Planta (2011) 234:293–304
DOI 10.1007/s00425-011-1399-7

ORIGINAL ARTICLE

Biochemical characterization, cDNA cloning, and molecular modeling of araujiain aII, a papain-like cysteine protease from *Araujia angustifolia* latex

Walter D. Obregón · Daniela Lufrano · Constanza S. Liggieri · Sebastián A. Trejo · Sandra E. Vairo-Cavalli · Francesc X. Avilés · Nora S. Priolo

Received: 29 December 2010 / Accepted: 6 March 2011 / Published online: 20 March 2011
© Springer-Verlag 2011

Abstract Araujiain aII, the protease with highest specific activity purified from latex of *Araujia angustifolia* (Apocynaceae), shows optimum proteolytic activity at alkaline pH, and it is completely inhibited by the irreversible inhibitor of cysteine proteases *trans*-epoxysucciny-L-leucyl-amido(4-guanidino) butane. It exhibits esterolytic activity on several *N*-α-Cbz-amino acid *p*-nitrophenyl esters with a preference for Gln, Ala, and Gly derivatives. Kinetic enzymatic assays were performed with the thiol proteinase substrate *p*-Glu-Phe-Leu-*p*-nitroanilide ($K_m = 0.18 \pm 0.03$ mM, $k_{cat} = 1.078 \pm 0.055\ s^{-1}$, $k_{cat}/K_m = 5.99 \pm 0.57\ s^{-1}\ mM^{-1}$). The enzyme has a p*I* value above 9.3 and a molecular mass of 23.528 kDa determined by mass spectrometry. cDNA of the peptidase was obtained by reverse transcription-PCR starting from total RNA isolated from latex. The deduced amino acid sequence was confirmed by peptide mass fingerprinting analysis. The N-terminus of the mature protein was determined by automated sequencing using Edman's degradation and compared with the sequence deduced from cDNA. The full araujiain aII sequence was thus obtained with a total of 213 amino acid residues. The peptidase, as well as other Apocynaceae latex peptidases, is a member of the subfamily C1A of cysteine proteases. The enzyme belongs to the alpha + beta class of proteins, with two disulfide bridges (Cys22–Cys63 and Cys56–Cys95) in the alpha domain, and another one (Cys150–Cys201) in the beta domain, as was suggested by molecular modeling.

Keywords *Araujia angustifolia* · Cysteine protease · Latex peptidase · Papain-like protease · Araujiain

Abbreviations

AaCPII	Amino acid sequence corresponding to the cysteine protease araujiain aII deduced from cDNA and N-terminal sequence
ACN	Acetonitrile
AMPSO	(*N*-(1,1-dimethyl-2-hydroxyethyl)-3-amino-2-hydroxy-3-[(1-hydroxy-2-methylpropan-2-yl) amino] propane-1-sulfonic acid
CAPS	3-[Cyclohexylamino]-1-propanesulfonic acid
CE	Crude extract
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
E-64	*trans*-Epoxysucciny-L-leucyl-amido(4-guanidino) butane
EDTA	Ethylenediaminetetraacetic acid
MES	4-Morpholineethanesulfonic acid
MOPS	3-Morpholinopropanesulfonic acid
PFLNA	*p*-Glu-Phe-Leu-*p*-nitroanilide
PLCPs	Papain-like cysteine proteases
PMF	Peptide mass fingerprinting
PVDF	Polyvinyl difluoride
RACE	Rapid amplification of cDNA ends
RMS	Root mean square deviation
RT-PCR	Reverse transcription-polymerase chain reaction

W. D. Obregón · D. Lufrano · C. S. Liggieri · S. E. Vairo-Cavalli (✉) · N. S. Priolo
Laboratorio de Investigación de Proteínas Vegetales, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, 47 y 115 s/N, C.C. 711, B1900AVW La Plata, Argentina
e-mail: svairo@biol.unlp.edu.ar

S. A. Trejo · F. X. Avilés
Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Campus Universitari, Bellaterra, Cerdanyola del Vallès, 08193 Barcelona, Spain

TAPS [(2-Hydroxy-1,1-bis(hydroxymethyl) ethyl) amino]-1-propanesulfonic acid sodium–potassium salt, *N*-tris (hydroxymethyl) methyl-3-aminopropanesulfonic acid
TFA Trifluoroacetic acid

Introduction

Araujia angustifolia [(Apocynaceae, subfamily Asclepiadoideae, formerly Asclepiadaceae family) Endress and Bruyn (2000)], a native species from South America, commonly known as "tasi", "taso" or "doca" (Bucciarelli et al. 2008) oozes sticky and toxic latex upon tissue damage. Latex is a milky fluid with a complex mixture of constituents, including proteins, vitamins, carbohydrates, lipids, terpenes, alkaloids, starches, oils, tannins, resins, gums and free amino acids. Laticifers not only serve as a repository for natural products but also exhibit unique proteomes. Many latex proteins are thought to have defensive roles in plants, and include proteases, chitinases and pathogenesis-related polypeptides (Hagel et al. 2008; Liggieri et al. 2009). Over 110 latices of different plant families are known to contain at least one proteolytic enzyme. Most of them belong to the cysteine or serine endopeptidase catalytic type (Domsalla and Melzig 2008). Species belonging to the Asclepiadoideae subfamily often contain papain-like cysteine proteases (PLCPs) in latex, which may be integrated in the defense scheme of those plants against a specialist herbivore playing crucial roles in plant-pathogen/pest interactions (Rasmann et al. 2010). During these parasitic interactions, PLCPs act on host pathogen substrates. In this manner proteolysis induces the selection of counteracting inhibitors, non-cleavable substrates and other means to evade proteolysis. Therefore, the interactions of proteases with their substrates and inhibitors can be seen as a molecular battlefield. Intriguingly, both plants and their invaders use PLCPs at these molecular battlefields (Shindo and van der Hoorn 2008). In *A. angustifolia*, latex is contained within unarticulated ramified laticifers creating a network throughout most of the plant.

From a pharmaceutical and biotechnological point of view, latex is an important source of plant peptidases used in industry due to their property of being active over wide ranges of temperature and pH (Domsalla and Melzig 2008; Liggieri et al. 2009). Applications include the use of these enzymes as meat tenderizers (such as Panol® Purified Papain, Liquipanol® T100), component of detergent formulations (Khaparde and Singhal 2001), or surfactants synthesis (Morcelle et al. 2009) among others. Papain alone or in combination with other proteolytic enzymes has long been available for diverse medical indications (Beuth 2008; Salas et al. 2008); as an example we can mention Papacarie, a gel available in the market, used for chemomechanical dental caries removal. Latices from Asclepiadoideae subfamily have been used in folk medicine as antiparasitic agents against gastrointestinal nematodes and in wound healing in addition to other ailments. The pharmacological actions of these milky latices have been attributed to the presence of cysteine peptidases (Stepek et al. 2004; Obregón et al. 2009a). Shivaprasad et al. (2009) have shown that the PLCPs present in *Asclepias curassavica* latex exhibited strong procoagulant action and were found to be specific in their action (thrombin like). This is considered as the basis for the traditional use of plant latex to stop bleeding on fresh cuts. On the other hand, PLCP activity of *Calotropis procera* latex is involved in the larvicidal action of latex proteins on *Aedes aegypti* larvae (Ramos et al. 2009).

PLCPs are usually 23–30 kDa in size, and use a catalytic cysteine residue to cleave peptide bonds in protein substrates. This catalytic cysteine is part of a catalytic triad situated in the middle of a cleft that binds the substrate through specific interactions (Shindo and van der Hoorn 2008). PLCPs are included in clan CA, family C1 and subfamily A in the MEROPS classification (http://www.merops.sanger.ac.uk). Subfamily C1A comprises proteases that contain disulfide bridges and accumulate in vesicles, vacuoles, or the apoplast (van der Hoorn 2008). In the current paper we report the biochemical, proteomic and molecular characterization of araujiain aII, the PLCP with higher specific activity, purified from latex of the milkweed *A. angustifolia*.

Materials and methods

Plant material

Fruits of *A. angustifolia* (Hook. et Arn.) Decaisne were obtained from plants grown in the town of Ringuelet, Province of Buenos Aires, Argentina (Obregón et al. 2006). The plant is an attractive, small, tender vine with small green twining stems, evergreen arrow-shaped leaves and white bell-shaped flowers with the petals reflexed leaving a pink pointed central structure; fruits are green oblong-fusiform follicles (Burkart 1979). Voucher specimens were deposited in the LPE herbarium (Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina).

Chemicals

Bovine serum albumin, *N*-α-carbobenzoxy-*p*-nitrophenyl esters of Ala, Asn, Asp, Gln, Gly, Ile, Leu, Lys, Phe, Pro,

Trp, Tyr and Val, MES, MOPS, TAPS, AMPSO and CAPS were purchased from Sigma–Aldrich Inc. (St. Louis, MO, USA). *p*-Glu-Phe-Leu-*p*-nitroanilide (PFLNA) and E-64 were products of Bachem AG (Torrance, CA, USA). Casein (Hammarsten type) was obtained from Research Organics (Cleveland, OH, USA). Coomassie Brilliant Blue R-250, acrylamide, bisacrylamide, and 3–10 carrier ampholytes were obtained from Bio-Rad (Hercules, CA, USA). SP-Sepharose Fast Flow was purchased from GE Healthcare Life Sciences (Uppsala, Sweden). EcoTaq DNA polymerase was obtained from ECOGEN Barcelona (Barcelona, Spain). Trypsin Gold and pGEM-T Easy vector were products from Promega Corporation (Madison, WI, USA). Restriction enzymes were obtained from Roche (Roche Diagnostics-Roche Applied Science, Indianapolis, IN, USA). Peptide calibrants and α-cyano-4-hydroxycinnamic acid were purchased from Bruker Daltonics (Bremen, Germany). *Escherichia coli* XL1-Blue was purchased from Stratagene (La Jolla, CA, USA). Pre-stained broad molecular mass standard was purchased from Invitrogen Corporation (Carlsbad, CA, USA). Other reagents used were of the highest grade available.

Crude extract preparation and protein determination

Latex obtained by superficial incisions of fruits, collected in 0.05 M citric–citrate buffer (pH 4.5) containing 5 mM EDTA, was first centrifuged at 5,000×*g* for 30 min at 4°C. Gums and other insoluble materials were discarded, and the supernatant was ultra centrifuged at 100,000×*g* for 30 min at 4°C. This new supernatant ("crude extract"), containing soluble proteins, was fractionated and conserved at −20°C for further studies.

Proteins present in the crude extract (CE) were determined by Bradford's method using bovine albumin as standard (Bradford 1976).

Protein purification

Purification was carried out by cation exchange chromatography according to the method of Obregón et al. (2009b). Briefly, 1 ml of the crude extract containing 1.5 mg of protein was loaded onto a Pharmacia XK 16/40 column having AK16 adaptors, packed with SP-Sepharose Fast Flow and equilibrated with 0.055 M Tris–HCl (pH 7.4). Chromatography was developed in an FPLC equipment (Pharmacia, Uppsala, Sweden) by washing with the equilibrating buffer and further elution of the bound material with a step gradient of sodium chloride (0–0.5, 0.5–0.8, and 0.8–2.0 M) in the same buffer. Cation exchange chromatography was spectrophotometrically monitored by absorbance measurement at 280 nm.

Caseinolytic activity and pH profile

Proteolytic assays were made using casein as substrate. The reaction mixture was prepared by mixing 0.1 ml of enzyme extract with 1.1 ml of 1% casein containing 12 mM cysteine, in a 0.1 M Tris–HCl buffer (pH 8.0). The reaction was performed at 45°C on an Agilent 8453 UV–visible spectrophotometer equipped with a thermostated cell and was stopped 2 min later by the addition of 1.8 ml of 5% TCA. Each test tube was centrifuged at 3,000×*g* for 30 min and the absorbance of the supernatant measured at 280 nm. An arbitrary enzyme unit (caseinolytic unit, U_{cas}) was defined as the amount of protease which produces an increment of one absorbance unit per min in the assay conditions (Priolo et al. 1991).

The effect of pH on enzyme activity of the purified protease was measured with casein (pH range 6.4–10.5) using 10 mM sodium salts of the following "Good" buffers: MES, MOPS, TAPS, AMPSO and CAPS (Liggieri et al. 2004).

Triplicate measurements were performed for each assay.

Endoesterolytic activity

These assays were performed under optimal conditions according to the method of Silverstein (1974) modified by Obregón et al. (2001) with *N*-α-carbobenzoxy-*p*-nitrophenyl esters of the following amino acids: Ala, Asn, Asp, Gln, Gly, Ile, Leu, Lys, Phe, Pro, Trp, Tyr and Val. Reaction mixture consisted of 1.8 ml of 0.1 M Tris–HCl buffer (pH 8.0), 0.1 ml of 1 mM substrate in ACN solution and 0.1 ml of purified peptidase. Continuous liberation of *p*-nitrophenol was followed at 37°C with an Agilent 8453 E UV–visible spectroscopy system at 405 nm every 3 s for 2 min. Triplicate measurements were performed for each assay and control assays were also made. In each case the non-enzymatic hydrolysis for each substrate was determined by replacing the enzyme solution with 0.1 ml of buffer. Enzyme kinetic data were analyzed for the most preferred *N*-Cbz-aa *p*-nitrophenyl ester derivative determined in the aforementioned assays; substrate concentrations range from 0.1 to 1 mM in the reaction mixture.

Amidolytic activity and enzyme kinetic

Amidasic activity was measured according to the method of Filippova et al. (1984) modified by Morcelle et al. (2004). The reaction mixture contained 1.5 ml of 0.1 M phosphate buffer pH 6.5, 0.3 M KCl, 10 mM EDTA, 3 mM DTT, 0.18 ml of substrate (stock solution: 4 mM PFLNA in DMSO) and 0.12 ml of enzyme. The *p*-nitroaniline released at 37°C was spectrophotometrically detected at 410 nm. Triplicate measurements were performed and

control assays were also made. An arbitrary enzyme activity unit (U_{PFLNA}) was defined as the amount of protease that released one micromole of *p*-nitroaniline per minute in the assay conditions. Kinetic parameters using this substrate were determined for the purified enzyme. K_m and k_{cat} were calculated by non-linear regression analysis of Michaelis–Menten equation.

Determination of catalytic type

The action of different cysteine protease inhibitors (50 mM sodium iodoacetate, 50 mM iodoacetamide and 100 μM E-64) was evaluated by incubating the purified enzyme for 30 min at 25°C (Salvesen and Nagase 2001). The residual caseinolytic activity was measured as indicated in "Caseinolytic activity and pH profile" (Obregón et al. 2006).

Gradient SDS-PAGE analysis

The active fraction, uninhibited and inhibited with 10 mM sodium iodoacetate, was submitted to denaturing SDS polyacrylamide gel electrophoresis according to Schägger and von Jagow (1987) with modifications. The electrophoresis was performed in a 10–16% linear gradient of polyacrylamide and 0–9% sucrose linear gradient to amplify resolution. Electrophoresis was run in two constant voltage steps (stacking gel: 30 V, resolution gel: 90 V). Gels were stained with Coomassie Brilliant Blue R-250. Silver staining was also performed to improve visualization of protein bands (O'Connell and Stults 1997).

Isoelectric point determination

Non-denaturing isoelectrofocusing (IEF) was carried out on immobilized pH gradient (3–10) in polyacrylamide gels (10%) (Biolyte 3-10 carrier ampholytes) in a Mini IEF Cell (Model 111, Bio-Rad) as described by Obregón et al. (2006).

Mass spectrometry analyses

Matrix-assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI-TOF MS) was used for the determination of endopeptidase molecular mass, as well as its purity degree. MALDI-TOF mass spectra was acquired on a BRUKER Ultraflex spectrometer equipped with a pulsed nitrogen laser (337 nm), in linear positive ion mode, using a 19 kV acceleration voltage. Samples were prepared by mixing equal volumes of a saturated solution of matrix (3,5-dimethoxy-4-hydroxycinnamic acid–sinapinic acid) in 0.1% TFA in water/ACN 2:1, and 1 μM protein solution. From this mixture, 1 μl was spotted on a sample slide (MP 384 Polished Steel, Bruker) and allowed to evaporate to dryness. Trypsinogen was used as standard for mass calibration (M+1 23981, M+2 11991).

PMF by MALDI-TOF MS

In situ tryptic digestion of an electrophoretically homogeneous peptidase band was performed following the protocol of Obregón et al. (2006). Peptides were dissolved in 5 μl 0.1% TFA (v/v) and analyzed by MALDI-TOF MS using a matrix of acid α-cyano-4-hydroxycinnamic (HCCA). For calibration a standard mixture was used composed of bradykinin [1–7 (757.39916)], angiotensin II (1,046.5418), angiotensin (1,296.6848), substance P (1,347.7354), bombesin (1,619.8223), renin substrate (1,758.93261), adrenocorticotropic hormone: ACTH clip [1–17 (2,093.0862)], adrenocorticotropic hormone: ACTH clip [18–39 (2,465.1983)] and somatostatin 28 (3,147.4710). The results were processed using the MASCOT search engine (http://www.matrixscience.com). Search parameters were (1) type of search, peptide mass fingerprint; (2) enzyme, trypsin; (3) database, SwissProt 55.2; (4) taxonomy, Viridiplantae; (5) Variable modifications, carbamidomethyl (C), oxidation (M); (6) mass values, monoisotopic; (7) peptide mass tolerance: ±100 ppm; (8) peptide charge state, 1+.

N-terminal sequence

The N-terminal sequence was determined by automated Edman's degradation. Samples were prepared by immobilizing 10–100 pmol of pure protein on polyvinylidene fluoride (PVDF) membrane in Ultrafree-Probind Filter tubes (Millipore Corporation, Billerica, MA, USA). Each sample was centrifuged at 10,000 rpm for 2 min and the PVDF filter was washed with MilliQ water. Finally, we introduced the immobilized protein in the reaction chamber of the Procise protein sequencer (Applied Biosystem, Life Technologies, Carlsbad, CA, USA).

Cloning of araujian aII

RNA of *A. angustifolia* was extracted from two latex drops, obtained by superficial incisions of petioles, using the RNAeasy Plant Mini Kit (Qiagen, Barcelona, Spain) following the manufacturer's protocol. To obtain the first cDNA chain reverse transcription reaction was carried out using a commercial kit from Roche (First Strand cDNA Synthesis Kit for RT-PCR, AMV) and the primers R_0R_1polidT (5′-CCGGAATTCACTGCAGGGTACCCAATACGACTCACTATAGGGCTTTTTTTTTTTTTTTTT-3′), being R_0 = 5′-CCGGAATTCACTGCAG-3′ and R_1 = 5′-GGTACCCAATACGACTCACTATAGGGC-3′. The reaction was carried out in a thermal cycler XPCycler Bioer Technology Co., Ltd. (Hangzhou, Zhejiang, China), program:

10 min at 25°C, 120 min at 42°C, 5 min at 99°C and 5 min at 10°C.

PCR was made using specific primers that recognize highly conserved sequences of latex peptidases of the family Apocynaceae (NTapo$_1$ 5′-GTTGAATTGCCAGATTCTGTAGATTGG-3′, NTapo$_2$ 5′-CCAGATTCTGTAGATTGGCGG-3′ (Obregón et al. 2009a) in combination with R$_0$. Nt$_{1-8}$ (5′-yTkCCdGATTCCGATGTTTGGmG-3′), effective for the isolation of cDNAs of other cysteine proteases from latex (Trejo et al. 2009) was also used combined with R$_0$. Reactions were carried out under the following conditions: [1× (5 min at 95°C, 15 min at 72°C), 30× (1 min at 94°C, 1 min at 46°C, 2 min at 72°C), 1× (15 min at 72°C), 1× (16 h at 10°C)]. An aliquot of total RNA was used as template to control RT-PCR.

To confirm that the PCR products corresponded to the sequence of the gene sought we conducted a nested PCR reaction (Frohman 1994). Reaction conditions were as specified above using two combinations of internal primers: CAapo$_1$-R$_1$, and CAapo$_2$-R$_1$ (CAapo$_1$ 5′-CCTATCAGAAATCAAGGAAAATGTGGGAGTTGCTGG-3′ and CAapo$_2$ 5′-ATCAAGGAAAATGTGGGAGTTGCTGG-3′). The nested-PCR products were separated by agarose gel (2%) electrophoresis, and selected DNA bands were excised and purified using a DNA extraction kit (QIAEX II Agarose Gel Extraction, QIAGEN GmbH). Finally, the purified DNA fragments were cloned into the pGEM-T Easy vector in heat shock competent *E. coli* XL1-Blue. Plasmids of selected colonies were purified using the GFX extraction kit Micro Plasmid Prep Kit (GE HealthCare Life Sciences) and sequenced by the Sequencing Services of Veterinary Faculty, Autonomous University of Barcelona (Barcelona, Spain). Chromatograms were analyzed with the software Chromas v2.13 (Technelysium Pty. Ltd.).

cDNA sequence analysis, homology modeling and validation of 3-D structure

cDNA sequences obtained from selected clones were analyzed with Clustal-W alignment software (Thompson et al. 1994) to obtain a consensus sequence. The consensus sequence was translated in all six reading frames with the *Translate Tool* software available on the ExPASy server to identify the presence of conserved cysteine endopeptidase elements in the amino acid sequence. At the same time, a PSI-BLAST search was conducted using the NCBI nr (non-redundant) database restricted to the *Viridiplanteae* Kingdom with the matrix Blosum 62. Distance map was constructed using the Neighbor-Joining Method. Physicochemical properties of araujiain aII were predicted by GPMAW v6.0 (Lighthouse data, DK-5230, Odense M, Denmark) and theoretical values were compared with experimental data.

The 3D model of araujiain aII from *A. angustifolia* was built by homology modeling using Modeller 9v7 software. Sequence similarity search was performed with the NCBI Basic Local Alignment Search Tool (BLASTprotein) and crystal structures of the closest homologues available in the Brookhaven Protein Data Bank (PDB: http://www.rcsb.org/pdb/home/home.do) were selected. The high score and low *E*-value were considered. Results yielded by NCBI BLAST revealed papaya proteinase omega with a resolution of 1.80 Å (PDB ID: 1PPO) of latex of *Carica papaya* as a suitable template. Models obtained were evaluated and validated with DOPE assessment score.

Results and discussion

Characterization of araujiain aII

Araujiain aII, the most active peptidase isolated from the latex of *A. angustifolia*, was purified by cation exchange chromatography as previously reported by Obregón et al. (2009b). Homogeneity of this peptidase was analyzed by electrophoresis on polyacrylamide gradient gel (Fig. 1) and mass spectrometry (Fig. 2). Unlike what happens with other peptidases when preparing electrophoresis samples (Priolo et al. 2000; Obregón et al. 2001) the presence of an inhibitor is not required to prevent autodigestion. Autolysis resistance proteases are very useful for industrial or therapeutic applications and frequently proteases are modified by protein engineering to avoid rapid autolysis (Singh et al. 2010).

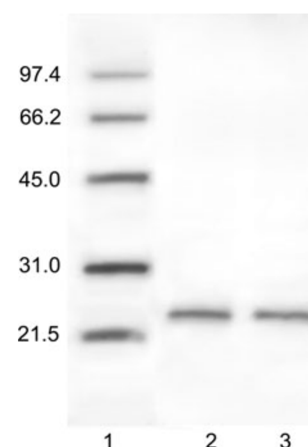
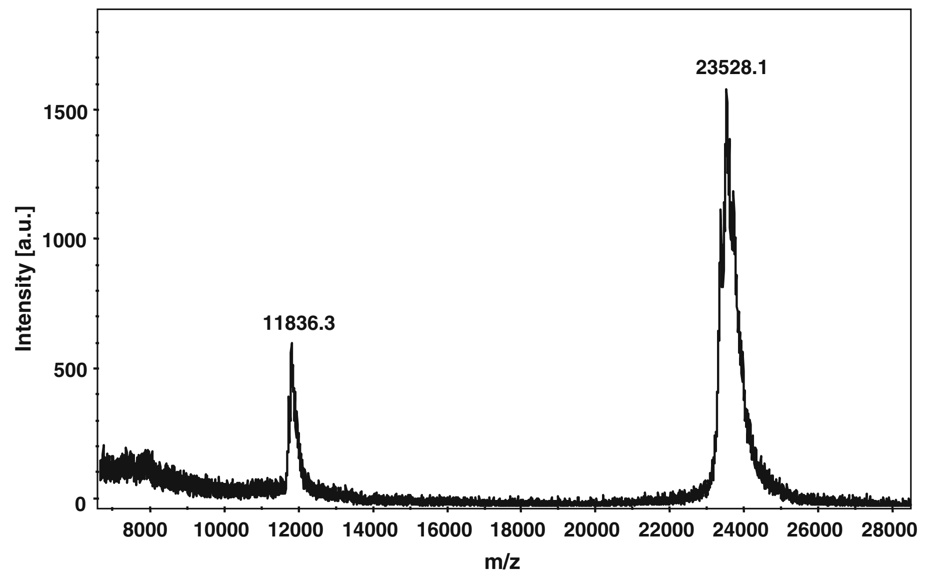


Fig. 1 Gradient SDS-PAGE. Molecular mass standards (Bio Rad): phosphorylase b (97.4 kDa), bovine serum albumin (66.2 kDa), ovoalbumin (45.0 kDa), carbonic anhydrase bovine (31.0 kDa), soybean trypsin inhibitor (21.5 kDa) (*Lane 1*); araujiain aII inhibited with sodium iodoacetate (*Lane 2*); araujiain aII without inhibitor (*Lane 3*)

Fig. 2 Araujiain aII mass spectrometry (MALDI-TOF)



Araujiain aII molecular mass measured by MALDI-TOF was 23.528 kDa (Fig. 2). This value is of the same order of magnitude than those obtained for other Apocynaceae peptidases (Obregón et al. 2009b) and as reviewed by Domsalla and Melzig (2008) all latex cysteine peptidases are in the range from 21 to 29 kDa.

A suspension of pure enzyme with a concentration of 55 µg protein/ml displayed a caseinolytic activity of 0.17 U_{cas}/ml (specific activity: 3.1 U_{cas}/g). The range of pH at which the enzyme had more than 10% activity on the same substrate was 7.5–9.0, while it showed more than 90% of activity in a very narrow pH range, between 8.0 and 8.5 (Fig. 3). These data were consistent with that observed for the crude extract of *A. angustifolia* latex with a pH optimum between 6.7 and 8.5 (Obregón et al. 2006). The action of cysteine protease inhibitors produced a total and irreversible inhibition of enzyme activity. While β-mercaptoethanol, DTT, and cysteine produced a significant activation of the enzyme confirming the cysteine peptidase catalytic mechanism of araujian aII (data not shown).

Resembling other cysteine proteases, araujiain aII in addition to proteolytic activity showed amidolytic (U_{PFLNA}: 0.0354) and esterolytic activities (Fig. 4). K_m of this peptidase towards PFLNA (amidolytic activity) was 0.18 ± 0.026 mM, about half of those obtained for papain, 0.34 mM, bromelain, 0.30 mM and ficain, 0.43 mM (Filippova et al. 1984). The maximum rate for the enzyme reaction was $9.3 \times 10^{-5} \pm 0.26 \times 10^{-5}$ mM/s^{-1}, and k_{cat} 1.078 ± 0.055 s^{-1}. Therefore, araujiain aII showed a significant k_{cat}/K_m value ($5.99 \times 10^3 \pm 0.57 \times 10^3$ M^{-1} s^{-1}) revealing the effectiveness of this enzyme toward PFLNA. These kinetics values were comparable with those reported by Morcelle et al. (2004) and Trejo (2005) for other latex cysteine peptidases (funastrain cII from latex of *Funastrum clausum*, and asclepain f from latex of *Asclepias fruticosa*, respectively).

As seen in Fig. 4 the highest preferences were shown for Gln and Ala derivatives, followed by Gly, Tyr, Phe and Asp derivatives in decreasing order. Both araujiain aII and asclepain f hydrolyse the same substrates (Trejo et al. 2001). Other peptidases of the Apocynaceae family display different substrate specificity. Asclepain cII and asclepain cI from latex of *A. curassavica* strongly prefer Asp and Tyr derivatives (Liggieri et al. 2004, 2009). Morrenain bI exhibits strong preference for Ala and Asp derivatives while morrenain bII prefers Asp and Gly, both enzymes isolated from latex of *Morrenia brachystephana* (Vairo-Cavalli et al. 2001, 2003). With regard to the enzyme kinetics, araujiain aII had a non-Michaelian behavior with these kinds of substrates.

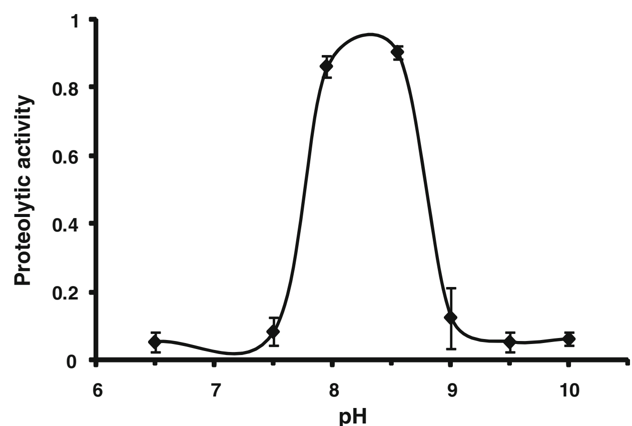


Fig. 3 Araujiain aII pH profile activity with 1% casein as substrate. *Bars* indicate standard deviation from triplicate determinations

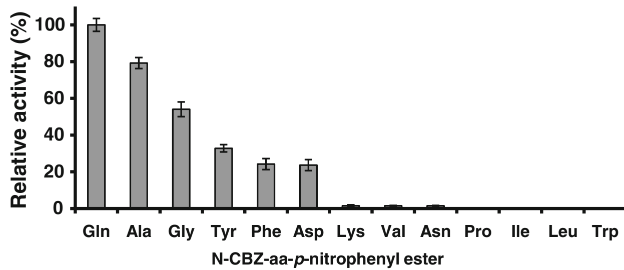


Fig. 4 Araujiain aII esterolytic preference on *N*-CBZ-aa-*p*-nitrophenyl esters. *Bars* indicate standard deviation from triplicate determinations

Native peptidase N-terminal amino acid sequence was determined up to 23 aminoacids by EDMAN sequencing: LPDSVDWRDKGVVFPIRRQGKCG. On the other hand, the araujiain aII peptide map is shown in Fig. 5. Using MASCOT search engine it was not possible to identify this enzyme; therefore, it may be considered a novel proteinase. Protein identification and differentiation by PMF has been adopted in our group as an excellent tool to differentiate, in a fast and unequivocal way, proteases with very similar physicochemical and functional properties (Obregón et al. 2009a; Torres et al. 2010). The robustness of the method even allowed the differentiation of the homologue isoenzymes of the latex of *A. curassavica* (Obregón et al. 2009a).

Peptidase cloning and sequence analysis

Araujiain aII was cloned with the template synthesized from *A. angustifolia* latex RNA and primers Nt_{1-8} and R_0. Identity of PCR products was confirmed by amplification with nested primers $CAapo_1$ and R_1. From nested PCR

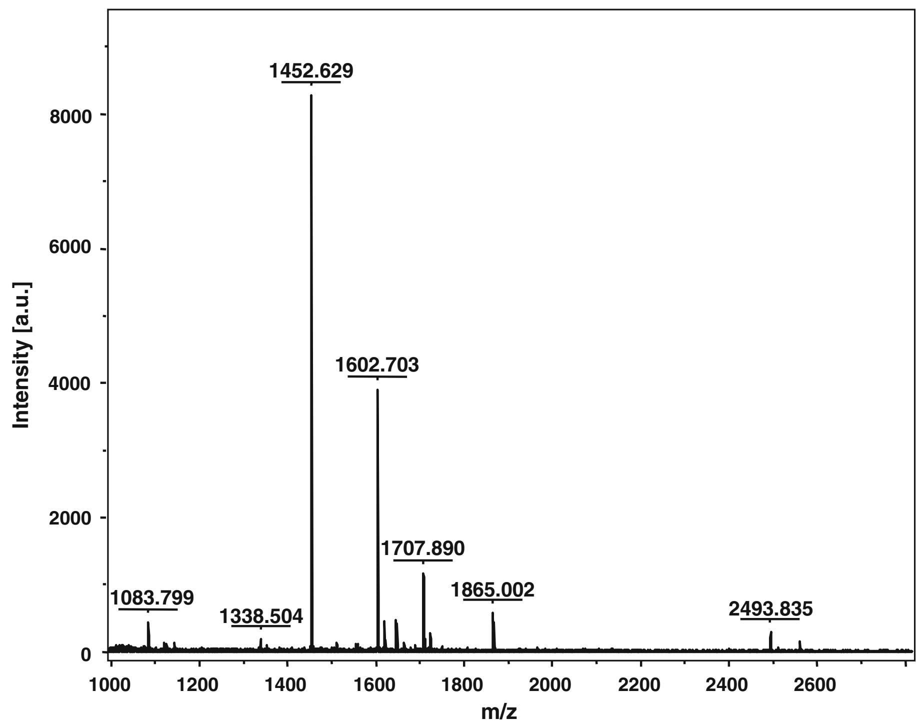


Fig. 5 MALDI-TOF mass spectra of tryptic digests from purified araujiain aII, *m/z* values of prominent peaks are indicated in the *graphic*

Table 1 Identification of trysin-digested peptides of araujiain aII with theoretical PMF

Experimental PMF (*m/z*)	Theoretical PMF (*m/z*)	Fragment position	Peptide fragment sequence
1,337.031	1,337.627	89–100	GQQGQCYQLQK(Carbamidomethyl(C)
1,452.625	1,451.695	64–77	GGFYEAAFTYVAR
1,602.703	1,601.680	175–189	NSWGTTWGEGGYMR
1,707.890	1,706.809	44–58	MIALSEQELLDCER(Carbamidomethyl(C)
1,865.002	1,865.853	136–153	DFQFYQSGLFTGACGPK

Cysteines have been treated with iodoacetamide to form carbamidomethyl-cysteine (Cys-CAM). For theoretical PMF Cys-CAM, oxidized Cys and monoisotopic masses of the occurring amino acid residues were considered. The experimental and theoretical peptide matches were selected with a mass tolerance of 0.6 Da

results, the partial cDNA of the peptidase that corresponds to 700 bp (including the stop codon: TAA, polyadenylation signal and polyA tail) was sequenced and analyzed (GenBank accession number: FN666432). The amino acid sequence deduced from the nucleic acid sequence started from 19 to 213. In order to obtain the full sequence of mature araujiain aII the theoretical mass map of the translated sequence was compared with empirical PMF

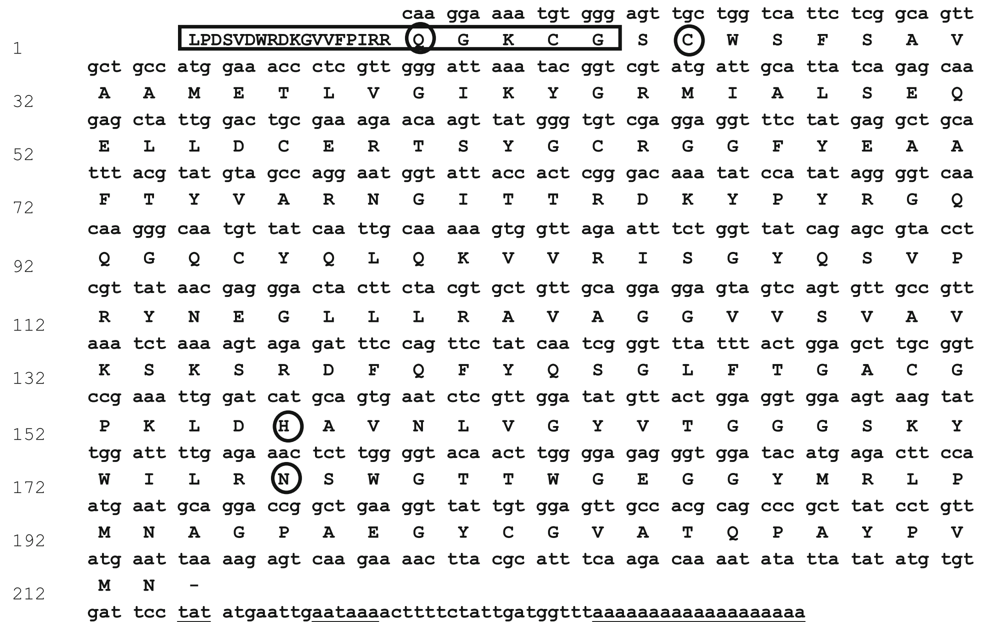


Fig. 6 cDNA sequence and deduced protein sequence of araujiain aII. AaCPII amino acid sequence determined at the protein level is *boxed*. Cys25, His156, Asn176—the catalytic triad—and Glu19, that helps to stabilize the triad, are *circled*. Stop codon, polyadenylation signal and polyA tail are *underlined*

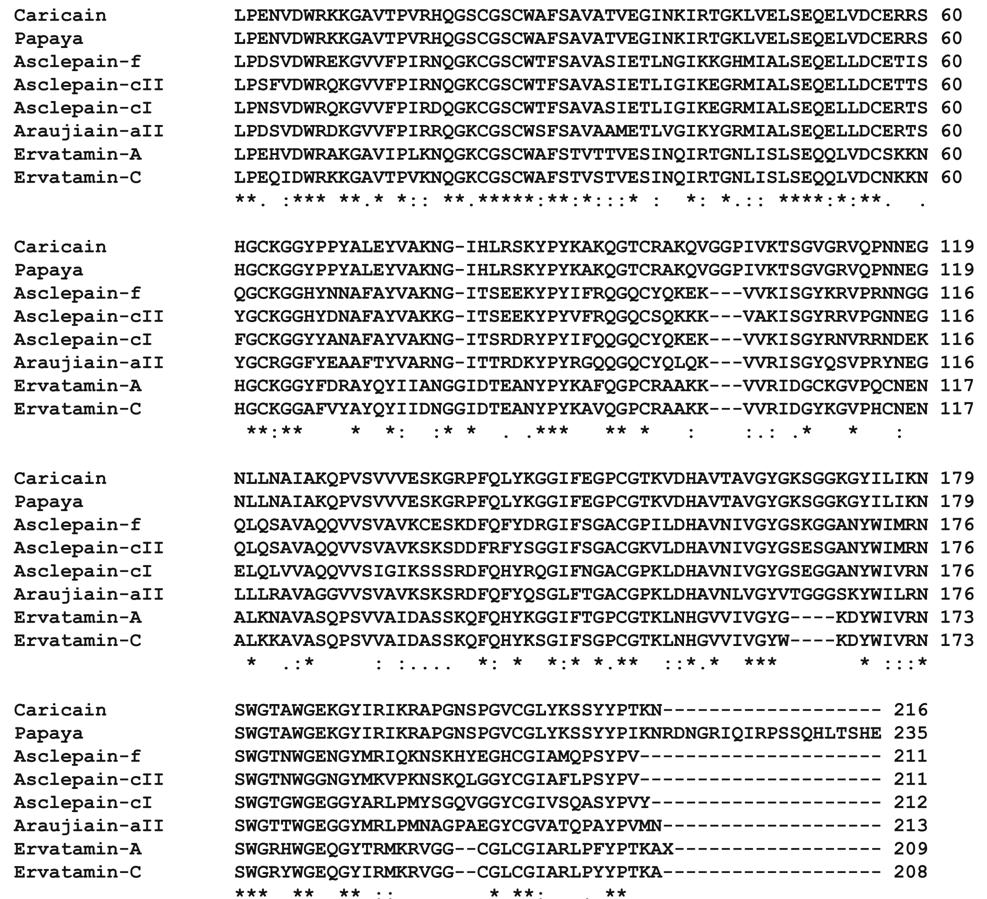


Fig. 7 Multiple sequence alignment by CLUSTAL-W of putative mature protein sequences of araujiain aII against asclepain cI (CAT00687), asclepain cII (CAT00688), asclepain f (B5BLP0), ervatamin A (PDB: 3BCN_A), ervatamin C (PDB: 2PNS_A), and caricain (P10056). *Asterisk* residues in the column are identical in all sequences in the alignment. *Colon* conserved substitutions. *Dot* residues in the column are semi-conserved substitutions

(Table 1). As shown in Table 1, five experimental tryptic fragments match peptides of the putative sequence generated by simulated digestion (mass error tolerance: 0.6 Da); the protein sequence coverage of peptide fragments was 33% with a fine peptide distribution. The proposed sequence for the mature enzyme is shown in Fig. 6, containing 213 aminoacids and including those residues coming from the N-terminus.

The deduced sequence (AaCPII) was analyzed by PSI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi, Altschul et al. 1997), finding 47–70% identity with 43 cysteine peptidases belonging to the C1A family, which groups PLPs. The database search showed similarity to latex peptidases of *C. papaya*: papaya proteinase omega (52% identities, $E = 3 \times e^{-60}$), caricain (52% identities, $E = 9 \times e^{-60}$), glycylendopeptidase (50% identities, $E = 2 \times e^{-54}$), chymopapain (49% identities, $E = 4 \times e^{-58}$), and papain (47% identities, $E = 3 \times e^{-54}$). Figure 7 shows a multiple sequence alignment of putative mature protein sequence of araujiain aII against primary structure of cysteine peptidases with crystal structures solved by X-ray: papaya proteinase omega (PDB: 1PPO) and caricain (PDB: 1MEG_A), both with 52% of identity; ervatamin A (PDB: 3BCN_A) and ervatamin C (PDB: 2PNS_A) both with 49% of identity. On the other hand, Fig. 7 also shows the comparison with asclepain f (B5BLP0), asclepain cI (CAT00687) and asclepain cII (CAT00688), the closest relatives of araujiain aII, with 70, 70 and 67% identity, respectively.

As could be expected, a high conservation degree was observed for those amino acid residues which are essential for catalytic activity (Cys25, His159 and Asn 175 according to mature papain numbering; Beers et al. 2004) and those important for maintaining the tertiary structure. Araujiain aII as well as asclepain cI, asclepain cII (Obregón et al. 2009a), asclepain f (Trejo et al. 2009), philibertain gA and philibertain gB (Sequeiros et al. 2005), all enzymes from latex of Apocynaceae family, presented

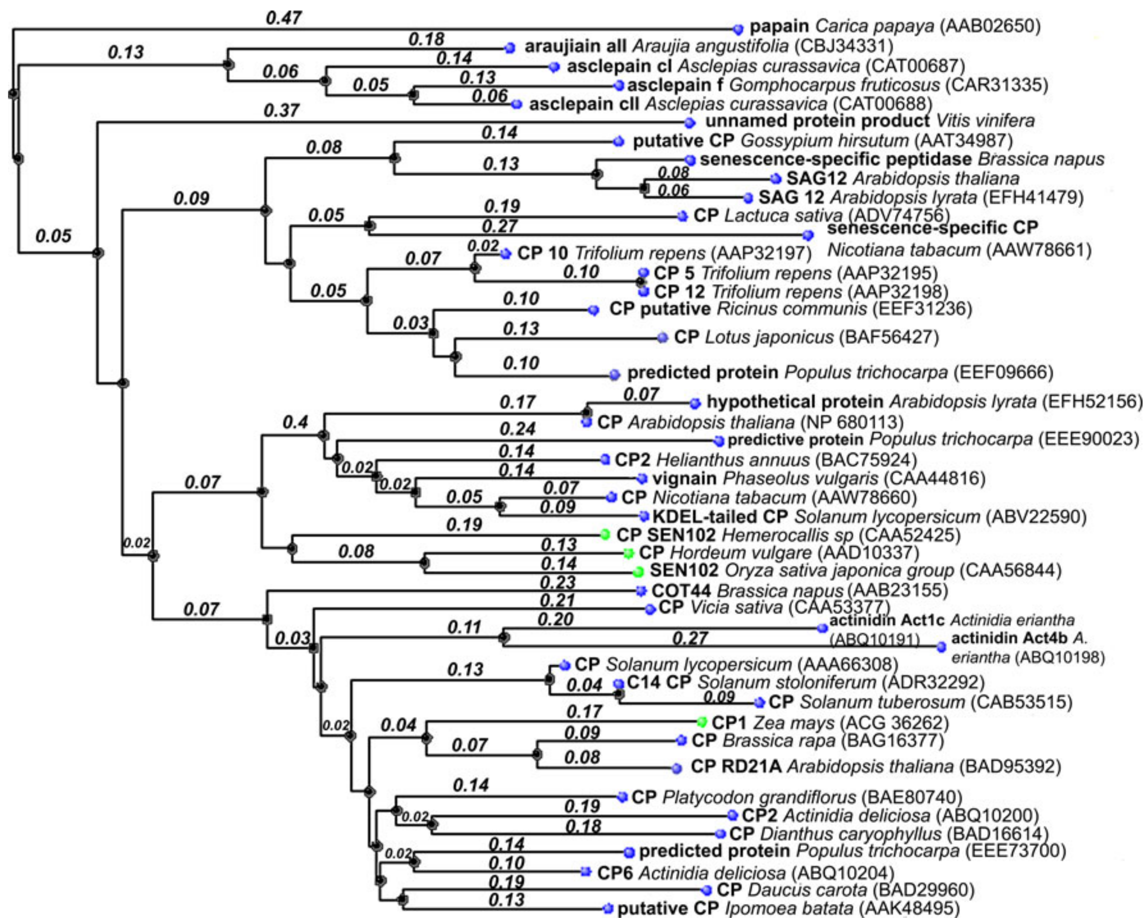


Fig. 8 Distance map of papain-like cysteine peptidases. Distance model was constructed with the PSI-BLAST tool; tree method: neighbor joining, distance according to Grishin (protein) restricted to 0.85. Peptidase names are indicated in **bold** followed by plant species in *italic* and accession numbers are represented in *parenthesis*. The tree includes peptidases most closely related in sequence to araujiain aII, and papain (the subfamily C1A type example). *Green circles* monocots, *blue circles* eudicots. *CP* cysteine protease, *SAG* senescence associated gene

the catalytic Cys at position 25, while His was located at position 156 (Fig. 6).

Neighbor joining analysis (Fig. 8) showed enzymes from Apocynaceae latex as a separate group. Despite their homology, as was discussed previously, each peptidase displayed unique substrate specificity.

AaCPII was analyzed by GPMAW v6.0 to estimate some physicochemical properties: relative molecular mass (23,390.66 Da); 280 nm molar extinction coefficient (48,010 M^{-1} cm^{-1}); and p*I* with reduced and oxidized thiols (9.36 and 10.05, respectively). The latter is slightly more basic than the experimental p*I*: 8.9 (Fig. 9). This experimental p*I* is similar to papain p*I* (8.75) and less basic than p*I* of other Apocynaceae peptidases (Dubey and Jagannadham 2003; Sequeiros et al. 2005; Liggieri et al. 2009). The estimated molecular mass of araujiain aII is fairly close to the value obtained by mass spectrometry and of the same order of molecular masses of other peptidases from Apocynaceae (Obregón et al. 2009a, b; Trejo et al. 2009).

Molecular modeling of araujiain aII

Among all the available molecular structures of PLCPs, papaya proteinase omega from latex of *C. papaya* was identified as the best structural template for modeling the molecular structure of araujiain aII. The root mean square deviation (RMS) of the Cα atoms between the modeled structure and the template structure was within a reasonable range (0.467 Å). The modeled overall structure of araujiain aII is shown in Fig. 10. The fold consisted of two domains (lobes) with the catalytic site lying between them.

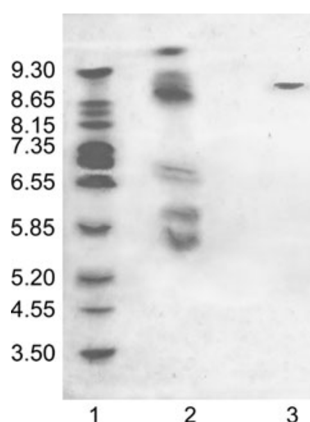


Fig. 9 Isoelectrofocusing pH gradient (3–10) in polyacrylamide gel (10%). Broad p*I* calibration kit (GE Healthcare; amylglucosidase (3.50); trypsin inhibitor (4.55); α-lactoglobulin A (5.20); bovine carbonic anhydrase B (5.85); human carbonic anhydrase B (6.55); myoglobin, acidic band (6.85); myoglobin, basic band (7.35); lentil lectin, acidic band (p*I* 8.15); lentil lectin, middle band (p*I* 8.45); lentil lectin, basic band (p*I* 8.65) and trypsinogen (p*I* 9.30)

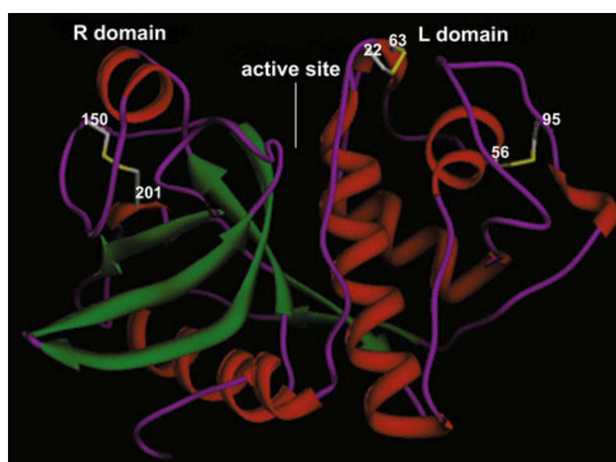


Fig. 10 The modeled overall structure of araujiain aII from *A. angustifolia* latex is shown in cartoon style. R-L domains are indicated and disulfide bridges are represented with *sticks*

The location of residues involved in catalysis is shown in Fig. 11. The L domain (N-terminal domain) composed of α helices is stabilized by the presence of two disulfide bonds and included Gln19 and Cys25. The R domain with a β-barrel structure and a disulfide bond contributes to the active site with His156 and Asn176 residues. PLCPs use Cys25 as nucleophile, activated by His 159 in the active site (papain numbering). Besides Cys and His, Gln19 helps in the formation of the electrophilic center that stabilizes the tetrahedral intermediate; while Asn175 is thought to orientate the imidazolium ring of catalytic His (Barrett and Rawlings 2004). Phe140, Trp178 and Trp182 (141, 177 and 181 according to papain numbering, respectively), involved in the hydrophobic pocket of the catalytic site in which the hydrogen bond between Asn176 and His156 is located, are also conserved in araujiain aII. In addition to the Cys25 residue of the active site, six Cys residues are

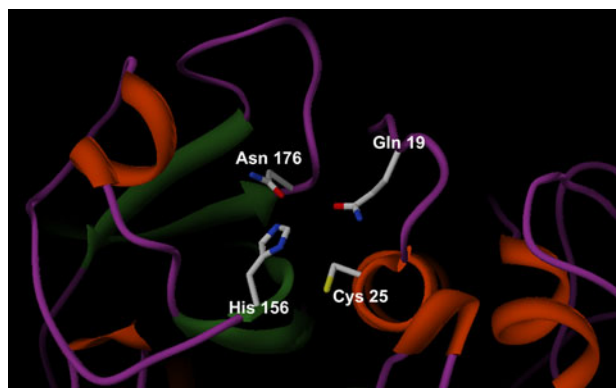


Fig. 11 Structure of catalytic site of araujiain aII, residues 19, 25, 156 and 176 are represented with *sticks*. Cys25 acts as nucleophile, activated by His156 in the active site; Gln19 helps in the formation of the electrophilic center that stabilizes the tetrahedral intermediate, while Asn176 orientates the imidazolium ring of catalytic His

located in araujiain aII at positions 22, 56, 63, 95, 150 and 201, as suggested by modeling probably involved in disulfide bridges(Cys22–Cys63, Cys56–Cys95, Cys150–Cys201). While asclepain f, chymopapain, *Carica candamarcensis* endopeptidase, and ficain have, in addition to the active site cysteine, a free thiol group (Cys133), AaCPII and papain both lack free sulfhydryl (Trejo et al. 2009).

Acknowledgments The present work was supported by grants from ANPCyT, Argentina (PICT 38088 and PICT 02224), University of La Plata, Argentina (Project X-576), and Spanish Ministry of Education and Science (BIO2010-22321-C02-01). WD Obregón and SE Vairo-Cavalli are members of CONICET Researcher Career; CS Liggieri belongs to CIC Support Professional Career Program. D. Lufrano is fellowship of CONICET. The IBB-UAB is a member of ProteoRed, funded by Genoma Spain and follows the quality criteria set up by ProteoRed standards. We are grateful for the excellent technical assistance of Sìlvia Bronsoms, Oscar Conchilla (both belong to the Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona), Analía Lanteri (División Entomología, Museo de La Plata, Facultad de Ciencias Naturales y Museo, UNLP), and Eugenia Ghirimoldi (CIC professional scientific editing service).

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Barrett AJ, Rawlings ND (2004) Introduction: the clans and families of cysteine peptidases. In: Barrett AJ, Rawlings ND, Woessner JF (eds) Handbook of proteolytic enzymes, 2nd edn edn. Elsevier, London, pp 1051–1071

Beers EP, Jones AM, Dickerman AW (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in Arabidopsis. Phytochemistry 65:43–58

Beuth J (2008) Proteolytic enzyme therapy in evidence-based complementary oncology: fact or fiction? Integr Cancer Ther 7(4):311–316

Bradford MM (1976) A rapid and sensitive method for the quantitation of micrograms quantities of protein utilizing the principle of protein dye binding. Anal Biochem 72:248–254

Bucciarelli A, Cambi VN, Villamil CB (2008) Morphoanatomical characters of *Araujia hortorum* E. Fourn (Asclepiadaceae), a native species of medicinal interest. Phyton (B. Aires). http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-56572008000100024. Accessed 19 Oct 2010

Burkart A (1979) Flora ilustrada de Entre Ríos (Argentina): Parte V. Dicotiledóneas Metaclamídeas (Gamopétalas) Colección Científica del INTA. Tomo VI–V. Buenos Aires, 173–175

Domsalla A, Melzig MF (2008) Occurrence and properties of proteases in plant latices. Plant Med 74:699–711

Dubey VK, Jagannadham MV (2003) Procerain, a stable cysteine protease from the latex of *Calotropis procera*. Phytochemistry 62(7):1057–1071

Endress ME, Bruyn PV (2000) A revised classification of the Apocynaceae. Bot Rev 66:1–56

Filippova IY, Lysogorskaya EN, Oksenoit ES, Rudenskaya GN, Stepanov VM (1984) L-Pyroglutamyl-L-phenylalanyl-L-leucine-*p*-nitroanilide: a chromogenic substrate for thiol proteinase assay. Anal Biochem 143:293–297

Frohman MA (1994) On beyond classic RACE (rapid amplification of cDNA ends). PCR Methods Appl 4(1):40–58

Hagel JM, Yeung EC, Facchini PJ (2008) Got milk? The secret life of laticifers. Trends Plant Sci 13:631–639

Khaparde SS, Singhal RS (2001) Chemically modified papain for applications in detergent formulations. Bioresour Technol 78(1):1–4

Liggieri CS, Arribére MC, Trejo SA, Canals F, Avilés F, Priolo NS (2004) Purification and biochemical characterization of asclepain c I from the latex *Asclepias curassavica* L. J Protein 23:403–411

Liggieri C, Obregón WD, Trejo SA, Priolo N (2009) Biochemical analysis of a papain-like protease isolated from the latex of *Asclepias curassavica* L. Acta Biochim Biophys Sin 41(2):154–162

Morcelle S, Caffini NO, Priolo NS (2004) Proteolytic properties of *Funastrum clausum* latex. Fitoterapia 75:480–490

Morcelle SR, Liggieri CS, Bruno MA, Priolo N, Clapés P (2009) Screening of phytoproteases for the synthesis of arginine-based surfactants. J Mol Catal B Enzyme 57:177–182

O'Connell K, Stults J (1997) Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests. Electrophoresis 18:349–359

Obregón WD, Arribére MC, Morcelle del Valle SR, Liggieri CS, Caffini NO, Priolo NS (2001) Two new cysteine endopeptidases obtained from the latex of *Araujia hortorum* fruits. J Protein Chem 20:17–25

Obregón WD, Curciarello R, Caffini NO, Priolo NS (2006) Hidrolytic profile and isolation of the proteolytic components of latex from *Araujia angustifolia* fruits. Acta Farm Bonaerense 25(2):206–212

Obregón WD, Liggieri CS, Morcelle del Valle SR, Trejo SA, Avilés XF, Priolo NS (2009a) Biochemical and PMF MALDI-TOF analyses of three novel papain-like plant proteinases. Protein Pept Lett 16:1323–1333

Obregón WD, Liggieri CS, Trejo SA, Avilés FX, Vairo-Cavalli SE, Priolo NS (2009b) Characterization of papain-like isoenzymes from latex of *Asclepias curassavica* by molecular biology validated by proteomic approach. Biochimie 91(11–12):1457–1464

Priolo NS, López LMI, Arribere MC, Natalucci CL, Caffini NO (1991) New purified plant proteinases for the food industry. Acta Alimentaria 20:189–196

Priolo N, Morcelle del Valle S, Arribére MC, López LMI, Caffini N (2000) Isolation and characterization of a cysteine protease from the latex of *Araujia hortorum* fruits. J Protein Chem 19:39–49

Ramos MV, Pereira DA, Souza DP, Araújo ES, Freitas CDT, Cavalheiro MG, Matos MPV, Carvalho AFU (2009) Potential of laticifer fluids for inhibiting *Aedes aegypti* larval development: evidence for the involvement of proteolytic activity. Mem Inst Oswaldo Cruz 104(6) Rio de Janeiro. doi:10.1590/S0074-02762009000600001

Rasmann S, Johnson MD, Agrawal AA (2010) Induced responses to herbivory and jasmonate in three milkweed species. J Chem Ecol 35(11):1326–1334

Salas CE, Gomes TR, Hernandez M, Lopes MTP (2008) Plant cysteine proteinases: evaluation of the pharmacological activity. Phytochemistry 69(12):2263–2269

Salvesen GS, Nagase H (2001) Inhibition of proteolytic enzymes. In: Beynon R, Bond JS (eds) Proteolytic enzymes, 2nd edn. Oxford University Press, Oxford

Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate polyacrilamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166:368–379

Sequeiros C, Torres MJ, Trejo SA, Esteves JL, Natalucci CL, López LMI (2005) Philibertain g I, the most basic cysteine endopeptidase purified from the latex of *Philibertia gilliesii* Hook. et Arn. (Apocynaceae). Protein J 24:445–453

Shindo T, van der Hoorn RA (2008) Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. Mol Plant Pathol 9(1):119–125

Shivaprasad HV, Rajesh R, Nanda BL, Dharmappa KK, Vishwanath BS (2009) Thrombin like activity of *Asclepias curassavica* L. latex: action of cysteine proteases. J Ethnopharmacol 123(1): 106–109

Silverstein RM (1974) The assay of the bromelains using *N*-CBZ-L-lysine *p*-nitrophenyl ester and *N*-CBZ-L-glycine *p*-nitrophenyl ester as substrates. Anal Biochem 62:478–484

Singh AN, Shukla AK, Jagannadham MV, Dubey VK (2010) Purification of a novel cysteine protease, procerain B, from *Calotropis procera* with distinct characteristics compared to procerain. Process Biochem 45:399–406

Stepek G, Behnke JM, Buttle DJ, Duce IR (2004) Natural plant cysteine proteinases as anthelmintics? Trends Parasitol 20(7): 322–327

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-speciWc gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Torres MJ, Trejo SA, Martín MI, Natalucci CL, Avilés FX, López LMI (2010) Purification and characterization of a cysteine endopeptidase from *Vasconcellea quercifolia* A. St.-Hil. Latex displaying high substrate specificity. J Agric Food Chem 58:11027–11035

Trejo SA (2005) Purificación, Caracterización Bioquímica y Estructural y Expresión de una Endopeptidasas Cisteínicas de Látex de *Asclepias fruticosa* L. (Apocynaceae). Ph.D. thesis, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina

Trejo SA, López LMI, Cimino CV, Caffini NO, Natalucci CL (2001) Purification and characterization of a new plant endopeptidase isolated from the latex of *Asclepias fruticosa* L. (Asclepiadaceae). J Protein Chem 20:469–477

Trejo SA, López LMI, Caffini NO, Natalucci CL, Canals F, Avilés FX (2009) Sequencing and characterization of asclepain f: the first cysteine peptidase cDNA cloned and expressed from *Asclepias fruticosa* latex. Planta 230:319–328

Vairo-Cavalli SE, Cortadi A, Arribére MC, Conforti P, Caffini NO, Priolo NS (2001) Comparison of two cysteine endopeptidases from latices of *Morrenia brachystephana* Griseb. and *Morrenia odorata* (Hook et Arn.) Lindley (Asclepiadaceae). Biol Chem 382:879–883

Vairo-Cavalli S, Arribére MC, Cortadi A, Caffini NO, Priolo NS (2003) Morrenain b I, a papain-like endopeptidase from the latex of *Morrenia brachystephana* Griseb. (Asclepiadaceae). J Protein Chem 22:15–22

van der Hoorn RAL (2008) Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol 59:191–223