Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification of a Cardosin B Homolog in Physcomitrium patens and preliminary evidence of Milk-Clotting Activity in overexpressing lines

Agustín Baroffio
Universidad de la República
Andrés Vivas Panario
Universidad de la República
Carolina Villadóniga
Universidad de la República
Leticia Perez Diaz
Universidad de la República
Cecilia Ruibal
cruibal@fcien.edu.uy

Universidad de la República

Research Article

Keywords: Physcomitrium patens, cardosin, aspartic endopeptidase, milk-clotting

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11651835/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Plant aspartic proteases are promising alternatives to animal-derived milk-clotting enzymes, although their potential in bryophytes remains largely unexplored. In this study, we investigated endogenous homologs of Cardosin B from *Cynara cardunculus* in the moss *Physcomitrium patens* and evaluated their potential milk-clotting activity. A BLASTp search against the *P. patens* proteome identified 16 candidate proteins, of which three exhibited high sequence similarity to Cardosin B (55–56%) and conserved structural features, including two catalytic aspartate residues and saposin-like domains. The closest homolog, *PpAP1* (PpV7_05g0093900-5), was selected for functional characterization. Stable transgenic *P. patens* lines overexpressing *PpAP1* were generated, and transgene integration and increased transcript accumulation were confirmed by genomic PCR and quantitative RT-PCR, respectively. Milk-clotting assays revealed coagulation exclusively in protein extracts from PpAP1-overexpressing plants, whereas wild-type extracts showed no detectable activity. Notably, coagulation was completely abolished in the presence of aspartic protease inhibitors, supporting the involvement of aspartic protease activity. These findings provide functional evidence linking PpAP1 overexpression to milk-clotting activity and identify this moss-derived protease as a promising candidate for further biochemical characterization and potential applications in dairy biotechnology.

Introduction

Plant-based expression systems have emerged as attractive platforms for the production of proteins of biotechnological interest, offering potential advantages in production costs, biosafety, and scalability. Among these systems, the moss *Physcomitrium patens* (formerly *Physcomitrella patens*) has gained increasing attention as a versatile platform for molecular farming. Its fully sequenced genome, efficient homologous recombination machinery, and well-established genetic transformation methods enable precise genetic manipulation and the development of stable engineered lines (1, 2). Furthermore, *P. patens* can be cultivated under controlled conditions in liquid suspension cultures and photobioreactors, with successful scale-up to 500-L systems (3, 4). These characteristics, together with the reduced risk of contamination by animal-derived pathogens and the possibility of controlled cultivation, make *P. patens* a promising platform for the production of proteins with pharmaceutical and industrial applications (3, 5).

Among industrially relevant proteins, proteolytic enzymes play an important role in food processing, particularly in cheese manufacturing, where milk-clotting enzymes are essential for curd formation. Traditionally, animal-derived rennet has been widely used for this purpose; however, plant-derived proteases represent attractive alternatives. Of particular interest are aspartic proteases (APs), a family of enzymes characterized by two catalytic aspartate residues and proteolytic activity generally favored by acidic conditions (6). The flowers of *Cynara cardunculus* L. (cardoon) contain several APs, including cardosins A and B, which have long been exploited as milk-clotting enzymes in traditional cheese production (7, 8). Both cardosins exhibit proteolytic activity toward milk caseins, although they differ in substrate specificity, potentially influencing the characteristics of the resulting dairy products (9).

Despite their industrial relevance, the use of cardoon extract derived enzymes presents limitations associated with variability in enzyme composition and activity among plant sources, which may affect the reproducibility of cheese production. Consequently, alternative strategies for obtaining milk-clotting proteases under controlled conditions have attracted considerable interest (7). In this context, plant-based production systems have shown promising results, including the production of biologically active cardosin B in transgenic tobacco BY-2 cell cultures (10). These advances highlight the potential of plant biotechnology to provide alternative sources of milk-clotting enzymes. However, beyond the heterologous production of characterized enzymes, the identification and enhanced expression of endogenous plant proteases with desirable catalytic properties may represent an additional strategy for obtaining enzymes of industrial interest.

In this study, we explored *P. patens* as a source of milk-clotting enzymes by searching for endogenous homologs of cardosin B from *C. cardunculus*. Sequence-based analysis identified a putative phytepsin belonging to the aspartic protease family, which we selected for functional investigation. We generated *P. patens* lines overexpressing this enzyme and evaluated the milk-clotting activity of the resulting protein extracts. Our results demonstrate milk-clotting activity in extracts from phytepsin-overexpressing lines, providing preliminary evidence supporting the potential of this moss-derived protease for biotechnological applications in dairy processing.

Materials and methods

Plant Material and Growth Conditions

P. patens (Gransden strain) wild-type plants were used in all experiments (11). Plants were grown and maintained axenically on cellophane-overlaid BCDAT medium, as described by Ashton and Cove (1977) (12). Protonemal cultures were generated by homogenizing moss colonies using an Ultra-Turrax T25 basic homogenizer (IKA, Stauffen, Germany) in sterile Milli-Q water. The resulting suspension was plated onto cellophane-overlaid BCDAT medium in Petri dishes. For micropropagation, plant fragments from protonemal cultures or established colonies were transferred to fresh agar medium. Plants were grown at 22°C under a 16-h light/8-h dark photoperiod, with a photosynthetic photon flux density of 50 µmol m^{-2} s^{-1}. Three-week-old colonies were used for experiments.

Identification and in silico characterization of Cardosin B homologs

The amino acid sequence of Cardosin B from *C. cardunculus* (UniProt accession Q9XFX4) was retrieved and used as a query for a BLASTp search against the *P. patens* resource (version 7.1) available in the Phytozome database (13). Candidate homologs were selected based on an amino acid sequence identity threshold of greater than 50%, and redundant sequences corresponding to alternative isoforms of the same gene were removed.

The selected protein sequences were aligned with Cardosin B using ClustalW (14), and multiple sequence alignments were visualized using ESPript 3.2 (15). Protein domain architecture and conserved

catalytic motifs were analyzed using InterProScan (16), focusing on aspartic peptidase A1 domains (PROSITE PS51767), aspartic protease active-site motifs (PROSITE PS00141), and associated Pfam domains, including the eukaryotic aspartyl protease domain (PF00026) and saposin-like regions (PF03489 and PF05184). Predicted signal peptides were also examined.

Based on sequence similarity to Cardosin B, conservation of catalytic aspartate residues, and protein domain organization, *PpAP1* (PpV7_05g0093900-5) was selected for overexpression in *P. patens* and subsequent evaluation of its milk-clotting activity.

Structural Prediction and Comparative Analysis of PpAP1 and Cardosin B

The amino acid sequence of *P. patens* aspartic protease PpAP1 (PpV7_05g0093900-5), retrieved from Phytozome v7.1, was used for three-dimensional structure prediction using ColabFold (17), an AlphaFold2-based protein structure prediction platform. The highest-ranked predicted model (rank 001) was selected for subsequent structural analyses.

The experimentally determined crystal structure of cardosin B (PDB ID: 5NFG) was used as a reference for structural comparison. The predicted PpAP1 model and the cardosin B crystal structure were visualized and superimposed using the Matchmaker algorithm implemented in UCSF ChimeraX (18). Structural similarity was assessed by calculating the root-mean-square deviation (RMSD) between corresponding atomic positions following structural alignment, considering both the atom pairs retained after structural pruning and all initially matched pairs.

The conserved catalytic aspartate residues of PpAP1 (Asp98 and Asp285) were identified through sequence analysis and compared with the corresponding residues of cardosin B (Asp80 and Asp267 in chain B of PDB 5NFG). Their relative spatial positions were examined following structural superposition.

Molecular representations were prepared using UCSF ChimeraX, with PpAP1 and cardosin B displayed as blue and gray ribbons, respectively. Catalytic aspartate residues were highlighted in red (PpAP1) and green (cardosin B).

Construct Design

The construct for PpAP1 (PpV7_05g0093900) overexpression was generated using the destination vector pTHUbi (19), containing the hygromycin resistance gene under the control of the ubiquitin promoter and the *nos* terminator. A 1,687-bp cDNA fragment encompassing the PpAP1 coding region (bases − 112 to 1575 relative to the start codon) was amplified using the forward primer 5'-ttgggtaccggtttgcgttcaactgcctt-3' and reverse primer 5'-gaactcgagggcggtatttcgtgttgagc-3', containing *KpnI* and *XhoI* restriction sites, respectively. The resulting PCR fragment was cloned into the pENTR2B

entry vector and subsequently transferred into the pTHUbi destination vector by Gateway LR recombination using LR Clonase (Invitrogen).

Generation and Molecular Characterization of PpAP1 overexpressed lines in P. patens

Protoplast preparation and transformation were performed according to Schaefer et al. (1991) (11). Thirty micrograms of *SwaI*-linearized plasmid DNA were used for polyethylene glycol (PEG)-mediated protoplast transformation. After two rounds of selection alternating with growth under nonselective conditions, resistant colonies were selected for molecular characterization. To assess transgene expression, total RNA was reverse-transcribed into cDNA, followed by PCR amplification using two primer pairs: forward primer 5'-ttgggtaccggtttgcgttcaactgcctt-3' and reverse primer 5'-gaactcgagggcggtatttcgtgttgagc-3', designed to amplify a 1,687-bp fragment encompassing the *PpAP1* coding region and upstream sequence; and forward primer 5'-gagcctgacctattgcatctc-3' and reverse primer 5'-gatgttggcgacctcgtatt-3', designed to amplify a 442-bp fragment of the hygromycin resistance gene. PpAP1 overexpression was assessed by quantitative reverse transcription PCR (qRT-PCR) using RNA isolated from wild-type and putative overexpressing lines.

For qRT-PCR Total RNA was isolated from *P. patens* tissues using the RNeasy Plant Mini Kit (Qiagen, Germany), according to the manufacturer's instructions. For cDNA synthesis, 2 µg of total RNA was reverse-transcribed using RevertAid Reverse Transcriptase (Thermo Scientific).

PpAP1 transcript abundance was quantified by qRT-PCR using gene-specific primers designed with Primer3Plus (20) on a DT Prime4 real-time PCR system (DNA-Technology).

For *PpAP1*, the forward and reverse primers were 5'-aatcaggcaagtcgagcaca-3' and 5'-cagcaaacacctggcctttc-3', respectively. For the reference gene (PpV7_10g0085200), the forward and reverse primers were 5'-accgagtccaacattctacc-3' and 5'-gtccacattagattctcgca-3', respectively.

Amplification efficiencies exceeded 90% for all primer pairs. Each 10-µL reaction contained 5 µL of 2× SYBR Green PCR Master Mix (Qiagen), a primer mixture at a final concentration of 0.5 µM, and 1 µL of fourfold-diluted cDNA.

The thermal cycling conditions consisted of an initial denaturation at 95°C for 2 min, followed by 40 cycles of denaturation at 95°C for 5 s and annealing/extension at 60°C for 10 s.

PpAP1 transcript levels were normalized to those of the reference gene PpV7_10g0085200. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method (21). Each data point represented the mean of three biological replicates, each analyzed in three technical replicates.

Milk-cloatting activity

Plant tissue was macerated in liquid nitrogen using a mortar and pestle and then homogenized in 100 mM Tris–HCl buffer (pH 7.0) at a 1:1 (w/v) tissue-to-buffer ratio. The homogenates were centrifuged at 16,000 × g for 15 min at 4°C, and the supernatants were collected. Spent culture medium was separately collected and centrifuged under the same conditions to remove cellular debris. Total soluble protein concentration was determined using the Bradford assay (Bio-Rad, USA), with bovine serum albumin (BSA) as the standard, according to the manufacturer's instructions.

Milk-clotting activity was determined according to the method described by Foltmann (1970), with modifications (22). The substrate was prepared by dissolving 12 g of skim milk powder in 100 mL of 10 mM $CaCl_2$ solution and stirring for 5 min at room temperature. Aliquots of 1 mL of reconstituted milk were preincubated at 37°C for 5 min. The reaction was initiated by adding 100 µL of protein extract from wild-type (WT) or PpAP1-overexpressing *P. patens* plants, and the time required for visible clot formation was recorded. To assess the involvement of aspartic protease activity, inhibition assays were performed by preincubating protein extracts with 4 µM pepstatin A for 30 min at room temperature before initiating the milk-clotting reaction.

One milk-clotting unit (MCU) was defined as the amount of enzyme required to coagulate 1 mL of reconstituted skim milk at 37°C within 100 s. Milk-clotting activity (MCU/mL) was calculated as (1000 × D)/t, where t is the clotting time (s) and D is the dilution factor of the enzyme sample.

Results

Identification of Cardosin B homologs in Physcomitrium patens

To identify potential Cardosin B homologs in *P. patens*, the amino acid sequence of Cardosin B from *C. cardunculus* (UniProt accession Q9XFX4) was retrieved and used as a query in a BLASTp search against the *P. patens* proteome (version 7.1) available in the Phytozome database (13).

The search yielded 71 sequence matches, which were subsequently filtered based on sequence identity, retaining candidates sharing greater than 50% amino acid identity with Cardosin B. After removing redundant sequences corresponding to alternative isoforms of the same gene, 16 distinct candidate protein sequences were retained for further analysis.

The selected protein sequences were aligned with Cardosin B using ClustalW (14) and the resulting multiple sequence alignment was visualized using ESPript 3.2 (15) to examine sequence conservation and identify regions of similarity to Cardosin B **(Supplementary Figure S1).**

To further characterize the candidate proteins identified by sequence similarity to Cardosin B, their domain architecture and conserved catalytic motifs were analyzed using InterProScan (16). Particular attention was given to the presence of aspartic peptidase A1 domains, the conservation of the two characteristic catalytic motifs containing aspartate residues, and similarities in protein architecture to Cardosin B from *C. cardunculus*.

All 16 *P. patens* candidate proteins contained regions matching the aspartic peptidase A1 domain profile (PROSITE PS51767). However, differences were observed in the conservation of catalytic motifs and domain organization. Seven proteins displayed both characteristic aspartic peptidase active-site motifs (PROSITE PS00141), eight contained only one detectable motif, and one lacked detectable PS00141 motifs **(Supplementary Figure S2).**

Among the candidates, three proteins (PpV7_05g0093900-5, PpV7_06g0053700-6, and PpV7_15g0051300-3) exhibited the closest resemblance in domain architecture to Cardosin B. All three contained the characteristic eukaryotic aspartyl protease domain (Pfam PF00026), two conserved catalytic motifs, a predicted N-terminal signal peptide, and saposin-like regions (Pfam PF03489 and PF05184), consistent with the typical architecture of plant aspartic proteases **(**Fig. 1A**)**. In contrast, the remaining candidates exhibited a different domain organization, predominantly characterized by regions annotated as xylanase inhibitor N-terminal (PF14543) and C-terminal (PF14541) domains.

These three proteins showed the highest sequence identity to Cardosin B (55–56%), with BLASTp alignments spanning 475–477 amino acids and highly significant E-values. Multiple sequence alignment with Cardosin B revealed extensive conservation, particularly around the two catalytic motifs, with both aspartate residues conserved at corresponding positions in all three *P. patens* proteins **(Supplementary Figure S3** and Fig. 1B**)**.

Together, the high amino acid sequence identity, conservation of both catalytic aspartate residues, and similarity in domain architecture identified these three proteins as the most closely related Cardosin B-like candidates in *P. patens*, supporting their selection for further functional characterization.

Notably, the closest homolog, PpV7_05g0093900-5 (56% identity), corresponds to PpAP1, an aspartic protease previously characterized by Schaaf et al. (2004), who demonstrated its targeting to the secretory pathway and vacuole using GFP fusion proteins (23). Based on its high sequence similarity to Cardosin B, conservation of both catalytic motifs, and characteristic domain architecture, PpAP1 was selected for further structural and functional characterization.

Structural conservation between PpAP1 and Cardosin B

To further characterize the structural relationship between PpAP1 and Cardosin B, the predicted three-dimensional structure of PpAP1 was compared with the experimentally determined crystal structure of Cardosin B (PDB ID: 5NFG). The predicted PpAP1 structure exhibited the characteristic bilobal fold of plant aspartic proteases, with the two catalytic aspartate residues positioned within the cleft separating the two lobes **(**Fig. 2A**)**. Analysis of the prediction confidence revealed high local confidence at both catalytic aspartate residues, Asp98 (pLDDT = 93.5) and Asp285 (pLDDT = 94.3), whereas lower-confidence regions were observed in other parts of the predicted structure **(Supplementary Figure S4A, B)**. Furthermore, sequence alignment of the regions surrounding the catalytic residues revealed conservation of both aspartate-containing motifs among PpAP1, two closely related P. patens proteases, and Cardosin B **(Supplementary Figure S4C).**

Structural superposition revealed a high degree of conservation within the protein core (Fig. 2A). The RMSD calculated for 278 atom pairs retained after structural pruning was 0.853 Å, indicating close structural similarity within the aligned core regions. In contrast, the RMSD calculated for all 327 initially matched atom pairs was 6.445 Å, indicating more pronounced structural differences outside the conserved core.

Closer examination of the catalytic region showed that the two catalytic aspartate residues of PpAP1 (Asp98 and Asp285) were spatially aligned with the corresponding residues of Cardosin B (Asp80 and Asp267, respectively, in chain B of PDB 5NFG) (Fig. 2B). The similar positioning of these residues within the catalytic cleft suggests conservation of the active-site architecture between the two proteins.

Overall, these results demonstrate substantial conservation of the three-dimensional structure of PpAP1 and Cardosin B, particularly within the catalytic core, while revealing structural divergence in other regions of the proteins.

Generation and molecular characterization of PpAP1-overexpressing P. patens lines

To investigate the potential milk-clotting activity of PpAP1, the corresponding coding sequence was amplified and cloned into the pTHUbi expression vector. The resulting construct was introduced into *P. patens* protoplasts, and stable transformants were obtained following antibiotic selection.

To confirm transgene expression in the selected transformants, total RNA was reverse-transcribed into cDNA, followed by PCR amplification using primers specific to the hygromycin resistance gene present in the construct and primers targeting the *PpAP1* gene (Fig. 3A, B). PCR analysis confirmed the expression of both the hygromycin resistance gene and *PpAP1* in the selected transformants.

PpAP1 transcript levels were subsequently quantified by quantitative real-time PCR (qRT-PCR). The transgenic line exhibited a 47-fold increase in *PpAP1* expression compared with wild-type plants (Fig. 3C).

Together, these results confirmed the successful generation of a stable *P. patens* transgenic line overexpressing *PpAP1* at the transcript level, which was subsequently used to evaluate its potential milk-clotting activity.

Milk-clotting activity of PpAP1-overexpressing P. patens

To evaluate the milk-clotting activity associated with *PpAP1* overexpression, soluble protein extracts from wild-type (WT) and *PpAP1*-overexpressing *P. patens* lines were subjected to a milk-clotting assay. Milk coagulation was observed exclusively in extracts from the *PpAP1*-overexpressing line, whereas WT extracts showed no detectable clotting activity under the same experimental conditions (Fig. 4).

Importantly, milk coagulation was completely abolished in the presence of aspartic protease inhibitors, indicating that the observed activity was dependent on aspartic protease activity.

Together, these results provide functional evidence that *PpAP1* overexpression confers milk-clotting activity to *P. patens* extracts, consistent with the milk-clotting properties of Cardosin B.

Discussion

Plant aspartic proteases have attracted considerable interest as alternatives to animal-derived milk-clotting enzymes, particularly those obtained from *C. cardunculus*, such as Cardosin B (24, 25). In this study, we investigated the potential of *P. patens* as a source of endogenous milk-clotting proteases. Through sequence similarity searches and domain analysis, we identified three proteins with a characteristic plant aspartic protease architecture. Among these, *PpAP1* (PpV7_05g0093900-5) exhibited the highest sequence identity to Cardosin B and was selected for functional characterization.

The identification of *PpAP1* as a Cardosin B homolog was supported by the conservation of the two catalytic aspartate residues and the presence of the characteristic aspartic protease domain and plant-specific insert (PSI). These structural features are associated with the catalytic mechanism and intracellular trafficking of typical plant aspartic proteases (26). Interestingly, PpAP1 had previously been characterized by Schaaf et al. (2004)(23), who demonstrated that its N-terminal signal peptide directs proteins into the secretory pathway and that the full-length protein is targeted to the vacuole. However, its proteolytic properties and potential milk-clotting activity were not investigated. Our findings therefore extend the previous characterization of PpAP1 by providing experimental evidence of its milk-clotting potential.

Beyond sequence similarity and domain conservation, the three-dimensional structural comparison provided additional evidence supporting the relationship between PpAP1 and Cardosin B. The predicted PpAP1 structure exhibited the characteristic bilobal organization of pepsin-like aspartic proteases, with the catalytic aspartate residues positioned within the interdomain cleft. Structural superposition with the experimentally determined crystal structure of Cardosin B revealed close agreement between their conserved protein cores, as indicated by the low RMSD obtained for the 278 atom pairs retained after structural pruning (0.853 Å). However, the substantially higher RMSD calculated across all 327 initially matched atom pairs (6.445 Å) indicates that structural similarity is not uniform throughout the proteins. These differences may reflect variations in flexible loops, terminal regions, or other structurally divergent segments, as well as differences between the predicted full-length PpAP1 model and the experimentally determined, processed Cardosin B structure.

Of particular relevance was the conservation of the spatial arrangement of the catalytic aspartate residues. In PpAP1, Asp98 and Asp285 occupied positions corresponding to Asp80 and Asp267 of Cardosin B, respectively, in the crystallographic structure (PDB 5NFG). This structural correspondence is consistent with conservation of the catalytic machinery characteristic of aspartic proteases and suggests that PpAP1 may operate through a similar catalytic mechanism. Nevertheless, the structural model does not establish equivalent substrate specificity or catalytic efficiency, which may depend on differences in substrate-binding regions, protein processing, and other structural features. Furthermore,

because the PpAP1 structure was computationally predicted, these observations should be interpreted as evidence supporting a conserved catalytic architecture rather than direct demonstration of identical enzymatic properties.

The most significant finding of this study was the detection of milk-clotting activity in soluble protein extracts from PpAP1-overexpressing *P. patens* plants, whereas no detectable coagulation was observed in wild-type extracts under the same experimental conditions. Furthermore, milk coagulation was completely abolished in the presence of pepstatin A, an inhibitor of aspartic proteases. Together, these results strongly suggest that the observed activity is mediated by an aspartic protease and is associated with PpAP1 overexpression. The conservation of the catalytic-site architecture between PpAP1 and Cardosin B provides a plausible structural basis for this activity. Although the use of crude protein extracts does not allow the contribution of PpAP1 to be established independently of other endogenous proteins, the combined evidence from transgene expression and inhibitor sensitivity supports its involvement in milk coagulation.

These findings are particularly relevant considering the established milk-clotting properties of Cardosin B (25, 27). Despite their evolutionary divergence, the conservation of both catalytic motifs and the three-dimensional organization of the active site suggests that these proteins may share certain functional properties. The milk-clotting activity detected in PpAP1-overexpressing extracts is consistent with this hypothesis. Nevertheless, further biochemical characterization is required to determine whether PpAP1 exhibits substrate specificity, catalytic efficiency, and milk-clotting properties comparable to those of Cardosin B. In particular, structural conservation alone cannot establish whether PpAP1 recognizes or cleaves milk caseins at the same sites as Cardosin B.

Beyond the characterization of PpAP1, this study highlights the potential of *P. patens* as a platform for the production of endogenous enzymes with biotechnological applications. The availability of efficient genetic transformation methods and controlled cultivation systems makes this moss an attractive alternative for producing proteins of industrial interest (4, 5). In contrast to heterologous expression strategies, overexpression of an endogenous protease may provide an approach for enhancing the production of a protein within its native cellular environment. However, the intracellular localization and processing of PpAP1, as well as their possible effects on enzyme accumulation and recovery, remain to be investigated in the overexpressing lines.

Further studies should focus on the purification of PpAP1 and the quantitative characterization of its enzymatic properties, including milk-clotting activity, general proteolytic activity, optimal pH and temperature, and substrate specificity toward milk caseins. In particular, determining the ratio between milk-clotting and general proteolytic activities will be essential for evaluating its suitability for cheese production, since excessive nonspecific proteolysis may negatively affect product quality (24). Additionally, structural and biochemical studies could also help identify substrate-binding determinants and establish whether the observed differences between PpAP1 and Cardosin B influence their catalytic

properties. Additionally, assessing protein accumulation and enzyme yield in the transgenic lines will help establish the feasibility of using *P. patens* as a production platform.

In conclusion, this study provides evidence that overexpression of the endogenous aspartic protease PpAP1 confers milk-clotting activity to *P. patens* protein extracts. The conservation of the catalytic motifs and three-dimensional active-site architecture shared with Cardosin B, together with the complete inhibition of coagulation by pepstatin A, supports the involvement of PpAP1-associated aspartic protease activity in this process. These findings expand the functional knowledge of moss aspartic proteases and identify PpAP1 as a promising candidate for further development as a plant-derived milk-clotting enzyme.

Declarations

Author Contribution

C.R., L.P.D., and C.V. contributed to the study conception and design. Experimental work and data collection were performed by A.B. and A.V.P. Data analysis was performed by C.R., C.V., and L.P.D. The first draft of the manuscript was written by L.P.D., C.R., and C.V. All authors reviewed and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors would like to thank Sabina Vidal for kindly providing access to her laboratory and facilities, which contributed to the experimental work.

References

1. Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., et al. (2008). The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science*, *319*(5859), 64–69.
2. Lienard, D., & Nogue, F. (2009). Physcomitrella patens: a non-vascular plant for recombinant protein production. *Methods In Molecular Biology*, *483*, 135–144.
3. Reski, R., Bae, H., & Simonsen, H. T. (2018). Physcomitrella patens, a versatile synthetic biology chassis. *Plant Cell Reports*, *37*(10), 1409–1417.
4. Reski, R., Parsons, J., & Decker, E. L. (2015). Moss-made pharmaceuticals: from bench to bedside. *Plant Biotechnology Journal*, *13*(8), 1191–1198.
5. Decker, E. L., & Reski, R. (2007). Moss bioreactors producing improved biopharmaceuticals. *Current Opinion In Biotechnology*, *18*(5), 393–398.
6. Simoes, I., & Faro, C. (2004). Structure and function of plant aspartic proteinases. *European Journal Of Biochemistry*, *271*(11), 2067–2075.

7. Folgado, A., Pires, A. S., Figueiredo, A. C., Pimentel, C., & Abranches, R. (2020). Toward alternative sources of milk coagulants for cheese manufacturing: establishment of hairy roots culture and protease characterization from Cynara cardunculus L. *Plant Cell Reports*, *39*(1), 89–100.
8. Vieira, M., Pissarr, J., Verissimo, P., Castanheira, P., Costa, Y., Pires, E., et al. (2001). Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of Cynara cardunculus L. *Plant Molecular Biology*, *45*(5), 529–539.
9. Ramalho-Santos, M., Verissimo, P., Faro, C., & Pires, E. (1996). Action on bovine alpha s1-casein of cardosins A and B, aspartic proteinases from the flowers of the cardoon Cynara cardunculus L. *Biochimica Et Biophysica Acta*, *1297*(1), 83–89.
10. Folgado, A., & Abranches, R. (2021). Tobacco BY2 cells expressing recombinant cardosin B as an alternative for production of active milk clotting enzymes. *Scientific Reports*, *11*(1), 14501.
11. Schaefer, D., Zryd, J. P., Knight, C. D., & Cove, D. J. (1991). Stable transformation of the moss Physcomitrella patens. *Molecular And General Genetics*, *226*(3), 418–424.
12. Ashton, N. W., & Cove, D. J. (1977). The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss, Physcomitrella patens. *Molecular and General Genetics MGG*, *154*, 87–95.
13. Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, *40*(Database issue), D1178–D1186.
14. Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680.
15. Gouet, P., & Courcelle, E. (2002). ENDscript: a workflow to display sequence and structure information. *Bioinformatics*, *18*(5), 767–768.
16. Blum, M., Hobbs, E., Florentino, L., & Bateman, A. (2026). InterProScan 6: a modern large-scale protein function annotation pipeline. *Bioinform Adv*, *6*(1), vbag141.
17. Kim, G., Lee, S., Levy Karin, E., Kim, H., Moriwaki, Y., Ovchinnikov, S., et al. (2025). Easy and accurate protein structure prediction using ColabFold. *Nature Protocols*, *20*(3), 620–642.
18. Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., et al. (2023). UCSF ChimeraX: Tools for structure building and analysis. *Protein Science*, *32*(11), e4792.
19. Vidali, L., Augustine, R. C., Kleinman, K. P., & Bezanilla, M. (2007). Profilin is essential for tip growth in the moss Physcomitrella patens. *The Plant Cell*, *19*(11), 3705–3722.
20. Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., & Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*. ;35(Web Server issue):W71–W74.
21. Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, *25*(4), 402–408.

22. Foltmann, B. (1970). The N-terminal and C-terminal amino acid sequence of calf rennin. *Philosophical Transactions Of The Royal Society Of London. Series B, Biological Sciences*, *257*(813), 147–151.
23. Schaaf, A., Reski, R., & Decker, E. L. (2004). A novel aspartic proteinase is targeted to the secretory pathway and to the vacuole in the moss Physcomitrella patens. *European Journal Of Cell Biology*, *83*(4), 145–152.
24. Shah, M. A., Mir, S. A., & Paray, M. A. (2014). Plant proteases as milk-clotting enzymes in cheesemaking: a review. *Dairy Science & Technology*, *94*(1), 5–16.
25. Silva, S. V., & Malcata, F. X. (1999). On the activity and specificity of cardosin B, a plant proteinase, on ovine caseins. *Food Chemistry*, *67*(4), 373–378.
26. Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A., & Zdanov, A. (1999). Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. *The EMBO Journal*, *18*(14), 3947–3955.
27. Veríssimo, P., Faro, C., Moir, A. J. G., Lin, Y., Tang, J., & Pires, E. (1996). Purification, Characterization and Partial Amino Acid Sequencing of Two New Aspartic Proteinases from Fresh Flowers of Cynara cardunculus L. *European Journal of Biochemistry*, *235*(3), 762–768.

Figures

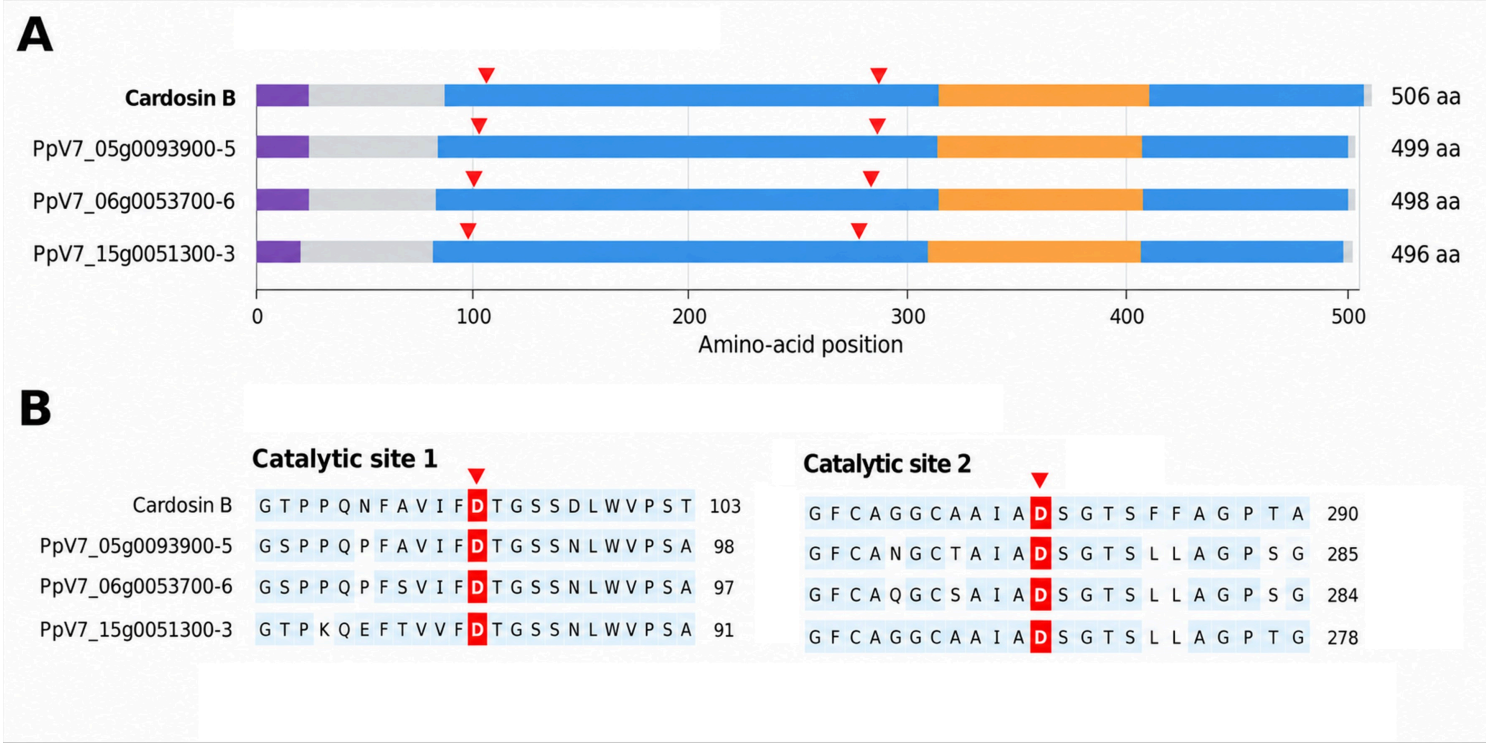


Figure 1

Domain architecture and conservation of catalytic residues in Cardosin B and its closest homologs from *P. patens*. (A) Predicted domain organization of Cardosin B from *C. cardunculus* and three *P.*

*patens*homologs, based on InterProScan analysis. Red arrowheads indicate the positions of catalytic aspartate residues. **(B)** Comparison of amino acid sequences surrounding the two conserved catalytic aspartates. Conserved residues are highlighted in light blue, while catalytic aspartates are shown in red. Protein lengths are indicated in amino acids.

These three proteins showed the highest sequence identity to Cardosin B (55–56%), with BLASTp alignments spanning 475–477 amino acids and highly significant E-values. Multiple sequence alignment with Cardosin B revealed extensive conservation, particularly around the two catalytic motifs, with both aspartate residues conserved at corresponding positions in all three *P. patens*proteins **(Supplementary Figure S3** and **Figure 1B)**.

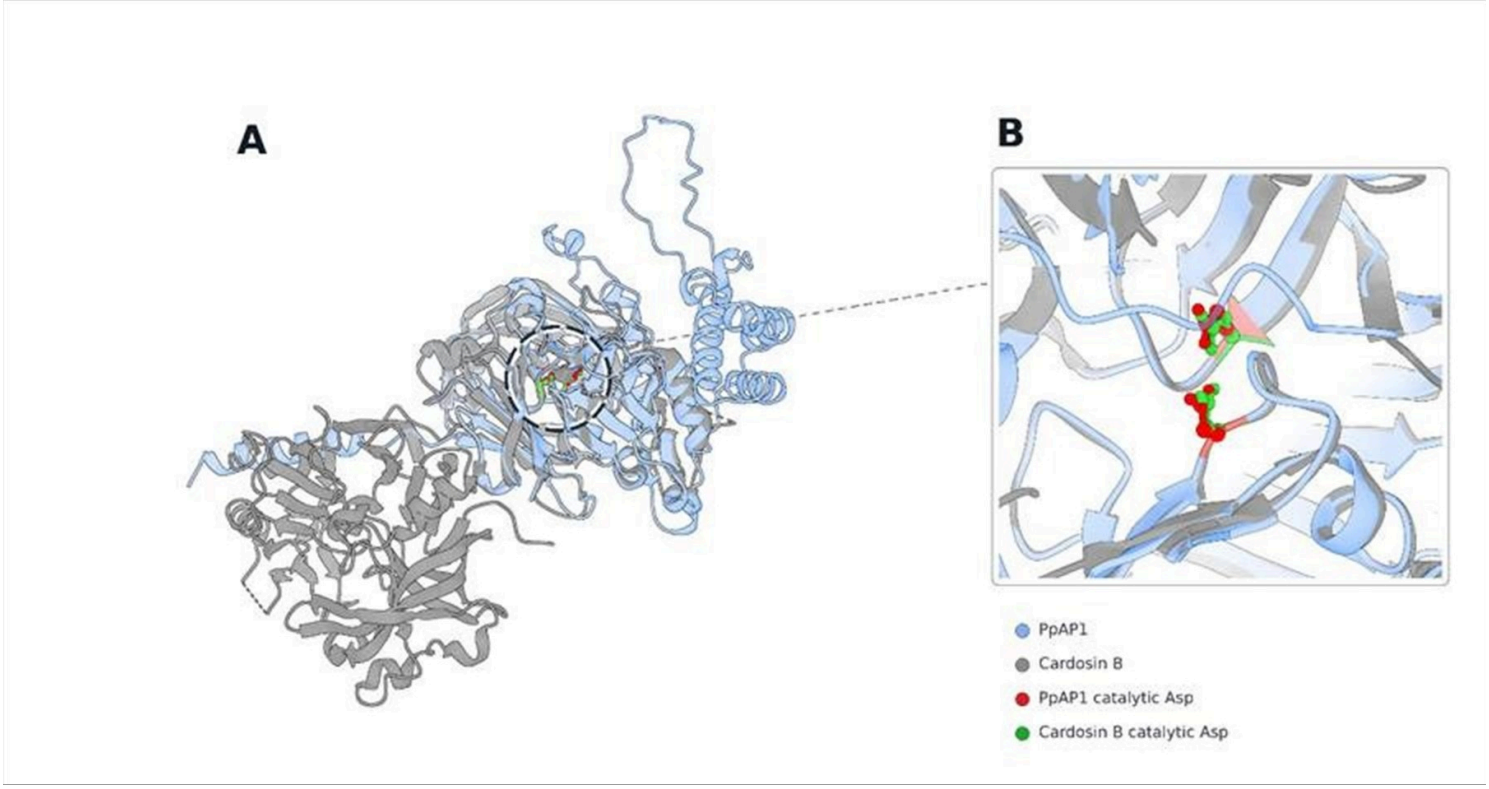


Figure 2

Structural comparison of PpAP1 from *P, patens* and Cardosin B. (A) Structural superposition of the ColabFold-predicted PpAP1 model (blue) and the experimentally determined crystal structure of Cardosin B (gray; PDB ID: 5NFG). The RMSD for 278 atom pairs retained after structural pruning was 0.853 Å.**(B)** Enlarged view of the catalytic region indicated in panel A, highlighting the conserved catalytic aspartate residues of PpAP1 (Asp98 and Asp285, red) and the corresponding residues of Cardosin B (Asp80 and Asp267 in chain B of PDB 5NFG, green). Structural superposition and molecular visualization were performed using UCSF ChimeraX.

Overall, these results demonstrate substantial conservation of the three-dimensional structure of PpAP1 and Cardosin B, particularly within the catalytic core, while revealing structural divergence in other regions of the proteins.

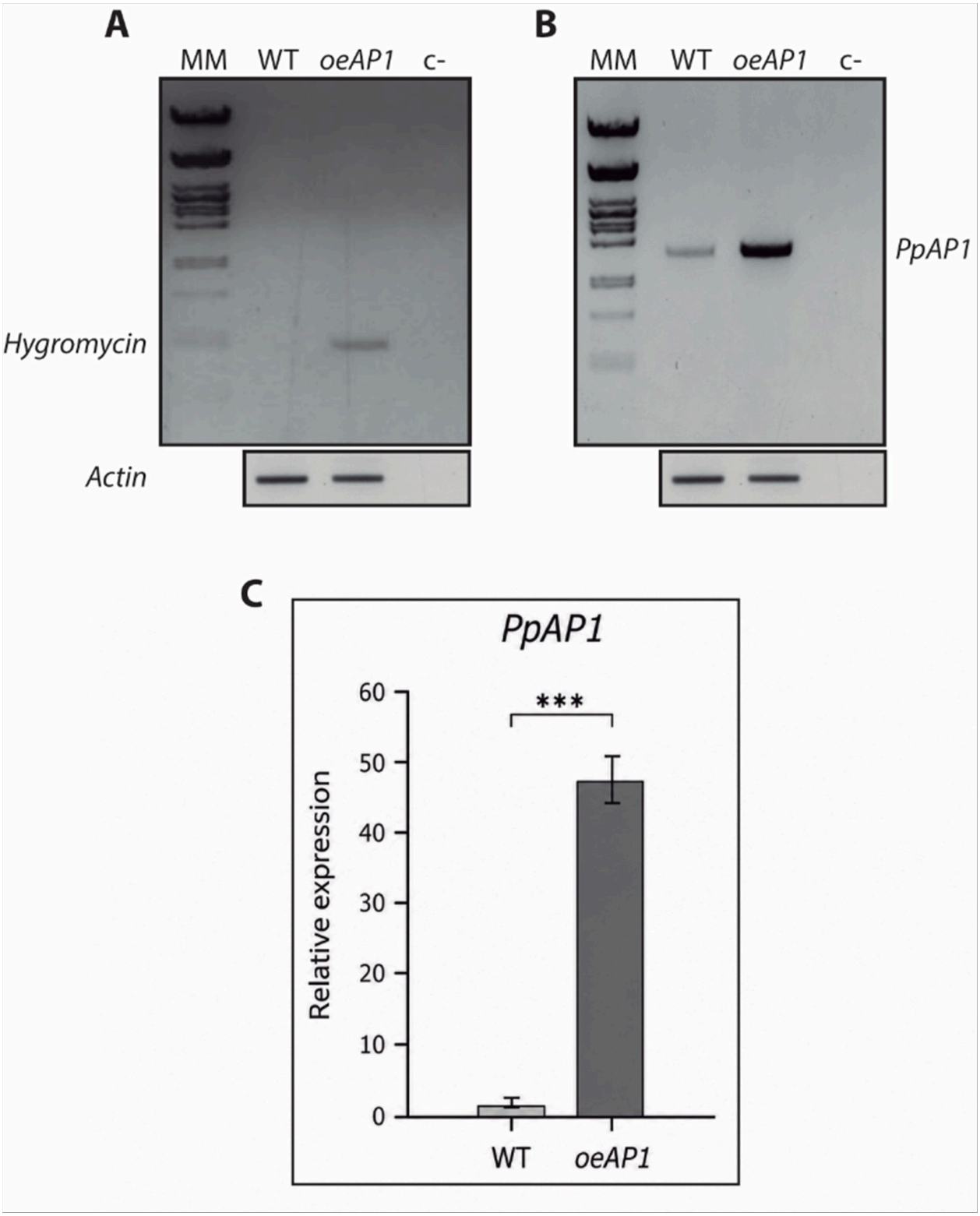


Figure 3

Molecular characterization of *PpAP1*-overexpressing *P. patens* plants. (A) Semi-quantitative RT-PCR analysis of the hygromycin resistance gene in wild-type (WT) and *PpAP1*-overexpressing (oeAP1) plants**. (B)** Semi-quantitative RT-PCR analysis of *PpAP1*expression in WT and oeAP1 plants. Actin was used as an internal reference gene in both analyses. MM, molecular weight marker; C−, negative control. **(C)**

Quantitative real-time PCR (RT-qPCR) analysis of *PpAP1* transcript levels in WT and oeAP1 plants. Relative expression levels are presented as mean ± standard deviation (SD).

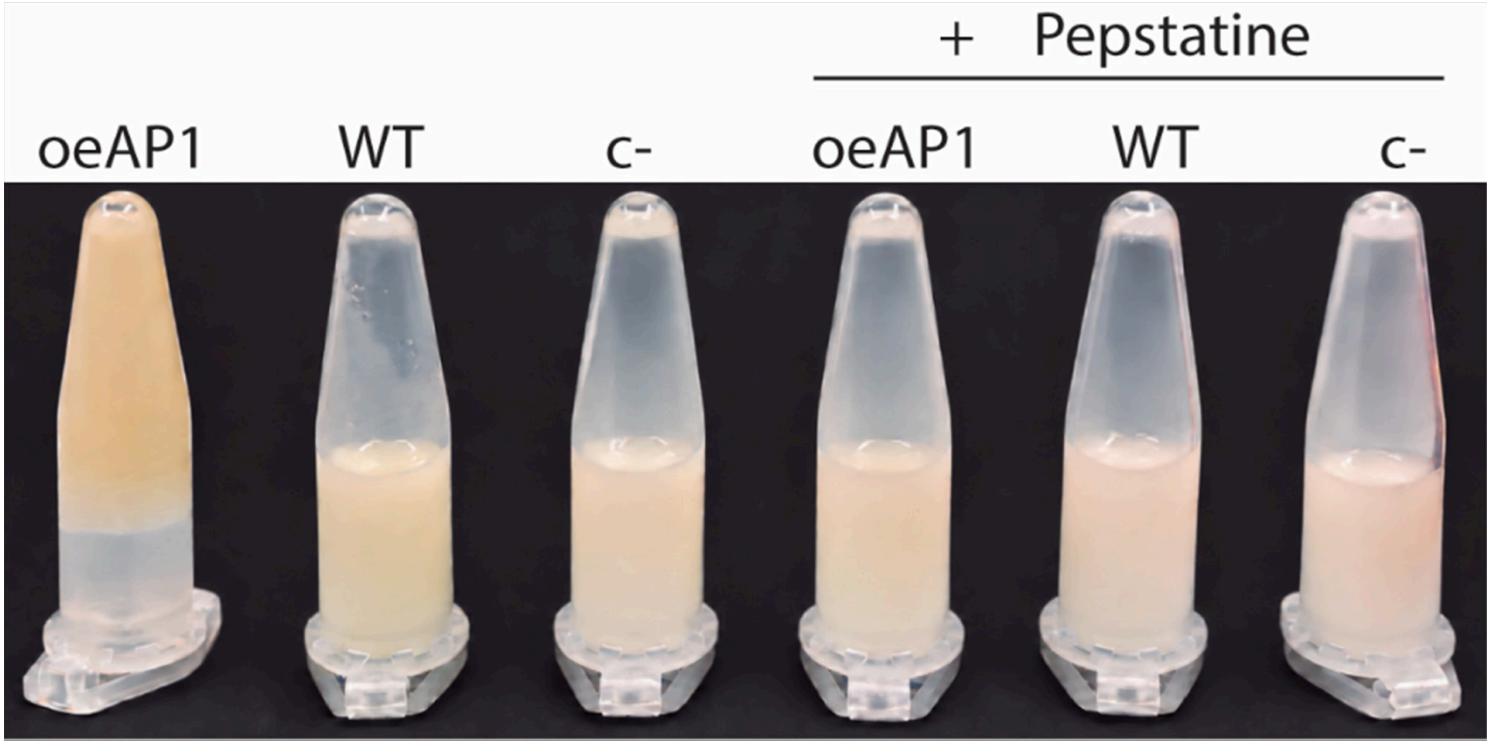


Figure 4

Milk-clotting activity of extracts from *P. patens* overexpressing *PpAP*1. Representative images of milk coagulation assays performed using extracts from *PpAP1*-overexpressing (oeAP1) and wild-type (WT) plants, in the absence (left) or presence (right) of pepstatin A. As negative control (C−), the extraction buffer was included under both conditions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryS1.pdf
- SupplementaryS4.tiff
- SupplementaryS3.tif
- SupplementaryS2.tiff
- Supplementaryinformationlegends.docx