Activation, Proteolytic Processing, and Peptide Specificity of Recombinant Cardosin A*

Pedro Castanheira‡‡, Bart Samyn††, Kjell Sergeant‡‡, José C. Clemente‡‡, Ben M. Dunn‡‡, Euclides Pires‡‡, Jozef Van Beeumen‡‡, and Carlos Faro§§§

From the ‡‡Departamento de Bioquímica, Faculdade de Ciências e Tecnologia, Universidade de Coimbra, Apartado 3126, 3001-401 Coimbra, Portugal, and the ††Department Biochemistry, Physiology and Microbiology, University Gent, Laboratory of Protein Biochemistry and Protein Engineering, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium, and the ‡‡‡Department of Biochemistry and Molecular Biology, University of Florida College of Medicine, Gainesville, Florida 32610-0245

Cardosins are model plant aspartic proteases, a group of proteases that are involved in cell death events associated with plant senescence and stress responses. They are synthesized as single-chain zymogens, and subsequent conversion into two-chain mature enzymes is a crucial step in the regulation of their activity. Here we describe the activation and proteolytic processing of recombinant procardosin A. The cleavage sites involved in this multi-step autocatalytic process were determined, some of them using a novel method for C-terminal sequence analysis. Even though the two-chain recombinant enzyme displayed similar properties as natural cardosin A, a single-chain mutant form was engineered based on the processing results and produced in Escherichia coli. Determination of its primary specificity using two combinatorial peptide libraries revealed that this mutant form behaved like the natural enzyme. The primary specificity of the enzyme closely resembles those of cathepsin D and plasminogen, suggesting that cardosin A shares the same peptide scissile bond preferences of its vacuolar/lysosomal mammalian and protozoan homologues.

Sequencing of the Arabidopsis genome unveiled over 50 genes coding for aspartic proteases of the pepsin type (1, 2). These genes have been assigned to three different categories: typical, nucellin-like, and atypical proteases (2). With the exception of the products of cnd41 (3–5) and cdr1 (6) genes, very little is known about the nucellin and the atypical subgroups. In fact, most of the knowledge acquired during the last decade is associated with plant senescence and stress responses. They display activity at acidic pH and are inhibited by pepstatin A, a natural hexapeptide from Streptomyces. Common features also include the overall three-dimensional structure of mature enzymes, the catalytic apparatus, and a conserved DTSGI motif. A distinguishing feature of typical plant aspartic proteases is the occurrence of an extra 100-amino acid-long internal segment, known as the plant-specific insert (PSI), in the sequence of their precursors. This domain displays structural and functional similarities to saposins, sphingolipid-activating proteins from animals, and some antimicrobial peptides such as NK-lysin, granulysin, and amoebapores. It folds as an independent domain in the precursor form and is subsequently excised to generate the mature two-chain form. Even though the function of PSI1 is not yet fully elucidated, it seems to play an important role in targeting plant aspartic protease precursors to the vacuole.

Little is known regarding the function of typical plant aspartic proteases. Colocalization studies with putative substrates and their temporal and spatial expression profiles have implicated members of this group in cell death events associated with plant senescence, stress responses, programmed cell death, and plant sexual reproduction. In the particular case of cardosin A, the major aspartic protease from the flowers of cardoon, gene expression occurs specifically at the early stages of flower development, and the protein accumulates in protein storage vacuoles until the later stages of flower senescence. As the flower matures, the single-chain precursor form (procardosin A) is converted into a two-chain mature cardosin A by removal of the internal PSI domain and the N-terminal propeptide (14). This event is a crucial step in regulating the activity of cardosin A. Therefore, an improved understanding of the activation mechanism, along with a more detailed analysis of the specificity of cardosin A, may give more clues about the molecular mechanism underlying its participation in the cell death events associated with flower senescence.

MATERIALS AND METHODS

The expression vector pET-23a (Novagen, Madison, WI) containing the procardosin A cDNA (pET_pCA-PsiI) has been previously constructed in our lab (14), as well as the PSI deleted procardosin A construct (pET_pCA-PsiI) (8). The Escherichia coli BL21(DE3) strain was pur-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Recipient of a Ph.D. fellowship from the PRAXIS XXI program, Fundação para a Ciência e a Tecnologia, Ministério da Ciência e Tecnologia.

¶ Postdoctoral Fellow of the Fund for Scientific Research-Flanders (F.W.O.-Vlaanderen, Belgium).

** Supported by a Ph.D. grant from the Institute for the promotion of Innovation through Science and Technology in Flanders (I.W.T.-Vlaanderen).

†§ To whom correspondence should be addressed: Dept. de Bioquímica, Faculdade de Ciências e Tecnologia, Universidade de Coimbra, Apartado 3126, 3000 Coimbra, Portugal. Tel.: 351239480210; Fax: 351239480217; E-mail: cfaro@imagei.ibil.uc.pt.
chased from Novagen. Prestained protein markers were purchased from Bio-Rad. The synthetic chromogenic peptide Lys-Pro-Ala-Glu-Phe-Phe(NO\textsubscript{2})\textsubscript{3}-Ala-Leu was purchased from Pepsyn Ltd. (University of Liverpool, Biological Sciences). Natural Cardosin A was purified from fresh flowers of \textit{Cynara cardunculus} L. as previously described (15).

**Expression, Refolding, and Purification of Procardosin A in E. coli—** 
\textit{pET-pCA-PSI} and \textit{pET-pCA-PSI} were introduced into \textit{E. coli} BL21 (DE3), and expression of procardosin A (pCA) or its mutant without the PSI (pCA-PSI) was induced by the addition of isopropyl-1-thio-\beta-D-galactopyranoside (final concentration, 0.5 mM) when the \textit{OD\textsubscript{600}} of the cells, grown at 37 °C, had reached 0.6. After 3 h, the cells were harvested, resuspended in 50 mM Tris-HCl, 50 mM NaCl (pH 7.4), and lysed by adding lysozyme (100 \mu g/ml). After freezing and thawing, deoxyribonuclease I (100 \mu g/ml) and MgCl\textsubscript{2} (100 mM) were added and incubated at 4 °C for 1 h. The inclusion bodies were then washed for 3 h with 50 mM Tris-HCl, 50 mM NaCl (pH 7.4), centrifuged at 10,000 \times g for 20 min at 4 °C, and then washed again for another 3 h with 50 mM Tris-HCl, 50 mM NaCl (pH 7.4). 0.1% Triton X-100 (v/v). Upon centrifugation at 10,000 \times g for 20 min at 4 °C, the purified inclusion bodies were dissolved in 8 mM urea containing 100 mM 2-mercaptoethanol. The protein was refolded by rapid dilution (20-fold) into 20 mM Tris base, and 124 mM diisopropyl ethylamine/ACN to derivatize the bovine \textit{PC1} (100 \textmu M) or without PSI) samples were first incubated with the activation buffers (0.1 M sodium acetate, pH 5.0; 0.1 M sodium phosphate, pH 6.0; and 0.1 M Tris-HCl, pH 8.0). After incubation at 37 °C for 1 h, the samples were then concentrated in a tangential flow ultrafiltration system (Pellicon 2; Millipore) and then purified by ion exchange chromatography on a Resource Q column (Amersham Biosciences) using a gradient of 0–0.5 M NaCl. Polyacrylamide Gel Electrophoresis—SDS-PAGE was performed using 15% gels in a Bio-Rad Mini Protean III electrophoresis apparatus and stained with Coomassie Brilliant Blue R-250 (Sigma). After SDS-PAGE, the protein samples for chemical N- and C-terminal sequence analysis were transferred onto polyvinylidene difluoride membranes by electrophoretic transfer (4). Approximately 10% mass of the sample was loaded onto an S-300 gel filtration chromatographic column (Amersham Biosciences) equilibrated in 20 mM Tris-HCl, 0.4 mM urea, pH 8.0 buffer. The fractions corresponding to the second protein peak, which corresponds to the nonaggregated forms of cardosin A, were then combined and further purified by ion exchange chromatography on a Resource Q column (Amersham Biosciences) using a gradient of 0–0.5 M NaCl.

**Activation and Enzymatic Properties—** Purified procardosin A (with or without PSI) samples were first incubated with the activation buffers in a 1:1 volume ratio. The samples were incubated overnight at 37 °C with the following buffers: 0.1 mM sodium citrate, pH 5.0 and 4.0; 0.1 mM sodium acetate, pH 5.0; 0.1 mM sodium phosphate, pH 6.0; and 0.1 mM Tris-HCl, pH 8.0. The samples were then assayed for activity and analyzed by SDS-PAGE.

**N- and C-Terminal Sequence Analysis of Electroblotted Samples—** N-terminal sequence analysis was performed in the gas-pulsed liquid phase using a model 476A protein sequencer with a microreaction chamber and an on-line HPLC system for phenylthiohydantoin analysis (Applied Biosystems, Foster City, CA). Absorbance was monitored at 300 nm in a Varian-Cary 100 UV-visible spectrophotometer using the operating software. A molar absorption coefficient of 1480 \textmu M cm\textsuperscript{-1} at 300 nm was used in the calculations (15). The enzyme concentration was determined by active site titration with pepstatin A. For activity studies at different pH, the following buffers were used: 0.1 M sodium citrate, pH 2.5–4.0; 0.1 M sodium phosphate, pH 4.5–5.5; 0.1 M sodium carbonate, pH 8.5–9.5; 0.1 M Tris-HCl, pH 7.0. The kinetic parameters were calculated from the Lineweaver-Burk plot using appropriate software, and the inhibition constant for pepstatin A was calculated as previously described (15).

For the activation studies at different protein concentrations, procardosin A was incubated at concentrations ranging from 0.1 to 1.44 \textmu M in 50 mM sodium acetate, pH 4.7, 0.2 mM NaCl, and the rate of hydrolysis was monitored at 300 nm in a Varian-Cary 100 UV-visible spectrophotometer using the operating software. A molar absorption coefficient of 1480 \textmu M cm\textsuperscript{-1} at 300 nm was used in the calculations (15). The enzyme concentration was determined by active site titration with pepstatin A. For activity studies at different pH, the following buffers were used: 0.1 M sodium citrate, pH 2.5–4.0; 0.1 M sodium phosphate, pH 4.5–5.5; 0.1 M sodium carbonate, pH 8.5–9.5; 0.1 M Tris-HCl, pH 7.0. The kinetic parameters were calculated from the Lineweaver-Burk plot using appropriate software, and the inhibition constant for pepstatin A was calculated as previously described (15).

Specificity Preferences of Recombinant and Native Cardosin A—The peptide substrate specificity was studied by using two combinatorial libraries based on the sequences Lys-Pro-Xaa-Glu-Phe-Phe(NO\textsubscript{2})\textsubscript{3}-Ala-Leu as substrate. The enzymes were incubated at 37 °C with 0.1 mM substrate in 50 mM sodium acetate, pH 4.7, 0.2 mM NaCl, and the rate of hydrolysis was monitored at 300 nm in a Varian-Cary 100 UV-visible spectrophotometer using the operating software. A molar absorption coefficient of 1480 \textmu M cm\textsuperscript{-1} at 300 nm was used in the calculations (15). The enzyme concentration was determined by active site titration with pepstatin A. For activity studies at different pH, the following buffers were used: 0.1 M sodium citrate, pH 2.5–4.0; 0.1 M sodium phosphate, pH 4.5–5.5; 0.1 M sodium carbonate, pH 8.5–9.5; 0.1 M Tris-HCl, pH 7.0. The kinetic parameters were calculated from the Lineweaver-Burk plot using appropriate software, and the inhibition constant for pepstatin A was calculated as previously described (15).

Specificity Preferences of Recombinant and Native Cardosin A—The peptide substrate specificity was studied by using two combinatorial libraries based on the sequences Lys-Pro-Xaa-Glu-Phe-Phe(NO\textsubscript{2})\textsubscript{3}-Ala-Leu as substrate. The enzymes were incubated at 37 °C with 0.1 mM substrate in 50 mM sodium acetate, pH 4.7, 0.2 mM NaCl, and the rate of hydrolysis was monitored at 300 nm in a Varian-Cary 100 UV-visible spectrophotometer using the operating software. A molar absorption coefficient of 1480 \textmu M cm\textsuperscript{-1} at 300 nm was used in the calculations (15). The enzyme concentration was determined by active site titration with pepstatin A. For activity studies at different pH, the following buffers were used: 0.1 M sodium citrate, pH 2.5–4.0; 0.1 M sodium phosphate, pH 4.5–5.5; 0.1 M sodium carbonate, pH 8.5–9.5; 0.1 M Tris-HCl, pH 7.0. The kinetic parameters were calculated from the Lineweaver-Burk plot using appropriate software, and the inhibition constant for pepstatin A was calculated as previously described (15).
RESULTS

Expression, Purification, and Characterization of Recombinant Procardosin A—The expression plasmid construct used in this work contains the cDNA encoding procardosin A starting at Ser3P (pepsin numbering) with an extra Met-Ala-Ser sequence at the N terminus, encoded by the pET-23a sequence. Expression of this construct in E. coli resulted in the accumulation as inclusion bodies of a protein with an apparent molecular mass of about 53 kDa (data not shown); the inclusion bodies were easily purified to greater than 90% purity with a simple freeze/thaw and detergent washing procedure, and after dissolving them in 8 M urea, the refolding was induced by a rapid dilution step followed by the slow adjustment of the pH to 8.0. Upon refolding, the high molecular mass soluble aggregates were removed by size exclusion chromatography in a Sephacryl S-300 column, and recombinant procardosin A was then purified by ion exchange chromatography on a Resource Q column. SDS-PAGE analysis of the protein fraction collected from the ion exchange chromatography confirmed the presence of a protein with an apparent molecular mass of 53 kDa (data not shown), the predicted mass for the nonglycosylated form of procardosin A. Edman degradation of the electroblotted protein indicated that the recombinant procardosin started with the expected N-terminal ASSDD sequence (the initial Met residue indicated that the recombinant procardosin A corresponded to full-length procardosin A. Edman degradation of the electroblotted protein not shown), the predicted mass for the nonglycosylated form of a protein with an apparent molecular mass of 53 kDa (data not shown), the predicted mass for the nonglycosylated form of procardosin A. The expression and purification procedure yielded about 4 mg of pure recombinant procardosin A.

Activation of Recombinant Procardosin A—To investigate the activation of recombinant procardosin A, the protein was incubated in a series of buffers with pH ranging from 3.0 to 7.0. Conversion of recombinant procardosin A into its intermediate and mature forms was observed at acidic pH values, with the highest rate of processing obtained at pH 4.0 (Fig. 1). A time course analysis of the activation process using the chromogenic substrate Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu as substrate revealed that maximal activity was obtained after 30 min of incubation (Fig. 2). SDS-PAGE analysis showed that at this time point full-length procardosin A was converted into a protein with an apparent molecular mass of 47 kDa. As the incubation time increases, this protein is converted into the two-chain form of cardosin A (Fig. 3).

To study the nature of the activation process, recombinant procardosin A was incubated at pH 4.0 in the presence of various protease inhibitors. SDS-PAGE analysis showed that the processing enzyme is an aspartic protease because only pepstatin A (Fig. 4, lane 1), the classic inhibitor of this class of proteases, was able to prevent the proteolytic cleavage of the recombinant procardosin A. To rule out the presence of a contaminant pepstatin-sensitive protease from E. coli, an active site D32A mutant of procardosin A was generated and produced. The inactive mutant could not be processed in the same conditions used to process procardosin A (data not shown), clearly ruling out the possibility of a contaminant protease responsible for the proteolytic events leading to zymogen activation. Taken together these results suggest that recombinant procardosin A activates itself at acidic pH values.

The activation of recombinant procardosin A may result from either an intramolecular or intermolecular processing mechanism or a combination of both. An intramolecular activation mechanism would imply a zero order reaction where the rate of procardosin A activation would be independent of procardosin
A concentration. An intermolecular activation mechanism would have a higher order of reaction.

The relative rates of activation of procardosin A at different protein concentration were determined, and the results are given in Fig. 5. As observed, the extrapolation of activation rate when the procardosin A concentration tends toward zero is clearly different from zero, in agreement with an intramolecular mechanism. However, when procardosin A concentration increases, the rate of activation increases, suggesting that the processing reaction can also follow an intermolecular mechanism.

When fully processed recombinant cardosin A was added to procardosin A in a 1:50 or 1:10 ratio, the initial rate of procardosin A proteolytic processing was not significantly affected (data not shown), indicating that under the conditions used, procardosin A processing appears to be primarily an intramolecular reaction.

**Proteolytic Processing of Recombinant Procardosin A**—To better understand the molecular events involved in the proteolytic activation of recombinant procardosin A, the peptide bonds cleaved during the process were identified by a combination of N- and C-terminal sequencing together with the use of a novel mass spectrometry based method for C-terminal sequence analysis. Although considerable progress has been made in recent years, the existing chemical method for C-terminal sequence analysis is still hampered by the relatively modest sensitivity (20–100 pmol) of this approach (17). Therefore, the C-terminal sequence was determined by using a novel, mass spectrometry-based method that was applied on the SDS-PAGE-separated fragments (21). In this method, a peptide mixture is generated by cleavage of the protein with CNBr in acid solution, a chemical cleavage method known to be specific for Met-Xaa peptide bonds. During the cleavage Met residues are converted to homoserine lactone, in equilibrium with homoserine. During subsequent incubation with carboxypeptidases only the original carboxyl containing fragment is accessible to enzymatic degradation. In this ladder sequencing approach, the sequence can be read, in the correct order, by

Sequence analysis of the 8-kDa protein revealed that the N terminus was that of the recombinant procardosin A (ASSD), whereas C-terminal terminal sequence analysis identified the sequence GF. These results indicate that the first intermediate form of procardosin A results from a cleavage between residues Phe40P-Arg41P (pepsin numbering) of the propeptide. The S-S bonds between the two subunits. Question marks indicate nondetermined sequences. B, SDS-PAGE analysis of recombinant procardosin A activation (activation was achieved by incubation of recombinant procardosin A with 100 mM citrate buffer, pH 4.0, at 37 °C). The band letters correspond to the scheme of A. C, SDS-PAGE analysis of fully processed recombinant procardosin A (72 h) denatured in the presence and absence of β-mercaptoethanol (β-ME). Both gels were stained with Coomassie Blue.
simply calculating the differences in mass of adjacent peptide peaks, representing the loss of C-terminal amino acids. The ladder read-out is performed using MALDI-time-of-flight MS, because this ionization technique produces ladders of singly charged ions and because of its relatively good salt tolerance (carboxypeptidase buffer). This approach is illustrated in Fig. 7 where it was applied on the B1 band. Upon alkylation of the Cys residues and CNBr cleavage, three fragments were observed in the mass spectrum. Upon incubation with carboxypeptidases, only one fragment (at \( m/z \) 2240.16) started to disappear and formed a ladder, from which the C-terminal sequence, GNLLVGFAEAA, was determined. MS-based C-terminal sequence analysis of the C1 band identified the sequence KLCTFDGARDVSSIIE as the C terminus of the 33-kDa protein. This indicates that cleavage occurs between residues Glu50 and Ser51 in the PSI domain.

Application of the MS-based procedure on the D1 gel spot resulted in a mass spectrum wherein only two peptide fragments were observed. Upon incubation with carboxypeptidases, both fragments (\( m/z \) 1233.68 and 1396.73) started to shift and formed a ladder (VS(R/Y)). This indicated that both fragments were the same C-terminal peptide with a ragged C terminus R(Y) (\( \Delta m = 163.06 \) Da) being produced by cleavage of the peptide bonds at Tyr13-Gly14 and Arg12-Tyr13 in the PSI domain. N-terminal sequence analysis showed that the sequence RGTVR is present at the N-terminal side of the D1 fragment, corresponding to the 31-kDa polypeptide of cardosin A (heavy chain) having the last 5 amino acids of the propeptide and the first 12/13 amino acids of the PSI still attached (Fig. 6A).

After CNBr cleavage, MALDI analysis of the less intense C2 and D2 band indicates the presence of a fragment at 2240.23 Da. The low intensity of the observed peak excluded incubation with carboxypeptidases Y and P. However, the mass is in perfect agreement with the calculated mass of the C-terminal fragment Arg109–Ala120 (pepsin numbering) of the intact protein (molecular mass, 2239.11 Da). As shown by the C-terminal sequence analysis of the larger B1 fragment, this indicates that the processing does not occur at the C terminus of the C2 or D2 fragments but at their N termini somewhere in the PSI domain. However, further experimental evidence will be required to identify the exact processing sites.

These results showed that the processed form of recombinant cardosin A is formed by the polypeptide chains corresponding to the D1 and D2 bands. The SDS-PAGE analysis of mature recombinant cardosin A was repeated, and the sample was denatured in the presence or absence of \( \beta \)-mercaptoethanol (Fig. 6C); the nonreduced sample showed a shift of the protein band to higher molecular mass in comparison with the heavy chain protein band of the reduced sample. The shift was estimated to be 11 kDa (from 26 to 37 kDa). This clearly indicates that this two-chain form of recombinant cardosin A remains associated by a disulfide bridge between Cys5 and Cys100 of the PSI domain that was not completely excised from the protein.

Expression, Purification, and Activation of a Single-chain Recombinant Cardosin A—Because removal of the propeptide by autoactivation and excision of PSI turned out to be a more complex event, a PSI deletion mutant of procardosin A was engineered and produced in \( E. \ coli \) using the same expression and purification procedure as described for recombinant procardosin A. A recombinant protein with an apparent molecular mass of 43 kDa was obtained, in agreement with the expected molecular mass for the mutant form of procardosin A. This precursor form was inactive, and activation was achieved under the conditions previously used for recombinant procardosin.
FIG. 8. Schematic representation of the active forms generated from recombinant procardosin A. proCA, recombinant procardosin A; rCA + PSI, intermediate form generated from recombinant procardosin A upon incubation at pH 4.0 for 30 min, which corresponds to a PSI-containing form without the propeptide; rCA, two-chain form of recombinant cardosin A with the polypeptide chains held together by a disulfide bridge (S-S); rCAΔPSI, engineered single-chain recombinant cardosin A activated upon incubation at pH 4.0.

A. Time course analysis revealed that maximum activity was obtained upon incubation in 0.1 M sodium citrate, pH 4.0, at 37 °C for 150 min. SDS-PAGE analysis showed that at this time point the mutant form was totally converted into a 37-kDa protein. No further processing or inactivation of the single-chain recombinant cardosin A was observed for at least 3 days. Conversion was not observed when the recombinant protein was incubated under the same experimental conditions in the presence of pepstatin A. Sequencing of the 37-kDa processed form revealed that the only proteolytic event that occurred during its autoactivation is the incomplete removal of the propeptide.

Enzymatic Properties of the Different Forms of Recombinant Cardosin A—In the course of this work, at least three different forms of active recombinant cardosin A were produced (Fig. 8): (i) an intermediate form generated from recombinant procardosin A upon incubation at pH 4.0 for 30 min, which corresponds to a PSI-containing form without the propeptide (rCA + PSI); (ii) a two-chain form of rCA with the polypeptide chains held together by disulfide bridges; and (iii) an engineered single-chain recombinant cardosin A (rCAΔPSI). To characterize the proteolytic activity of these three forms, their enzymatic properties were determined and compared with those of natural cardosin A. The proteolytic activity assays were carried out using the synthetic chromogenic peptide Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu as substrate.

The temperature and pH dependence of natural and recombinant forms of cardosin A are shown in Fig. 9. Similar profiles were obtained in all cases. The proteolytic activity increases with the temperature with a maximum between 40 and 45 °C and then decreases abruptly for temperatures over 50 °C (Fig. 9A). In respect to their pH optimum, natural and recombinant forms of cardosin A are active between pH 2.5 and 6.0, with a maximal activity achieved between pH 4.0–4.5 (Fig. 9B).

The activity of the recombinant cardosin A forms is inhibited by the specific inhibitor of aspartic proteases, pepstatin A. Inhibitors of other classes of proteases (20 mM phenylmethylsulfonyl fluoride, 0.1 mM E-64, 20 mM EDTA, or 40 mM Pefabloc SC) did not affect its proteolytic activity (data not shown). The inhibition constant, Ki, for pepstatin A was determined for the wild type cardosin A and the three recombinant forms prepared in this study and was between 1 and 1.5 mM for all.

The kcat and Km parameters for the hydrolysis of the peptide substrate Lys-Pro-Ala-Glu-Phe*Nph-Ala-Leu were also determined and are summarized in Table I. Even though there are some differences in the kcat and Km parameters, the values of their specificity constant (kcat/Km) indicate that all forms have similar catalytic efficiencies to cleave the synthetic peptide used as substrate.

Primary Peptide Specificity of Single-chain Recombinant Cardosin A—Given the results reported in the previous section, further characterization of the catalytical properties of recombinant cardosin A was carried out using the single-chain engineered form. The primary specificity of recombinant and natural cardosin A was investigated using two peptide libraries, following a method described elsewhere (18). The results are given in Fig. 10, and all of the values in the P1 and P1′ positions were normalized to a maximum of 100%. The higher the peak for a given amino acid, the higher the preference the enzymes exhibited for that substitution in the substrate.

The results for the primary specificity of recombinant and natural cardosin A for P1 showed that phenylalanine is the only amino acid readily accepted in the S1 pocket, with leucine and norleucine being moderately accepted. For the P1′ position, the results have shown a broader specificity in the S1′ subsite in comparison with the S1 subsite; both enzymes have a preference for tyrosine and phenylalanine in the P1′ position. Concerning the moderate and less tolerated amino acids, there are some differences, mainly in the case of alanine and tryptophan by native cardosin A compared with the low tolerance of recombinant cardosin A for the same amino acids in the P1′ position.

DISCUSSION

Plant aspartic proteases have been implicated in a variety of physiological processes where cell death events play a key role (7, 22–25). Activation of these proteases therefore triggers the onset of these events and determines the fate of the plant cell. In particular the case of cardosin A, the major milk-clotting enzyme of the flowers of cardoon and a model plant aspartic
protease, characterization of the activation process has been hampered by the difficulty in isolating the precursor form from its natural source. Therefore, production of milligram amounts of recombinant procardosin A was important, not only for structural studies, but also to study its activation and proteolytic processing in more detail.

In this work, we applied a novel MS-based C-terminal sequencing method to study the proteolytic processing of a recombinant protein. Even though details of the method are given elsewhere (21), it proved to be highly efficient to obtain C-terminal sequence information using low picomol amounts of protein. Using a combination of CNBr cleavage, enzymatic degradation, and MALDI MS/MS analysis, it is possible to get the C-terminal sequence of proteins at this sensitivity.

Processing of recombinant procardosin A turned out to be a multi-step process. Our data clearly indicate that the first step is the removal of the propeptide, generating an active intermediate form present in the structure. This suggests that inactivation of recombinant cardosin A is accomplished by the presence of a propeptide, as in many other aspartic proteases (26–28), and that the PSI has apparently no relevant role in the inhibition. Cleavage occurs only at the Phe-Arg bond of the propeptide, leaving five amino acids extra at the N terminus of the mature form. Because cleavage was inhibited by pepstatin A and an active site mutant form of cardosin A was unable to undergo the same processing as the wild type form, in vitro activation of recombinant procardosin A appears to be an autocatalytic process. In fact, cardosin A showed a strong preference for Phe in position P1 and was able to cleave the synthetic peptide of the combinatorial library having Arg in position P1’, although not at an optimal rate, further supporting the hypothesis that the enzyme is responsible for its own activation. The mechanism occurs through a mixed intramolecular/intermolecular mechanism, in a similar fashion to what has been proposed for the activation of recombinant rhizopuspepsinogen (29) and pepsinogen (30).

The second step in the proteolytic processing of cardosin A is the removal of the PSI, which is sequentially excised from the precursor. The initial cleavage occurs in the middle of the PSI and then proceeds bidirectionally from that point toward the sequence boundaries between the PSI and the two polypeptide chains of mature cardosin A. In vitro activation of recombinant procardosin A leads ultimately to the generation of an active form where the two polypeptide chains still remain associated by a disulfide bond. This has also been reported for cyrosin and the sunflower seed aspartic protease (31, 32). The complete removal of the PSI described in this work and for cyrosin and sunflower seed aspartic protease suggests therefore that completion of in vivo maturation might require the action of other protease/exopeptidase(s). Nevertheless, the production of an active form of cardosin A requires only the removal of the propeptide, which can be accomplished under acidic conditions. Most likely this occurs inside the vacuole at a slow rate and is accelerated by sudden decreases in pH.

Even though the two-chain recombinant cardosin A is active and displays properties similar to the natural form, the complexity of the processing reaction makes this form of enzyme less attractive for large scale production. Generation of the single-chain form of procardosin A without the PSI domain circumvented this problem because it can easily be autoactivated and shows similar properties when compared with the natural enzyme. The specificity constants are quite similar for all of the recombinant forms of cardosin A tested and are also similar to those determined for native cardosin A. Nevertheless, the comparison of individual constants showed some differences, namely between rCA and rCAΔPSI, with the latter showing increased kcat and K_m values. The difference between these two molecules resides in the PSI domain. Because the PSI domain is located in the C-terminal part of the enzyme, it most likely influences the specificity of the subsites dictated by the C-terminal domain of the enzyme, namely P1’ and P3’ (33); in fact the small differences between nCA and rCAΔPSI specificity are only observed for the amino acid at the P1’ position.

Despite these differences, the results show that both recombinant and natural forms of cardosin A are considerably more selective for the P1 than for the P1’ position, a typical observation for aspartic proteases (18, 33–36), and that both have a preference to cleave peptide bonds between hydrophobic amino acids, like most aspartic proteases. According to the results, cardosin A has a high preference to cleave peptide bonds with Phe at P1, and a moderate preference with Leu or Nle. For P1’, the results obtained with the peptide libraries have shown a high preference toward Phe and Tyr, and a moderate preference for Nle, Trp, and Ala. Taken together, these results indicate that the primary specificity closely resembles those of the mammalian and protozoan homologues.
REFERENCES

1. Beers, E. P., Jones, A. M., and Dickerman, A. W. (2004) *Phytochemistry* 65, 43–58
2. Faro, C., and Gal, S. (2005) *Curr. Protein Pept. Sci.*, in press
3. Murakami, S., Kondo, Y., Nakano, T., and Sato, F. (2000) *FEBS Lett.* 468, 15–18
4. Kato, Y., Murakami, S., Yamamoto, Y., Chatani, H., Kondo, Y., Nakano, T., Yokota, A., and Sato, F. (2005) *Plant Cell* 17, 1673–1682
5. Kato, Y., Murakami, S., Yamamoto, Y., Chatani, H., Kondo, Y., Nakano, T., Yokota, A., and Sato, F. (2005) *Planta* 220, 97–104
6. Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., Dixon, R. A., and Lamb, C. (2004) *EMBO J.* 23, 980–988
7. Simões, I., and Faro, C. (2004) *Eur. J. Biochem.* 271, 2067–2075
8. Egas, C., Lavoura, N., Resende, R., Brito, R. M., Pires, E., de Lima, M. C., and Faro, C. (2000) *J. Biol. Chem.* 275, 38190–38196
9. Guruprasad, K., Tormakangas, K., Kervinen, J., and Blundell, T. L. (1994) *FEBS Lett.* 352, 131–136
10. Munford, R. S., Sheppard, P. O., and O’Hara, P. J. (1995) *J. Biol. Chem.* 270, 1653–1663
11. Koelsch, G., Mares, M., Metcalf, P., and Fusek, M. (1994) *FEBS Lett.* 343, 6–10
12. Tormakangas, K., Hadlington, J. L., Pimpl, P., Hillmer, S., Brandizzi, F., Teeri, T. H., and Blundell, T. L. (1994) *EMBO J.* 13, 3947–3955
13. Ramalho-Santos, M., Pissarra, J., Verissimo, P., Pereira, S., Salema, R., Pires, E., and Faro, C. (1999) *Plant Cell* 11, 1673–1682
14. Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A., and Zdanov, A. (1999) *EMBO J.* 18, 3947–3955
15. Mutlu, A., and Gal, S. (1999) *Physiol. Plant.* 106, 569–576
16. Richter, C., Tanaka, T., and Yada, R. Y. (1998) *Biochem. J.* 335, 481–490
17. Bernstein, N. K., and James, M. N. (1999) *Curr. Opin. Struct. Biol.* 9, 684–689
18. Chen, Z., Koelsch, G., Han, H. P., Wang, X. J., Lin, X. L., Hartsuck, J. A., and Tang, J. (1991) *J. Biol. Chem.* 266, 11718–11725
19. Bhatt, D., and Dunn, B. M. (2000) *Biosens. Chem.* 28, 374–393
20. Turner, R. T., III, Koelsch, G., Hartsuck, J. A., and Tang, J. (1975) *J. Biol. Chem.* 250, 2635–2639
21. White, P. C., Cordeiro, M. C., Arnold, D., Brodelius, P. E., and Kay, J. (1999) *J. Biol. Chem.* 274, 16685–16693
22. Park, H., Kusakabe, I., Sakakibara, Y., and Kobayashi, H. (2001) *Biosci. Biotechnol. Biochem.* 65, 792–795
23. Bhattacharya, D., and Dunn, B. M. (2000) *Biosens. Chem.* 18, 374–393
24. Turner, R. T., III, Koelsch, G., Hartsuck, J. A., and Tang, J. (1975) *J. Biol. Chem.* 250, 2635–2639
25. White, P. C., Cordeiro, M. C., Arnold, D., Brodelius, P. E., and Kay, J. (1999) *J. Biol. Chem.* 274, 16685–16693
26. Park, H., Kusakabe, I., Sakakibara, Y., and Kobayashi, H. (2001) *Biosci. Biotechnol. Biochem.* 65, 792–795
27. Bhattacharya, D., and Dunn, B. M. (2000) *Biosens. Chem.* 18, 374–393
28. Turner, R. T., III, Koelsch, G., Hartsuck, J. A., and Tang, J. (1975) *J. Biol. Chem.* 250, 2635–2639
29. White, P. C., Cordeiro, M. C., Arnold, D., Brodelius, P. E., and Kay, J. (1999) *J. Biol. Chem.* 274, 16685–16693
30. Park, H., Kusakabe, I., Sakakibara, Y., and Kobayashi, H. (2001) *Biosci. Biotechnol. Biochem.* 65, 792–795
31. Bhattacharya, D., and Dunn, B. M. (2000) *Biosens. Chem.* 18, 374–393
32. Turner, R. T., III, Koelsch, G., Hartsuck, J. A., and Tang, J. (1975) *J. Biol. Chem.* 250, 2635–2639
33. White, P. C., Cordeiro, M. C., Arnold, D., Brodelius, P. E., and Kay, J. (1999) *J. Biol. Chem.* 274, 16685–16693
34. Park, H., Kusakabe, I., Sakakibara, Y., and Kobayashi, H. (2001) *Biosci. Biotechnol. Biochem.* 65, 792–795
35. Bhattacharya, D., and Dunn, B. M. (2000) *Biosens. Chem.* 18, 374–393
36. Turner, R. T., III, Koelsch, G., Hartsuck, J. A., and Tang, J. (1975) *J. Biol. Chem.* 250, 2635–2639