Available online at www.sciencedirect.com
SCIENCE DIRECT
Protein Expression and Purification 47 (2006) 551–561
Protein Expression & Purification
www.elsevier.com/locate/yprep

Cloning, sequencing, expression, and characterization of protealysin, a novel neutral proteinase from *Serratia proteamaculans* representing a new group of thermolysin-like proteases with short N-terminal region of precursor

Ilya V. Demidyuk [a,*], Alexander E. Kalashnikov [a], Tatiana Yu. Gromova [b], Eugene V. Gasanov [a], Dina R. Safina [a], Maria V. Zabolotskaya [a], Galina N. Rudenskaya [c], Sergey V. Kostrov [a]

[a] *Laboratory of Protein Engineering, Institute of Molecular Genetics, Russian Academy of Sciences, Kurchatov Sq. 2, Moscow 123182, Russia*
[b] *Department of Biotechnology, Lomonosov Moscow State Academy of Fine Chemical Technology, Moscow 119571, Russia*
[c] *Chemistry Department, Lomonosov Moscow State University, Moscow 119992, Russia*

Received 19 October 2005, and in revised form 6 December 2005
Available online 5 January 2006

Abstract

The gene of *Serratia proteamaculans* proteinase, protealysin, was cloned, sequenced, and expressed in *Escherichia coli*. The gene encoded a precursor of 341 amino acids (AAs) with a significant homology to thermolysin-like proteinases (TLPs). The molecular weight of the purified mature active recombinant protein was 32 kDa, the N-terminal amino acid sequence was AKTSTGGEVI. The optimum pH for azocasein hydrolysis by protealysin was seven and it was completely inhibited by *o*-phenanthroline. The enzyme hydrolyzed 3-(2-furyl)acryloyl-glycyl-L-leucine amide, the standard substrate for TLPs, with k_{cat}/K_m ratio of $(2.52 \pm 0.02) \times 10^2\ M^{-1}\ s^{-1}$. Protealysin maturation removes 50 AA from the N-terminus of the precursor. The removed region had no similarity with the preprosequence of thermolysin (232 AA) but was homologous to some other TLPs. These proteins shared a conserved 7-AA motif near the initial methionine. Such motif was also found in some nonproteolytic putative proteins; two of them were eukaryotic.

Keywords: Thermolysin-like proteinase; Protein precursor; Signal peptide; Propeptide; Protein motif; Molecular cloning; Gene expression; Protein purification

Peptidase family M4 [1]—thermolysin-like neutral metalloproteinases (TLPs)[1]—is one of the best studied groups of enzymes. Presently, over 100 primary structures of proteins of this family are available [2], several crystal structures were determined [3–6], and ample data were obtained on the structural and functional relationships in the TLP molecules [7,8]. Peptidases of the M4 family were found in about 70 bacterial species of the phyla Firmicutes, Proteobacteria, Actinobacteria, Cyanobacteria, and Bacteroidetes. In addition, TLPs are known in the Fungi and archaebacterium *Methanosarcina acetivorans* [2]. At the same time, the bacillary enzymes remain the best studied despite the wide range of organisms where proteinases of the M4 family were found.

Similar to many other bacterial proteinases, TLPs are essentially synthesized as a precursor containing a propeptide removed during formation of the mature active protein [9,10]. The mechanisms of propeptide activity in the M4

* Corresponding author. Fax: +7 495 196 02 21.
E-mail address: duk@img.ras.ru (I.V. Demidyuk).

[1] *Abbreviations used:* AAs, amino acids; TLPs, thermolysin-like proteinases; IPTG, isopropyl-β-D-thiogalactopyranoside; FAGLA, *N*-[3-(2-furyl)acryloyl]-glycyl-L-leucine amide; PMSF, phenylmethyl sulfonyl fluoride; EDTA, ethylenediaminetetraacetic acid; PMSF, phenylmethyl sulfonyl fluoride; Pln, protealysin; Eca, extracellular proteinase from *Erwinia carotovora* subsp. *carotovora*; Sma, minor protease from *Serratia marcescens*.

1046-5928/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2005.12.005

family are less well understood than those of mature enzymes of the family and propeptides of other proteinase groups. At the same time, propeptides of certain TLPs can act as intramolecular chaperones [10] that modulate protein folding [11–13]. In addition, the propeptide domains of M4 peptidases proved to inhibit the corresponding mature proteins [13–16] and to mediate their secretion [17,18].

Previously, we have isolated *Serratia proteamaculans* strain 94 from spoiled meat and characterized it as a proteolytic enzyme producer [19]. In this work, a metalloproteinase from this organism was studied. Its gene was cloned, sequenced, and expressed and the corresponding protein was isolated and described. This enzyme is the first isolated *S. proteamaculans* proteinase. (Although two proteolytic bands with approximate masses 52–48 kDa have been recently identified on a zymogram of *S. proteamaculans* culture supernatant [20].) No data on the structure and enzymatic properties of other proteolytic enzymes produced by this bacterium as well as on the structure of other proteinase genes of *S. proteamaculans* are presently available. According to the determined primary structure and characteristics, the studied metalloproteinase named protealysin is a typical member of the M4 family. At the same time, the N-terminal region of protealysin precursor has a particular structure.

It is interesting to note that *S. proteomaculans* has been previously described as an insect-pathogenic bacterium, which is the causative agent of the fatal disease of New Zealand grass grub *Costelytra zealandica* (Coleoptera: Scarabaeidae) [21]. Although *S. proteamaculans* is rarely found in clinical samples, it can cause disease in humans [22], thus, suggesting a pathogenic relevance of proteases from *S. proteamaculans*.

Materials and methods

Materials

In this work, we used oligonucleotides (Syntol, Russia), tris–(hydroxymethyl)aminomethane (Tris), ampicillin, azocasein, *N*-[3-(2-furyl)acryloyl]-glycyl-L-leucine amide (FAGLA) (Sigma, USA), Coomassie brilliant blue G-250 (Loba, Austria), bromophenol blue, glycine, ammonium persulfate, Coomassie brilliant blue R-250, bovine IgG (Reanal, Hungary), *Pfu* DNA polymerase, *Taq* DNA polymerase, T4 DNA ligase, restriction endonucleases and the corresponding buffers (MBI Fermentas, Lithuania), acrylamide, bisacrylamide (Promega, USA), isopropyl-β-D-thiogalactopyranoside (IPTG) (Amresco, USA), LE agarose (USB, USA), yeast extract, tryptone (Difco, USA), edible gelatin (Santus, Russia), and $CaCl_2$, NaCl (Merck, Germany). All other reagents were of reagent grade and were purchased from Dia-M (Russia).

General methods

Transformation of *Escherichia coli* cells with plasmid DNA was carried out as described in [23].

Plasmid DNA used for cloning and sequencing was isolated using alkaline lysis method [23].

N-Terminal amino acid sequence was determined by automated Edman degradation using a Model 477A Protein Sequencer (Applied Biosystems, USA) equipped with a Model 120A PTH Analyzer (Applied Biosystems, USA).

The nucleotide sequences were determined by the dideoxy chain termination method on an automated sequencer ABI-373A (Applied Biosystems, USA).

Polymerase chain reaction was performed using a Tertsik MS2 Multichannel DNA Amplifier (DNA-Technology, Russia).

Protein electrophoresis (SDS–PAGE) was performed in vertical 12.5% polyacrylamide gels with 0.1% sodium dodecylsulfate [24]. Proteins were visualized by Coomassie brilliant blue staining. Cellulase from *Aspergillus niger* (94.6 kDa), bovine serum albumin (66 kDa), albumin from chicken egg white (45 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa), trypsin inhibitor from soybean (20.1 kDa), and lysozyme from chicken egg white (14.4 kDa) (Helicon, Russia) were used as molecular weight standards.

Analytical gel permeation chromatography was performed on a Superdex 75 HR 10/30 column (Amersham Biosciences) in 50 mM Tris–HCl buffer containing 1 mM $CaCl_2$ and 0.15 M NaCl, pH 7.1, at a flow rate of 0.4 ml/min. Bovine serum albumin (66 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa), cytochrome C from horse heart (12 kDa), and aprotinin from bovine lung (6.5 kDa) (Sigma, USA) were used as molecular weight standards.

Protein concentration was assayed after Bradford [25] with modifications [26] using IgG as a standard.

Nucleic acids were separated by horizontal electrophoresis in 0.8–2% agarose gel using 89 mM Tris–borate buffer with 2 mM EDTA, pH 7.6. DNA was visualized by ethidium bromide staining.

Total proteolytic activity was evaluated by azocasein cleavage [27]: 100 μl of 1% solution of azocasein in 100 mM Tris–HCl buffer containing 5 mM $CaCl_2$, pH 7.1, was incubated with 50 μl of the enzyme for 15 min at 37 °C unless stated otherwise. The reaction was stopped by adding 200 μl of 10% trichloroacetic acid. After centrifugation at 9500*g* for 5 min, 250 μl of the supernatant was mixed with 50 μl of 5 M NaOH and the optical density was measured at 450 nm on a Model 2550 Microplate Reader (Bio-Rad, USA). The activity unit was defined as the amount of enzyme that changed the absorbance by 1 optical density unit per min.

Genomic library construction and screening

The total genomic DNA of *S. proteamaculans* strain 94 [19] was isolated following the procedure described earlier [28]. The genomic DNA (5 μg) was partially digested with the restriction endonuclease *Bsp*143I (*Sau*3AI). The reaction time and enzyme quantities were selected for the maximum yield of 5–10 kb fragments. The reaction was stopped by isopropanol precipitation. The restriction fragments were ligated into 8 μg *Bam*HI digested pUC19 vector. The ligated mixture was transformed into *E. coli* TG1 strain. The trans-