Appl Microbiol Biotechnol (2006) 73:356–365
DOI 10.1007/s00253-006-0498-z

BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Genetic characterization and expression of the novel fungal protease, EPg222 active in dry-cured meat products

María J. Benito · Ian F. Connerton · Juan J. Córdoba

Received: 13 February 2006 / Revised: 19 April 2006 / Accepted: 8 May 2006 / Published online: 29 August 2006
© Springer-Verlag 2006

Abstract EPg222 protease is a novel extracellular enzyme produced by *Penicillium chrysogenum* (Pg222) isolated from dry-cured hams that has the potential for use over a broad range of applications in industries that produce dry-cured meat products. The gene encoding EPg222 protease has been identified. Peptide sequences of EPg222 were obtained by de novo sequencing of tryptic peptides using mass spectrometry. The corresponding gene was amplified by PCR using degenerated primers based on a combination of conserved serine protease-encoding sequences and reverse translation of the peptide sequences. EPg222 is encoded as a gene of 1,361 bp interrupted by two introns. The deduced amino acid sequence indicated that the enzyme is synthesized as a preproenzyme with a putative signal sequence of 19 amino acids (aa), a prosequence of 96 aa and a mature protein of 283 aa. A cDNA encoding EPg222 has been cloned and expressed as a functionally active enzyme in *Pichia pastoris*. The recombinant enzyme exhibits similar activities to the native enzyme against a wide range of protein substrates including muscle myofibrillar protein. The mature sequence contains conserved aa residues characteristic of those forming the catalytic triad of serine proteases (Asp42, His76 and Ser228) but notably the food enzyme exhibits specific aa substitutions in the immunoglobulin-E recognition regions that have been identified in protein homologues that are allergenic.

Keywords *Penicillium chrysogenum* · Protease · Proteolysis · Dry-cured meat products

M. J. Benito · I. F. Connerton
Division of Food Sciences, School of Biosciences,
University of Nottingham,
Sutton Bonnington Campus,
Loughborough LE12 5RD, UK

Present address:
M. J. Benito (✉)
Nutrición y Bromatología, Escuela de Ingenierías Agrarias,
Universidad de Extremadura,
Carretera de Cáceres s/n,
06071-Badajoz, Spain
e-mail: mjbenito@unex.es

J. J. Córdoba
Higiene de los Alimentos, Facultad de Veterinaria,
Universidad de Extremadura,
10071 Cáceres, Spain

Introduction

Proteolysis is considered to be one of the major processes involved in aroma development in dry-cured meat products. In these kinds of products, proteolysis has been attributed to microorganisms growing on the surface, mainly mould species (Bruna et al. 2001; Martín et al. 2003, 2004; Rodríguez et al. 1998). The use of selected proteolytic moulds or their purified enzymes could be of great interest to the dry-cured meat industry as they have the potential to accelerate the ripening process and flavour development. Using an enzyme is an attractive option, as it could eliminate the need for mould growth and the inherent variation in product quality. Protease EPg222 purified from a non-toxigenic strain of *Penicillium chrysogenum* has shown a high proteolytic activity against myofibrillar proteins under the conditions of temperature, pH and NaCl concentration present in dry-cured meat products (Benito et al. 2002) and also in sterile ripened pork and dry fermented sausages (Benito et al. 2003b,c – 2005a). To achieve these benefits, genetic characterization and efficient expression of EPg222 protease is required in sufficient amounts to assess the utility of the protease, and thereafter, make recommendation to the meat industry.

The methylotrophic yeast, *Pichia pastoris*, has been successfully used for expression of heterologous proteins (Goh et al. 2001; Higgins and Cregg 1998; Rodríguez et al. 2000). This expression system uses relatively simple and inexpensive medium to produce high yields of extracellular proteins mediated by highly inducible alcohol oxidase (AOX1) promoter or the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter. As several proteases from bacterial, fungal or mammalian origin have been successfully expressed in *P. pastoris* (Sun et al. 1997), and as the organism is Generally Regarded As Safe (United States FDA, Food Additive Status List October 2005) we opted to use this organism for expression of a food compatible enzyme.

In this paper, we report the molecular cloning of the gene encoding the serine protease, EPg222, from *P. chrysogenum* strain, Pg222. The functional expression of the gene product in *P. pastoris* has a potential for application in the meat industry.

Materials and methods

Microbial strain: The *P. chrysogenum* strain, Pg222, isolated from dry-cured ham and demonstrated to possess high proteolytic activity (Benito et al. 2003a, 2005b), was used in this study.

Exogenous protease: EPg222 is a serine protease obtained from a strain of *P. chrysogenum*, Pg222. The optimum working conditions of this protease are 45–55 °C, pH 6 and 0.25-M NaCl (Benito et al. 2002).

Induction of protease production and purification: Culture medium was obtained by mixing myofibrillar protein extract (Benito et al. 2002) with sterile nutrient broth and NaCl to reach final concentrations of 1.6 mg of protein per milliliter, 0.1% (w/v) nutrient broth, and 5% (w/v) NaCl. The pH of the culture medium was 5.8. Next, the medium was inoculated with 100 µl of a suspension of spores from *P. chrysogenum*, Pg222, containing approximately 10^6 spores/milliliter and incubated at 25 °C for 4 days under continuous shaking (200×*g*) in an orbital incubator (Certomat IS model; B. Braun Biotech International, Melsungen, Germany).

After incubation, the mycelium was removed by filtration, and the filtrate was treated as indicated in the following steps, performed at 4 °C: (1) Ammonium sulfate precipitation. Ammonium sulfate was added to the filtrate (250 ml) to reach 80% saturation. After centrifugation at 9,000×*g* for 15 min, the precipitate was dissolved in 50 ml of 50-mM potassium phosphate buffer (pH 7.2) and dialyzed thoroughly against this buffer overnight. (2) Ultrafiltration. The protein was concentrated by ultrafiltration with centrifugal filters with 50- and 30-kDa cutoffs (Centricon Plus-20; Millipore, Bedford, Mass. USA).

In preparation for protein sequencing, the protease samples were separated on 1-mm thick homogeneous 12% (w/v) acrylamide resolving gels and 4% (w/v) acrylamide stacking gels with the buffer systems (Laemmli 1970) using the Bio-Rad mini Protean 3 apparatus (Bio-Rad Laboratories, Hemel Hempstead, UK). Polymerized acrylamide gels were left overnight to eliminate cross-linking of unpolymerized acrylamide to proteins during electrophoresis. Equal volume of sample buffer that contained 65-mM Tris–HCl, pH 6.8, 25% (v/v) glycerol, 2% (w/v) SDS, 5% (v/v) 2-mercaptoethanol and 0.2% (w/v) bromophenol blue was added to the protein sample and boiled for 5 min before loading on the gels. The proteins were separated at constant voltage of 100 V in running buffer containing 25-mM Tris, 192-mM glycine and 0.1% (w/v) sodium dodecyl sulfate (SDS), pH 8.3. Separated proteins were stained with colloidal Coomassie blue G-250 from which the bands were excised for protein sequencing. Alternatively, the 35-kDa band was transferred to polyvinylidene difluoride membrane (Amersham Biosciences UK, Little Chalfont, UK) and the N-terminal sequences determined using an Applied Biosystems Procise 494 sequencer (PE Applied Biosystems, Warrington, UK).

Mass spectrometry analysis of tryptic peptides: Excised protein spots were destained in buffer (50% acetonitrile, 50% digestion buffer: 100-mM ammonium bicarbonate) for 10 min. Two dehydration steps were performed, followed by either trypsin (6 ng/µl in 50% digestion buffer) or glutamic-C endopeptidase (10 ng/µl in 50-mM Tris–HCl pH 8.0) digestion in 25 µl and incubated at 8 °C 40 min, and after this at 37 °C. Ten microliters aliquot were removed after 2.5 h, and the reminder was incubated overnight at 37 °C. Extraction was performed from samples after 2.5 h of incubation by the addition of formic acid at 0.1% from which, 1 µl of extracted peptide solution was injected into the Micromass ES Q-TOF Ultima for MS/MS analysis. For data analysis, PKL files generated were used to perform a Mascot MS/MS ion search from http://www.matrixscience.com/. Search parameters included: Database, MSDB; taxonomy, all entries; enzyme, trypsin or glutamyl endopeptidase; fixed modifications, carbamidomethyl (C); variable modifications, oxidation (M); peptide charge, 2+ and 3+; data format; Micromass (.PKL); instrument, ESI-QUAD-TOF. Significant sequence homologies to known proteins were noted.

DNA isolation: In liquid nitrogen, 2 g of frozen mycelium was ground to a fine powder, homogenized in 4-ml 0.05 M

Tris–Cl, pH 8; 0.005-M ethylenediamine tetraacetic acid (EDTA); 0.05-M NaCl (TES) buffer and sodium dodecyl sulfate (SDS) to a final concentration of 1% (w/v). Then, 200-μl proteinase K (10 μg/μl) was added, incubated at 60 °C for 35 min, cooled on ice and extracted with phenol–chloroform–isoamyl alcohol (25:24:1). This suspension was centrifuged at 3,000×*g* for 2 min. The upper phase containing DNA was transferred to another centrifuge tube and precipitated by adding 3-M sodium acetate to a final concentration of 10% (w/v) and two volumes of cold ethanol. After centrifugation, the pellet was resuspended with sterile water and treated with 50-μl RNase (10 μg/μl). Finally, DNA was precipitated again as indicated above.

Genomic DNA amplification: Polymerase chain reactions (PCR) were performed with genomic DNA (20 ng) as template in 50-μl reaction mixtures containing 50-pmol each of forward and reverse primers, 500 mM each of deoxyribonucleotide triphosphate (dNTP), 0.1 vol. of ×10 PCR buffer (500-mM Tris–HCl, pH 9.2, 140-mM $(NH_4)_2SO_4$, 22.5-mM $MgCl_2$) and 2.5-U pfu polymerase (Stratagene, La Jolla, CA, USA). DNA amplification was carried out in a Progene thermocycler (Techne, Cambridge, UK) through 28 cycles of: denaturation (30 s, 94 °C), annealing (40 s, 45 °C), and extension (2 min and 30 s, 70 °C) followed by 70 °C for 5 min. PCR reactions were performed using eight degenerate primer combinations Fwd 1, Rev 1, Rev 1a, Fwd 2, Rev 2 designed on the basis of reverse translation of the peptide sequences determined by mass spectrometry (Table 1). The resulting PCR products were gel-purified before cloning into the pCR2.1-TOPO vector (Invitrogen, Paisley, UK) and sequenced using a Perkin-Elmer ABI Prism 310 fluorescent DNA analyser (PE Applied Biosystems, Warrington, UK).

Specific primers were designed from the sequences determined (Fig. 1), in the middle of the gene (Fwd half EPg, Rev half EPg) and the primer for the end, Rev1a, and some high homology sequences for the signal sequence at the beginning (Fwd B) (Table 1). These were used to amplify the complete coding sequence from *P. chrysogenum* Pg222 genomic DNA, involving 28 cycles of: denaturation (30 s, 94 °C), annealing (40 s, 58 and 50 °C for the second reaction) and extension (1 min, 72 °C), followed by 10 min at 70 °C. The obtained PCR products using the nested primers were sequenced and these cloned in pCR2.1-TOPO vector were also sequenced, obtaining the whole sequence.

*Southern blotting: Eco*RI, *Bam*HI and *Sma*I restriction digests of *P. chrysogenum* Pg222 genomic DNA were electrophoresed in 0.8% agarose gels and blotted by capillary transfer onto Hybond-N+ (Amersham Biosciences) nylon membrane before fixation with UV light. The blots were probed with a specific PCR fragment (361 bp, see above) labelled using the DIG-labeling system (Roche Diagnostics, Lewes, UK), according to the manufacturer's instructions. Colorimetric detection was performed in the presence of the nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) substrates (Roche), according to the manufacturer's instructions.

RNA isolation and RT-PCR: P. chrysogenum Pg222 was grown in the media for induction of protease production in 4 days, as described above. Total RNA was extracted from

Table 1 Names and sequences of oligonucleotides used to identify the EPg222 gene

Primer names	Primer sequences with their translations[a]
Fwd 1	5′-GTN GTN CAR CCN AAY GCN CC-3′
	V V Q P N A P
Rev 1:	3′-TT NGC RTT NAG NCC RTT RAA-5′
	K A N L G N F
Rev 1a	3′-TT NGC RTT YAA NCC RTT RAA-5′
	K A N L G N F
Fwd 2	5′-TGG GCN GTN AAR GAY GCN AA-3′
	W A V K D A K
Rev 2	3′-TT NGC RTC YTT NAC NGC CCA-5′
	K A D K V A W
Fwd half EPg	5′-GGATTGACATTAGGCATTCCGACTT-3′
Rev half EPg	3′-AAGTCGGAATGCCTAATGTCAATCC-5′
Fwd B	5′-ATGGGTTTCCTCAAGGTCCTCGCTA-3′
WEPg1F	5′-ATGGGTTTCCTCAAGGTCCTCGCTACGTCCCTTGCGACTCTGGCGGTC-3′
WEPg2R	3′-TTACTTCGCGTTTAAACCGTTGAAGAGAAGCTTGCTAGTAGTGCCGAA-5′
Rev intron1	3′-GTACTTGACCGCCGGGTCACTGGCAATGTC-5′
Fwd-intron1	5′-GACCCGGCGGTCAAGTACGTTGAACCAGAC-3′
Rev-intron2	3′-AGAGGTCGACAACGGAACCGAAGTTGCTGA-5′
Fwd-intron2	3′-GTTCCGTTGTCGACCTCTATGCTCCTGGGG-5′

[a] *R* for purine, *Y* for pirimidine, *N* for A, C, T or G

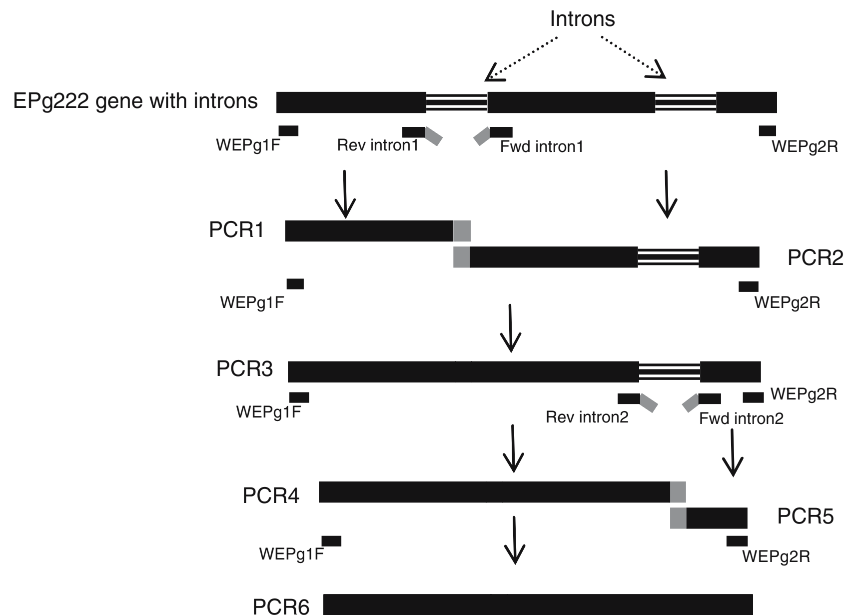

Fig. 1 Schematic representation of splice-overlap extension strategy performed using the polymerase chain reaction (SOE by PCR). Introns are represented by the *open bars* and exons by the *solid bars*

deep-frozen mycelium employing the RNeasy Kit from Qiagen according to the manufacturer's instructions. Reverse transcriptase-polymerase chain reaction (RT-PCR) was performed using the ImProm-II Reverse Transcription System (Promega). Total RNA (3 µg) and Oligo dT (0.5 µg) in nuclease-free water for a final volume of 5 µl were preheated 70 °C for 5 min and chilled in ice-water for 5 min. The following reagents were then added: ImProm-II ×5 Buffer (4 µl), $MgCl_2$, (25 mM; 4.8 µl), 10-mM dNTP mix (1.0 µl), recombinant RNasin ribonuclease inhibitor (20 U), ImProm-II reverse transcriptase (1.0 µl) and nuclease-free water to a final volume of 20 µl. The reaction was incubated at 42 °C for 50 min and inactivated at 70 °C for 15 min. The RT reaction (1 µl) was used as template for PCR with WEPg1F forward (100 ng/µl; 1 µl) and WEPg2R reverse primers (100 ng/µl; 1 µl) (Table 1). cDNA amplification was carried out through 40 cycles of denaturation (30 s at 94 °C), annealing (40 s at 60 °C) and extension (2 min 30 s at 70 °C), followed by 10 min at 70 °C. The resulting PCR product was cloned into the plasmid vector pCR2.1-TOPO and sequenced.

Cloning and expression of the esp1 gene in P. pastoris strain GS115: To format the *esp1* gene for expression in *P. pastoris* the intron DNA sequences were removed from the genomic clone by splice-overlap extension PCR and restriction sites appended for cloning into the expression vector. Specific oligonucleotides are given in Table 1. EPg222 gene fragments were assembled according to the protocol of SOE-PCR (Horton et al. 1990) as indicated in Fig. 1. The 5′-ends of the appropriate primers (Rev intron1–Fwd intron1 and Rev intron2–Fwd intron2) were designed in a way that both PCR products shared the same sequence at one end. In a subsequent step, PCR products are allowed to hybridize to one another, forming an overlap. Extension of this overlap by DNA polymerase yields a recombinant DNA molecule. For these reactions, pfu turbo polymerase (Stratagene) was used with 1 µl of PCR fragment as template in 50-µl reaction mixtures containing 50 pmol each of forward and reverse primers, 500 mM each of dNTP, 0.1 vol. of 10× PCR buffer (500-mM Tris–HCl, pH 9.2, 140 mM $(NH_4)_2SO_4$, 22.5-mM $MgCl_2$) and 2.5-U pfu turbo polymerase (Stratagene). DNA amplification involved 28 cycles of: denaturation (30 s, 94 °C), annealing (40 s, 60 °C in PCR1, 3, and 5 or 57 ° in PCR 2 and 4 or 55 ° in PCR6) and extension (70 °C during 30 s in PCR1 and 5, 1 min in PCR2 and 4 or 1 min 30 s in PCR 3 and 6) followed by 70 °C for 5 min. The resulting PCR6 product was gel-purified before cloning into the pCR2.1-TOPO vector and sequenced as described above.

Primers FwdEPg222-*Eco*RI (5′-TTTTGAATTCATGGGTTTCCTCAAGGTCCTCG-3′) and RevEPg222-*Not*I (5′-TTTTCAGCGGCCGCTTACTTCGCGTTTAAACCGT-3′) were used to introduce the restriction sites *Eco*RI and *Not*I at the 5′ and 3′ ends, respectively, by PCR amplification of the intronless gene sequence. The PCR product encoding the EPg222 preproprotein was digested with the restriction endonucleases, *Eco*RI and *Not*I. The fragment was subsequently gel-purifed and ligated into the corresponding restriction sites of the vector, pPIC3.5K, resulting in the expression vector, pPIC3.5K/esp1. For examination of the construct sequence, the vector,

pPIC3.5K/esp1, was transformed into *Escherichia coli* strain DH5α (Invitrogen).

P. pastoris strain GS115 (*his*4) was used to produce the recombinant enzymes with the expression vector, pPIC3.5K/esp1. *P. pastoris* was electroporated using the method described by Becker and Guarente (1991). The expression vector was linearized with *Dra*I to allow gene replacement at *P. pastorisAOX*1 gene and electroporated (Bio-Rad Gene Pulser II) into electrocompetent *P. pastoris* cells (Becker and Guarente 1991). Transformants were selected by their ability to grow at 30 °C on histidine-deficient RD plates.

P. pastoris was grown in 10-g yeast extract, 20-g peptone and 20-g dextrose per liter (YPD) medium or 10-g yeast extract, 20-g peptone, 13.4-g yeast nitrogen base (BMGY) medium, 0.2-g biotin, 10-ml glycerol and 100 ml of 1-M K_2HPO_4/KH_2PO_4, pH 6.0, per liter) and transformants were recovered on regeneration dextrose (RD) plates (186-g sorbitol, 20-g agar, 20-g dextrose, 13.4-g yeast nitrogen base, 0.2-g biotin and 50-mg amino acid mix without histidine per liter). Minimal dextrose (MD: 20-g dextrose, 13.4-g yeast nitrogen base and 0.2-g biotin) and Minimal methanol (MM: 5-ml methanol, 13.4-g yeast nitrogen base and 0.2-g biotin) plates were used to assess the Methanol Utilization phenotype of *Pichia* transformants. For induction, BMMY medium was used (10-g yeast extract, 20-g peptone, 13.4-g yeast nitrogen base, 0.2-g biotin, 5-ml methanol and 100 ml of 1-M K_2HPO_4/KH_2PO_4, pH 6.0, per liter). Protein expression was monitored by SDS-polyacrylamide gel electrophoresis (PAGE) and any candidate protein bands were excised manually for the analysis by Mass spectrometry, as described above.

Hydrolysis of different sources of proteins by recombinant EPg222: The effectiveness of the recombinant EPg222 was estimated in hydrolyzing bovine serum albumin (BSA) (1 mg/ml) (Sigma) and pork myofibrillar and sarcoplasmic proteins (4 mg/ml). To obtain the sarcoplasmic proteins extract, 2 g of sample was homogenised with 20 ml of 0.03 M, pH 7.4 sodium phosphate buffer in a Sorvall omnimixer (Omni Corporation International Instruments, Waterbury, CT, USA). The extract was centrifuged at 8,000×*g* for 15 min at 4 °C, and the supernatant was filtered through a 0.45-μm filter. Myofibrillar proteins extract was obtained after extraction of the resultant pellet with 20 ml of 1.1-M potassium iodide, 0.1-M sodium phosphate buffer, pH 7.4, following the procedure described by Benito et al. (2003b).

A final digestion volume of 250 μl contained 1.15 μg of EPg222. Protein hydrolysis was carried out in 0.05-M sodium phosphate buffer (pH 6) at 45 °C for 4 h. From these digestions, 15 μl of the digestion products were analysed by either 12 or 5% SDS-PAGE, run against undigested protein controls. Protein concentration was evaluated using $Abs_{280\ nm}$ and the method of Bradford (1976).

Kinetic constants: The synthetic substrate Nα-benzoyl-*dl*-arginine-*p*-nitroanilide (BAPNA) (Sigma, Dorset UK) was used. Each reaction mixture, contained 0.1 ml of a solution of purified enzyme (1 mg/ml), 0.99 ml of 1-mM BAPNA in 50 mM Tris–HCl, pH 8.0, 10-mM $CaCl_2$ buffer at 45 °C. Production of *p*-nitroaniline was measured by monitoring the increment in $Abs_{410\ nm}$ every 30 s for 10 min. BAPNA hydrolysis units (*U*) were calculated with the following equation: $U=(A_{410}/\mathrm{min})(1{,}000)(1/8{,}800)(\mathrm{mg\ enzyme})$, where 8,800 is the *p*-nitroaniline molar extinction coefficient. The Michaelis–Menten constant (K_m) and maximum velocity (V_{max}) were evaluated. The initial velocity of the enzymatic reaction was evaluated at 30 °C by varying BAPNA substrate concentration between 1 and 0.01 mM. K_m and V_{max} were evaluated by Lineweaver–Burke plot using Graphit and Excel software.

Results

Protein sequence analysis of the EPg222 protease

Extracts of the *P. chrysogenum* Pg222 culture were subject to ammonium sulfate precipitation and ultrafiltration and demonstrated to hydrolyze pork myofibrillar protein as demonstrated by Benito et al. (2002). These preparations were electrophoresed on 12% SDS-PAGE, and a candidate band migrating at 35 kDa, corresponding to EPg222, was excised for protein sequence analysis (Fig. 2). Lower molecular weight bands were also evident in the preparation as possible autolysis products; amongst these was a major product migrating at 30 kDa. Separation of the 30-kDa protein by ultrafiltration demonstrated that this protein was stable but retains no proteolytic activity. To confirm the origin of this protein, the 30-kDa band was also excised for sequence analysis. Protein sequences of peptides derived from the 35-kDa band were determined using mass spectrometry: AMNDAAANVVK, LLFNG LNAK, SVMNMSLGGAFSR, ENVVQPNAPSWGL and WAVKDAK. These sequences exhibited similarity to the protein sequences of *Penicillium citrinum* Pen c1 and *P. chrysogenum* Pen n13 (formerly ascribed a *Penicillium notatum*) proteases present in the database (Chow et al. 2000; Su et al. 1999). N-terminal sequence analysis of the product migrating at 35 kDa yielded a sequence of ENVVQP indicating the N-terminal peptide was represented in mass spectrum. Analysis of the 30-kDa band identified the peptide sequences: AMNDAAANVVK, LLFNGLNAK and SVMNMSLGGAFSR. These sequences

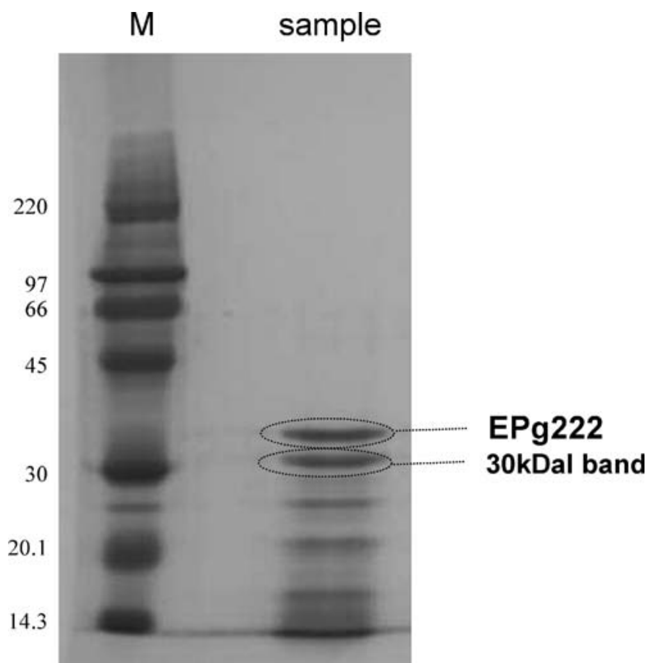

Fig. 2 SDS-PAGE (12.5%) of exogenous proteins obtained after ammonium sulfate precipitation and concentration by ultrafiltration with centrifugal filters with 50- and 30-kDa cut-offs (marked sample). *M* shows the standard protein molecular weight markers in kDa

were identical to those found in the protein migrating at 35 kDa. However, the absence of the molecular ions representing the N-terminal peptide was notable.

Molecular cloning of the gene encoding EPg222 protease

PCR reactions were performed with *P. chrysogenum* Pg222 genomic DNA using degenerate primers designed following reverse translation of the protein sequences of EPg222. Of the primer combinations attempted, PCR amplification was successful with the primes Fwd1 and Rev2 to yield a DNA fragment with 361 bp. The DNA sequence of this fragment showed significant identity (84–92%) with database sequences encoding serine proteases from *P. chrysogenum* (AF321100, AF084546, AF193420 and AY596924). A Southern blot of *Eco*RI, *Bam*HI and *Sma*I restriction digests of *P. chrysogenum* Pg222 genomic DNA was probed with the 361-bp DNA fragment to reveal the sequence was present as a single copy in the *P. chrysogenum* Pg222 genome. The unique nucleotide sequences together with conserved sequences noted in the database permitted the design of nested and flanking primers (Table 1). Using these primers, the corresponding gene was amplified as a 1,344-nucleotide sequence that contains a reading frame of 398 codons interrupted by two introns. The assignment of the reading frame and the locations of the introns were confirmed by RT-PCR amplification of a full-length cDNA sequence from RNA extracted from *P. chrysogenum* Pg222 grown on medium containing pork myofibrillar protein extract. These nucleotide sequences appear in the GenBank database under the accession number AJ870492. The protease-encoding gene was designated as *esp1* (extracellular serine protease) and represents a unique allele of *P. chrysogenum*.

The EPg222 protein is synthesized as a 398 amino acid protein that is cleaved at residue 115 to leave a mature product of 283 amino acids. Analysis of the signal peptidase cleavage site predicted by the program, SignalP (Nielsen et al. 1997), suggests that the pre-sequence consists of the first 19 amino acids, and is therefore, followed by a 96-amino acid pro-sequence, which precedes the cleavage site to produce the N-terminus of the native protein. The predicted mass of the mature protein is 28,413 Da which is somewhat less than its estimation of 35 kDa from SDS-PAGE. The N-terminal sequence of the 35-kDa band excised from the gel, however, confirms that the band represents a post processed product from which the 96 amino acid pro-sequence has been cleaved. Protein sequence alignments with related serine proteases indicate the key amino acids involved in the catalytic mechanism are located at residues Asp42, His76 and Ser228 (Fig. 3).

Recombinant enzyme expression

We removed the intronic sequences from *esp1* by splice overlap PCR to create a continuous reading frame flanked by restriction sites to enable cloning into suitable expression vectors. We were able to express functional EPg222 in the methylotrophic yeast *P. pastoris*. We have previously demonstrated that *P. pastoris* will support the efficient secretion of filamentous fungal proteins using their native

```
                 Asp (D)                    His (H)              Ser (S)
                   *                           *                    *
Epg222    TAGEGI VVYGVDTGIDIRHSDFE-----  TDGNGHGTHTASTAVGSKF ---KTLSGTSMAAPHVAGA
AAT09329  TAGEGI VIYGVDTGIDIGHADFG-----  TDGNGHGTHTASTAAGRKF ---KTLSGTSMAAPRVAGA
AAD25926  TAGEGV VFYGVDTGIDISHSDFG-----  TDGNGHGTHTASTAAGSKY ---KTLSGTSMAAPHVAGV
AAF23726  TAGEGV VFYGVDTGIDISHSDFG-----  TDGNGHGTHTASTAAGSKY ---KTLSGTSMAAPHVAGV
AAR11460  SAGEGI TVYSVDTGVDVNHEDFE--- DDRDGSGHGTHTSGTMVGKMY --SQSMSGTSMAAPHVAGL
AAS45673  SAGEGI TVYSVDTGVDVNHEDFE--- DDRDGSGHGTHTSGTMVGKEF --SQSMSGTSMAAPHVAGL
```

Fig. 3 Conserved regions around the three catalytic residues Asp(D)-His(H)-Ser(S). *Letters in white* represent the identical positions found in the six sequences, and *letters shaded in gray* represent the conservative amino acids. The conserved catalytic residues are marked with an *asterisk*

signal sequences (Crepin et al. 2003). We, therefore, inserted the EPg222 reading frame (including the prepro-encoding sequences) into the methanol-inducible expression vector, pPIC3.5 K, and used a linearized form of the DNA construct to transform *P. pastoris*. Two hundred transformants were selected by their ability to grow on histidine-deficient medium and His^+ transformants were assessed for their methanol utilization phenotype. Fifteen transformants exhibiting slow growth on methanol (Mut^S phenotype) were selected to screen for protein expression in small-scale cultures under methanol induction, along with a control culture transformed with the parental vector to assess background protein secretion levels. Culture supernatants of the *P. pastoris* transformants were analysed by SDS-PAGE for protein expression to reveal that three of them generated a major secreted protein migrating at 35 kDa that was absent in the vector control (Fig. 4). The native mature EPg222 protein running alongside the recombinant protein also migrates at 35 kDa in SDS-PAGE, which suggests that the *P. pastoris* product had undergone post-translational cleavage. It was notable that a band at 30 kDa was also evident similar to that observed for the native protein. Protein sequencing of tryptic peptides derived from these bands was performed to confirm expression of the EPg222 protein and to confirm the identity of the 30-kDa product. Identical protein sequences to those obtained with native protein confirmed the secretion of mature EPg222 from *P. pastoris* and the identity of the protein migrating at 30 kDa as an autolysis product. The tryptic peptide, AEWGTNTVDNDNTDGNGHGTHTAS TAVGSK, was present in the 35- and 30-kDa protein bands when excised from the gel after SDS-PAGE but the mature N-terminal peptide, ENVVQPNAPSWGLPR, and internal peptide, HSDFEGR, were only present in the 35-kDa band.

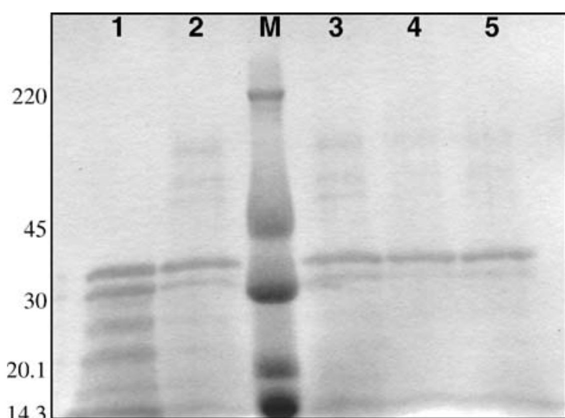

Fig. 4 SDS-PAGE analysis of supernatants from small-scale cultures of *P. pastoris* transformants expressing recombinant EPg222 stained with Coomassie brilliant blue. *M* corresponds to the molecular weight markers (kDa). *lane 1*: Native EPg222, *lane 2*: recombinant EPg222 from transformant 20. lanes 3–5: recombinant EPg222 from transformant 20, 33 and 47, respectively

These data indicate that the cleavage event that gave rise to the 30-kDa protein had taken place after residue 47 but before residue 56 within the sequence HSDFEGR. The loss of the N-terminal sequence up to residue 55 would result in a mass reduction 5.9 kDa, which is consistent with the difference in migration observed on SDS-PAGE.

Activity of recombinant EPg222

To confirm that the proteins were enzymatically active, culture supernatants were assayed for activity against BSA and pork myofibrillar and sarcoplasmic proteins. BSA, myofibrillar proteins and sarcoplasmic proteins extract were completely hydrolyzed within 4 h of incubation by recombinant EPg222 enzyme (Fig. 5). Based on the protein concentration of the recombinant EPg222 enzyme, both the turnover and utilization of complex protein substrates were similar to that observed for the native EPg222 protein.

Kinetic constants (k_{cat}, K_m,) were calculated from initial rates of the recombinant EPg222 against the synthetic substrate BAPNA. The values for K_m and K_{cat} were 2.10±0.0708 mM and 5.54±0.22 s^{-1} respectively. The kinetic constants were also similar to those recorded for the native enzyme (Benito et al. 2002).

Discussion

Several proteolytic enzymes from fungi have been cloned and characterized but this represents the first derived from dry-cured meat products. Sequence similarities between the deduced amino acid sequence of mature EPg222 and other serine proteases from *Penicillum* species revealed sequence identities in the range of 92–81% to the alkaline serine proteases of *P. nordicum* (AAT65816) and *P. chrysogenum* (AAM33821, AAT09329, AF321100, AAD25926 and AAF23726), respectively. The amino acid residues Asp, His and Ser that form the functional catalytic triad in the subtilisin-like serine proteinase family are generally located in conserved sequence environments (Estell et al. 1986; Kraut 1977). These sequences are recognizable in mature EPg222, which are centred on residues Asp42, His76 and Ser228 of the mature polypeptide.

In common with other fungal serine proteases, the EPg222 precursor contains three distinct domains, a 19-amino acid putative signal peptide, a 96-amino acid propolypeptide and a 283-amino acid mature polypeptide. The signal peptide cleavage site is predicted to lie between Ala19 and Gly20 (Nielsen et al. 1997). The 19-amino acid segment is a secretory signal sequence with a high content of hydrophobic amino acids that is expected to direct transport of the polypeptide chain across the ER membrane into the ER lumen (Hattori et al. 1987). Like many members of the

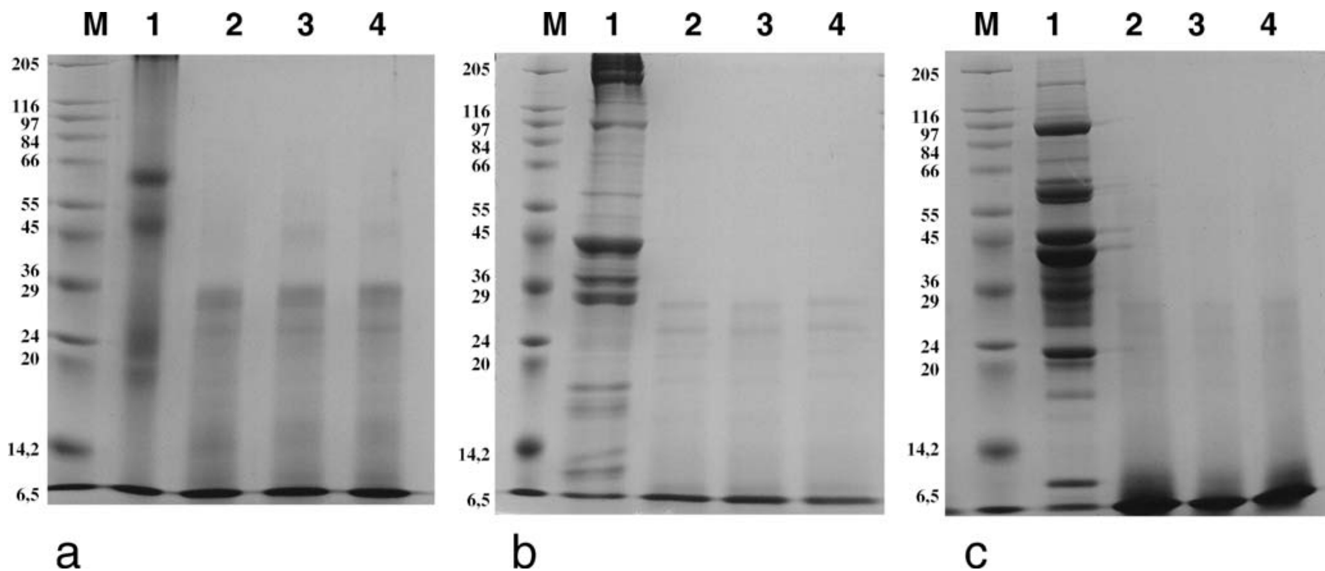

Fig. 5 SDS-PAGE analysis to show the hydrolysis of **a** bovine serum albumin (BSA), **b** pork myofibrillar proteins and **c** pork sarcoplasmic proteins by EPg222 expressed in *P. pastoris*. The gel was stained with Coomassie brilliant blue where *M* corresponds to the molecular weight markers (kDa). *lane 1*: undigested protein control. *lanes 2–4*: hydrolysis of the proteins for 4 h with recombinant EPg222 from *P. pastoris* transformants 20, 33 and 47, respectively

serine protease family, EPg222 is synthesized as precursor structure with a 96-amino acid N-terminal propeptide sequence (Ikemura et al. 1987). These propeptides have been proposed to act as intramolecular chaperones (Sienzen et al. 1995). However, a clear biological role is to prevent the protease from autolysis until it is processed (Shinde and Inouye 1994), as removal of this prosequence from the catalytic domain is necessary to generate the active enzyme (Sienzen and Leunissen 1997). It has been demonstrated that the uncleaved prosequence interacts with the active site of the catalytic domain, and thereby, inhibits its activity (Sienzen and Leunissen 1997). The prosequence may be removed from the catalytic domains of these enzymes by autoprocessing or through the action of another protease. EPg222 is clearly subject to autolysis as evidenced by the appearance of the protein migrating at 30 kDa upon SDS-PAGE from native and recombinant protein preparations. Analysis of the 30-kDa band has shown that it shares peptide sequences with the mature active protein. Analysis of these tryptic peptide fragments indicate that the cleavage position is between residues 48 and 55 of mature EPg222 that would result in the loss of Asp42, which is putatively a key component of the catalytic mechanism. This finding is consistent with the observation that the protein migrating at 30 kDa can be separated by ultrafiltration and demonstrated to retain no catalytic activity. Collectively, these observations explain the autolytic action of EPg222 and the loss in proteolytic activity of purified protein preparations noted previously (Benito et al. 2002).

It has been established that the *P. chrysogenum* alkaline serine proteases Pen c 1 (Su et al. 1999), Pen ch13 (Chou et al. 2002) and Pen ch18 (Shen et al. 2003) are allergens that provoke extrinsic bronchial asthma associated with airborne fungal spores. The immunoglobulin E-binding epitopes of Pen ch13 have recently been identified using serum from patients with clinical histories of bronchial asthma. The immuno-dominant epitopes could be discriminated on the basis of IgE binding to overlapping synthetic peptides derived from the sequence of Pen ch13 (Lai et al. 2004). As cases of allergy associated with the ingestion of dry-fermented sausage have recently been reported (Morriset et al. 2003) and incidences of allergic response have been reported in workers making salami (Marchisio et al. 1999), we decided to compare the epitope sequences responsible for IgE recognition of Pen ch13 with those found in EPg222. The overall amino acid identity between mature Pen ch13 and EPg222 was 90%, but in all four IgE recognition regions, EPg222 contains amino acid substitutions. Most notably, the octapeptide GHADFGGR (residues 48–55 of Pen ch13) is predicted to lie in a loop structure at the protein surface and leads to the highest intensity and frequency of patient response. The corresponding amino acid sequence in EPg222 contains three substitutions to present the peptide RHSDFEGR. Specific changes in the immunogenic domains of EPg222 coupled with the possibility of precise dosing through its liquid application to meat products could make the use of EPg222 an attractive alternative to relatively uncontrolled mould growth and the occupational exposure of workers to fungal spores.

It is clear from the present study that *P. pastoris* can be used as a host to express Epg222 protease from *P. chrysogenum*. The expression system used methanol-induc-

ible *AOX1* promoter to drive high yields of the enzyme. The biochemical properties of recombinant and native EPg222 are similar exhibiting activity against a wide range of protein substrates including bovine serum albumin, pork myofibrillar and sarcoplasmic protein extracts. These proteins are hydrolyzed during the ripening of dry-cured meats, which lead to the development of compounds responsible for flavour and aroma that are characteristic of these products (Martín et al. 2003, 2004).

In conclusion, we have demonstrated EPg222 constitutes a novel enzyme from *P. chrysogenum* and report the development of an expression system capable of delivering commercially viable quantities of EPg222 with activities that are indistinguishable from the native enzyme. The enzyme may be used to improve the sensorial characteristics of dry-cured meat products and accelerate the ripening process. Control of this process without the hazards associated with mould growth would provide a significant economic advantage to production of high-value meat products.

Acknowledgements The authors would like to thank Dr. Susan Liddell (Division of Agricultural and Environmental Sciences, University of Nottingham) for performing the amino acid sequence determinations. M. J. Benito was the recipient of a postdoctoral grant from the Spanish Ministerio de Educación y Ciencia. This work was supported by a grant from the Consejería de Educación, Ciencia y Tecnología de la Junta de Extremadura (2PR04C022) and Spanish Comisión Interministerial de Ciencia y Tecnología (AGL01-0804).

References

Becker DM, Guarente L (1991) High-efficiency transformation of yeast by electroporation. In: Guthrie C, Fink GR (eds) Guide to yeast genetics and molecular biology, methods in enzymology, vol. 194. Academic, London, UK, pp 182–186

Benito MJ, Rodríguez M, Núñez F, Asensio MA, Bermúdez ME, Córdoba JJ (2002) Purification and characterization of an extracellular protease from *Penicillium chrysogenum* Pg222 active against meat proteins. Appl Environ Microb 68:3532–3536

Benito MJ, Córdoba JJ, Alonso M, Asensio MA, Núñez F (2003a) Hydrolytic activity of *Penicillium chrysogenum* Pg222 on pork myofibrillar proteins. Int J Food Microbiol 89:155–161

Benito MJ, Rodríguez MM, Sosa MJ, Martín A, Córdoba JJ (2003b) Effect of Protease EPg222 obtained from *Penicillium chrysogenum* isolated from dry-cured ham in pieces of pork loins. J Agric Food Chem 51:106–111

Benito MJ, Rodríguez MM, Acosta R, Córdoba JJ (2003c) Effect of the fungal extracellular protease EPg222 on texture of whole pieces of pork loin. Meat Sci 65:877–884

Benito MJ, Rodríguez MM, Martín A, Aranda E, Córdoba JJ (2004) Effect of the fungal protease EPg222 on the sensory characteristics of dry fermented sausage “salchichón” ripened with commercial starter cultures. Meat Sci 67:497–505

Benito MJ, Rodríguez MM, Córdoba MG, Andrade MJ, Córdoba JJ (2005a) Effect of the fungal protease EPg222 on proteolysis and texture in the dry fermented sausage ‘salchichón’. J Sci Food Agric 85:273–280

Benito MJ, Núñez F, Córdoba MG, Martín A, Córdoba JJ (2005b) Generation of non-protein nitrogen and volatile compounds by *Penicillium chrysogenum* Pg222 activity on pork myofibrillar proteins. Food Microbiol 22:513–519

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

Bruna JM, Ordóñez JA, Fernández M, Herranz B, de la Hoz L (2001) Microbial and physico-chemical changes during the ripening of dry fermented sausages superficially inoculated with or having added an intracellular cell-free extract of *Penicillium aurantiogriseum*. Meat Sci 59:87–96

Chou H, Lai HY, Tam MF, Chou MY, Wang SR, Han SH, Shen HD (2002) cDNA cloning, biological and immunological characterization of the alkaline serine protease major allergen from *Penicillium chrysogenum*. Int Arch Allergy Immunol 127:15–26

Chow LP, Chiou SH, Hsiao MC, Yu CJ, Chiang BL (2000) Characterization of Pen n13, a major allergen from the mold *Penicillium notatum*. Biochem Biophys Res Commun 261: 669–675

Crepin VF, Faulds CB, Connerton IF (2003) Production and characterization of the *Talaromyces stipitatus* feruloyl esterase FAEC in *Pichia pastoris*: identification of the nucleophilic serine. Protein Expr Purif 29:176–184

Estell DA, Graycar TP, Miller JV, Powers DB, Burnier JP, Ng PG, Wells JA (1986) Probing steric and hydrophobic effects on enzyme–substrate interactions by protein engineering. Science 233:659–663

Goh L, Kuo I, Luo S, Chua KY, White M (2001) Production and purification of recombinant *Blomia tropicalis* group 5 allergen from *Pichia pastoris* culture. Biotechnol Lett 23:661–665

Hattori T, Ichihara S, Nakamura K (1987) Processing of a plant vacuolar protein in precursor in vitro. Eur J Biochem 166:533–538

Higgins DR, Cregg JM (1998) Introduction to *Pichia pastoris*. In: Higgins DR, Cregg JM (eds) *Pichia* protocols. methods in molecular biology series, vol 103. Humana, New York

Horton RM, Cai Z, Ho SN, Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques 8:528–535

Ikemura H, Takagi H, Inouye M (1987) Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli*. J Biol Chem 262:7859–7864

Kraut J (1977) Serine protease: structure and mechanism of catalysis. Annu Rev Biochem 46:331–358

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Lai H-Y, Tamw MF, Chou H, Leez S-S, Tai H-Y, Shen H-D (2004) Molecular and structural analysis of immunoglobulin E-binding epitopes of Pen ch 13, an alkaline serine protease major allergen from *Penicillium chrysogenum*. Clin Exp Allergy 34:1926–1933

Marchisio VF, Sulotto F, Botta GC, Chiesa A, Airaudi D, Anastasi A (1999) Aerobiological analysis in a salami factory: a possible case of extrinsic allergic alveolitis by *Penicillium camembertii*. Med Mycol 37:285–289

Martín A, Córdoba JJ, Benito MJ, Aranda E, Asensio MA (2003) Effect of *Penicillium chrysogenum* and *Debaryomyces hansenii* on the volatile compounds during controlled ripening of pork loins. Inter J Food Microbiol 84:327–338

Martín A, Córdoba JJ, Núñez F, Benito MJ, Asensio MA (2004) Contribution of a selected fungal population to proteolysis on dry-cured ham. Inter J Food Microbiol 94:55–66

Morriset M, Parisot L, Kanny G, Moneret-Vautrin DA (2003) Food allergy to moulds: two cases observed after dry fermented sausage ingestion. Allergy 38:1203–1204

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Analysis identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6

Rodríguez MM, Núñez F, Córdoba JJ, Bermúdez ME, Asensio MA (1998) Evaluation of proteolytic activity of microorganisms isolated from dry-cured ham. J Appl Microbiol 85:905–912

Rodriguez E, Mullaney EJ, Lei XG (2000) Expression of the *Aspergillus fumigatus* phytase gene in *Pichia pastoris* and characterization of the recombinant enzyme. Biochem Biophys Res Commun 268:373–378

Shen H-D, Chou H, Tam MF, Chang C-Y, Lai H-Y, Wang S-R (2003) Molecular and immunological characterization of Pen ch 18, the vacuolar serine protease major allergen of *Penicillium chrysogenum*. Allergy 58:992–1002

Shinde U, Inouye M (1994) The structural and functional organization of intramolecular chaperones: the N-terminal propeptides which mediate protein folding. J Biochem 115:629–636

Sienzen RJ, Leunissen JAM (1997) Subtilases: the superfamily of subtilisin-like serine proteases. Protein Sci 6:501–523

Sienzen RJ, Leunissen JAM, Shinde U (1995) Homology analysis of the propeptides of subtilisin-like serine proteases (subtilases). In: Shinde U, Inouye M (eds) Intramolecular chaperones and protein folding. RG Landes Company, Austin, Texas, pp 233–256

Su NY, Yu CJ, Shen HD, Pan FM, Chow LP (1999) Pen c 1, a novel enzymic allergen protein from *Penicillium citrinum*. Purification, characterization, cloning and expression. Eur J Biochem 261:115–123

Sun J, Bottomley SP, Kumar S, Bird PI (1997) Recombinant caspase-3 expressed in *Pichia pastoris* is fully activated and kinetically indistinguishable from the native enzyme. Biochem Biophys Res Commun 238:920–924