Molecular Cloning, Sequencing, and Expression of a Fibrinolytic Serine-protease Gene from the Earthworm *Lumbricus rubellus*

Il Hwan Cho[†], **Eui Sung Choi**[‡] **and Hyung Hoan Lee**[*]

Department of Biological Sciences, Konkuk University, Seoul 143-701, Korea
[†]Shinpoong Pharmaceutical Co., Moknae-dong, Ansan, Kyunggido 425-100, Korea
[‡]Lab of Microbial Functions, Korea Research Institute of BioSciences and Biotechnology, Daejon 305-333, Korea

The full-length cDNA of the lumbrokinase fraction 6 (F6) protease gene of *Lumbricus rubellus* was amplified using an mRNA template, sequenced and expressed in *E. coli* cells. The F6 protease gene consisted of pro- and mature sequences by gene sequence analysis, and the protease was translated and modified into active mature polypeptide by N-terminal amino acid sequence analysis of the F6 protease. The pro-region of F6 protease consisted of the 44 residues from methionine-1 to lysine-44, and the mature polypeptide sequence (239 amino acid residues and one stop codon; 720 bp) started from isoleucine-45 and continued to the terminal residue. F6 protease gene clones having pro-mature sequence and mature sequence produced inclusion bodies in *E. coli* cells. When inclusion bodies were orally administrated rats, generated thrombus weight in the rat's venous was reduced by approximately 60% versus controls. When the inclusion bodies were solubilized in pepsin and/or trypsin solutions, the solubilized enzymes showed hemolytic activity *in vitro*. It was concluded the F6 protease has hemolytic activity, and that it is composed of pro- and mature regions.

Keywords: Gene cloning, *Lumbricus rubellus*, Lumbrokinase, Protease

Introduction

Fibrinolytic proteases in the earthworm, *Lumbricus rubellus* have been purified and characterized by several investigators (Park *et al.*, 1989; Mihara *et al.,* 1991, 1993; Nakajima *et al.*, 1993; Jeon *et al.*, 1995; Cho *et al.*, 2004) and found to hydrolyze plasmogen-rich fibrin and plasmogen-free fibrin. Earthworm protease appeared a mixture of six iso-enzyme proteins of molecular weight 25 to 32 kDa. Fibrinolytic proteases dissolve blood fibrin clots, and are important chemotherapeutic agents (Mihara *et al.*, 1989; Ryu *et al.*, 1994, 1995; Park *et al.*, 1999).

The sequences of the F-III-2 and F-III-1 fraction iso-enzyme genes of *L. rubellus* were reported by Sugimoto and Nakajima (2001), and Sun *et al.* (2002) and Xu *et al.* (2002), reported the sequences of the lumbrokinase genes PI_{239} and PV_{242} of *L. bimastus*. The other four iso-proteases and differences between the two iso-proteases of *L. rubellus* produced by different strains have not been determined. Thus we undertook to clone and express the protease genes of a Korean *L. rubellus* strain, which it was hoped would be useful for clinical applications. The aims of this study were; (1) to clone and sequence one of the fibrinolytic protease genes, (2) to determine the levels of expression of the protease gene in *E. coli,* and (3) to examine the fibrinolytic activity of the cloned prokaryotic products.

Materials and Methods

Animals Earthworms (*Lumbricus rubellus*) obtained from the Giheung Farmer School, Korea, were used as the lumbrokinase source. A specific pathogen free, eight-week-old Sprague-Dawley male rat strain (300-350 g), was maintained at 23.1°C in 55.5% humidity with 10-18 aeration changes per hour under a 12 h light and dark cycle of 300-500 luxes, and used to examine blood clot hydrolysis.

Oligonucleotide primers Nine oligonucleotide primers recognizing the antisense and sense strands of the 5- and 3-ends of the fibrinolytic F6 protease coding sequence were used for PCR (94, 1 min; 60, 2 min; 72, 3 min; 35 cycles) to generate a clonable fragment encompassing only the coding region. These primers were synthesized at Bionics, Seoul, Korea, and are illustrated in Table 1.

*To whom correspondence should be addressed.
Tel: 82-2-450-3426; Fax: 82-2-452-9715
E-mail: hhlee@konkuk.ac.kr

Table 1. List of oligonucleotides

Primers	Sequences (5' → 3')
AP	5'-GGCCACGCGTCGACTAGTAC(T)$_{17}$3'
UAP	5'-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3'
GSP	5'-ATTGTCGGAGGAATTGAAGCCAGACCATACGAGTTC-3'
AFA	5-GGAGACTTCCAAGGTCTTAGCTATCACTTAAGCAC-5'
AFAP	5'-CTGGTTCGGCCCACCTCTGAAGGTTCCAGAATCGATAG-3'
5'RGSP	5'-TAGATATCATCGCAGAAGGC-3'
NCP	5'-AGGAGGAAACA<u>CATATG</u>TTACTTCTCGCCCTTGCA-3'
FCP	5'-GCGCGC<u>CATATG</u>TAAATGTGTCATTCGTTATAA-3'
MNCP	5'-AGGAGGAAAC<u>CATATG</u>ATTGTCGGAGGAATTGAAG-3'

Nine oligonucleotide primers that recognized the antisense and sense strands of the 5'- and 3'-ends of the fibrinolytic F6 protease coding sequence were used in a PCR reaction to generate a clonable fragment encompassing only the coding region.

The primers, and AP and UAP sequences, were determined using a 5'-RGSP kit (Clontech, Palo Alto, USA). The AP primer contained oligo-T sequences, *Mlu* I, *Sal* I and *Spe* I sites. The AFA primer was the AmpliFINDER anchor primer, and the AFAP primer was the counterpart of the AFA primer. The GSP primer, which was used for DNA amplification, was designed from the N-terminal amino acid sequences of F6 protease (Cho *et al.*, 2004). The primer 5'-RGSP, which was used for PCR amplification, was based on the 450 bp upstream sequence of the deduced N-terminal amino acid sequence of the mature region of the F6 protease gene (Fig. 2). The primer (NCP) was prepared using the N-terminal amino acid sequences of the F6 pro-mature sequence (Fig. 2), and contained the *Nde* I cleavage site (underlined). When this cleavage site was digested with the *Nde* I, a starting ATG codon was generated. The primer (FCP), the counterpart primer of NCP, was deduced from an area near the stop codon of the F6 mature protease gene and contained a *Nde* I cleavage site (underlined) (Fig. 2). The primer (MNCP) was deduced from the N-terminal amino acid sequence of the mature region of the F6 protease gene and contained the *Nde* I site (underlined) (Fig. 2).

Molecular cloning of the LKF6 protein gene The full-length fibrinolytic F6 protease gene was constructed by combining two cDNA fragments of 0.9 kb, and 0.4 kb, which were synthesized from mRNA using the procedures for 3'- or 5'-rapid amplification of cDNA ends and PCR (Fig. 1). A recombinant plasmid was constructed using standard protocols (Sambrook *et al.*, 1989; Lee *et al.*, 1998a, 1998b; Uh *et al.*, 2001).

Construction of the p3LK6 clone One gram of *L. rubellus* was crushed and homogenized for 20 sec in 10 ml of guanidium thiocyanate buffer (4 M guanidium thiocyanate, 0.1 M Tris-HCl. 1% -mercaptoethanol, pH 7.5), and then 0.5 ml of 10% sarkosyl (Sigma, St. Louis, USA) was added. This solution was mixed serially with 1.0 ml of 2 M sodium acetate (pH 4.0), 10 ml of phenol and 2 ml of chloroform/isoamyl alcohol (49 : 1) and then centrifuged for 20 min at 4ºC and 10,000 × *g*. The supernatants were collected, 10 ml of isopropanol added, and the solution was stored for 1 h to allow precipitation. The mRNA precipitates were then harvested by centrifugation for 20 min at 4ºC and 10,000 × *g*, and the pellets were washed with 80% ethanol twice and dried;

1.3 mg of total RNA was obtained from the *L. rubellus* lysates. Six micrograms of the mRNA were purified from 300 µg of the total RNA using oligo (dT) 25 Beads (Novagen, Milwaukee, USA). First strand cDNAs were synthesized using mRNA templates and a 3-RACE kit (Gibco, Gaithersburg, USA) (Fig. 1). One µg of the mRNAs and 1.0 pmol of adapter primer (AP) were suspended in 13 µl of H$_2$O and heated for 3 min at 80ºC. Then 2.0 µl of 10 X cDNA synthesis buffer [1.0 µl of 10 mM dNTPs, 2.0 µl of 0.1 M DTT, and 1.0 µl of 10 units/µl superscript RT (reverse transcriptase)] were added and the resultant mixture was heated for 30 min at 42ºC. After heating, 1.0 µl of (1000 units/ml) RNaseH was added to the solution, which was reheated for 10 min to remove the mRNAs. The first cDNA templates were amplified by PCR with primers AP and UAP, and a vent DNA polymerase (NEB, Beverly, USA). PCR was done over 35 cycles of denaturation for 1 min at 94ºC, annealing for 2 min at 60ºC, and extension for 3 min at 72ºC. After PCR, the resulting products were run in 1.0% agarose gels, and the band of interest was eluted, purified, and extracted with a QiaEX kit (Qiagen, Hilden, Germany). The purified DNA fragment was named LK6N cDNAs (0.9 kb). The LK6N DNA fragment (100 ng) was the blunt-ligated into pT7 Blue vector (50 ng) using 2 units of T4 ligase (Novagen), and the resulting recombinant, which was named p3LK6 plasmid, was transformed into *E. coli* DH5α (Sambrook *et al.*, 1989).

Construction of p5LK6 clone First strand cDNA was synthesized using the mRNA templates of *L. rubellus* and a 5'-RACE kit (Clontech, Palo Alto, USA) with the primers AFA and AFAP (Fig. 1). The cDNA fragments were amplified using PCR as in the above-mentioned procedure, and the resultant double stranded cDNA was named LK6C cDNA (0.4 kb). LK6C was cloned into pT7 Blue vector and the resulting clone was named p5LK6, and transformed into *E. coli* DH5α.

Construction of pLK6T clone p3LK6 was digested with *Nco* I and *Sac* I enzymes, and the resultant *Nco* I/*Sac* I fragment was cloned into the *Nco* I and *Sac* I sites of the p5LK6 clone. The resulting plasmid was called pLK6T recombinant (Fig. 1).

Expression construct, pLK-prom clone Using the pLK6T recombinant, the pro-mature gene sequences of the LKF6 gene were
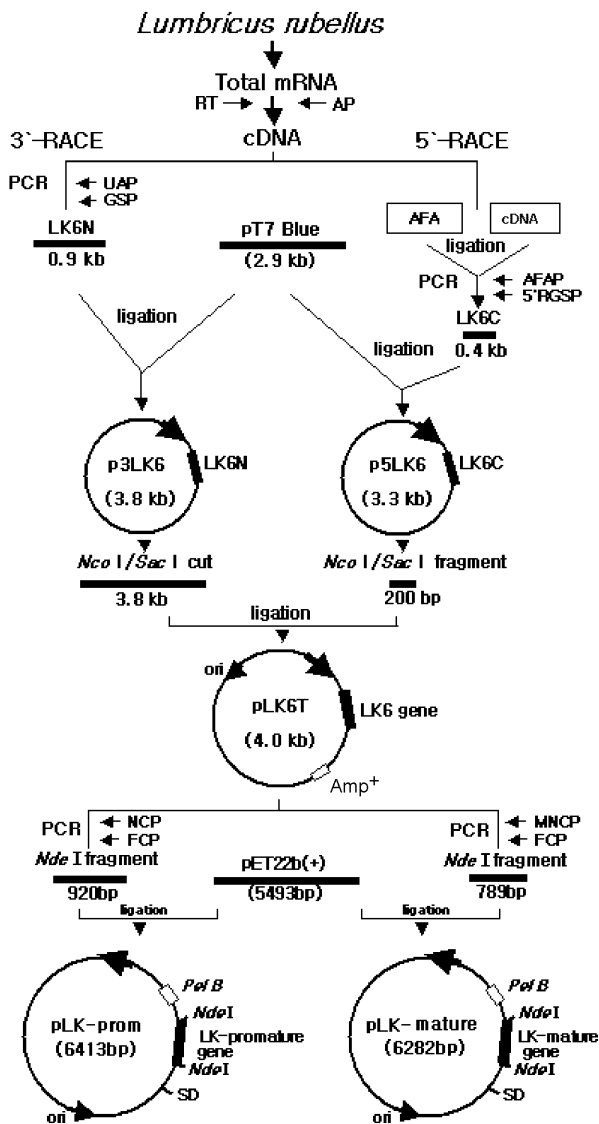
Fig. 1. Construction of pLK6T, pLK-prom and pLK-mature recombinant plasmids containing the F6 gene. Using mRNA templates of the F6 gene in *L. rubellus* lysates, the cDNAs were synthesized with 3- or 5-rapid amplification of cDNA end kits, primers (AP, UAP, GSP, AFA, in Table 1) and a PCR. The cDNAs were cloned into pT7 plasmids, and the full-length cDNAs of the F6 gene clone were generated by combining the *Nco* I and *Sac* I 200 bp DNA fragments in p5LK plasmids into the same sites of the p3LK6. The resulting clone was called the pLK6T recombinant. Two expression clones, containing the sequences of the pro-mature or mature regions, were synthesized and cloned into pET22b(+) vector using primers containing *Nde* I sites. PCR was then performed on the resulting recombinant plasmids, which were named pLK-prom and pLK-mature.

amplified by PCR using the primers NCP and FCP containing *Nde* I sites. Eight base pairs of space were placed between the Shine-Dalgarno sequence (SD) and the ATG codon in the pET22b(+) vector. The PCR product was digested with *Nde* I and cloned into the *Nde* I site of pET22b(+) vector, called pLK-prom recombinant plasmid (6.413 kb) (Fig. 1), and then transformed into the *E. coli*.

Fig. 2. Analysis of the nucleotide sequences of the ORF of the F6 protease gene and its deduced amino acid sequence (GenBank accession No. LK-6 AF304199). The sequences of the proand of the mature region of the F6 protease gene are given. The gene sequences were determined using an automatic DNA sequencer. The 44 pro-region amino acids are shaded. The sequence of the mature protein starts from the 45-isoleucine through to the terminal residue. The N-terminal amino acid sequences of the F6 mature protein are underlined. Italic bases represent conserved residues around a/the serine protease active site. The initiation codon is boldfaced, and ● represents the termination codon. *, sequence of primer GSP. **, sequence of primer 5RGSP.

Expression construct, pLK-mature clone Using the pLK6T recombinant, the mature gene sequence of the LKF6 gene was amplified by PCR with FCP and MNCP primers containing *Nde* I sites. The resulting PCR product was cloned into the *Nde* I site of pET22b(+) vector and called pLK-mature recombinant plasmid (6.282 kb) (Fig. 1), which was transformed into *E. coli*.

Sequencing and computer analysis of LK6 gene cDNAs The LK6N cDNA fragment in the p3LK6 clone, LK6C cDNA fragment in the p5LK6 clone, and LK6 gene DNA in the pLK6T clone were purified through Qiagen tip to determine their DNA sequences.

Initially, an Exo/Mung Bean unidirectional deletion kit (Qiagen) was used to obtain deletion fragments. These DNA fragments were amplified using a PRISM Ready Reaction dye primer cycle sequencing kit (Qiagen). The sequences of the amplified cDNA fragments were determined using an automatic DNA sequencer (Bio-Rad Lab, Richmond, USA) (Crothers and Dreak, 1992). DNA fragments were sequenced on both strands, and the LK6N, LK6C

and LK6 cDNA sequences were fully determined. The whole amino acid sequence of the LKF6 protein was deduced from the complete LK6 cDNA sequence. The DNA base sequences of all fragments were analyzed using the MacDNASIS program (Apple Computer Ltd., Cupertino, USA).

Preparation of fibrinolytic protease from *E. coli*   *E. coli* cells containing the recombinant plasmids, pLK-prom or pLK-mature (Fig. 2), were cultured overnight on LB agar plate containing ampicillin (100 μg/ml), and a single colony was then transferred into 20 ml of LB broth and allowed to grow overnight at 37°C with shaking. The seed cultures obtained were then inoculated into 500 ml of LB broth containing ampicillin (100 μg/ml) and allowed to grow at 37°C until reaching an optical density of 1.2 at 600 nm. At this point, 1.0 mM of IPTG (isopropyl-α-thio-galactopyranoside) (Stratagene, Heidelbergh, Germany) was added, and the cells were re-incubated for 4 hrs at 37°C, and centrifuged at $12,000 \times g$ for 30 min at 4°C. The insoluble protein pellets and supernatants were then separated, and the supernatants were concentrated with 1.0% of trichloroacetic acid and freeze-dried. The insoluble protein pellets were sonicated and centrifuged at $12,000 \times g$ for 30 min at 4°C. The insoluble protein pellets (LK inclusion bodies) were washed twice with 1% Triton X-100, resuspended in 150 μl of 1.0 mM Tris · HCl buffer (pH 7.6), and used for SDS-PAGE (Laemmli, 1970) and to determine fibrinolytic activity (Kumada *et al.,* 1980; Cho *et al.,* 2004).

In vitro* fibrinolytic activity of the protease from *E. coli   To estimate the *in vitro* activity of the inclusion bodies, 1.0 mg of the inclusion bodies of lumbrokinase produced in *E. coli* cells was suspended in 12 ml of 0.05 M HCl-KCl buffer, pH 2.0. The suspension was then ultrasonicated for 10 min at 37°C and 100 μl of pepsin solution (500 units/ml) was added. The mixture was then incubated in a shaking water bath at 37°C for 30 min. To create an intestinal environment, the pH of the mixture was neutralized by adding 1.0 N NaOH. Trypsin solution (500 units/ml) was then added to the neutralized mixture, and incubated in a shaking water bath at 37°C for 60 min. Then 10 μl of the solubilized enzyme solution was plated on the fibrin plate (Cho *et al*., 2004), incubated for 15 h at 37°C, and the hydrolyzed clear zone was measured (Cho *et al*., 2004).

In vivo* fibrinolytic activity of the protease from *E. coli   The fibrinolytic activity of the lumbrokinase (LK) inclusion bodies produced by the *E. coli* system, was evaluated using a rat model of venous thrombosis (Kumada *et al*., 1980) with some modification. Briefly, the inclusion bodies were resuspended in 50 mM phosphate buffer (pH 7.4), washed twice with 1.0% Triton X-100, rewashed with phosphate buffer, and freeze-dried. Eight rats were anaesthetized with ketamine (Yuhan Pharmaceutical Co, Seoul, Korea) and a midline incision was made in the abdomen. The inferior vena cava was exposed and a stainless steel wire coil was inserted 15 mm into the lumen of the inferior vena cava at the left renal vein branching. The inclusion bodies were then suspended for two hours in a 0.5% carboxylmethyl cellulose solution, after which, the inclusion body solution (at a dose of 50 mg/10 ml/kg body wt/ d) was administered orally for 5 d. The control group of eight rats received only 10 ml of 0.5% carboxylmethyl cellulose solution for

5 d. One hour after the final administration, the rats were laparotomized under ketamine anesthesia. Immediately after clamping the vena cava, the wire, together with its thrombus, was removed carefully. The weight of the thrombus on the steel wire was determined as total protein.

Results and Discussion

Cloning and sequence analyses of the F6 protease gene
The full-length of the F6 protease gene was cloned by combining two cDNA fragments, and the resulting clone was named pLK6T recombinant plasmid (Fig. 1). The F6 protease gene in the pLK6T plasmid was sequenced using an automatic DNA sequencer (ABI Prism 310) (Perkin Elmer, Foster, USA), according to the manufactures directions. The open reading frame (ORF) sequence of the F6 protease gene and its deduced amino acid residues (GenBank No. LK-6 AF304199) are illustrated in Fig. 2. The ORF sequence contains 284 codons (852 bp), which encode a polypeptide of 283 amino acid residues. These results were consistent with a report (AF433650) on *L. bimastus* (Sun *et al*., 2002) (Fig. 3). Sugimoto and Nakajima (2001) reported only the mature protein sequence of the ORF of *L. rubellus* strain, and found that the LK gene sequence is highly conserved in the genus *Lumbricus*. The differences in the LK gene and protein sequences in reported *L. bimastus* (Sun *et al*., 2002) and *L. rubellus* strains (Sugimoto and Nakajima, 2001) are shown in Fig. 3 and 4. When our data is compared our data with the previously reported LK gene sequence (AF433650) of *L. bimastus* by Sun *et al*. (2002), eight nucleotide differences were found, which corresponded two codon were changes in the coding region of 852 base pairs (Fig. 3 and 4). However, compared our data with the previously reported mature sequence of the LK fraction of the F-III-1 gene (AB045720) of *L. rubellus* by Sugimoto and Nakajima (2001), 76 nucleotide differences (10.5%) were found, corresponding to which 20 codon changes in the mature coding region of 720 base pairs (Fig. 3 and 4). Sugimoto and Nakajima (2002) did not report the pro-region sequence of the LK gene. In addition, when the mature protein gene sequence of *L. rubellus* reported by Sugimoto and Nakajima (2001) was compared with the previously reported LK gene sequence of *L. bimastus* by Sun *et al*. (2002), 73 nucleotide differences (10.1%) were found, corresponding to 19 codon changes in the mature region of 720 base pairs (Fig. 3 and 4). Also, in the mature protein region (720 base pair) reported by Sugimoto and Nakajima one codon (AGC) between nucleotide 612 and nucleotide 616 was deleted (Fig. 3). In the report by Sun *et al*. (2002) and according to our data one codon (AGC) between nucleotide 612 and 616 was inserted and one amino acid residue (serine) at the amino acid residue-205 added when compared with Sugimoto and Nakajima's report (Fig. 3 and 4). These results indicate that the LK gene sequence of *L. rubellus* strains is less conserved than *L. bimastus* strain. The termination codon was TAG according to our data (Fig. 2), but

```
AF304199    1  ATGTTACTTCTCGCCCTTGCATCGTTGGTAGCGGTGGGTTTTGCGCAACCACCAGTCTGGTACCCCGGTGGTCAATGCAG
AF433650    1  ATGTTACTTCTCGCTCTTGCATCGTTGGTAGCGGTGGGCTTTGCGCAACCACCAGTCTGGTACCCCGGTGGTCAATGCCG
AB045720    1  --------------------------------------------------------------------------------

AF304199   81  TGTCAGCCAGTACTCAGATGCTGGTGACATGGAACTTCCTCCCGGAACAAAAATTGTCGGAGGAATTGAAGCCAGACCAT
AF433650   81  TGTCAGCCAGTACTCAGATGCTGGTGACATGGAACTTCCTCCCGGAACAAAAATTGTCGGAGGAATTGAAGCCAGACCAT
AB045720    1  -------------------------ATGGAACTTCCTCCCGGAACAAAAATTGTCGGAGGAATTGAAGCCAGACCAT

AF304199  161  ACGAGTTCCCATGGCAGGTGTCCGTCCGAAGGAAGTCTTCCGATTCCCATTTCTGCGGAGGTAGCATCATCAACGATCGT
AF433650  161  ACGAGTTCCCATGGCAGGTGTCCGTCCGAAGGAAGTCTTCCGATTCCCATTTCTGCGGAGGTAGCATCATCAACGATCGT
AB045720   53  ACGAGTTCCCATGGCAGGTGTCCGTCCGAAGGAAGTCTTCCGATTCCCATTTCTGCGGAGGTAGCATCATCAACGATCGT

AF304199  241  TGGGTTGTCTGCGCTGCTCACTGCATGCAGGGAGAGAGCCCTGCCCTGGTTTCATTGGTCGTCGGTGAGCACGATAGCAG
AF433650  241  TGGGTTGTCTGCGCTGCTCACTGCATGCAGGGAGAGAGCCCTGCCCTGGTTTCATTGGTCGTCGGTGAGCACGATAGCAG
AB045720  133  TGGGTTGTCTGCGCTGCTCACTGCATGCAGGGAGAGGCCCCGGCTCTGGTTTCATTGGTCGTGGGTGAGCACGACAGGAG

AF304199  321  CGCTGCGAGTACAGTACGTCAGACTCATGACGTTGACAGCATCTTCGTCCACGAGGACTACAACGGAAATACCTTTGAGA
AF433650  321  CGCTGCGAGTACAGTACGTCAGACTCATGACGTTGACAGCATCTTCGTCCACGAGGACTACAACGGAAATACCTTTGAGA
AB045720  213  TGCAGCGAGTACAGTACGTCAGACTCATGACGTTGATAGCATCTTCGTCCACGAGGACTACAACGCAAATACCGTAGAGA

AF304199  401  ACGACGTTTCTGTCATCAAGACAGTTAACGCCATCGCCATCGACATCAACGATGGGCCAATCTGCGCTCCAGATCCAGCC
AF433650  401  ACGACGTTTCTGTCATCAAGACAGTTAACGCCATCGCCATCGACATCAACGTTGGGCCAATCTGCGCTCCAGATCCAGCC
AB045720  293  ACGACGTTTCTGTCATCAAGACATCTGTTGCCATCACTTTCGACATCAACGTTGGTCCAATCTGCGCGCCCAGATCCGGCT

AF304199  481  AACGATTACGTCTACCGTAAGAGCCAGTGCTCCGGATGGGGAACTATCAACTCAGGTGGAGTCTGCTGCCCCAACGTTCT
AF433650  481  AACGATTACGTCTACCGTAAGAGCCAGTGCTCCGGATGGGGAACTATCAACTCAGGTGGAGTCTGCTGCCCCAACGTTCT
AB045720  373  AACGAGTACGTCTACCGTAAGAGCCAGTGTTCCGGATGGGGAACTATCAATTCAGGTGGAATCTGCTGTCCCAACGTTCT

AF304199  561  GCGATATGTGACACTGAACGTCACAACCAACGCCTTCTGCGATGATATCTACAGCCCATTATATACAATTACCAGCGACA
AF433650  561  GCGATATGTGACACTGAACGTCACAACCAACGCCTTCTGCGATGATATCTACAGCCCATTATATACAATTACCAGCGACA
AB045720  453  GCGATACGTGACGCTGAATGTCACAACCAACCAATTCTGCGAAGATGTATAC---CCACTAAATTCAATCTACGACGATA

AF304199  641  TGATCTGCGCCACGGACAACACCGGACAGAACGAGAGAGACTCTTGCCAGGGTGACTCTGGCGGCCCTCTGAGCGTCAAG
AF433650  641  TGATCTGCGCCACGGACAACACCGGACAGAACGAGAGAGACTCTTGCCAGGGTGACTCTGGCGGCCCTCTGAGCGTCAAG
AB045720  530  TGATTTGCGCGTCGGACAACACTGGGGGTAACGACAGAGACTCCTGCCAGGGTGACTCGGCGGCCCTCTGAGCGTCAAG

AF304199  721  GATGGCAGCGGAATCTTCAGCCTCATTGGTATTGTGTCTTGGGGAATCGGTTGCGCATCTGGCTATCCAGGAGTCTACGC
AF433650  721  GATGGCAGCGGAATCTTCAGCCTCATTGGTATTGTGTCTTGGGGAATCGGTTGCGCATCTGGATATCCAGGAGTCTACGC
AB045720  610  GATGGCAGTGGAATCTTCAGCCTGATTGGTATTGTGTCTTGGGGAATTGGTTGCGCTTCTGGCTATCCAGGAGTCTACTC

AF304199  801  CCGCGTCGGATCCCAAACTGGATGGATCACAGACATTATTACCAACAACTAG
AF433650  801  CCGCGTCGGGTCCCAAACTGGATGGATCACAGACATCATCACCAACAACTAA
AB045720  690  CCGCGTCGGATTCCATGCTGCATGGATCACGACATCATCACCAACAACTAA
```

Fig. 3. Comparison of the LK gene sequences of *L. rubellus* and *L. bimastus* strains. The base sequence with GenBank accession number AF304199 is the LK gene of *L. rubellus* isolated in Korea, the sequence AF433650 (Sun *et al.*, 2002) is the LK gene of *L. bimastus* strain, and the sequence AB045720 (Sugimoto and Nakajima, 2001) is LK gene of a *L. rubellus* strain isolated in Japan. Black bars with white text indicate different bases. The numbers on the left indicate the nucleotide numbers. (codon: ---) between nucleotides 612 and 616 of AB045720 means a deleted codon versus the other two sequences.

TAA according to Sugimoto and Nakajima (2001) and Sun *et al.* (2002).

The F6 protein gene might consist of duplicate regions of the full-length amino acid sequence of the F6 protein, based on the N-terminal amino acid sequence (Cho *et al.*, 2004). The active form (the mature protein) of F6 polypeptides is initiated from isoleucine-45 and not methionine. This means that the polypeptides produced may be modified after translation. The F6 gene has a pro-region sequence upstream from its mature sequence. The amino acid sequence of the pro-region (44 amino acid residues) in the F6 polypeptide starts at methionine-1 and continues to lysine-44 (shaded box). The active mature polypeptide sequence (239 amino acid residues and one stop codon; 720 bp) starts from isoleucine-45 to the terminal residue. Thus the pro-region consists of 132 bp (44 amino acid residues), and the mature

protein sequence of the F6 protein gene is of 720 bp, encoding a polypeptide of 239 amino acid residues. The F6 polypeptide enzymes have a potential *N*-glycosylation site at amino acid residues 149-152 and a serine protease active site at 231-235-amino acid residues. The deduced sequence of the amino acid residues in the mature polypeptide contains a high percentage (11-14%) of asparagine and aspartic acid, and the percentage of lysine was lower than former amino acid.

Fibrinolytic activity of F6 proteases produced by *E. coli* clones The pLK-prom and the pLK-mature clones were constructed according to the scheme shown in Fig. 1. The two clones were transformed in *E. coli* system, and produced inclusion body types of F6 proteases, were identified by SDS-PAGE (Fig. 5 lanes 2 and 3). Lane 2 is the protein produced by the pLK-mature clone, which has a molecular weight of

```
AF304199    1 MLLLALASLVAVGFAQPPVWYPGGQCSVSQYSDAGDMELPPGTKIVGGIEARPYEFPWQVSVRRKSSDSHFCGGSIINDR
AF433650    1 MLLLALASLVAVGFAQPPVWYPGGQCGVSQYSDAGDMELPPGTKIVGGIEARPYEFPWQVSVRRKSSDSHFCGGSIINDR
AB045720    1 ------------------------------------MELPPGTKIVGGIEARPYEFPWQVSVRRKSSDSHFCGGSIINDR


AF304199   81 WVVCAAHCMQGESPALVSLVVGEHDSSAASTVRQTHDVDSIFVHEDYNGNTFENDVSVIKTVNAIAIDINDGPICAPDPA
AF433650   81 WVVCAAHCMQGESPALVSLVVGEHDSSAASTVRQTHDVDSIFVHEDYNGNTFENDVSVIKTVNAIAIDINVGPICAPDPA
AB045720   45 WVVCAAHCMQGEAPALVSLVVGEHDESAASTVRQTHDVDSIFVHEDYNANTLENDVSVIKTSVAITFDINVGPICAPDPA


AF304199  161 NDYVYRKSQCSGWGTINSGGVCCPNVLRYVTLNVTTNAFCDDIYSPLYTITSDMICATDNTGQNERDSCQGDSGGPLSVK
AF433650  161 NDYVYRKSQCSGWGTINSGGVCCPNVLRYVTLNVTTNAFCDDIYSPLYTITSDMICATDNTGQNERDSCQGDSGGPLSVK
AB045720  125 NDYVYRKSQCSGWGTINSGGICCPNVLRYVTLNVTTNQFCEDVYEPLNSIYDDMICASDNTGGNDRDSCQGDSGGPLSVK


AF304199  241 DGSGIFSLIGIVSWGIGCASGYPGVYARVGSQTGWITDIITNN
AF433650  241 DGSGIFSLIGIVSWGIGCASGYPGVYARVGSQTGWITDIITNN
AB045720  204 DGSGIFSLIGIVSWGIGCASGYPGVYSRVGFHAAWITDIITNN
```

Fig. 4. Comparison of the LK protein sequences of *L. rubellus* and *L. bimastus* strains. The amino acid sequence with GenBank accession number AF304199 is the LK protein of *L. rubellus* isolated in Korea, the sequence AF433650 is the LK protein of *L. bimastus* strain, and the sequence AB045720 is LK protein of *L. rubellus* strain isolated in Japan. Black bars with white text indicate different amino acids. The numbers on the left indicate amino acid residues. (-) at amino acid residues-245 on AB045720 indicates a deleted amino acid versus the other two sequences.
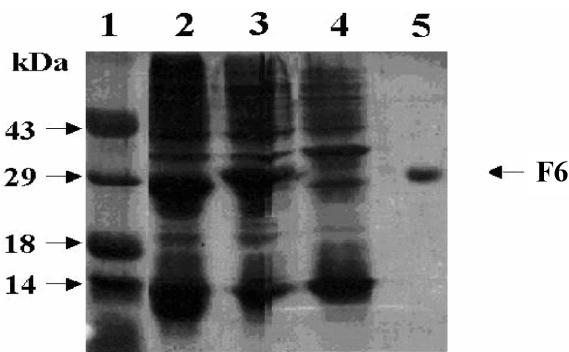

Fig. 5. SDS-PAGE patterns of F6 inclusion bodies produced by clones. *E. coli* containing the clones (pLK-prom and pLK-mature) produced inclusion bodies of F6 protein. The F6 inclusion bodies were dissolved in the 50 mM phosphate buffer and run on 12% SDS-PAGE. The lanes shown are: 1, molecular weight maker; 2, proteins produced by pLK-mature clone; 3, proteins produced by pLK-prom clone; 4, proteins produced by pET22b vector (control), and 5, the F6 fraction protein. Numbers on the left correspond to the positions of the molecular weight markers (in kDa) (ovalbumin 43 kDa, carbonic anhydrase 29 kDa, b-lactogloblin 18 kDa, and lysozyme 14 kDa). F6 indicates F6 protein.

Fig. 6. *In vitro* fibrinolytic activity on fibrin plates of the LK inclusion bodies from *E. coli*. A 5% fibrinogen solution was prepared using a fibrin plate buffer. 10 ml of the mixture was distributed into sterile Petri-dishes, and then 50 µl of 20 NIH units/ml of thrombin solution were slowly mixed into each dish and allowed to solidify for 1 h at room temperature. 10 µl of each LK inclusion body (IB) solution treated with trypsin and/or pepsin was dropped on the fibrin plates and incubated for 15 h at 37°C. After incubation, the mean diameters of the hydrolyzed clear zones were measured. Trypsin, IB treated with trypsin; trypsin + pepsin + IB spot, IB treated with trypsin and pepsin.

approximately 28 kDa, and lane 3 is the pro-mature protein produced by the pLK-prom clone, which has a molecular weight of approximately 31 kDa.

Lane 4 shows the proteins produced by the *E. coli* control, and lane 5 the protein from fraction 6, which has a molecular weight of approximately 33 kDa. The mature and pro-mature proteins showed lower molecular weights than the proteases from fraction 6. This means that the mature and promature proteins were not modified in *E. coli* cells. These results indicate that the two clones produced different molecular weights gene products due to the different lengths of their genes.
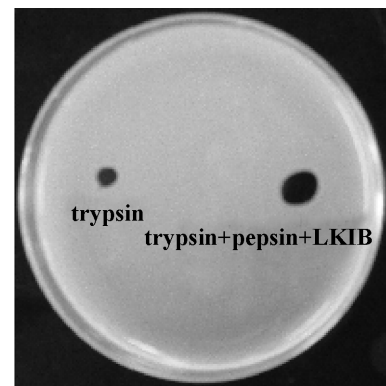
The inclusion bodies, which were solubilized in pepsin and/or trypsin solutions (described in "Materials and Methods") showed high hemolytic activity *in vitro* (Table 2 and Fig. 6). The IB solutions treated with pepsin and trypsin together showed much stronger hemolytic activity than those treated with only trypsin. The IB solution alone was inactive. The inclusion bodies suspended in a 0.5% carboxylmethyl cellulose solution were administered orally for 5 days. After last administration, the steel wires and attached thrombus were removed, and the thrombus weights determined (described in the Method). The F6 proteases hydrolyzed 62-

Table 2. *In vitro* fibrinolytic activity of inclusion bodies

Biochemical reactions	Diameter of clear zone (mm)
LK-IB	ND
pepsin	ND
LK-IB + trypsin	5.7 ± 0.4
LK-IB + pepsin	ND
LK-IB + pepsin + trypsin	11.2 ± 1.6*

LK-IB, the inclusion body of lumbrokinase F6 produced in *E. coli* cells; ND, not detected; *, $p < 0.05$ vs trypsin. Data were represented mean ± SD (n = 5).

Table 3. Thrombus weight change after oral administration of the LK inclusion bodies produced in *E. coli*

LK IB sources	Treated dose	Thrombus weight (μg) ± S.D
Control	CMC 10 ml/kg/d	1726 ± 852
pLK-prom	50 mg/kg/d	659 ± 528
pLK-mature	50 mg/kg/d	610 ± 429

Fibrinolytic activity of the LK inclusion bodies (IB) was evaluated using a rat model of venous thrombosis (Kumada *et al.*, 1980). The inclusion bodies suspended in a 0.5% carboxylmethyl cellulose (CMC) solution was administered orally for 5 days. One hour after the last administration, the rats were laparotomized. Immediately after clamping the vena cava, the wire with its thrombus was removed carefully. The weight of the thrombus on the steel wire was measured as total protein content.

65% of the blood clots on the stainless wire coils in rats when fed orally (Table 3). The thrombus weights were reduced to 659 (62%) by the pro-mature enzymes produced by the pLK-prom clone and to 610 (65%) by the mature form produced by the pLK-mature clone versus the control group (Table 3). The pro-mature and mature proteases produced in the *E. coli* system had similar fibrinolytic activities *in vivo* and *in vitro,* which means that the pro-mature and the mature regions of the F6 proteases may be involved in proteolytic activity. We assumed that when the inclusion body solution was fed to rats that were converted in the intestine to active forms. Further studies are necessary.

We conclude that the F6 proteases consist of pro- and mature regions. The ORF sequence of the pro-mature protease was composed of 852 bp and the mature region was of 720 bp. The mature and pro-mature enzymes produced by *E. coli* were produced as inclusion bodies and showed high fibrinolytic activity *in vivo* and *in vitro*.

References

Cho, I. H., Choi, E. S., Lim, H. G. and Lee, H. H. (2004) Purification and characterization of six fibrinolytic serine-proteases from earthworm *Lumbricus rubellus. J. Biochem. Mol. Biol.* **37**, 199-205.

Crothers, D. M. and Drak, J. (1992) Global features of DNA structure by comparative gel electrophoresis. *Methods in Enzymology* **212**, 46-71.

Jeon, O. H., Moon, W. J. and Kim, D. S. (1995) An anticoagulant/fibrinolytic protease from *Lumbricus rubellus. J. Biochem. Mol. Biol.* **28**, 138-1452.

Kumada, T., Ishihara, M., Ogawa, H. and Abiko, Y. (1980) Experimental model of venous thrombosis in rats and effect of some agents. *Thrombosis Res.* **18**, 189-203.

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

Lee, H. H., Kang, B. J. and Park, K. J. (1998a) Construction of a Baculovirus expression system using *H. cunea* nuclear polyhedrosis virus for eukaryotic cell. *J. Microbiol. Biotechnol.* **8**, 676-684.

Lee, H. H., Moon, E. S., Lee, S. T., Hwang, S. H., Cha, S. C. and Yoo, K. H. (1998b) Construction of a Baculovirus *Hyphantria cunea* NPV insecticide containing the insecticidal protein gene of *Bacillus thuringiensis* subsp. *kurstaki* HD1. *J. Microbiol. Biotechnol.* **8**, 685-691.

Mihara, H., Sumi, H., Yoneta, T., Mizumoto, H., Ikeda, R., Seikl, M. and Maruyama, M. (1991) A novel fibrinolytic enzyme extracted from the earthworm, *Lumbricus rubellus. Jpn. J. Phgysiol.* **41**, 461-472.

Mihara, H., Yineta, T., Sumi, H., Soeda, M. and Maruyama, M. (1989) A possibility of earth worm powder as therapeutic agent for thrombosis. *Thromb. Haemosta.* **62**, 545-549.

Nakajima, N., Mihara, H. and Sumi, H. (1993) Characterization of potent fibrinolytic enzymes in Earthworm, *Lumbricus rubellus. Biosci. Biotech. Biochem.* **57**, 1726-1730.

Park, S. Y., Kye, K. C., Lee, M. H., Sumi, H. and Mihara, H. (1989) Fibrinolytic activity of the earth worm extract. *Thromb. Haemosta.* **62**, 545-550.

Park, Y., Ryu, E., Kim, H., Jeong, J., Kim, J., Shim, J., Jeon, S., Jo, Y., Kim, W. and Min, B. (1999) Characterization of antithrombotic activity of lumbrokinase-immobilized Polyurethane valves in the total artificial heart. *Artif. Organs.* **23**, 210-214.

Ryu, G. H., Han, D. K., Park, S. Y., Kim, M., Kim, Y. H. and Min, B. G. (1995) Surface characteristics and properties of lumbrokinase-immobilized polyurethane. *J. Biomed. Mater. Res.* **29**, 403-409.

Ryu, G. H., Park, S., Kim, M., Han, D. K., Kim, Y. H. and Min, B. G. (1994) Antithrombogenicity of lumbrokinase immobilized polyurethane. *J. Biomed. Mater. Res.* **28**, 1069-1077.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, USA.

Sugimoto, M. and Nakajima, N. (2001) Molecular cloning, sequencing, and expression of cDNA encoding serine protease with fibrinolytic activity from earthworm. *Biosci. Biotechnol. Biochem.* **65**, 1575-1580.

Sun, Z-J., Liang, G-D., Chen, F., Fu, S-H., Shen, Y., Chai, Y-B., Li, X-Y., Xu, Y-H. and Hou, Y-D. (2002) Lumbrukinase from Earthworm-Cloning and its effecxt to BHK cells. *Chinese J.*

Biochem. Mol. Biol. **18**, 776-779.

Uh, H. S., Choi, J. H., Byun, S. M., Kim, S. Y. and Lee, H. H. (2001) Cloning, sequencing and Baculovirus-based expression of fusion-glycoprotein D gene of *Herpes simplex* virus type 1(F). *J. Biochem. Mol. Biol.* **34**, 371-378.

Xu, Y.-H., Liang G-D., Sun Z-J., Chem, F., Fu, S-H., Chai, Y. B. and Hou, Y-D. (2002) Cloning and expression of the novel gene-pV242 of earthworm fibrinolytic enzyme. *Prog. Biochem. Biophys.* **29**, 610-614.