Review

P1 peptidase – a mysterious protein of family Potyviridae

JANA ROHOŽKOVÁ* and MILAN NAVRÁTIL

Faculty of Science, Palacký University in Olomouc, Šlechtitelů 11, 783 71, Olomouc-Holice, Czech Republic

**Corresponding author (Fax, +420-241-062219; Email, jana.rohozkova@img.cas.cz)*

The Potyviridae family, named after its type member, *Potato virus Y* (PVY), is the largest of the 65 plant virus groups and families currently recognized. The coding region for P1 peptidase is located at the very beginning of the viral genome of the family Potyviridae. Until recently P1 was thought of as serine peptidase with RNA-binding activity and with possible influence in cell-to-cell viral spreading. This N-terminal protein, among all of the potyviruses, is the most divergent protein: varying in length and in its amino acid sequence. Nevertheless, P1 peptidase in many ways is still a mysterious viral protein. In this review, we would like to offer a comprehensive overview, discussing the proteomic, biochemical and phylogenetic views of the P1 protein.

[Rohožková J and Navrátil M 2011 P1 peptidase – a mysterious protein of family Potyviridae. *J. Biosci.* **36** 189–200] **DOI** 10.1007/s12038-011-9020-6

1. Introduction

The Potyviridae family [named after its type member, *Potato virus Y* (PVY)] is the largest of the 65 plant virus groups and families currently recognized (Virus taxonomy list in 2009, ICTV). This family contains over 100 different species in six genera (Berger *et al*. 2000; López-Moya and García 2008). The viruses of this family cause significant losses in agricultural, pastoral, horticultural and ornamental crops (Ward and Shukla 1991). Within the Potyviridae family, the following genera are recognized: *Brambyvirus*, *Potyvirus*, *Rymovirus*, *Macluravirus*, *Ipomovirus* and *Tritimovirus*. A seventh genus, *Bymovirus*, contains viruses with bipartite genomes. Recent attempts to expand the genera include *Blackyvirus* (Susaimuthu *et al*. 2008), *Poacevirus* (Tatineni *et al*. 2009) and *Susmovirus* (Xu *et al*. 2010). Virions are flexuous and rod-shaped, 80 to 900 nm long and 11 to 15 nm wide. The nucleocapsid is composed of around 2000 units of a single structural protein, surrounding one molecule of nucleic acid (Carrington and Dougherty 1988). A feature shared by all of the potyviruses is the induction of characteristic pinwheel or scroll-shaped inclusion bodies in the cytoplasm of infected cells (Edwardson 1974). The predominant way of transmission of potyviruses is through aphids. Other possible means of the virus' spread are by mites and whiteflies. The genome is created by single-stranded RNA of positive sense, around 10 000 bp long. In most potyvirid species with a monopartite genome, their genomic sequence encodes nine multifunctional proteins plus the capsid protein. Each viral protein matures after its proteolytic cleavage from the polyprotein by nuclear inclusion-a peptidase (NIa), helper component proteinase (HC-Pro) and P1 peptidase. NIa and HC-Pro are multifunctional proteins, containing a C-terminal proteolytic domain as well as an N-terminal domain, each of which plays different roles in the replication cycle. Whether the topological quality of P1 protease should be compared to these is still not clear (Verchot *et al*. 1991).

The coding region for P1 proteinase is located at the very beginning of the viral genome. This terminal protein, among all

Keywords. Genetic diversity; P1 serine peptidase; Potyviridae; viral infection

Abbreviations used: BCMNV, *Bean common mosaic necrosis virus*; CI, cylindrical inclusion; CMV, *Cucumber mosaic virus*; CVYV, *Cucumber vein yellowing virus*; HC-Pro, helper component proteinase; MP, movement protein; NIa, nuclear inclusion-a peptidase; NBRF, National Biomedical Research Foundation; PPV, *Plum pox virus*; PSbMV, *Pea seed-borne mosaic virus*; PVA, *Potato virus A*; PVV, *Potato virus V*; PVX, *Potato virus X*; PVY, *Potato virus Y*; SMV, *Soybean mosaic virus*; TEV, Tobacco etch virus; TuMV, *Turnip mosaic virus*; TVMV, *Tobacco vein mottling virus*; WMV, *Watermelon mosaic virus*; YMV, *Yam mosaic virus*; YTHS, yeast two-hybrid system; ZYMV, *Zucchini yellow mosaic virus*

of the potyviruses, is the most divergent protein: varying in length (39–85) and in its amino acid sequence (NCBI Protein Database). This mostly non-conservative potyviral protein is thought to have contributed to the successful adaptation of the potyviruses into a wide range of host species (Valli *et al*. 2007).

With the accumulation of new information that has been assembled about potyviral P1, it is now an appropriate time to review the research progress on this protein. The aim of this review is to offer a comprehensive overview, discussing the proteomic, biochemical, and phylogenetic views of the P1 protein. Additionally, this review attempts to offer the latest research directions, following current theories on the function of the P1 protein.

2. Potyviral P1 protein: function and genetic diversity

It was not very long ago when the molecular properties of potyviruses were a mystery. The first breakthrough was achieved in 1986, when the complete genomic sequences of two members of this group were reported – for *Tobacco etch virus* (TEV) (Allison *et al*. 1986) and the *Tobacco vein mottling virus* (TVMV) (Domier *et al*. 1986). A year later, a step closer to uncovering the mystery occurred. In order to understand the functions of the potyviral proteins, the working group of Domier *et al*. (1986) made use of the then newly determined nucleotide sequences of the RNAs of TEV and TVMV. Subsequently, the predicted amino acid sequences of TVMV- and TEV-encoded proteins were compared with the protein sequence bank of the National Biomedical Research Foundation (NBRF). This study regarded as one of the significant studies that dealt with protein similarities among the potyviruses. After that came the useful achievement of identifying the functional properties of each genome's coding region, which brought cDNA clones of three potyviruses: TVMV (Domier *et al*. 1989), *Plum pox virus* (PPV) (Riechmann *et al*. 1990) and *Zucchini yellow mosaic virus* (ZYMV) (Gal-On *et al*. 1991). The ability to generate a virus infection from the cloned cDNA unlocked the possibility of applying the techniques of genetic engineering, at the molecular level, into their biology. The experiments initially performed on the previously mentioned group of viruses allowed for the beginnings of direct mutational tests and analyses at specific coding regions of the potyviruses, which had direct impact on both its phenotype as well as the process of its pathogenesis.

At this time, we can recognize two different groups of experiments dealing with the function of the P1 protein. On the one hand, there are those trying to solve the real function of P1 protease in several potyviruses: *Plum pox virus* (PPV) (Salvador *et al*. 2008), *Tobacco vein mottling virus* (TVMV) (Brantley and Hunt 1993; Klein *et al*. 1994; Salvador *et al*. 2008), *Cucumber vein yellowing virus* (CVYV) (Valli *et al*. 2006, 2008), *Soybean mosaic virus* (SMV) (Gagarinova *et al*. 2008), *Pea seed-borne mosaic virus* (PSbMV) (Aparicio *et al*. 2005), *Potato virus A* (PVA) (Merits *et al*. 1999; Kekarainen *et al*. 2002), *Potato virus V* (PVV) (Oruetxebarria and Valkonen 2001), *Potato virus Y* (PVY) (Arbatova *et al*. 1998; Pehu *et al*. 1995), *Potato virus X* (PVX) (Vance *et al*. 1995; Pruss *et al*. 1997), *Yam mosaic virus* (YMV) (Aleman-Verdaguer *et al*. 1997), *Tobacco etch virus* (TEV) (Carrington and Dougherty 1988; Carrington and Freed 1990; Verchot *et al*. 1992; Verchot and Carrington 1995a, b) and *Turnip mosaic virus* (TuMV) (Soumounou and Laliberté 1994). On the other hand, there are those experiments concerned with the sequence variability of the P1 coding region in YMV (Aleman-Verdaguer *et al*. 1997), PVA (Kekarainen *et al*. 1999), PVV (Oruetxebarria *et al*. 2000), PVY (Tordo *et al*. 1995; Mukherjee *et al*. 2004), ZYMV (Lee and Wong 1998; Wisler *et al*. 1995), TVMV and TEV (Domier *et al*. 1987), within the group of Potyviruses (Riechmann *et al*. 1992), within the Potyviridae family (Adams *et al*. 2005b; Valli *et al*. 2008), or among several viral families, including Potyviridae (Chare and Holmes 2006). There has been a dedicated effort on the phylogenetic studies of P1. Intensive research during the subsequent years led to a better understanding of the genomic structure of the potyvirus expression, quality, as well as the activity of specific viral proteins. Nevertheless, the P1 protein in many ways is still a mysterious viral protein, showing direct evidence of having some role in host range definition; however, this is yet to be demonstrated.

3. P1 serine peptidase

P1 peptidase was the last of the three peptidases to have been identified as processing the potyviral polyprotein (Verchot *et al*. 1991). In the microbial world, peptidases are not uncommon. Serine, aspartic and cystein peptidases (although not metallopeptidases) have been found in various viruses (Rawlings and Barrett 1993). They have acquired their importance because of their functional involvement in the processing of proteins of viruses that cause certain fatal diseases such as AIDS and cancer. Consequently, we have taken this opportunity to consolidate information about the progress in plant virus research as one of the crucial steps in understanding the activity of RNA viruses, not only in the realm of plants but also with the wider application of those viruses concerning and affecting humans.

All virus-encoded proteases are endopeptidases. The mature protease is released by the autolysis of the precursor (Rao *et al*. 1998). Serine proteases are characterized by the presence of a serine group within their active site. The numerous examples of serine proteases among viruses suggest that they are vital to the organisms (Barret 1994). With the intent of preventing serious losses to several major crop plants, strong efforts continue into research on the viral

proteases, precisely because this portion of the viral infection still offers likely prospects for answers leading to improvements in the defence mechanisms of the host plants.

The first research works dealing with the processing of viral polyproteins with proteases were published between 1979 and 1990 (Korant *et al*. 1979; Ghysdael *et al*. 1981; Franssen *et al*. 1982; Nicklin *et al*. 1986; Krausslich and Wimmer 1988; Wellink and van Kammen 1988; Skalka 1989; Oroszlan and Luftig 1990; Palmberg 1990; Strauss and Strass 1990). The reader is referred to these articles for additional specific information, and a summary of the earlier work on this subject. Of the known potyviral proteinases, NIa and HC-Pro were discovered in their multifunctional characteristics. The first N-terminal peptidase taking part in polyprotein processing (P1 protein) remains the last discovered, without any clear elucidation as to its function during viral infection and propagation. As listed earlier, there have been several interesting works that indirectly brought results which might ultimately be linked to the explanation of P1's true activities.

3.1 *Serine peptidase domain is located at the C-terminus of P1*

Verchot *et al*. (1991) were the first to mention P1 as a third proteinase in the processing of the viral polyprotein. In a wheat germ system, they had synthesized a polyprotein containing the 35 kDa protein and HC-Pro. Proteolysis generates products that resemble fully processed proteins. Furthermore, this protease was classified as a serine-like protease. All then-known facts on Potyviridae were compiled in an extensive review (Dougherty and Semler 1993).

Serine and serine-like peptidases are found in both prokaryotic and eukaryotic organisms, and participate in a wide spectrum of biological reactions. As a group, these enzymes are characterized by the presence of an active site domain that contains a reactive Ser. The active site also comprises two additional amino acids, Asp and His. These three residues make up a catalytic triad (for reviews, *see* Kraut 1977; Bond and Butler 1987; Craik *et al*. 1987a, b; Higaki *et al*. 1987; Sprang *et al*. 1987). The serine-like peptidase domain in the P1 coding region was identified in the C-terminal, highly conserved portion for all serine peptidases, associated with proteolytic activity and containing the Gly-Xaa-Ser-Gly motif (Barret 1986). It has been identified in all of the Potyviridae family members examined (Adams *et al*. 2005a; Valli *et al*. 2007). The conserved His and Asp residues are present upstream of the putative active-site Ser residue (for precise positioning in each virus, *see* Adams *et al*. 2005a; Valli *et al*. 2007); further, their spacing is not typical of those observed in other cellular serine proteases. In experiments involving the substitution of His or Ser in TEV (within positions His-215 and Ser-256), the results of Verchot *et al*. (1991) achieved the elimination of proteolytic activity.

3.2 *Processing of potyviral P1*

The processing of the serine protease, out of HC-Pro and polyprotein, is secured by a specific motif on the border of P1/HC-Pro. Mavankal and Rhoads (1991), with research on *Tobacco vein mottling virus* (TVMV), brought the first approaches in this research field of P1, which released itself from HC-Pro as the N-terminal 34 kDa protein. These facts have since been proven by the immunoprecipitation of proteins with antibody against HC protein. With the specific identification of the cleavage location by the automated Edman radiochemical degradation at positions Phe-256 and Ser-257 (of the TMVM polyprotein) confirmed its necessity for the proper functioning of HC-Pro in TMVM (Mavankal and Rhoads 1991). Later, Verchot *et al*. (1992) confirmed this theory with the specification of Tyr-304 and Ser-305 motif (of the TEV polyprotein) undergoing proteolysis with supplementary His-303. In deletion tests with this specific motif, they concluded that the expressed protein adjacent to His-303 was unable to undergo proteolysis, suggesting that Tyr-304 was not essential for protein processing but only for substrate and/or protein activity. From this finding arose the hypothesis that the cleavage occurs between the Tyr304 and Ser305, because the P1 cleavage site position (Tyr-304) is crucial for substrate recognition by several types of viral and cellular proteases (Verchot *et al*. 1992).

As mentioned earlier, potyviral P1 proteinase is an N-terminal product, derived from a genome-length polyprotein cleaving itself from HC-Pro. Verchot and Carrington (1995b) added a further contribution that this separation, but not P1's own, appears to be essential for viral infectivity. The relevance of P1 integrity for virus amplification and movement from cell to cell was assessed by mutagenesis. Using the coding region of *Tobacco etch virus* (TEV), they produced clones with the entire P1 coding region (DP1) deleted, further having investigated the reporter gene's expression of β-glucuronidase replication as well as the movement of the virus in both tobacco protoplasts and plants. In the protoplasts, this DP1 mutant accumulated to approximately 2% to 3% of the level of parental amplification, with both cell-to-cell and systemic movement, registering decreased values. On the contrary, the clone with a point mutation affecting the Ser of the proteolytic triad (S256A), which abolished the proteolytic activity of the protein and prevented its separation from the rest of polyprotein, yielded a non-viable mutant. A comparison of the infection phenotypes of DP1 and S256A mutants further suggests that the lack of proteolytic separation of P1 and HC-Pro to be more detrimental than the deletion of the P1 sequence altogether.

Experiments in 1992 with a specific motif of the P1 cleavage activity, resulting in it releasing itself from the rest of viral polyprotein, led to the specific information that these experiments had been successful in wheat germ extracts and transgenic plants, but not in rabbit reticulocyte lysate (Verchot *et al*. 1992). Furthermore, there had been previous efforts (Carrington and Freed 1990; Mavankal and Rhoads 1991; Verchot *et al*. 1991) that gave puzzling results. Based on the processing characteristics of P1-proteinase-containing polyprotein, within a mixture of reticulocyte lysate and wheat germ extract, Verchot *et al*. (1991) concluded that some positive-acting factor resides within the wheat germ system. Confirmation of this idea came from an experiment that involved adding small amounts of wheat germ extract to the rabbit reticulocyte cell-free expression system, which led to a revival of P1's protease activity. On the other hand, the addition of reticulocyte to wheat germ extract had no significant effect. Furthermore, there was an experiment using denatured wheat germ extract, which resulted in the assertion that the essential protein for proteinase activity had to be a heat-labile protein. The explanation of this phenomenon does not seem to reside solely in an essential co-factor in the P1 pathway but rather in a factor required for protein synthesis, as well as the proteins involved in the proper co-translational folding of the proteinase in the arrangement of the cleavage site. It is not the first time that a theory suggested this as a possible explanation of P1's ability to process itself under specific conditions. The involvement of accessory factors in polyprotein processing at specific cleavage sites has precedents in several other viral groups, including the comoviruses (Goldbach 1990), the picornaviruses (Harris *et al*. 1990) and flaviviruses (Rice and Strauss 1990).

3.3 *P1's N-terminal portion – the most variable part of the genome in all potyviruses*

From the first announcements regarding the sequence of the P1 coding region, this protein has been mentioned as the least conserved protein among the potyviruses (Domier *et al*. 1987; Vance *et al*. 1992), ranging in size from 30 to 63 kDa. Owing to this great divergence, and the discovery of the conserved motif in the C-terminal portion, it is obvious that the increase in the size is created by the N-terminal portion of P1 of each potyvirus. In a study by Verchot and Carrington (1995a), it was shown to be non-essential for most functions such as viral propagation, cell-to-cell movement, and viral replication. By use of the GUS reporter system, which is appropriate for quantitative measurements of the parental and mutant TEV genome amplifications in inoculated protoplasts (Carrington *et al*. 1993; Dolja *et al*. 1992; Restrepo-Hartwig and Carrington 1994), Verchot and Carrington (1995a) characterized three truncated forms of P1 protein with the deletion in the N-terminal portion. The P1 coding region of the TEV lacked the region from 1 to 139 nt in the Δ5 mutant, and the region of 1–157 nt in the Δ6 mutant, i.e. in almost half of the P1 coding region, but not affecting the catalytic triad of proteolytic activity. Interestingly, most dramatic mutations in the Δ5 and Δ6 constructs actually had only small effects on viral replication and cell-to-cell movement. Experiments by Kekarainen *et al*. (2002) and Rajamäki *et al*. (2005) gave results inconsistent with those of deletions. Both of these two working groups have confirmed the high tolerance of P1 to insertions. The aim of the work by Kekarainen *et al*. (2002) was to map the genome regions (both essential and non-essential) for PVA propagation.

To that end, they generated 15-bp-insertion mutants. The C-terminal region of P1 tolerated most of the insertions. Based upon those results of tolerance to short insertions, they constructed a map of the identified essential, as well as non-essential, regions of P1. From this viewpoint, we can note a conflict in the categorization of the importance of P1's N-terminal portion. Seventy percent of the regions described as essential, resulting in non-viable mutants, were located in the N-terminal half of P1. Rajamäki *et al*. (2005) also demonstrated a similar tolerance. They confirmed this ability by the insertion (insertion of up to 783 nucleotides) and expression of heterologous protein in PVA's N-terminal portion of the P1-coding region.

The next efforts to identify an even-lesser-conserved motif in the P1 N-terminal portion (which could help in understanding its evolution), and at first sight a non-conservative portion of the potyviral genome, were by Gagarinova *et al*. (2008) and Valli *et al*. (2008). They both focused on the sequencing analysis of the whole coding region of the P1 protein in batches of potyviruses. Their interesting results and conclusions are analysed in section 4 of this article.

3.4 *P1 is able to bind RNA*

The net charge of the P1 protein was computationally evaluated to be +30 (TVMV; Brantley and Hunt 1993) and +38 (TuMV; Soumounou and Laliberté 1994), which are much higher than those of any other potyviral protein. This charge has been characterized as a z value. It has been proposed that the binding of a ligand of charge $+z$ to a linear nucleic acid would neutralize z phosphates, resulting in the release of the counter-ions that are thermodynamically associated with those z phosphates (Record *et al*. 1976). The existence of a net charge for P1, of the magnitude mentioned above, led to the suggestion that it represents nucleic-acid-binding activity. Brantley and Hunt (1993) were the first to assert that P1 generally binds RNA, and they subsequently decided to characterize the possible RNA-binding properties. They had shown that P1 is able to bind 35-nt-long RNA, and that the size of the binding site is consistent with the number of phosphate groups that

would be neutralized by the binding of a protein of the +30 charge previously mentioned (Mascotti and Lohman 1990). The next set of experiments were led by Soumounou and Laliberté (1994), who had analysed P1 of TuMV, and contributed further with pronouncements of the binding capabilities to ss- and ds-RNA. Under *in vitro* conditions, it was confirmed that P1 binds to ss- but not to ds-DNA. The potential interaction domain of P1 in TuMV was characterized as the very basic domain, RSSRAMKQKRARERR RAQQ (spanning residues 150 to 168), which has the potential to interact with nucleic acids. The basic residues (Lys and Arg) can form ionic bonds with the negatively charged phosphate groups; whereas amide (Q), acidic (E) or hydroxylated (S) amino-acids could interact, via hydrogen bonds, with the nucleic acid bases (Soumounou and Laliberté 1994). Contributing factors that have led to the speculation of P1's silencing activity during the viral infection cycle should be mentioned. Concluding with the finding that the depletion of the entire P1 protein had not disrupted the infection cycle, and that even P1's ability to bind RNA, could suggest that this protein acts as a helper protein, improving the expression of the viral genome; yet, on the other hand, P1 could subdue the defense activities of the host plant.

In order to prove P1's activity in overcoming host plant defense mechanisms, several experiments have been performed that dealt with transgenic plants expressing the short-sequence homologue to P1 (Mäkki-Valkama *et al*. 2000a, b; Tavert-Roudet *et al*. 1998). However, more extensive research has attempted to define the extent of P1's involvement in the processes of interactions between virus and its host. Pruss *et al*. (1997) reported that the potyviral sequence P1/HC-Pro enhanced the pathogenicity and accumulation of two heterologous viruses [*Cucumber mosaic virus* (CMV) and TMV]. Kasschau and Carrington (1998) dealt with P1 and its influence on transcription activity, performing the nuclear transcription assay. These experiments indicated that the silencing suppression activity of P1/HC-Pro could possibly be on the post-transcriptional level. These two results are in agreement, and could reveal that this combination of molecules can condition an enhanced susceptibility within a host through the interdiction of the defense response. In this connection, Valli *et al*. (2006) proved the crucial role of P1 in the HC-Pro gene's silencing activity during the infectious, providing further evidence that P1 enhances the activity of HC-Pro in the *Potyvirus* genus.

One very special situation within the P1 proteinase coding region and the activity of this protein has been described in the ipomovirus (family Potyviridae) *Cucumber vein yellowing virus* (CVYV). It was recently reported that this virus lacks HC-Pro, but has a duplicated P1 coding sequence (Janssen *et al*. 2005). After cleavage, two mature proteins are produced, P1a and P1b. Both of these are serine proteases (Valli *et al*. 2007). HC-Pro in potyviruses is the typical silencing suppressor of viruses of the family Potyviridae (Anandalakshmi *et al*. 1998; Brigneti *et al*. 1998; Kasschau and Carrington 1998). The crucial point of this P1 study was a demonstration of P1b protease activity, as well as the RNA-silencing suppression ability, which is consistent with that of HC-Pro (Valli *et al*. 2008). According to the subsequent discoveries for the RNA-silencing quality, the specific LXKA motif is crucial – it has been suggested as the zinc finger motif (Cox and McLendon 2000; Klug and Rhodes 1987); after its point mutation, P1b's silencing activity was abolished. Nevertheless, disruption of protease activity has no acute effects on this silencing quality. Additionally, there were informational data sets with complementary findings about P1 ability to bind siRNA in crude plant extracts. This piece of information supports the suggestion that CVYV P1b uses a strategy of siRNA sequestration, with the interference of viral RNA degradation. Within the context of this interesting discovery, it is very important to mention that the positive charge of the P1 protein is universal in all of the members within the genus *Potyvirus*. In contrast, the ipomoviral P1bs, or P1b-like P1 protein from ipomoviruses (as well as tritimoviruses), either has a neutral or a negative charge.

3.5 *Atypical P1s – Tritimovirus and Ipomovirus*

Because of phylogenetic discoveries, we now know that rymoviruses are more closely related to potyviruses than to ipomoviruses. Tritimoviral P1 protein constitutes an independent branch in the tree (Adams *et al*. 2005b). The most conspicuous difference between the two types of P1 is represented by their pI difference, which probably reflects not only a great phylogenetic distance but also some functional divergence. As has been mentioned before, CVYV lacks the HC-Pro gene and its tritimo-like P1b protein appears to compensate for this defect. This could be conclusive evidence for the occurrence of the functional diversification between poty-like and tritimo-like P1s', P1a's, as well as P1b's own protease activities. If these two putative P1 protease domains are included in the phylogenetic analysis, P1a clustered with the potyviral and rymovila P1s and P1b was more closely related to the tritimoviral P1s (Valli *et al*. 2007; data not shown)

3.6 *P1's localization in infected plant cell*

Arbatova *et al*. (1998) completed one of the first attempts to localize the P1 protein on the ultra-structural level of the infected cell. They produced polyclonal antibody against the P1 protein of PVY-O. Fusing this antibody to gold particles, they conducted their experiments with the immunogold localization of the P1 protein in the cells of infected plants.

This protein has been found in association with cytoplasmic inclusion bodies, characteristic for several potyviruses. Particles marking the P1 protein have also been found freely diffused in the cytoplasm. However, no significant P1 antibody binding with other plant cell organelles, or with the cell wall and plasmodesmata, was detected with the immunogold labelling. Arbatova *et al*. (1998) described their findings of P1 associated with pinwheel-shaped inclusion bodies with two possible explanations: (1) P1 could stay associated with this complex due to its simultaneous synthesis there as well, as it has been shown by the P3 protein (Ammar *et al*. 1993, 1994; Rodríguez-Cerezo *et al*. 1991, 1993). (2) As suggested by Ammar *et al*. (1994), all potyviral proteins might by synthesized in, on or near the inclusions, which are associated with a rough endoplasmic reticulum. Speculative, yet conclusive, seems to be the inference stemming from P1's ability to bind RNA, and the fact that several point mutations have led to a decrease of viral particles in infected plants. The piece of evidence behind the development of this theory is the newly found helicase activity of cylindrical inclusion protein (CI); and related to this matter, replication of viral RNA is likely to also be associated with the connections found with cytoplasmic inclusions. This might be the missing piece accounting for the P1 cytoplasmic localization. In regard to this information, it would be appropriate for someone to supplement such conclusions with experiments that confirm any biochemical pathways that might help account for an explanation of the localization, i.e. the finding of molecular interaction partners of P1 protein during the infection cycle.

3.7 *P1 and its activity in viral cell-to-cell movements and systemic spreading*

P1 activity, influencing virus cell-to-cell movement, has been hypothetically suggested (Domier *et al*. 1987; Hull 1989; Maiss *et al*. 1989; Atabekov and Taliansky 1990; Ward and Shukla 1991; Riechmann *et al*. 1992). Later, such P1 activity was refuted by Verchot and Carrington (1995a, b). As previously mentioned, experiments with point-mutated P1 protein of TEV (Verchot and Carrington 1995b) resulted in a group of nonviable viruses (three mutations termed S256A, F and Δ304). Nevertheless, a fourth proteinase-debilitating mutation (termed C) caused a slow-infection phenotype. According to these findings, the speculation about a direct movement function of the P1 protein now appears to have little basis; however, a movement-enhancing function cannot be totally excluded. Several independent researchers (Klein *et al*. 1994; Cronin *et al*. 1995; Verchot and Carrington 1995b; Kasschau *et al*. 1997) have found that P1 and HC-Pro are both active enhancers of genome amplification at the single-cell level, and have additionally demonstrated the long-distance movement. However, parallel to this information, it has also been asserted that neither of these two proteins appears to be required for cell-to-cell movement. Arbatova *et al*. (1998) later indirectly confirmed this assertion. Immunogold localization of PVY P1 protein in infected plants showed an absence of this protein in the cell wall and plasmodesmata (Arbatova *et al*. 1998). On the other hand, we need to mention the situation in the sobemoviruses. Two open reading frames (ORF1 and ORF2) create their genome, while P1 is located in ORF1. Due to the theory about the non-essentiality of P1 in virus spread, there have been conflicts with previous findings. Bonneau *et al*. (1998) proclaimed, working on *Rice yellow mottle virus* (RYMV), that the creation of deleted and frame-shifted mutants in the P1 coding region (resulting in truncations of 83 amino acids from the C-terminus of P1) were incapable of replicating in protoplasts. Furthermore, a mutant not expressing P1 at all replicated in a protoplast at a reduced level (0.5–2× less), when compared with the replication of the wild-type RNA. In the case of cell-to-cell movement as a system infection, none of the mutants caused systemic infection in the host rice plants. Their results demonstrated that in the case of the sobemoviral infection, P1 of RYMV is indispensable for virus replication; however, nucleotide deletions or additions in ORF1 are lethal for virus replication. Furthermore, P1 of RYMV is required for the infection of plants and is important for the spread of the virus.

3.8 *P1's interaction partners in the infected cell*

Further efforts to study the P1 molecular interaction partners were led by Merits *et al*. (1999). However, they tried to analyse the possible P1 pathway during an infection cycle; thus, their work involved an interaction study of only the viral proteins. Consistent with the balance between P1's affinity for one or more viral proteins, they had expected P1's activity as a helper factor of the presently identified mechanisms during the infection cycle. Using *Escherichia coli*–expressed recombinant proteins in two *in vitro* interaction assays, and a genetic yeast two-hybrid system (YTHS), they tried to analyse the protein–protein interactions of the P1 and P3 proteins of PVA with six other viral proteins, creating presumed replication complexes. In these *in vitro* experiments, they concluded that P1 and P3 interacted with each other, as well as with proteins of the putative replication complex of potyvirus: RNA helicase (CI), viral protein genome linked (VPg), NIa peptidase part (NIa-pro) and RNA-dependent RNA-polymerase (NIb). P1 also interacted with itself and with HC-Pro. In this situation, it seems to be reasonably confirmed that the interaction between the P1 and CI proteins form the main superficial component of cytoplasmic inclusion bodies, as shown by Arbatova *et al*. (1998) and Rodríguez-Cerezo *et al*. (1993, 1997).

Experiments that resulted in the interaction of P1 with HC-Pro have endeavoured to suggest that these interactions are associated with events that help proteinase domains to recognize their cleavage sites. The creation of a self-dimer of P1, and its possible interaction with P3, has contributed to the proposed role of the protein's interaction with the replication complex of the potyvirus. Other interactions explored in regard to viral proteins are yet to be explained. However, it is necessary to add that Merits *et al*. (1999) have been cautious in the interpretation of their results reached, because protein expression in bacterial cells possibly lacking post-translational modification (phosphorylation and glycosylation) could have a large influence on real protein activity.

A novel identification of the interaction partner of P1 was discovered using yeast two-hybrid screen coupling P1 from SMV-P (*Pinellia* isolate) and a c-DNA expression library of its host, the aroid *Pinellia ternate*. Shi *et al*. (2007) identified an interesting 23.7-kDa-sized molecule that was interacting with P1 and was closely related to the cytochrome b6/f complex Rieske Fe/S genes of plants. The topology of the interaction of P1 to Rieske Fe/S was found to be located at the N-terminal part of the protein (1–33 amino acids), interacting with the transitional form of the Rieske Fe/S protein; whereas, a 34- to 82-amino-acid-long motif was able to interact with the entire Rieske Fe/S protein. Owing to these results, some authors have suggested that the P1–Rieske Fe/S pair is likely to be involved in symptom development, and that the very variable N-terminus of P1 may play an important role in host adaptation.

A specific feature about P1, recently discovered, is its ability to induce production of the HSP70 heat-shock protein. Experiments dealing with this activity had been performed by Aparicio *et al*. (2005), using *GUS* as a reporter gene fused to the HSP70 coding region of *Arabidopsis*. This construct, infiltrated into the leaf of *Nicotiana benthamiana*, allowed them to see the different responses of the promoter–reporter system. A few of the tested constructs have been used for testing the hypothesis that viral induction of pHSP70-GUS represents a broad response to the expression of the virus' proteins. For their experiment, Aparicio *et al*. selected *Tobacco mosaic virus* (TMV, genus *Tobamovirus*) movement protein (MP), the P1 protein of *Pea seed borne mosaic virus* (PSbMV, genus *Potyvirus*) and a 6 kDa protein (6K1) fused to the cylindrical inclusion (CI) protein from PSbMV. All of these proteins showed a significant induction of pHSP70-GUS as increased *GUS* activity; the P1 protein was the most activating one. Increase in the activity of heat-shock protein have been demonstrated by a wide range of viruses in diverse host plants including pea (*Pisum sativum*), *Nicotiana benthamiana*, squash (*Cucurbita pepo*), tobacco (*Nicotiana tabacum*) and *Arabidopsis thaliana* (Aranda *et al*. 1996, 1999; Escaler *et al*. 2000; Havelda and Maule 2000; Jockusch *et al*. 2001; Whitham *et al*. 2003). These authors have conceding that viral proteins crucial to the viral infection cycle may have a specific induction mechanism that would meet the particular needs of viral genome replication and expression. In any case, there remains a possible alternative interpretation, where the P1 activity is just like the response of pHSP70 in the ability of cytosol to sense the individual properties of particular proteins when expressed at a high level. This phenomenon is reminiscent of the unfolded protein response observed when the induced accumulation of proteins in the endoplasmic reticulum also induces a specific suite of chaperons.

4. Evolution of P1 serine proteinase

According to the MEROPS databases, P1 proteinase has been classified within the Clan SA, S30 family of serine proteases. It is possible to attempt an overview of the evolution of P1 proteinase. Although there are publications that have concluded that the evolution of P1 in Potyviridae emerged from the sequences of the coding region, with their differences in evaluating possible conserved motifs (listed further ahead), we can speculate about phylogenesis inside the protease family, leading to a theory about its selection under the functional properties of proteinase and its adaptation in the different metabolic pathways of new host organisms (Barrett and Rawlings 2007). Initially, the first announcements about this phenomenon were described by Rawlings and Barrett (1993), who tried to identify specific groups of peptidase, according to scored similarities in a sequence of over 600 peptidases. They suggested that the classification by peptidase families could be used as an extension of the present classification by catalytic type. No studies dealing with the evolutionary origins of family S30 have been done; however, several families of serine proteinases present in viral organisms have been found, for example, the S3-*Togavirus* endopeptidase with the catalytic triad His/Asp/Ser as well as the family S14-ClpP, identified by *Potato leafroll luteovirus* genomic RNA (Luteoviridae, the genome is created in a similar manner to Potyviridae, by single-stranded RNA of positive sense) with the catalytic motif Ser/His (Asp not known). Rawlings and Barrett (1993) concluded that due to the highly conserved catalytic triad motif of serine peptidase activity, these proteins likely share a common evolutionary origin with family S1 (most of the eukaryotic and bacterial serine peptidases), despite the differences in their sequences. Due to the very small conserved portions, there might be some doubts about the actual relationships between the first site-distinct viral species. Structural/functional studies in many of the peptidase families have highlighted particular amino acid residues that contribute to the specificity sites of the enzymes. Variations of these, between sequences, can be taken as strong indications of the

differences in substrate specificity, and therefore in their function. Such differences would surely require the assignment of the peptidases concerned to be used by the distinct species (Barrett and Rawlings 2007).

4.1 *Phylogenesis of potyviral P1 protein*

Recombination is one of the main forces driving plant virus evolution (García-Arenal *et al*. 2003; Roossinck 2003). Both interspecies and intraspecies recombination events are involved in potyviral evolution, and some of these affect the P1 sequence (Glais *et al*. 2002; Desbiez and Lecoq 2004; Tan *et al*. 2004; Larsen *et al*. 2005; Valli *et al*. 2006). Up to now, there have been several attempts at phylogenetic analysis of the most variable coding region of P1. Each of these has been made achievable from the RNA or amino acid viral sequences, which were taken from theoretical cDNA translations. Despite these functional analyses, phylogenetic analyses have been performed on TVMV and TEV (Domier *et al*. 1987), WMV (Desbiez and Lecoq 2004; Ali *et al*. 2006) and *Soybean mosaic virus* (SMV) (Gagarinova *et al*. 2008). Additionally, Valli *et al*. (2007) attempted the widest overview and most extensive endeavours to explain the origin and phylogeny path of potyviral P1. They performed extensive computational analysis of P1 proteins from 53 virus species in the family Potyviridae. In their last, most extensive assay (Valli *et al*. 2007), they have tried to explain recombination and gene duplication in the evolutionary diversification of P1 proteins.

The first effort to specify some homology or conserved motifs in the P1 coding region, shared by two potyviruses (TVMV and TEV), was described by Domier *et al*. (1987). However, after the TVMV and TEV polyprotein sequence comparison, they announced that the regions of the N-terminal and the 42 kDa region (also involving the P1 coding region) clearly showed less homology. They justified these findings by suggesting that the apparent lack of sequence conservation found in these two regions was due to the involvement of these proteins in specific virus–host interactions.

An extension of the theory regarding the presence of specific motifs in correlation with the host–virus interaction came from the work of Desbiez and Lecoq (2004). They had focused their attention on the N-terminal region of P1, and had assumed it was especially relevant in regard to the previously mentioned interaction between the virus and host organism. This, together with the theory about the interspecific recombination between two related potyviruses WMV and SMV, revealed the specific nucleotide sequence variability of WMV in the 5′ part of the genome (Desbiez and Lecoq 2004). Using that paradigm, Larsen *et al*. (2005) working on insertion mutants of *Bean common mosaic necrosis virus* (BCMNV) demonstrated by the differences found within the P1 N-terminus that although P1 played some significant role in the pathogenicity and virulence, it did not appear to affect the virus–host range (Larsen *et al*. 2005).

The results from comparison analysis suggested that not only intraspecies and intragenus but also intergenus recombination within the P1 gene contributed to potyvirus evolution. Well-conserved motifs were identified within the C-proximal protease domain of all *Potyvirus* P1s (H-D/E-G-x-S-G-I/V-I/V-R-G). The cleavage site between P1 and the next protein, HC-Pro, is also well conserved. It is 22 to 28 amino acids downstream of the strictly conserved RG dipeptide, and it has as its consensus sequence I/V/L/M-x-H/E/Q-F/Y ↓ S. Short, but also frequently present, conserved motifs have been demonstrated to be spread all over the P1 coding region. Motifs IXFG and VELI have been shown to be represented in most potyviruses and rymoviruses. Conversely, there are also motifs only represented within a group of potyviruses, for example, ISIXGG, TPS and FLXG in the small group of potyviruses (for a definitive review, *see* Valli *et al*. 2007). All of those are located approximately in the middle of the P1 coding region. Finally, the identified cystein-rich domain is represented by 13 potyviral viruses usually found in the N-terminal portion of P1. Despite their certain identification, these short motifs have no simple phylogenetic relationships justifying either their presence or absence. Interestingly, all of these results suggest that the potyviral P1 gene has undergone extensive and uneven evolutionary diversification, which has not always paralleled the evolution of the complete genome (Valli *et al*. 2007).

Gagarinova *et al*. (2008) conducted the latest experiments dedicated to gathering evidence of recombination breakage points in P1 on selected potyviruses (which could be helpful in an explanation of P1 phylogenesis). In contrast to previous studies, they used the Recombination Detection Programme 3.3.1 (RDP3), which provided automated analysis using the RDP, GENECONV, Bootscan, MaxiChi, Chimera and SiScan methods (Martin *et al*. 2005) instead of the Genetic Algorithms for Recombination Detection (GARD, Kosakovsky Pond *et al*. 2006) used by Valli *et al*. (2007) for mapping phylogenetic detection of recombination using a genetic algorithm, and Chare and Holmes's (2006) application of Sawyer's Runs Test for the detection of recombination events. Gagarinova *et al*. (2008) demonstrated for the first time that recombination occurred during SMV evolution among distinct viral isolates, thus providing evidence that at least two distinct viral SMV pathotypes could simultaneously infect a host cell and exchange genetic materials through RNA recombination.

5. Concluding remarks

The understanding of P1 peptidase has expanded rapidly in the genomic as well as in proteomic area. The brief overview of known facts about this peptidase presented herein demonstrates

that much is known. However, P1 remains the mysterious protein. In our opinion, a possible way to uncover P1 properties is to identify proteins interacting with it during infection cycle and to try to co-localize them on the level of light microscopy as well as electron microscopy. From the proteomic point of view, it is also important to identify the concrete structure of this protein. In spite of its high variability in the coding sequence all over Potyviridae, there have to be conserved structural motif (even several amino acids) responsible for P1's interaction with other active molecules of virus or host plant. This identification seems to be crucial for uncovering the real activity of P1. Biochemical and proteomic studies of P1 peptidase are essential for this. Such work will offer us a deeper understanding of the involvement of P1 in viral process, spreading and its involvement during viral infection.

Acknowledgements

This work was supported by grant MSM 6198959215 (Ministry of Education, Youth and Sports of the Czech Republic). The excellent critical review by Juan Antonio Garcia Alvarez is gratefully acknowledged. We thank Peter Lemkin for editing the manuscript.

References

Adams MJ, Antoniw JF and Fauquet CM 2005a Molecular criteria for genus and species discrimination within the family Potyviridae. *Arch. Virol.* **150** 459–579

Adams MJ, Antoniw JF and Beaudoin F 2005b Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. *Mol. Plant Pathol.* **6** 471–487

Aleman-Verdaguer ME, Goudou-Urbino C, Dubern J, Beachy RN and Fauquet C 1997 Analysis of sequence diversity of the P1, HC, P3 NIb and CP genomic regions of several yam mosaic potyvirus isolates: implications for the intra-species molecular diversity of potyviruses. *J. Gen. Virol.* **78** 1253–1264

Ali A, Natsuaki T and Okuda S 2006 The complete nucleotide sequence of a Pakistani isolate of *Watermelon mosaic virus* provides further insights into the taxonomic status in the *Bean common mosaic virus* subgroup. *Virus Genes* **32** 307–311

Allison R, Johnston RE and Dougherty WG 1986 The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence of the synthesis of a single polyprotein. *Virology* **154** 9–20

Ammar ED, Rodriguez-Cerezo E, Shaw JG and Pirone TP 1993 Association of virions and coat protein of tobacco vein mottling potyvirus with cylindrical inclusions in tobacco leaves and protoplasts. *Phytopathology* **83** 1374–1375

Ammar ED, Rodriguez-Cerezo E, Shaw JG and Pirone TP 1994 Association of virions and coat protein of tobacco vein mottling potyvirus with cylindrical inclusion in tobacco cells. *Phytopathology* **84** 520–524

Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH and Vance VB 1998 A viral suppressor of gene silencing in plants. *Proc. Natl. Acad. Sci. USA* **95** 13079–13084

Aparicio F, Thomas CL, Lederer C, Niu Y, Wang D and Maule AJ 2005 Virus induction of Heat shock protein 70 reflects a general response to protein accumulation in the plant cytosol. *Plant Physiol*. **138** 529–536

Aranda MA, Escaler M, Thomas CL and Maule AJ 1996 Induction of HSP70 and polyubiquitin expression associated with plant virus replication. *Proc. Natl. Acad. Sci. USA* **93** 15289–15293

Aranda MA, Escaler M, Thomas CL and Maule AJ 1999 A heat shock transcription factor in pea is differentially controlled by heat and virus replication. *Plant J.* **20** 153–161

Arbatova J, Lehto K, Pehu E and Pehu T 1998 Localization of the P1 protein of *potato Y potyvirus* in association with cytoplasmic inclusion bodies and in the cytoplasm of infected cells. *J. Gen. Virol.* **79** 2319–2323

Atabekov JG and Taliansky ME 1990 Expression of plant virus-coded transport function by different viral genomes. *Adv. Virus Res.* **38** 201–248

Barret AJ 1986 An introduction to the proteinases; in *Proteinase inhibitors* (eds.) AJ Barrett and G Salvesen (New York: Elsevier) pp 3–22

Barret AJ 1994 Proteolytic enzymes: serine and cystein peptidases. *Method. Enzymol.* **244** 1–15

Barrett AJ and Rawlings ND 2007 'Species' of peptidases. *Biol. Chem*. **388** 1151–1157

Berger PH, Banett OW, Brunt AA, Colinet D, Edwardson JR, Hammond J, Hill J, Jordan RL, Kashiwazaki S, *et al.* 2000 Family Potyviridae; in *Virus taxonomy: Seventh report of International Committee on Taxonomy of Viruses* (Academic Press) pp 703–724

Bond JS and Butler PE 1987 Intracellular proteases. *Annu. Rev. Biochem.* **56** 333–364

Bonneau C, Brugidou C, Chen L, Beachy R and Fuaquet C 1998 Expression of the Rice yellow mottling virus P1 protein *in vitro* and in vivo and its involvement in virus spread. *Virology* **244** 79–86

Brantley JD and Hunt AG 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

Brigneti G, Voinnet O, Li WX, Ji L H, Ding SW and Baulcombe DL 1998 Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana bethamiana*. *EMBO J.* **17** 6739–6746

Carrington JC and Dougherty WG 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 JC and Freed DD 1990 Cap-independent enhancement of translation by a plant potyvirus 5′ nontranslated region. *J. Virol.* **64** 1590–1597

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

Chare ER and Holmes EC 2006 A phylogenetic survey of recombinant frequency in plant RNA viruses. *Arch. Virol.* **151** 933–946

Cox EH and McLendon GL 2000 Zinc-dependent protein folding. *Curr. Opin. Chem. Biol.* **4** 162–165

Craik CS, Roczniak S, Lagman C and Rutter RJ 1987a The catalytic role of the active site aspartic acid in serine proteases. *Science* **237** 909–913

Craik CS, Roczniak S, Sprang S, Fletterick R and Rutter W 1987b Redesigning trypsin via genetic engineering. *J. Cell Biochem.* **33** 199–211

Cronin S, Verchot J, Haldeman-Cahill R, Shaad MC and Carrington JC 1995 Long-distance movement factor: a transport functions of the potyvirus helper component-proteinase. *Plant Cell* **7** 549–559

Desbiez C and Lecoq H 2004 The nucleotide sequence of *Watermelon mosaic virus* (WMV, Potyvirus) reveals interspecific recombination between two related potyviruses in the 5′ part of the genome. *Arch. Virol.* **149** 1619–1632

Dolja VV, McBride HJ and Carrington JC 1992 Tagging of plant potyvirus replication and movement by insertion of β-glucuronidase into the viral polyprotein. *Proc. Natl. Acad. Sci. USA* **89** 10208–10212

Domier LL, Franklin KM, Shahabuddin M, Hellmann GM, Overmeyer JH, Hiremath ST, Siaw MFE, Lomonossoff GP, *et al*. 1986 The nucleotide sequence of tobacco vein mottling virus RNA. *Nucleic Acids Res.* **14** 5417–5430

Domier LL, Shaw JG and Rhoads RE 1987 Potyviral proteins share amino acid sequence homology with *picorna-*, *como-* and *caulimoviral* proteins. *Virology* **158** 20–27

Domier LL, Franklin KM, Hunt AG, Rhoads RE and Shaw JG 1989 Infectious *in vitro* transcripts from cloned cDNA of a potyvirus, tobacco vein mottling virus. *Proc. Natl. Acad. Sci. USA* **86** 3509–3513

Dougherty WG and Semler BL 1993 Viral proteinases; *Microbiol. Rev.* **57** 781–822

Edwardson JR 1974 Some properties of the potato virus Y-group. *Fla. Agric. Exp. Stn. Monogr. Ser.* **4** 225

Escaler M, Aranda MA, Thomas CL and Maule AJ 2000 Pea embryonic tissues show common responses to the replication of a wide range of viruses. *Virology* **267** 318–325

Franssen H, Goldbach R, Broekhuijsen M, Moerman M and van Kammen A 1982 Expression of middle-component RNA of *Cowpea mosaic virus*: *in vitro* generation of a precursor to both capsid proteins by a bottom-component RNA-encoded protease from infected cells. *J. Virol.* **41** 8–17

Gagarinova AG, Babu M, Strömvik M and Wang A 2008 Recombination analysis of Soybean mosaic virus sequences reveals evidence of RNA recombination between distinct pathotypes. *Virol. J.* **5** 143–151

Gal-On A, Antignus Y, Rosner A and Raccah B 1991 Infectious *in vitro* RNA transcripts derived from cloned cDNA of the cucurbit potyvirus *Zucchini yellow mosaic virus*. *J. Gen. Virol.* **72** 2639–2643

García-Arenal F, Fraile A and Malpica JM 2003 Variation and evolution of plant virus populations. *Int. Microbiol.* **6** 225–232

Ghysdael J, Neil JC and Vogt PK 1981 Cleavage of four avian sarcoma virus polyproteins with virion protease p15 removes gag sequences and yields large fragments that function as tyrosine phosphoacceptors *in vitro*. *Proc. Natl. Acad. Sci. USA* **78** 5847–5851

Glais L, Tribodet M and Kerlan C 2002 Genomic variability in *Potato potyvirus Y* (PVY): evidence that $PVY^{N}W$ and PVY^{NTN} variants are single to multiple recombinants between PVY^{O} and PVY^{N} isolates. *Arch. Virol.* **147** 363–378

Goldbach H 1990 Plant viral proteinases; *Semin. Virol.* **1** 335–346

Harris KS, Hellen CUT and Wimmer E 1990 Proteolytic processing in the replication of picornaviruses. *Semin. Virol.* **1** 323–333

Havelda Z and Maule AJ 2000 Complex spatial responses to cucumber mosaic virus infection in susceptible *Cucurbita pepo* cotyledons. *Plant Cell* **12** 1975–1985

Higaki JN, Gibson BW and Craik CS 1987 Evolution of catalysis in the serine proteases. *Cold Spring Harb. Symp.* **52** 615–621

Hull R 1989 The movement of viruses in plants; *Annu. Rev. Phytopathol.* **27** 213–240

Janssen D, Martín G, Velasco L, Gómez P, Segundo E, Ruiz L and Cuadrado IM 2005 Absence of coding region for the helper component proteinase in the genome of a cucumber vein yellowing virus, a whitefly transmitted member of Potyviridae. *Arch. Virol.* **150** 1439–1447

Jockusch H, Wiegand C, Mersch B and Rajes D 2001 Mutants of tobacco mosaic virus with temperature-sensitive coat proteins induce heat shock response in tobacco leaves. *Mol. Plant-Microb. Interact.* **14** 914–917

Kasschau KD and Carrington JC 1998 A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell* **95** 461–470

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

Kekarainen T, Merits A, Oruetcebarria I, Rajamäki M and Valkonen JPT 1999 Comparison of the complete sequences of five different isolates of *Potato virus A* (PVA), genus Potyvirus. *Arch. Virol.* **144** 2355–2366

Kekarainen T, Savilahti H and Valkonen JPT 2002 Functional genomics on *Potato virus A*: Virus genome-wide map of sites essential for virus propagation. *Genome Res.* **12** 584–594

Klein PG, Klein RR, Rodríguez-Cerezo E, Hunt AG and Shaw JG 1994 Mutational analysis of tobacco vein mottling virus genome. *Virology* **204** 759–769

Klug A and Rhodes D 1987 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12** 464–469

Korant B, Chow N, Lively M and Powers J 1979 Virus-specified protease in poliovirus infected HeLa cells. *Proc. Natl. Acad. Sci. USA* **76** 2992–2995

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH and Frost SD 2006 GARD: a genetic algorithm for recombination detection. *Bioinformatics* **22** 3096–3098

Krausslich H and Wimmer E 1988 Viral proteinases. *Annu. Rev. Biochem.* **57** 701–754

Kraut J 1977 Serine proteases: structure and mechanism of catalysis; *Annu. Rev. Biochem.* **46** 331–358

Larsen RC, Miklas PN, Druffel KL and Wyatt SD 2005 NL-3 K strain is s stable and naturally occurring interspecific recombinant derived from *Bean common mosaic necrosis virus* and *Bean common mosaic virus*. *Phytopathology* **95** 1037–1042

Lee KC and Wong SM 1998 Variability of P1 protein of *Zucchini yellow mosaic virus* for strains differentiation and phylogenetic analysis with other potyviruses. *DNA Sequence* **9** 275–293

López-Moya JJ and García JA 2008 Potyviruses; in *Encyclopedia of virology* (eds.) BWJ Mahy and MHV Van Regenmortel (Oxford: Elsevier) pp 313–322

Maiss E, Timpe U, Brisske A, Jelkmann W, Casper R, Himmler D, Mattanovich D and Katinger HWD 1989 The complete nucleotide sequence of *Plum pox virus* RNA. *J. Gen. Virol.* **70** 513–524

Mäkki-Valkama T, Valkonen JPT, Kreuze JF and Pehu E 2000a Transgenic resistance to PVY(O) associated with post-transcriptional silencing of P1 transgene is overcome by PVY(N) strains that carry highly homologous P1 sequences and recover transgene expression at infection. *Mol. Plant-Microb. Interact.* **13** 366–373

Mäkki-Valkama T, Pehu T, Santala A, Valkonen JPT, Koivu K, Lehto K and Pehu E 2000b High level of resistance to *Potato virus Y* by expressing P1 sequence in antisense orientation in transgenic potato. *Mol. Breed.* **6** 95–104

Martin DP, Williamson C and Posada D 2005 RDP2: recombination detection and analysis from sequence alignments; *Bioinformatics* **21** 260–262

Mascotti DP and Lohman TM 1990 Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids. *Proc. Natl. Acad. Sci. USA* **87** 3142–3146

Mavankal G and Rhoads RE 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

Merits A, Guo D, Järvekülg L and Saarma M 1999 Biochemical and genetic evidence of interactions between *Potato A Potyvirus*-encoded proteins P1 and P3 and proteins of the putative replication complex. *Virology* **263** 15–22

Mukherjee K, Verma Y, Chakrabarti SK and Khurana SM P 2004 Phylogenetic analysis of 5′-UTR and P1 protein of Indian common strain of *Potato virus Y* reveals its possible introduction in India; *Virus Genes* **29** 229–237

Nicklin MJH, Toyoda H, Murray MG and Wimmer E 1986 Proteolytic processing in the replication of polio and related viruses. *Nat. Biotechnol.* **4** 33–34

Oroszlan S and Luftig RB 1990 Retroviral proteinases; *Curr. Top. Microbiol.* **157** 153–185

Oruetxebarria I and Valkonen JPT 2001 Analysis of the P1 gene sequences and the 3′ terminal sequences and secondary structures of the single-stranded RNA genome of *Potato Virus V*. *Virus Genes* **23** 335–343

Oruetxebarria I, Kekarainen T, Spetz C and Valkonen JPT 2000 Molecular chracterization of *Potato virus V* genomes from Europe indicates limited spatiotemporal strain differentiation. *Phytopathology* **90** 437–444

Palmberg AC 1990 Proteolytic processing of picornaviral polyprotein. *Annu. Rev. Microbiol.* **44** 603–623

Pehu TM, Mäki-Valkama TK, Valkonen JPT, Koivu KT, Lehto KM and Pehu EP 1995 Potato plants transformed with potato virus Y P1 gene sequence are resistant to PVY^0. *Am. Potato. J.* **72** 523–532

Pruss G, Ge X, Shi XM, Carrington JC and Vance VB 1997 Plant viral synergism: The Potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell* **9** 859–868

Rajamäki ML, Kelloniemi J, Alminaite A, Kekarainen T, Rabenstein F and Valkonen JP 2005 A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions. *Virology* **342** 88–101

Rao MB, Tanksale AM, Ghatge MS and Deshpande VV 1998 Molecular and biotechnological aspects of microbial proteases. *Microbiol. Mol. Biol. Rev.* **62** 597–635

Rawlings NS and Barrett AJ 1993 Evolutionary families of the peptidases; *Biochem. J.* **290** 205–218

Record MT, Lohman TM and deHaseth PL 1976 Ions effects on ligand-nucleic acid interactions. *J. Mol. Biol.* **107** 145–158

Restrepo-Hartwig MA and Carrington JC 1994 The Tobacco etch potyvirus 6-kilodalton protein is a membrane associated and involved in viral replication. *J. Virol.* **68** 2388–2397

Rice CM and Strauss JH 1990 Production of flavivirus polypeptides by proteolytic processing; *Semin. Virol.* **1** 357–367

Riechmann JL, Laín S and García JA 1990 Infectious *in vitro* transcripts from a plum pox potyvirus cDNA clone. *Virology* **177** 710–716

Riechmann JL, Laín S, García JA 1992 Highlights and prospects of potyvirus molecular biology. *J. Gen. Virol.* **73** 1–16

Rodriguez-Cezero E and Shaw JG 1991 Two newly-detected non-structural viral proteins in poty virus infected cells. *Virology* **185** 572–579

Rodríguez-Cerezo E, Ammar ED, Pirone TP and Shaw JG 1993 Association of the non-structural P3 viral protein with cylindrical inclusions in potyvirus-infected cells. *J. Gen. Virol.* **74** 1845–1949

Rodríguez-Cerezo E, Findlay K, Shaw JG, Lomonossoff GP, Qiu SG, Linstead P, Shanks M and Risco C 1997 The coat and cylindrical inclusion proteins of a potyvirus are associated with connections between plant cells. *Virology* **236** 296–306

Roossinck MJ 2003 Plant RNA virus evolution. *Curr. Opin. Microbiol.* **6** 406–409

Salvador B, Saénz P, Yangüez E, Quiot JB, Quiot L, Delgadillo MO, García JA and Simón-Mateo C 2008 Host-specific effect of P1 exchange between two potyviruses. *Mol. Plant Pathol.* **9** 147–155

Shi Y, Chen J, Hong X, Chen J and Adams MJ 2007 A potyvirus P1 protein interacts with the Rieske Fe/S protein of its host. *Mol. Plant Pathol.* **8** 785–790

Skalka AM 1989 Retroviral proteases: first glimpses at the anatomy of a processing machine. *Cell* **56** 911–913

Soumounou Y and Laliberté JF 1994 Nucleic acid-binding properties of the P1 protein of *Turnip mosaic potyvirus* produced in *Escherichia coli*. *J. Gen. Virol.* **75** 2567–2573

Sprang S, Standing T, Fletterick RJ, Stroud RM, Finer-Moore J, Xoung N, Hamlin R, Rutter WJ, and Craik CS 1987 The three-dimensional structure of Asn^{102} mutant of trypsin: role of Asp^{102} in serine protease catalysis. *Science* **237** 905–909

Strauss JH and Strass EG 1990 Alphavirus proteases. *Semin. Virol.* **1** 347–356

Susaimuthu J, Tzanetakis IE, Gergerich RC and Martin RR 2008 A member of a new genus in the Potyviridae infects Rubus. *Virus Res.* **131** 145–151

Tan Z, Wada Y, Chen J and Ohshima K 2004 Inter-and intralineage recombinants are common in natural population of *Turnip mosaic virus*. *J. Gen. Virol.* **85** 2683–2696

Tatineni S, Ziems AD, Wegulo SN and French R 2009 *Triticum mosaic virus*: a distinct member of a family potyviridae with an unusually long leader sequence. *Phytoplathology* **99** 943–950

Tavert-Roudet G, Ravelonandro M, Bachelier JC and Dunez J 1998 Transgenic *Nicotiana benthamiana* plants containing the P1 gene of *Plum pox virus* are resistant to virus challenge. *Eur. J. Plant Pathol.* **104** 103–107

Tordo VMJ, Chachulska AM, Fakhfakh H, Romancer ML, Robaglia C and Astier-Manifacier S 1995 Sequence polymorphism in the 5′NTR and n the P1 coding region of *Potato virus Y* genomic RNA. *J. Gen. Virol.* **76** 939–949

Valli A, Martin-Hernández AM, López-Moya JJ and García JA 2006 RNA silencing suppression by second copy of the P1 serine protease of *Cucumber vein yellow ipomovirus* (CVYV), a member of the family Potyviridae that lacks the cysteine protease HC-Pro. *J. Virol.* **80** 10055–10063

Valli A, López-Moya JJ and García JA 2007 Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. *J. Gen. Virol.* **88** 1016–1028

Valli A, Dujovny G and García JA 2008 Protease activity, self interaction and small interfering RNA binding of the silencing suppressor P1b from *Cucumber vein yellowing ipomovirus*. *J. Virol.* **82** 974–986

Vance VB, Moore D, Turpen TH, Bracker A and Hollowell VC 1992 The complete nucleotide sequence of pepper mottle virus genomic RNA: comparison of the encoded polyprotein with those of other sequences potyviruses. *Virology* **191** 19–30

Vance VB, Berger PH, Carrington JC, Hunt AG and Shi XM 1995 5′ proximal potyviral sequences mediate *Potato virus X* potyviral synergistic disease in transgenic tobacco. *Virology* **206** 583–590

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

Verchot JM and Carrington JC 1995b Evidence that the Potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification. *J. Virol.* **69** 3668–3674

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

Verchot JM, Herndon KL and Carrington JC 1992 Mutational analysis of the Tobacco Etch Potyviral 35-kDa proteinase: Identification of essential residues and requirements for autoproteolysis. *Virology* **190** 298–306

Ward CW and Shukla DD 1991 Taxonomy of potyviruses: current problems and some solutions. *Intervirology* **32** 269–296

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

Whitham S, Quan S, Chang HS, Cooper B, Estes B, Zhu T, Wang X and Hou YM 2003 Diverse RNA viruses elicit the expression of common rests of genes in susceptible *Arabidopsis thaliana* plants. *Plant J.* **33** 271–283

Wisler GC, Porciful DE and Hiebert E 1995 Characterization of the P1 protein and coding region of the *Zucchini yellow mosaic virus*. *J. Gen. Virol.* **76** 37–45

Xu DL, Zhou DH, Xie YJ, Mock R and Li R 2010 Complete nucleotide sequence and taxonomy of *Sugarcane streak mosaic virus*, member of a novel genus in the family Potyviridae. *Virus Genes* **40** 432–439

MS received 28 September 2010; accepted 10 Febuary 2011

*e*Publication: 14 March 2011

Corresponding editor: INDRANIL DASGUPTA