The Journal of Microbiology (2011) Vol. 49, No. 1, pp. 121-129
DOI 10.1007/s12275-011-9355-6

Cloning, Expression, and Characterization of Serine Protease from Thermophilic Fungus *Thermoascus aurantiacus* var. *levisporus*

An-Na Li†, Chen Xie†, Jie Zhang, Jia Zhang, and Duo-Chuan Li*

Department of Environmental Biology, Shandong Agricultural University, Tai'an, Shandong, 271018, P. R. China

(Received November 17, 2009 / Accepted September 2, 2010)

The serine protease gene from a thermophilic fungus *Thermoascus aurantiacus* var. *levisporus*, was cloned, sequenced, and expressed in *Pichia pastoris* and the recombinant protein was characterized. The full-length cDNA of 2,592 bp contains an ORF of 1,482 bp encoding 494 amino acids. Sequence analysis of the deduced amino acid sequence revealed high homology with subtilisin serine proteases. The putative enzyme contained catalytic domain with active sites formed by three residues of Asp183, His215, and Ser384. The molecular mass of the recombinant enzyme was estimated to be 59.1 kDa after overexpression in *P. pastoris*. The activity of recombinant protein was 115.58 U/mg. The protease exhibited its maximal activity at 50°C and pH 8.0 and kept thermostable at 60°C, and retained 60% activity after 60 min at 70°C. The protease activity was found to be inhibited by PMSF, but not by DTT or EDTA. The enzyme has broad substrate specificity such as gelatin, casein and pure milk, and exhibiting highest activity towards casein.

Keywords: *Thermoascus aurantiacus* var. *levisporus*, serine protease, cloning, expression, *Pichia pastoris*

Proteases or proteolytic enzymes catalyse the cleavage of peptide bonds in proteins. Proteases are one of the most important groups of industrial enzyme, they are one of the three largest group of industrial enzymes, accounting for about 60% of the total worldwide sale of enzymes (Niehaus *et al.*, 1999). They have been widely used in the detergent, beer, meat, leather and dairy industries for a long time (Kumar *et al*., 2005). Microbial proteases constitute approximately 40% of the total worldwide production of enzymes and bacteria belonging to the genus Bacillus produce most commercial proteases used today (Genckal and Tari, 2006; Shikha *et al*., 2007). These enzymes are also involved in essential biological processes like blood clotting, controlled cell death, and tissue differentiation.

In recent research, the potential use of fungal proteases is being increasingly realized (Gupta *et al*., 2002), since they offer distinct advantages over the bacterial enzymes in terms of easing the downstream processing (Laxman *et al*., 2005). In industry, a major requirement for commercial protease is thermal stability, and thermal denaturation is a common cause of enzyme inactivation. Furthermore thermostable proteases are more effective reduce the incidence of microbial contamination from mesophilic organisms during the protease production (Chen *et al*., 2004).

Thermophilic fungi belongs to extremophiles which can grow at a high temperature, therefore, their extracellular enzymes might be more active and stable under extreme conditions. Studies on thermostable enzymes from thermophilic fungi have been paid more attention with the expectation to produce thermostable enzymes (Maheshwari *et al*., 2000). Numerous protease-encoding genes have been cloned and characterized from mesophilic fungi, such as *Penicillium oxalicum* (Shen *et al.*, 2001), *Trichoderma harzianum* (Liu and Yang, 2007), *Aspergillus niger* (Gomi *et al*., 1993; Pel *et al.*, 2007), *Aspergillus fumigatus* (Vickers *et al*., 2007). However studies on the proteases from thermophilic fungi are very limited, only few proteases gene have been cloned, such as, protease from thermophilic fungi *Thermomonospora fusca* YX (TfpA), *Chaetomium thermophilum* and have been expressed in *P. pastoris*, respectively (Kim and Lei, 2005; Li and Li, 2009).

T. aurantiacus var. *levisporus* is a widely distributed soil inhibiting fungus with the potential of thermostable enzymes production. It is reported that in a medium containing casein, *T. aurantiacus* var. *levisporus* also produced thermostable protease (Marcy *et al*., 1984; Merheb *et al*., 2007), however, the serine protease gene has not been cloned yet. In the present study, we report the cloning and sequencing of the new protease gene from *T. aurantiacus* var. *levisporus* and its expression in *P. pastoris*. Some characterizations of the recombinant enzyme are also shown.

Materials and Methods

Strains, plasmids, media, and culture growth

T. aurantiacus var. *levisporus* was isolated and preserved by our laboratory, and used as an mRNA donor. *Escherichia coli* DH5α and JM109 were used for gene cloning. Plasmid pMD18-T was obtained from TaKaRa (Japan). The expression vector pPIC9K was purchased from Invitrogen (USA). *P. pastoris* GS115 was used as a host for protease expression. To induce the protease gene expression, actively growing mycelium of *T. aurantiacus* var. *levisporus* was transferred from a potato dextrose agar (PDA) plate to the inducing medium containing the following (g/L): 4.0 g of casein, 4.0 g of agar, 4.0 g

† These authors contributed equally to this work.
* For correspondence. E-mail: lidc20@sdau.edu.cn; Tel: +86-538-824-9071; Fax: +86-538-822-6399

Table 1. The primers for PCR used in the study

Primer name	Sequence (5'→3')	Purpose
PS1	5'-GGYCAYGGHACTCACTGC-3'	First partial fragment
PA1	5'-GGRGTGGCCATRGARGTACC-3'	First partial fragment
PS2	5'-GGYTACTCSGGACACTTC-3'	Second partial fragment
PA2	5'-TGAGCCTGAGCCGCCCATTC-3'	Second partial fragment
P3-1	5'-CTCGGTGGAGGCAAATC-3'	3'-RACE
P3-2	5'-GGCATCCACTTCGCTGT-3'	3'-RACE
M13PrimerM4	5'-GTTTTCCCAGTCACGAC-3'	3'-RACE
P5-1	5'-ACCGGCACCAACATCGACCATG-3'	5'-TAIL
P5-2	5'-ACGCTGCCAATGACGGTGAGG-3'	5'-TAIL
P5-3	5'-AGATTCGCAGGCATCCGGACG-3'	5'-TAIL
AD4	5'-TGWGNAGSANCASAGA-3'	5'-TAIL
PQ5	5'-CGTCTCTCCAACTTCCATCGTT-3'	Full length cDNA
PQ3	5'-GGTCGTGACATGCTCCAATCAT-3'	Full length cDNA
PU	5'-CCGGAATTCTCACCCGTCGTCGTAGACTCGAT-3'	Expression
PD	5'-TTGCGGCCGCCTACGCAGTAACAGCATCCTTGAT-3'	Expression

of yeast extract, 1.0 g of $K_2HPO_4{\cdot}3H_2O$, and 0.5 g of $MgSO_4{\cdot}7H_2O$ dissolved in 1 L (distilled and tap water 3:1). The media used for culture of *E. coli* and *P. pastoris* GS115 were referred to standardized methods (Sambrook *et al*., 1989) or *Pichia* Expression kit (Invitrogen).

cDNA and genomic DNA cloning

After two days growth at 50℃ on inducing medium, the total RNA was obtained from the mycelia of *T. aurantiacus* var. *levisporus* using TRIZOL reagent (Gibco BRL) following the manufacture's instructions. cDNA template of *T. aurantiacus* var. *levisporus* was obtained by the procedures of reverse transcription polymerase chain reaction (RT-PCR). Genomic DNA isolated from *T. aurantiacus* var. *levisporus* mycelia growing on PDA plate in the late growth phase according to Lebaron *et al.* (1998). All primers used for generating *T. aurantiacus* var. *levisporus* protease are shown in Table 1.

The first partial fragment sequence of protease gene was performed by RT-PCR using the degenerated primer PS1 and PA1, synthesized based on the highly conserved amino acid sequences GHGTHV and GTSMASPH of other fungal serine proteases. The second partial fragment sequence of protease gene was obtained by primer PS2 and PA2 which were synthesized based on conserved amino acid sequences GYSGHF and the first partial gene sequence obtained above. The PCR amplified DNA fragment was cloned into a pMD18-T easy vector (TaKaRa) and sequenced by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) using an automated DNA sequencer (Applied Biosystems, Model 370A, Shanghai Bioasia Engineering, China).

RACE product of protease cDNA Rapid amplification of cDNA ends (RACE) and thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) were used to get the full length of the protease gene. The gene-specific primers for RACE-PCR and TAIL-PCR were designed based on the partial gene sequence obtained above. PCR for 3'-RACE were carried out with the primer M13M4 and P3-1. The primers for 5'-TAIL were random degenerate primer AD4 and three internal primers, P5-1, P5-2, and P5-3 (Liu and Huang, 1998; Terauchi and Kahl, 2000; Singer and Burke, 2003). The full-length cDNA and genomic DNA encoding *Tapro* of *T. aurantiacus* var. *levisporus* were generated using the sense primer PQ5 and anti-sense primer PQ3 designed from the sequence of the 5'-TAIL and 3'-RACE product respectively.

Nucleotide sequence accession number

The nucleotide sequence accession number in GenBank of the protease cDNA and DNA gene from *T. aurantiacus* var. *levisporus* is EU364816 and GU119907 respectively.

Analysis of nucleotide and amino acid sequence

In order to search for nucleotide and amino acid sequences similar to the *T. aurantiacus* var. *levisporus* protease, the NCBI BLAST program (http://www.ncbi.nlm.nih.gov) was used. The deduced amino acid sequences of protease cDNA were analyzed using ExPASy search program (http://au.expasy.org/tools/). The signal peptide was predicted by Signal P 3.0 program (http://www.cbs.dtu.dk/services/Signal P).

Construction of expression plasmid pPIC9K/*Tapro*

The gene *Tapro* encoding sequence was PCR-amplified using a 5' end primers: PU and a 3' end primer PD (Table 1, restriction enzyme *Eco*RI and *Not*I sites are underlined) with the full-length cDNA as a template. The amplified DNA fragment was digested with *Eco*RI and *Not*I and ligated to pPIC9K vector which is also digested by *Eco*RI and *Not*I. The recombinant plasmid was termed pPIC9K/*Tapro* and was subsequently confirmed by PCR, restriction digestion and sequencing to conform its integrity.

Transformation of *P. pastoris* GS115

Electrocompetent *P. pastoris* GS115 cells were prepared as described by Chen *et al.* (2007). A total of 10 μg of X*bal*I-linearized recombinant plasmid pPIC9K/T*apro* was transformed into *P. pastoris* strain GS115 by electroporation using an Eppendorf Electroporator 2510. 1 ml of ice-cold 1 mol sorbitol was added into the cells after the pulse immediately, and the tranformants were spread on MD plates and incubated at 30℃ until colonies appeared (2-3 days). Integration screening was performed by streaking single colonies on MD and MM plates. The multi-copy integrants were selected on YPD plates containing various concentrations of defined G418 sulfate (Amresco).

To verify the gene integrated into the *P. pastoris* genome and the determination of the phenotype, a colony from the YPD-G418 plates was picked up and grown in 5 ml minimal medium overnight and genomic DNA was isolated from numbers of transformants as described by Lee (1992). PCR amplifications were carried out with 2 μl of the genomic DNA and a pair of primers corresponding to α-factor sequencing primer (5'-TACTATTGCCAGC ATTGCTGC-3') and 3'*AOX1* sequencing primer (5'-GCAAATGGCATTCTGACATCC-3').

Expression and purification of *T. aurantiacus* var. *levisporus* protease

The transformants containing G418 were inoculated into 25 ml of BMGY medium and subsequently incubated for 16-18 h until to the log phase growth (A_{600nm}=2-6). The cells were harvested, washed and suspended in 50 ml of BMMY medium. The expression was induced in BMMY medium at 30°C for 7 days by adding methanol daily to a final concentration of 1%. Aliquots of each culture were sampled at 24 h intervals and protein expression was analysed by SDS-PAGE (Laemmli, 1970).

To purify the expressed protease, cells were removed from the culture by centrifugation at 10,000×g for 10 min at 4°C. Solid ammonium sulfate was added to the supernatant to reach 90% saturation. After deposition for 12 h at 4°C, the precipitate was collected by centrifugation at 8,000×g for 15 min, dissolved in 50 mmol Tris-HCl (pH 8.0) and dialyzed overnight against three changes of the same buffer. The dialyzed sample was clarified by centrifugation and then loaded onto an anion exchange column of DEAE Sepharose column (1×20 cm) equilibrated with 50 mmol Tris-HCl (pH 8.0). Bound proteins were eluted using 200 ml of equilibration buffer with a linear gradient from 0 to 0.3 mol NaCl at a flow rate of 40 ml/h. Purity of the recombinant protease was determined by SDS-PAGE (Laemmli, 1970). The protease activity was detected by the method of active staining described by Li *et al*. (2007a). The slab gel with 1% gelatin after electrophoresis was washed in 50 mmol Tris-HCl (pH 8.0), containing 5% (v/v) Triton X-100, twice for 15 min at 4°C, followed by washing in the same buffer without Triton X-100 for 15 min to remove SDS. The gel was then incubated in 50 mmol Tris-HCl (pH 8.0) for 12 h at 50°C to allow the degradation of the gelatin, and stained and destained in the same solution with SDS-PAGE. The activity band was observed as a clear colorless area depleted of gelatin from the gel against the blue background.

Protease activity assay and protein characterization

The protease activity was determined by Rick's method (1974) with some modifications. The enzyme (0.1 ml, 1 μg of protein) was added to 0.2 ml of 0.5% casein in 0.2 mol Tris-HCl (pH 8.0) buffer and incubated at 50°C for 30 min, 0.2 ml 10% TCA (Trichloroacetic acid) was then added to terminate the reaction. After staying at room temperature for 30 min, the solution was filtered. To 0.5 ml of filtrate, 2.5 ml of water was added, the optical density (OD) at 280 nm was determined spectrophotometrically in UV-160A (Shimadzu). Sample prepared by adding 10% TCA to the enzyme before the addition of casein was used as a blank. One unit of protease activity was defined as the amount of enzyme producing 1 μg of tyrosine per minute under the assay conditions. Protein content was determined by the method of Bradford (1976) with crystalline bovine serum albumin as the standard.

The optimum temperature was measured in 0.2 mol Tris-HCl (pH 8.0) at various temperatures from 10 to 90°C. The activity was estimated as a percentage of the maximum. The thermal stability of protease was examined in the range of 50-90°C, and samples were removed at fixed time intervals and allowed to cool on ice before residual activities were determined under standard conditions. The experiment was conducted in triplicate.

The optimum reaction pH of purified protease was measured under different pH conditions. The buffers of HCl-KCl (pH 3.0-4.0), CH_3COOH-CH_3COONa (pH 4.0-6.0), NaH_2PO_4-Na_2HPO_4 (pH 6.0-7.0), Tris-HCl (pH 7.0-10.0), and Na_2HPO_4-NaOH (pH 10.0-12.0) at 0.2 mol were used. Activity was estimated as a percentage of the maximum. pH stability was investigated by measuring the residual activity of the enzyme after it had been kept for 1 h in various pH (3.0-12.0) conditions at room temperature, then adjusting the pH to 8.0, and the residual protease activities were tested under standard conditions. The experiment was conducted in triplicate. For the inhibition and denaturant assay, EDTA (10 mmol), PMSF (10 mmol), DTT (10 mmol) were used, respectively.

Either a colorimetric method or SDS-PAGE method was used to measure protease activity with different protein substrates. For a colorimetric method, protease were reacted with 0.1 ml different substrate, 1% bovine serum albumin (BSA), 1% bovine hemoglobin (BHb), 1% gelatin, 1% casein and 1% pure milk. Activity was measured under standard assay condition. For SDS-PAGE method, the reaction mixture consisted of 10 μl of enzyme sample and 10 μl of 1% different substrate dissolved in 50 μl 0.2 mol Tris-HCl buffer (pH 8.0). After incubation at 50°C for 30 min, the mixture was detected by SDS-PAGE. Incubated substrate without protease was also detected as a control.

Results

Cloning aanalyzing *T. aurantiacus* var. *levisporus* protease

Using the total RNA isolated from *T. aurantiacus* var. *levisporus*, 515 bp and 476 bp internal fragments of the gene *Tapro* were firstly amplified by RT-PCR. Sequence analysis using BLAST revealed that the two fragments are part of the serine protease gene, and the overlap of them is 190 bp. The full-length cDNA was obtained by RACE-PCR and TAIL-PCR. The resulting fragments of 3' and 5' were 840 bp and 875 bp. The full-length cDNA was generated with a length of 1,599 bp. Protease-encoding sequence on the genomic DNA was also amplified and sequenced. The gene- encoding sequence on the genomic DNA spans 1711 bp with two introns of 55 bp and 57 bp. The *T. aurantiacus* var. *levisporus* protease cDNA sequence analysis resulted in an open reading frame of 1,482 bp encoding 494 amino acids. There is a 5' end noncoding sequence of 85 bp and a 3' end noncoding sequence of 240 bp followed by a 15 bp polyA tail (Fig. 1). Sequence of protease gene *Tapro* was aligned with the known sequences of protease genes from other fungi using Nucleotide-nucleotide BLAST (blastn) at http://www.ncbi.nlm.nih.gov/Blast. It was found that the sequence of *Tapro* gene had high similarity to that from *Aspergillus niger* (XM 001391433) and *Penicillium marneffei* (XM 002147391). Similar to other subtilisins reported previously (Catara *et al.*, 2003; Shi *et al*., 2008; Li and Li, 2009), *T. aurantiacus* var. *levisporus* protease gene *Tapro* encodes a putative enzyme that consists of a putative catalytic domain with an active sites formed by three residues of Asp183, His215, and Ser384. Many extracellular subtilisin family proteases are translated as the prepro form containing

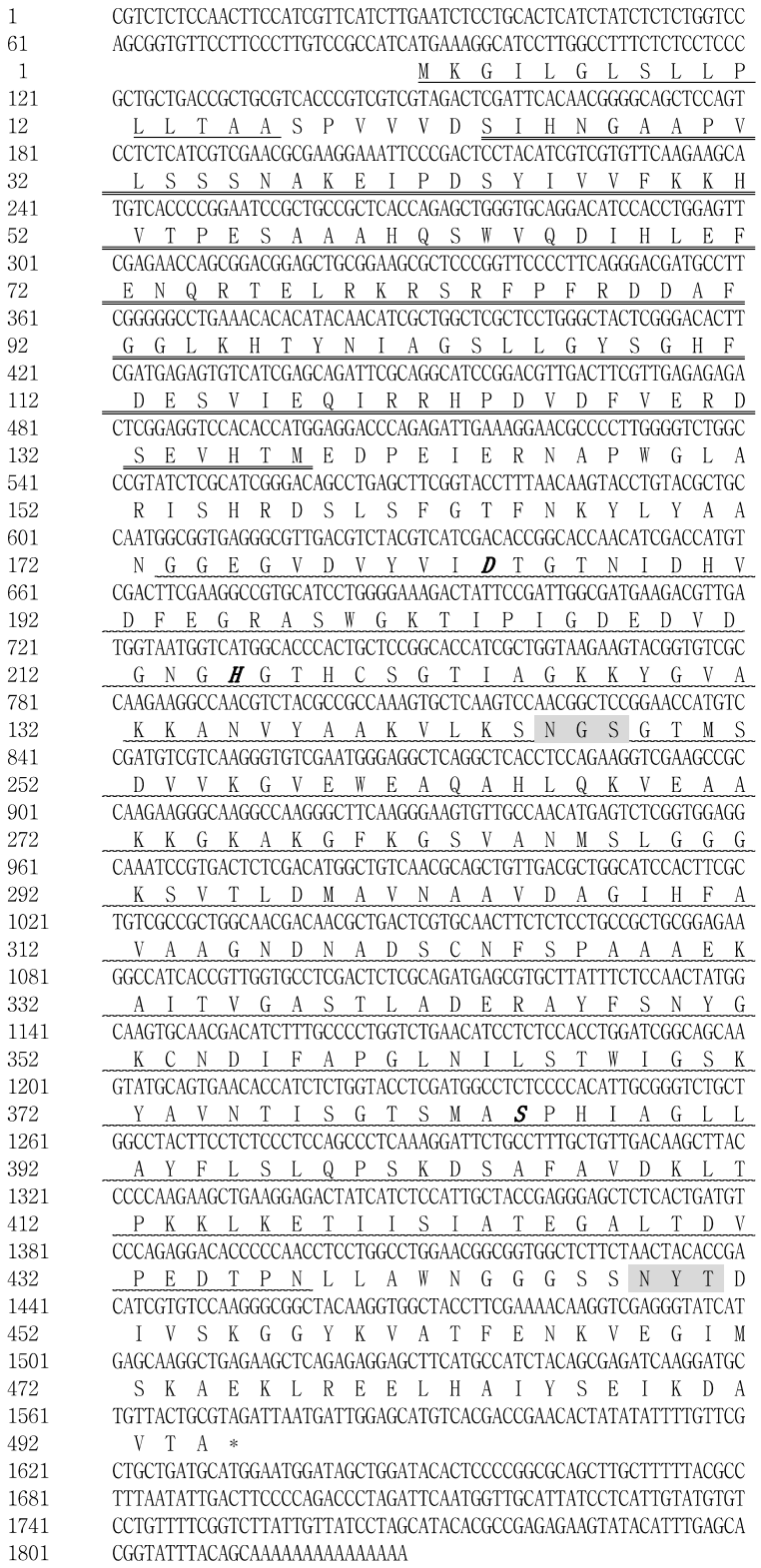

Fig. 1. Nucleotide and deduced amino acid sequence of protease gene *Tapro* (GenBank accession no. ABY71274). The nucleotide and/or amino acid sequence is numbered on the left. The upper lines show the nucleotide sequence with two introns (in small letter) and a CAAT box (boxed). The putative signal peptide, prepeptide, and the catalytic domain are underlined, double and wave underlined, respectively. The potential glycosylation sites are indicated by black shadow, and the three conservative catalytic triad (aspartic acid, histidine, and serine) are indicated in italic and boldface. The stop codon is indicated with an asterisk (*).

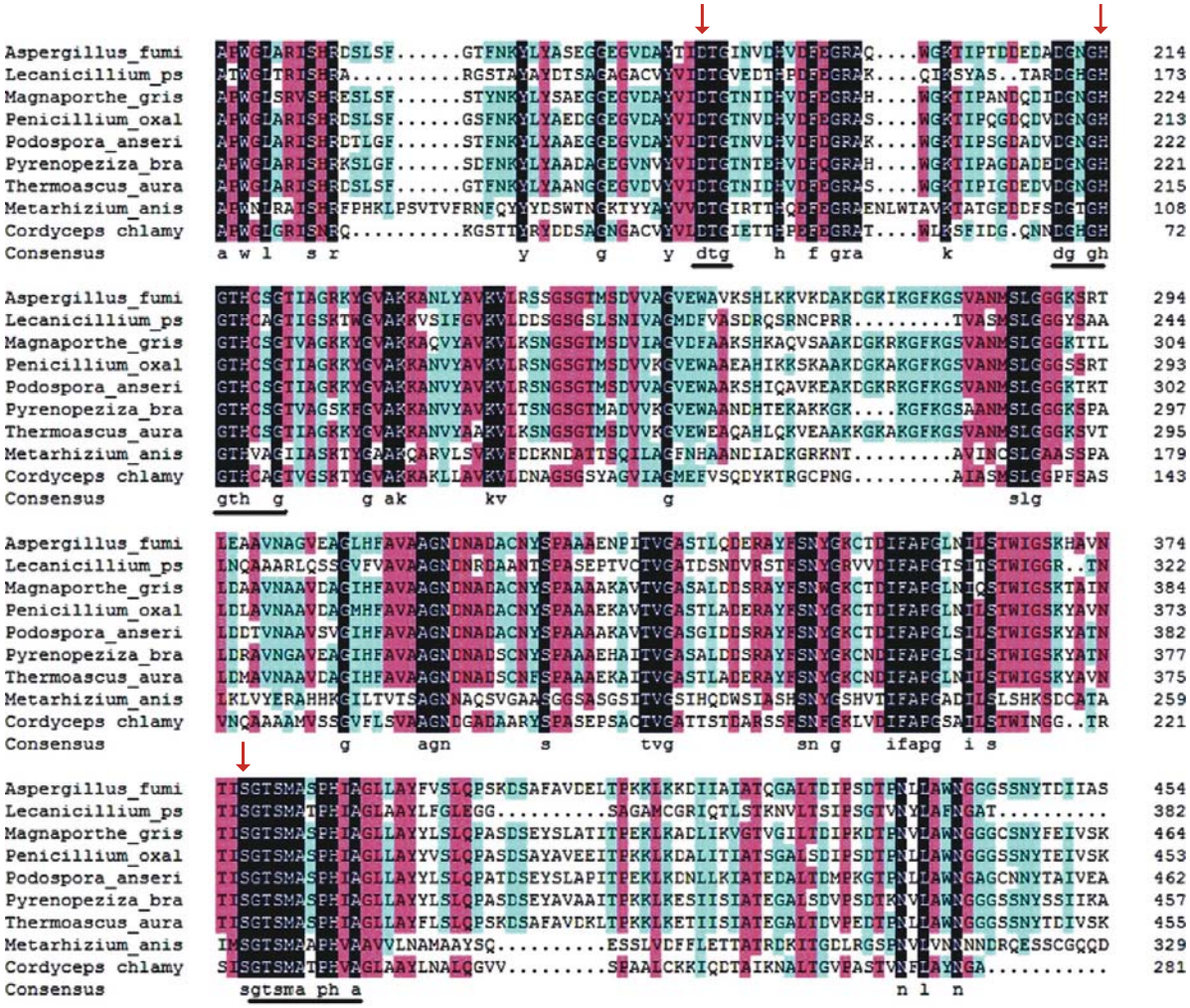

Fig. 2. Alignment of the deduced amino acids sequence with other fungal proteases. The conserved amino acids are shaded in black and the consensuses are shown. *Aspergillus fumigatus* (XP753718); *Lecanicillium psalliotae* (AY692148); *Magnaporthe grisea* (AB070268); *Penicillium oxalicum* (AF243425); *Podospora anserine* (AF047689); *Pyrenopeziza brassicae* (AJ319745); *Thermoascus aurantiacus* var. *levisporus* (EU364816); *Metarhizium anisopliae* var. *acridum* (AJ431637); *Cordyceps chlamydosporia* (AJ427459). The three canonical amino acid residues (D, H, and S) are indicated by arrows.

a signal sequence. (Choi *et al.*, 1999; Epstein and Wensink, 1998; Kitadokoro *et al.*, 1993; Li *et al.*, 2007b; Wei *et al.*, 2008). Signal P 3.0 program analysis revealed that it contained a putative signal peptide of 16 amino acids **MKGILGLSLLPLL TAA.** N-glycosylation sites of the protein were also analyzed at http://cbs.dtu.dk/services/NetNGlyc and the results indicated that there were two potential N-linked glycosylation sites in the protein (Fig. 1). Calculated molecular weight of the mature protease is 52.77 kDa with a PI of 5.91. Based on the analysis of the conserved sequence with other fungal protease by Blast, the protein can be assigned to serine protease family S8.

The sequence around the catalytic site of *T. aurantiacus* var. *levisporus* protease shows a high homology with that of *Bacillus* proteases of the subtilisin family (Rao *et al.*, 1998; Pan *et al.*, 2004; Ni *et al.*, 2008). Comparison of the deduced peptide sequence with other serine proteases revealed an identity of 79%, 79%, 68% to that of *Aspergillus fumigatus* (XP753718), *Penicillium oxalicum* (AF243425), and *Pyrenopeziza brassicae* (AJ319745), respectively. We found that three motifs DTG, DG-GHGTH and GTSMA-PH were highly conserved among the fungi protease (Fig. 2). Therefore, the results demonstrated that protease gene obtained in this study belonged to serine protease of the subtilisin family.

Expression and purification of recombinant protease

After transformed with plasmid Ppic9K/*Tapro*, 96 recombinants of *P. pastoris* GS115 were identified by screening on MD and MM plates. The PCR screening of the positive recombinants using α-factor and 3′*AOX1* sequencing primers produced a 1.7 kb fragment while the control yeast transformed with pPIC9K produced a 195 bp product. The cDNA was introduced into the yeast expression system and heterologous expression successfully under transcriptional control of

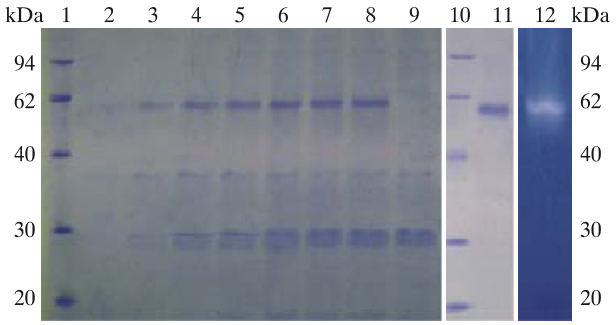

Fig. 3. SDS-PAGE analysis of the recombinant protease from *P. pastoris* and purified protein. Lanes: 1, 10: protein marker; 2-8: 1, 2, 3, 4, 5, 6 and 7 day, respectively; 9: *P. pastoris* GS115 control; 11: purified protease; 12: activity staining of the recombinant protease.

the *AOX1* promoter in *P. pastoris* GS115. Fig. 3 showed SDS-PAGE analysis of the induced recombinant protease from a transformant strain TA-18. After seven days induction, the transformant strain had the maximum activity of 115.58 U/mg and the expression level was 1.15 mg/ml. The expressed protein was easily purified with DEAE-Sepharose column protein purification system. Electrophoresis of the enzyme on SDS-PAGE gave a single band with a molecular weight of 59.1 kDa which is slightly bigger than that calculated from the amino acid sequence. The molecular weight variation of the protease expressed in *P. pastoris* might be a result of N-linked glycosylation. The possible explanation is that in many enzymes, glycosylation is involved in protein rigid structure formation that increases the thermostability (Chen *et al.*, 2007).

Effect of temperature and pH on protease activity and stability

The effect of temperature and pH on the enzyme activity and stability are shown in Fig. 4. The recombinant protease had its maximal activity at 50°C. The enzyme was thermostable at 50 and 60°C, and retained 60% activity after 60 min at 70°C. The half-life of the enzyme at 80°C was approximately 50 min. The protease is active in the pH range of 6.0-12.0, with an optimum at pH 8.0. The relative activity at pH10.0 was 83%. The protease was stable in a pH range of 5.0-12.0.

Inhibitors effect on the enzyme activity

Enzyme activity was measured in the presence of different enzyme inhibitors, in order to determine the nature of the protease. The expressed protease was found to be inhibited by PMSF, but not by DTT or EDTA (Fig. 5A), indicating

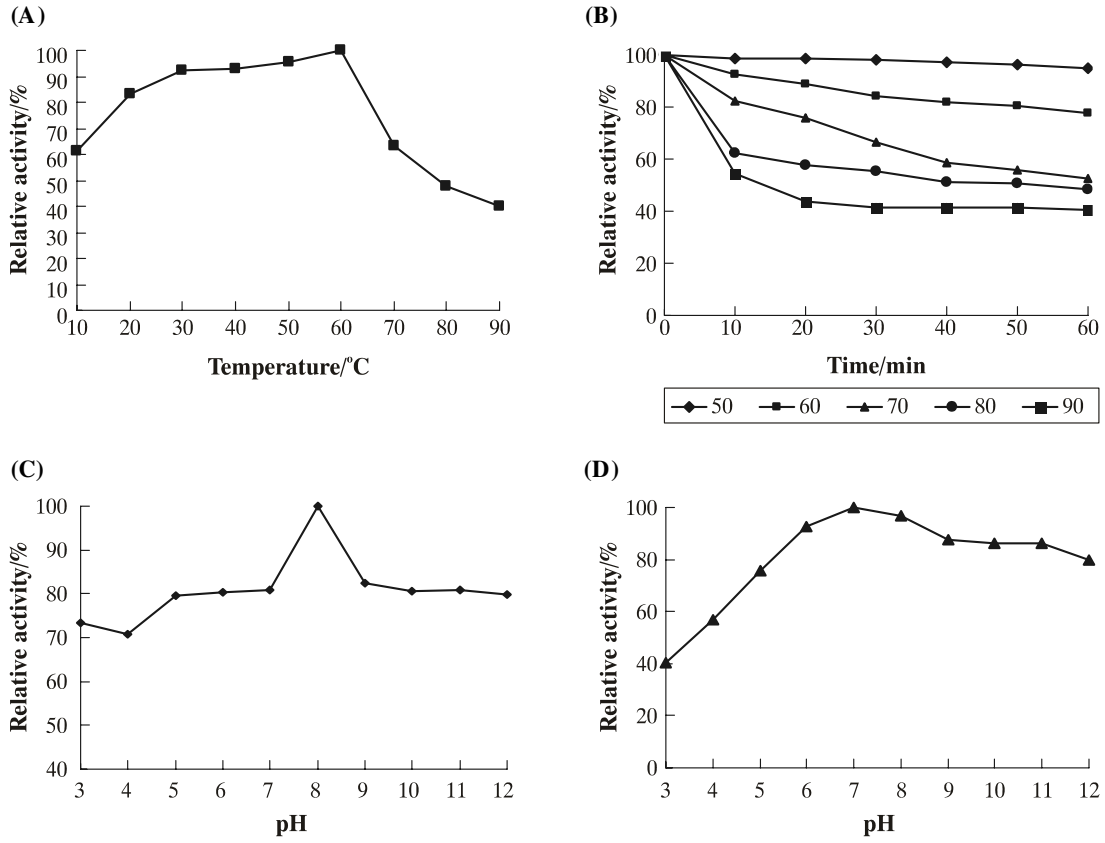

Fig. 4. Activity and stability of the purified protease. (A) The optimum reaction temperature of the purified protease, (B) the stability of the purified protease at different temperature, (C) the activity of the purified protease at different pH and (D) the stability of the purified protease at different pH.

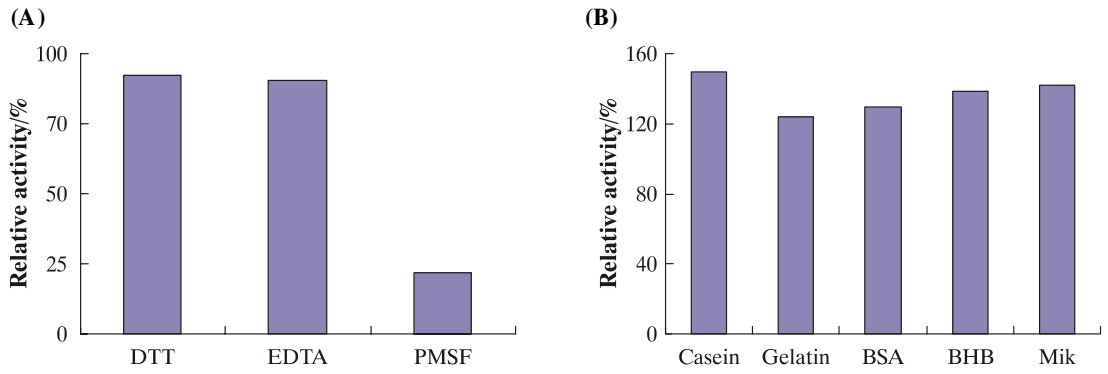

Fig. 5. Detection of protease from *T. aurantiacus* var. *levisporus* by inhibitors (A) and the hydrolysis of protein substrates (B).

that the recombinant protease gene was serine protease (Adinarayana and Ellaiah, 2003; Manavalan, 2007).

Hydrolysis of protein substrates

It has been reported that subtilisins are hydrolytic enzymes with broad substrate specificity. The substrate specificity of protease was found high when casein (100%) was used as substrate followed by BHb (92%), BSA (86%), and gelatin (82%), respectively (Fig. 5B). Similarly result was showed by SDS-PAGE, the recombinant protease have strong activity towards bovine serum albumin (BSA), bovine hemoglobin, gelatin, casein, and pure milk (Fig. 6). It appears as smearing in the lane of untreated gelatin and casein due to the heterogeneous molecular weight of the two substrates. The lane of protein substrate treated by the recombinant enzyme came to almost clear in BSA, gelatin, casein, and pure milk, showed broad substrate specificity of the enzyme. The enzyme exhibit highest activity towards casein, which is similar to protease reported before (Kobayashi *et al.*, 1995; Manavalan *et al.*, 2007).

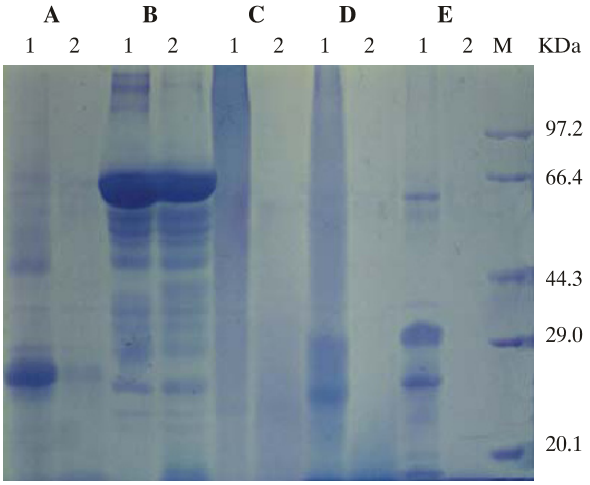

Fig. 6. SDS-PAGE analysis of the specificity the recombinant protease towards different protein substrates. Lines: 1, incubated proteins at 50°C for 30 min without enzyme; 2, incubated proteins with enzyme. A, Bovine hemoglobin; B, Bovine serum albumin; C, gelatin; D, casein; E, pure milk. Lane M, protein marker.

Discussion

Methylotrophic yeast *P. pastoris* has proven to offer significant advantages the high level production of certain proteins. It has been widely used to express proteins requiring post-translational modifications of higher eukaryotes and to secrete high levels of heterologous protein into the culture supernatant (Berrin *et al*., 2000; Wei *et al*., 2008). In our study the recombinant protease was expressed and secreted into the culture medium in an active form by the transformants of *P. pastoris.*

The protease gene *Tapro* was cloned from the thermophilic fungus *T. aurantiacus* var. *levisporus.* It was reported that nine residues usually regarded as sufficient to span a membrane (Watson, 1984), and the gene is consistent with the empirical rules of typical pre-secretory sequences. It was also reported that the mature extracellular alkaline protease was formed after N-glycosylation, signal peptide cleavage and other modifications of the propreprotein (Matoba *et al.*, 1988; Ni *et al.*, 2008). The pro-sequence is likely to be required for proper protein folding. The pro-peptide of subtilisin are known to function as molecular chaperones, assisting in folding of the mature peptidase. Deletion the complete removal of the pro-peptide or only 15 amino acids of pre-propeptide would result in the loss of activity of subtilisin (Ikemura *et al.*, 1987; Silen and Agard, 1989; Saul *et al*., 1996). We also cloned the cDNA of the protease without the first 8 amino acids of pro-peptide into the pPIC9K expression vector, but the activity of transformants is lower than that includes the full pro-peptide (data not shown). This suggests the importance of the pro-peptide for TAPRO post-translational processing.

In the activity assays, the expressed protease exhibited remarkable thermal stability and a high activity and it is one of the most thermostable proteases isolated from fungi (Shenolikar and Stevenson, 1982; Gaur *et al*., 1989; Morton *et al*., 2003; Wang *et al*., 2006; Merheb *et al*., 2007; Liu and Yang, 2007; Yang *et al*., 2007; Peña-Montes *et al*., 2008; Li and Li, 2009). Most of proteases reported before are not stable when treated at a high temperature or a wide range of pH, and their optimum reaction temperature is relatively lower, for example, the protease from *Stenotrophomonas maltophilia* was found to be stable up to 40°C for only 5 min (Miyaji *et al*., 2005), the purified protease of *Aspergillus nidulans* was

stable below 60°C, while the inactivation rate increase dramatically above this temperature (Peña-Montes *et al*., 2008), the protease of *T. aurantiacus* remained stable only up to 50°C (Merheb *et al*., 2007), the *Malbranchea pulchella* protease, named thermomycolase, was stable at 45°C when concentrated from the culture medium (Maheshwari *et al*., 2000). Though the optimum reaction temperature of this protease from *T. aurantiacus* var. *levisporus* is still lower than that of proteases isolated from thermophile bacteria (Manavalan *et al*., 2007) and hyperthermophilic archaea, most applications of enzymatic action occur at 40-60°C. The thermostable enzymes from thermophilic fungi may be more suitable for commercially applications than those from hyperthermophiles (Maheshwari *et al*., 2000; Li and Li, 2009).

Protease TAPRO showed its highest activity at an alkaline pH. Some alkaline proteases were also produced by other thermophilic fungi such as *Malbranchea pulchella* var. *sulfurea* (Ong and Gaucher, 1973), *Humicola lanuginosa* (Shenolikar and Stevenson, 1982), and *Cheatomium thermophilum* (Li and Li, 2009), which are different from the fungi proteases whose optimum pH at 3.0-5.5, such as *T. aurantiacus* (Merheb *et al*., 2007), *Penicillium duponti*, *Mucor pusillus*, and *M. miehei* (Maheshwari *et al*., 2000). It is worthwhile to note that alkaline proteases are more advantageous due to some of the detergent and food industries applications.

Nowadays, investigations have been focused mainly on the identification of suitable thermophilic fungal sources for desired enzymes, and the enzymes that are secreted in the growth media have been studied more frequently than cell-associated enzymes, since culture filtrates can be obtained in substantial quantities (Maheshwari *et al*., 2000). This study is the first report of a gene cloned from thermophilic fungi *T. aurantiacus* var. *levisporus,* encoding a serine alkaline protease and over expressed successfully. The thermostability characteristic and the broad substrate specificity of the protease make the TAPRO suitable for use in industry. Characterization analyzing of the enzymes would facilitate the understanding on the biochemical and structural basis of the molecular stability under extreme conditions. It is necessary to make further studies on the protease TAPRO such as amino acid residues involved in substrate binding, packing of the hydrophobic core, electrostatic interactions, and stabilization of helices on the protease.

References

Adinarayana, K. and P. Ellaiah. 2003. Production of alkaline protease by immobilized cells of alkalophilic *Bacillus* sp. *J. Sci. Indust. Res. (India)* 62, 589-592.

Berrin, J.G., G. Williamson, A. Puigserver, J.C. Chaix, W.R. McLauchlan, and N. Juge. 2000. High-level production of recombinant fungal endo-beta-1, 4-xylanase in the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif* 19, 179-187.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254.

Catara, G., G. Ruggiero, F.L. a Cara, F.A. Digilio, A. Capasso, and A.M. Rossi. 2003. A novel extracellular subtilisin-like protease from the hyperthermophile *Aeropyrum pernix* K1: biochemical properties, cloning, and expression. *Extremophiles* 7, 391-399.

Chen, X.G., O. Stabnikova, J.H. Tay, J.Y. Wang, and S.T.L. Tay. 2004. Thermoactive extracellular proteases of *Geobacillus caldoproteolyticus*, sp. nov., from sewage sludge. *Extremophiles* 8, 489-498.

Chen, J., Y.Q. Zhang, C.Q. Zhao, A.N. Li, Q.X. Zhou, and D.C. Li. 2007. Cloning of a gene encoding thermostable glucoamylase from *Chaetomium thermophilum* and its expression in *Pichia pastoris*. *J. Appl. Microbiol.* 103, 2277-2284.

Choi, I.G., W.G. Bang, S.H. Kim, and Y.G. Yu. 1999. Extremely thermostable serine-type protease from *Aquifex pyrophilus*. Molecular cloning, expression, and characterization. *J. Biol. Chem.* 274, 881-888.

Epstein, D.M. and P.C. Wensink. 1998. The α-lytic protease gene of *Lysobacter* enzyme genes. *J. Biol. Chem.* 263, 16568-16590.

Gaur, R., J. Yadav, and L. Pandey. 1989. Thermostability of extracellular protease enzyme produced by *Spicaria fusispora*, a thermophilic fungus. *Hindustan Antibiot. Bull.* 31, 36-37.

Genckal, H. and C. Tari. 2006. Alkaline protease production from alkalophilic *Bacillus* sp. isolated from natural habitats. *Enzyme Microb. Technol.* 39, 703-710.

Gomi, K., K. Arikawa, N. Kamiya, K. Kitamoto, C. Kumagai. A. Kenji, K. Naokata, K. Katsuhiko, and K. Chieko. 1993. Cloning and nucleotide sequence of the acid protease-encoding gene (*pepA*) from *Aspergillus oryzae*. *Biosci. Biotech. Biochem.* 57, 1095-1100.

Gupta, R., Q.K. Beg, S. Khan, and B. Chauhan. 2002. An overview on fermentation, downstream processing and properties of microbial alkaline proteases. *Appl. Microbiol. Biotechnol* 60, 381-395.

Ikemura, H., H. Takagi, and M.J. Inouye. 1987. *In vitro* processing of pro-subtilisin produced in *Escherichia coli* CJ7. *J. Biol. Chem.* 262, 7859-7864.

Kim, T. and X.G. Lei. 2005. Expression and characterization of a thermostable serine protease (TfpA) from *Thermomonospora fusca* YX in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* 68, 355-359.

Kitadokoro, K., E. Nakamura, M. Tamaki, T. Horii, H. Okamoto, M. Shin, T. Sato, and *et al*. 1993. Purification, characterization and molecular cloning of an acidic amino acidspecific proteinase from *Streptomyces fradiae* ATCC 14544. *Biochim. Biophys. Acta.* 1163, 149-157.

Kobayashi, T., Y. Hakamada, S. Adachi, J. Hitomi, T. Yoshimatsu, K. Koike, S. Kawai, and S. Ito. 1995. Purification and properties of an alkaline protease from alkalophilic *Bacillus* sp. KSMK16. *Appl. Microbiol. Biotechnol.* 43, 473-481.

Kumar, S., N.S. Sharma, M.R. Saharan, and R. Singh. 2005. Extracellular acid protease from *Rhizopus oryzae*: purification and characterization. *Proc. Biochem.* 40, 1701-1705.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.

Laxman, R.S., A.P. Sonawane, S.V. More, B.S. Rao, M.V. Rele, V.V. Jogdand, V.V. Deshpande, and M.B. Rao. 2005. Optimization and scale up of production of alkaline protease from *Conidiobolus coronatus*. *Process Biochem.* 40, 3152-3158.

Lebaron. P., J.F. Ghiglione, C. Fajon, N. Batailler, and P. Normand. 1998. Phenotypic and genetic diversity within a colony morphotype. *FEMS Microbiol. Lett*. 160, 137-143.

Lee, F.S. 1992. Modified protocol for Yeast DNA mini-preparation. *Biotechnique* 12, 679.

Li, A.N., A.Y. Ding, J. Chen, S.A. Liu, M. Zhang, and D.C. Li. 2007a. Purification and characterization of two thermostable proteases from the thermophilic fungus *Chaetomium thermophilum. J. Microbiol. Biotechnol.* 17, 624-631.

Li, A.N. and D.C. Li. 2009. Cloning, expression and characterization of the serine protease gene from *Chaetomium thermophilum*. *J. Appl. Microbiol.* 106, 369-380.

Li, J., P.J. Shi, X.Y. Han, K. Meng, P.L. Yang, Y.R. Wang, H.Y. Luo, N.F. Wu, B. Yao, and Y.L. Fan. 2007b. Functional expression of the keratinolytic serine protease gene sfp2 from *Streptomyces fradiae* var. k11 in *Pichia pastoris*. *Protein Expr. Purif.* 54, 79-86.

Liu, Y.G. and N. Huang. 1998. Efficient amplification of inserted

sequences from bacterial artificial chromo some clones thermal asymmetric interlaced PCR. *Plant Mol. Biol. Reporter* 16, 175-181.

Liu, Y. and Q. Yang. 2007. Cloning and heterologous expression of aspartic protease SA76 related to biocontrol in *Trichoderma harzianum*. *FEMS Microbiol. Lett.* 277, 173-181.

Maheshwari, R., G. Bharadwaj, and M.K. Bhat. 2000. Thermophilic Fungi: their physiology and enzymes. *Microbiol. Mol. Biol. Rev.* 3, 461-488.

Manavalan, A., A. Kalaichelvan, V. Kalyanasundaram, A. Perumal, and A. Manavalan. 2007. Purification and partial characterization of serine protease from thermostable alkalophilic *Bacillus laterosporus*-AK1. *World J. Microbiol. Biotechnol.* 23, 475-481.

Marcy, R.M., T.C. Engelhardt, and J.M. Upadhyay. 1984. Isolation, partial purification, and some properties of protease I from a thermophilic mold *Thermoascus aurantiacus* var *levisporus*. *Mycopathologia* 87, 57-65.

Matoba, S., J. Fukuyama, R.A. Wing, and D.M. Ogrydziak. 1988. Intracellular precursors and secretion of alkaline extracellular protease of *Yarrowia lipolytica*. *Mol. Cell. Biol.* 8, 4904-4916.

Merheb, C.W., H. Cabrala, E. Gomesa, and R. Da-Silva. 2007. Partial characterization of protease from a thermophilic fungus, *Thermoascus aurantiacus*, and its hydrolytic activity on bovine casein. *Food Chemistry* 104, 127-131.

Miyaji, T., Y. Otta, T. Shibata, K. Mitsui, T. Nakagawa, T. Watanabe, Y. Niimura, and N. Tomizuka. 2005. Purification and characterization of extracellular alkaline serine protease from *Stenotrophomonas maltophilia* strain S-1. *Lett. Appl. Microbiol.* 41, 253-257.

Morton, C.O., P.R. Hirsch, J.P. Peberdy, and B.R. Kerry. 2003. Cloning of and genetic variation in protease VCP1 from the nematophagous fungus *Pochonia chlamydosporia*. *Mycol. Res.* 107, 38-46.

Ni, X., Z. Chi, C. Ma, and C. Madzak. 2008. Cloning, characterization, and expression of the gene encoding alkaline protease in the marine yeast *Aureobasidium pullulans* 10. *Mar. Biotechnol.* 10, 319-327.

Niehaus, F., C. Bertoldo, M. Kahler, and G. Antranikian. 1999. Extremophiles as a source of novel enzymes for industrial application. *Appl. Microbiol. Biotechnol.* 51, 711-729.

Ong, P.S. and G.M. Gaucher. 1973. Protease production by thermophilic fungi. *Can. J. Microbiol*. 19, 129-133.

Pan, J., Q. Huang, and Y.Z. Zhang. 2004. Gene cloning and expression of an alkaline serine protease with dehairing function from *Bacillus pumilus. Curr. Microbiol.* 49, 165-169.

Pel, H.J., J.H. de Winde, D.B. Archer, P.S. Dyer, G. Hofmann, P.J. Schaap, and G. Turner. 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 25, 221-231.

Peña-Montes, C., A. González, D. Castro-Ochoa, and A. Farrés. 2008. Purification and biochemical characterization of a broad substrate specificity thermostable alkaline protease from *Aspergillus nidulans*. *Appl. Microbiol. Biotechnol.* 78, 603-612.

Rao, M.B., A.M. Tanksale, M.S. Ghatge, and V.V. Deshpande. 1998. Molecular and biotechnological aspects of microbial proteases. *Microbiol. Mol. Biol. Rev.* 62, 597-635.

Rick, W. 1974. Chymotrypsin. *In* H.U. Bergmeyer (ed.) Methods of Enzymatic Analysis. Academic Press, Berlin, New York, and London.

Sambrook, J., E.F. Frotsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd (ed.). Cold Spring Harbor, Cold Spring Harbor Laboratory, New York, N.Y., USA.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.

Saul, D.J., L.C. Williams, J.S. David, H.S. Toogood, R.M. Daniel, and P.L. Bergquist. 1996. Sequence of the gene encoding a highly thermostable neutral proteinase from *Bacillus* sp. Strain EA1: expression in *Escherichia coli* and characterization. *Biochimica et Biophysica Acta* 1308, 74-80.

Shen, H.D., C.W. Wang, W.L. Lin, H.Y. Lai, M.F. Tam, H. Chou, S.R. Wang, and S.H. Han. 2001. cDNA cloning and immunologic characterization of Peno18, the vacuolar serine protease major allergen of *Penicillium oxalicum. J. Lab. Clin. Med*. 137, 115-124.

Shenolikar, S. and K.J. Stevenson. 1982. Purification and partial characterization of a thiol proteinase from the thermophilic fungus *Humicola lanuginosa*. *Biochem. J.* 205, 147-152.

Shi, X.Z., X.F. Zhao, and J.X. Wang. 2008. Molecular cloning and expression analysis of chymotrypsin-like serine protease from the Chinese shrimp, *Fenneropenaeus chinensis. Fish Shellfish Immunol.* 25, 589-597.

Shikha, A. Sharan, and N.S. Darmwal. 2007. Improved production of alkaline protease from a mutant of alkalophilic *Bacillus pantotheneticus* using molasses as a substrate. *Bioresour. Technol.* 98, 881-885.

Silen, J.L. and D.A. Agard. 1989. The a-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*. *Nature* 341, 462-464.

Singer, T. and E. Burke. 2003. High-throughput TAIL PCR as a tool to identify DNA flanking insertions. *Methods Mol. Biol.* 8, 241-272.

Terauchi, R. and G. Kahl. 2000. Rapid isolation of promoter sequences by TAIL PCR: the 5' flanking regions of Pal and Pgigenes from yams (Dioscorea). *Mol. Gen. Genet.* 4, 554-560.

Vickers, I., E.P. Reeves, K.A. Kavanagh, and S. Doyle. 2007. Isolation, activity and immunological characterisation of a secreted aspartic protease, CtsD, from *Aspergillus fumigatus*. *Protein Expr. Purif.* 53, 216-224.

Wang, M., J. Yang, and K.Q. Zhang. 2006. Characterization of an extracellular protease and its cDNA from the nematode-trapping fungus *Monacrosporium microscaphoides*. *Can. J. Microbiol.* 52, 130-139.

Watson, M.E. 1984. Compilation of published signal sequences. *Nucleic Acids Res.* 12, 5145-5164.

Wei, Z., Y. Cao, and Y.X. Xia. 2008. Cloning of the subtilisin Pr1A gene from a strain of locust specific fungus, *Metarhizium anisopliae*, and functional expression of the protein in *Pichia pastoris*. *World J. Microbiol. Biotechnol.* 24, 2481-2488.

Yang, J., L. Liang, Y. Zhang, J. Li, L. Zhang, F. Ye, Z. Gan, and K.Q. Zhang. 2007. Purification and cloning of a novel serine protease from the nematode-trapping fungus *Dactylellina varietas* and its potential roles in infection against nematodes. *Appl. Microbiol. Biotechnol.* 75, 557-565.