Mechanistic divergence between P1 proteases of the family Potyviridae

Bernardo Rodamilans, Adrián Valli† and Juan Antonio García

P1a and P1b are two serine proteases of Cucumber vein yellowing virus (an ipomovirus). They belong to the group of P1 factors present at the N terminus of the polyproteins of most members of the family Potyviridae. The present work compares the protease activities of P1a and P1b in different experimental systems. The findings made regarding how these two proteases work, such as the requirement for a host factor by P1a but not by P1b, underscore important differences in their catalytic activity that point towards their undergoing divergent evolution involving the acquisition of mechanistic variations. The expression of several truncated forms of P1b in bacteria and in planta helped define the protease domain of P1b, along with other important features such as its apparently in cis mode of action. Recent phylogenetic data, together with the present results, allow an appealing hypothesis to be proposed regarding P1 evolution and its involvement in potyvirid speciation.

INTRODUCTION

Cucumber vein yellowing virus (CVYV) is an ipomovirus that infects cucurbit plants, causing economic losses in Asia, Africa and Europe. Infected plants may remain symptomless or present a variety of external features ranging from vein clearing and chlorosis to general necrosis and death (OEPP/EPPO, 2005). CVYV belongs to the family Potyviridae, whose members show a conserved organization of the genome and of their structural and non-structural proteins (Shukla et al., 1998). The N-terminal part of the polyprotein of CVYV, however, is quite different to that of other members of Potyviridae in that it lacks a coding sequence for HCPro, an RNA-silencing suppressor (RSS) with cysteine protease activity (Janssen et al., 2005). Rather, it has an extra copy of serine protease P1, making a P1a–P1b tandem at the N terminus of the polyprotein (Valli et al., 2006). The second cistron (P1b), which may have appeared through gene duplication or interspecific RNA recombination (Valli et al., 2007), is also present in Squash vein yellowing virus (another ipomovirus) (Li et al., 2008). Sequence comparisons have shown that P1a in both these two viruses is similar to the typical P1 of members of the genus Potyvirus, whereas P1b aligns with the P1 proteins of other ipomoviruses, as well as with those belonging to tritimoviruses and poaceviruses (Abraham et al., 2012; Giner et al., 2010). P1b has been described as an RSS that specifically sequesters 21 nt small RNAs (Valli et al., 2011), and this function is independent of its protease activity (Valli et al., 2008). It therefore at least partially assumes the role of HCPro, the usual RSS of potyviruses, in CVYV. It has also been shown that P1b can replace HCPro of the chimerical potyvirus Plum pox virus (PPV), investing it with full infective functionality (Carbonell et al., 2011).

HCPro is a cysteine protease with functions including viral defence, genome amplification, movement and aphid transmission (Anandalakshmi et al., 1998; Berger et al., 1989; Brigneti et al., 1998; Cronin et al., 1995; Kasschau & Carrington, 1998; Kasschau et al., 1997; Rojas et al., 1997). It works in cis to release its C-end from the potyviral polyprotein. In Turnip mosaic virus (TuMV), the structure of its protease domain has been solved by X-ray crystallography, providing an important insight into its unique folding and catalysis mechanism (Guo et al., 2011). The P1 protein of potyviruses, however, is a serine protease with no clearly defined function besides its proteolytic activity (Rohožková & Navrátil, 2011), although there is evidence that it might be involved in host range definition, genome amplification and RNA-silencing reinforcement, among other activities (Salvador et al., 2008; Verchot & Carrington, 1995). Interestingly, P1 needs a factor from the plant to release itself from HCPro; no such factor would appear to be required for HCPro protease activity (Carrington et al., 1990; Thornbury et al., 1993; Verchot et al., 1992).

The two P1 proteins of CVYV, P1a (525 aa) and P1b (318 aa), show proteolytic activity in plants and bear the catalytic triad formed by histidine, aspartic and serine that is characteristic of the serine protease family (Rohožková &
Navrátil, 2011; Valli et al., 2007). No structural or mechanistic information exists regarding these two proteins, and the functional importance behind their sequence similarity to other potyvirid P1 proteins remains unclear. The aim of the present work was to shed light on their enzyme activities and on the divergent evolutionary paths taken by these seemingly duplicated proteins.

**RESULTS**

**CVYV P1b, but not P1a, conserves its protease activity in a bacterial system**

The ability of P1a and P1b to cut themselves free at their C-ends in bacteria was examined first. The idea of using a simple prokaryotic system was to facilitate further analyses should both proteins be shown to retain their protease activity. Viral-derived constructs were expressed from an inducible T7 promoter in pET21 vectors (Fig. 1a). As well as using P1b alone and the P1aP1b fusion construct, a construct including an inactive mutant of P1b, in which the catalytic serine was replaced by alanine (P1bS264A) (Valli et al., 2008), was prepared as an inactive protease control.

No antibody was available against P1a, so a previously developed anti-P1b polyclonal antibody (Valli et al., 2008), which also reacts with the tandem affinity purification (TAP) tag (Puig et al., 2001; Rigaut et al., 1999) fused to P1b and P1bS264A, was used in Western blot detections (Fig. 1b). No band corresponding to the size of P1b was seen for the P1aP1b fusion construct. Rather, a single band appeared at a position that fitted with the expected size of the entire P1aP1b construct (Fig. 1b, lane 4). This indicates that P1a is unable to release itself from P1b when produced in bacteria as a recombinant fusion protein. These analyses were performed using *E. coli* strain ArcticExpressBL21 (DE3), but strains TunerBL21(DE3) and BL21(DE3) pLysS returned identical results (data not shown). The inability of P1a to function in bacteria might be due to improper folding, or to the fact that P1a needs a factor from a plant that is not provided in this prokaryotic environment, as is the case for P1 proteins from potyviruses.

In contrast, P1b appears to be functional in bacteria, maintaining high proteolytic activity. Hence, when P1bTAP was expressed, a clear band corresponding to a protein of the expected size for P1b appeared in the Western blot (Fig. 1b, lane 3). A higher mobility band corresponding to the free TAP tag was also seen. A faint band was also noticed that seemed to correspond to the uncut P1bTAP protein, since it matched the electrophoretic mobility of the band detected when the TAP-tagged protease mutant P1bS264A was expressed (Fig. 1b, lane 2).

**P1a and P1b protease activities differ in their dependence on a component of wheatgerm extract that is not supplied by a rabbit reticulocyte lysate**

To test the hypothesis that P1a might need a plant factor to undertake its protease activity, *in vitro* translation systems were used, simulating the experiments performed by Verchot et al. (1992) and García et al. (1993) with the P1
and HCPro proteases of the potyviruses *Tobacco etch virus* (TEV) and PPV.

Although the pattern of proteins generated by the PPV-derived P1HCP3\* construct in the wheatgerm extract (WGE) system was intricate, it was clear that a free form of P1 was produced in this system that was not produced in the rabbit reticulocyte lysate (RRL) system (Fig. 2b, lanes 1 and 7). This agrees with previous reports of the proteolytic activity of potyviral P1 proteases (García et al., 1993; Verchot et al., 1992). Similarly, free P1a is a major product of *in vitro* translation of P1aP1bP3\* in WGE, while a band with the expected mobility of the unprocessed P1aP1b was the main translation product of this construct in RRL (Fig. 2b, lanes 2 and 8). This indicates that the protease activity of ipomoviral P1a, like that of potyviral P1, relies on a component of the WGE that cannot be supplied by the RRL system. The accumulation of products with the mobility expected for P1b and P1aP1b after *in vitro* translation of P1aP1bP3\* in WGE and RRL, respectively (Fig. 2b, lanes 2 and 8), suggests that P1b protease is active not only in WGE but also in the RRL system. This was confirmed by the fact that free P1b was the main product of *in vitro* synthesis when P1bTAP was subjected to *in vitro* translation in both systems (Fig. 2b, lanes 3 and 9). As expected, a band corresponding to the unprocessed product was observed when the mutant P1bS264A fused to TAP was synthesized (Fig. 2b, lanes 4 and 10). Together, these results support the idea that P1b protease activity, in contrast with that of P1a, does not rely on a component supplied by the WGE system.

**P1b activity mainly takes place intramolecularly**

Experiments using *E. coli* strains with leaky expression, such as ArcticExpressBL21(DE3), showed that the expression of P1b fused to the TAP tag produced a protein with the expected size of the unprocessed P1bTAP at the time of induction (Fig. 3a, lane 2). The presence of this fusion protein, however, was drastically reduced after overnight incubation (Figs 1b, lane 3; 3a lane 5; and 4b, lane 2). This might indicate that the protease activity of P1b does not occur co-translationally in bacteria, but later. Interestingly, a band corresponding to the uncut product of P1bTAP was also observed after 1 h when this protein was translated in WGE (Fig. 2b, lane 3), but not after 6 days when the same construct was expressed in plants (Fig. 4c, lane 2).

To determine whether the protease activity of P1b works only in *cis* to release itself from the rest of the polyprotein, or whether it also works in *trans* thus opening the field of action for P1b, a protease inactive mutant (P1bS264A) was fused to the TAP tag and co-expressed with WT P1b by agroinfiltration in *Nicotiana benthamiana* plants. If P1b worked in *trans*, the release of the TAP from the TAP-tagged P1bS264A would be expected. However, no TAP band for the co-agroinfiltrated plants was detectable by Western blotting (Fig. 3b, lane 4). This suggests that P1b protease activity does not function in *trans*, although it is possible that a substrate other than itself is required for this mode of action.

---

**Fig. 2.** *In vitro* translation in WGE and RRL. (a) Diagram of the translated constructs. The numbers on top of the boxes indicate the total size (in kDa) of the expressed protein as well as the size of the smaller fragment that might be released upon proteolysis. The asterisk after P3 indicates that a part of viral 6k1 is being translated and is attached. (b) *In vitro* synthesis and processing of proteins. Two different extracts were used: WGE (lanes 1–6) and RRL (lanes 7–12). Luciferase was used as a positive translation control (lanes 5 and 11); water was used as negative control (lanes 6 and 12). Individual proteins (in roman text) and unprocessed products (in italics) are marked with solid and dotted arrows respectively on the left side of the autoradiograph. The molecular mass ladder is marked on the right side.
The protease domain of P1b maps to its C-terminal 118 amino acids

Although the overall similarity between CVYV P1b (318 amino acids) and the typical P1 proteins of potyviruses is very low, their C-terminal regions are quite conserved, showing a shared evolutionary origin (Valli et al., 2007). Deletion analysis of the TEV P1 mapped the border of the domain essential for protease activity between lysine 158 and glutamic acid 188 (Verchot et al., 1991). Glutamic acid 188 belongs to the VELI motif, which is quite conserved in the P1 proteins of potyviruses (Valli et al., 2007), and aligns at around phenylalanine 193 to asparagine 198 of CVYV P1b (Fig. 4a). Three constructs were thus prepared to help define the actual protease domain of P1b (Fig. 4a). These were tested both in E. coli and in N. benthamiana plants. Western blot analysis with an antibody against P1b showed that even with the largest deletion, which reached valine 200, the trimmed P1b retained its proteolytic activity (Fig. 4b, lanes 4–6, and 4c, lanes 4 and 5). It is interesting to note that, in bacteria, although all deletion mutants released the TAP tag from their C-end, they did it less effectively than the WT protein (Fig. 4b, compare lane 2 with lanes 4–6). In the plants, no uncut TAP-tagged Δ176P1b was detected, but the cleavage of the Δ200P1bTAP fusion was only partial (Fig. 4c, lanes 4 and 5). These results indicate that, although only the C-terminal 118 amino acids are essential for the protease activity of P1b, upstream sequences also contribute to the efficiency of self-cleavage.

DISCUSSION

A common trait in the Potyviridae family is the presence of a P1 protein at the N terminus of the viral polyprotein, followed by HCPro. Both P1 and HCPro are proteases that release themselves from the rest of the polyprotein by autocatalysis. This standard design, which has been much studied in the genus Potyvirus, is also seen in other genera of the Potyviridae family, such as Rymovirus, Brambyvirus, Poacevirus and Tritimovirus (Adams et al., 2012). However, notable deviations from this pattern exist. The members of the Bymovirus genus (the only genus in the family with a bipartite genome) lack P1 and HCPro; these are replaced by two proteins of unclear function in a second RNA molecule (You & Shirako, 2010). More striking is the lack of a P1 protein in the genus Maclurovirus (Kondo & Fujita, 2012). The genus Ipomovirus includes members that conserve the binominal P1-HCPro (Abraham et al., 2012; Giner et al., 2010), as well as viruses that lack HCPro but which have either one (Mbanzibwa et al., 2009a, b) or two (P1a and P1b) P1-like proteases (Li et al., 2008; Valli et al., 2006).

Alignment of the conserved protease domain of different P1 proteins of the family Potyviridae (Fig. S1, available in
JGV Online) provided a phylogenetic tree with two main branches, in agreement with information in the literature (Valli et al., 2007) (Fig. 5). One branch places the P1 proteins of potyviruses and rymoviruses with the P1a proteins of ipomoviruses. The second branch clusters P1 proteins of tritimoviruses and poaceviruses along with the P1b proteins and single P1 proteins of ipomoviruses. A functional importance is inferred from the fact that, despite their wide divergence, all members of the P1-P1a branch share high isoelectric point values, a feature not conserved in proteins of the P1-P1b group (Valli et al., 2007; Fig. S1).

Even more important, whereas P1-P1a proteins are followed by an RSS in every species, proteins from the P1-P1b branch suppress RNA silencing themselves, irrespective of the accompanying proteins (Giner et al., 2010; Mbanzibwa et al., 2009a; Tatineni et al., 2012; Valli et al., 2006; Young et al., 2012).

The functional divergence envisaged for the two main groups of P1 proteins of the family Potyviridae might involve mechanistic differences. The present work with CVYV P1a and P1b examines this possibility, and shows the divergent paths that these proteins have taken from their predicted common origin. The results obtained in E. coli show the drastic differences between P1a and P1b, with P1b being active and P1a inactive in this heterologous expression system (Fig. 1). These differences were confirmed when moving to eukaryotic in vitro translation systems, in which P1a showed a mode of action resembling that described for potyviral P1 proteins that rely on a component(s) of WGE, and that is (are) not available in RRL, while P1b acts in a manner independent of such host factor(s) (Fig. 2).

The expression of P1b in E. coli provides another clue regarding its protease activity: this protein appears not to act immediately after its translation (Fig. 3a). This contrasts with the behaviour of potyviral P1, which has been suggested to act co-translationally (Verchot et al., 1992). In addition, the present data suggest that P1b only functions in cis (Fig. 3b), similar to the activity described for the potyviral P1 and HCPro proteins (Carrington et al., 1989; Verchot et al., 1992). In the structure solved by Guo et al. (2011) for the
HCPro protease domain of TuMV, the final amino acid of the protein is embedded in the catalytic pocket of the cysteine protease triad. These authors postulate this to be the reason why HCPro acts only in cis. Although the molecular structure of P1b remains unknown, the present results suggest that it might behave like HCPro in this regard, sharing with this protein, and P1, the inability to act in trans.

Despite the low-level similarity among P1 proteins, specific motifs inside their protease domains are well conserved (Fig. S1). Deletion mutants of P1b were therefore designed using the presumed histidine residue of the catalytic active site and the upstream VELI motif, which is conserved in the group of P1-P1a proteins (Valli et al., 2007; Fig. 4a), as references. Even though a proper VELI motif was absent from P1b (Fig. 4a), constructs in the corresponding area showed differences in protease activity in planta (Fig. 4c), signalling this region to be the proteolysis limiting point (as seen in other P1 proteins). Moreover, these experiments suggest that protease activity is not an all-or-nothing process, but a reaction that can be modulated by amino acids surrounding the canonical protease motifs.

Together, these data suggest that P1a and P1b are proteins with mechanistic specializations derived from a common ancestor protein. These specializations affect their protease activity and may be important in the evolutionary history of potyviruses. Interestingly, neither P1 nor HCPro seem to be indispensable for viral infection by Potyviridae family members. The genomes of macluraviruses (Kondo & Fujita, 2012) and bymoviruses (You & Shirako, 2010) suggest them to have a common origin with an ancestral HCPro but no P1. Following this line of thought, the emergence of the P1 precursor of the present P1-P1a group, which has been shown to be involved in adaptation to the host (Valli et al., 2007), together with the imprinting of RNA-silencing suppression activity in HCPro, might be responsible for the extraordinary radiation of potyviruses (of which there are almost 150 different species) (Adams et al., 2012; Gibbs & Ohshima, 2010; Gibbs et al., 2008). Another main potyvirus lineage ending in poaceviruses, tritimoviruses and ipomoviruses, probably derived from the appearance of a P1-P1b precursor with RSS activity. Although this scenario is rather speculative, the importance of P1 as a player in speciation seems unquestionable. Taking into account the crucial role that P1 protease activity plays in viral infection, further study should aim to clarify how these important proteolytic events take place.

**METHODS**

**Plasmids.** All constructs prepared for expression in E. coli cells were cloned into plasmid pET21d (Novagen). The only exception was the P1aP1b construct, which was cloned into pET21a (Novagen). pGGSN (Garcia et al., 1993) was used for the expression of
P1-HCPro-P3* and as a backbone for the rest of the constructs to be used in in vitro translation. Gateway technology (Invitrogen) was used to build plasmids expressing the desired constructs for use in plants. pDONR-207 (Invitrogen) was used as the donor vector and pmDC32 (provided by Mark Curtis, University of Zurich) (Curtis & Grossniklaus, 2003) as the destination vector. BP clonase reactions were performed to introduce PCR fragments into the entry vector, followed by LR clonase reactions, adhering to the enzyme manufacturer’s instructions (Invitrogen). Viral sequences, with restriction sites or Gateway recombination sequences to facilitate cloning, were amplified by PCR using Phusion DNA polymerase (Fermentas). The primers used for PCR amplification are listed in Table S2. E. coli DH5α cells were used for the cloning and selection of the appropriate plasmids.

**Protein expression in bacteria.** Plasmids were transformed into ArcticExpressBL21(DE3) (Stratagene), BL21(DE3)pLysS (Promega) and TunerBL21(DE3) (EMD Millipore) E. coli strains by heat shock. Transformed cells were grown in LB media with the appropriate antibiotics at 37 °C overnight with strong agitation. These pre-cultures were used the day after to grow fresh cultures until reaching an OD600 of 0.4–0.6. Induction was begun by the addition of 50 μM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Cultures were then maintained at 20 °C overnight. Samples of the cultures were taken before and after induction for Western blot analysis.

**Protein expression by in vitro translation in WGE and RRL.** mRNA was prepared by in vitro transcription using the T7-Scribe Standard RNA IVT kit (Cellscript) following the manufacturer’s instructions. Luciferase mRNA was provided with the translation kit. mRNAs were translated in WGE and RRL (Promega) and the products analysed by SDS-PAGE according to Promega’s instructions. A 35S methionine–cysteine protein labelling mix (Perkin Elmer) was included in the reaction mixtures to label the synthesized proteins.

**Protein expression in planta.** pmDC32-plasmid plasmids were transformed by electroporation into Agrobacterium tumefaciens strain C58C1. *N. benthamiana* plants were infiltrated with the agrobacterium carrying the appropriate plasmid as previously described (Valli et al., 2006). Six days after agroinfiltration, tissue was collected, crushed in liquid nitrogen and stored at −80 °C until Western blot analysis.

**Electrophoresis and Western blot analysis.** The preparation of protein samples from *N. benthamiana* plants, SDS-PAGE and electroblotting were performed as previously described (Valli et al., 2006). Samples from bacteria were prepared from 200 μl and 50 μl of non-induced and induced cultures, respectively. Samples were centrifuged at 18,000 g for 2 min, and the pellets resuspended in a mixture of 40 μl of 10 mM Tris/HC1 pH 8, 200 mM NaCl and 10 μl of loading buffer 5X (200 mM Tris/HC1 pH 8.8, 10 % glycerol, 5 mM EDTA [EDTA], 4 % SDS, 7 % β-mercaptoethanol, 50 μg ml⁻¹ bromophenol blue), and heated at 95 °C for 5 min prior to undergoing SDS-PAGE and electroblotting. Pib was detected using anti-CVYV Pib polyclonal serum and horseradish peroxidase-conjugated goat anti-rabbit IgG (Jackson). The immunostained proteins were visualized by enhanced chemiluminescence detection using a LifeABlot kit (Euroclone) according to the manufacturer’s instructions. Ponceau red staining was used to check the total protein content of the samples.

**Phylogenetic analysis.** Sequence alignments were performed initially by CLUSTAL W (Goujon et al., 2010; Larkin et al., 2007), and refined by manual editing using Jalview software (Waterhouse et al., 2009). A phylogenetic tree was built using Genebee software following the provided cluster algorithm. One hundred replicas were used to obtain the percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (Felsenstein, 1985). The Protein computer program (part of the DNASTAR Lasergene® software suite) was used for isoelectric point (pl) calculations.

**ACKNOWLEDGEMENTS**

We are grateful to Beatriz García for technical assistance. This work was funded by the Spanish MCI (grant BIO2010-18541) and the European Union (grant KBBE-204429).

**REFERENCES**


binding through conserved WG/GW motifs. PLoS Pathog 6, e1000996.


Valli, A., Oliveros, J. C., Molnar, A., Baulcombe, D. & García, J. A. (2011). The specific binding to 21-nt double-stranded RNAs is crucial for the anti-silencing activity of Cucumber vein yellowing virus P1b and perturbs endogenous small RNA populations. RNA 17, 1148–1158.


