Original Article

Molecular cloning, expression, and anti-tumor activity of a novel serine protease from Arenicola cristata

Chunling Zhao¹ and Jiyu Ju²*

¹College of Pharmacy and Biological Science, Weifang Medical University, Weifang 261053, China
²College of Basic Medicine, Weifang Medical University, Weifang 261053, China
*Correspondence address. Tel: +86-536-8462053; Fax: +86-536-2602083; E-mail: juimmu@163.com

Arenicola cristata, a marine annelid, is a well-known and prized traditional Chinese medicine. However, the serine protease gene of A. cristata has not been cloned yet. In this study, a novel protease of A. cristata was cloned, sequenced, and expressed in Escherichia coli, and the functions of this recombinant protease were also investigated. The whole complementary DNA (cDNA) of this novel protease was of 980 bp in length and consisted of an open reading frame of 861 bp encoding 286 aa. Sequence analysis of the deduced amino acid sequence revealed that the protease belongs to the serine protease family. The active enzyme of the proposed A. cristata protease is composed of a signal peptide, a propeptide, and a mature polypeptide. The molecular weight of the recombinant mature protein was ∼26 kDa after over-expression in E. coli. The recombinant protein significantly inhibited cell growth and induced cell apoptosis of esophageal squamous cell carcinoma (ESCC) in vitro, and reduced tumorigenicity in vivo. Furthermore, administration of the recombinant protein led to the activation of caspase-9 as well as down-regulation of Mcl-1 and Bcl-2. Taken together, our findings indicated that the recombinant serine protease of A. cristata could inhibit ESCC cell growth by mitochondrial apoptotic pathway and might act as a potential pharmacological agent for ESCC therapy.

Keywords Arenicola cristata; serine protease; esophageal squamous cell carcinoma (ESCC); apoptosis; mitochondrial pathway

Received: December 12, 2013     Accepted: March 4, 2014

Introduction

Proteases are a class of enzymes that can catalyze the cleavage of peptide bonds in protein. These enzymes are widely used in many fields, i.e., in the food and pharmaceutical fields [1,2]. They are also involved in essential biological processes, such as embryonic development, tissue differentiation, immune response, and blood clotting [3–7]. Proteases have been isolated from diverse organisms, including animals [8,9], plants [10], and microorganisms [11,12]. Arenicola cristata, a well-known and prized traditional Chinese medicine, is a widely distributed polychaete annelid in the eastern coastland of China. Several investigators have focused on A. cristata [13–15]. Hemoglobin was isolated from A. cristata and its structure was characterized [13]. Parker and Lin [14] isolated four proteases from A. cristata that activate its cyclic AMP phosphodiesterase. Wang et al. [15] investigated the anti-proliferative activity of arenicolsterol A, a novel cytotoxic enolic sulfated sterol isolated from A. cristata on tumors.

However, up to date, the serine protease gene from A. cristata has not been cloned yet. In the present study, we reported the molecular cloning and sequencing of the novel protease gene from A. cristata, as well as its expression in E. coli, and the functions of the recombinant protease were also investigated.

Materials and Methods

Arenicola cristata, bacterial strains, and vectors
Arenicola cristata was collected in BoHai Bay (Qingdao, China) in October 2010. Escherichia coli strain JM109 and plasmid vector pGEM-T vector (Promega, Madison, USA) were used in the DNA manipulation. Strain BL21(DE3) (Invitrogen, Grand Island, USA) and plasmid vector pET-21a(+) were used for protein expression. Escherichia coli was grown in Luria broth (LB) medium with 100 μg/ml ampicillin.

Full-length cDNA cloning and sequencing
Total RNA was extracted from the digestive tract of A. cristata using Trizol reagent (Gibco BRL, Gaithersburg, USA) following the manufacturer’s instructions. All the primers used for generating the A. cristata protease were synthesized (SBS Genetech, Peking, China) and illustrated in Table 1. In order to get the full-length protease gene, rapid amplification
of cDNA ends (RACE) was done using the cDNA amplification kits (Genebiotech, Peking, China).

The 3' end flanking sequence of the protease gene was performed by 3'-RACE. The gene-specific 3'-RACE primer (3F/3GSP) was designed on the basis of the highly conserved amino acid sequence by alignment of the other proteases from annelids such as *Eisenia fetida* (DQ836917.1), *Eisenia fetida* (DQ418454.1), *Lumbricus rubellus* (U25647.1), and *Urechis unicinctus* (HM623463.1). As reverse transcription primer, 3'-RACE adapter primer (3AP) was used to conduct the first-strand cDNA synthesis. Using the first-strand cDNA as a template, polymerization chain reaction (PCR) amplification was performed using 3F (3GSP) as an upstream primer and 3R as a downstream primer. The PCR amplification was carried out. The initial denaturation reaction was performed at 94°C for 5 min and then 35 cycles were performed with denaturation (94°C, 40 s), annealing (60°C, 30 s), and extension (72°C, 1 min), followed by 72°C for 10 min.

The 5' flanking sequence of the protease gene was cloned by 5'-RACE. The gene-specific 5'-RACE primers (5R1/5GSP1, 5R2/5GSP2) were designed according to the obtained sequence of 3'-RACE fragment. As a reverse transcription primer, oligodT (5RP) was used to conduct the first-strand cDNA synthesis. The reverse transcription product was purified and the tail was added with dCTP and TdT. The nested PCR amplified dC-tailed cDNA using the adapter primer 5F1/5AP, 5F2 as forward primers, and 5R1/5GSP1, 5R2/5GSP2 as reverse primers. Two rounds of PCR amplifications were respectively performed as follows: 94°C for 5 min, 35 cycles with denaturation (94°C, 40 s), annealing (60°C, 30 s), and extension (72°C, 1 min), and the final extension at 72°C for 10 min.

All the above-mentioned PCR products were subcloned into the pGEM-T vector and were transformed into *E. coli* JM109 and sequenced. The determined nucleotide sequence of protease gene from *A. cristata* has been submitted to the GenBank database under accession number KF830712.

### Sequence analysis
DNAMAN version 5.2.2 software was used for full-length cDNA sequence assembly and analysis. The protease sequences of other annelid animals, such as *E. fetida*, *L. rubellus*, *Nereis aibuhitensis*, and *U. unicinctus*, were obtained from NCBI. Similar searches for sequences were performed with the NCBI BLAST program (http://www.ncbi.nlm.nih.gov). The signal peptide was predicted by Signal P 3.0 program (http://www.cbs.dtu.dk/services/SignalP). N-glycosylation sites of the protein were analyzed at http://web.expasy.org/glycomod. The alignment of amino acid sequences was done with the DNAMAN software.

### Phylogenetic analysis
The phylogenetic tree was constructed using a neighboring-joining method [16] and the MEGA 4.0 software [17] with 1000 bootstrap replicates.

### Construction of expression plasmid for the mature form of the protease gene
The mature form at aa position 58–286 of the protease was amplified by PCR using *A. cristata* cDNA obtained by using the protein reverse transcription primer (PRP/oligodT) as a template. The PCR primers PF and PR (*Table 1*, BamHI and Xhol sites are underlined) were designed on the basis of the determined nucleotide sequences. The amplified DNA fragment was digested with BamHI and XhoI, and ligated to pET-21a(+) vector which was digested with BamHI and XhoI. The recombinant plasmid was transformed into *E. coli* JM109 and subsequently confirmed by sequencing.

### Expression and purification of the recombinant protein
The recombinant plasmid was transformed into *E. coli* BL21(DE3) and cultured in the LB broth containing ampicillin, grown at 37°C with shaking until the absorbance at 600 nm reached 1.2. The cultures were then induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG)
to a final concentration of 1.0 mM, and were incubated at 30°C for another 4 h. Cells were harvested at 8000 g for 10 min at 4°C, suspended in lysis buffer (20 mM Tris-HCl, pH 7.4, 10 mM imidazole, 0.5 M sodium chloride), and disrupted by sonication. The insoluble protein pellets and the supernatants were then separated, and the insoluble protein pellets were resuspended in the buffer. The supernatants and the insoluble protein pellets were used for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) analysis. The protease was purified from the supernatant by adding 1 ml of a 50% Ni-NTA slurry (Qiagen, Frankfurt, Germany) pre-equilibrated in lysis buffer to 4 ml of the lysate. After incubation for 1 h at 4°C, the lysate–Ni-NTA mixture was loaded onto a column. Unbound proteins were removed by washing the column with two column volumes of lysis buffer. The column was washed twice with 5 ml of 20 mM Tris-HCl (pH 7.4) containing 20 mM imidazole and 0.5 M sodium chloride. Finally, the recombinant protein was eluted with four column volumes of elution buffer. The purified recombinant protein was quantified using the Bradford assay reagents (Bio-Rad, Hercules, USA). Fractions were analyzed by SDS–PAGE.

Cell culture
Human esophageal squamous cell carcinoma (ESCC) cell lines Eca-109, TE-8, TE-10, and TE-15, which were purchased from the Shanghai Cell Bank, Chinese Academy of Sciences (Shanghai, China), were grown in RPMI 1640 (Invitrogen) supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 mg/ml streptomycin, maintained at 37°C in a humidified atmosphere of 5% CO2.

Cell proliferation assay
The cell proliferation was measured by a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) assay. The Eca-109, TE-8, TE-10 and TE-15 cells were respectively seeded in 96-well flat-bottomed plates at 3 × 10^3 cells/well in RPMI 1640 with 10% fetal bovine serum. After culturing for 24 h, the medium in the wells was replaced with the new medium containing the recombinant protease at final concentrations of 2.5–30 μg/ml. Each concentration was set for four wells. After 72 h, cells were stained with 100 μl of 0.5 mg/ml MTT (Sigma, St Louis, USA) for 4 h at 37°C, then culture medium was removed and 200 μl of dimethylsulfoxide was added and completely lysed for 20 min. The absorbance was measured at 570 nm using a microplate reader (Bio-Rad).

Colony formation assay
The recombinant protease-treated cells were trypsinized and 400 cells were inoculated into 60-mm plates. Cultures were maintained for 2 weeks, and the cells were washed with phosphate-buffered saline (PBS) and fixed with methanol and 0.1% crystal violet. The colonies were manually counted and then photographed.

Apoptosis morphology analysis
Cells were grown on coverslips in RPMI 1640 complete medium for 24 h. The cells were changed with new medium containing the recombinant protease at a final concentration of 20 μg/ml. After incubation for 48 h, the coverslips were rinsed twice with ice-cold PBS and fixed and permeabilized by immersion in cold methanol for 20 min. Then, the coverslips were rinsed with PBS and stained with 0.1 μg/ml 4′,6′-diamidino-2-phenylindole (DAPI) (Sigma). The slides were examined with an ultraviolet confocal microscopy system.

Assessment of apoptosis by FACS
After 48 h of induction by the recombinant protease (at a final concentration of 20 μg/ml), cells were harvested for apoptosis staining using the Annexin V-FITC apoptosis kit (Sigma) and propidium iodide apoptosis kit (Becton Dickinson, New York, USA). Samples were analyzed with a fluorescence-activated cell sorter (FACS) (Becton Dickinson).

Tumorigenicity in nude mice
Four-week old BALB/c nude mice were injected subcutaneously with 1 × 10^6 TE-10 cells in 0.1 ml of PBS into the right upper back. During the subsequent 4 weeks, the mice were administered orally with 800 μg/kg recombinant protease once each day, while the mice in the negative control group received only saline. Every week, the mice were examined for tumor formation at the sites of injection. Tumors were measured with linear calipers, and tumor volumes were calculated by the formula 4πr_1^2r_2/3 (r_1 < r_2). On the 28th day of the experiment, all mice were killed, and the tumor volumes and weights were measured. Tumor tissues were kept at −80°C or in formalin for further analysis.

Cellular protein extraction and western blot analysis
Cells were washed with ice-cold PBS and lysed in an RIPA buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% SDS, 1% NP40) containing PMSF (100 μg/ml) and protease inhibitors (2 μg/ml aprotinin and 2 μg/ml leupeptin) on ice for 30 min. The lysates were centrifuged at 12,000 g and 4°C for 20 min. The supernatant was collected and the total protein concentration was estimated using the Bradford assay reagents (Bio-Rad). Cellular proteins (60–100 μg) were separated using 10% SDS–PAGE and were then transferred to Immobilon-P membranes (Millipore, Bedford, USA). Membranes were blocked in 5% milk, washed with PBS with 0.1% Tween (PBST), and incubated with the indicated antibodies for 1 h in PBS containing 5% milk, washed with PBS, incubated subsequently with secondary antibodies for 1 h. Finally, the protein bands were visualized with the
enhanced chemiluminescent reagent (ECL). Antibodies including anti-cleaved poly ADP-ribose polymerase (PARP) (1:1000), anti-procaspase-3 (1:1000), anti-procaspase-8 (1:1000), anti-procaspase-9 (1:1000), anti-Bcl-2 (1:500), anti-Mcl-1 (1:500), and anti-β-actin (1:5000) were purchased from Santa Cruz Biotechnology (California, USA).

Statistical analysis
Data were presented as mean ± SD. Statistical analysis was performed with SPSS12.0 using Student’s t-test. The difference was considered statistically significant when P < 0.05.

Results

Cloning and sequencing result analysis
The performance of RACE has led to the identification of the full-length cDNA encoding the serine protease from A. cristata (Table 1 and Fig. 1). The specific primer (3F) for 3’-RACE was designed on the basis of the highly conserved amino acid sequence of the serine proteases from other annelid animals, such as E. fetida, L. rubellus, N. aibuhiten-sis, U. unicinctus (Fig. 1A). After 3’-RACE and 5’-RACE,

Figure 1. cDNA cloning of the A. cristata protease gene by RACE (A) The designed region for primer 3F in 3’-RACE by alignment of the other proteases from annelid. DNAMAN 1, 2, 3, 4 stands for E. fetida (DQ836917.1), E. fetida (DQ418454.1), L. rubellus (U25647.1), and U. unicinctus (HM623463.1), respectively. (B) The amplification of the A. cristata protease gene by RACE. Lanes 1 and 4, DNA markers; lane 2, the amplified product of 3’-RACE; lane 3, the amplified product of 5’-RACE.

Figure 2. Nucleotide and deduced amino acid sequences of cDNA encoding the serine protease from A. cristata The nucleotide and amino acids are numbered on the left. The putative signal peptide and prepeptide are underlined and double underlined, respectively. The conservative domain is indicated in boldface, and the conservative catalytic triad (histidine, aspartic acid, and serine) are indicated in italic and boldface. The stop codon is indicated with an asterisk.

A novel serine protease from A. cristata


Downloaded from https://academic.oup.com/abbs/article-abstract/46/6/450/1348 by guest on 14 April 2019
two kinds of cDNA fragments were obtained. The resulting fragments of 3′-RACE and 5′-RACE were ~400 and 720 bp (Fig. 1B, lanes 2 and 3). The full-length cDNA was assembled with 980 bp. The cDNA sequence analysis showed an open reading frame (ORF) of 861 bp encoding 286 aa. There was a 5′-end non-coding sequence of 59 bp and a 3′-end non-coding sequence of 43 bp followed by a 17-bp polyA tail (Fig. 2).

Based on the analysis of the conserved sequence with other proteases by BLASTp, this novel protease could be
assigned to trypsin-like serine protease family. As shown in Fig. 2, we proposed that the active form of this *A. cristata* protease consisted of a 16-aa-long signal peptide (from aa 1 to 16), a 41-aa-long propeptide (from aa 17 to 57), and a 229-aa-long mature polypeptide (from aa 58 to 286). The deduced amino acid sequence for the polypeptide enzyme had a potential N-glycosylation site at aa 137–139 and the conserved catalytic triad (H98, D146, and S239) of the serine proteases [18]. The aspartate at aa position 233 corresponded to the bottom of the primary substrate specificity pouch in serine proteases [19] and is compatible with a trypsin-like primary substrate specificity. The *A. cristata* protease also contained a highly conserved N-terminal amino acid sequence (aa 58–61: Ile-Val-Cys-Gly) in the active form [20], and the highly conserved sequence GDSGGP which contained a serine protease active site at aa 239.

In addition, no significant homologies were found for the amino acid sequence of the *A. cristata* protease. Comparison of the deduced peptide sequence of *A. cristata* protease gene with other serine proteases showed identities of 40%. In comparison with serine proteases from the annelid creature, *A. cristata* protease showed an identity of 39%, 38%, 40%, and 36%, to that of *E. fetida* (ABW04905.1), *L. rubellus* (AAQ13829.1), *N. aibuhitensis* (ACL12061.1), and *U. unicinctus* (ADL28819.1), respectively (Fig. 3), which was confirmed by the phylogenetic analysis (Fig. 4).

Expression and purification of the recombinant protein

The ORF encoding the mature form of the *A. cristata* protease was amplified by specific primers PF and PR (Table 1), and then cloned into prokaryotic expression vector pET-21a(+). The recombinant plasmid was transformed into *E. coli* BL21(DE3), and the over-expression was induced by 1.0 mM IPTG (Fig. 5, lane 3). The recombinant protein mainly existed in inclusion body and a little in soluble protein (Fig. 5, lanes 5 and 4). The His-tag fused recombinant protease was purified by Ni-charged affinity chromatography. SDS–PAGE analysis showed that the purified product appeared as a single band with the apparent molecular mass of \( \sim 26 \text{kDa} \) (Fig. 5, lane 6), which was basically consistent with that calculated from the deduced amino acid sequence. This indicated that the mature protein was not modified in *E. coli*.

The recombinant protease inhibits ESCC cell growth

To study the effect of the recombinant protease on ESCC cell growth, cell proliferation assays were performed using Eca-109, TE-8, TE-10, and TE-15 cell lines. After the recombinant protease treatment at several concentrations for 72 h, the survival rates of ESCC cells were examined (Fig. 6A). The survival rates of all these cell lines were
decreased in proportion to the concentration of the recombinant protease although with some difference. Among them, the cell growth of TE-10 was inhibited most significantly, and this cell line was used for further study. The colony formation assay verified that the recombinant protease could effectively reduce foci formation ($P < 0.05$) (Fig. 6B).

To further explore the in vivo tumor-suppressive ability of the recombinant protease, nude mice with tumors were orally administrated with the recombinant protease. The results (Fig. 6C) showed that tumor formation in nude mice was significantly inhibited in the recombinant protease-treated group compared with the control group ($P < 0.05$, week 3 or 4). The tumor weight on day 28 was significantly decreased (up to 48.7%) in the recombinant protease-treated group compared with the control group ($P < 0.05$) (Fig. 6D). These findings indicated that the recombinant protease could suppress the cell growth of ESCC in vitro and in vivo and had tumor-suppressive ability.

The recombinant protease induced ESCC cell apoptosis

In view that tumor growth is determined by the balance of cell proliferation and programmed cell death, the potential role of the recombinant protease in apoptosis was analyzed by nucleus staining. As shown in Fig. 7A, compared with control cells, the nuclei in ESCC cells with the recombinant protease administration were either relatively small, chromatin condensed or fragmented. More importantly, the apoptotic body was also observed. Furthermore, FACS results showed that the apoptotic indexes were significantly increased in the recombinant protease-treated cells compared with control cells ($P < 0.05$) (Fig. 7B). Thus, decreased tumorigenesis in vivo may partly correlate with ESCC cell apoptosis induced by the recombinant protease.

Because apoptosis is usually mediated by the activation of caspases that result in the death substrates, such as PARP, binding to fragmented DNA. Western blot was then conducted to detect caspase activation. Figure 7C shows that the cleavages of caspase-3 and PARP were dramatically increased in TE-10 cells after the recombinant protease treatment, compared with control cells. To further determine which signaling pathway was mainly involved in the recombinant protease-inducing apoptosis, the activation of caspase-9 (mitochondria pathway) and caspase-8 (death receptor pathway) after the protease induction was examined. Caspase-9 is activated and the cleavage of caspase-9 was increased. However, no obvious changes in caspase-8 were observed between the protease-treating cells and control.
cells (Fig. 7D). These data indicated that the mitochondria apoptotic pathway might be responsible for the diminished tumorigenicity of the recombinant protease administration. Furthermore, we observed that two negatively regulating molecules Bcl-2 and Mcl-1, protein members of Bcl-2 family, were down-regulated (Fig. 7E).

**Discussion**

In this work, we presented the full-length cDNA sequence encoding a novel serine protease from *A. cristata*. The sequence was obtained by RACE from the digestive tract of *A. cristata*. The amino acid sequence deduced from the cDNA could be assigned to trypsin-like serine protease family. The active form of the *A. cristata* protease was proposed to consist of a 16-aa-long signal peptide, a 41-aa-long propeptide and a 229-aa-long mature polypeptide. The protease was also predicted to have a potential N-glycosylation site at aa 137–139.

It has been reported that the mature extracellular protease was formed after N-glycosylation, signal peptide cleavage, and other modifications of the proproteose. *Arenicola cristata* protease might be modified by N-glycosylation in vivo, although in our study the recombinant protease was expressed in *E. coli* and showed no N-glycosylation. The proposed sequence of the active enzyme contained the conserved sequence of the catalytic triad (H98, D146, and S239) typical of serine proteases of chymotrypsin family [19]. The amino acid at the N-terminal position 234 of the active serine is an aspartate residue, indicating that the protease might have a trypsin-like primary substrate specificity [19,21]. *Arenicola cristata* protease also contains a highly conserved N-terminal amino acid sequence (aa 58–61: Ile-Val-Cys-Gly) in the active form. Trypsin-like serine proteases are usually synthesized as proproteases that contain an N-terminal signal peptide (11–15 aa), and have a highly conserved N-terminal amino acid sequence (Ile-Val-Gly-Gly) in the active form [20].

Homology search was performed, but no significant homologies were found between the *A. cristata* protease and other serine proteases. Phylogenetic analysis further confirmed that *A. cristata* protease gene obtained in this study is a novel gene and the protease it encodes might be a new member of serine protease family.

The recombinant *A. cristata* protease was purified to homogeneity by affinity chromatography. It appeared as a single band on SDS–PAGE stained by Coomassie brilliant blue R250, confirming that by this purification procedure, this protease could be isolated from *E. coli*.

Cell proliferation assay, colony formation assay, and tumorigenicity in nude mice proved that the recombinant protease could suppress the growth of ESCC cells in vitro.
and in vivo. Furthermore, this tumor-suppressive ability was related to the apoptosis of ESCC cells.

As another annelid animal, earthworm has been focused early in 1980s. Earthworm proteases could directly kill the cancer cells in vitro [22]. Earthworm proteases also enhanced the curative effects by both radiation therapy and chemotherapy [23]. Recently, an apoptosis-related serine protease from earthworm was purified and identified. The earthworm protease had obvious anti-tumor activity in hepatoma cells in vitro and in vivo. The earthworm protease could induce apoptosis of hepatoma cells and down-regulate the expression of MMP-2 [24]. Neanthes japonica also belongs to annelid animals. An acidic serine protease purified from N. japonica could inhibit proliferation, induce apoptosis, and enhance chemo-susceptibility of acute promyelocytic leukemia cells [25]. Therefore, A. cristata protease showed a similar anti-tumor activity with proteases from earthworm and N. japonica.

To further confirm the induction of apoptosis after the recombinant protease administration, the activation of caspase-3 and its substrate were detected. A sharp decrease in procaspase-3 and the cleavage of PARP were found. Apoptosis could be induced by either the intrinsic/mitochondrial pathway or the extrinsic/death receptor pathway. In our study, the recombinant protease induced apoptosis through the mitochondrial pathway, since the recombinant protease administration resulted in caspase-9, but not caspase-8, activation.

In addition, down-regulation of Bcl-2 and Mcl-1 was found. Members of the Bcl-2 family are key messengers for delivering the apoptotic signal to the mitochondria. As a correlative link in the initiation of mitochondrial apoptosis induced by the recombinant protease administration, down-regulation of Bcl-2 and Mcl-1 protein level was followed by activation of caspase-9. The detailed molecular mechanism by which the recombinant protease down-regulates Bcl-2 and Mcl-1 remains to be elucidated in further studies.

Taken together, a novel serine protease gene with anti-tumor activity was cloned and expressed from A. cristata. This recombinant protease resulted in ESCC cell growth inhibition and apoptosis via the mitochondrial pathway. The recombinant protease might be a potential pharmacological agent in ESCC therapy.

### Funding

This work was in part supported by grants from the National Natural Science Foundation of China (30901779) and the Natural Science Foundation of Shandong, China (ZR 2009CM019).

### References


