Electronic Journal of Biotechnology ISSN: 0717-3458
Vol.11 No.4, Issue of October 15, 2008

Received October 17, 2007 / Accepted May 22, 2008
DOI: 10.2225/vol11-issue4-fulltext-5
RESEARCH ARTICLE

Molecular cloning, expression and characterization of a serine proteinase from Japanese edible mushroom, *Grifola frondosa*: solving the structure - function anomaly of a reported aminopeptidase

Mohammed M. Islam[#]
Enzyme Laboratory
National Food Research Institute
Tsukuba, 2-1-2 Kannondai
Ibaraki 305-8642, Japan
Tel: 61-3-5320-2041
Fax: 61-3-5320-2408
E-mail: moniruli61@yahoo.com

Financial support: This work was supported by the Science and Technology Agency (STA) fellowship sponsored by Japan International Science & Technology Exchange Centre (JISTEC).

Keywords: 5'-RACE, cDNA, genomic DNA, maitake, refolding.

[#]**Present address:** Ballarat Cancer Research Centre, University of Ballarat, c/o-St. John of God Hospital, 101 Drummond Street, Ballarat, VIC 3350, Australia. Tel: 61-3-5320-2041; Fax: 61-3-5320-2408.

Abbreviations: CAPS: N-cyclohexyl-3-aminopropanesulfonic acid
cDNA: omplementary DNA to mRNA
CHES: N-cyclohexyl-2-aminoethanesulfonic acid.
DTT: dithiothreitol
FPLC: fast protein liquid chromatography
IPTG: isopropylthio-β-D-galactoside
LB: Luria-Bertani
MES: 2-(N-morpholino)ethanesulfonic acid
Ni-NTA: Nickel-nitriloacetic acid
ORF: open reading frame
PCR: polymerase chain reaction
PMSF: phenylmethanesulphonylfluoride
PVDF: polyvinylidene fluoride
RACE: rapid amplification of cDNA ends
SDS-PAGE: sodium dodecyl sulfate - polyacrylamide gel electrophoresis
TAPS: N-tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
UTR: untranslated region

The N-terminal amino acid sequence of an aminopeptidase from Japanese edible mushroom, *Grifola frondosa*, was reported to have high similarity with that of a serine proteinase from basidiomycete, *Agaricus bisporous* (Nishiwaki and Hayashi, 2001). The full-length cDNA and the corresponding genomic DNA of the enzyme were cloned, based on the reported N-terminal amino acid sequence. The predicted open reading frame (ORF) of the cloned cDNA, encoding a product of 379 amino acids, was expressed in *E. coli* using pET expression vector. The expressed pro-enzyme (40 kDa) underwent autolysis to produce the mature protein (30 kDa) and a pro-peptide (10 kDa). The mature protein and the pro-peptide remained tightly bound to each other and could not be separated by Ni-NTA metal affinity chromatography or Q-Sepharose ion-exchange chromatography. The enzyme was inactive in the bound form. Upon treatment with subtilisin, the bound pro-peptide was further hydrolyzed and a high serine proteinase activity was recovered. No aminopeptidase activity was detected at any stage of the protein processing. These results clearly indicated that the N-terminal amino acid sequence and the function of the reported aminopeptidase were not derived from the same protein entity and hence caused the structure-function anomaly.

Aminopeptidases are a class of proteolytic enzymes that catalyze the cleavage of N-terminal amino acids from polypeptides or proteins. They are widely distributed in animals, plants, bacteria, and fungi. They play important roles in protein maturation, activation, modulation, the degradation of bioactive peptides, and the determination of protein stability (Taylor, 1993). Aminopeptidases also play roles in several biological processes such as angiogenesis, cell cycle regulation, reproduction, blood pressure control,

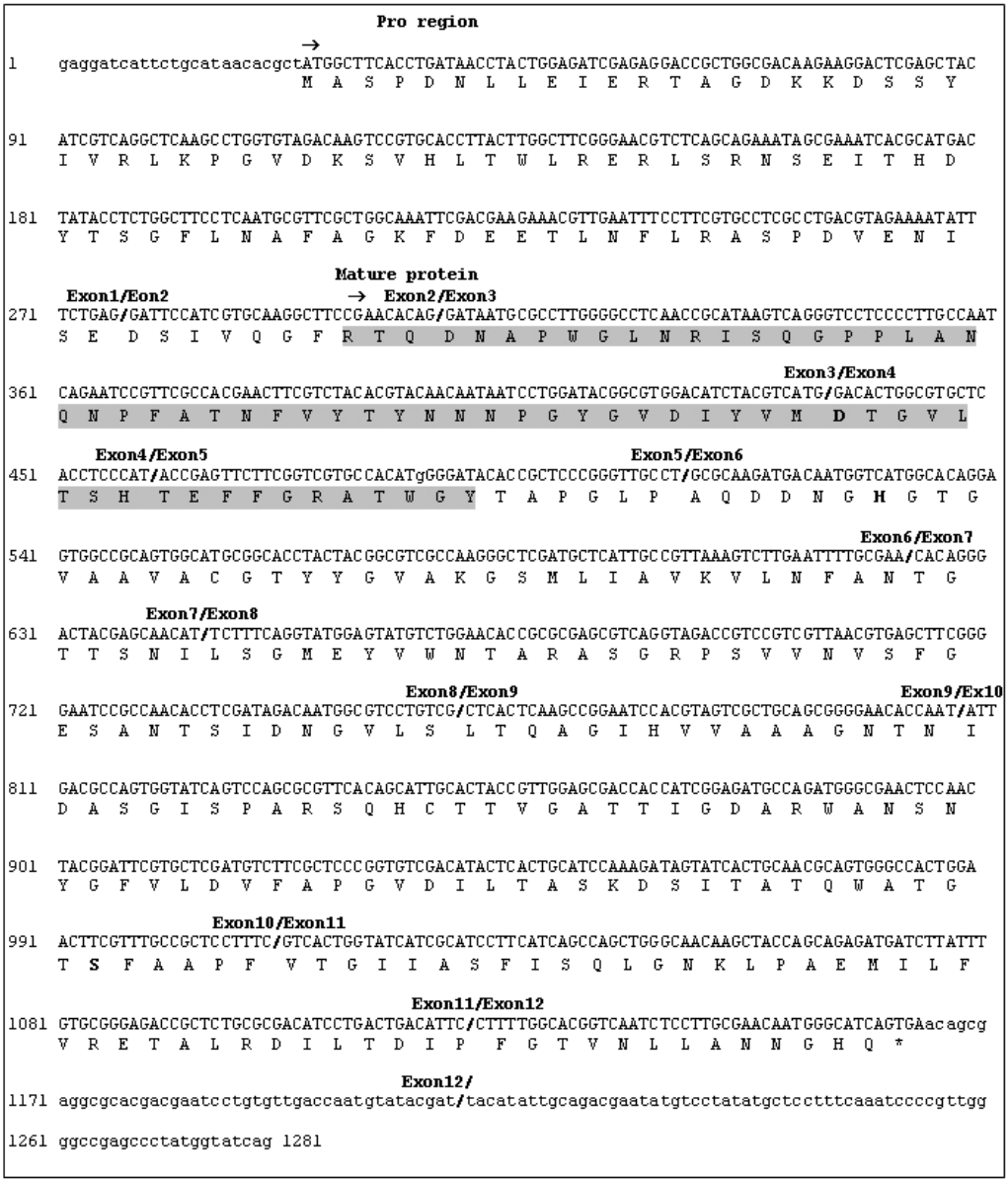

Figure 1. Nucleotide and predicted amino acid sequences of gf-spr1. The nucleotide sequence is available in the DDBJ/EMBL/GenBank database with accession no. AB090242. The amino acid sequence is shown in single-letter code below the nucleotide sequence. The coding sequence is shown in uppercase characters, while the 3'-UTR and 5'-UTR are shown in lowercase. The pro-sequence and the start of the mature protein are marked with arrows. The aspartic acid, histidine and serine residues of the active site are shown in boldface larger type. The stop codon is indicated with an asterisk. Exon-intron junctions are indicated by slashes (/) between bases. The N-terminal 65 amino acid residues of the target aminopeptidase are shaded in the amino acid sequence.

and antigen presentation to MHC class I molecule (Wright et al. 1990; Reaux et al. 1999; Albiston et al. 2001; Osada et al. 2001; Saric et al. 2002; Serwold et al. 2002; Sato, 2004; Goto et al. 2006a; Goto et al. 2006b; Maruyama et al. 2007). Some of the aminopeptidases require metal ions for their catalytic activity and are hence called metallopeptidases. They can be inactivated with chelating agents such as EDTA or o-phenanthroline; the later being more specific for Zn^{2+} metallopeptidases. Bestatin is also an inhibitor of aminopeptidases but is specific for aminopeptidase B and leucine aminopeptidase. On the other hand, serine proteinases are composed of a large group of endopeptidases that require a serine residue for their catalytic activity. They are equally important for biological systems (Kalafatis et al. 1997; Rooprai and McCornick, 1997; Turgeon and Houenou, 1997; Salamonsen, 1999; Diamandis and Yousef, 2001; Yousef and Diamandis, 2002; Blaber et al. 2004; Kini, 2005; Voswinkel et al. 2007; Whitcomb and Lowe, 2007; Yoon et al. 2007) as well as for commercial purposes (Horikoshi, 1999; Saeki et al. 2007). Serine proteinases do not need a metal ion for their catalytic activity, but a conserved catalytic triad Ser-His-Asp is

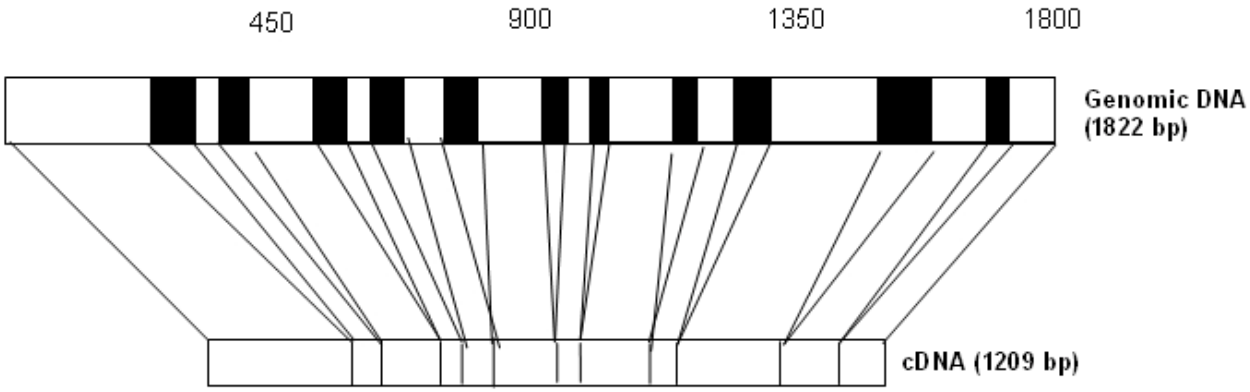

Figure 2. Schematic diagram for the intron-exon arrangement in the gf-spr1 gene. The nucleotide sequence of *gf-spr*1 gene is available in the DDBJ/EMBL/GenBank database with accession no. AB090947. The intron and exon regions are shown in black and white boxes, respectively, in the genomic DNA (1822 bp). The vertical lines on the cDNA box (1209 bp) represent the intron-exon boundaries. Nucleotide positions of exons are presented in the Results section.

essential for their activity. PMSF is a typical inhibitor of serine proteinases. In addition, these two groups of enzymes are conspicuously different in their primary structures.

Nishiwaki and Hayashi (2001) reported an aminopeptidase from a Japanese edible mushroom, *Grifola frondosa,* commonly known as maitake. The N-terminal amino acid sequence of the enzyme was determined up to 65 residues but no significant homology was observed with that of any known aminopeptidase. In contrast, very high homology was observed with the N-terminal amino acid sequence of a serine proteinase from basidiomycete, *Agaricus bisporous*. They could not explain the cause of the inconsistency between the sequence information and the activity of the enzyme, rather they emphasized on their inabilities to detect the serine proteinase activity using the substrate, N-Succinyl-Ala-Ala-Pro-Phe-pNA. These contrasting results inspired us to clone and express the full-length cDNA of the enzyme in *E. coli* so that we could further characterize it.

This report presents the comprehensive methods for cloning, expression, and characterization of the target aminopeptidase, based on the reported N-terminal amino acid sequence. The expressed enzyme was initially inactive and was very difficult to characterize. Two step refolding analysis also failed to revive the enzyme activity. But after treatment with subtilisin, a new serine proteinase activity was generated. No aminopeptidase activity was detected at any stage of the protein processing. These results clearly contradicted the report of Nishiwaki and Hayashi (2001) who characterized the enzyme as an aminopeptidase instead of a serine proteinase.

A genomic DNA sequence of the enzyme consisting of 11 introns and 12 exons was also cloned and is reported in this article. This is the first report of a cloned serine proteinase from *Grifola frondosa*.

MATERIALS AND METHODS

PCR-ready crude DNA preparation

Genomic DNA was extracted from maitake fruiting bodies using an InstaGene™ Matrix kit (BIO-RAD) according to the manufacturer's instructions.

Total RNA purification from maitake

Samples of about 100 mg of fresh maitake fruiting body were crushed in liquid nitrogen using a hand-held super homogenizer (Toyobo, Osaka, Japan) according to the manufacturer's directions. Total RNA was then extracted from the powdered maitake and purified using an RNeasy Plant Mini kit (Qiagen) in accordance with the manufacturer's protocol with slight modifications. At the final step of elution, 1 µl of the RNase inhibitor RNasin (40 U/µl) (Promega) and 1 µl of RNase-free dithiothreitol (DTT) (0.1 M) were added to the bottom of the collection tube before elution of the total RNA with 50 µl of RNase-free water. Extracted RNA was then treated with 20 units (2µl) of RNase-free DNase-I (Takara, Osaka, Japan) in high salt restriction enzyme buffer (Takara) for 30 min at room temperature to remove any contaminating DNA. Treated RNA was then purified again using the RNeasy Mini Protocol for RNA cleanup (Qiagen) and the final elute was collected in 50 µl of RNase-free water in microcentrifuge tubes containing 1 µl of RNasin (40 U/µl) and 1 µl of DTT (0.1 M).

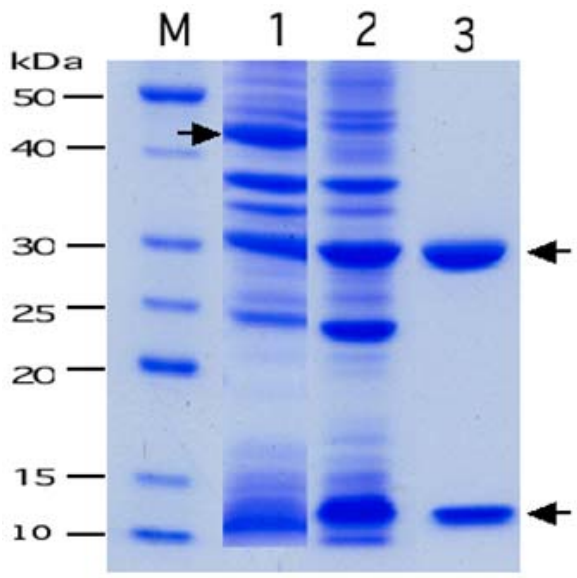

Figure 3. SDS-PAGE of GF-Spr1 at various stages of purification. Lane-M, Molecular weight marker; Lane-1, Insoluble fraction of crude extract; Lane-2, Soluble fraction of crude extract; Lane-3, Ni-NTA purified enzyme. The pro-protein (40 kDa), mature protein (30 kDa), and pro-peptide (10 kDa) are indicated by arrows.

Preparation of double-stranded cDNA

Double-stranded cDNA was prepared using a SMART™ PCR cDNA Synthesis kit (Clontech) according to the manufacturer's protocol with 2 µg of total maitake RNA. First strand cDNA synthesis was carried out using Powerscript Reverse Transcriptase. Second strand synthesis and amplification of both strands were performed with Advantage 2 DNA polymerase mix. An Oligo dT-linked CDS primer (cDNA synthesis primer, SMART™ PCR cDNA Synthesis kit component) was used for first strand cDNA synthesis and a specially designed long distance (LD) primer was used for second strand synthesis.

Full-length cDNA cloning

Full-length cDNA cloning of the target protein was performed in three phases. In the first phase, a gene-specific PCR product (140 bp) was amplified from PCR-ready crude maitake DNA using two degenerate primers: FWD 5'- CAR GAY AAY GCT CCN TGG GG-3' and REV 5'- CCG TRT CCA TGA CRT ADA TRT C-3' (where R = A, G; Y = C, T; W = A, T; N = A, C, G, T; D = A, G, T). These primers were designed based on amino acid positions (3-9) and (42-49), respectively, of the known N-terminal 65 amino acid residues of the target enzyme. A GeneAmp PCR System 9600 thermal cycler (Perkin-Elmer) was used with a thermal profile: 96ºC for 5 min followed by 25 cycles of 96ºC for 1 min, 60ºC for 1 min, and 72ºC for 10 sec. A final 5 min elongation step at 72ºC was performed to complete the reaction. The PCR product was then cloned into the TOPO-TA cloning vector (Invitrogen) and sequenced.

In the second phase, 5' rapid amplification of cDNA ends (5'-RACE) of the target gene was performed using a SMART™ PCR cDNA Synthesis kit. Two nested anti-sense primers, R2 (5'-GTGTCAAACGGATTCTGGTT-3') and R3 (5'-GCTTGACTTATGCGGTTGAG-3'), were designed from the known DNA sequence of the 140 bp PCR product using Oligo 4.0 Primer Analysis Software (National Bioscience Inc., MN, USA). The first strand 5'cDNA end was synthesized from 1.0 µg of total RNA using Power Script Reverse Transcriptase in a reaction volume of 10 µl containing 1 mM R2 primer, 1 mM SMART II Oligonucleotide, 2 mM DTT, 2 mM dNTP mixture, and 1x First Strand Buffer. The reaction was performed in two steps. In the first step, total RNA, R2 primer, and SMART II Oligonucleotide were mixed in a volume of 5 µl, incubated for 2 min at 70ºC, and then allowed to cool to room temperature. The rest of the components were then added to a final volume of 10 µl and then incubated at 42ºC for 1 hr. Second strand synthesis and amplification of both strands were performed by PCR using Advantage 2 DNA Polymerase in a reaction volume of 50 µl containing 1 µl of five fold diluted first strand cDNA, 0.1 µM LD (long distance) primer, 0.1 µM R3 primer, and other PCR components. The thermal profile used for PCR was: pre-PCR treatment at 95ºC for 1 min, 30 cycles of 95ºC for 5 sec, 60ºC for 5 sec, and 68ºC for 2 min. A 390 bp 5' RACE product was obtained, which was then cloned into the pDrive vector using a Qiagen PCR Cloning kit (Qiagen) and then sequenced.

In the third phase, a full-length cDNA of the target protein was amplified by PCR from the double-stranded cDNA using the primer pair cDNA 5' END primer (5'-GAGGATCATTCTGCATAACACGC- 3') and CDS primer with Advantage 2 Polymerase Mix. The thermal profile used was: pre-PCR heating at 95ºC for 1 min, 30 cycles of 95ºC denaturation for 30 sec, 60ºC annealing for 5 sec, 68ºC elongation for 2 min, and a final post-PCR elongation step at 68ºC for 5 min. The PCR product was then cloned into the pDrive vector and sequenced.

Full-length genomic DNA cloning

Full-length genomic DNA, corresponding to the full-length cDNA, was amplified from maitake genomic DNA using a gene-specific primer pair, sense-U16 (5'- TAA CAC GCT ATG GCT TCA CCT -3') and antisense-L1188 (5'-CGT ATA CAT TGG TCA ACA CA- 3'). Primers were designed based on the sequences of each end of the full-length cDNA un-translated regions (UTR). PCR was performed using a proofreading DNA polymerase, KOD-plus (TAKARA). The thermal profile used for PCR was: pre-PCR heating at 98ºC for 5 min, followed by 32 cycles of 98ºC denaturation for 30 sec, 55ºC annealing for 30 sec, and 68ºC elongation for 4 min, with a final post-PCR

elongation step at 68ºC for 5 min. The PCR product was cloned into the pDrive vector and then sequenced.

DNA sequencing of PCR products

PCR products were either treated with *Taq* DNA polymerase for 10 min at 70ºC, when the proofreading DNA polymerase was used for PCR, or cloned directly into the TOPO TA vector (Invitrogen) or the pDrive UA vector (Qiagen) when non-proofreading DNA polymerase was used. The respective manufacturer's protocols were strictly followed for cloning into these vectors. Recombinant plasmids were then transformed into appropriate host cells by electroporation using an *E. coli* Pulser^{TM} (Bio-Rad) and grown on selective antibiotic LB plates. Positive clones were determined by colony PCR. Recombinant plasmids from the positive clones were purified from overnight cultures using a QIAprep^{R} Spin Miniprep kit (Qiagen). Sequencing reactions with appropriate primer(s) were performed using a BigDye Terminator Cycle Sequencing kit (Perkin-Elmer) and reaction products were analyzed using a 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Primers for sequencing were either pre-designed, in the vector, or designed using primer analysis software, 'Oligo' (Version 4). DNA sequences were further analyzed using GENETYX program (Software Development Corporation, Tokyo, Japan).

Expression vector construction

The ORF of the cloned enzyme, termed pro-GF-spr1, was amplified from the recombinant DNA construct (cDNA-pDrive) by PCR using a pair of primers: 5' sense primer (5'-GGAGATATACCATGGCTTCACCTGATAACCTA-3') and 3' anti-sense primer (5'-CTCGAGTGCGGCCGCCTGATGCCCATTGTTCGCAA G-3'). The sense primer contained an *Nco*I restriction site (underlined) and the anti-sense primer contained a *Not*I restriction site (underlined). PCR was performed using a proofreading DNA polymerase, KOD-plus (TAKARA). The thermal profile used for the PCR was: pre-PCR heating at 98ºC for 5 min, 25 cycles of 98ºC for 30 sec, 58ºC for 30 sec, 68ºC for 2 min, and an extra elongation step at 68ºC for 5 min. The PCR product was then digested with *Nco*I and *Not*I and cloned into the pET-28(+) vector, which was previously digested with both of these enzymes. High 4 DNA ligase (Toyobo) was used for the ligation reaction. Recombinant (pET-gf-spr1) was then transformed into *E. coli* BL21-Codon-Plus-RIL competent cells by electroporation. Positive recombinant clones were selected on LB plates containing kanamycin, and the correct insertion of the ORF into the cloning sites was verified by sequencing.

Expression and purification of the expressed protein

Transformed recombinant cells harboring the correct expression construct were grown in 2L LB broth containing kanamycin (50 µg/ml) at 37ºC until the growth reached log phase (OD_{600} = 0.4 - 1.0). IPTG at a final concentration of 0.4 mM was added to the cultures to induce protein expression. Growth was continued for a further 20 hrs at 20ºC, at which time the OD_{600} reached around 1.5. Cells were then harvested by centrifugation and washed with 200 ml of 20 mM Tris-HCl buffer (pH 7.0). The cell pellet was then re-suspended in 200 ml of the same buffer and disrupted by ultra-sonication with nine bursts at a power level of 5 and 50% duty-output (BRANSON ULTRASONICS, Danbury, CT, USA). The disrupted cells were then centrifuged at 10,000 x g for 10 min and the clear supernatant was collected as the crude enzyme fraction.

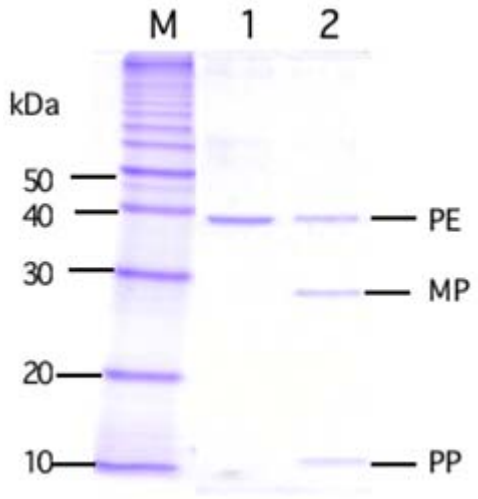

Figure 4. Refolding and the auto processing of pro-GF-Spr1. The two stage refolding method (Marston, 1986; Tang et al. 2002) was employed for refolding of the insoluble pro-form of the enzyme as described in Experimental Procedures section. Lane-M, Molecular weight marker; Lane-1, Refolded pro-GF-Spr1; Lane-2, Auto-processed GF-Spr1. PE stands for pro-enzyme Pro-GF-Spr1; MP stands for mature protein GF-Spr1; PP stands for N-terminal pro-peptide.

The first step of protein purification was performed by Ni-NTA metal chelate affinity chromatography. About 5.0 mg of crude protein was mixed with 1 ml of Ni-NTA resin pre-equilibrated with 50 mM MOPS buffer (pH 7.5) containing 300 mM NaCl and 20 mM imidazole. The resin was then packed into a 1 ml column and washed with 10 volumes of the same MOPS buffer. The enzyme was then eluted in a linear gradient of 20 - 250 mM imidazole in 50 mM MOPS buffer (pH 7.5) containing 300 mM NaCl using an FPLC purifier system (Pharmacia). The major protein peaks were pooled together, dialyzed overnight in 20 mM Tris-HCl (pH 8.0) and then diluted two-fold with the same Tris-HCl buffer. A second round of purification was performed by Q-Sepharose anion exchange column chromatography. The dialyzed and diluted sample was loaded onto a pre-equilibrated Q-Sepharose column (1 ml) with 50 mM Tris-HCl buffer (pH 8.0), and the enzymes were eluted in a linear gradient of 0 - 250 mM NaCl at a flow rate of 0.5

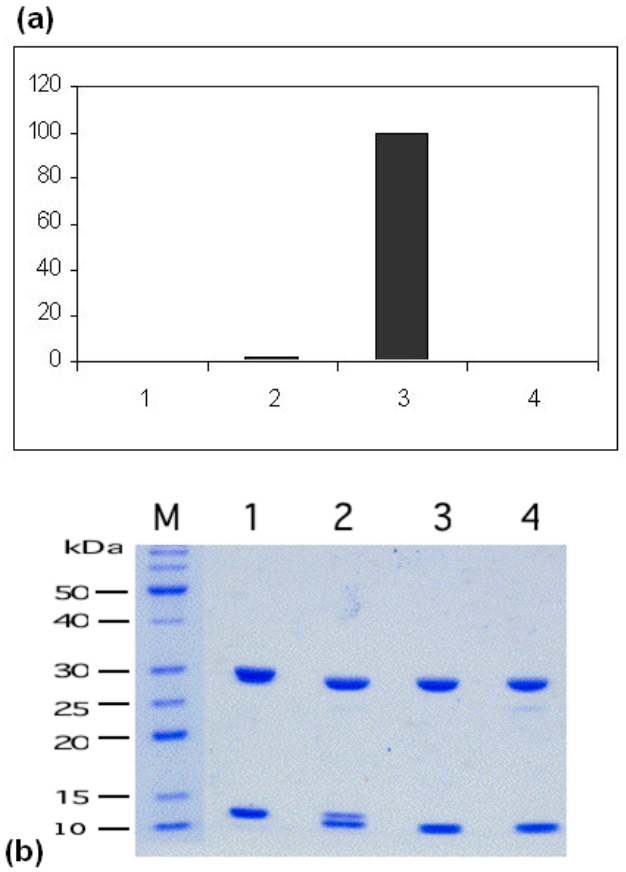

Figure 5. Subtilisin treatment and revival of proteolytic activity of GF-Spr1.

(a) Enzyme activity profiles at different stages of subtilisin treatment. The proteolytic activity of the purified enzyme was assayed before and after subtilisin treatment using Succinyl-Ala-Ala-Pro-Phe-*p*NA as the substrate, as described in Experimental Procedures. Lane-1, enzyme activity before subtilisin treatment; Lane-2, activity of subtilisin alone; Lane-3, enzyme after subtilisin treatment; Lane-4, activity after inactivation of the subtilisin treated enzyme with PMSF. The proteolytic activity of the subtilisin treated sample was defined as 100%.

(b) SDS-PAGE of subtilisin-treated GF-Spr1. Lane-M, molecular weight marker; Lane-1, Ni-NTA-purified enzyme; Lane-2, 1 hr subtilisin-treated enzyme; Lane-3, 3 hrs subtilisin-treated enzyme; Lane- 4, 16 hrs subtilisin-treated enzyme. The 27 kDa faint band in lanes 2-4 was derived from the subtilisin itself.

ml/min. Homogeneity of the purified enzyme fractions was monitored by SDS-PAGE.

Amino acid sequencing

Protein bands separated by SDS-PAGE were transferred onto PVDF membranes (Bio-rad) by Western blotting. After staining with Coomassie Brilliant Blue R250, the target bands were excised and subjected to N-terminal amino acid sequencing by the Edman degradation method using a G1000A protein sequencer (Hewlett Packard).

Preparation of inclusion bodies

After addition of 0.4 mM IPTG, cells were grown at 37ºC for 3-4 hrs and then harvested by centrifugation. Cells were washed three times with 200 ml of 20 mM Tris-HCl buffer (pH 7.0) and centrifuged at 3000 x g for 5 min. Washed cells were then disrupted by ultra-sonication in a manner similar to that described in previous section. The cell lysate was then centrifuged and the supernatant collected as the soluble fraction. The insoluble pellet was suspended in 200 ml of wash buffer (20 mM Tris-HCl, pH 7.5, 1% Triton X-100) and then centrifuged at 20,000 x g for 10 min. The supernatant was discarded and the pellet was re-suspended in 200 ml of wash buffer. After three successive washes, the pellet was stored as the inclusion body fraction at -20ºC until it was used for refolding.

Refolding and processing of pro-enzyme into mature enzyme

A revised two-stage refolding method (Marston, 1986; Tang et al. 2002) was employed for refolding the insoluble pro-form of the enzyme. The inclusion bodies were solubilized in 20 mM CAPS buffer (pH 11) containing 8 M urea and 10 mM dithiothreitol (DTT). After incubation at 30ºC for 1 hr, the solution was centrifuged at 20,000 x g for 10 min and the supernatant was recovered as the denatured pro-form enzyme solution. One volume of the denatured pro-protein was then diluted with 9 volumes of 20 mM CAPS (pH 11.0) to give a final protein concentration of 500 µg/ml, and the resulting solution was maintained at 25ºC for 2 hrs. The solution was then dialyzed against 20 mM Tris-HCl (pH 7.0) at 4ºC. The pro-protein underwent refolding during the dialysis step and samples were removed at various time intervals for SDS-PAGE analysis.

Subtilisin treatment and activation of the inactive enzyme

The purified protein was diluted with 20 mM Tris-HCl buffer (pH 7.0) to a final concentration of 20 mM. Subtilisin (Bacillus) was added to the mixture at a final concentration of 0.05 mM and incubated for 1-16 hrs at 30ºC. Aliquots of enzyme mixtures were withdrawn at different time intervals to check digestion by SDS-PAGE and activity assay.

Enzyme assay

Serine proteinase activity of the enzyme was measured by the release of p-nitroanilidine (pNA) from the typical serine proteinase substrate N-Succinyl-Ala-Ala-Pro-Phe-pNA and monitored by spectrophotometry at 405 nm. The assay was performed in 50 mM Tris-HCl buffer (pH 8.2) at 30ºC for 20 min in a reaction volume of 50 µl containing 0.2 mM substrate and 5 µl of the appropriately diluted enzyme. The reaction was stopped by the addition of an equal volume of 1 M acetic acid followed by 10 min incubation at room temperature.

Aminopeptidase activity was measured at 45ºC for 5 min in 50 mM TAPS buffer, pH 8.5 using 2 mM Leu- pNA as a substrate, in a reaction volume of 50 µl. The reaction was stopped by the addition of 0.5 ml of 10% acetic acid and the absorbance of the reaction mixture was measured at 405 nm.

Effect of temperature and pH

The optimum temperature was determined under standard assay conditions by incubating the reaction mixtures at temperatures ranging from 25ºC to 70ºC. To determine thermal stability, the enzyme was pre-incubated for 15 min in the standard assay buffer at temperatures ranging from 25ºC to 70ºC and then chilled on ice for 10 min. The remaining enzyme activity was then determined by the standard assay method. To determine the optimum pH for the enzyme activity, standard assay mixtures in 50 mM of the following buffers were used: sodium acetate (pH 3.4 - 5.66), 2-(N-morpholino) ethane sulfonic acid (MES; pH 5.15 - 7.22), Tris-HCl (pH 7.13 - 8.9), 2-(N-cyclohexylamino)ethane sulfonic acid (CHES; pH 8.17 - 10.28) and 3-(cyclohexylamino)-1-propanasulfonic acid (CAPS; pH 9.4 - 11.39). The pH stability was determined by pre-incubating the enzyme in the above buffers as well as in sodium citrate buffer (pH 2.17 - 4.1; all 10 mM) for 30 min at 30ºC. The remaining activity was then measured under standard assay conditions.

Study of kinetic parameters

The minor activity due to subtilisin (about 1% of total activity) was first inactivated to almost zero by incubating the enzyme mixture at 40ºC for 30 min. The remaining activity was solely due to the cloned serine proteinase. The kinetic parameters K_m and K_{cat} were then determined against the substrate N-Succinyl-Ala-Ala-Pro-Phe-pNA. Data was obtained by measuring the initial rate of hydrolysis by incubating the enzyme with appropriate concentrations of the substrate in 50 mM Tris-HCl (pH 8.5) at 30ºC. The reaction was monitored at 405 nm on a Beckman spectrophotometer (model DU 640) equipped with a temperature-controlled cell holder. Initial hydrolysis rates were determined at six different concentrations ranging from approximately 0.5 to 4.0 x the K_m value. Values for K_m and K_{cat} and their standard errors were obtained using the nonlinear regression analysis program GraFit (Leatherbarrow, 1996).

RESULTS

Full-length cDNA cloning

A cDNA consisting of 1320 nucleotides was cloned following three phases of cloning procedures. The predicted amino acid sequence showed that the cDNA encoded a protein consisted of 379 amino acids including a pro-peptide of 91 amino acids (Figure 1). It contained a methionine start codon at nucleotide position 25-27 and a stop codon at nucleotide position 1162-1164. The poly A tail started at nucleotide position 1282 but no consensus polyadenylation signal sequence was identified. The N-terminal amino acid sequence of the predicted mature protein showed 100% identity to the N-terminal 65 amino acid residues of the target aminopeptidase. These results indicated that the full-length cDNA for the target enzyme was cloned.

Full-length genomic DNA cloning

A full-length genomic DNA consisting of 1822 nucleotides was cloned from crude maitake DNA following the method described in the Materials and Methods. Alignment of the genomic sequence, with that of the cDNA, showed that the genomic DNA consisted of 11 introns and 12 exons, and that the exon regions were 100% identical to the cDNA sequences (Figure 2). The nucleotide positions of exons were exon 1: 9 - 260; Exon 2: 321 - 349; exon 3: 407 - 537; exon 4: 596 - 618; exon 5: 672 - 722; exon 6: 776 - 888; exon 7: 947 - 967; exon 8: 1028 - 1142; exon 9: 1194 - 1241; exon 10: 1308 - 1511; exon 11: 1570 - 1678 and exon 12: 1734 - 1777. Exons ranged 21 to 260 bp in length and introns ranged 53 to 67 bp in length. We designated this gene as gf-spr1 and the gene product (protein) as GF-Spr1.

Expression and purification of GF-Spr1

IPTG induced recombinant bacteria, harboring the expression construct pET-gf-spr1, were disrupted by ultra-sonication and expression products were collected from the cell lysate by centrifugation. SDS-PAGE analysis of the cell extracts showed (Figure 3) two major bands of 30 kDa and 10 kDa in the soluble portion of the cell extract and a minor band of 40 kDa in the insoluble fraction. The estimated molecular weights of these protein bands corresponded to the predicted molecular masses of the mature protein (30,162), pro-peptide (10,290), and pro-protein (40,435) of the cloned enzyme, respectively. None of these bands were expressed in control cells carrying only the pET-28b (+) vector (without any insert), even after IPTG induction.

The first step in the purification of the soluble expressed proteins was performed by Ni-NTA affinity chromatography. The purified proteins contained equimolecular proportions of about 95% pure 10 kDa protein and 30 kDa protein (Figure 3). A second round of purification was performed with Q-Sepharose ion-exchange chromatography, but this procedure failed to separate these two bands, indicating that they were strongly bound to each other. N-terminal amino acid sequencing of up to 5 residues of each band showed that the 30 kDa protein exactly matched the mature target protein and the 10 kDa protein matched the predicted pro-peptide. These findings clearly indicated that we successfully expressed the target protein but that it underwent processing into the pro-peptide and mature peptide after expression.

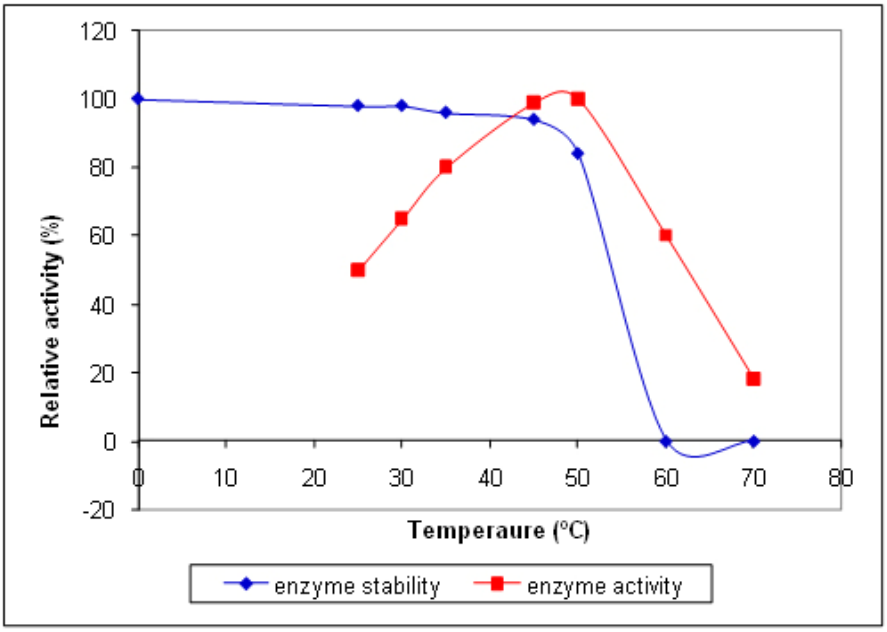

Figure 6. Effects of temperature on activity and stability of GF-Spr1. The effects of temperature on enzyme activity were determined by measuring the enzyme activity at different temperatures. The thermal stability of the enzyme was determined by measuring the remaining activity of the enzyme after incubating the enzyme at different temperatures for 15 min. In both cases, the standard assay buffer was used. The optimum or highest enzyme activity was defined as 100% activity and other activity levels were calculated relative to this value.

Refolding analysis of insoluble inclusion bodies

As described in the Materials and Methods section, inclusion bodies were isolated from the insoluble fraction of the cell extract and then refolded using a two-stage oxidative refolding method (Marston, 1986). No processing of the pro-enzyme was observed at the end of the first refolding stage. However, processing of pro-GF-Spr1 was observed during the second refolding stage, whereby the 40-kDa pro-enzyme was divided into two parts with the molecular mass of 30 kDa and 10 kDa, respectively (Figure 4). N-terminal amino acid sequencing of the pro-enzyme and the divided peptides confirmed that the 30 kDa protein was the N-terminal pro-peptide-deleted mature protein and the 10 kDa protein was the cleaved N-terminal pro-peptide.

Role of pro-peptide in enzyme activity

None of the refolded protein, crude protein or pure protein showed any proteinase or aminopeptidase activity even after 24 hrs incubation under the experimental assay conditions. However, after the treatment of the pure protein for 3 hrs at 30ºC with 0.05 mM (final concentration) subtilisin, a 90-100 fold increase in hydrolytic activity was observed (Figure 5a). This activity was inhibited by the addition of 1 µM PMSF, which is a typical serine proteinase inhibitor, indicating that the expressed protein was indeed a serine proteinase. SDS-PAGE analysis of the subtilisin-treated pure enzyme showed that the pro-peptide fragment was hydrolyzed to a smaller peptide leaving the active mature protein intact (Figure 5b). These results indicated that the pro-peptide, which was strongly bound to the mature protein, strongly inhibited the enzyme activity, and high enzyme activity was recovered after digestion of the pro-peptide with subtilisin. No aminopeptidase activity was detected at any stage of the subtilisin treatment.

Enzyme characterization

The optimum temperature of the subtilisin treated enzyme was determined to be at 50ºC under standard assay conditions. Enzyme activity increased with temperature from 25ºC to 50ºC, and then decreased with further elevation of temperature. At 70ºC, less than 5% of the maximum activity was found. Thermal stability of the enzyme was determined by measuring the residual enzyme activity after incubation at different temperatures ranging from 0ºC to 70ºC for 30 min. The enzyme was stable within the temperature range from 0ºC to 45ºC and then declined sharply with increasing temperature (Figure 6).

Maximum enzyme activity was observed at pH 8.5 (Figure 7a). The enzyme was incubated at 30ºC for 30 min in different pH buffers and then the residual activity was measured under standard conditions. These experiments indicated that the enzyme was stable from pH 3.5-10.0 (Figure 7b). Background subtilisin activity was always subtracted from the total activity in all the activity measurements to obtain the activity of the enzyme alone (redundant activity).

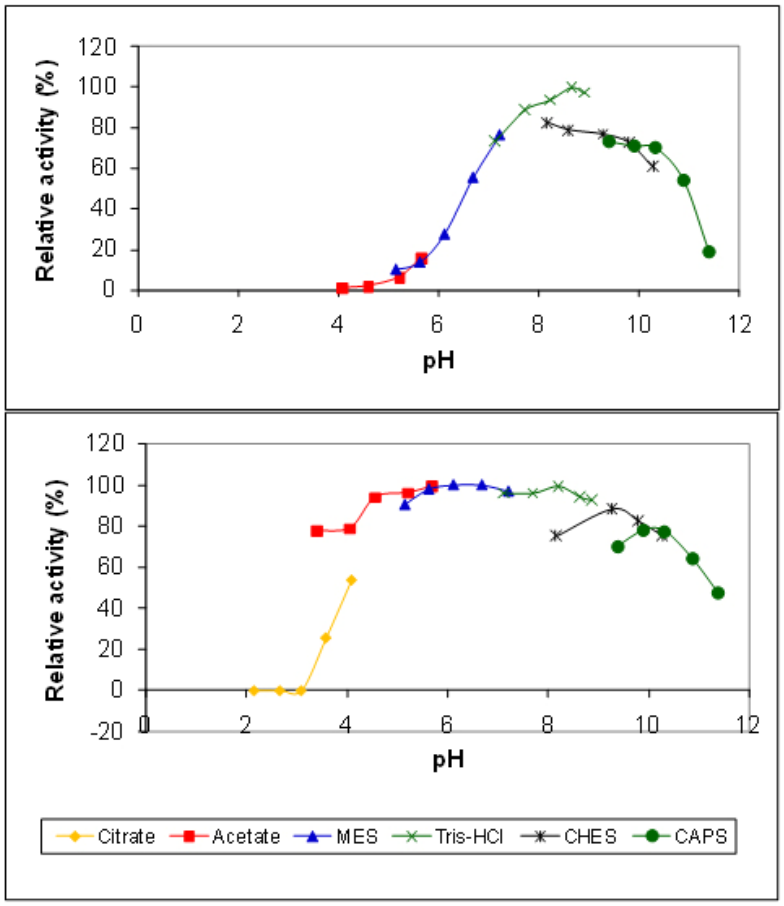

Figure 7. Effect of pH on the activity and stability of GF-Spr1.

(a) Effect of pH on enzyme activity. To determine the optimum pH for the enzyme activity, standard assay mixtures in 50 mM of the following buffers were used at 30ºC: sodium acetate, pH 3.4 - 5.66; MES, pH 5.15 - 7.22; Tris-HCl, pH 7.13 - 8.9; CHES, pH 8.17 - 10.28 and CAPS, pH 9.4 - 11.39. The highest activity observed at pH 8.5 was defined as 100% activity and other activity levels were calculated relative to this value.

(b) Effect of pH on enzyme stability. The pH stability was determined by measuring the remaining activity of the enzyme after incubating the enzyme at 30ºC for 30 min with 10 mM of all the above mentioned buffers and an additional buffer, Citrate, pH 2.17 - 4.1.

The value of the kinetic parameters K_m and K_{cat}, determined against only the substrate N-Succinyl-Ala-Ala-Pro-Phe-pNA, were 0.0931 ± 0.0109 mM^{-1} and 0.496 ± 0.02 s^{-1}, respectively. The ratio of K_m / K_{cat} was 5.32 mM^{-1} s^{-1}.

DISCUSSION

The nucleotide sequence of the cloned full-length cDNA was analyzed to characterize the enzyme. At least three clones containing the full-length cDNA were subjected to sequence analysis. Sequence alignment (data not shown) revealed that all three sequences were almost identical except in the length of their 3'-untranslated regions (3-UTR), which ranged from 117 to 130 bp. Differences in the lengths of the 3'-UTR have also been reported in serine proteinases from other fungi (Kingsnorth et al. 2001). Determination of the polyadenylation signal sequence at the 3'-UTR was rather difficult because it did not follow the general consensus sequence AAUAAA (Mogen et al. 1990). Plant genes have also been reported to have no sharply defined consensus sequence for polyadenylation signal (Li and Hunt, 1995). We speculated that mushroom genes might be similar with regard to their polyadenylation signals. The polyadenylation sites for all the three cDNAs started with a pyrimidine nucleotide, as also observed in *A. bisporous* Spr1 cDNA sequences (Kingsnorth et al. 2001). Among the three cDNAs, Clone 5-2 contained the fewest mutations and was considered to be the most representative full-length cDNA (Figure 1). The predicted ORF of the cloned protein consisted of 379 amino acids, including a pro-peptide of 91 amino acids. The pro-peptide sequence was determined by comparing the ORF with the known N-terminal sequence of the mature protein. No signal peptide sequence was identified in the pro-peptide region by the method described by Nielsen et al. (1997).

A database search using PSI-BLAST (Altschul et al. 1997) revealed about 800 proteins showing at least partial homology with the predicted ORF (data not shown), almost all of which belonged to the serine proteinase family. Not a single aminopeptidase was identified among the 800 proteins. The highest identity of 42%, was found with *Agaricus bisporous* serine proteinase. Similarity was most prominent within the mature part of the protein with lower levels observed in pre- and pro-regions. Conserved Domain Database analysis (NCBI) showed that the predicted protein belonged to the subtilase family (family S8) of serine proteinases. Catalytic triad analysis showed that it contained identical catalytic residues, *i.e.*, Asp138, His 169 and Ser 323, to subtilisin Carlsberg from *Bacillus lechiniformis* (Jacobs et al. 1985). These results indicated that we have cloned a serine proteinase and not the aminopeptidase gene from *Grifola frondosa*.

Serine proteinases belong to a multifamily gene (Gan et al. 2000). The predicted amino acid sequence of the 140 bp gene fragment that was obtained in the first step of cDNA cloning showed 93% homology with the target enzyme. Only 3 of the 46 predicted amino acids, which accounted for the remaining 7% non-identical amino acid residues, were initially thought to have been erroneously introduced during protein sequencing. However, after cloning the full-length cDNA we found that the fragment was most probably derived from a homologous gene and not from the target gene. In fact, we have already cloned the cDNA of another serine proteinase-homologous gene (unpublished), which is 80% homologous with this protein at the predicted amino acid level.

Exon-intron boundaries of the cloned full-length genomic DNA were identified by the loss of sequence identity between cDNA and genomic nucleotide sequences and also by the presence of consensus donor and acceptor signals at the points of divergence, which were in accordance with GT-AG rule (Breathnach and Chambon, 1981). Intron 2 was the only exception, with GC at the donor site instead of GT. This may be a mutation that could be verified by sequencing several clones. Unfortunately, we only successfully isolated one clone for this protein.

N-terminal amino acid sequencing of the expressed pro-peptide, mature peptide and pro-protein has clearly indicated that the correct enzyme was expressed but no activity was detected by using the substrate, N-Succinyl-Ala-Ala-Pro-Phe-pNA. The pro-peptides of subtilisin family of serine proteases are known to exhibit inhibitory activity towards the cognate protease in addition to functioning as an intra-molecular chaperon (Hu et al. 1996; Kojima et al. 1997). Although it is initially potent, the pro-peptide is gradually degraded by subtilisin through a specific intermediate (Hu et al. 1996) and thus has been termed a temporary inhibitor. However, we found the pro-peptide of the expressed protein remained intact even after prolonged incubation at 30ºC for 24 hrs and no activity was detected (data not shown). The refolding analysis also failed to activate the enzyme activity. At this stage, the enzyme was analyzed with aminopeptidase specific substrate, Leu-pNA and several co-factors such as magnesium, manganese, and cobalt. None of the enzyme activation methods worked (data not shown). However, the enzyme was activated only after treatment with a trace amount of another proteolytic enzyme, subtilisin (Figure 5a). The SDS-PAGE analysis (Figure 5b) has clearly indicated that the pro-peptide was inhibiting the enzyme activity. After the hydrolysis of the pro-peptide into smaller peptides by subtilisin, the enzyme activity was revived. That the inhibiting pro-peptide needs a second serine proteinase for its hydrolysis and activation of the enzyme is the unique property of this enzyme.

The nucleotide sequence information, as well as, the expressed enzyme activity has undoubtedly proved that the cloned enzyme is a serine proteinase and not an aminopeptidase. It is very hard to explain why Nishiwaki and Hayashi (2001) found aminopeptidase activity in a purified protein which has 100% N-terminal amino acid sequence homology (up to 65 residues) with a serine proteinase. There are two explanations; either a contaminating aminopeptidase had contributed the enzyme activity in the purified but inactive serine proteinase or the incorrect serine proteinase was mistakenly sequenced for the characterized aminopeptidase. Regardless of the explanation, our findings have clearly indicated that the reported N-terminal amino acid sequence was derived from a serine proteinase, both structurally and functionally and not from the reported aminopeptidase. Hence they do not belong to the same protein entity.

REFERENCES

ALBISTON, A.L.; MCDOWALL, S.G.; MATSACOS, D.; SIM, P.; CLUNE, E.; MUSTAFA, T.; LEE, J.; MENDELSOHN, F.A.; SIMPSON, R.J.; CONNOLLY, L.M. and CHAI, S.Y. Evidence that the angiotensin IV (AT(4)) receptor is the enzyme insulin-regulated aminopeptidase. *Journal of Biological Chemistry*, December 2001, vol. 276, no. 52, p. 48623-48626.

ALTSCHUL, Stephen F.; MADDEN, Thomas L.; SCHÄFFER, Alejandro A.; ZHANG, Jinghui; ZHANG, Zheng; MILLER, Webb and LIPMAN, David J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, July 1997, vol. 25, no. 17, p. 3389-3402.

BLABER, S.I.; CIRIC, B.; CHRISTOPHI, G.P.; BERNETT, M.J.; BLABER, M.; RODRIGUEZ, M. and SCARISBRICK, I.A. Targeting kallikrein 6 proteolysis attenuates CNS inflammatory disease. *FASEB Journal,* May 2004, vol. 18, no. 7, p. 920-922.

BREATHNACH, R. and CHAMBON, P. Organization and expression of eukaryotic split genes coding for proteins. *Annual Review of Biochemistry*, 1981, vol. 50, p. 349-383.

DIAMANDIS, E.P. and YOUSEF, G.M. Human tissue kallikrein gene family: a rich source of novel disease biomarkers. *Expert Review of Molecular Diagnostics*, July 2001, vol. 1, no. 2, p. 182-190.

GAN, L.; LEE, I.; SMITH, R.; ARGONZA-BARRETT, R.; LEI, H.; MCCUAIG, J.; MOSS, P.; PAEPER, B. and WANG, K. Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region. *Gene,* October 2000, vol. 257, no. 1, p. 119-130.

GOTO, Y.; HATTORI, A.; ISHII, Y. and TSUJIMOTO, M. Reduced activity of the hypertension-associated Lys528Arg mutant of human adipocyte-derived leucine aminopeptidase (A-LAP)/ER-aminopeptidase-1. *FEBS Letters,* March 2006a, vol. 580, no. 7, p. 1833-1838.

GOTO, Y.; HATTORI, A.; ISHII, Y.; MIZUTANI, S. and TSUJIMOTO, M. Enzymatic properties of human aminopeptidase A. Regulation of its enzymatic activity by calcium and angiotensin IV. *Journal of Biological Chemistry*, August 2006b, vol. 281, no. 33, p. 23503-23513.

HORIKOSHI, K. Alkaliphiles: some applications of their products for biotechnology. *Microbiology and Molecular Biology Reviews,* December 1999, vol. 63, no. 4, p. 735-750.

HU, Z.; HAGHJOO, K. and JORDAN, F. Further evidence for the structure of the subtilisin propeptide and for its interactions with mature subtilisin. *Journal of Biological Chemistry*, February 1996, vol. 271, no. 7, p. 3375-3384.

JACOBS, M.; ELIASSON, M.; UHLEN, M. and FLOCK, J.I. Cloning, sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*. *Nucleic Acids Research,* December 1985, vol. 13, no. 24, p. 8913-8926.

KALAFATIS, M.; EGAN, J.O.; VAN'TVEER, C.; CAWTHERN, K.M. and MANN, K.G. The regulation of clotting factors. *Critical Review in Eukaryotic Gene Expression,* 1997, vol. 7, no. 3, p. 241-280.

KINGSNORTH, C.S.; EASTWOOD, D.C. and BUTTON, K.S. Cloning and postharvest expression of serine proteinase transcripts in the cultivated mushroom *Agaricus bisporus*. *Fungal Genetics and Biology*, April 2001, vol. 32, no. 3, p. 135-144.

KINI, R.M. Serine proteases affecting blood coagulation and fibrinolysis from snake venoms. *Pathophysiology of Haemostasis and Thrombosis,* 2005, vol. 34, no. 4-5, p. 200-204.

KOJIMA, S.; MINAGAWA, T. and MIURA, K. The propeptide of subtilisin BPN' as a temporary inhibitor and effect of an amino acid replacement on its inhibitory activity. *FEBS Letters,* July 1997, vol. 411, no. 1, p. 128-132.

LEATHERBARROW, Robin J. *GraFit*, Version 3.09b, UK, Erithacus Software Limited, 1996.

LI, Q. and HUNT, A.G. A near-upstream element in a plant polyadenylation signal consists of more than six nucleotides. *Plant Molecular Biology,* August 1995, vol. 28, no. 5, p. 927-934.

MARSTON, F.A.O. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochemical Journal,* November 1986, vol. 240, no. 1, p. 1-12.

MARUYAMA, M.; HATTORI, A.; GOTO, Y.; UEDA, M.; MAEDA, M.; FUJIWARA, H. and TSUJIMOTO, M. Laeverin/aminopeptidase Q, a novel bestatin-sensitive leucine aminopeptidase belonging to the M1 family of aminopeptidases. *Journal of Biological Chemistry,* July 2007, vol. 282, no. 28, p. 20088-20096.

MOGEN, B.D.; GRAYBOSCH, R.; MCDONALD, M. and HUNT, A.G. Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. *Plant Cell*, December 1990, vol. 2, no. 12, p. 1261-1272.

NIELSEN, H.; ENGELBRECHT, J.; BRUNAK, S. and VON HEIJNE, G. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *International Journal of Neural Systems,* October-December 1997, vol. 8, no. 5-6, p. 581-599.

NISHIWAKI, Thoshikazu and HAYASHI, Kiyoshi. Purification and characterization an aminopeptidase from edible basidiomycete *Grifola frondosa*. *Bioscience Biotechnology and Biochemistry,* 2001, vol. 65, no. 2, p. 424-427.

OSADA, T.; WATANABE, G.; SAKAKI, Y. and TAKEUCHI, T. Puromycin-sensitive aminopeptidase is essential for the maternal recognition of pregnancy in mice. *Molecular Endocrinology,* June 2001, vol. 15, no. 6, p. 882-893.

REAUX, A.; FOURNIE-ZALUSKI, M.C.; DAVID, C.; Zini, S.; ROQUES, B.P.; CORVOL, P. and LLORENS-CORTES, C. Aminopeptidase A inhibitors as potential central antihypertensive agents. *Proceedings of the National Academy of Sciences of the United States of America,* November 1999, vol. 96, no. 23, p. 13415-13420.

ROOPRAI, H.K. and MCCORMICK, D. Proteases and their inhibitors in human brain tumours: a review. *Anticancer Research,* November-December 1997, vol. 7, no. 6B, p. 4151-4162.

SAEKI, K.; OZAKI, K.; KOBAYASHI, T. and ITO, S. Detergent alkaline proteases: enzymatic properties, genes, and crystal structures. *Journal of Bioscience and Bioengineering*, 2007, vol. 103, no. 6, p. 501-508.

SALAMONSEN, L.A. Role of proteases in implantation. *Reviews of Reproduction,* January 1999, vol. 4, no. 1, p. 11-22.

SARIC, T.; CHANG, S.C.; HATTORI, A.; YORK, I.A.; MARKANT, S.; ROCK, K.L.; TSUJIMOTO, M. and GOLDBERG, A.L. An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nature Immunology,* December 2002, vol. 3, no. 12, p. 1169-1176.

SATO, Y. Role of aminopeptidase in angiogenesis. *Biological and Pharmaceutical Bulletin,* June 2004, vol. 27, no. 6, p. 772-776.

SERWOLD, T.; GONZALEZ, F.; KIM, J.; JACOB, R. and SHASTRI, N. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature*, October 2002, vol. 419, no. 6906, p. 480-483.

TANG, B.; NIRASAWA, S.; KITAOKA, M. and HAYASHI, K. In vitro stepwise autoprocessing of the proform of pro-aminopeptidase processing protease from *Aeromonas caviae* T-64. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, April 2002, vol. 1596, no. 1, p. 16-27.

TAYLOR, A. Aminopeptidases: structure and function. The *FASEB Journal,* 1993, vol. 7, p. 290-298.

TURGEON, V.L. and HOUENOU, L.J. The role of thrombin-like (serine) proteases in the development, plasticity and pathology of the nervous system. *Brain Research Review*, September 1997, vol. 25, no. 1, p. 85-95.

VOSWINKEL, J.; GOTTSCHLICH, S. and CSERNOK, E. Human proteinase 3 (PR3) and its binding molecules: implications for inflammatory and PR3-related autoimmune responses. *Annals of the New York Academy of Sciences,* August 2007, vol. 1109, p. 84-92.

WHITCOMB, D.C. and LOWE, M.E. Human pancreatic digestive enzymes. *Digestive Diseases and Sciences,* January 2007, vol. 52, no. 1, p. 1-17.

WRIGHT, J.W.; MIZUTANI, S.; MURRAY, C.E.; AMIR, H.Z. and HARDING, J.W. Aminopeptidase-induced elevations and reductions in blood pressure in the spontaneously hypertensive rat. *Journal of Hypertension*, October 1990, vol. 8, no. 10, p. 969-974.

YOON, H.; LAXMIKANTHAN, G.; LEE, J.; BLABER, S.I.; RODRIGUEZ, A.; KOGOT, J.M.; SCARISBRICK, I.A. and BLABER, M. Activation profiles and regulatory cascades of the human kallikrein-related peptidases. *Journal of Biological Chemistry,* November 2007, vol. 282, no. 44, p. 31852-31864.

YOUSEF, G.M. and DIAMANDIS, E.P. Human tissue kallikreins: a new enzymatic cascade pathway? *Biological Chemistry,* July-August 2002, vol. 383, no. 7-8, p. 1045-1057.