J Sci Food Agric 85:549–562 (2005)
DOI: 10.1002/jsfa.1777

Identification and characterisation of the major aspartic proteinase activity in *Theobroma cacao* seeds

Martine Guilloteau,[1] Maryse Laloi,[2] Stephan Michaux,[1] Peter Bucheli[3] and James McCarthy[1]*

[1]Department of Plant Science, Nestlé Research Centre, 101 Avenue Gustave Eiffel, BP 9716, F-37097 Tours Cedex 2, France
[2]Laboratoire de Biogenèse Membranaire, UMR-CNRS 5544, Boite 92, Université Victor Segalen Bordeaux 2, 146 Rue Léo Saignat, F-33076 Bordeaux, France
[3]Nestlé R&D Center Shanghai Ltd, No 13 Qiao Nan, Cao An Road, Jia Ding District, Shanghai 201812, China

Abstract: ***Theobroma cacao*** **seeds contain an unusually high level of aspartic proteinase activity. Although this activity is central to the development of high-quality cocoa flavour, the *T cacao* polypeptide responsible has not yet been definitively identified. Here we report the identification and characterisation of an active protein complex from *T cacao* seeds with an apparent molecular weight of approximately 50 kDa. This active complex contains at least two polypeptides: an approximately 30.5 kDa aspartic proteinase, the product of the TcAP2 gene, and an associated polypeptide, the 20.5 kDa trypsin inhibitor protein. The active complex co-eluted off a size exclusion column with another complex containing the trypsin inhibitor and a putative acid chitinase. The 30.5 kDa TcAP2 proteinase is apparently a monomeric aspartic proteinase with optimal activity between 42 and 47 °C and an optimal pH of 3.0. Significant inactivation of the TcAP2 activity occurs at acid pH around 47–52 °C, a temperature potentially obtained during cocoa bean fermentation. SDS-PAGE analysis showed that the purified TcAP2 complex efficiently degrades the cacao seed storage protein vicilin into peptides smaller than 10 kDa. In addition, high-resolution size exclusion chromatography showed that this proteinase is capable of degrading proteins into peptides as small as di- and tripeptides, indicating for the first time that the main *T cacao* seed aspartic proteinase can produce very small peptide products. Our results demonstrate that the aspartic proteinase encoded by the TcAP2 gene plays a critical role in the production of cocoa flavour precursor peptides during cocoa bean fermentation.**

Keywords: aspartic proteinase; proteolysis; *Theobroma cacao*; trypsin inhibitor; flavour precursor peptides

INTRODUCTION

Proteases perform a number of vital cellular processes in plants, such as protein maturation and the protein degradation associated with tissue restructuring and cell maintenance.[1,2] One important set of proteases comprises the aspartic proteinases (EC 3.4.23), which have been purified and characterised from a number of monocots, dicots and gymnosperms.[2,3] Plant aspartic proteinases are generally synthesised as precursor proteins which are subsequently processed to produce the mature active polypeptides. Precursor protein processing yields either a monomeric or a dimeric protein, with the dimeric form being generated by the removal of an internal peptide sequence of approximately 50–100 amino acids called the plant specific insert (PSI).[2] All aspartic proteinases are optimally active at acid pH, contain two aspartic residues at their active sites and are specifically inhibited by pepstatin A.[4]

In arabidopsis the expression of an aspartic proteinase called AtAP has been detected in seeds, flowers, stems and roots.[2,5] In barley an aspartic proteinase mRNA has been detected in developing flowers, root tips and leaves and in both the developing and mature grain.[6,7] In both these plants the aspartic proteinase is localised in the vacuole.[5,8] Most of the plant aspartic proteinases characterised to date have been purified from seeds, and the purification procedures often include a low-pH extraction process combined with a pepstatin affinity chromatography step. Although many of the biochemically characterised plant aspartic proteinases are heterodimeric proteins with a large subunit of 28–35 kDa and a small subunit of 11–16 kDa, a few

* Correspondence to: James McCarthy, Department of Plant Science, Nestlé Research Centre, 101 Avenue Gustave Eiffel, BP 9716, F-37097 Tours Cedex 2, France
E-mail: james.mccarthy@rdto.nestle.com
(*Received 3 February 2003; revised version received 10 December 2003; accepted 8 March 2004*)

plant aspartic proteinases are monomeric proteins of between 36 and 65 kDa.[2] Where the cellular location has been determined, the heterodimeric aspartic proteinases were generally found to be associated with vacuoles, while the larger and less processed monomers were located extracellularly.[2,5,8]

Aspartic proteinase activity in the seeds of *Theobroma cacao* has been extensively studied by Biehl and co-workers. Biehl *et al*[9] partially purified an aspartic proteinase from seeds which had optimal activity at 55 °C and a pH optimum of 3.5. Subsequently, Voigt *et al*[10] obtained a more purified *T cacao* seed aspartic proteinase and showed that this protein was a heterodimer of two polypeptides with approximate molecular weights of 29 and 13 kDa respectively; however, no biochemical characterisation or recovery data were reported. The partially purified *T cacao* seed aspartic proteinase efficiently degraded cacao vicilin and trypsin inhibitor polypeptides of *T cacao* seeds into peptides, although the latter seemed to be a less favoured substrate.[11] More recently, two cDNA encoding cacao seed aspartic proteinases, called TcAP1 and TcAP2 respectively, have been cloned and characterised.[12] These cDNA encode relatively different proteins which appear to be associated with distinct groups in the plant aspartic proteinase gene family. Both genes are induced early in seed development, and the expression of both decreases significantly as the seeds reach maturity. However, TcAP2 expression is induced to a much higher degree in seeds than TcAP1 expression, strongly suggesting that TcAP2 encodes the main aspartic proteinase activity in the mature seed. It is also important to note that *T cacao* seeds have unusually high levels of aspartic proteinase activity compared with other plant seeds.[13]

In order to produce cocoa beans with good cocoa flavour, it is necessary to carry out a natural fermentation step, during which the sugars in the pulp are fermented, ultimately generating high levels of acids, particularly acetic acid.[13,14] As the fermentation continues, the pH in the seed decreases and the cell structure becomes disrupted.[15] This triggers mobilisation and/or activation of the main cacao seed aspartic proteinase, resulting in massive degradation of cellular protein.[15,16] Because peptides and amino acids represent an important group of cocoa flavour precursors,[17,18] the *T cacao* seed aspartic proteinase, together with a seed serine carboxypeptidase, has been proposed as being critical for the generation of cocoa flavour precursors during fermentation.[18,19] Here we describe the purification and characterisation of a protein complex containing the major *T cacao* seed aspartic proteinase. Using SDS-PAGE analysis and N-terminal protein sequencing analysis, we show that the activity in the purified preparation is associated with a single processed polypeptide of approximately 30.5 kDa whose N-terminal end corresponds to amino acid 73 of the polypeptide encoded by the TcAP2 gene.[12] The most purified aspartic proteinase preparation also contains a very abundant seed protein, the trypsin inhibitor protein, and a moderately abundant seed protein which has been tentatively identified as an acid chitinase. These three proteins probably exist in the form of two distinct complexes that co-purify. Finally we show that the purified aspartic proteinase complex can hydrolyse cacao seed proteins into peptides which range in size from approximately 80 amino acids down to approximately two or three amino acids in length. This latter property is believed to be central to the production of cocoa flavour precursor peptides and amino acids during the fermentation of cacao beans.

MATERIALS AND METHODS

Production of acetone-dried powder of cacao seeds

Fresh mature cacao pods (genotype EET 95) were obtained from the former Nestlé farm in Quevedo, Ecuador. The pulp, testa and main part of the embryo were removed and the remaining material was rapidly frozen in liquid nitrogen and stored at −80 °C. Approximately 25 g of the frozen material was then ground to a fine powder using liquid nitrogen and a Model 6800 Freeze Mill (Spex CertiPrep, Metuchen, NJ, USA). Subsequently an acetone extraction step[20] was carried out to remove most of the fat and phenolic compounds, resulting in the production of a fine whitish/grey protein-containing powder.

Measurement of aspartic proteinase activity

The routine assay was conducted at 42 °C using 240 µl of BHM solution (10 mg ml^{-1} stock solution of bovine haemoglobin (Sigma-Aldrich Chimie, St Quentin Fallavier, France) in 0.2 M sodium citrate pH 3), 1–120 µl of extract, and water to a total volume of 300 µl. At various times, 80 µl of the reaction mixture was added to 80 µl of 80 g l^{-1} TCA (trichloroacetic acid). This mixture was left at room temperature for 10 min and then centrifuged at 10 000 × *g* for 10 min in a microcentrifuge. A 30 µl aliquot of the supernatant was added to 250 µl of freshly prepared OPA reaction solution (50 mM sodium tetraborate, 10 mg ml^{-1} SDS (sodium dodecyl sulphate), 6 mM OPA (*o*-phthaldialdehyde), 28.6 mM β-mercaptoethanol) in the wells of a microtitre plate.[21] The reaction mixture was incubated for 15 min with shaking at room temperature, then the absorbance was read at 340 nm. All measurements were in the linear range of the reaction. Absorbance values were subsequently converted to leucine equivalents by using a standard curve generated by reacting various known amounts of leucine with the OPA reagent under the standard reaction conditions. One unit of activity is defined as 100 ng leucine equivalents produced min^{-1} at 42 °C.

Purification of native TcAP2 aspartic proteinase complex

A 5 g sample of acetone powder was added to 500 ml of buffer 1 (10 mM sodium phosphate pH 7.8, 2 mM EDTA (ethylene diamine tetra-acetic acid), 10 mM sodium ascorbate). This mixture was agitated for 1 h at 4 °C and then centrifuged at 7840 × *g* for 25 min at 4 °C using 50 ml polycarbonate tubes (Beckman Coulter France SA, Villepinte, France). The supernatant was filtered through four layers of Miracloth (Calbiochem-Merck Biosciences, Beeston, Nottingham, UK) to produce SU1. The pellet from centrifugation was resuspended in 500 ml of buffer 1, and the extraction was repeated as above to produce SU2. Subsequent steps were carried out at 4 °C unless otherwise noted. The two supernatants SU1 and SU2 were pooled and two steps of protein precipitation were done using ammonium sulphate at 30 and 60% saturation respectively. The 60% saturation ammonium sulphate precipitate, which contained the highest level of activity, was dissolved in 6 ml of buffer 2 (50 mM sodium phosphate pH 7.8, 1 mM EDTA) and placed in 9–12 ml 3.5 K 'Slide-a-Lyzer' dialysis cassettes (Perbio Science France, Brebières, France). The dialysis was carried out for 3 h in 1.5 l of buffer 2; then the buffer was changed, and the dialysis was continued for 1.5 h in another 1.5 l of buffer 2. After this relatively short dialysis step, a significant amount of insoluble material was formed. Therefore, to obtain only the soluble protein, it was necessary to centrifuge this solution for 10 min at 7840 × *g* in glass Corex™ (Bibby Sterilin, Nemours, France) tubes at 10 °C, and the supernatant was then put on ice. To increase the recovery of soluble material/activity, the pellet obtained was resuspended in 7 ml of buffer 2, vortexed vigorously and centrifuged as above. The two solutions were pooled and loaded onto a Q Sepharose Fast Flow column (Pharmacia, now Amersham Biosciences Europe, Saclay, France). This column was then washed with buffer A(20 mM Tris/HCl pH 8) and then eluted with a linear gradient of buffer B (20 mM Tris/HCl pH 8, 1 M NaCl).

Aliquots from various selected fractions were assayed for aspartic proteinase activity. Fractions showing the highest level of protease activity were pooled, concentrated using an Ultrafree Biomax filter (5 kDa molecular weight cut-off (MWCO); Millipore, Molsheim, France) and loaded onto a Sephacryl S-200 size exclusion column (Pharmacia) equilibrated with 10 mM Tris/HCl pH 8 and 500 mM NaCl. Aliquots from the various fractions were assayed for aspartic proteinase activity. The most active fractions (57–64) were pooled and concentrated using an Ultrafree Biomax filter (5 kDa MWCO). The concentrated pools were mixed with 500 g l^{-1} glycerol stock solution to give a glycerol concentration of approximately 250 g l^{-1}; aliquots were frozen at −80 °C and the specific activities of the final pools were determined. Protein concentration was determined using the micro BCA protein assay kit of Pierce, Inc with the protocol for microtitre plates and BSA as standard. To determine the total protein concentration in the crude extract SU1 + SU2 more accurately, it was necessary to carry out a dialysis step using a 5 kDa MWCO membrane to remove low-molecular-weight compounds which react strongly with the BCA reagent (Pierce, Inc). However, we believe that the calculated total protein in the crude SU1 + SU2 extract was still overestimated at least twofold owing to a residual low level of interfering molecules.

Characterisation of purified aspartic proteinase

Inhibitor sensitivity

The inhibitor sensitivity of the aspartic proteinase was determined in 300 µl reaction mixtures containing 200 mM sodium citrate pH 3, 10 mg ml^{-1} bovine haemoglobin and 5 µl of purified aspartic proteinase (2.4 µg protein $µl^{-1}$). Inhibitors were added to give final concentrations of 2 µM pepstatin, 2 mM 1,10-phenanthroline, 100 µM dichloroisocoumarin (DCI) and 10 µM E-64. Inhibitor stocks were stored at −20 °C and thawed for single use. The reactions were run at 42 °C, and 30 or 40 µl samples were taken at various times up to 75 or 90 min and added to an equal volume of 80 g l^{-1} TCA. After centrifugation, 30 µl of each reaction supernatant was added to 250 µl of OPA and further processed as described earlier. For the test with the cacao trypsin inhibitor protein, 13 or 130 µl of Fast Flow Q (FFQ) fraction 54 (enriched for the trypsin inhibitor; 1 µg $µl^{-1}$ protein concentration) was added to the 300 µl reaction mixtures. To test the effect of β-mercaptoethanol, the 300 µl reaction mixtures were made either 0.5 or 5.0 mM in β-mercaptoethanol.

Determination of optimal pH

Reactions (300 µl) were set up in 200 mM sodium citrate buffer at pH 2, 2.5, 3.0 and 4.0. Each reaction mixture also had 5 mg ml^{-1} bovine haemoglobin (from a 50 mg ml^{-1} stock solution made in Milli-Q water (Millipore)). As addition of the haemoglobin solution raised the pH, each reaction mixture was made up as a larger volume than necessary for the reaction, and the pH was measured. This gave final reaction pH values of 2.5, 3.0, 3.4 and 4.3 respectively. The reactions were started by adding 2 µl of purified aspartic proteinase (2.4 µg protein $µl^{-1}$). Triplicate reactions were set up for each pH at 42 °C and then sampled and processed as described for the inhibitor sensitivity study.

Determination of optimal reaction temperature and temperature stability

Reactions (300 µl) were set up at the indicated temperatures and sampled as described for the inhibitor sensitivity study. At least three reactions were run for each temperature. For the stability tests the purified aspartic proteinase was examined under two different conditions. The enzyme was left in the storage buffer (10 mM sodium phosphate pH 8, 250 mM NaCl, 250 g l^{-1} glycerol) at room temperature (18–22 °C), and the activities of aliquots were measured after

different lengths of time using the routine aspartic proteinase assay. Alternatively, the stability of the enzyme was tested under acid conditions at room temperature as follows: 60 µl of purified aspartic proteinase was put into 30 µl of 250 mM sodium citrate pH 4 (85 mM sodium citrate final concentration), and the activities of aliquots were measured after different lengths of time using the standard assay.

Degradation of partially purified cacao vicilin by purified aspartic proteinase

The course of degradation of vicilin by the purified aspartic proteinase was followed by setting up 200 µl reaction mixtures containing 250 µg of partially purified vicilin (prepared as described in Ref 19) and 3.9 µg of purified aspartic proteinase in 100 mM sodium citrate pH 3, and incubating at 42 °C. At the times indicated, 5 µl samples were taken for SDS-PAGE analysis. As one control, an identical reaction was set up without the purified aspartic proteinase. As another control, a 100 µl reaction was set up with only 15.6 µg of purified aspartic proteinase in 100 mM sodium citrate pH 3. In each case, samples plus molecular weight markers (Precision™ Markers, Bio-Rad, Marnes-la-Coquette, France) were taken and analysed on a 100–200 g l^{-1} SDS-PAGE gel as described above.

Size exclusion chromatography of products formed by autodigestion with *Theobroma cacao* aspartic proteinase

A reaction was set up with 120 µl of Q Sepharose Fast Flow column-purified TcAP2 (197 µg of protein, specific activity 821 units mg^{-1} protein) and 30 µl of 1 M sodium citrate pH 3. A 4 µl sample was taken immediately after mixing ($T = 1$ min) for SDS-PAGE analysis. At the same time a 70 µl sample was also taken, stopped by adding SDS to a final concentration of 10 mg ml^{-1} and frozen. The remaining reaction mixture was held at 42 °C. After incubation for 7 h, another 4 µl sample was taken for SDS-PAGE analysis, and the rest of the reaction mixture was made 10 mg ml^{-1} in SDS and frozen. Subsequently the two large frozen SDS-treated samples were freeze-dried, solubilised with 100 µl of 6 M urea and 20 mM sodium phosphate pH 7. These samples were loaded onto a Superdex Peptide HR 10/30 column (Amersham Biosciences Europe, Saclay, France) and eluted with 6 M urea and 20 mM sodium phosphate pH 7 at ambient temperature. The following markers were obtained from Sigma and run separately under the same conditions as molecular weight standards: ribonuclease A, 13 700 Da; aprotinin, 6500 Da; substance P, 1347 Da; *N*-benzoyl-gly-phe (hippuryl-phe), 326 Da; phenylalanine, 165 Da.

N-terminal protein sequencing

Purified protein samples were loaded on either a 150 g l^{-1} SDS-PAGE gel or a Tris/tricine gel for low-molecular-weight peptides.[22] The gels were run in a Biorad electrophoresis apparatus with water cooling. After electrophoresis the separated proteins were blotted onto PROBLOTT™ PVDF membrane (Applied Biosystems, Courtaboeuf, France) using a Bio-Rad (Marnes-la-Coquette, France) Trans-Blot cell and transfer buffer (10 mM CAPS (3-(cyclohexylamino)-1-propanesulphonic acid buffer) titrated to pH 11 with NaOH and containing 0.05 mg ml^{-1} SDS). The transfer was carried out at 420 A constant current, with cooling, at 6 °C for 2 h. The blot was then rinsed in Milli-Q water for 1 min and stained for 1 min in 450 ml l^{-1} methanol containing 10 mg ml^{-1} acetic acid and 1 mg ml^{-1} coomassie brilliant blue 250 (Biorad). The blot was subsequently washed five times in 500 ml l^{-1} methanol to reduce the background and then air dried before storage at −20 °C. Strips of PVDF paper with the appropriate bands were cut out and N-terminal protein micro-sequencing was performed on a model 494 or 473A protein-sequencing machine (Applied Biosystems, Courtaboeuf, France) using the Edman degradation method according to the manufacturer's instructions.

RESULTS

Purification of major aspartic proteinase from cacao seeds

Low-temperature milled cacao seeds from the genotype EET 95 were extracted with acetone to remove most of the polyphenols and fats from the seed material. The acetone powder prepared as described above was then extracted with a low-salt buffer (10 mM sodium phosphate pH 7.8, 2 mM EDTA, 10 mM sodium ascorbate). In order to limit the activity of the acid proteases, the pH of the extraction buffer was 7.8. The total acid protease activity in the crude extract (SU1 + SU2) was assayed using the standard method described above. More than 95% of the activity found using acidic assay conditions was inhibited by pepstatin, indicating that the bulk of the activity detected in the crude extract was probably due to an aspartic proteinase. The crude extract was then subjected to ammonium sulphate precipitation at 30 and 60% saturation respectively. The majority of the activity was found in the 60% ammonium sulphate precipitate fraction, and this fraction was dialysed into the buffer used for the subsequent Q Sepharose Fast Flow anion exchange column. The length of the dialysis step was kept to a minimum, because a very significant amount of protein precipitation was found to occur during dialysis. The soluble protein recovered after dialysis was loaded onto the Q Sepharose column. The aspartic proteinase activity eluted in one peak well after most of the protein had already been washed off the column, demonstrating the high purification capability of this column. The most active Q Sepharose column fractions were pooled, concentrated and loaded on a Sephacryl S-200 size exclusion column. Three protein peaks eluted from this column at pH 8.0 and 1 M

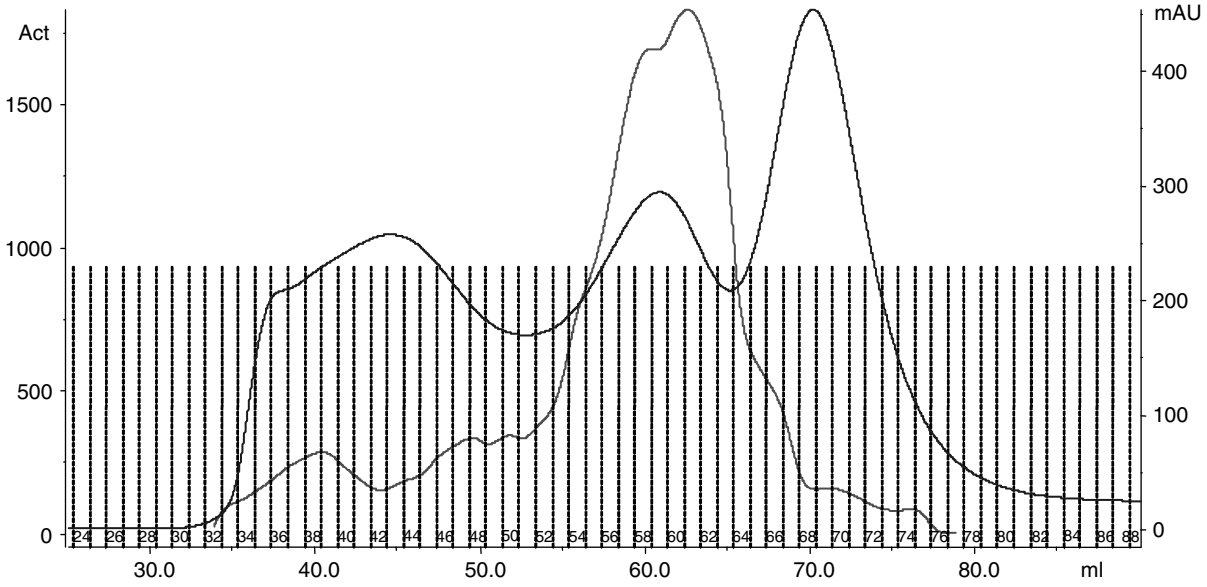

Figure 1. Elution profile of Sephacryl S-200 size exclusion column. The main peak of activity off the Q Sepharose Fast Flow column was run on a Sephacryl S-200 column as described in the 'Materials and Methods' section. The profile with three peaks corresponds to the UV 280 nm scan (right-hand scale). The second profile, which has one major peak, is the activity (ng leucine equivalents produced min^{-1}) profile (left-hand scale).

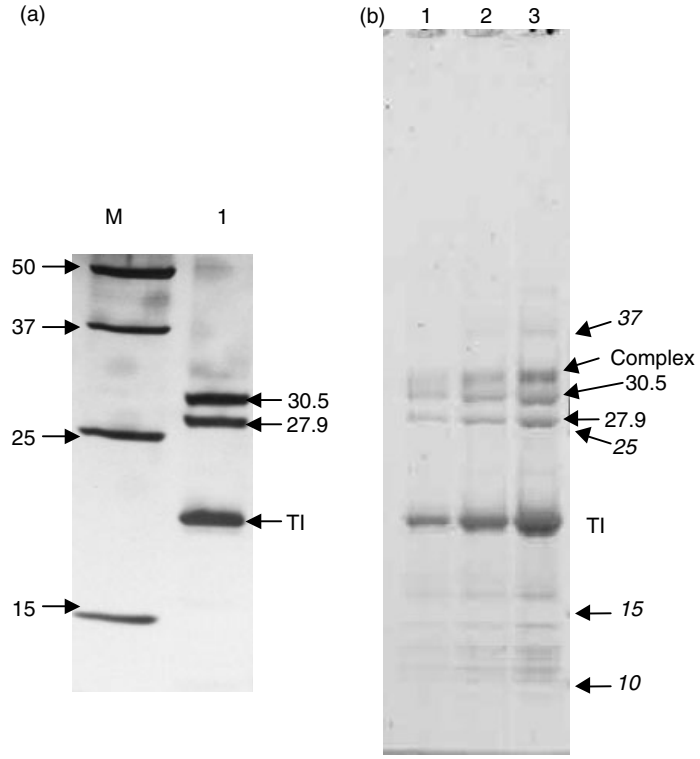

Figure 2. SDS-PAGE analysis of pooled active fractions from size exclusion column. Fractions 57–64 from the size exclusion column were run on a 100–200 g l^{-1} SDS-PAGE gel and either (a) silver stained or (b) coomassie blue stained. For (a), 5 ng of protein was loaded in lane 1. For (b), 12, 24 and 40.8 μg of protein respectively was loaded in lanes 1–3. TI, trypsin inhibitor, Complex, a covalent complex between the aspartic proteinase and a fragment of the trypsin inhibitor protein. The molecular weights of the markers are on the left in (a) and are in italics in (b).

NaCl, and the active fractions (57–64) co-eluted with the second protein peak (Fig 1). The active fractions, which eluted at a molecular weight of approximately 50 kDa associated with globular proteins, were pooled and used for further analysis.

The polypeptides present in the purified fraction were examined by SDS-PAGE using both silver and coomassie blue staining (Fig 2). Three strong bands were apparent at relatively similar intensities on the silver-stained gel (Fig 2(a)), corresponding to polypeptides of approximately 30.5, 27.9 and 20.2 kDa respectively. The same pattern of bands was also seen on the coomassie blue-stained gel. However, the coomassie blue-stained gel, which was loaded with significantly higher amounts of the purified protein, exhibited several additional features. In contrast to the silver-stained gel, the band at 20.2 kDa appeared much stronger than those for the larger polypeptides, suggesting that

it reacts more strongly than the other polypeptides with coomassie blue (Fig 2(b), lane 1). Coomassie blue staining also revealed the presence of a strong new band at approximately 33.5 kDa (see below), and five or six much weaker bands corresponding to polypeptides with molecular weights in the range of approximately 16–11 kDa. To determine the identity of these polypeptides, each was subjected to analysis by N-terminal protein sequencing after blotting onto PVDF membranes. N-terminal sequencing of the 30.5 kDa polypeptide produced the sequence DSEETDIVAL. This sequence corresponds precisely to amino acid positions 73–82 in the sequence of TcAP2,[12] while it has three amino acid differences from the protein sequence of TcAP1 in the same region.[12] Thus the 30.5 kDa polypeptide was the product of the recently discovered TcAP2 gene. The N-terminal sequence obtained from the 27.9 kDa protein was TVISTYWGQNGFEGT. Various BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST/) using this sequence against the various protein databases found that the strongest similarity was to the plant class III acid chitinases. For example, this sequence showed 76.9% identity beginning at position 29 in the translated sequence of a *Glycine max* acid chitinase chibl-b (a class III chitinase, accession # AB007127). Importantly, this region of the *G max* polypeptide is also likely to be at, or very near, the N-terminal sequence of the mature *G max* chitinase after the signal peptide is removed. Because of the high level of sequence identity, and the conservation of this peptide sequence position, we propose that the 27.9 kDa polypeptide is a cacao acid chitinase.

Based on its size and abundance, together with some internal peptide sequence data, it had been determined earlier that the 20.2 kDa polypeptide was the abundant trypsin inhibitor protein. To verify the identity of the 20.2 kDa band in the size exclusion column-purified material, N-terminal sequencing was carried out on this polypeptide for several cycles. The peptide sequence obtained was ANSP, confirming both that this band was the cacao trypsin inhibitor protein (accession # X56 509) and that the N-terminal position of the mature cacao trypsin inhibitor protein was at amino acid 27 of the precursor protein.[23] It is noted that, from amino acid 27 to the end of the trypsin inhibitor protein, ORF represents a polypeptide of 195 amino acids with a predicted molecular weight of 21.3 kDa. This molecular weight is very close to the molecular weight determined here from the SDS-PAGE analysis for the trypsin inhibitor protein and to the previous LC/ESI-MS-determined value.[23] N-terminal peptide sequencing of the material in the slightly diffuse but strong coomassie blue-stained band at 33.5 kDa detected both the N-terminal sequence of the 30.5 kDa polypeptide and the N-terminal sequence of the trypsin inhibitor. Because this band was not detected by silver staining (Fig 2(a)), the discovery of a mixture of 30.5 kDa peptide and trypsin inhibitor peptide sequences in the 33.5 kDa band suggests that heating a sample of the purified activity at high protein concentration induces formation of a covalent complex between these two polypeptides.

The aspartic proteinase previously purified from cacao seeds using pepstatin A affinity chromatography was shown to be composed of two subunits with molecular weights of 29 and 13 kDa respectively.[10] However, from the coomassie blue-stained SDS-PAGE gel shown in Fig 2(b), it was not obvious which of the lower-molecular-weight polypeptides seen at low levels in the purified fractions could potentially correspond to the small cacao aspartic proteinase subunit previously observed. In an attempt to determine if any of the low-molecular-weight polypeptides present in the purified activity corresponded to the expected 13 kDa polypeptide, we carried out N-terminal sequencing on the four strongest bands seen below 15.6 kDa. The polypeptides in the size exclusion column-purified activity were run on a Tris/tricine gel capable of separating small peptides[22] and then blotted onto a PVDF membrane. Each of the bands sequenced was found to correspond to progressively smaller fragments of the trypsin inhibitor protein (the N-terminal ends of the progressively smaller fragments were ANSP, RL(C)ST, SAGKW and TLCSWFK respectively). This result strongly suggests that some of the trypsin inhibitor polypeptides have one or more of their peptide bonds cleaved, but that the complex containing the trypsin inhibitor remains structurally intact. All the data obtained to date strongly indicate that there is only a monomeric aspartic proteinase in the purest fractions.

Overall, although the TcAP2 activity had several associated polypeptides on SDS-PAGE analysis (Fig 2(b)), all the N-terminal sequencing information obtained indicates that there are only three polypeptides in the purest fractions: the 30.5 kDa TcAP2 product, the 27.9 kDa putative acid chitinase and the trypsin inhibitor polypeptide (and cleavage fragments thereof). We have attempted to further separate the aspartic proteinase activity from the trypsin inhibitor and putative chitinase polypeptides with other chromatographic steps such as hydrophobic and cation exchange chromatography. However, no further significant purification of the aspartic proteinase activity from the trypsin inhibitor has been accomplished to date even after several trials with different columns. This observation implies that there is a very strong interaction between the TcAP2 polypeptide and the trypsin inhibitor and/or possibly the putative acid chitinase (see below).

A summary of the purification scheme and the total recovery of activity during the purification procedure is shown in Table 1. The aspartic proteinase activity has been purified at least 16-fold. However, it is clear that the recovery of total activity is poor, especially during the first purification step when the activity is concentrated using ammonium sulphate precipitation, and there is a loss of approximately 83% of the activity. Although low amounts of aspartic proteinase

Table 1. Purification summary showing percentage recoveries of protein and protease activity at each step

	Activity units[a]	Activity recovery (%)	Protein (mg)	Protein recovery (%)	Specific activity (units mg^{-1} protein)
SU1 + SU2	98 799 (non-dialysed)	100	1516 (dialysed, 5 kDa)[b]	100	65.2
60% ammonium sulphate precipitate	16 210	16.4	184.3	12.2	87.9
FFQ column concentrated pool (fractions 65–80)	6006	6.1	11.5	0.76	522.3
Size exclusion column pools (fractions 53–56, 57–64 and 65–68 combined)	6647	6.7	11.4	0.75	583.0

[a] 1unit = 100 ng leucine equivalents produced min^{-1} at 42 °C.
[b] See 'Materials and Methods' section.

activity were present in the other ammonium sulphate fractions, the majority of the activity was simply lost during this step. Because the TcAP2 activity was found to be quite stable at pH 8 (see below), it is unlikely that the activity loss seen at this step was due to inactivation. It is possible that the loss of activity observed in our experiments was due to the loss of the small subunit found in the previously characterised *T cacao* aspartic proteinase.[10] However, it is more likely that the activity lost during the ammonium precipitation/dialysis step and other steps where protein precipitates formed (concentration on Biomax filters and during dialysis) was due to irreversible protein precipitate formation. Consistent with the idea that the aspartic proteinase complex is very susceptible to precipitation, we have found that simply dialysing the crude SU1 + SU2 extract against 25 mM sodium phosphate pH 8.0 and 1 mM EDTA produces both a significant precipitate and a loss of >40% of the activity in the extract. In the future it will be interesting to search for buffer conditions which minimise this precipitation. It will also be important to determine if the precipitate formed consists of a specific group of proteins and to try to recover the activity from these precipitates by using various solubilisation/refolding methods.

Characterisation of activity in purified aspartic proteinase complex

Table 2 shows the effects of various inhibitors on the purified TcAP2. Only marginal inhibition was detected with 1,10-phenanthroline, DCI and E-64, which are specific inhibitors of metallo-proteases, serine proteases and cysteine proteases respectively. The very weak inhibition seen with these three inhibitors was probably due to a small non-specific effect from the carrier solvents and/or the high concentration of the inhibitors themselves, because the activity was completely inhibited by a low concentration of pepstatin, a very specific inhibitor of the aspartic proteinases. Overall, the inhibition data clearly showed that all of the purified protease activity was due to a pepstatin-sensitive aspartic proteinase. Because of the apparent strong binding of the trypsin inhibitor to the purified protein, we also tested the effects of adding more trypsin inhibitor protein to the purified TcAP2. Approximately 13 or 130 μg of a trypsin inhibitor-enriched fraction from the FFQ column was added to two assay samples each containing 12 μg of purified TcAP2. The results obtained showed that at the lower level the cacao trypsin inhibitor had no effect on the activity. Even at the 10-fold higher level, 81% of the control activity was detected, indicating that even at very high levels the trypsin inhibitor protein only marginally inhibits the purified TcAP2 activity. These results are in agreement with earlier work showing that the trypsin inhibitor protein does not inhibit the cacao seed aspartic proteinase activity dramatically.[23] Protein extracts from the flowers of cardoon (*Cynara cardunculus* L) are used in Portugal for the production of a traditional cheese.[24] The major protease activity of this flower was purified and was shown to be an aspartic proteinase that efficiently cleaved κ-casein at the same peptide bond as chymosin.[24] In order to test the milk-clotting capability of the TcAP2 activity, milk was treated with FFQ-purified TcAP2 at 42 °C as described previously.[25] However, even at a relatively high level of added proteinase activity the major cacao aspartic proteinase TcAP2 showed only a very limited capability to induce milk clotting in the range of pH from 5.0 to 5.5.

The pH optimum of the purified TcAP2 activity was determined as described in the 'Materials and methods' section. The results obtained (Fig 3(a)) showed that the TcAP2 activity was optimal at pH 3.0 when

Table 2. Inhibitor sensitivity of TcAP2 Activity

Inhibitor	Conc (mM)	Remaining activity (%)
—	—	100
Pepstatin A	0.002	0
1,10-Phenanthroline	2.0	86
E-64	0.01	88
DCI	0.1	90

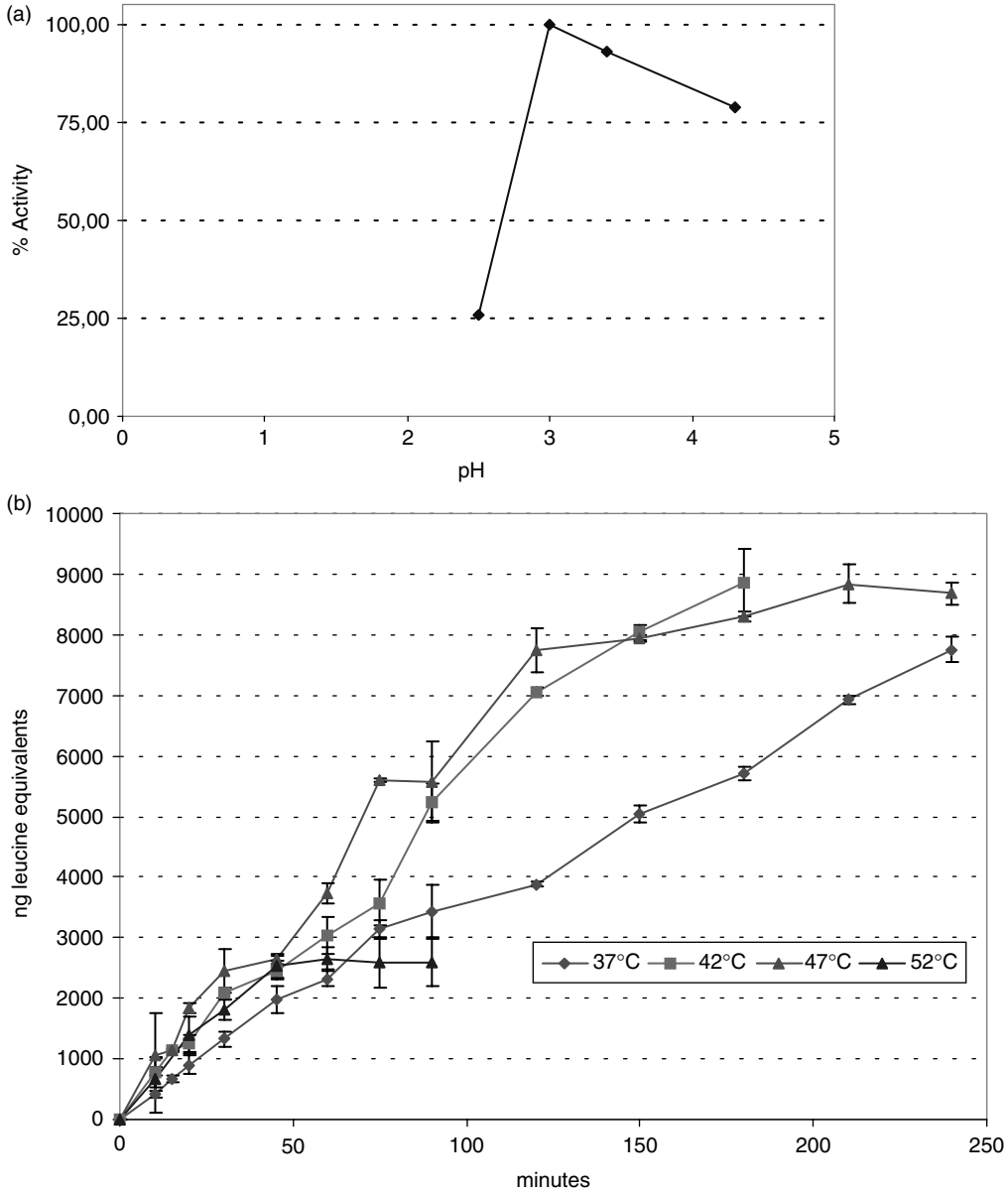

Figure 3. Optimal pH and reaction temperature of purified TcAP2 complex: (a) pH optimum of purified TcAP2 complex; (b) rate and extent of haemoglobin hydrolysis at different temperatures by purified TcAP2 complex as a function of time.

haemoglobin was used as substrate. The pH optimum for the *T cacao* seed aspartic proteinase described earlier by Biehl *et al*[9] was 3.5 using BSA as substrate. To determine the optimal reaction temperature, the purified TcAP2 activity was reacted with haemoglobin at different temperatures. The results of this experiment are shown in Fig 3(b). The enzyme activity was highest between 42 and 47 °C. At 52 °C, although the initial reaction rate was approximately the same as that seen at 42 °C, the enzyme rapidly lost activity at the higher temperature. This loss was manifested by the absence of activity after approximately 45 min (Fig 3(b)). This effect was even more pronounced at 57 °C, where most of the activity was lost after only 20 min of reaction (data not shown). Again, the optimal reaction temperature for the TcAP2 activity (42–47 °C) differs from the temperature optimum found for the *T cacao* aspartic proteinase activity characterised by Biehl *et al*,[9] which was reported to be optimal at 55 °C. We also directly tested the stability of the purified TcAP2 activity at room temperature under two different pH conditions. The purified activity was relatively stable for over 9 h at pH 8.0 (in storage buffer), while it was lost slowly when incubated at room temperature at pH 4 (storage buffer titrated to pH 4.0 with citric acid; data not shown).

Native PAGE gel analysis of size exclusion-purified TcAP2

TcAP2 remained strongly associated with the trypsin inhibitor protein on all the chromatography columns tested to date. To study further the interactions between the polypeptides in the purified TcAP2 preparation, we decided to examine their mobility on a native PAGE gel under different conditions. For this experiment, two sets of fractions eluting off the size exclusion column were run on a native 100–200 g l^{-1}

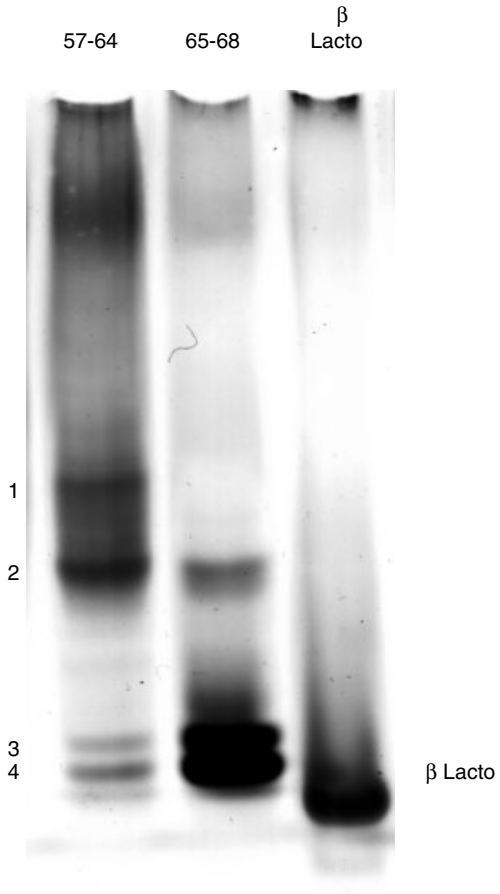

Figure 4. Native PAGE gel analysis of two size exclusion column-purified pools containing different levels of TcAP2. The pool of fractions 57–64 (6 μg loaded) has a high level of AP activity, the pool of fractions 65–68 (6 μg loaded) has a low level of AP activity. The gel was silver stained. ß-lacto, β-lactoglobulin 18 400 kDa.

polyacrylamide gel. These two sets of fractions differed substantially in their specific activities and in the relative levels of the three major polypeptides. The results (Fig 4) demonstrated that purified TcAP2 (fractions 57–64), which had both a very high specific activity and a high level of the 30.5 kDa TcAP2 polypeptide, also exhibited significant amounts of two distinct high-molecular-weight bands (complexes marked 1 and 2), in addition to some less stable higher-molecular-weight material which ran as a faint smear between band 2 and the top of the gel. This pool also exhibited low levels of two other bands which migrated faster (Fig 4, complexes marked 3 and 4). The other group of fractions analysed (65–68, representing approximately the first half of the third peak eluting off the size exclusion column in Fig 1) had a much lower specific activity, a very low level of the 30.5 kDa TcAP2 polypeptide and high levels of the trypsin inhibitor protein and the 27.9 kDa putative acid chitinase. When this latter sample was run on the native gel, a low level of the stable high-molecular-weight band 2 was seen, together with high levels of the two smaller stable bands (3 and 4). Because the pool containing fractions 65–68 was almost exclusively composed of the trypsin inhibitor and the 27.9 kDa polypeptide, and control experiments showed that the trypsin inhibitor protein alone formed a new band in a different region of the gel, the two smaller bands are likely to be due to two different forms of complex containing the trypsin inhibitor and the 27.9 kDa polypeptide. Given the fact that bands 1 and 2 are the main stable bands in the pool with the highest specific activity (fractions 57–64), it is likely that these bands represent specific complexes between the TcAP2 polypeptide and the trypsin inhibitor and/or the 27.9 kDa polypeptide. In the future it should be possible to test the composition of the different complexes detected on native PAGE gels directly by probing western blots of these complexes with antibodies specific for each of the three polypeptides present.

To study further the stability of the different complexes, the purified TcAP2 was subjected to a number of treatments before being run on a PAGE gel under native conditions. For example, we tested the effects of treatments which would be expected to inhibit (pepstatin addition) or destroy (heating the samples at 52 °C for 60 min) the activity before running the samples on native gels. Neither of these treatments significantly changed the native gel patterns described above, suggesting that the interactions between the TcAP 2 polypeptide and the trypsin inhibitor protein and/or the 27.9 kDa polypeptide do not require the proteolytically active conformation of TcAP2. Other attempts to break up both complexes detected on native gels, by adding a large amount of a non-specific protein (β-lactoglobulin) as a competitor and by adding Triton X-100 (up to 20 g l^{-1} in final loaded sample), were also not successful. Finally, addition of high levels of a Fast Flow Q fraction which was highly enriched in the trypsin inhibitor protein did not cause any change in the native gel pattern, except that the added trypsin inhibitor protein migrated as a new band just ahead of band 2 (data not shown). Overall, these results indicate that the various complexes which form between the TcAP2 polypeptide and the trypsin inhibitor protein and/or the 27.9 kDa polypeptide each have significant specificity and stability.

Partial purification of cacao seed vicilin polypeptides and their hydrolysis by purified TcAP2

To study the generation of cocoa flavour precursor peptides and amino acids from cacao vicilin, it was necessary first to obtain an enriched preparation of the cacao vicilin. Using a similar method to that described by Voigt *et al*,[19] a polypeptide preparation that was highly enriched in the vicilin polypeptides of 48.5 and 34.1 kDa was obtained (Fig 5(a), lane V). To determine the sensitivity of cacao seed vicilin to the purified aspartic proteinase, the partially purified vicilin preparation was reacted with the purified TcAP2. The data obtained showed that the two vicilin polypeptides, and all the other less abundant proteins present in the reaction mixture,

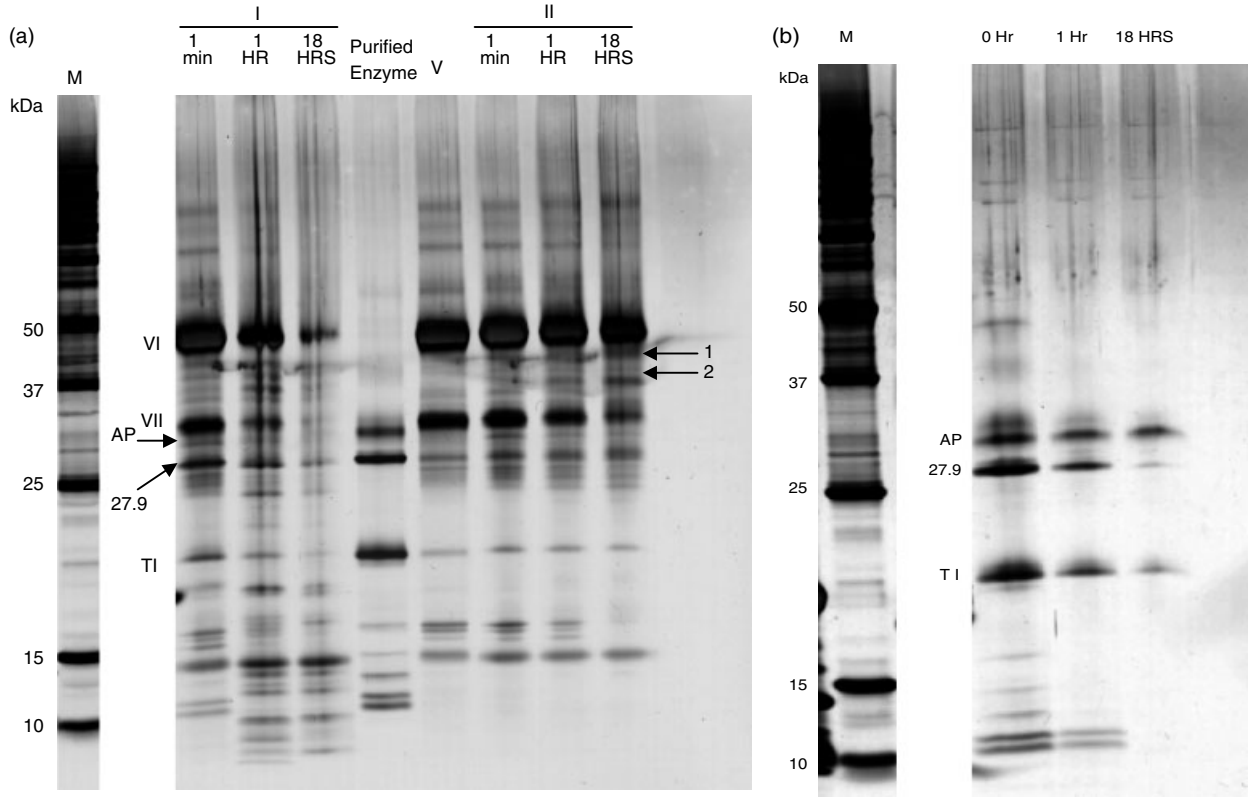

Figure 5. Protein hydrolysis by purified TcAP2. (a) Set I, partially purified vicilin was incubated with purified TcAP2 complex for the times indicated and the reaction products were analysed by SDS-PAGE; set II, partially purified vicilin incubated alone; V, partially purified vicilin preparation; VI and VII, 48.5 and 34.1 kDa vicilin polypeptides respectively; AP, 30.5 kDa TcAP2; TI, trypsin inhibitor; 1 and 2, major degradation fragments of the larger vicilin polypeptide VI; M, molecular weight markers. (b) The purified TcAP2 complex has been allowed to autodigest for the times indicated. The symbols are as noted for (a); 27.9 marks the position of the 27.9 kDa polypeptide.

were substantially degraded by the TcAP2 activity after an 18 h treatment. The majority of the peptide products generated were less than 10 kDa in size (Fig 5(a), set I). In fact, many of the peptide digestion products are likely to be much smaller than 10 kDa and were thus lost from the gel during the fixation and staining steps or, if present, were stained poorly owing to their small size. Further attempts to resolve the small peptides on Tris/tricine gels specific for peptide separation[22] did not give significantly better detection of the small vicilin derived peptides.

A control reaction without the enzyme showed that there was only a small amount of vicilin degradation in the absence of added protease during an 18 h incubation (Fig 5(a), set II). This result demonstrates that the majority of the vicilin degradation seen in this experiment was the result of specific digestion by the purified TcAP2. The low level of vicilin degradation seen in the enzyme-free control after 18 h was due to a very low level of aspartic proteinase activity in the vicilin preparation, because, when pepstatin was added in a control experiment, all the vicilin hydrolysis was inhibited. In other experiments we have confirmed that a low level of aspartic proteinase activity often exists in our partially purified vicilin preparations. This low level of activity cannot be detected by the standard protease assay but can be detected by SDS-PAGE analysis, which shows both the loss of pre-existing bands and the generation of new bands (products) after a long incubation time. We have also found that this residual activity can be eliminated entirely by a short heat treatment. In a further control reaction, incubation of the purified TcAP2 alone at pH 3.0 resulted in slow autohydrolysis of TcAP2 as well as a significant level of degradation of the trypsin inhibitor and 27.9 kDa polypeptides (Fig 5(b)). We noted that the 30.5 kDa TcAP2 polypeptide was much more resistant to degradation than the other co-purifying polypeptides.

The experiments just described show directly that the purified TcAP2 activity hydrolyses vicilin, the trypsin inhibitor and the 27.9 kDa polypeptides to peptides smaller than 10 kDa. To produce a better estimation of the size of the peptides generated by the purified TcAP2, further analysis of the digestion products was carried out using denaturing size exclusion column chromatography with a column specific for the resolution of very short peptides. However, for this experiment it was necessary to use significantly more enzyme activity than that used for the analysis of vicilin degradation by SDS-PAGE. Given the difficulty of obtaining large amounts of highly purified TcAP2, this particular experiment was carried out with the slightly less pure TcAP2 preparation, ie the pooled active fractions off the Q Sepharose Fast Flow column. The SDS-PAGE analysis of the proteins before and after the reaction is seen in Fig 6(a). It is clear from this experiment that the majority of the starting protein had been

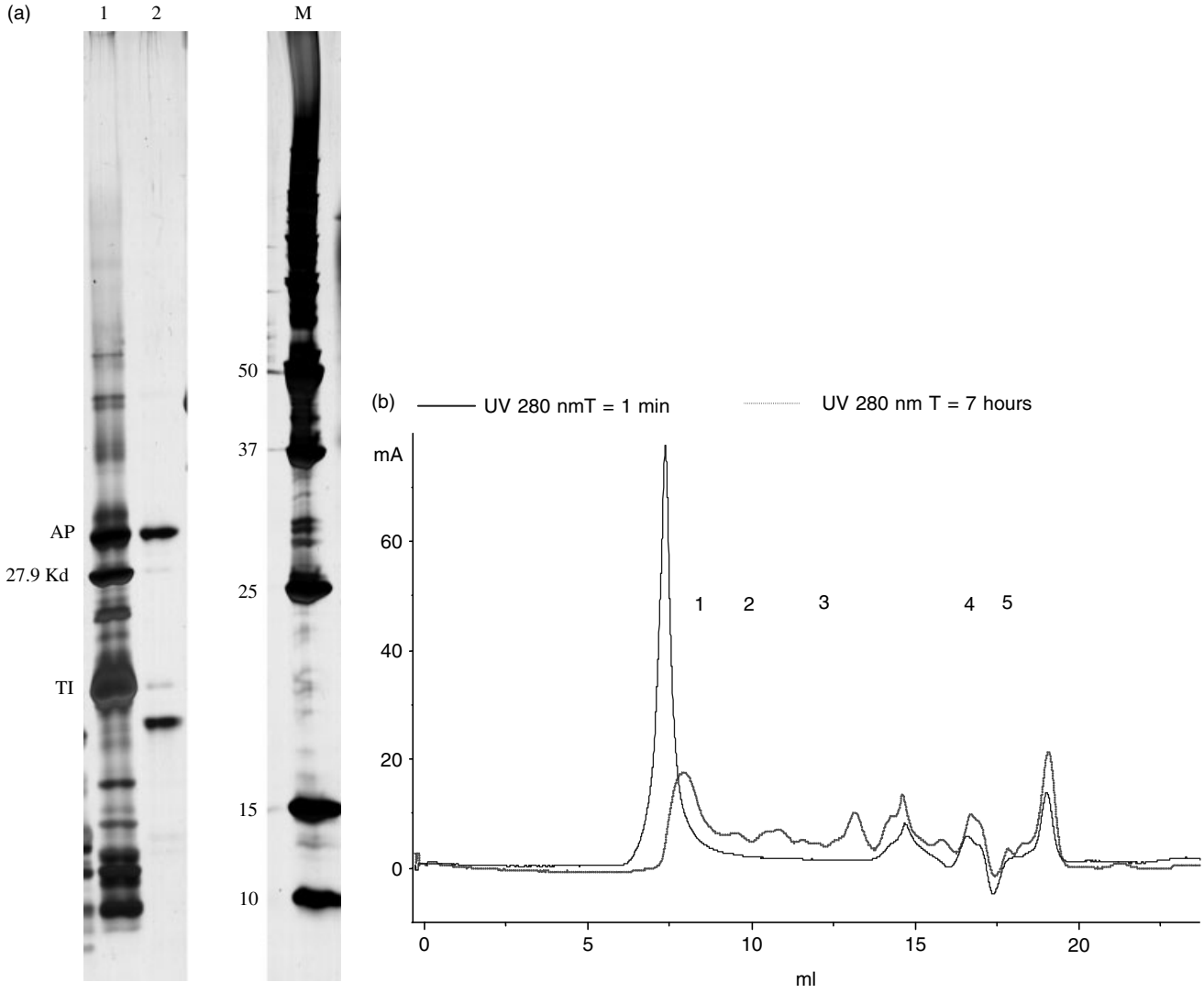

Figure 6. Autodigestion (for 1 min or 7 h) of Fast Flow Q-purified TcAP2. (a) Silver-stained 100–200 g l^{-1} SDS-PAGE gel analysis of the reaction products. AP, 30.5 kDa TcAP2 aspartic proteinase; 27.9, 27.9 kDa polypeptide; TI, trypsin inhibitor; M, molecular weight markers. (b) Denaturing size exclusion chromatography of the reaction products. The molecular weight size markers are: 1, ribonuclease A 13.7 kDa; 2, aprotinin 6.5 kDa; 3, substance P 1347 Da; 4, *N*-benzoyl-gly-phe (hippuryl-phe) 326 Da; 5, phenylalanine 165 Da.

extensively hydrolysed and that the resulting peptides were generally lost from the gel and/or stained poorly owing to their small size. The only polypeptides remaining after the 7 h reaction time were the 30.5 kDa TcAP2 polypeptide and another polypeptide at approximately 17–19 kDa that was not seen in the 1 min sample and was probably a degradation product of the trypsin inhibitor protein. Closer examination of the original gel showed that very low levels of the 27.9 kDa and trypsin inhibitor polypeptides remained after 7 h and that there were also a few weak bands between 10 and 15 kDa. Importantly, the fact that there was no significant and selective enrichment of a low-molecular weight aspartic proteinase associated polypeptide after the autodigestion reactions in Fig 5(b) or 6(a) is consistent with our proposal that the 30.5 kDa TcAP2 polypeptide is active as a monomeric aspartic proteinase.

These two TcAP2 hydrolysed protein samples were further examined by denaturing size exclusion chromatography using a column specific for separating small peptides. The results are presented in Fig 6(b). The majority of the material absorbing at 280 nm in the 1 min sample corresponded to polypeptides significantly greater in size than 13.7 kDa, in agreement with the gel analysis, showing that the majority of the protein in this sample was either 20.2 kDa (trypsin inhibitor) or larger. However, after 7 h at 42 °C the main peak shifted significantly in the direction of the lower-molecular weight material and was markedly reduced in total area. The remaining material absorbing at 280 nm appeared to be distributed relatively evenly over a range of peptide sizes stretching from approximately 45–60 amino acids to approximately di- and tripeptides. It is important to note that the amount of small peptides present after the 7 h reaction was clearly underestimated, because peptides without a phenolic ring (phenylalanine, tyrosine and tryptophan) are very poorly detected at 280 nm. Unfortunately, the scan at 220 nm (which

detects peptide bonds in all peptides) was not useful, as there appeared to be a large background peak at this wavelength in the region of the very small peptides for the 1 min sample. It is possible that this background absorbance was due to small amounts of polyphenols bound non-covalently, and probably to a lesser extent covalently, to the proteins in the Q Sepharose Fast Flow purified TcAP2 sample. Nonetheless, despite the fact that detection at 280 nm does not give a complete analysis of the hydrolysis products, the results obtained confirm the conclusion made from SDS-PAGE analysis; that is, the TcAP2 aspartic proteinase can hydrolyse cacao seed proteins into very small peptides.

DISCUSSION

The results presented here describe the purification of TcAP2, the major aspartic proteinase activity in *T cacao* seeds. Despite considerable effort using several columns, it has not been possible to obtain significant quantities of a homogeneously pure TcAP2 protein. In addition to the single 30.5 kDa TcAP2 aspartic proteinase polypeptide, the most highly purified preparation contained two other polypeptides, the trypsin inhibitor protein (plus a low level of trypsin inhibitor fragments) and a 27.9 kDa putative class III acid chitinase. Based on the results presented here, we propose that the purified aspartic proteinase is probably a mixture of two co-purifying complexes. One, the active complex, contains the monomeric 30.5 kDa kDa TcAP2 aspartic proteinase polypeptide and the trypsin inhibitor. The other contains the trypsin inhibitor and the 27.9 kDa putative acid chitinase. Supporting evidence for the presence of two complexes comes from the size exclusion column elution profile (Fig 1). This shows that three distinct protein peaks eluted from this column. The third and last peak eluting contained two polypeptides, the trypsin inhibitor and the 27.9 kDa putative acid chitinase, and these polypeptides form two stable complexes which were visualised on a native gel (Fig 4). The second peak eluting from the size exclusion column had most of the activity and had three polypeptides, the 30.5 kDa TcAP2 polypeptide and the two polypeptides seen in the third peak, together with a low level of trypsin inhibitor cleavage fragments. Native PAGE gel analysis showed that the second peak contained a low level of the two complexes between the trypsin inhibitor and the 27.9 kDa polypeptide which were seen in peak three, plus two new complexes which we propose are composed of the TcAP2 30.5 kDa polypeptide and the trypsin inhibitor (Fig 4). The possibility that the three polypeptides found in the second peak exist in a single complex cannot be ruled out by the data available. However, the apparent size of this complex on the native size exclusion column (calculated to be approximately 50 kDa) does not support this model. Another idea is that the activity in the second peak was simply a non-specific aggregate between the TcAP2 polypeptide, the trypsin inhibitor protein and the 27.9 kDa putative acid chitinase. We do not believe that this is likely for a number of reasons. First, if the purified activity were simply a non-specific aggregate, one would expect most of the aspartic proteinase activity to elute from the size exclusion column in a region corresponding to a molecular weight considerably greater than 50 kDa. Second, aggregated proteins would also be likely to merely smear on a native gel rather than produce the specific bands seen in Fig 4. Finally, when large amounts of partially purified trypsin inhibitor protein were added to the purified TcAP2 activity, no change in the native PAGE gel pattern was observed. This result demonstrates that no aggregation was produced when substantial amounts of trypsin inhibitor protein were added exogenously.

The existence of a complex between a seed aspartic proteinase and abundant seed proteins has been reported previously by Bourgeois and Malek.[26] They isolated a complex of >300 kDa which contained seven polypeptides from dried jack pine seeds. The proteins in this complex were not characterised by protein sequencing, although at least two of them were abundant on an SDS-PAGE gel analysis of the original crude extract and thus are likely to be storage proteins. Aspartic proteinases and trypsin inhibitor polypeptides have both been shown to reside in the storage vacuole.[5,6,27] Also, a very abundant acid chitinase polypeptide has been found in the storage vacuole of banana, and it has been proposed that this protein functions as a storage protein.[28] Because it appears that the three polypeptides associated with the highly purified aspartic proteinase activity are vacuolar proteins, and owing to the fact that they co-purified together over several columns, it is possible that they are normally associated in the vacuole, perhaps to facilitate protein packaging. It is also possible that the existence of a complex between TcAP2 and the trypsin inhibitor may somehow limit the mobility and/or the activity of the TcAP2 polypeptide in *T cacao* seed vacuoles *in vivo* and thus protect these seeds from the unusually high aspartic protease activity which exists in them.[13]

Using a similar purification strategy, plus an additional affinity step on a pepstatin column, Biehl and co-workers[10] have previously purified an aspartic proteinase from *T cacao* seeds that has a different structure (heterodimer of 29 and 13 kDa) and different pH and temperature optima from the activity described here. Because no protein sequence data were obtained for either polypeptide of this purified dimeric enzyme, the relationship between that protein and the TcAP2 polypeptide described here is not known. It should be noted that the majority of the plant seed aspartic proteinases characterised to date have been purified using pepstatin affinity chromatography. However, only a small number of the members of this potentially large protein family

have been purified for any one plant species. As discussed earlier, when we tested pepstatin affinity chromatography to purify the TcAP2 activity further, we were unable to recover significant levels of activity from the column. Interestingly, a similar observation has been made recently by Domingos *et al*[29] while purifying an aspartic proteinase from the flowers of *Centura calcitrapa*. This suggests that certain aspartic proteinases may have an unusual interaction with the pepstatin column matrix and are lost during this type of chromatographic step. Considering the fact that it is believed that there could be >60 members of the aspartic proteinase family in plants (BeersE, personnal communication), it is easy to imagine that one or more of the plant aspartic proteinases could interact in a substantially different fashion with the pepstatin column matrix. Further work with other purified plant aspartic proteinase family members is clearly necessary to answer this and other questions concerning the functions of the different members.

Beihl and co-workers have proposed that most of the peptides and amino acids generated during the fermentation of cacao seeds are due to the concerted action of an aspartic proteinase and a carboxypeptidase (see Ref 18 and references cited therein). Cacao seeds that are not fermented have very poor flavour quality. Although it has been shown that a partially purified cacao seed aspartic proteinase could hydrolyse cacao seed storage proteins,[11] and that this protease activity was directly involved in the production of cocoa flavour precursors,[19] the size range of the peptides produced has not yet been determined. Here we show that the monomeric TcAP2 polypeptide, the main aspartic proteinase activity of cacao seeds, is capable of efficiently degrading seed storage proteins into peptides that range in size from 65 amino acids down to di- and tripeptides *in vitro*. Because small peptides are believed to be important cocoa flavour precursors, it is possible that the aspartic proteinase activity actually has a more significant role in the production of cocoa flavour precursors than has been previously recognised. This possibility is supported by the observation that the carboxypeptidase activity in the mature seeds is quite weak relative to the aspartic proteinase activity (Tazi H and Laloi M, unpublished observations). Further work on the *in vitro* digestion of storage proteins with purified TcAP2 and the seed carboxypeptidase, followed by a model flavour reaction,[19] should help to clarify the precise role of the TcAP2 protein in the production of cocoa flavour precursors during cocoa bean fermentation. Finally, it should be noted that the TcAP2 activity is relatively heat labile. This property could play a signficant role in the outcome of a cocoa bean fermentation. For example, if the internal temperature of a cocoa bean heap were to rise too quickly to the 50–55 °C range during fermentation, it is possible that there could be a substantial loss of aspartic proteinase activity in the beans before enough acid has been produced to induce significant acid protease mediated protein degradation. In this case, although a relatively 'normal' fermentation process could have ostensibly occurred, the flavour potential of the resulting beans would be less than optimal. To investigate this point further, it will be necessary to follow the heap temperature, acid production, total protein degradation and the overall aspartic proteinase activity in beans during the fermentation process under various fermentation conditions. These data could then be correlated with the flavour qualities of the cocoa beans obtained after each different fermentation condition examined.

CONCLUSIONS

The major aspartic proteinase activity of *T cacao* seeds has been purified in the form of a protein complex and characterised. The purified complex contains a 30.5 kDa aspartic proteinase polypeptide (product of the TcAP2 gene) and an associated 20.5 kDa trypsin inhibitor polypeptide. The physical and biochemical properties of the active *T cacao* seed TcAP2 aspartic proteinase complex are different from a previously purified *T cacao* seed aspartic proteinase, suggesting that the highly expressed product of the TcAP2 gene may represent a previously uncharacterised activity. The purified TcAP2 complex efficiently degrades cacao seed storage proteins such as vicilin into peptides as small as di- and tripeptides, implying that the TcAP2 gene product plays an important role in cocoa flavour precursor generation during cacao seed fermentation. It will be interesting in the future to determine if the expression of TcAP2 varies in different *T cacao* varieties and whether a high level of TcAP2 expression is associated with higher cocoa flavour potential of specific varieties.

ACKNOWLEDGEMENTS

We would like to thank J Rogers, C Hansen and S Kochhar for helpful discussions. We would also like to thank M Ben-Amor for his comments on the manuscript and, especially, V Petiard for his advice and support. Finally, we wish to thank J d'Alayer of the Institut Pasteur for his skilled performance of the N-terminal sequencing.

REFERENCES

1 Callis J, Regulation of protein degradation. *Plant Cell* 7:845–857 (1995).
2 Mutlu A and Gal S, Plant aspartic proteinases; enzymes on the way to function. *Physiol Plant* **105**:569–576 (1999).
3 Kervinen J, Tormakangas K, Runeberg-Roos P, Guruprasad K, Blundell T and Teeri T, Structure and possible function of aspartic proteinases in barley and other plants, in *Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications*, ed by Takahashi K. Plenum, New York, pp 241–254 (1994).
4 Rawlings ND and Barret AJ, Families of aspartic peptidases, and those of unknown catalytic mechanism. *Methods Enzymol* **248**:105–120 (1995).

5 Mutlu A, Chen X, Reddy SM and Gal S, The aspartic proteinase is expressed in *Arabidopsis thaliana* seeds and localized in the protein bodies. *Seed Science Res* **9**:75–84 (1999).
6 Runeberg-Roos P, Tormakangas K and Ostman A, Primary structure of a barley-grain aspartic proteinase. *Eur J Biochem* **202**:1021–1027 (1991).
7 Tormakangas K, Kervinen J, Ostman A and Teeri T, Tissue specific localization of aspartic proteinase in developing and germinating barley grains. *Planta* **195**:116–125 (1994).
8 Runeberg-Roos P, Kervinen J, Kovaleva V, Gal S and Raikhel NV, The aspartic proteinase of barley is a vacuolar enzyme that processes probarley lectin in-vitro. *Plant Physiol* **105**:321–329 (1994).
9 Biehl B, Heinrichs H, Ziegeler-Berghausen H, Srivastava S, Xiong Q, Passern D, Senyuk V and Hammoor M, The proteases of ungerminated cocoa seeds and their role in the fermentation process. *Angew Bot* **67**:59–65 (1993).
10 Voigt J, Kamaruddin S, Heinrichs H, Wrann D, Senyuk V and Biehl B, Developmental stage-dependent variation of the levels of globular storage protein and aspartic endoproteinase during ripening and germination of *Theobroma cacao* seeds. *J Plant Physiol* **145**:299–307 (1995).
11 Voigt J, Voigt G, Heinrichs D, Wrann D and Biehl B, *In vitro* studies on the proteolytic formation of the characteristic aroma precursors of fermented cocoa seeds: the significance of endoprotease specificity. *Food Chem* **51**:7–14 (1994).
12 Laloi M, McCarthy J, Morandi O, Gysler C and Bucheli P, Molecular and biochemical characterization of two proteinases TcAP1 and TcAP2 from *Theobroma cacao* seeds. *Planta* **215**:754–762 (2002).
13 Biehl B, Voigt J, Voigt G, Heinrichs H, Senyuk V and Bytof G, pH dependent enzymatic formation of oligopeptides and amino acids, the aroma precursors in raw cocoa beans. *Proc 11th Int Cocoa Research Conf, 18–24 July, 1993, Yamoussoukro, Ivory Coast*, pp 717–721 (1994).
14 Carr JG, Cocoa, in *Fermented Foods*, Vol 7 of *Economic Microbiology*, ed by Rose AH. Academic Press, New York, pp 275–292 (1982).
15 Biehl B, Passern D and Sagemann W, Effect of acetic acid on subcellular structures of cocoa bean cotyledons. *J Sci Food Agric* **33**:1101–1109 (1982).
16 Biehl B, Brunner E, Passern D, Quesnel VC and Adomako D, Acidification, proteolysis and flavour potential in fermenting cocoa beans. *J Sci Food Agric* **36**:583–598 (1985).
17 Rohan T, The precursors of chocolate aroma: a comparative study of fermented and unfermented cocoa beans. *J Food Sci* **29**:456–459 (1964).
18 Voigt J and Biehl B, Precursors of the cocoa specific aroma components are derived from the vicilin-class (7S) globulin of the cocoa seeds by proteolytic processing. *Bot Acta* **108**:283–289 (1995).
19 Voigt J, Heinrichs H, Voigt G and Biehl B, Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. *Food Chem* **50**:177–184 (1994).
20 Hansen C, del Olmo M and Burri C, Enzyme activities in cocoa beans during fermentation. *J Sci Food Agric* **77**:273–281 (1998).
21 Church F, Swaisgood H, Porter D and Catignani G, Spectrophotometric assay using *o*-phthaldialdehyde for determination of proteolysis in milk and isolated milk proteins. *J Dairy Sci* **66**:1219–1227 (1983).
22 Schagger H and von Jagow G, Tricine–sodium dodecyl sulphate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**:368–379 (1987).
23 Kochhar S, Gartenmenn K and Juillerat M, Primary structure of the abundant seed albumin of *Theobroma cacao* by mass spectrometry. *J Agric Food Chem* **48**:5593–5599 (2000).
24 Faro C, Moir A and Pires E, Specificity of a milk clotting enzyme extracted from the thistle *Cynara cardunculus* L.; action on oxidized insulin and K-casein. *Biotechnol Lett* **14**:841–846 (1992).
25 Park H, Yamanaka N, Mikkonen A, Kusakabe I and Kobayashi H, Purification and characterization of aspartic proteinase from sunflower seeds. *Biosci, Biotechnol Biochem* **64**:931–939 (2000).
26 Bourgeois J and Malek L, Purification and characterization of an aspartyl proteinase from dry jack pine seeds. *Seed Sci Res* **1**:139–147 (1991).
27 Horisberger M and Tacchini-Vonlanthen M, Ultrastructural localization of Bowman–Birk inhibitor on thin sections of *Glycine max* (soyabean) cv. Maple Arrow by the gold method. *Histochemistry* **77**:37–50 (1983).
28 Clendennen S, Lopez-Gomez R, Gomez-Lim M, Arntzen C and May G, The abundant 31-kilodalton banana pulp protein is homologous to class-III acidic chitinases. *Phytochemistry* **47**:613–619 (1998).
29 Domingos A, Cardoso P, Xue Z, Clemente A, Brodelius P and Pais M, Purification, cloning, and autoproteolytic processing of an aspartic proteinase from *Centaurea calcitrapa*. *Eur J Biochem* **267**:6824–6831 (2000).