Construction, expression and characterization of a chimaeric mammalian–plant aspartic proteinase

Kenneth G. PAYIE*¹, Takuji TANAKA*, Susannah GAL† and Rickey Y. YADA*²

*Department of Food Science, University of Guelph, Guelph, ON N1G 2W1, Canada, and †Department of Biological Sciences, The State University of New York at Binghamton, Binghamton, NY 13902-6000, U.S.A.

Aspartic proteinases are a well-characterized class of proteinases. In plants, all nascent aspartic proteinases possess a 100-amino-acid, plant-specific sequence (PSS) within their C-terminal lobe, presumed to possess a targeting role *in vivo*. In this study, the PSS domain from the *Arabidopsis thaliana* aspartic proteinase was inserted into porcine pepsinogen at the identical location found in nascent plant aspartic proteinases, to create a chimaeric mammalian–plant enzyme. Based on enzymic activity, this chimaeric enzyme demonstrated increases in pH stability above 6 and temperature stability above 60 °C compared with commercial pepsin. Differential scanning calorimetry of the chimaeric enzyme illustrated an approx. 2 °C increase in denaturation temperature (T_m), with increases in co-operativity and similar enthalpy values.

Kinetic analysis indicated an increase in K_m and decreased k_{cat} compared with pepsin, but with a catalytic efficiency similar to the monomeric plant aspartic proteinase from wheat. Using oxidized insulin B-chain, the chimaeric enzyme demonstrated more restricted substrate specificity in comparison with commercial pepsin. This study highlights the use of a chimaeric enzyme strategy in order to characterize unique protein domains within enzyme families, and, for the first time, a putative structure–function role for the PSS as it pertains to plant aspartic proteinases.

Key words: aspartic proteinase, chimaeric enzyme, pepsin, structure–function.

INTRODUCTION

The aspartic proteinases (E.C. 3.4.23.X) are a well-studied class of enzymes, and knowledge of their structure and mechanism is extensively documented [1,2]. Proposed to be ubiquitous, the aspartic proteinases constitute a heterologous class of proteinases with a varied range of activity and specificity. A number of common characteristics are shared: they possess two catalytic aspartate residues localized in two short amino-acid stretches with sequence homology with one another, similar three-dimensional structures, a general catalytic activity at acidic pH values, a scission preference between large hydrophobic amino acids and inhibition by the naturally occurring compound pepstatin [1,2].

Unlike their animal and microbial counterparts, the plant aspartic proteinases have received limited attention [3]. Few have been characterized biochemically, but common within the nascent sequence of those that have is the presence of a 100-amino-acid plant-specific sequence (PSS) completely novel to the plant kingdom. The discovery of the PSS has prompted the elucidation of its three-dimensional structure; a bilobar fold with a highly α -helical and glycosylated PSS within the C-terminal portion of the enzyme [4]. All active plant aspartic proteinases possess a PSS in their nascent polypeptide sequence [5] but, surprisingly, the PSS has shown limited homology between plant sources [3,6,7].

Plant aspartic proteinases are found in either heterodimeric or monomeric forms [3,5]. Heterodimer formation is a result of post-translational proteolytic modification of the nascent enzyme resulting in the removal of most, or all, of the PSS prior to activation [8,9]. In contrast, monomeric plant aspartic proteinases retain their PSS structure. It is unclear as to the reason(s) for excision of the PSS. The PSS shares approx. 30 % identity with

a class of small activator proteins called saposins that associate with procathepsin D (human aspartic proteinase). It has been speculated that saposins are involved in targeting procathepsin D to the heavy lysosome for processing [10]. Due to the putative targeting function and relative homology of the saposin to the PSS, the PSS domain may perform a similar role for delivery of the plant aspartic proteinase to the vacuole [4,9,11–13]. However, contrary views have been suggested [5,8,14].

Evidence suggesting that this domain possesses a structure–function relationship for these enzymes has not been presented in the literature. To this end, the construction of a chimaeric aspartic proteinase was undertaken to assess a putative structure–function role of this domain in these enzymes. The chimaeric enzyme was designed using the PSS derived from the *Arabidopsis thaliana* aspartic proteinase [15] and inserting it into porcine pepsinogen. The heterologous expression and characterization of the chimaeric enzyme demonstrated moderate increases in stability with respect to pH and temperature, as well as modified and restricted substrate specificity compared with wild-type pepsin. Accordingly, these data suggest that the PSS could provide a structure–function role within plant aspartic proteinases and highlights the relationship of unique protein domains within enzyme families.

EXPERIMENTAL

Reagents and organisms

Restriction enzymes, *Pwo* polymerase, DNA ligase and peptide N-glycosidase F (PNGase F) were purchased from Roche Biochemicals (Laval, Quebec, Canada). Pepsinogen,

Abbreviations used: SS1, synthetic substrate 1; PSS, plant-specific sequence; DSC, differential scanning calorimetry; T_m , denaturation temperature; RP-HPLC, reversed-phase HPLC; PNGase F, peptide N-glycosidase F.

¹ Present address: Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, U.S.A.

² To whom correspondence should be addressed (e-mail ryada@uoguelph.ca).

oxidized insulin B-chain and methyl-D- α -mannopyranoside were purchased from Sigma (St. Louis, MO, U.S.A.). Concanavalin A was purchased from Amersham Biosciences (Baie D'Urfé, Quebec). *Escherichia coli* TOP10F' and the *Pichia pastoris* gene-expression system, containing the host strain (KM71) and vector pHIL-S1, were purchased from Invitrogen (San Diego, CA, U.S.A.). All other chemicals were of analytical grade for biochemical use.

Cloning and mutagenesis

The following oligonucleotide primers were synthesized for the QuikChange mutagenesis procedure (Stratagene, La Jolla, CA, U.S.A.). Porcine-N-1, 5'-C GAC ATC GGA GCG GCC GCG ACC TCA TAT GG-3'; Porcine-N-2, 5'-CC ATA TGA GGT CGC GGC CGC TCC GAT GTC G-3'; Porcine-C-1, 5'-GGA GCA GCT TAA GAC CAT CTC GCC-3'; Porcine-C-2, 5'-GGC GAG ATG GTC TTA AGC TGC TCC-3'; PSS-N-1, 5'-CAT GCT ATT GGA GCG GCC GGA GTT GTT AGC-3'; PSS-N-2, 5'-GCT AAC AAC TCC GGC CGC TCC AAT AGC ATG-3'; PSS-C-1, 5'-G GGA GAG TCT GCC TTA AGC TGT GCA CAA CTG-3'; PSS-C-2, 5'-CAG TTG TGC ACA GCT TAA GGC AGA CTC TCC-3'.

Non-mutagenesis oligonucleotide primers used for amplification of the porcine pepsinogen and the mutated porcine pepsinogen sequence from the pPSP2000 and pPSP2000KP vectors, respectively, were: *Xho*I-R, 5'-CGC TCG AGT ATT CAT GC-3'; *Eco*R-F, 5'-TCA GAA TTC TCT ACA TGG AAG-3'. The cDNA for the porcine pepsinogen was subcloned into the *Xho*I and *Eco*RI restriction sites of the expression vector pHIL-S1 resulting in plasmid pPSP2000, which was used to transform *E. coli* TOP10F' cells [16]. Selection of the transformants containing the pHIL-S1 plasmid were made using LB agar containing ampicillin (150 μ g/ml). Plasmid DNA was purified using a plasmid-isolation kit (Roche Biochemicals) from the selected colonies and mutagenesis was performed using the QuikChange methodology as per the manufacturer's instructions. Plasmid DNA was isolated and screened for successful mutagenesis by linearization using the engineered restriction-endonuclease sites. Once successfully screened by restriction digestion, plasmids were sequenced for confirmation of mutagenesis. The plasmid was renamed pPSP2000KP.

The *A. thaliana* aspartic proteinase was cloned previously into pBlueScript SK⁻ vector (Stratagene) and named pBSG2J2 [15]. The plasmid was used as a template for the QuikChange method using the complimentary primers PSS-N-1 and PSS-N-2, and PSS-C-1 and PSS-C-2. Once successfully screened by restriction digestion, plasmids were sequenced for confirmation of mutagenesis. The plasmid was renamed pBSG2J2KP.

PSS insertion into porcine pepsinogen

Chimaeric construction was performed by ligating three independent DNA portions, corresponding to the N-terminal fragment of porcine pepsinogen, the PSS and the C-terminal fragment of porcine pepsinogen, in a stepwise fashion. The successful ligation was purified and amplified using the *Eco*R-F and *Xho*I-R primers. The chimaeric DNA fragment was purified from the agarose gel and double digested with *Eco*RI and *Xho*I restriction endonucleases. The chimaeric sequence was ligated into the pHIL-S1 vector. The ligation reaction was transformed into *E. coli* TOP10F' and selected using LB agar with ampicillin (150 μ g/ml). Colonies were selected, screened by restriction

digestion and sequenced. The plasmid containing the chimaeric aspartic proteinase sequence was renamed pKGP2.

Expression using *P. pastoris*

The pKGP2 plasmid was linearized using the restriction endonuclease *Sal*I, electroporated into *P. pastoris* (KM71) cells according to the manufacturer's instructions (Invitrogen) and expressed as described in [14]. The clones that produced supernatants that were positive upon activation for milk-clotting activity and via Western blot, using the anti-pepsinogen antibody, were frozen at -86°C in the presence of sterile glycerol (25 % final concentration). One clone (KP1-2), which was positive using both screening methods, was grown on a large scale and used for further purification and characterization (see below). The genomic DNA of *P. pastoris* KP1-2 was isolated using a Qiagen DNA-extraction kit. Amplification of the integrated region containing the chimaeric aspartic proteinase sequence was performed using the 5' AOX1 and 3' AOX1 primers (Invitrogen). The amplified fragment was purified from the agarose gel and confirmed to contain the complete chimaeric sequence by sequencing.

Chimaeric aspartic proteinase purification

After 72 h of incubation in BMMY medium (buffered methanol complex medium) at 30°C (supplemented with 0.75 % methanol every 24 h), the *P. pastoris* cells were pelleted at 3000 g for 5 min at 4°C . The supernatant (containing the chimaeric enzyme) was removed and concentrated via ultrafiltration using a 30 kDa molecular-mass-cut-off membrane and washed with 1 litre of 20 mM Tris/HCl. The concentrated post-(NH_4)₂SO₄ material (50 ml) was applied five times to a 20 ml concanavalin A-Sepharose column and the column washed with 20 mM Tris/HCl, pH 7.5, with 0.5 M NaCl. Elution was performed using 0.4 M methyl-D- α -mannopyranoside in 20 mM Tris/HCl, pH 7.5, with 0.5 M NaCl. The eluted material was concentrated using a Centricon YM-30 with 20 mM Tris/HCl, pH 7.5. Anion-exchange chromatography was performed on the post-concanavalin A-Sepharose sample using an HR 5/5 Source Q column (Amersham Biosciences) equilibrated with 20 mM Tris/HCl and 30 mM NaCl, pH 7.5, at a flow rate of 1 ml·min⁻¹. Elution from the column was performed using an increasing NaCl gradient from 0.03 to 1 M over 80 min. Fractions were collected manually, activated and assessed for milk-clotting activity. Those fractions possessing activity were concentrated (Centricon YM-30) and tested via Western blot using an anti-pepsinogen antibody, and used for further enzyme characterization.

SDS/PAGE and Western blotting

All steps of the purification were assessed by SDS/PAGE [17] using 10 % Bio-Rad Ready Gels and stained using either SYPRO Ruby Gel stain or SYPRO Orange G stain (Bio-Rad), or transferred to a PVDF membrane (Bio-Rad) for Western blotting [18]. For N-terminal analysis, the post-electroblotting membrane was stained with Coomassie Brilliant Blue R solution and prepared as described in [19].

Enzyme assays and activation

Activation of both the chimaeric aspartic proteinases and commercial porcine pepsinogen was initiated unimolecularly as described in [20] for 30 min periods, and the released

prosegment removed using a Centricon YM-10 filtration unit (10 kDa molecular-mass-cut-off). Protein concentrations were assessed in duplicate using the Bio-Rad DC assay and absorbance calculations at 280 nm [21]. Preparations of the chimaeric aspartic proteinase and commercial pepsinogen were activated and kinetic measurements were made using the synthetic substrate 1 [SS1; Lys-Pro-Ala-Glu-Phe-Phe(NO₂)-Ala-Leu], synthesized at Queen's University, Kingston, ON, Canada [20]. For kinetic studies, a minimum of 10 SS1 concentrations in the range 0.02–0.35 mM, pH 1.5, at 25 °C (determined to be the pH optimum of the chimaeric enzyme) were used. The analysis consisted of monitoring the change in absorbance at 300 nm using a DU 640 Spectrophotometer (Beckman Instruments, Fullerton, CA, U.S.A.), where progress curves were measured to give a change in absorbance/min. K_m and V_{max} values were estimated using non-linear regression [20]. The k_{cat} values were estimated using the estimated V_{max} values and the protein concentration from spectrophotometric measurements at 280 nm [21]. Three independent samples assessed in triplicate were used for kinetic constant determination with means and S.D. calculated.

Stability studies

The purified pepsin, or chimaeric enzyme, was activated as described in the previous section and tested for pH or temperature stability. Activity measurements using SS1 were obtained by independently incubating each enzyme at pH 6.5, 7.0 and 7.5 using 0.4 M phosphate buffer for 10 min. The pH was adjusted to 5.2, using 3 M sodium acetate. Temperature stability was determined by incubation at specified temperatures (35, 45, 55, 65 and 75 °C) for 10 min and subsequent adjustment of the temperature back to 25 °C. All measurements were performed in triplicate with three replicates and comparisons were made to the control conditions (pH 5.2 and 25 °C, respectively).

Hydrolytic study

Oxidized insulin B-chain (100 µg) was dissolved in 1 ml of 0.1 M sodium citrate, pH 1.5, and incubated with activated chimaeric enzyme or activated pepsin (100:1, w/w) at 37 °C [22]. Aliquots were removed at 1, 3, 7, 24 and 48 h intervals and frozen at –25 °C to stop the reaction. The peptides were analysed by reversed-phase HPLC (RP-HPLC) with a µBondapak C₁₈ RP column (Waters Associates, Milford, MA, U.S.A.) using a linear gradient of acetonitrile from 3 to 80 % in 0.1 % trifluoroacetic acid over 75 min at 0.7 ml·min^{–1}. The eluant was monitored at 215 nm and the peaks detected were manually collected and dried using a vacuum concentrator and submitted for amino acid analysis. Control reactions were run using identical conditions without the enzyme and monitored from 1 to 24 h.

Differential scanning calorimetry (DSC)

Calorimetric measurements were carried out using a MicroCal MC-2 differential scanning calorimeter (MicroCal, Northampton, MA, U.S.A.). The activated chimaeric aspartic proteinase and pepsin samples (0.6–0.8 mg/ml final concentrations) were subjected to calorimetric measurements from 20 to 100 °C at a heating rate of 1.5 °C/min. Data analysis was performed using the Origin DSC ITC software (MicroCal). Runs were performed in duplicate.

Figure 1 Schematic and amino acid sequence for the chimaeric aspartic proteinase construct using pepsinogen and the aspartic proteinase from *A. thaliana*

Top panel: the chimaeric enzyme is composed of three parts. Part A includes the prosegment (44 amino acids) and both catalytic sites (DTG) of the porcine pepsinogen gene, part B is the PSS region of the *A. thaliana* aspartic proteinase and part C is the C-terminal lobe of porcine pepsinogen. The prosegment is removed upon activation to result in an active chimaeric enzyme. Bottom panel: the specific sequences involved in the creation of the chimaeric aspartic proteinase. The PSS region from *A. thaliana* was engineered into porcine pepsinogen at the identical location in which it is found in *A. thaliana* (approx. 100 amino acids from the C-terminus). The 100-amino-acid PSS region is boxed and labelled. Asterisks represent identical residues between the porcine pepsinogen and *A. thaliana* aspartic proteinase and were used to identify the insertion site of the PSS region into pepsinogen. The sole amino acid foreign to the sequence was Leu, in the N-terminal portion of porcine pepsinogen part B (light grey).

RESULTS AND DISCUSSION

Construction, expression and purification of the chimaeric aspartic proteinase

The general scheme for the construction of the chimaeric enzyme is indicated in Figure 1 (upper panel), i.e. the catalytic aspartates from pepsin, the PSS from the *A. thaliana* aspartic proteinase and the C-terminal fragment from pepsin. With the exception of one amino acid change of Ile to Leu, both hydrophobic residues, to create one unique restriction endonuclease site, the entire chimaeric enzyme was designed from the available sequences of both aspartic proteinases (Figure 1, lower panel). Thus minimal loss or change to the naturally occurring amino acids for both polypeptides was achieved.

The PSS possesses a conserved N-glycosylation site modified *in vivo* with an oligomannose-type moiety containing proximal fucose and xylose residues [23]. In order to mimic conditions *in vivo*, the chimaeric enzyme was expressed in *P. pastoris*, as this lower eukaryote is capable of both N- and O-glycosylation using high-mannose residues [24]. Porcine pepsinogen, which does not possess the typical Asn-Xaa-Ser/Thr glycosylation sequence, has been successfully expressed in *P. pastoris* and characterized [16]. The kinetic characterization has shown similar kinetics, pH stability and thermostability [16] to thioredoxin-fused porcine pepsinogen (*E. coli* expression system) and commercially available pepsinogen [20].

The chimaeric aspartic proteinase was screened for expression from *P. pastoris* using milk-clotting ability, recognition by the anti-pepsinogen antibody [20] and successful genomic integration. Purification of the chimaeric enzyme from the supernatant of *P. pastoris* involved Source Q anion-exchange chromatography and concanavalin A-Sepharose affinity chromatography. Approx. 0.8 mg/l purified enzyme was obtained from the *P. pastoris* expression system. Assessment of the purification was performed via SDS/PAGE, Western blotting (using an anti-pepsinogen antibody) and N-terminal sequencing. A broad peak from Source Q anion-exchange chromatography showed a major band at 68 kDa

Figure 2 SDS/PAGE and Western-blot analysis of the purified chimaeric aspartic proteinase

Top panels: SDS/PAGE analysis performed under reduced conditions, stained with SYPRO Ruby Red and photographed using a Syngene gel documentation system. All protein amounts were approx. 1 μ g. Molecular-mass markers are indicated on the figure. Lane 1, commercial pepsinogen; lane 2, zymogenic chimaeric enzyme; lane 3, activated chimaeric enzyme; lane 4, PNGase F-treated zymogenic chimaeric enzyme; lane 5, PNGase F-treated activated chimaeric enzyme. Bottom panels: Western-blot analysis using the anti-pepsinogen antibody. All protein amounts were approx. 0.1 μ g. Lane designations were as described for the top panel. Molecular-mass markers are indicated.

and a minor band at 58 kDa by SDS/PAGE and Western blot analyses (Figure 2, both panels, lanes 2). Based on the amino acid sequence, the predicted molecular mass of the zymogenic form of the chimaeric aspartic proteinase was predicted to be 51 100 Da. The cause of the heterogeneity (production of 68 and 58 kDa forms) of the chimaeric enzyme was determined to be 2-fold. First, the primary sequence was predicted to possess an N-glycosylation site within the PSS region and it was determined that the amount of glycosylation differed between the two forms. Both forms indicated, through the periodic acid-Schiff method, the presence of glycosylation (results not shown). Deglycosylation with PNGase F resulted in an SDS/PAGE and Western-blot profile of the chimaeric enzyme as a single band at the predicted 51 kDa (Figure 2, both panels, lanes 4). Secondly, the lower band (58 kDa) also possessed a truncated N-terminus (a loss of a 26 amino acids from the N-terminal region of the enzyme; Figure 3). The full (68 kDa) and truncated (58 kDa) forms of the chimaeric enzyme (from the Source Q anion-exchange purification step) were successfully separated by Superose 12 gel filtration; however, there was a substantial loss in yield (from approx. 20% to 1% of total protein). Since amino acid sequencing determined that the full and truncated forms of the chimaeric enzyme activated to the proper N-terminus of pepsin, and that the 68 and 58 kDa forms displayed similar properties to one another upon activation, e.g. stability and activity, to maximize yield of the chimaeric enzyme source, the heterogeneous preparation (68 and 58 kDa forms) was activated and used for stability and enzymic studies.

No evidence of contaminating yeast aspartic proteinases from the *P. pastoris* expression system was apparent. Aspartic proteinase activity was not detected in the medium (assayed using milk-clotting assay) until it was activated as described in

A: N-terminal sequence of pepsinogen:

1p 10p 20p 30p 40p 1 10
* * * * * * *
MLVKVPLVRKKSLRQNLIKNGKLDKFLKTHKHNPASKYFPEAAALIGDEPLENYLD

B: M[D]L[K]VKVPLV
C: LK[D]THKLV[E]K
D: I[S]G[A]DE[V]P[S]L[D]E[I]NY[G]LA?
E: I[D]GD[I]E[I]P[G]L[N]E[I]N[V]YL

Figure 3 N-terminal sequence data for the zymogenic and active forms of the chimaeric aspartic proteinase

Sequencing was performed as described in the Experimental section: SDS/PAGE and Western blotting. (A) The N-terminal sequence of pepsinogen (zymogen) (accession number J04601). Underlining indicates the 44-amino-acid prosegment sequence, while bold letters indicate the N-terminal sequence of the active form of the enzyme, pepsin. (B) The N-terminal sequence of the 68 kDa chimaeric aspartic proteinase (zymogenic form). (C) The N-terminal sequence of the truncated 58 kDa chimaeric aspartic proteinase (zymogenic form). The sequence of the truncated 58 kDa form indicates that the sequence starts 26 amino acids from the anticipated N-terminus of the pepsinogen sequence. (D) and (E) indicate the N-terminal sequences derived from the activated forms of the 68 and 58 kDa chimaeric enzymes, respectively. [] Indicates a minor amount of amino acid detected during the cycle.

the Experimental section. Secondly, N-terminal analysis of the major and minor bands isolated from Source Q anion-exchange and Concanavalin A affinity chromatography revealed the full and truncated N-terminus of porcine pepsinogen. The absence of endogenous yeast aspartic proteinases in our chimaeric enzyme preparation is consistent with the findings of other groups that have used the *P. pastoris* system for over-expressing, purifying and characterizing aspartic proteinases [14,16].

All eukaryotic aspartic proteinases are produced in a stable, zymogenic form [7]. The prosegment of the zymogenic form of the chimaeric enzyme was predicted to consist of a 44-amino-acid stretch that resides in the active site of the enzyme cleft and prevents premature functioning [25] until either a unimolecular or bimolecular reaction occurs resulting in an active protease [26]. Activity was not present until the chimaeric enzyme was activated unimolecularly (lowering the pH below 2), resulting in loss of the prosegment (confirmed via N-terminal analysis; Figure 3). The reasons for the production of a minor amount of the chimaeric enzyme with a truncated prosegment were not clear. It is possible that an endogenous proteinase may have cleaved the 26 amino acids of the prosegment, as postulated for the activation of the cyprosin aspartic proteinase produced in *P. pastoris* [14]. However, porcine pepsinogen produced in this system does not result in a prosegment cleavage event [16]. The loss of the 26 amino acids could indicate the first stage of intermolecular activation from the zymogenic form to an intermediate form of the enzyme, as found for the conversion of procathepsin D into pseudocathepsin D *in vitro* [27]. However, production of the mature active enzyme was not achieved after prolonged incubations at room temperature (> 48 h). Alternatively, the *P. pastoris* overexpression system is documented to produce N-terminal heterogeneity [28] and the truncated form of the prosegment may simply be an inherent trait of the expression system.

Structural properties of the chimaeric protein

The inclusion of the PSS resulted in a chimaeric enzyme with increased pH stability above pH 6. After incubation at pH 6.5, the chimaeric enzyme possessed almost twice the relative activity compared with pepsin (Figure 4, upper panel) and after incubation at pH 7, pepsin was without activity, while the chimaeric enzyme retained approx. 20% of its original activity. The chimaeric enzyme was found to be marginally stable at pH 7.5, where approx. 5% residual activity remained. Temperature studies indicated that the chimaeric enzyme demonstrated an increase in stability in comparison with pepsin (Figure 4, bottom panel). The chimaeric enzyme showed almost 40% of residual activity after incubation at 65 °C, whereas porcine pepsin retained only 10%.

DSC was used to assess the thermal properties of the chimaeric enzyme and pepsin, where higher denaturation temperatures (T_m) would indicate a more thermostable structure. The chimaeric enzyme demonstrated an increased T_m (68.3 °C) compared with pepsin (66.5 °C; Figure 5). Analysis also indicated that the temperature range at which the chimaeric enzyme began unfolding and was completely unfolded (64.2–69.5 °C) was narrower than that of pepsin (60.3–69.4 °C). The narrower temperature range for unfolding represented a greater co-operativity during denaturation for the chimaeric enzyme compared with pepsin. Enthalpy calculations determined for the chimaeric enzyme and pepsin were 2.7 ± 0.1 compared with 2.64 ± 0.08 kcal/g (1 kcal = 4.184 kJ) respectively. The DSC data would tend to indicate that the thermostability of pepsin was lower than that of chimaeric enzyme. An initial loss of co-operativity in pepsin structure occurs at a lower temperature (60.3 °C) compared with the chimaera (64.2 °C) and can explain the higher relative activity of the chimaera at 65 °C (Figure 4, bottom panel), where pepsin had lost most of its activity due to a loss of co-operativity.

The described stability changes of pepsin, through the introduction of the PSS, could be attributed to the α -helical design of the PSS domain. DSC studies of α -helical domains indicate that a single α -helical structure is not highly stable in aqueous media.

Figure 4 Stability curves of the chimaeric enzyme and porcine pepsin

Top panel: pH stability of the chimaeric enzyme and porcine pepsin. Activity was measured after incubation at pH 6.5, 7.0 and 7.5 using 0.4 M phosphate buffer for 10 min and then the pH adjusted back to pH 5.2 using 3 M sodium acetate. Three replicates were performed in triplicate using 0.025 mM SS1 at 25 °C, with the residual mean activity compared with the mean activity at pH 5.2 in 3 M sodium acetate (control) with S.D. plotted. Bottom panel: temperature stability of the chimaeric enzyme and porcine pepsin. After incubation at specified temperatures for 10 min, the enzyme was returned to 4 °C. Three replicate activity measurements were performed in triplicate using 0.025 mM SS1, pH 5.2, at 25 °C. The mean residual activity was compared with the mean activity at 25 °C (control), with S.D. plotted.

Figure 5 DSC characterization of the chimaeric enzyme and porcine pepsin

DSC measurements were performed in duplicate using a MicroCal MC-2 differential scanning calorimeter with the T_m values for the chimaeric enzyme and pepsin as indicated by arrows. Cp stands for specific heat capacity.

Figure 6 CD spectra of the chimaeric enzyme and pepsin

The CD spectrum was scanned four times between 190 and 250 nm and obtained using a JASCO 600 spectropolarimeter at 25 °C in 20 mM sodium acetate buffer, pH 5.2. The final spectra obtained were the average of the four scans corrected by subtracting the four scans of the buffer. Analyses were performed in duplicate using 0.2 mg/ml enzyme.

However, as the number of helices increases, lateral interactions can lead to increases in co-operativity [29]. The PSS is characterized by five amphipathic α -helices folded into a single compact entity with three disulphide linkages [4], and CD spectra supported the increase in α -helical content (via the presence of a minimum in the 207–222 nm region) for the chimaeric enzyme (Figure 6). DSC demonstrated that the inclusion of the PSS has made the pepsin enzyme a more thermodynamically compact structure, i.e. greater co-operativity during denaturation. Glycosylation of the PSS region in the chimaeric enzyme did not appear to contribute to the stability of the enzyme as deglycosylating the chimaeric enzyme did not demonstrate altered pH-activity profiles, nor thermostability changes when compared with the glycosylated chimaera (results not shown).

Kinetic properties of the chimaeric protein

The K_m and k_{cat} values of the chimaeric enzyme were approx. 0.117 mM and 8.3 s⁻¹, respectively, using SS1. This K_m value was found to be approx. 2-fold higher than the reported literature values of pepsin and a reduced catalytic-centre activity ('turnover number') in comparison with pepsin (Table 1). The kinetic characterization was performed at the optimal pH for the chimaeric enzyme (pH 1.5) and compared with the kinetic values of several plant aspartic proteinases from the literature at various pH values (Table 1). Monomeric aspartic proteinases are poorly characterized and a search among the monomeric aspartic proteinases resulted in a single plant aspartic proteinase (isolated from the bran of soft wheat) being characterized using a chromophoric synthetic substrate [30]. This proteinase, at pH 3.3, possessed an elevated K_m compared with our enzyme and a lower k_{cat} [30] that could be ascribed to a different peptide substrate (Table 1). However, comparing catalytic efficiencies, our monomeric chimaeric enzyme demonstrated a low catalytic efficiency (k_{cat}/K_m) of 70.5 mM⁻¹ · s⁻¹ compared with pepsin, and this may be supported by the overall lower catalytic efficiency of the monomeric aspartic proteinase from wheat [30]. While it is speculative to conclude that this chimaeric enzyme has acquired the properties of a monomeric plant aspartic proteinase, there is evidence to suggest that chimaeric proteins can exhibit activities found in one or both parent proteins [31]. In the present study, an

Table 1 Kinetic analysis of the chimaeric aspartic proteinase and comparison with recombinant pepsin, commercial pepsin and four plant aspartic proteinases

r-Pepsin denotes recombinant pepsin produced from the *P. pastoris* expression system. c-Pepsin denotes activated, purified commercial pepsinogen (Sigma). NR, not reported; AP, aspartic proteinase.

	K_m (μ M)	k_{cat} (s ⁻¹)	k_{cat}/K_m (mM ⁻¹ · s ⁻¹)
Chimaeric enzyme*	117 ± 37	8.3 ± 2.9	70.5 ± 7.0
De-N-glycosylated*	131 ± 9	1.2 ± 0.5	8.9 ± 3.3
r-Pepsin* [16]	62 ± 3	162 ± 5.8	2600 ± 47
c-Pepsin* [16]	54 ± 13	106 ± 10.4	2010 ± 260
Cardosin A* [41]	108 ± 86	55.6 ± 6.1	515 ± 57
Cardosin B* [41]	110 ± 12	89.4 ± 3.5	808 ± 32
Wheat AP†	375 ± NR	25.5 ± NR	68 ± NR
Barley AP‡	205 ± NR	31 ± NR	149 ± NR

* Kinetics determined using SS1 [Lys-Pro-Ala-Glu-Phe-Phe(NO₂)-Ala-Leu].

† Kinetics determined using synthetic substrate Pro-Thr-Glu-Phe-Phe(NO₂)-Arg-Leu [30].

‡ Kinetics determined using synthetic substrate Lys-Pro-Ile-Glu-Phe-Phe(NO₂)-Arg-Leu [22].

entire domain has been incorporated within pepsin that is naturally present in some of the enzymes within the same family. The plant-like changes described herein, with respect to kinetic parameters, may be due to the incorporation of this domain. However, we recognize that the changes determined may simply be due to the disruption of the catalytic residues. Additionally for the chimaeric enzyme, and unlike pepsin, the insertion of the PSS resulted in an enzyme with inconsistent pepstatin binding, i.e. the pepstatin-inhibition curve was non-linear. A PSS-deficient monomeric aspartic proteinase (from rice) has been produced *in vitro*, and while kinetic and specificity data were not presented [32], it would be of interest to determine if this modified plant enzyme possessed altered specificity and kinetic parameters from its parent enzyme.

Glycosylation of the PSS may interfere with the catalytic efficiency, yet crystallographic data of phytepsin indicated that the glycosylation site within the PSS is far removed from the active site [4]. Speculation is that the role of the glycosylation in the PSS may be to prevent premature proteolysis of the expressed zymogen [8]. In order to define the role of glycosylation in reference to kinetic values, a deglycosylated (PNGase F-treated) form of the chimaeric enzyme was explored, which demonstrated similar K_m values with an approx. 10-fold loss in k_{cat} (Table 1). This illustrated that the chimaeric enzyme was less efficient without glycosylation. Regardless, the function of glycosylation in the PSS is not well understood and requires further characterization.

To elucidate the hydrolytic specificity of the chimaeric enzyme and pepsin, digestion of oxidized insulin B-chain and detection of the peptide products using RP-HPLC was employed. Initial analysis of the chimaeric enzyme cleavage profile indicated a modified profile compared with pepsin. Pepsin is known to result in extensive hydrolysis of this substrate [2,33] and similar peptide cleavage products were determined in our study. However, the chimaeric enzyme demonstrated greater specificity and appeared to exhibit exo-protease activity. Figure 7 compares the cleavage profiles of some well-studied aspartic proteinases, and the chimaeric aspartic proteinase demonstrated a profile similar to phytepsin [22]. Extended time studies of the chimaeric enzyme (beyond 24 h and up to 48 h) indicated that the cleavage profile of the oxidized insulin B-chain did not change. However, it should be noted that peak amplitudes did change over the course of the experiment; amplitude changes have been detected

Cleavage profiles for the chimaeric enzyme and pepsin were determined via RP-HPLC with amino acid analysis. Arrows represent cleavage sites determined within the oxidized insulin B-chain for the given aspartic proteinase. Superscript numbers indicate references for each enzyme.

Role of the PSS in pepsin and plant aspartic proteinases

The exact role of the PSS is unknown, but much speculated upon. Recent work demonstrated that the PSS might be responsible for membrane targeting, as the presence of the PSS resulted in a bifunctional molecule containing a proteolytic portion and a membrane-associating-/destabilizing domain [5,12]. However, the present study suggests that the PSS can contribute to changes in stability, activity and structure, in an enzyme that does not normally possess this domain. Thus we suggest adding a putative structure–function role to the list. The tertiary structure of the PSS appears to be conserved between the fully sequenced plant aspartic proteinases; however, the PSS shows limited identity to other plant aspartic proteinase

It is our belief that the PSS is an operationally relevant domain [39], as the present study has demonstrated that it can be transferred into a related structure and used to create a functional protein that has altered substrate specificity and moderately enhanced stability. Typically, more successful enzyme engineering advances to construct enzymes with desired functionality have been achieved via error-prone PCR and DNA shuffling. Combining both methodologies, this chimaeric enzyme may be a point of introduction to directed evolution studies [40] that could provide insight into eukaryotic aspartic proteinase evolution and/or improved kinetic parameters for optimizing aspartic proteinases in industrial endeavours.

REFERENCES

- 1 Davies, D. R. (1990) The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Chem.* **19**, 189–215
- 2 Chitpinityoi, S. and Crabbe, M. J. C. (1998) Chymosin and aspartic proteinases. *Food Chem.* **61**, 395–418
- 3 Mutlu, A. and Gal, S. (1999) Plant aspartic proteinases: enzymes on the way to a function. *Physiol. Plant.* **105**, 569–576
- 4 Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A. and Zdanov, A. (1999) Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. *EMBO J.* **18**, 3947–3955
- 5 Egas, C., Lavoura, N., Resende, R., Brito, R. M., Pires, E., de Lima, M. C. and Faro, C. (2000) The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *J. Biol. Chem.* **275**, 38190–38196
- 6 Cordeiro, M. C., Xue, Z. T., Pietrzak, M., Pais, M. S. and Brodelius, P. E. (1994) Isolation and characterization of a cDNA from the flowers of *Cynara cardunculus* encoding cyprosin (an aspartic proteinase) and its use to study the organ-specific expression of cyprosin. *Plant Mol. Biol.* **24**, 733–741
- 7 Domingos, A., Cardoso, P. C., Xue, Z. T., Clemente, A., Brodelius, P. E. and Pais, M. S. (2000) Purification, cloning and autotoproteolytic processing of an aspartic proteinase from *Centaurea calcitrapa*. *Eur. J. Biochem.* **267**, 6824–6831
- 8 Glathe, S., Kervinen, J., Nimtz, M., Li, G. H., Tobin, G. J., Copeland, T. D., Ashford, D. A., Wlodawer, A. and Costa, J. (1998) Transport and activation of the vacuolar aspartic proteinase phytepsin in barley (*Hordeum vulgare* L.). *J. Biol. Chem.* **273**, 31230–31236
- 9 Ramalho-Santos, M., Veríssimo, P., Cortes, L., Samyn, B., Van Beeumen, J., Pires, E. and Faro, C. (1998) Identification and proteolytic processing of procarnosin A. *Eur. J. Biochem.* **255**, 133–138
- 10 Zhu, Y. and Conner, G. E. (1994) Intermolecular association of lysosomal protein precursors during biosynthesis. *J. Biol. Chem.* **269**, 3846–3851
- 11 Guruprasad, K., Törmäkangas, K., Kervinen, J. and Blundell, T. L. (1994) Comparative modelling of barley-grain aspartic proteinase: a structural rationale for observed hydrolytic specificity. *FEBS Lett.* **352**, 131–136
- 12 Faro, C., Ramalho-Santos, M., Vieira, M., Mendes, A., Simoes, I., Andrade, R., Veríssimo, P., Lin, X., Tang, J. and Pires, E. (1999) Cloning and characterization of cDNA encoding cardosin A, an RGD-containing plant aspartic proteinase. *J. Biol. Chem.* **274**, 28724–28729
- 13 Frazão, C., Bento, I., Costa, J., Soares, C. M., Veríssimo, P., Faro, C., Pires, E., Cooper, J. and Carrondo, M. A. (1999) Crystal structure of cardosin A, a glycosylated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L. *J. Biol. Chem.* **274**, 27694–27701

- 14 White, P. C., Cordeiro, M. C., Arnold, D., Brodelius, P. E., and Kay, J. (1999) Processing, activity, and inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). *J. Biol. Chem.* **274**, 16685–16693
- 15 D'Hondt, K., Stack, S., Gutteridge, S., Vandekerckhove, J., Krebbers, E. and Gal, S. (1997) Aspartic proteinase genes in the Brassicaceae *Arabidopsis thaliana* and *Brassica napus*. *Plant Mol. Biol.* **33**, 187–192
- 16 Yoshimasu, A. K., Ahn, J.-K., Tanaka, T. and Yada, R. Y. (2002) A new system for the soluble expression of porcine pepsinogen. *Protein Expression Purif.* **25**, 229–236
- 17 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680–685
- 18 Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350–4354
- 19 Payie, K. G., Weadge, J. T., Tanaka, T. and Yada, R. Y. (2000) Purification, N-terminal sequencing and partial characterization of a novel aspartic proteinase from the leaves of *Medicago sativa* L. (alfalfa). *Biotechnol. Lett.* **22**, 1515–1520
- 20 Tanaka, T. and Yada, R. Y. (1996) Expression of soluble cloned porcine pepsinogen A in *Escherichia coli*. *Biochem. J.* **315**, 443–446
- 21 Stoscheck, C. M. (1990) Quantitation of protein. *Methods Enzymol.* **182**, 50–68
- 22 Kervinen, J., Sarkkinen, P., Kalkkinen, N., Mikola, N. and Saarna, M. (1993) Hydrolytic specificity of the barley grain aspartic proteinase. *Phytochemistry* **32**, 799–803
- 23 Costa, J., Ashford, D. A., Nimtz, M., Bento, I., Frazao, C., Esteves, C. L., Faro, C. J., Kervinen, J., Pires, E., Verissimo, P., Wlodawer, A. and Carrondo, M. A. (1997) The glycosylation of the aspartic proteinases from barley (*Hordeum vulgare* L.) and cardoon (*Cynara cardunculus* L.). *Eur. J. Biochem.* **243**, 695–700
- 24 Cereghino, J. L. and Clegg, J. M. (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* **24**, 45–66
- 25 Bernstein, N. K. and James, M. N. G. (1999) Novel ways to prevent proteolysis – prophytepsin and proplasmepsin II. *Curr. Opin. Struct. Biol.* **9**, 684–689
- 26 Richter, C., Tanaka, T. and Yada, R. Y. (1998) Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. *Biochem. J.* **335**, 481–490
- 27 Richo, G. R. and Conner, G. E. (1994) Structural requirements of procathepsin D activation and maturation. *J. Biol. Chem.* **269**, 14806–14812
- 28 Letourneur, O., Gervasi, G., Gaia, S., Pagès, J., Watelet, B. and Jolivet, M. (2001) Characterization of *Toxoplasma gondii* surface antigen I (SAGI) secreted from *Pichia pastoris*: evidence for hyper O-glycosylation. *Biotechnol. Appl. Biochem.* **33**, 35–45
- 29 Privalov, P. L. (1982) Stability of proteins: proteins which do not present a single cooperative system. *Adv. Protein Chem.* **35**, 1–104
- 30 Galleschi, L. and Felicioli, F. (1994) Purification, characterization and activation by anions of an aspartic proteinase isolated from bran of soft wheat. *Plant Sci.* **98**, 15–24
- 31 Campbell, R. K., Bergert, E. R., Wang, Y., Morris, J. C. and Moyle, W. R. (1997) Chimeric proteins can exceed the sum of their parts: implications for evolution and protein design. *Nat. Biotechnol.* **15**, 439–443
- 32 Asakura, T., Matsumoto, I., Funaki, J., Arai, S. and Abe, K. (2000) The plant aspartic proteinase-specific polypeptide insert is not directly related to the activity of oryzasin 1. *Eur. J. Biochem.* **267**, 5115–5122
- 33 Faro, C. J., Moir, A. J. G. and Pires, E. V. (1992) Specificity of a milk-clotting enzyme extracted from the thistle *Cynara cardunculus* L.: action on oxidised insulin and κ -casein. *Biotechnol. Lett.* **14**, 841–846
- 34 Bleux, W. and Delcour, J. A. (2000) A second aspartic proteinase associated with wheat gluten. *J. Cereal Sci.* **32**, 31–42
- 35 Favilla, R., Parisoli, A. and Mazzini, A. (1997) Alkaline denaturation and partial refolding of pepsin investigated with DAPI as an extrinsic probe. *Biophys. Chem.* **67**, 75–83
- 36 Lin, X., Loy, J. A., Sussman, F. and Tang, J. (1993) Conformational instability of the N- and C-terminal lobes of porcine pepsin in neutral and alkaline solutions. *Protein Sci.* **2**, 1383–1390
- 37 Konno, T., Kamatari, Y. O., Tanaka, N., Kamikubo, H., Dobson, C. M. and Nagayama, K. (2000) A partially unfolded structure of the alkaline-denatured state of pepsin and its implication for stability of the zymogen-derived protein. *Biochemistry* **39**, 4182–4190
- 38 Tanaka, T. and Yada, R. Y. (2001) N-terminal portion acts as an initiator of the inactivation of pepsin at neutral pH. *Protein Eng.* **14**, 669–674
- 39 Voigt, C. A., Martinez, C., Wang, Z. G., Mayo, S. L. and Arnold, F. H. (2002) Protein building blocks preserved by recombination. *Nat. Struct. Biol.* **9**, 553–558
- 40 Petrounia, I. P. and Arnold, F. H. (2000) Designed evolution of enzymatic properties. *Curr. Opin. Biotechnol.* **11**, 325–330
- 41 Verissimo, P., Faro, C., Moir, A. J., Lin, Y., Tang, J. and Pires, E. (1996) Purification, characterization and partial amino acid sequencing of two new aspartic proteinases from fresh flowers of *Cynara cardunculus* L. *Eur. J. Biochem.* **235**, 762–768
- 42 Verissimo, P., Ramahlo-Santos, M., Faro, C. and Pires, E. (1998) A comparative study on the aspartic proteinases from different species of *Cynara*. *Adv. Exp. Med. Biol.* **436**, 459–463

Received 17 July 2002/26 February 2003; accepted 12 March 2003

Published as BJ Immediate Publication 12 March 2003, DOI 10.1042/BJ20021126