Eur. J. Biochem. **270**, 3026–3035 (2003) © FEBS 2003 doi:10.1046/j.1432-1033.2003.03681.x

Procarboxypeptidase A from the insect pest *Helicoverpa armigera* and its derived enzyme

Two forms with new functional properties

Alex Bayés[1], Anka Sonnenschein[2], Xavier Daura[3], Josep Vendrell[1] and Francesc X. Aviles[1]

[1]*Departament de Bioquímica i Biologia Molecular, Facultat de Ciències and Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Spain;* [2]*Klinik für Neurologie, Universitätsklinikum der Technischen Universität, Dresden, Germany;* [3]*Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Spain*

Although there is a significant knowledge about mammalian metallocarboxypeptidases, the data available on this family of enzymes is very poor for invertebrate forms. Here we present the biochemical characterization of a metallocarboxypeptidase from the insect *Helicoverpa armigera* (Lepidoptera: Noctuidae), a devastating pest spread in subtropical regions of Europe, Asia, Africa and Oceania. The zymogen of this carboxypeptidase (PCPAHa) has been expressed at high levels in a *Pichia pastoris* system and shown to display the characteristics of the enzyme purified from the insect midgut. The *in vitro* activation process of the proenzyme differs significantly from the mammalian ones. The lysine-specific endoprotease LysC activates PCPAHa four times more efficiently than trypsin, the general activating enzyme for all previously studied metalloprocarboxypeptidases. LysC and trypsin independently use two different activation targets and the presence of sugars in the vicinity of the LysC activation point affects the activation process, indicating a possible modulation of the activation mechanism. During the activation with LysC the prodomain is degraded, while the carboxypeptidase moiety remains intact except for a C-terminal octapeptide that is rapidly released. Interestingly, the sequence at the cleavage point for the release of the octapeptide is also found at the boundary between the activation peptide and the enzyme moieties. The active enzyme (CPAHa) is shown to have a very broad substrate specificity, as it appears to be the only known metallocarboxypeptidase capable of efficiently hydrolysing basic and aliphatic residues and, to a much lower extent, acidic residues. Two carboxypeptidase inhibitors, from potato and leech, were tested against CPAHa. The former, of vegetal origin, is the most efficient metallocarboxypeptidase inhibitor described so far, with a K_i in the pM range.

Keywords: metallocarboxypeptidase; zymogen; proteolytic activation; substrate specificity; protein inhibitor.

Correspondence to J. Vendrell, Departament de Bioquímica i Biologia Molecular, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain.
Fax: + 34 93 5811264, Tel.: + 34 93 5812375,
E-mail: josep.vendrell@uab.es
or F. X. Aviles, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, E-08193, Bellaterra, Spain.
Fax: + 34 93 5812011, Tel.: + 34 93 5811315,
E-mail: FrancescXavier.Aviles@uab.es

Abbreviations: PCPAHa, procarboxypeptidase from *Helicoverpa armigera*; PCPAHaα, procarboxypeptidase α from *Helicoverpa armigera*; CPAHa, carboxypeptidase from *Helicoverpa armigera*; CPA1h, human carboxypeptidase A1; CPA2h, human carboxypeptidase A2; CPBh, human carboxypeptidase B; CPAb, bovine carboxypeptidase A; LysC, lysyl endopeptidase; PCI, potato carboxypeptidase inhibitor; LCI, leech carboxypeptidase inhibitor; AOX1, alcohol oxidase gene; AAFP, *N*-(4-methoxyphenylazoformyl)-L-phenylalanine; Cbz, carbobenzoxy; *N*-(3-(2-furyl), acryloyl)-L-phenylalanyl-L-phenylalanine.

Enzymes: CPAHa SWP: O97434 (E.C. 3.4.17.1); CPA1 h SWP: P15085 (E.C. 3.4.17.1); CPA2 h SWP: P48052 (E.C. 3.4.17.15); CPBh SWP: P15086 (E.C. 3.4.17.2); CPAb SWP: P00730 (E.C. 3.4.17.1).

(Received 20 March 2003, revised 8 May 2003, accepted 21 May 2003)

The understanding of the digestive process in pest insects is a key step in the design of many insecticides, including insect-resistant transgenic plants [1]. Exopeptidases are supposed to play a major role in protein digestion, as peptides and proteins have to be converted into dipeptides or single amino acids in order to be taken up efficiently by the gut. These proteases are well described in mammals, but little is known about the exoproteases of insect origin. *Helicoverpa armigera* (Lepidoptera: Noctuidae), also known as cotton worm or boll worm, has a widespread distribution in tropical, subtropical and warm temperature regions in Europe, Asia, Africa and Oceania. It is an important pest of many crop plants, including cotton, corn, maize, tomato, bean, sorghum, tobacco and certain flower plants such as chrysanthemum or carnation. The losses due to *Helicoverpa zea* and *Helicoverpa virescens*, two butterflies that belong to the genre of *Helicoverpa armigera*, were calculated to be one thousand million dollars per year in the USA [2]. A midgut carboxypeptidase from this lepidopter, first described by Bown *et al.* [3], is the subject of the present work.

From a mechanistic point of view, two major types of carboxypeptidases can be distinguished: serinecarboxypeptidases and metallocarboxypeptidases. In mammals the metallocarboxypeptidase family is divided into subfamilies

A/B and N/E [4] which include, respectively, the formerly called pancreatic-like and regulatory forms, the latter referring to a number of enzymes involved in the processing of bioactive peptides and hormones [4,5]. The carboxypeptidase of *H. armigera* belongs to the A/B subfamily and contains a Zn^{2+} atom directly involved in catalysis. From its localization in the gut of the larvae, it is thought to participate in the digestive process of the insect.

Two forms of pancreatic-like carboxypeptidases, CPA and CPB, are involved in the degradation of dietary proteins. The two isoforms of CPA, A1 and A2, differ in specificity with the former having a preference for aliphatic and aromatic C-terminal residues and the latter being more restrictive for aromatic residues, particularly tryptophan [5–7]. The B form is highly specific for basic residues. Pancreatic-like carboxypeptidases are synthesized as proenzymes. Upon tryptic activation, a 92–95 residue N-terminal activation segment, that shields the entrance of substrates to the active site, is released. This proregion, besides acting as a potent inhibitor of the enzyme (in the nM range) [5,8], also behaves as an intramolecular chaperone for the folding of the enzyme.

We have recently described [9] the three-dimensional crystal structure of an A-type metalloprocarboxypeptidase from *H. armigera* (PCPAHa), showing that its overall fold and conformation is very much similar to other known zinc procarboxypeptidases, indicating the conservation of these features through evolution.

In the present study we report the production of this zymogen at high yield in the methylotrophic yeast *Pichia pastoris*, a fact that allowed the study of the biochemical properties of both the proenzyme and enzyme forms. Through the description of the proenzyme activation process, the substrate specificity of the active enzyme and the behaviour upon inhibition by two well known natural inhibitors, a number of specific, distinctive features can be deduced for this new member of the family of pancreatic-like procarboxypeptidases.

Experimental procedures

Materials

Restriction endonucleases *Avr*II, *Sac*I and *Xho*I, T4 DNA ligase, *Taq* polymerase, deoxynucleotide stocks and N-glycosidase F were purchased from Roche. Salts and media for *E. coli* and *P. pastoris* growth were obtained from Sigma and Hispanlab (Alcobendas, Spain), respectively. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanyl chloromethyl ketone) was from Worthington (Lakewood, USA) and Lysyl endopeptidase (from *Achromobacter lyticus*) from Waco. Chymotrypsin was from Merck. The peptides V14R, V15E and V14W where synthesized by Diverdrugs (Barcelona, Spain). Poly(vinylidene difluoride) (PVDF) membrane was from Waters. Elastase, trifluoroacetic acid, cyanogen bromide, synapinic acid, *N*-(3-(2-furyl) acryloyl)-L-phenylalanyl-L-phenylalanine (FAPP) and Cbz-Gly-Gly-Ser were from Sigma. *N*-(4-Methoxyphenylazoformyl)-L-phenylalanine (AAFP) and the rest of substrates used for the kinetic measurements were from Bachem (Bubendorf, Switzerland).

Plasmids constructs

DNA manipulations were carried out essentially as described by Sambrook *et al.* [10] using *E. coli* strain MC1061 as host. Primers were synthesized to amplify the cDNA containing the procarboxypeptidase by PCR. Sense primer 5′-GATTCTCTCGAGAAAAGAAAACATGAAATTT ATGATGG-3′; antisense primer 5′-CTTCTTTGAGT TATGACGAATTGGATCCTAC-3′. The original signal peptide from this molecule is not included in the construct. The underlined sequences indicate the restriction sites for *Xho*I and *Avr*II introduced to be able to subclone the cDNA into the *P. pastoris* expression vector pPIC9. The cDNA was introduced between the 5′ promoter and 3′ terminator of the alcohol oxidase gene (AOX1), resulting in a new vector called pPIC9-PCPAHa. This vector provides the α-mating factor signal for secretion of the recombinant protein.

Transformation and selection of the productive clones

Prior to the transformation the vector was linearized with *Sac*I. The KM 71 strain of *P. pastoris*, which produces only the slow growing phenotype, was transformed using the spheroplasts method with the linearized vector. The cells where then plated on minimal dextrose medium (MD) agar (1.34% yeast nitrogen base, 0.00004% biotin, and 1% dextrose) a medium devoid of histidine where only the transformed cells can grow. To find a highly producing clone, over 60 colonies were grown in 10 mL buffered glycerol-complex (BMGY) medium (1% yeast extract, 2% peptone, 90 mM potassium phosphate, pH 6.0, 0.00004% biotin and 1% glycerol) at 30 °C for 3 days. Cells were collected by centrifugation and resuspended in 2 mL buffered methanol-complex medium (BMMY) medium (same as BMGY but containing 1% methanol instead of 1% glycerol) and grown for 3 days more to induce the production of recombinant protein. The supernatant of all the clones was analysed by SDS/PAGE, followed by densitometry to identify the most productive ones. The functionality of the recombinant protein was tested with the specific substrate FAPP(*N*-(3-(2-furyl) acryloyl)-L-phenylalanyl-L-phenylalanine) [11] after activation of the proenzyme with trypsin.

Expression and purification of the recombinant enzyme

Expression and purification procedures were carried out essentially as described in [9]. In short, 1 L of BMGY medium was grown at 30 °C and at 300 r.p.m. constant shaking for 2 days until D_{600} reached 20 units. The cells were then collected by centrifugation at 1500 ***g*** and gently resuspended in 200 mL of BMMY medium. In a first step, the protein secreted to the supernatant was purified by hydrophobic interaction chromatography in a butyl-Toyopearl 650M column. The sample was loaded onto the column after equilibration of its ionic strength to 30% saturation with ammonium sulphate, and the protein was eluted with a decreasing gradient of the same salt. After overnight dialysis of the selected fractions, the protein was finally purified on an FPLC system using a preparative anion-exchange column (TSK-DEAE 5PW; TOSOH,

Tokyo, Japan) and applying a 65 min gradient from 100% buffer A (20 mM Tris, pH 7.0) to 15% buffer B (buffer A plus 0.8 M ammonium acetate).

Activity assays

Two different synthetic substrates were used to analyse carboxypeptidase A activity. *N*-(3-(2-furyl) acryloyl)-L-phenylalanyl-L-phenylalanine (FAPP) was used routinely to measure carboxypeptidase activity and AAFP was used to calculate the inhibition constants [12]. FAPP was prepared at a 0.2 mM concentration in 50 mM Tris, 0.45 M NaCl, pH 7.5, and the A_{330} decrease at 25 °C. A stock solution of 50 mM *N*-(4-methoxyphenylazoformyl)-L-phenylalanine (AAFP) in dimethylsulfoxide was diluted immediately before use to 10 mM with 50 mM Tris, 0.1 M NaCl, pH 8.0. From this solution, 10 μL were added to 1 mL of 50 mM Tris, 0.1 M NaCl, pH 8.0. CPA activity was measured by following the A_{350} decrease at 25 °C.

Deglycosylation assay

Samples were deglycosylated with N-glycosidase F, an enzyme that removes N-linked sugars by cleaving the bond between the asparagines from the polypeptide chain and the first *N*-acetylglucosamine. Glycosylated molecules were concentrated at 1 mg·mL^{-1} in Tris 5 mM pH 8.0 and appropriate volumes of N-glycosidase F at 1 unit μL^{-1} were added to achieve a final ratio of 100 : 1 v/v. The reaction was left to proceed overnight at 37 °C.

Kinetic measurements

The rate of hydrolysis of the different substrates were measured spectrophotometrically in 50 mM Tris, 0.5 M NaCl, 1 μM $ZnCl_2$, pH 8.0, at 25 °C. The wavelengths used to monitor the various reactions were as follows: 226 nm for Cbz-Gly-Gly-Ser, Cbz-Gly-Gly-Ala, Cbz-Gly-Gly-Leu, Cbz-Gly-Gly-Val, Cbz-Gly-Gly-Phe and Cbz-Gly-Phe; 236 nm for Cbz-Gly-Gly-Tyr and Cbz-Gly-Tyr; and 302 nm for Cbz-Gly-Gly-Trp and Cbz-Gly-Trp. Initial rates, determined from the first 5–10% of the time-trace of each reaction, were obtained at substrate concentrations close to the K_m value whenever possible. The kinetic parameters, k_{cat} and K_m, were obtained using 6–8 experimental points by direct fit to a Michaelis–Menten curve using the ENZFITTER program [13].

Activation studies of recombinant *H. armigera* PCPAHa

Recombinant enzyme at 1 mg·mL^{-1} in 5 mM Tris, 1 μM $ZnCl_2$, pH 8.0, was treated with lysyl endopeptidase (LysC) at a PCPAHa : LysC ratio of 40 : 1 (w/w) and at 37 °C. To avoid the action of active carboxypeptidase upon the fragments generated, the potato carboxypeptidase inhibitor (PCI) was also added to the mixture at a 1 : 4 PCPHa/PCI molar ratio when enzymatic activity was not going to be measured. During the activation process, aliquots were taken for reverse-phase HPLC analysis and activity measurements. Seventy microlitres of the reaction mixture, with trifluoroacetic acid added to a concentration of 0.05% (v/v) to stop the activation reaction, were analysed in a Vydac C_4 column (250 × 4.6 mm, 5 μm particle size and 0.3 μm pore size). The chromatographies were performed in the presence of 0.1% trifluoroacetic acid with an elution gradient between water (solvent A) and 90% acetonitrile (solvent B) according to the following steps: 10% solvent B from 0 to 10 min, 10–60% solvent B from 10 to 130 min. Elution was followed by measuring the A_{214} and the isolated fractions were concentrated in an Speed-Vac (Savant) and further analysed by MALDI-TOF spectrometry, SDS/PAGE and N-terminal sequencing. Parallel 10 μL aliquots of the activation mixture were added to 190 μL of aprotinin (bovine pancreas trypsin inhibitor) at 0.1 mg·mL^{-1} in 20 mM Tris, 0.1 M NaCl, 1 μM $ZnCl_2$, pH 8.0, and 10 μL of the resulting mixture were used to measure enzyme activities using FAPP as a substrate.

To analyse the effect of sugars on the activation of PCPAHa by LysC and bovine trypsin, glycosylated and nonglycosylated PCPAHa were activated with increasing PCPAHa/activating enzyme ratios at 37 °C for 2 h. One microlitre of the reaction mixture was assayed against FAPP. Triplicate measures were obtained for each data point.

Cyanogen bromide cleavage of PCPAHa

One hundred micrograms of PCPAHa and PCPAHa-α were lyophilized separately in eppendorf tubes and resuspended with 50 μL of 70% formic acid, containing CNBr at 100 mg·mL^{-1} and tryptophan at 0.1 mg·mL^{-1}. The tube was protected from light and the reaction left to proceed for 10 h at room temperature. The sample was subsequently diluted 10 times with Milli-Q water (Millipore, France), frozen and lyophilized. The resuspension-freezening-lyophilization cycle was repeated once. The sample was finally dissolved in 5 μL of Milli-Q water and analysed by MALDI-TOF.

Activity measurements using peptide substrates

The hydrolytic activity of CPAHa against the three different peptide substrates V14R (VKKKARKAAGGAKR), V14W [VKKKARKAAGC(Acm)AW] and V15E [VKKKARKAAGC(Acm)AWE] was analysed by HPLC in a Vydac C_{18} column (250 × 4.6 mm, 5 μm particle size and 0.3 μm pore size). Human carboxypeptidases A1 (CPA1 h), A2 (CPA2 h), B (CPABh), a CPBh mutant (CPBh S251T, D253K) [14] which hydrolyses acid C-terminal residues and the *H. armigera* carboxypeptidase (CPAHa) were used in this assay at an enzyme/substrate ratio of 1 : 1 (w/w) at 37 °C. At desired times, the reactions were stopped by the addition of trifluoroacetic acid to a final concentration of 0.05%. The reaction products were analysed by HPLC using the same column and solvents described for the activation studies, but applying a linear gradient from 10–30% solvent B in 60 min.

Mass spectrometry and N-terminal sequence analysis

A MALDI-TOF spectrometer (Bruker; Bremen, Germany) was used to analyse peptides and proteins. The matrix used was synapinic acid and samples were mixed 1 : 1 (v/v). All N-terminal sequences were obtained in a Beckman CF3000

sequencer. Samples were analysed in solution or blotted onto PVDF membranes and detected by Coomassie staining.

Measurement of equilibrium dissociation constant (K_i)

To calculate the K_i values, the method for reversible tight-binding inhibitors described by Bieth [15] was used. Carboxypeptidase concentration was left constant at 0.8 nM and increasing amounts of inhibitor were added. At each point, the activity (v_i) was measured against the substrate AAFP. The activity of CPAHa in the absence of inhibitor is defined as v_o and the parameter a is defined as v_i/v_o. By plotting [I]/1 − a against 1/a, a line is obtained that follows the equation: $[I] = [E](1 - a) + K_{iapp}\ (1 - a)/a$. To correct for the effect of the substrate on the formation of the complex EI, the following equation is applied: $K_i = K_{iapp}/(1 + [S]/K_m)$, resulting in the final K_i value.

Computational methods

The simulations were carried out using the GROMOS96 package of programs [16,17] and GROMOS96 45A3 force field [16,18]. The ionisable groups were set to their protonated or deprotonated state according to standard pK_a values of amino acids and a pH of 7. The SPC water model [19] was used as solvent.

The CPAHa-PCI complex was modelled using the coordinates of the CPAb-PCI complex (Protein Data Bank entry 4CPA) as a template. The coordinates for the apo form of CPAHa were obtained simply by removing the prosegment in the Protein Data Bank entry 1JQG. CPAHa was then superimposed onto the CPAb-PCI complex by least-squares fitting of the two enzyme structures using the C^{α} atoms in conserved helices (residues 14–28, 74–88, 98–102, 112–121, 173–186, 215–231, 253–262, 285–306), the catalytic triad, and the Zn^{2+} atom.

A 500 ps molecular dynamics (MD) simulation at 298 K and 1 atm under truncated–octahedron periodic boundary conditions was carried out for each system (CPAb-PCI: 39372 atoms; CPAHa-PCI: 38703 atoms). Trajectory coordinates and energies were stored at 0.5 ps intervals from the time frame 100–500 ps and used for analysis. Least-squares translational and rotational fitting of trajectory structures from the two complexes was based on the C^{α} atoms found in conserved helical regions (residues 14–28, 74–88, 98–102, 112–121, 173–186, 215–231, 253–262, 285–306), the catalytic triad, and the Zn^{2+} atom. The atom-positional rmsd was calculated for the backbone atoms (N-C^{α}-C) of PCI.

Results

Overexpression, purification and initial characterization of recombinant Pro-CPA from *H. armigera*

Analysis of more than 60 transformant colonies led to the identification of a clone able to produce up to 40 mg of pure protein from 1 L of initial culture. The product was highly homogenous as assessed by SDS/PAGE and had the expected molecular mass of 46.6 kDa (Fig. 1). In the first purification step, the use of a hydrophobic interaction

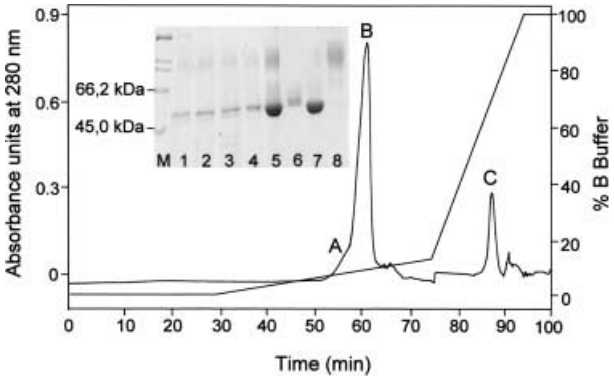


Fig. 1. Purification of PCPAHa. Electrophoretic analysis and anionic exchange chromatography showing, respectively, the evolution of the recombinant expression and the purification to homogeneity of PCPAHa. Lanes 1–4 in the electrophoresis correspond to the analysis of the protein expression culture supernatant at 16, 24, 36 and 46 h. Lane 5 corresponds to the eluate of the hydrophobic interaction chromatography and lanes 6, 7 and 8 correspond, respectively, to peaks A, B and C from the anionic exchange chromatography shown.

chromatography partially eliminates components from the culture supernatant and an additional anionic exchange chromatography is sufficient to obtain a highly purified enzyme that elutes at 6% of B buffer (0.8 M ammonium acetate). The N-terminal sequence determined for the sample in peak B corresponded to the first 10 N-terminal residues of the proenzyme, indicating that the α-mating factor had been completely removed by KEX2, the endoprotease from *P. pastoris* responsible for this action.

In initial activation tests, the purified recombinant PCPAHa was activated with trypsin, the general activator of mammalian pancreatic procarboxypeptidases. Peptidase assays with the synthetic substrate FAPP showed that the enzyme was completely activated at a 4 : 1 (w/w) PCPAHa/trypsin ratio at 25 °C, and its specific activity was calculated to be 150 μmol of substrate per minute and per mg of protein. PCI, leech carboxypeptidase inhibitor (LCI), benzylsuccinic acid and *o*-phenantroline completely inhibit the active enzyme at concentrations of 5 μM, 8 μM, 2 mM and 5 mM, respectively (results not shown), although no inhibitory effect of EDTA could be detected. This is in agreement with previous data obtained with the *H. armigera* gut extracts for the first and the last of the tested inhibitors [3]. Thus, the N-terminal sequence of the recombinant enzyme, its ability to be activated by trypsin and its response to different inhibitors together suggest that the protein is properly folded and very similar to the native form.

Elucidation of the activating enzyme

Four serine proteases (elastase, chymotrypsin, LysC and trypsin) were tested in the search for the type of proteolytic activity that might be responsible for the physiological activation of PCPAHa. A PCPAHa/activating enzyme ratio of 8 : 1 (w/w) was used in all four cases and activation was left to proceed at 23 °C for 60 min. As shown in Fig. 2A, elastase and chymotrypsin were not able to activate the enzyme, while LysC behaved as the best activator, as it only needed half the time used by trypsin to generate a maximum

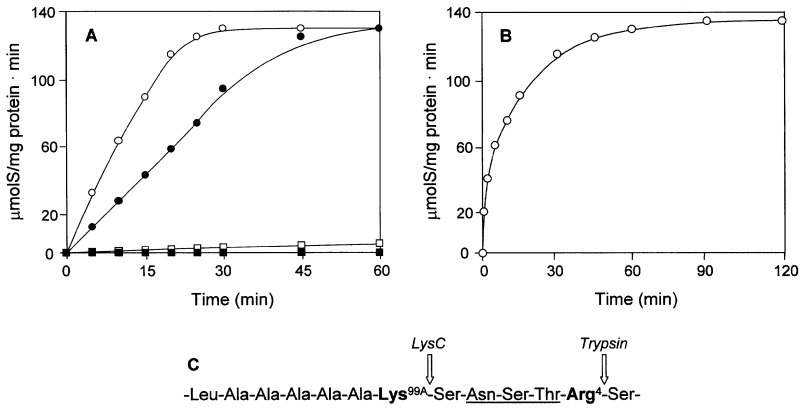


Fig. 2. Activation of PCPAHa by different serine proteases. (A) Activation was carried out at 23 °C at a PCPAHa:activating enzyme ratio of 8 : 1 (w/w) for 60 min. The amount of mature CPAHa produced at different times was detected with the substrate FAPP. The activating enzymes are: (○) LysC (●) bovine trypsin (□) porcine elastase and (■) bovine chymotrypsin. (B) Generation of CPA activity from PCPAHa after activation with LysC at a PCPAHa : LysC ratio of 40 : 1 (w/w) and at 37 °C. (C) Amino acid sequence of PCPA-Ha at the limit between the activation segment and the mature enzyme, where cleavage is produced. The activation points for LysC (Lys99A) and trypsin (Arg4) are shown and the consensus site for N-glycosylation is underlined.

activity. The search for the mildest activating conditions for LysC resulted in a PCPAHa/LysC ratio of 40 : 1 (w/w) at 37 °C (see below), while the mildest activation conditions required for trypsin to reach the maximum CPA activity required a fourfold higher ratio (10 : 1, w/w), also at 37 °C. The activation of PCPAHa by LysC in those conditions is shown in Fig. 2B. LysC and trypsin produced different N-terminal sequences for the mature protein, as determined by N-terminal sequencing. LysC activates the zymogen by cleaving at position 99A after the motif $(A)_5$-K, while trypsin cleaves after R4, five residues downstream (Fig. 2C).

Activation studies; effect of glycosylation and determination of species produced during activation

Some mammalian pancreatic procarboxypeptidases are not able to release a full carboxypeptidase activity upon tryptic activation even after complete cleavage of the limited proteolysis target bond and full release of the mature enzyme. This is due to the inhibitory capacity kept by the activation segment fragment before it suffers extensive and sufficient degradation. In these instances a biphasic curve is obtained when representing the time-course of activity generation [8]. In other cases the propiece is unable to interact with the enzyme moiety in *trans* and a hyperbolic curve is observed [20]. PCPAHa belongs to this second class of zymogens as seen from the shape of the activation course presented in Fig. 2B, in which the generation of activity closely reflects the appearance of the mature enzyme as followed by SDS/PAGE. Furthermore, no trace of activation domain of PCPAHa could be observed during the course of LysC activation by SDS/PAGE analysis, and a parallel HPLC follow-up confirmed that it is extensively fragmented by cleavage at its seven internal Lys residues (results not shown). In contrast to this, the enzyme moiety is resistant to further proteolysis beyond the activating event.

The presence of a unique consensus glycosylation site in the PCPAHa sequence at the border of the activation targets for both LysC and trypsin (Fig. 2C) suggested that glycosylation might affect the activation rate of the zymogen. In order to study this, PCPAHa was treated with N-glycosidase F, and both deglycosylated and glycosylated PCPAHa were activated with decreasing amounts of LysC. Figure 3 shows that PCPAHa is indeed glycosylated and that this modification affects activation, because treated and nontreated samples reach different levels of CPA activity depending on the quantity of activating protease used. Deglycosylated PCPAHa is fully activated at a ratio of 200 : 1 (w/w) whilst the glycosylated enzyme needs five times more LysC to reach the maximum activity, evidence that the presence of the sugar chain makes the access of LysC more difficult for activation. The shift in electrophoretic mobility produced by the deglycosylation is also clearly observed in Fig. 4B. A similar experiment performed using trypsin as the

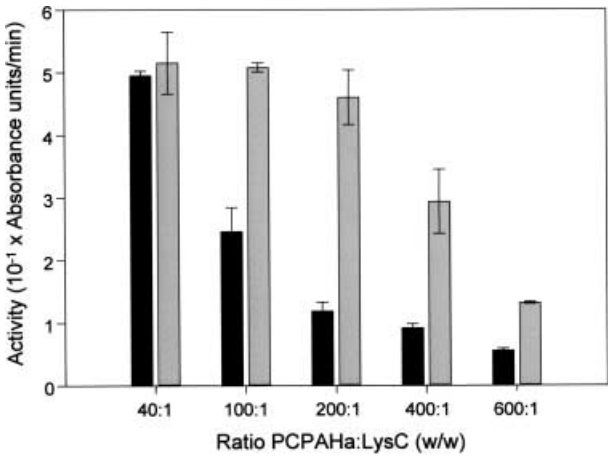
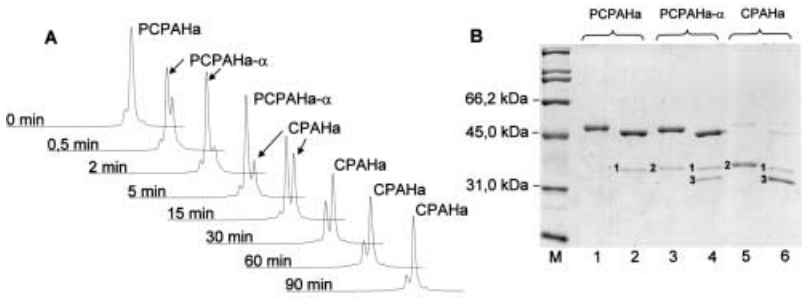


Fig. 3. Effect of glycosylation on the activation of PCPAHa with LysC. Nonglycosylated and glycosylated PCPAHa were activated with decreasing ratios of LysC for 2 h at 37 °C. Subsequently, 1 µL of the reaction mixture was assayed with the FAPP substrate to detect the CPAHa activity generated, which is expressed as absorbance units per min. Dark columns correspond to glycosylated PCPAHa, light columns correspond to nonglycosylated. The data shown are the mean of three measurements ± SD.



C

	Total Mass (Da)	N-t Sequence	Mass (Da) of Q356-A417 fragment.
PCPAHa	48.488 ± 12	KHEIYDGHAVY	6.795 ± 22
PCPAHa-α	47.557 ± 23	KHEIYDGHAVY	5.922 ± 10
CPAHa	36.744 ± 34	S-STRSRLSFD	5.924 ± 14

Fig. 4. Analysis of the species generated during the activation process. The activation of PCPHa with LysC was performed in the conditions of Fig. 2B with the addition of the carboxypeptidase inhibitor from potato (PCI) at a 1–4 molar ratio. (A) At given times, samples from the reaction mixture were made 0.05% in trifluoroacetic acid to stop the reaction and subsequently analysed by HPLC on a Vydac C_4 column. (B) SDS/PAGE electrophoresis of the 3 species isolated from the chromatograms shown in part A of the figure; lanes 1 and 2, PCPAHa; lanes 2 and 4, PCPAHa-α; lanes 5 and 6, CPAHa. Samples from lanes 2, 4 and 6 were treated with N-glycosidase F as described in the experimental procedures. Some bands are numbered: 1, N-glycosidase F; 2, glycosylated CPAHa; 3, deglycosilated CPAHa. (C) Table containing the results of the N-terminal and mass spectrometry analysis of all molecules and also the mass of the C-terminal fragment of the enzyme produced with cyanogen bromide fragmentation, as determined by MALDI-TOF spectrometry.

activating enzyme showed that it was not affected by the presence of the sugar chain (data not shown).

From the activation curve depicted in Fig. 2B it is clear that active CPAHa is produced from the very beginning of the process and that approximately 90 min are needed to attain full activity and thus to generate a maximum of mature enzyme. Analysis of the activation process over time by HPLC shows that, besides the PCPAHa precursor and the final CPAHa product, a third protein species is also detected. This form, marked as PCPAHaα in Fig. 4A, corresponds to a truncated proenzyme which has no CPA activity. The generation of PCPAHaα also starts immediately after activation, but it reaches a maximum in only 5 min, thereafter gently decreasing until complete disappearance in a process that generates the fully mature CPAHa. N-terminal sequencing and MALDI-TOF analysis (Fig. 4C) showed that PCPAHaα shares the N-terminal of the original proenzyme but has a molecular mass about 900 Da smaller. An interesting feature of the insect proenzyme studied here is, as commented above, the presence of an $(A)_nK$ sequence at the LysC activating point, which is also repeated at the end of the protein. A cleavage after this second motif would result in a decrease of 897 Da of mass and be responsible for the generation of PCPAHaα. To assess this possibility, PCPAHA and PCPAHaα were fragmented with cyanogen bromide and the peptides produced analysed by MALDI-TOF spectrometry. This fragmentation generates 11 peptides, Q356-A417 being the one containing the C-terminal peptide in the uncleaved proenzyme. The masses observed for the corresponding fragments in PCPAHa and PCPAHa-α were, respectively, 6.795 ± 22 and 5.922 ± 10, displaying a difference of 873 Da, close enough to 897 Da to demonstrate that the variation is due to the removal of the C-terminal octapeptide. The mass of the corresponding fragment observed for the active, mature enzyme was 5.924 ± 14, confirming that the final product of the activation is also lacking the C-terminal peptide.

In Fig. 4B the analysis of the species isolated from the chromatograms in part A of the figure confirms that the protein expressed in the *Pichia pastoris* system is glycosylated, and that the glycosylation takes place downstream of the cleavage point for LysC, since the electrophoretic mobility is affected in all three forms upon the addition of N-glycosidase F.

Characterization of substrate specificities of CPA from *H. armigera*

A series of synthetic substrates with the same spectrophotometric characteristics were used in the kinetic measurements to calculate the values of K_m, k_{cat} and k_{cat}/K_m for CPAHa and compare them to those of bovine CPA and human CPA2, two A-type enzymes from mammals (Table 1). These studies, as well as the inhibition kinetics measurements (see below), were always performed with the active enzyme generated by LysC, even though the enzyme generated by trypsin showed similar enzymatic properties. CPAHa is unable to hydrolyse synthetic substrates containing C-terminal Trp residues, in contrast to CPA2. This, together with its capability to cleave substrates containing Phe or Tyr as C-terminal, allows to classify CPAHa as an enzyme of the A1 subtype. In most instances, the insect enzyme appears to be less efficient than the mammal enzymes as judged by the k_{cat}/K_m values but, on the other hand, displays a broader substrate specificity. Its ability to hydrolyse Cbz-Gly-Gly-Ala is similar to rat CPA1 [6] the only carboxypeptidase known able to hydrolyse this substrate. It also displays activity against Cbz-Gly-Gly-Leu

Table 1. Kinetic constants for peptide substrate hydroysis by *H. armigera* CPA (CPAHa), bovine CPA (CPAb) and human CPA2 (CPA2 h). NM: not measurable.

	CPAHa			CPAb			CPA2 h		
Substrate	k_{cat} (s^{-1})	K_m (μM) ($M^{-1}.s^{-1}$)	(k_{cat}/K_m) $\times 10^{-5}$	k_{cat} (s^{-1})	K_m (μM) ($M^{-1}.s^{-1}$)	(k_{cat}/K_m) $\times 10^{-5}$	k_{cat} (s^{-1})	K_m (μM) ($M^{-1}.s^{-1}$)	(k_{cat}/K_m) $\times 10^{-5}$
Cbz-Gly-Gly-Phe	35.6 ± 1.8	506 ± 31	0.704	131.5 ± 3.1[a]	172 ± 12	7.62	58.3 ± 2.4	372 ± 30	1.57
Cbz-Gly-Gly-Tyr	57.0 ± 3.8	238.71 ± 40	2.38	56.3 ± 2.0[a]	102 ± 2	5.51	70.0 ± 5.3	125 ± 15	5.6
Cbz-Gly-Gly-Trp	NM	NM	NM	NM[b]	NM	NM	90.3 ± 7.0	146 ± 9	6.18
Cbz-Gly-Gly-Leu	49.4 ± 2.8	746 ± 150	0.662	63.4 ± 2.5[b]	1180 ± 93	0.54	11.8 ± 1.1	5300 ± 1400	0.03
Cbz-Gly-Gly-Val	0.3 ± 0.013	1748 ± 321	$1.72E10^{-3}$	19.5 ± 2.0[c]	3720 ± 390	0.052	NM	NM	NM
Cbz-Gly-Gly-Ala	6.25 ± 0.2	2618 ± 580	0.024	NM[c]	NM	NM	NM	NM	NM
Cbz-Gly-Gly-Ser	NM	NM	NM	NM[c]	NM	NM	NM	NM	NM
Cbz-Gly-Phe	35.6 ± 1.8	328 ± 40	1.095	41.7 ± 2.8[a]	1093 ± 154	0.38	16.1 ± 1.3	2270 ± 200	0.07
Cbz-Gly-Tyr	58.2 ± 4.2	289 ± 68	2.01	16.0 ± 0.6[a]	394 ± 29	0.41	9.7 ± 1.3	175 ± 10	0.56
Cbz-Gly-Trp	NM	NM	NM	50.0 ± 4.3[a]	3310 ± 430	0.15	33.8 ± 1.1	261 ± 12	1.29

Taken from [a] [6], [b] [23] and [c] [7].

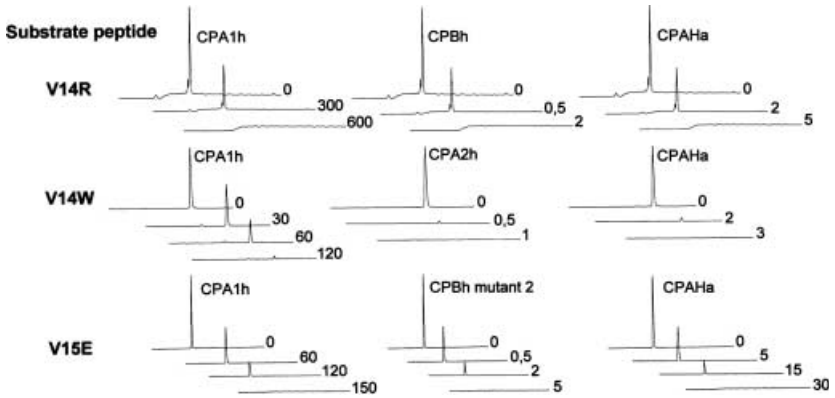


Fig. 5. Analysis of substrate specificity of PCPAHa with peptides. Comparative analysis by reverse-phase HPLC of the degradation of three synthetic substrates by CPAHa. Degradation of V15E (VKKKARKAAGC(Acm)AWE) by CPA1 h, CPBh mutant 2 (cleaves acidic C-ter residues) and CPAHa. Degradation of V14R (VKKKARKAAGGAKR) by CPA1 h, CPBh and CPAHa. Degradation of V14W (VKKKARKAAGC(Acm)AW) by CPA1 h, CPA2 h and CPAHa. The numbers beside the chromatograms indicate the reaction times for each enzyme-substrate combination. The chromatographic conditions are explained in Materials and methods.

and a measurable k_{cat}/K_m for the substrate with a C-terminal valine.

A remarkable difference between the insect enzyme and the mammalian ones is the ability of the former to hydrolyse short substrates. In contrast to the mammalian enzymes, CPAHa hydrolytic efficiency for Cbz-Gly-X is very similar to that displayed against Cbz-Gly-Gly-X, suggesting that the importance of the secondary substrate binding subsites is reduced in the insect enzyme.

Three different peptides were used as substrate models to analyse the ability of CPAHa for cleaving acid, basic and tryptophan C-terminal residues. In each assay, CPAHa was compared with a similar carboxypeptidase, human CPA1, and a second one chosen according to its specificity for the residue being analysed (Fig. 5). Relative cleavage rates were calculated on the basis of the time needed by each carboxypeptidase to fully degrade the same amount of initial substrate.

The ability of CPAHa to cleave all three peptides was always better than that of CPA1 h, an observation specially clear in the case of the peptide with a C-terminal arginine (V14R), which is cleaved 120 times faster by CPAHa. Compared with human CPB, a prototype enzyme for basic residue specificity, CPAHa showed a relatively high affinity for C-terminal arginine in peptide V14R since its relative cleavage rate is only 2.5 times smaller, whilst human CPA1 can hardly hydrolyse this substrate at all. A similar result is observed for V14W, where the relative cleavage rate for human CPA2, a very specific enzyme for peptides with tryptophan at the C-terminus, is three times larger than that for CPAHa but 120 times larger than that of human CPA1. Finally the cleavage of V15E by the mutant human CPB was six times faster than that of CPAHa.

Measurement of equilibrium dissociation constant (K_i) for protein inhibitors

The K_i value was calculated for the recombinant forms of two different carboxypeptidase inhibitors, PCI [21] and LCI [22] (Table 2). The inhibition constant of LCI,

Table 2. K_i values of PCI and LCI against CPAHa compared to previous data obtained for CPAb [21,22].

	K_i (pM)	
Carboxypeptidase	PCI	LCI
CPAHa	65 ± 7.3	260 ± 32.5
CPAb	1500 ± 600	250–480

260 ± 32 pM, is similar to that of LCI for bovine CPA, 250–480 pM. However, the K_i of PCI for CPAHa, 65 ± 7 pM, is 23 times lower than the K_i of PCI for the bovine homologue, which is 1.5 ± 0.6 nM.

Molecular modelling and dynamics simulation

To further investigate the nature of the important K_i difference between the CPAHa–PCI and CPAb–PCI complexes, a model structure of the former was generated. The CPAHa–PCI complex was modelled using the known crystal structure of the CPAb–PCI complex as template (Protein Data Bank entry 4CPA). To avoid the presence of unrealistic interactions in the model, the structure of the CPAHa–PCI complex was relaxed under the conditions of a molecular force field by means of a 500 ps molecular dynamics simulation in aqueous solution. A reference simulation of the CPAb–PCI complex was also carried out.

A cartoon representation of the superimposed complexes, at simulation time $t = 0$, is shown in Fig. 6. The atom-position rmsd of the PCI backbone from its initial structure in each of the complexes ($t = 0$) is also given in Fig. 6 as a function of time. The calculated rmsd values of the inhibitor's structure, which contain information about both internal motions and motions relative to the CPA moiety, are similar in the two systems. Although the amplitudes of the rmsd fluctuations are smaller for the CPAHa–PCI complex, they should not be considered statistically significant because the timescale of the simulations is not sufficient to draw conclusions about relative stabilities. The purpose of the simulations was to relax the experimentally determined structure of CPAb–PCI and the model structure of CPAHa–PCI under the same molecular force field and conditions, in order to facilitate the comparison of the corresponding molecular interfaces.

In spite of the similar binding geometries (imposed by the modelling strategy), the two complexes appear fairly different in terms of specific interactions between enzyme and inhibitor (results not shown). The difference in average interaction energy between enzyme and inhibitor in the simulation (−1074 kJ·mol^{-1} CPAHa–PCI vs. −1257 kJ mol^{-1} for CPAb-PCI) is not sufficient to explain the remarkably lower K_i of the CPAHa–PCI complex. However, we note that the free energy of binding is equal to the work required to bring the two molecules from free solution to the solvated complex, and the above-mentioned interaction energy is only one of the components of this free energy.

Discussion

The high expression yield of the procarboxypeptidase from the insect pest *Helicoverpa armigera* attained in the methylotrophic yeast *Pichia pastoris* indicates both the suitability of this organism to host the heterologous expression of this class of enzymes [23,24] and the correct folding of the proenzyme. The latter is further confirmed by its activation by trypsin, its capability to degrade synthetic CP substrates and its susceptibility to protein inhibitors, proved to be effective on related metallocarboxypeptidases. Overall, the

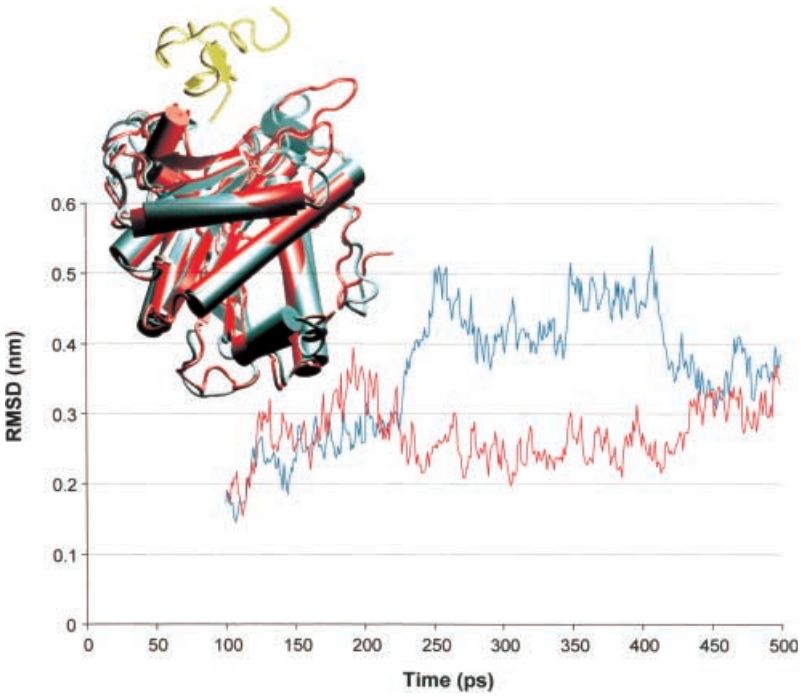


Fig. 6. Molecular modelling and dynamics simulation. Cartoon representation of a the least-squares fitted complexes CPAb–PCI and CPAHa–PCI, at simulation time $t = 0$; CPAb in cyan, CPAHa in red, PCI in yellow. Atom-position rmsd of the PCI backbone from its initial structure in each of the complexes ($t = 0$) as a function of time; CPAbPCI in cyan, CPAHa–PCI in red.

recombinant protein is thus indistinguishable from the natural one and constitutes a good model to study it.

Although trypsin can activate PCPAHa, as with many other procarboxypeptidases a lysine specific endopeptidase (LysC) can activate it four times more efficiently. However, activation with either protease releases an enzyme with identical activity against the synthetic substrates FAPP and AAFP. This is the first member of this family of proenzymes that can be activated more efficiently by a protease other than trypsin. The fact that several trypsin-like proteases from *H. armigera* have been cloned and sequenced [25] and that some of them show a higher degree of identity with LysC than with bovine trypsin suggests that this insect might possess a specific enzyme able to activate PCPAHa, as LysC does *in vitro*.

The activation point for trypsin *in vitro* is R4, an accessible residue located in an unstructured loop at the end of the connecting region between the activation domain and the enzyme moieties, a position very similar to that of most mammal procarboxypeptidases [9]. LysC, a serine protease that only recognizes lysine at the P1 position, is unable to activate any human pancreatic procarboxypeptidases, as previously observed in our laboratory. However, when acting on PCPAHa, it generates the active enzyme by cleavage at the carbonyl end of a lysine located four residues upstream of R4 and after five consecutive alanines (Fig. 2C), a sequence that could be a recognition motif for a highly specific activating protease. This motif, not found in any other protein, is repeated a second time near the C-terminal end of this molecule, and also in this case LysC has also been shown to be able to release the C-terminal peptide after specific cleavage. Whether the dual presence of the specific sequence motif is related to some hypothetical mechanism of control of the activity will require further investigation.

Between the two activation points described there is a consensus target for glycosylation (Asn-Ser-Thr) which does become glycosylated in the *P. pastoris* system, adding a mass of around 1900 Da. The presence of sugars seems to affect LysC activation, as demonstrated by the easier activation of deglycosylated PCPAHa. This adds a further possible regulatory mechanism which has never been observed before in enzymes of this family.

To achieve a complete *in vitro* activation of PCPAHa in a period of time similar to other activation studies performed with mammalian procarboxypeptidases it was found that the PCPAHa/LysC ratio needed was 40 : 1 (w/w) at 37 °C, and the activation process was studied in detail in these conditions. The timecourse of activity generation is hyperbolic and coincides with those described for procarboxypeptidases with a proregion that does not inhibit the enzyme after cleavage [20,23]. This is consistent with the observation that the prodomain is completely degraded during activation because LysC cleaves after all of the seven internal lysines. Besides the removal and degradation of the prodomain, LysC also causes the removal of a C-terminal octapeptide, which is placed after an $(A)_6K$ motif, almost identical to the sequence recognized by LysC at the border between the activation peptide and the enzyme moiety. The cleavage of the C-terminal peptide is much faster than the elimination of the proregion because disappearance of full-length PCPAHa occurs only 5 min after activation, while a complete CPAHa activity is only reached after 90 min (Figs 2B and 4). The parallel release of the active enzyme and the C-terminal peptide due to the highly specific action of an enzyme able to cleave after $(A)_nK$ might have some physiological relevance.

From the analysis with a series of carbobenzoxy (Cbz) substrates, and in a first instance, the mature enzyme derived from PCPAHa should be classified as A1, as it cleaves aliphatic and aromatic C-terminal residues but not tryptophan. Surprisingly, further analysis shows that the enzyme is also able to cleave C-terminal E, W and R residues, with a particularly good efficiency for the latter. In all cases, the insect enzyme was much more efficient than human pancreatic CPA1. This is the first reported case of a metallocarboxypeptidase showing such a wide specificity spectrum. S255, located in the S1′ pocket, which replaces a conserved isoleucine in the A-type carboxypeptidases and an equally conserved aspartate residue in the B forms might be responsible for this change in specificity [9].

The plant carboxypeptidase inhibitor PCI shows K_i values in the pM range with CPAHa, in contrast to the nM values displayed against mammalian carboxypeptidases. This supports the theory that PCI, which is expressed in potato leaves in response to wounding [26], may inhibit the digestive carboxypeptidases of potential insect pests. The impact of *H. armigera* in many different crops makes this efficient protein inhibitor very interesting in the design of new insecticide strategies. To investigate the structural bases of the strong binding of PCI to CPAHa, a molecular model of the CPAHa–PCI complex was generated based on the known structures of PCPAHa and the CPAb–PCI complex, and it was submitted to relaxation and structural analysis by a molecular dynamics approach. From these studies, the differences in the K_i values observed for CPAHa–PCI and CPAb–PCI cannot apparently be readily explained in terms of specific interactions in the model. This suggests that there may be local conformational differences between the structure of CPAHa in the proenzyme and in the complex which are not reproduced by the model, that the geometry of binding of PCI to CPAHa may differ from that assumed with the model and that the difference in the binding free energies of the two complexes may be dominated by other than the intermolecular interaction energy (e.g. enthalpy and/or entropy associated with desolvation and conformational changes upon complex formation). To evaluate these three possibilities, further computational studies are in progress.

Overall, the procarboxypeptidase A from *Helicoverpa armigera* and its derived enzyme, although apparently very similar both functionally and structurally to their mammalian counterparts, have some unique properties in terms of activation, specificity and regulation, which make them an interesting system that settles new questions on this family of enzymes, both in the basic and applied fields.

Acknowledgements

This work was supported by grant BIO2001-2046 (MCYT, Ministerio de Ciencia y Tecnología, Spain) and by the Centre de Referència en Biotecnologia (Generalitat de Catalunya, Spain). X. D. is grateful to W. F. van Gunsteren for granting access to computational resources at

the ETH Zurich. We wish to thank Drs John A. Gatehouse and David P. Bown, from the University of Durham, UK, for kindly providing us with the cDNA of PCPAHa and to Dr Sonia Segura for providing us with purified CPBh mutant2. We also thank Dr Salvador Bartolomé (LAFEAL-UAB) and Dr Francesc Canals (IBB-UAB) for technical assistance.

References

1. Terra, W.R. & Ferreira, C. (1994) Insect digestive enzymes: properties, compartimentalization and function. *Comp. Biochem. Physiol.* **109B**, 1–62.
2. Johnson, S.J., King, E.G. & Bradley, J.R., eds. (1986) Theory and tactics of *Heliothis* population management. I. Cultural and biological control. *South. Coop. Series Bull.* 316, 161.
3. Bown, D.P., Wilkinson, S. & Gatehouse, J.A. (1998) Midgut carboxypeptidase from Helicoverpa armigera (Lepidoptera: *Noctuidae*) larvae: enzyme characterization, cDNA cloning and expression. *Insect Biochem. Mol. Biol.* **28**, 739–749.
4. Wei, S., Segura, S., Vendrell, J., Aviles, F.X., Lanoue, E., Day, R., Feng, Y. & Fricker, L.D. (2002) Identification and characterization of three members of the human metallocarboxypeptidase gene family. *J. Biol. Chem.* **277**, 14954–14964.
5. Vendrell, J. & Aviles, F.X. (2000) Metallocarboxypeptidases and their protein inhibitors. *Biochem. Biophys. Acta.* **1477**, 284–298.
6. Gardell, S.J., Craik, C.S., Clauser, E., Goldsmith, E.J., Steward, C.B., Graf, M. & Rutter, W.J. (1988) A novel rat carboxypeptidase, CPA2: characterization, molecular cloning and evolutionary implications on substrate specificity in the carboxypeptidase gene family. *J. Biol. Chem.* **263**, 17828–17836.
7. Reverter, D., García-Saez, I., Catasús, L., Vendrell, J., Coll, M. & Avilés, F.X. (1997) Characterisation and preliminary X-ray diffraction analysis of human pancreatic procarboxypeptidase A2. *FEBS Lett.* **420**, 7–10.
8. Vendrell, J., Cuchillo, C. & Avilés, F.X. (1991) The trypsin activation pathway of monomeric procarboxypeptidase A. *J. Biol. Chem.* **265**, 6949–6953.
9. Estébanez-Perpiñá, E., Bayés, A., Vendrell, J., Jongsma, M.A., Bown, D., Gatehouse, J.A., Huber, R., Bode, W., Aviles, F.X. & Reverter, D. (2001) Crystal structure of a novel mid-gut procarboxypeptidase from the cotton pest *Helicoverpa armigera*. *J. Mol. Biol.* **313**, 629–638.
10. Sambrook, J., Fritsch, E.F. & Maniatis, T. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
11. Peterson, L.M., Holmquist, B. & Bethune, J.L. (1982) Unique activity assay for carboxypeptidase A in human serum. *Anal. Biochem.* **125**, 420–426.
12. Mock, W.L. (1996) Arazoformyl peptide surrogates as spectrophotometric kinetic assay substrates for carboxypeptidase A. *Anal. Biochem.* **239**, 218–222.
13. Leatherbarrow, R.J. (1987) Enzfitter: a non-linear regression data analysis program for the IBM-PC. Elsvier Biosoft, Cambridge, UK.
14. Edge, M., Forder, C., Hennam, J., Lee, I., Tonge, D., Harden, I., Fitton, J., Eckersley, K., East, S., Shuffebotham, A., Blakey, D. & Slater, A. (1998) Engineered human carboxypeptidase B enzymes that hydrolyse hippuryl-L-glutamic acid: reversed-polarity mutants. *Protein Eng.* **11**, 1229–1234.
15. Bieth, J.G. (1995) Theoretical and practical aspects of proteinase inhibition kinetics. *Methods Enzymol.* **248**, 59–84.
16. van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P. & Tironi, I.G. (1996) *Biomolecular simulation: The GROMOS96 manual and user guide*. vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v., Zürich, Groningen.
17. Scott, W.R.P., Hünenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Torda, A.E., Huber, T., Krüger, P. & van Gunsteren, W.F. (1999) The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103**, 3596–3607.
18. Schuler, L.D., Daura, X. & van Gunsteren, W.F. (2001) An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comp. Chem.* **22**, 1205–1218.
19. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. & Hermans, J. (1981) Interaction models for water in relation to protein hydration. In *Intermolecular Forces*. (Pullman, B., ed.), pp. 331–342. D. Reidel Publishing Co., Dordrecht.
20. Burgos, F.J., Salvà, M., Villegas, V., Soriano, F., Méndez, E. & Aviles, F.X. (1991) *Biochemistry* **30**, 4082–4089.
21. Molina, M.A., Marino, C., Oliva, B., Avilés, F.X. & Querol, E. (1994) C-tail valine Is a key residue for stabilitzation of complex between potato inhibitor and carboxypeptidase A. *J. Biol. Chem.* **269**, 21467–21472.
22. Reverter, D., Vendrell, J., Canals, F., Horstmann, J., Aviles, F.X., Fritz, H. & Sommerhoff, C.P. (1998a) A carboxypeptidase inhibitor from the medical leech *Hirudo medicinalis*. Isolation, sequence analysis, cDNA cloning, recombinant expression and characteritzation. *J. Biol. Chem.* **273**, 32927–32933.
23. Reverter, D., Ventura, S., Villegas, V., Vendrell, J. & Aviles, F.X. (1998b) Overexpression of human procarboxypeptidase A2 in *Pichia pastoris* and detailed characteritzation of its activation pathway. *J. Biol. Chem.* **273**, 3535–3541.
24. Ventura, S., Villegas, V., Sterner, J., Larson, J., Vendrell, J., Hershberger, C.L. & Aviles, F.X. (1999) Mapping the pro-region of carboxypeptidase B by protein engineering. Cloning overexpression and mutagenesis of the porcine proenzyme. *J. Biol. Chem.* **274**, 19925–19933.
25. Bown, D.P., Wilkinson, H.S. & Gatehouse, J.A. (1997) Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagus insect pest, Helicoverpa armigera, are members of complex multigene families. *Insect Biochem. Molec. Biol.* **27**, 625–638.
26. Villanueva, J., Canals, F., Prat, S., Ludevid, D., Querol, E. & Aviles, F.X. (1998) Characterization of the wound-induced metallocarboxypeptidase inhibitor from potato. *FEBS Lett.* **440**, 175–182.

Supplementary material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/EJB/EJB3681/EJB3681sm.htm

Appendix S1. Procarboxypeptidase A from the insect pest *Helicoverpa armigera* and its derived enzyme. Two forms with new functional properties.