Reddick *et al.*, 1997). The genomic RNA of parsnip yellow fleck sequivirus (PYFV) lacks a poly-A tail (Turnbull-Ross *et al.*, 1992).

Although until now experimental data on the activity of the sequiviridae 3C-like proteinases have not been obtained, they can be clearly identified by sequence alignments (Fig. 4) (Shen *et al.*, 1993; Turnbull-Ross *et al.*, 1993; Reddick *et al.*, 1997). The predicted catalytic triads of PYFV sequivirus and rice tungro spherical waikavirus (RTSV) are formed by His, Glu, and Cys (Figs. 3 and 4). However, the proposed acidic active site residue of maize chlorotic dwarf waikavirus (MCDV) is an Asp residue (Reddick *et al.*, 1997). Differences are also observed at the putative substrate-binding pockets. The typical His of proteinases that cleave after a Gln residue can be identified in the RTSV and MCDV waikavirus sequences, but it is replaced by Leu (like in most nepoviruses) in

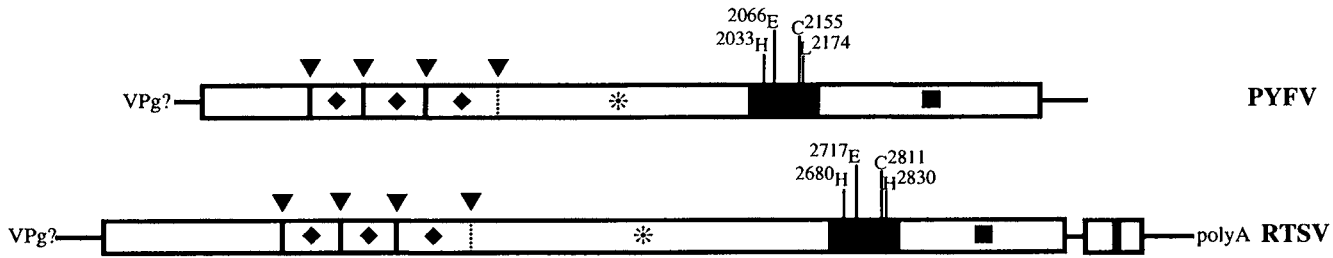

FIGURE 3 Genome maps of a sequivirus (PYFV) and a waikavirus (RTSV) depicted as explained in the legend to Fig. 1.

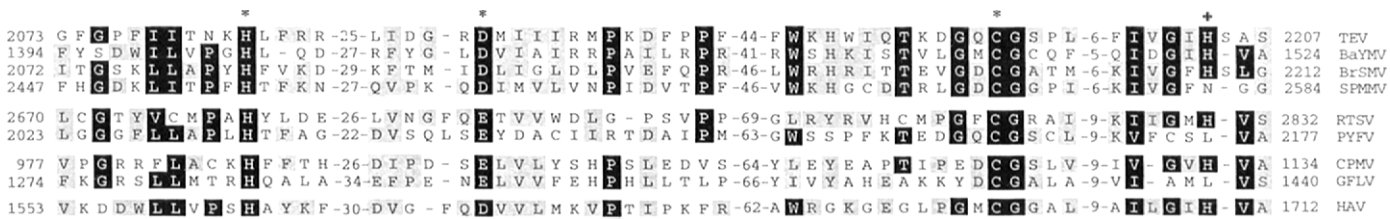

FIGURE 4 Sequence alignments of regions around the catalytic residues (signaled by *) of plant virus 3C-like proteinases. The 3C protein from hepatitis A virus (HAV) was also included in the alignment. The position of the residues proposed to interact with the substrate P1 amino acid is marked (✚). Black or gray backgrounds indicate highly or moderately conserved residues, respectively.

PYFV. In agreement with these data, N-terminal amino acid sequencing has revealed that the RTSV (Shen *et al.*, 1993) and MCDV (Reddick *et al.*, 1997) CPs probably arise from proteolytic cleavages at Gln-x dipeptide, whereas no consensus sequence can be found for the cleavage sites that originate the PYFV CPs (Ser–Pro, Asn–Ala, and Gln–Ala) (Turnbull-Ross *et al.*, 1993). However, it cannot be ruled out that the observed CP N-termini might result from secondary proteolytic degradation rather than from the primary cleavage event carried out by the virus proteinase.

III. ALPHALIKE SUPERGROUP

A. CLOSTEROVIRUS

The closterovirus genomes, which can be mono- or bipartites, are the largest among all (+)ssRNA plant viruses (Murphy *et al.*, 1995). The organization and expression of the closterovirus genome resembles that of coronaviruses, with polyprotein processing, translational frameshifting, and multiple sgRNA formation. However, the closterovirus RNA replicase belongs to the alphalike lineage and the mechanism of sgRNA transcription of citrus tristeza closterovirus (CTV) is similar to that of other alphalike viruses, differing clearly from that of coronaviruses (Karasev *et al.*, 1997). The closterovirus large replication proteins are encoded by ORF1a and ORF1b, the second one being probably translated by ribosomal frameshifting as a fusion product with the upstream protein (Fig. 5).

The 5′ end of beet yellows closterovirus (BYV) ORF1a encodes a papainlike cysteine proteinase domain which has been proven to be active in an *in vitro* assay (Agranovsky *et al.*, 1994). A similar domain has been identified in other closteroviruses (Agranovsky, 1996) and it is duplicated in CTV (Karasev *et al.*, 1995). The size of these closterovirus leader proteinases is around 60K. In spite of the similarity of their proteinase domains, the regions upstream of these domains are badly conserved among different closterovirus leader proteins. Leader papainlike proteinases have been described for very different (+)ssRNA viruses, such as coronaviruses, arteriviruses, aphtoviruses, and hypoviruses (Gorbalenya *et al.*, 1991; Rozanov *et al.*, 1995). Moreover, duplication of leader

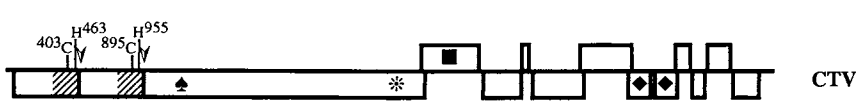

FIGURE 5 Genome map of citrus tristeza closterovirus depicted as explained in the legend to Fig. 1. (♠) Indicates a methyl transferase-like domain.

papainlike proteinases also seems to be a common phenomenon, whose selective advantage is still unknown (Lee *et al.*, 1991; Shapira and Nuss, 1991; Godeny *et al.*, 1993). The proteinase domain of the closterovirus leader protein is more similar to those of HC proteins of different genera of the family Potyviridae than to those of other papainlike proteinases (Section II,A,2; Fig. 7) (Agranovsky *et al.*, 1994).

Sequence alignments and site-directed mutagenesis experiments indicate that Cys509 and His569 constitute the catalytic dyad of the BYV leader proteinase (Agranovsky *et al.*, 1994). The closterovirus cysteine proteinases resemble other viral papainlike proteinases in cleaving at $uG\downarrow x$ sites, where u is a bulky hydrophobic residue and x is usually a Gly. Some conservation of a negatively charged residue at P4′ position has also been observed (Jelkmann *et al.*, 1997).

Little is known about the functional role of closterovirus leader proteins. A putative function in aphid transmission has been suggested on the basis of analogy to potyvirus HC protein. However, the multifunctional nature of HC protein and the conservation of the cysteine proteinase domain in potyviridae (see Section II,A,2) and closteroviruses (Jelkmann *et al.*, 1997) that are not transmitted by aphids poses concerns on this hypothesis. Recently, it has been described that the leader proteinase of BYV suppresses potyvirus infection in the BYV nonhost *Nicotiana tabacum*, but does not affect potyvirus replication in *N. tabacum* protoplasts or systemic infection in the BYV host *Nicotiana benthamiana* (Dolja *et al.*, 1997). The authors suggest that the potyviral HC and the BYV leader proteinases have analogous structural and functional organization and that the two proteins may compete for interaction with the same cellular target. The complex formed by BYV protein would be functional in the BYV host plant but would interfere with normal potyvirus infection in the BYV nonhost plant.

B. Tymovirus-like

The plant alphalike viruses are characterized by replication proteins that contain an ordered series of domains: methyltransferaselike (MTr), helicaselike (Hel), and polymeraselike (pol) (Goldbach *et al.*, 1991). In some cases two proteins, one containing the MTr and Hel domains and the other containing the Pol domain, are encoded by separated RNA molecules. Other viruses encode a single polyprotein that contains the three domains. However, most of the alphalike viruses have developed special strategies to produce different amounts of a protein containing the three domains and a Pol-containing protein. Suppression of termination at leaky stop codons (readthrough), ribosomal frameshifting, and proteolytic processing can play that role.

The proteinases involved in the proteolytic processing of the plant alphalike replication proteins have been denoted tymolike because they share some

particular features with the best studied member of the group, the tymovirus proteinase (for a recent review, see Rozanov *et al.*, 1995). The proteinase responsible for the intramolecular cleavage of the 206K replication protein of turnip yellow mosaic virus (TYMV) has been mapped just upstream of the Hel domain (Fig. 6), and deletion analysis has delimited the proteinase domain to residues 731–885 (Bransom and Dreher, 1994). Sequence alignments and site-directed mutagenesis analysis indicate that the TYMV proteinase is a papainlike cysteine proteinase with a predicted catalytic dyad formed by Cys783 and His869 (Bransom and Dreher, 1994; Rozanov *et al.*, 1995). In contrast with the closterovirus leader and the potyvirus HC proteinases, the TYMV proteinase does not cleave at the end of the proteinase domain but further downstream between the Hel and the Pol domains. N-Terminal sequencing of the C-terminal

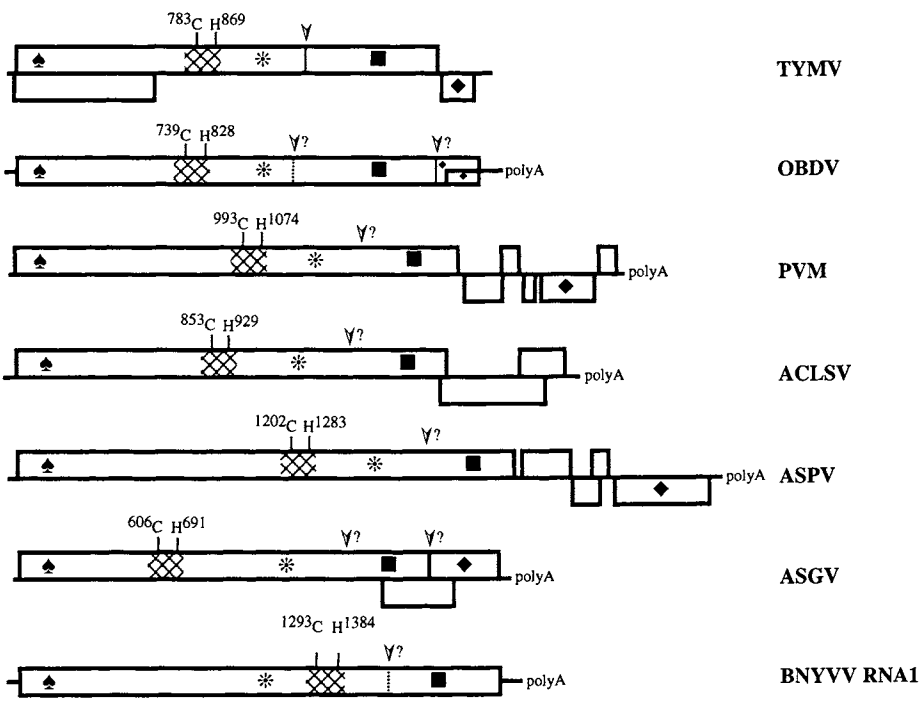

FIGURE 6 Genome maps of members of virus groups that have tymolike proteinases depicted as explained in the legends to Figs. 1 and 5. Viruses not mentioned in the text are: potato carlavirus M (PVM), apple chlorotic leaf spot trichovirus (ACLSV), and apple stem grooving capillovirus (ASGV). Tymolike proteinase sequences are represented by crossed patterns. Dashed vertical lines indicate putative cleavage sites predicted by sequence analysis. A box inside another one indicates sequences that are thought to be expressed both as a fusion with preceding in-frame ones and by a subgenomic RNA.

cleavage product derived from autoprocessing of the TYMV 206K polyprotein synthesized *in vitro* (Bransom *et al.*, 1996) or in *E. coli* (Kadaré *et al.*, 1995) has shown that cleavage occurs between Ala1259 and Thr1260. The sequences at the cleavage sites predicted for different tymoviruses are poorly conserved, although a small amino acid is always present at position P1 (resembling other viral papainlike cysteine proteinases, see Sections II,A,2 and III,A) and P2 (Kadaré *et al.*, 1995).

Predicted cysteine proteinase domains similar to the TYMV one have been identified in the replication proteins of carlaviruses, capilloviruses, tricho-viruses, marafiviruses, and apple stem pitting virus (ASPV) (Rozanov *et al.*, 1995; Edwards *et al.*, 1997) (Figs. 6 and 7). Although the genome structures of these viruses are very different, their large replication proteins always show the same modular organization: MTr–proteinase–Hel–Pol. Interestingly, the capsid protein of capilloviruses (Ohira *et al.*, 1995) and one of the capsid proteins of oat blue dwarf marafivirus (OBDV) (Edwards *et al.*, 1997) are translated as the C-terminal part of their large polyproteins, but it is unknown if they are proteolytically processed by their tymolike cysteine proteinases. Experimental evidence for proteolytic activity of tymolike proteinases other than the TYMV one has only been reported for that of blueberry scorch carlavirus (BBScV) (Lawrence *et al.*, 1995).

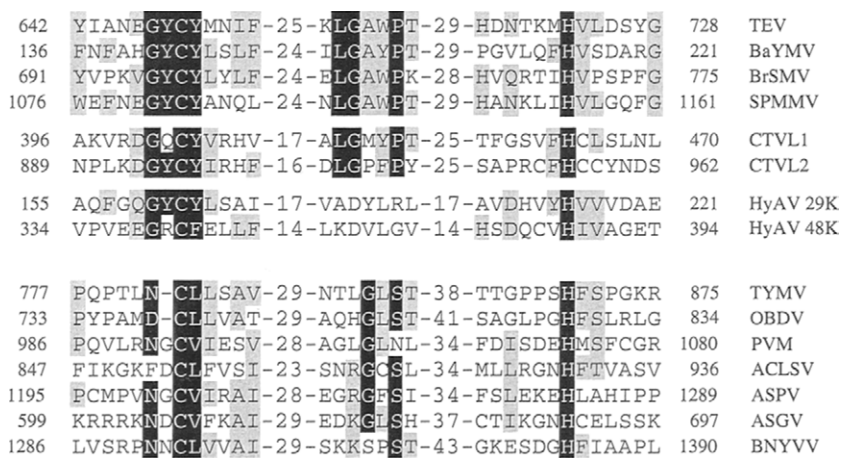

FIGURE 7 Sequence alignments of regions around the catalytic residues (signaled by *) of plant virus papainlike proteinases. The HyAV papainlike proteinases were also included in the alignment. Black or gray backgrounds indicate highly or moderately conserved residues, respectively, either in the HC-like or in the tymolike groups. Virus names are explained in the text and in the legend to Fig. 6.

The protein encoded by beet necrotic yellow vein benyvirus (BNYVV) RNA-1, which contains the information necessary for replication of the viral genome, has been shown to undergo autocatalytic processing (Hehn *et al.*, 1997). A domain with sequence similarity to the papainlike TYMV proteinase has been found in the BNYVV RNA-1-encoded product (Figs. 6 and 7). In contrast with other tymolike proteinase domains, that of BNYVV is located between the Hel and Pol domains of the polyprotein; that is, closed to the cleavage site upstream of the Pol domain (Rozanov *et al.*, 1995; Hehn *et al.*, 1997).

IV. SOBEMOLIKE SUPERGROUP

The sobemolike is a small supergroup formed by the genus *Sobemovirus* and the subgroup II of the genus *Luteovirus*, which have been proposed to have emerged by a recombination event between a sobemovirus and a subgroup I luteovirus (Goldbach *et al.*, 1991). The genome organization of these rather small viruses is quite complex, and its genome expression employs sgRNAs, frameshifting, readthrough, and, probably, proteolytic processing.

Gorbalenya *et al.* (1988) proposed some years ago that a serine proteinase is encoded by the sobemovirus genome. Although the proteinase domain has been tentatively identified in newly sequenced sobemoviruses and luteoviruses, direct experimental evidence of proteolytic activity associated with a gene product from these viruses is not yet available.

Recently, indirect evidence for a proteinase activity associated to potato leafroll luteovirus (PLRV) has been reported. The experimentally determined N-terminal amino acid sequence of the PLRV VPg has been shown to map to the carboxyl region of the PLRV ORF1 product downstream of the putative proteinase domain (van der Wilk *et al.*, 1997) (Fig. 8). Since the RNA replicase is thought to form the carboxyl part of the readthrough product of ORF1 and ORF2, the position of the proteinase on the luteoviral polyprotein would differ from the picornalike VPg–proteinase–Pol arrangement, which prevails in all other ssRNA viruses with a VPg. According to sequence alignment analysis, the proteolytic processing site at the N-terminus of the subgroup II luteovirus VPg is predicted to be E↓S/T (van der Wilk *et al.*, 1997).

V. PLANT PARARETROVIRUSES

Aspartyl proteinases from animal retroviruses have been studied in great detail in the last years (Dougherty and Semler, 1993). In contrast, few experimental data are available on plant pararetrovirus proteinases.

Plant pararetroviruses are now classified as a novel family Caulimoviridae

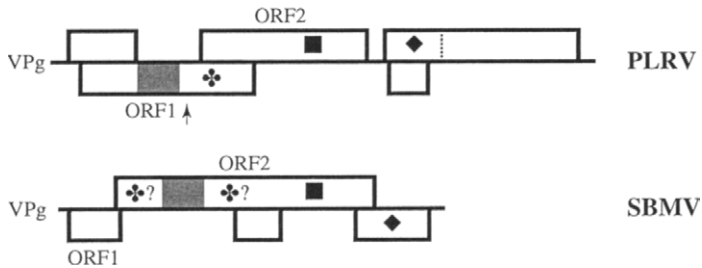

FIGURE 8 Genome maps of a luteovirus (PLRV) and a sobemovirus (SBMV) depicted as explained in Figs. 1 and 5.

(Pringle, 1998). Two genera, *Caulimovirus* and *Badnavirus*, were first defined using the capsid morphology as main classification criterion. However, the larger number of different plant pararetrovirus genome organizations demand a more complex classification (Hohn and Fütterer, 1997; Pringle, 1998). All plant pararetroviruses encapsidate dsDNA, contain genes homologous to *gag* and *pol* genes of animal retroviruses, and seem to share the same replication strategy involving reverse transcription. Conversely, different plant pararetroviruses use quite diverse gene expression mechanisms (Hohn and Fütterer, 1997).

Cauliflower mosaic caulimovirus (CaMV) is the best-studied plant pararetrovirus (Fig. 9). The CaMV genome encodes CP, proteinase, reverse transcriptase, and RNase H in the same order than animal retroviruses; however, CaMV differs from them in producing independent polyproteins for the CP (ORF4) and the enzymatic functions (ORF5) (Schultze *et al.*, 1990). *In vitro* translation of ORF5 has shown that its primary translation product is processed to yield an N-terminal protein containing the proteinase domain and a C-terminal one containing the reverse transcriptase and RNase H domains (Torruella *et al.*, 1989). The CaMV proteinase has the characteristic DTG active site (involvement of the Asp residue in the proteolytic activity has been shown by site-directed mutagenesis) and the conserved Gly in the typical IIGD context of aspartyl proteinases. Its 20K size and the fact that it contains only one copy of the proteinase motifs suggest that, like its animal retrovirus counterparts, the CaMV proteinase is active as a dimer. Experiments in plant protoplasts and in *E. coli* have demonstrated that the ORF5-encoded proteinase is also involved in the processing of the ORF4 product. The resulting 44K protein undergoes further, not well characterized, posttranslational modifications and forms the viral capsids (Martínez-Izquierdo and Hohn, 1987).

There are also experimental data on the aspartyl proteinase of rice tungro bacilliform virus (RTBV). In this virus, the CP, proteinase, reverse transcriptase, and RNase H are synthesized as part of a single polyprotein that includes addi-

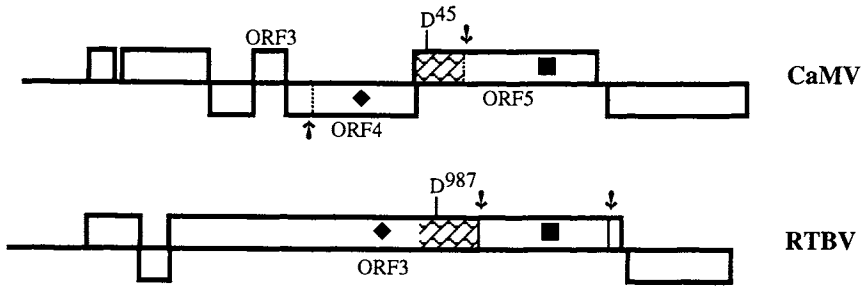

FIGURE 9 Linear representation of the circular genomes of CaMV caulimovirus and RTBV depicted as explained in the legends to Figs. 1 and 5. The aspartyl proteinase is represented by a bricked pattern. Only cleavage sites for which there is experimental evidence are indicated. Dashed vertical lines indicate that the exact place of the cleavage site has not been identified. The cleavage site of the noncharacterized cysteine proteinase in the ORF3 product is not shown.

tional sequences of unknown function (Qu *et al.*, 1991). Making use of a baculovirus expression system, Laco *et al.* (1994, 1995) have demonstrated that the RTBV proteinase is able to cleave upstream of the reverse transcriptase domain and downstream of the RNaseH domain. Cleavage at this second site is not required for reverse transcriptase activity but it is needed for RNase H activity. The sequences at the two cleavage sites, GY↓KN and LK↓CL, are not similar to those described for animal retroviruses. Immunoelectron microscopy experiments have shown the presence of the RTBV proteinase in the surface of virus particles; however, it is not known if it is present as a free protein or as part of a larger precursor (Hay *et al.*, 1994).

It has been suggested that a virion-associated cysteine proteinase is involved in the processing of the CaMV ORF3, a minor component of the virus particles (Guidasci *et al.*, 1992; Dautel *et al.*, 1994). More information is required to know the relevance of this second proteinase, probably of cellular origin, in the infection cycle of CaMV.

VI. CONCLUDING REMARKS AND PERSPECTIVES

In the last years, advances in plant virus genome sequencing and the availability of *in vitro* and *in vivo* heterologous experimental systems have permitted the identification and characterization of a large number of plant virus-encoded proteinases. The development of full-length cDNA clones from which infectious transcripts can be produced either *in vitro* or *in vivo*, has facilitated the functional analysis of the plant virus proteinases. However, at present nearly

nothing is known about how the different proteolytic processing pathways are controlled (by viral and host factors) to engender the required protein products in the appropriate place, amount, and time.

In spite of the high specificity of the viral proteinases, cellular substrates for animal virus proteinases have been described (for instance, Devaney *et al.*, 1988; Clark *et al.*, 1993; Novoa *et al.*, 1997). At least some of these cellular substrates are proteins involved in the control of cell transcription and translation. Thus, the activity of the viral proteinases can interfere with important cellular processes to favor virus replication. Although many plant virus encoded proteinases only act in cis, and cleavage of plant cell proteins by viral proteinases has not been described, it is tempting to speculate that virus-induced proteolytic activities could affect the basic plant cell machinery and/or its defensive responses. These events could be specially relevant to explain the ability of the virus to infect particular hosts and the development of disease symptoms.

Finally, the high specificity of the plant virus-encoded proteinases confers upon them very high interest as potential biotechnological tools and targets. As an example of their use in biotechnology, the potyvirus NIa proteinase has been reported to be helpful for the purification of tag-linked proteins synthesized in heterologous systems (Parks *et al.*, 1994) and for the production in transgenic plants of multiple proteins through tranlation of single self-processing polypeptide (Marcos and Beachy, 1997). On the other hand, the recent use of proteinase inhibitors in AIDS therapy has emphasized the convenience of virus-encoded proteinases as targets of antiviral action. Van Rompaey *et al.* (1995) have designed a mutant protein able to inhibit the activity of the TEV proteinase by manipulation of the α_2-macroglobulin bait region. The expression of appropriately designed proteinase inhibitors might provide to transgenic plants suitable virus resistance.

REFERENCES

Agranovsky, A. A. (1996). Principles of molecular organization, expression and evolution of closteroviruses: over the barriers. *Adv.Virus Res.* **47**, 119–158.

Agranovsky, A. A., Koonin, E. V., Boyko, V. P., Maiss, E., Frötschl, R., Lunina, N. A., and Atabekov, J. G. (1994). Beet yellows closterovirus: complete genome structure and identification of a leader papain-like thiol protease. *Virology* **198**, 311–324.

Atreya, C. D., Atreya, P. L., Thornbury, D. W., and Pirone, T. P. (1992). Site-directed mutations in the potyvirus HC-Pro gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. *Virology* **191**, 106–111.

Barrett, A. J. (1986). An introduction to the proteinases. *In* "Proteinase Inhibitors" (A. J. Barret and G. Salvessen, Eds.), pp. 3–22. Elsevier, New York.

Bazan, J. F., and Fletterick, R. J. (1988). Viral cysteine proteases are homologous to the trypsin-like

family of serine proteases: Structural and functional implications. *Proc. Natl. Acad. Sci. USA* **85**, 7872–7876.

Bazan, J. F., and Fletterick, R. J. (1990). Structural and catalytic models of trypsin-like viral proteases. *Semin. Virol.* **1**, 311–322.

Berger, P. H., Hunt, A. G., Domier, G. M., Hellman, G. M., Stram, Y., Thornbury, D. W., and Pirone, T. P. (1989). Expression in transgenic plants of a viral gene product that mediates insect transmission of potyviruses. *Proc. Natl. Acad. Sci. USA* **86**, 8402–8406.

Bergmann, E. M., Mosimann, S. C., Chernaia, M. M., Malcolm, B. A., and James, M. N. G. (1997). The refined crystal structure of the 3C gene product from hepatitis A virus: Specific proteinase activity and RNA recognition. *J. Virol.* **71**, 2436–2448.

Bransom, K. L., and Dreher, T. W. (1994). Identification of the essential cysteine and histidine residues of the turnip yellow mosaic virus protease. *Virology* **198**, 148–154.

Bransom, K. L., Wallace, E., and Dreher, T. W. (1996). Identification of the cleavage site recognized by the turnip yellow mosaic virus protease. *Virology* **217**, 404–406.

Brantley, J. D., and Hunt, A. G. (1993). The N-terminal protein of the polyprotein encoded by the potyvirus tobacco vein mottling virus is an RNA-binding protein. *J. Gen. Virol.* **74**, 1157–1162.

Brault, V., Hibrand, L., Candresse, T., Le Gall, O., and Dunez, J. (1989). Nucleotide sequence and genetic organization of Hungarian grapevine chrome mosaic nepovirus RNA2. *Nucleic Acids Res.* **17**, 7809–7819.

Carrington, J. C., and Dougherty, W. G. (1987). Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. *J. Virol.* **61**, 2540–2548.

Carrington, J. C., and Dougherty, W. G. (1988). A viral cleavage site cassette: Identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci. USA* **85**, 3391–3395.

Carrington, J. C., and Herndon, K. L. (1992). Characterization of the potyviral HC-Pro autoproteolytic cleavage site. *Virology* **187**, 308–315.

Carrington, J. C., Cary, S. M., and Dougherty, W. G. (1988). Mutational analysis of tobacco etch virus polyprotein processing: *Cis* and *trans* proteolytic activities of polyproteins containing the 49-kilodalton proteinase. *J. Virol.* **62**, 2313–2320.

Carrington, J. C., Cary, S. M., Parks, T. D., and Dougherty, W. G. (1989a). A second proteinase encoded by a plant potyvirus genome. *EMBO J.* **8**, 365–370.

Carrington, J. C., Freed, D. D., and Sanders, T. C. (1989b). Autocatalytic processing of the potyvirus helper component proteinase in *Escherichia coli* and *in vitro*. *J. Virol.* **63**, 4459–4463.

Carrington, J. C., Freed, D. D., and Oh, C.-S. (1990). Expression of potyviral polyproteins in transgenic plants reveals three proteolytic activities required for complete processing. *EMBO J.* **9**, 1347–1353.

Carrington, J. C., Haldeman, R., Dolja, V. V., and Restrepo-Hartwig, M. A. (1993). Internal cleavage and *trans*-proteolytic activities of the VPg-proteinase (NIa) of tobacco etch potyvirus *in vivo*. *J. Virol.* **67**, 6995–7000.

Chen, X., and Bruening, G. (1992a). Cloned DNA copies of cowpea severe mosaic virus genomic RNAs: Infectious transcripts and complete nucleotide sequence of RNA 1. *Virology* **191**, 607–618.

Chen, X., and Bruening, G. (1992b). Nucleotide sequence and genetic map of cowpea severe mosaic virus RNA 2 and comparisons with RNA 2 of other comoviruses. *Virology* **187**, 682–692.

Choi, G. H., Pawlyk, D. M., and Nuss, D. L. (1991). The autocatalytic protease p29 encoded by a hypovirulence-associated virus of the chesnut blight fungus resembles the potyvirus-encoded protease HC-Pro. *Virology* **183**, 747–752.

Clark, M. E., Lieberman, P. M., Berk, A. J., and Dasgupta, A. (1993). Direct cleavage of human TATA-binding protein by poliovirus protease 3C *in vivo* and *in vitro*. *Mol. Cell. Biol.* **13**, 1232–1237.

Cronin, S., Verchot, J., Haldeman-Cahill, R., Schaad, M. C., and Carrington, J. C. (1995). Long-distance movement factor: A transport function of the potyvirus helper component proteinase. *Plant Cell* **7**, 549–559.

Daros, J. A., and Carrington, J. C. (1997). RNA binding activity of Nla proteinase of tobacco etch potyvirus. *Virology* **237**, 327–336.

Dautel, S., Guidasci, T., Pique, M., Mougeot, J.-L., Lebeurier, G., Yot, P., and Mesnard, J. B. (1994). The full-length product of cauliflower mosaic virus open reading frame III is associated with the viral particle. *Virology* **202**, 1043–1045.

De Mejia, M. V. G., Hiebert, E., Purcifull, D. E., Thornbury, D. W., and Pirone, T. P. (1985). Identification of potyviral amorphous inclusion protein as a nonstructural virus-specific protein related to helper component. *Virology* **142**, 34–43.

Dessens, J. T., and Lomonossoff, G. P. (1991). Mutational analysis of the putative catalytic triad of the cowpea mosaic virus 24K protease. *Virology* **184**, 738–746.

Dessens, J. T., and Lomonossoff, G. P. (1992). Sequence upstream of the 24K protease enhances cleavage of the cowpea mosaic virus B-RNA-encoded polyprotein at the junction between the 24K and 87K proteins. *Virology* **189**, 225–232.

Devaney, M. A., Vakharia, V. N., Lloyd, R. E., Ehrenfeld, E., and Grubman, M. J. (1988). Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. *J. Virol.* **62**, 4407–4409.

Dolja, V. V., Hong, J., Keller, K. E., Martin, R. R., and Peremyslov, V. V. (1997). Suppression of potyvirus infection by coexpressed closterovirus protein. *Virology* **234**, 243–252.

Dougherty, W. G., and Parks, T. D. (1989). Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage sequence in determining the reaction profile at two tobacco etch virus cleavage sites in cell-free assays. *Virology* **172**, 145–155.

Dougherty, W. G., and Parks, T. D. (1991). Post-translational processing of the tobacco etch virus 49-kDa small nuclear inclusion polyprotein: Identification of an internal cleavage site and delimitation of VPg and proteinase domains. *Virology* **183**, 449–456.

Dougherty, W. G., and Semler, B. L. (1993). Expression of virus-encoded proteinases: Functional and structural similarities with cellular enzymes. *Microbiol. Rev.* **57**, 781–822.

Dougherty, W. G., Carrington, J. C., Cary, S. M., and Parks, T. D. (1988). Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J.* **7**, 1281–1287.

Dougherty, W. G., Cary, S. M., and Parks, T. D. (1989a). Molecular genetic analysis of a plant virus polyprotein cleavage site: A model. *Virology* **171**, 356–364.

Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F., and Fletterick, R. J. (1989b). Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. *Virology* **172**, 302–310.

Edwards, M. C., Zhang, Z., and Weiland, J. J. (1997). Oat blue dwarf marafivirus resembles the tymoviruses in sequence, genome organization, and expression strategy. *Virology* **232**, 217–229.

García, J. A., and Laín, S. (1991). Proteolytic activity of plum pox virus-tobacco etch virus chimeric proteins. *FEBS Lett.* **281**, 67–72.

García, J. A., Laín, S., Cervera, M. T., Riechmann, J. L., and Martín, M. T. (1990). Mutational analysis of plum pox potyvirus polyprotein processing by the NIa protease in *Escherichia coli*. *J. Gen. Virol.* **71**, 2773–2779.

García, J. A., Martín, M. T., Cervera, M. T., and Riechmann, J. L. (1992). Proteolytic processing of the plum pox potyvirus polyprotein by the NIa protease at a novel cleavage site. *Virology* **188**, 697–703.

García, J. A., Riechmann, J. L., and Laín, S. (1989a). Artificial cleavage site recognized by plum pox potyvirus protease in *Escherichia coli*. *J. Virol.* **63**, 2457–2460.

García, J. A., Riechmann, J. L., and Laín, S. (1989b). Proteolytic activity of the plum pox potyvirus NIa-like protein in *Escherichia coli*. *Virology* **170**, 362–369.

Ghabrial, S. A., Smith, H. A., Parks, T. D., and Dougherty, W. G. (1990). Molecular genetic analyses of the soybean mosaic virus Nla protease. *J. Gen. Virol.* **71**, 1921–1927.

Godeny, E. K., Chen, L., Kumar, S. N., Methven, S. L., Koonin, E. V., and Brinton, M. A. (1993). Complete sequence and phylogenetic analysis of the lactate dehydrogenase-elevating virus. *Virology* **194**, 585–596.

Goldbach, R., and Krijt, J. (1982). Cowpea mosaic virus-encoded protease does not recognise primary translation products of mRNAs from other comoviruses. *J. Virol.* **43**, 1151–1154.

Goldbach, R., Le Gall, O., and Wellink, J. (1991). Alpha-like viruses in plants. *Semin. Virol.* **2**, 19–25.

Gorbalenya, A. E., Donchenko, A. P., Blinov, V. M., and Koonin, E. V. (1989). Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases: A distinct protein superfamily with a common structural fold. *FEBS Lett.* **243**, 103–114.

Gorbalenya, A. E., Koonin, E., Blinov, V. M., and Donchenko, A. P. (1988). Sobemovirus genome appears to encode a serine protease related to cysteine proteases of picornaviruses. *FEBS Lett.* **236**, 287–290.

Gorbalenya, A. E., Koonin, E., and Lai, M. M.-C. (1991). Putative papain-related thiol proteases of positive-strand RNA viruses: Identification of rubi- and aphtovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, α-, and coronaviruses. *FEBS Lett.* **288**, 201–205.

Guidasci, T., Mougeot, J. L., Lebeurier, G., and Mesnard, J. M. (1992). Processing of the minor capsid protein of the cauliflower mosaic virus requires a cysteine proteinase. *Res.Virol.* **143**, 361–370.

Hay, J., Grieco, F., Druka, A., Pinner, M., Lee, S.-C., and Hull, R. (1994). Detection of rice tungro bacilliform virus gene products *in vivo*. *Virology* **205**, 430–437.

Hehn, A., Fritsch, C., Richards, K. E., Guilley, H., and Jonard, G. (1997). Evidence for *in vitro* and *in vivo* autocatalytic processing of the primary translation product of beet necrotic yellow vein virus RNA 1 by a papain-like proteinase. *Arch. Virol.* **142**, 1051–1058.

Hellen, C. U. T., and Wimmer, E. (1992). Maturation of poliovirus capsid proteins. *Virology* **187**, 391–397.

Hellmann, G. M., Shaw, J. G., and Rhoads, R. E. (1988). *In vitro* analysis of tobacco vein mottling virus Nla cistron: Evidence for a virus encoded protease. *Virology* **163**, 554–562.

Hiebert, E., Purcifull, D. E., and Christie, R. G. (1984). Purification and immunological analysis of plant viral inclusion bodies. *In* "Methods in Virology" (K. Maramorosch and H. Koprowski, Eds.), Vol. 8, pp. 225–279. Academic Press, New York.

Hohn, T., and Fütterer, J. (1997). The proteins and functions of plant pararetroviruses: Knowns and unknowns. *Crit. Rev. Plant Sci.* **16**, 133–161.

Jelkmann, W., Fechtner, B., and Agranovsky, A. A. (1997). Complete genome structure and phylogenetic analysis of little cherry virus, a mealybug-transmissible closterovirus. *J. Gen. Virol.* **78**, 2067–2071.

Kadaré, G., Rozanov, M., and Haenni, A.-L. (1995). Expression of the turnip yellow mosaic virus proteinase in *Escherichia coli* and determination of the cleavage site within the 206 kDa protein. *J. Gen. Virol.* **76**, 2853–2857.

Karasev, A. V., Boyko, V. P., Gowda, S., Nikolaeva, O. V., Hilf, M. E., Koonin, E. V., Niblett, C. L., Cline, K., Gumpf, D. J., Lee, R. F., Garnsey, S. M., Lewandowski, D. J., and Dawson, W. O. (1995). Complete sequence of the citrus tristeza virus RNA genome. *Virology* **208**, 511–520.

Karasev, A. V., Hilf, M. E., Garnsey, S. M., and Dawson, W. O. (1997). Transcriptional strategy of closteroviruses: Mapping the 5′ termini of the citrus tristeza virus subgenomic RNAs. *J. Virol.* **71**, 6233–6236.

Kasschau, K. D., and Carrington, J. C. (1995). Requirement for HC-Pro processing during genome amplification of tobacco etch potyvirus. *Virology* **209**, 268–273.

Kasschau, K. D., Cronin, S., and Carrington, J. C. (1997). Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology* **228**, 251–262.

Kim, D. H., Hwang, D. C., Kang, B. H., Lew, J., Han, J. S., Song, B. O. D., and Choi, K. Y. (1996). Effects of internal cleavages and mutations in the C-terminal region of NIa protease of turnip mosaic potyvirus on the catalytic activity. *Virology* **226**, 183–190.

Kim, D.-H., Park, Y. S., Kim, S. S., Lew, J., Nam, H. G., and Choi, K. Y. (1995). Expression, purification, and identification of a novel self-cleavage site of the NIa C-terminal 27-kDa protease of turnip mosaic potyvirus C5. *Virology* **213**, 517–525.

Laco, G. S., Kent, S. B. H., and Beachy, R. N. (1994). Rice tungro bacilliform encodes reverse transcriptase, DNA polymerase, and ribonuclease H activities. *Proc. Natl. Acad. Sci. USA* **91**, 2654–2658.

Laco, G. S., Kent, S. B. H., and Beachy, R. N. (1995). Analysis of the proteolytic processing and activation of the rice tungro bacilliform virus gene products reverse transcriptase. *Virology* **208**, 207–214.

Laliberté, J.-F., Nicolas, O., Chatel, H., Lazure, C., and Morosoli, R. (1992). Release of a 22-kDa protein derived from the amino-terminal domain of the 49-kDa NIa of turnip mosaic potyvirus in *Escherichia coli*. *Virology* **190**, 510–514.

Lawrence, D. M., Rozanov, M. N., and Hillman, B. I. (1995). Autocatalytic processing of the 223-kDa protein of blueberry scorch carlavirus by a papain-like proteinase. *Virology* **207**, 127–135.

Lee, C.-J., Shieh, C.-K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerases. *Virology* **180**, 567–582.

Maia, I. G., and Bernardi, F. (1996). Nucleic acid-binding properties of a bacterially expressed potato virus Y helper component-proteinase. *J. Gen. Virol.* **77**, 869–877.

Maia, I. G., Haenni, A.-L., and Bernardi, F. (1996a). Potyviral HC-Pro: A multifunctional protein. *J. Gen. Virol.* **77**, 1335–1341.

Maia, I. G., Seron, K., Haenni, A. L., and Bernardi, F. (1996b). Gene expression from viral RNA genomes. *Plant Mol. Biol.* **32**, 367–391.

Marcos, J. F., and Beachy, R. N. (1997). Transgenic accumulation of two plant virus coat proteins on a single self-processing polypeptide. *J. Gen. Virol.* **78**, 1771–1778.

Margis, R., and Pinck, L. (1992). Effects of site-directed mutagenesis on the presumed catalytic triad and substrate-binding pocket of grapevine fanleaf nepovirus 24-kDa proteinase. *Virology* **190**, 884–888.

Margis, R., Ritzenthaler, C., Reinbolt, J., Pinck, M., and Pinck, L. (1993). Genome organization of grapevine fanleaf nepovirus RNA2 deduced from the 122K polyprotein P2 *in vitro* cleavage products. *J. Gen. Virol.* **74**, 1919–1926.

Margis, R., Viry, M., Pinck, M., Bardonnet, N., and Pinck, L. (1994). Differential proteolytic activities of precursor and mature forms of the 24K proteinase of grapevine fanleaf nepovirus. *Virology* **200**, 79–86.

Martín, M. T., López-Otín, C., Laín, S., and García, J. A. (1990). Determination of polyprotein processing sites by amino terminal sequencing of nonstructural proteins encoded by plum pox potyvirus. *Virus Res.* **15**, 97–106.

Martínez-Izquierdo, J., and Hohn, T. (1987). Cauliflower mosaic virus coat protein is phosphorylated *in vitro* by a virion-associated protein-kinase. *Proc. Natl. Acad. Sci. USA* **84**, 1825–1828.

Mavankal, G., and Rhoads, R. (1991). *In vitro* cleavage at or near the N-terminus of the helper component protein in the tobacco vein mottling virus polyprotein. *Virology* **185**, 721–731.

Ménard, R., Chatel, H., Dupras, R., Plouffe, C., and Laliberté, J.-F. (1995). Purification of turnip mosaic potyvirus viral protein genome-linked proteinase expressed in *Escherichia coli* and development of a quantitative assay. *Eur. J. Biochem.* **229**, 107–112.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (1995). Virus taxonomy: Sixth Report of the International Committee on Taxonomy of Viruses. *In* "Archives of Virology: Supplement 10." Springer-Verlag, Wien/New York.

Murphy, J. F., Rhoads, R. E., Hunt, A. G., and Shaw, J. G. (1990). The VPg of tobacco etch virus RNA is the 49-kDa proteinase or the N-terminal 24-kDa part of the proteinase. *Virology* **178**, 285–288.

Novoa, I., Martinez Abarca, F., Fortes, P., Ortin, J., and Carrasco, L. (1997). Cleavage of p220 by purified poliovirus 2A(pro) in cell-free systems: Effects on translation of capped and uncapped mRNAs. *Biochemistry* **36**, 7802–7809.

Oh, C. S., and Carrington, J. C. (1989). Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. *Virology* **173**, 692–699.

Ohira, K., Namba, S., Rozanov, M., Kusumi, T., and Tsuchizaki, T. (1995). Complete sequence of an infectious full-length cDNA clone of citrus tatter leaf capillovirus: Comparative sequence analysis analysis of capillovirus genomes. *J. Gen. Virol.* **76**, 2305–2309.

Parks, T. D., and Dougherty, W. G. (1991). Substrate recognition by the NIa proteinase of two potyviruses involves multiple domains: Characterization using genetically engineered hybrid proteinase molecules. *Virology* **182**, 17–27.

Parks, T. D., Howard, E. D., Wolpert, T. J., Arp, D. J., and Dougherty, W. G. (1995). Expression and purification of a recombinant tobacco etch virus NIa proteinase: Biochemical analysis of the full-length and a naturally occurring truncated proteinase form. *Virology* **210**, 194–201.

Parks, T. D., Leuther, K. K., Howard, E. D., Johnston, S. A., and Dougherty, W. G. (1994). Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal. Biochem.* **216**, 413–417.

Parks, T. D., Smith, H. A., and Dougherty, W. G. (1992). Cleavage profiles of tobacco etch virus (TEV)-derived substrates mediated by precursor and processed forms of the TEV proteinase. *J. Gen. Virol.* **73**, 149–155.

Peters, S. A., Voorhorst, W. G. B., Wellink, J., and van Kammen, A. (1992a). Processing of VPg-containing polyproteins encoded by the B-RNA from cowpea mosaic virus. *Virology* **191**, 90–97.

Peters, S. A., Voorhost, W. G. B., Wery, J., Wellink, J., and van Kammen, A. (1992b). A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. *Virology* **191**, 81–89.

Pinck, M., Reinbolt, J., Loudes, A. M., Le Ret, M., and Pinck, L. (1991). Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. *FEBS Lett.* **284**, 117–119.

Pringle, C. R. (1998). The universal system of virus taxonomy of the International Committee on Virus Taxonomy (ICTV), including new proposals ratified since publication of the sixth ICTV report in 1995. *Arch. Virol.* **143**, 203–210.

Pruss, G., Ge, X., Shi, X. M., Carrington, J. C., and Vance, V. B. (1997). Plant viral synergism: The potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell* **9**, 859–868.

Qu, R., Bhattacharyya-Pakrasi, M., Laco, G. S., De Kochko, A., Subba Rao, B. L., Kaniewska, M. B., Elmer, J. S., Rochester, D. E., Smith, C., and Beachy, R. N. (1991). Characterization of the genome of rice tungro bacilliform virus: Comparison with commelina yellow mottle virus and caulimoviruses. *Virology* **185**, 354–364.

Reddick, B. B., Habera, L. F., and Law, M. D. (1997). Nucleotide sequence and taxonomy of maize chlorotic dwarf virus within the family *Sequiviridae*. *J. Gen. Virol.* **78**, 1165–1174.

Restrepo-Hartwig, M. A., and Carrington, J. C. (1992). Regulation of nuclear transport of a plant potyvirus protein by autoproteolysis. *J. Virol.* **66**, 5662–5666.

Restrepo-Hartwig, M. A., and Carrington, J. C. (1994). The tobacco etch potyvirus 6-kilodalton protein is membrane associated and involved in viral replication. *J. Virol.* **68**, 2388–2397.

Riechmann, J. L., Laín, S., and García, J. A. (1992). Highlights and prospects of potyvirus molecular biology. *J. Gen. Virol.* **73**, 1–16.

Riechmann, J. L., Cervera, M. T., and García, J. A. (1995). Processing of the plum pox virus polyprotein at the P3–$6K_1$ junction is not required for virus viability. *J. Gen. Virol.* **76**, 951–956.

Ritzenhaler, C., Viry, M., Pink, M., Margis, R., Fuchs, M., and Pinck, L. (1991). Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. *J. Gen. Virol.* **72**, 2357–2365.

Rojas, M. R., Zerbini, F. M., Allison, R. F., Gilbertson, R. L., and Lucas, W. J. (1997). Capsid protein and helper component proteinase function as potyvirus cell-to-cell movement proteins. *Virology* **237**, 283–295.

Rott, M. E., Gilchrist, A., Lee, L., and Rochon, D. (1995). Nucleotide sequence of tomato ringspot virus RNA1. *J. Gen. Virol.* **76**, 465–473.

Rozanov, M. N., Drugeon, G., and Haenni, A.-L. (1995). Papain-like proteinase of turnip yellow mosaic virus: A prototype of a new viral proteinase group. *Arch. Virol.* **140**, 273–288.

Ryan, M. D., and Flint, M. (1997). Virus-encoded proteinases of the picornavirus supergroup. *J. Gen. Virol.* **78**, 699–723.

Schaad, M. C., Haldeman-Cahill, R., Cronin, S., and Carrington, J. C. (1996). Analysis of the VPg-proteinase (NIa) encoded by tobacco etch potyvirus: Effects of mutations on subcellular transport, proteolytic processing, and genome amplification. *J. Virol.* **70**, 7039–7048.

Schaad, M. C., Jensen, P. E., and Carrington, J. C. (1997). Formation of plant RNA virus replication complexes on membranes: Role of an endoplasmic reticulum-targeted viral protein. *EMBO J.* **16**, 4049–4059.

Schultze, M., Hohn, T., and Jiricny, J. (1990). The reverse transcriptase gene of cauliflower mosaic virus is translated separately from the capsid gene *EMBO J.* **9**, 1177–1185.

Shanks, M., and Lomonossof, G. P. (1990). The primary structure of the 24K protease from red clover mottle virus: Implications for the mode of action of comovirus proteases. *J. Gen. Virol.* **71**, 735–738.

Shanks, M., and Lomonossoff, G. P. (1992). The nucleotide sequence of red clover mottle virus bottom component RNA. *J. Gen. Virol.* **73**, 2473–2477.

Shanks, M., Dessens, J. T., and Lomonossoff, G. P. (1996). The 24 kDa proteinases of comoviruses are virus-specific in *cis* as well as in *trans*. *J. Gen. Virol.* **77**, 2365–2369.

Shapira, R., and Nuss, D. L. (1991). Gene expression by a hypovirulence-associated virus of the chesnut blight fungus involves two papain-like protease activities. *J. Biol. Chem.* **266**, 19419–19425.

Shen, P., Kaniewska, M., Smith, C., and Beachy, R. N. (1993). Nucleotide sequence and genomic organization of rice tungro spherical virus. *Virology* **193**, 621–630.

Shi, X. M., Miller, H., Verchot, J., Carrington, J. C., and Vance, V. B. (1997). Mutations in the region encoding the central domain of helper component-proteinase (HC-Pro) eliminate potato virus X/potyviral synergism. *Virology* **231**, 35–42.

Shirako, Y., and Wilson, T. M. A. (1993). Complete nucleotide sequence and organization of the bipartite RNA genome of soil borne wheat mosaic virus. *Virology* **195**, 16–32.

Shukla, D. D., Ward, C. W., and Brunt, A. A. (1994). Genome structure, variation and function. *In* "The Potyviridae" (D. D. Shukla, C. W. Ward, and A. A. Brunt, Eds.), pp. 74–112. CAB International, Cambridge.

Thornbury, D. W., van den Heuvel, J. F. J. M., Lesnaw, J. A., and Pirone, T. P. (1993). Expression of potyvirus proteins in insect cells infected with a recombinant baculovirus. *J. Gen. Virol.* **74**, 2731–2735.

Torruella, M., Gordon, K., and Hohn, T. (1989). Cauliflower mosaic virus produces and aspartic proteinase to cleave its polyproteins. *EMBO J.* **8**, 2819–2825.

Turnbull-Ross, A. D., Mayo, M. A., Reavy, B., and Murant, A. F. (1993). Sequence analysis of the parsnip yellow fleck virus polyprotein: Evidence of affinities with picornaviruses. *J. Gen. Virol.* **74**, 555–561.

Turnbull-Ross, A. D., Reavy, B., Mayo, M. A., and Murant, A. F. (1992). The nucleotide sequence of parsnip yellow fleck virus: A plant picorna-like virus. *J. Gen. Virol.* **73**, 3203–3211.

van der Wilk, F., Verbeek, M., Dullemans, A. M., and van den Heuvel, J. F. J. M. (1997). The genome-linked protein of potato leafroll virus is located downstream of the putative protease domain of the ORF1 product. *Virology* **234**, 300–303.

Van Rompaey, L., Proost, P., Van den Berghe, H., and Marynen, P. (1995). Design of a new protease inhibitor by the manipulation of the bait region of α_2-macroglobulin: Inhibition of the tobacco etch virus protease by mutant α_2-macroglobulin. *Biochem. J.* **312**, 191–195.

Verchot, J., and Carrington, J. C. (1995a). Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. *J. Virol.* **69**, 1582–1590.

Verchot, J., and Carrington, J. C. (1995b). Evidence that the potyvirus P1 proteinase functions in trans as an accesory factor for genome amplification. *J. Virol.* **69**, 3668–3674.

Verchot, J., Herndon, K. L., and Carrington, J. C. (1992). Mutational analysis of the tobacco etch potyviral 35-kDa proteinase: Identification of essential residues and requirements for autoproteolysis. *Virology* **190**, 298–306.

Verchot, J., Koonin, E. V., and Carrington, J. C. (1991). The 35-kDa protein from the N-terminus of a potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* **185**, 527–535.

Vos, P., Verver, J., Jaegle, M., Wellink, J., van Kammen, A., and Goldbach, R. (1988). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16**, 1967–1985.

Wellink, J., and van Kammen, A. (1988). Proteases involved in the processing of viral polyproteins. *Arch. Virol.* **98**, 1–26.

Zaccomer, B., Haenni, A.-L., and Macaya, G. (1995). The remarkable variety of plant RNA virus genomes. *J. Gen. Virol.* **76**, 231–247.