A THESIS FOR THE DEGREE OF MASTER OF SCIENCE

Molecular cloning and inhibitory characterization of papain-like cysteine protease propeptide from *Calotropis procera* R. Br.

Calotropis procera R. Br. 유래 파파인 유사 cysteine protease 프로펩타이드의 유전자 클로닝 및 저해 특성 규명

August, 2018

Department of Agricultural Biotechnology

Seoul National University

Chung, Bokyong

석사학위논문

Molecular cloning and inhibitory characterization of papain-like cysteine protease propeptide from *Calotropis procera* R. Br.

지도교수 장 판 식

이 논문을 석사학위 논문으로 제출함

2018년 8월

서울대학교 대학원

농 생 명 공 학 부

정 보 경

정보경의 석사 학위논문을 인준함

2018년 8월

위원장 최 영 진 (인)

부위원장 장 판 식 (인)

위원 최 상 호 (인)

Abstract

Cysteine proteases are present in all living organisms and cause a wide range of diseases as well as bacterial and parasitic infections. Therefore, much effort has been made to search for inhibitors to control cysteine protease activity. Recently, twenty kinds of cysteine proteases from *Calotropis procera* R. Br., which were identified by transcriptome analysis, have been reported. Based on the inhibitory activity of propeptide to their cognate enzymes, eight propeptides (SnuCalCpIs) were selected and SnuCalCpI03 and SnuCalCpI15 could be used as cysteine protease inhibitors. In this study, a novel inhibitor among SnuCalCpI08, SnuCalCpI14, and SnuCalCpI17, which were not expressed in previous studies, has been secured and the analysis on inhibitory activity and kinetics of SnuCalCpIs for papain was performed to develop inhibitor material.

cDNA was synthesized using total RNA extract from *Calotropis procera* R. Br., and genes of SnuCalCpI08, SnuCalCpI14 and SnuCalCpI17 were amplified by PCR. Analysis of codon usage in SnuCalCpI genes revealed that SnuCalCpI08 and SnuCalCpI17 among SnuCalCpI08, SnuCalCpI14, and SnuCalCpI17, which were not expressed in the previous studies, were found

to hold 10% or more of rare codons, which were rarely used by *E. coli*, unlike the other inhibitor genes. Each gene was inserted into pET29b(+) vector and expressed in *E. coli* Rosetta(DE3) providing tRNAs for six rare codons. SnuCalCpI08 and SnuCalCpI17 were overexpressed, but only SnuCalCpI17 was expressed in soluble form.

Characterization of SnuCalCpI17 including SnuCalCpI03 and SnuCalCpI15, which have been previously obtained, was performed. According to results of papain inhibition assay, the half-maximal inhibitory concentrations (IC_{50}) of SnuCalCpI03, SnuCalCpI15, and SnuCalCpI17 were 103.702±2.060 μg/mL, 23.900±0.654 μg/mL, and 105.671±9.857 μg/mL, respectively. SnuCalCpIs exhibited the kinetic characteristics of a slow-binding inhibition for papain, and K_i^{app} values were 28.70 μg/mL, 13.63 μg/mL, and 75.80 μg/mL, respectively.

Slow-binding inhibitor has the slow dissociation of inhibitor from the inhibitor-target complex, resulting in long residence time *in vivo* and effective inhibition of the target at doses far below the IC_{50} of the inhibitor. Also, high-molecular weight protein inhibitors show higher enzyme selectivity than chemical inhibitors which bind to cysteine residues at the active site due to the stereospecificity of the protein-protein interaction. In addition, peptides from

natural sources do not accumulate in body tissues, and there are few reports about negative side effects. Therefore, SnuCalCpI, which was derived from natural resource and has characteristic of slow-binding inhibition, can be used as an effective inhibitor for papain-like cysteine protease.

Keywords: *Calotropis procera* R. Br., cysteine protease inhibitor, propeptide, rare codon, slow-binding inhibition

Student number: 2016-26913

Contents

List of tables

List of figures

1. Introduction

Cysteine proteases have catalytic cysteine residues and hydrolyze proteins by initiating a nucleophilic attack on the carbonyl carbon of peptide bonds (Ovat, Muindi, Fagan, Brouner, Hansell, Dvořák, et al., 2009). Most cysteine proteases belong to the papain-like cysteine protease classified as clan CA, C1 family (McGrath, 1999). Cysteine protease exists in eukaryotes such as plant and animal and prokaryotes such as bacteria and viruses (TAKAYUKI & L., 2008). Cysteine proteases affect a wide range of diseases. For example, it is known that overexpression of cathepsin K is associated with bone atrophy (osteoporosis), and overexpression of cathepsin B, L, and K is associated with cartilage destruction in rheumatoid arthritis (Hummel, Petrow, Franz, Müller-Ladner, Aicher, Gay, et al., 1998). Moreover, cathepsin L, B or S is overexpressed in cancer cells and is associated with carcinoma progression, metastasis, and angiogenesis (Radim, Matthias, Knut, & Tanja, 2006). Cysteine protease also functions as the cause of bacterial and parasitic infections as well as diseases (Lukomski, Montgomery, Rurangirwa, Geske, Barrish, Adams, et al., 1999; Kailash C. Pandey, 2013; Rosenthal, 2004). Thus, much effort has been made to search for inhibitors to control cysteine protease

activity.

There are many ways to develop inhibitors. However, the possibility could be found in the structural and functional properties of cysteine protease itself. All papain-like cysteine proteases are commonly expressed in the form of precursors called zymogens, which consist of a signal peptide, a propeptide, and a mature catalytic domain. Propeptide controls the activity of protease as an inhibitor as well as the role of intramolecular chaperone which catalyzes the folding of the catalytic domain of cognate protease (Khan & James, 1998; Turk, Stoka, Vasiljeva, Renko, Sun, Turk, et al., 2012). This shows that the propeptide of cysteine protease itself can be used as an inhibitor, and research on this issue has been actively carried out to develop inhibitors (F., W., & C., 2005; Kwon, Park, Kang, Kweon, Kim, Shin, et al., 2015; K. C. Pandey, Barkan, Sali, & Rosenthal, 2009; Scott & Taggart, 2010).

Although the peptide inhibitor is high-molecular weight protein, it can specifically bind well to cysteine residues at the active site because it stereospecifically engages in protein-protein interactions compared with a chemical inhibitor (Kwon, Yang, Yeo, Park, Jeong, Lee, et al., 2018). In addition, bioactive peptides obtained from natural sources are not accumulated in bodies (Uhlig, Kyprianou, Martinelli, Oppici, Heiligers, Hills, et al., 2014),

and there are few examples of side effects when it is used for preventative healthcare (Hayes, 2014; Hayes & Tiwari, 2015). Despite the advantages mentioned above, these natural peptides are difficult to extract, isolate, and purify, making them less economical. Therefore, it is necessary to conduct genetic engineering to produce the target protein in a large amount.

Previous studies found *Calotropis procera* R. Br. as a source of cysteine protease inhibitor. It is known as a traditional medicinal plant in the tropics and the latex of these plants contains various therapeutic functions such as hepatoprotective, anti-arthritic, anti-inflammatory, antipyretic, and anticancer treatment (Choedon, Mathan, Arya, Kumar, & Kumar, 2006; Dewan, Sangraula, & Kumar, 2000; V. L. Kumar & Basu, 1994; Vijay L. Kumar & Roy, 2007). From transcriptome analysis of *Calotropis procera* R. Br., twenty kinds of cysteine protease were identified (Kwon, et al., 2015). Eight propeptide domains (SnuCalCpIs) out of twenty cysteine proteases were selected, and genetic engineering was conducted to express them in *Escherichia coli* system. Five propeptide domains were expressed, but only SnuCalCpI03 and SnuCalCpI15 could be used as cysteine protease inhibitors. This result shows that the propeptide of cysteine protease obtained from *Calotropis procera* R. Br. can be used as an actual inhibitor agent. Therefore,

propeptide domains that were not expressed are also likely to act as similar cysteine protease inhibitors, and it is necessary to secure a novel inhibitor from natural material and broaden the spectrum of the inhibitor agent.

In this study, SnuCalCpI08, SnuCalCpI14, and SnuCalCpI17, which were not expressed in *E. coli* system, have been secured as a novel cysteine protease inhibitor in *E. coli* Rosetta(DE3). In addition, the analysis of inhibitory activity and kinetics of SnuCalCpIs was performed to develop novel inhibitors against cysteine protease (papain).

2. Materials and Methods

2.1. Genetic engineering of SnuCalCpI production

2.1.1. Strains, plasmids, and media

Escherichia coli (*E. coli*) DH5α, used as host strain for cloning, was obtained from Invitrogen Co. (Carlsbad, CA, USA). pET29b(+) was used as a plasmid for recombinant protein expression in *E. coli* Rosetta(DE3) (Novagen Co., CA, USA). Luria-Bertani (LB) medium was used for growing *E. coli* strains at 37°C. LB agar medium supplemented with 50 μg/mL kanamycin and 34 μg/mL chloramphenicol was used to screen the recombinant clones.

2.1.2. Cloning of SnuCalCpI genes

2.1.2.1. Homology modeling

Deduction of the amino acid sequence of cysteine protease inhibitors (SnuCalCpIs) from *Calotropis procera* R. Br. was performed by Translate tool (https://web.expasy.org/translate/). The theoretical molecular weights of SnuCalCpIs were determined using the online ProtParam tool at ExPasy (https://web.expasy.org/protparam/). Homology modeling of SnuCalCpIs was implemented using the BLASTp program on the servers of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov), and

sequences and structures of target proteins were compared to papain propeptide from *Carcia papaya*.

2.1.2.2. Codon usage analysis

To assess preferences for particular synonymous codons, the number and frequency of each codon in SnuCalCpI genes were calculated using Sequence Manipulation Suit (http://www.bioinformatics.org/sms2/reference.html).

2.1.2.3. Total RNA extraction and cDNA synthesis

Total RNA was isolated after grinding *Calotropis procera* R. Br. leaf in liquid nitrogen with RNeasy® plant mini kit (Qiagen Co., CA, USA). The tissues were lysed in lysis buffer containing guanidine hydrochloride, and total RNA in cell lysate was bound to the column supplied with the kit. The extracted RNA was eluted with 30 μL of RNase-free water from the column and purified according to the manufacturer's instructions. The purity and amount of RNA were calculated from the absorbance of RNA measured at 260 nm by NanoVue Plus NanoDrop spectrophotometer (GE Healthcare Co., NJ, USA).

The synthesis of single-stranded cDNA was carried out using PrimeScript™ 1st strand cDNA synthesis kit (Takara Korea Biomedical Co., Seoul, Korea).

Total RNA was mixed with oligo dT primer, dNTP mixture, and RNase free water and incubated for 50 min at 65°C, then cooled immediately on ice. To synthesize cDNA, 5X PrimeScript buffer and RTase with the kit were added to the mixture and incubated for 30 min at 42°C.

2.1.2.4. Polymerase chain reaction (PCR)

To amplify the SnuCalCpI genes, specific primers were designed including the in-fusion cloning sites (Table 1). An aliquot of single-stranded cDNA was used as the template for the synthesis of the second-stranded cDNA in a PCR reaction with 2X TOPsimple^{TM} PreMIX-Forte (Enzynomics Co., Daejeon, Korea). T100^{TM} Thermal Cycler (Bio-Rad Co., CA, USA) was used for the PCR. PCR conditions were 1 cycle initial denaturing at 95°C for 2 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 53°C for 30 s, and elongation at 72°C for 1 min, with a final elongation at 72°C for 5 min. The amplified PCR products were purified using MEGAquick-spin^{TM} Total Fragment DNA Purification Kit (iNtRON Biotechnology Co., Seongnam, Korea) according to the instructions of the manufacturer, dyed with a 6X Loading STAR solution (DyneBio Co., Seongnam, Korea), and identified by electrophoresis on a 1.0% agarose gel.

Table 1. Primers used for cloning of recombinant SnuCalCpIs

Primer name	**Sequence (5’ → 3’)**
SnuCalCpI08_F	5’-AAG GAG ATA TAC ATA TGG TTG ACG ACG GAT CAT CAG-3’
SnuCalCpI08_R	5’-GGT GGT GGT GCT CGA GGA CAA CGT TGG TTA GCT TG-3’
SnuCalCpI14_F	5’-AAG GAG ATA TAC ATA TGT CAT TTT CAT CTT CTT CTT CTT-3’
SnuCalCpI14_R	5’-GGT GGT GGT GCT CGA GAT CAA AAT CAT CAA AGA CAT CT-3’
SnuCalCpI17_F	5’-AAG GAG ATA TAC ATA TGT CTG AGA TCA CGT CGG TTA-3’
SnuCalCpI17_R	5’-GGT GGT GGT GCT CGA GAT TAT CTT CGG CTT TAG GAA G-3’

2.1.2.5. Construction of recombinant plasmids

To construct the expression plasmids pET29b(+)-SnuCalCpI-C6His (vector-gene-tag), genes of SnuCalCpI08, SnuCalCpI14, and SnuCalCpI17 were amplified by PCR and pET29b(+) expression vector was linearized by double digestion using *Nde*I and *Xho*I (Takara Bio Co.). All SnuCalCpI DNA fragments were inserted into the *Nde*I and *Xho*I site of pET29b(+) expression vector by In-Fusion® HD Cloning Kit (Takara Bio Co.). Recombinant plasmids were transformed into *E. coli* DH5α. To confirm the presence of the recombinant strains, cells were screened by kanamycin LB agar medium and colony PCR. The recombinant plasmids were precipitated with DNA-spin™ Plasmid DNA Purification Kit (iNtRON Biotechnology Co.) and confirmed by DNA sequencing with ABI 3730xl DNA analyzer (Macrogen Co., Seoul, Korea). Verified recombinant plasmids were transformed into *E. coli* Rosetta(DE3) for SnuCalCpI expression.

.

2.1.3. Expression and purification of SnuCalCpIs

2.1.3.1. Expression of SnuCalCpIs

Each recombinant cell was cultured in LB medium containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37°C with shaking overnight. The overnight culture was subcultured using 1% inoculum to fresh LB medium containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37°C with shaking until the optical density at 600 nm (OD_{600}) reached approximately 0.4-0.8. To induce protein expression, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to final concentration of 0.1 mM and growth was continued for 16 h. Effects of a temperature gradient (18°C, 30°C, and 37°C) and time variation (4, 6, and 16 h) to protein expression were considered.

2.1.3.2. Purification of SnuCalCpIs

Cells were harvested and centrifugated at 4,000xg for 20 min at 4°C. The pellet was resuspended in 50 mM sodium phosphate buffer (pH 7.0) with 300 mM NaCl and 10 mM imidazole and then sonicated on ice (40 cycles of 10 s pulses with 10 s pause) to disrupt the cell wall. The cell lysates were separated into supernatant and cell pellet by centrifugation at 15,000xg for 20 min, and the supernatant was analyzed by sodium dodecyl sulfate polyacrylamide gel

electrophoresis (SDS-PAGE).

The recombinant SnuCalCpI proteins that contain a 6xHis tag at the C-terminal were purified from the supernatant using an affinity Ni-NTA column (Qiagen Co.). Ni-NTA slurry was packed in econo-columns, and the resin was washed with the lysis buffer (50 mM sodium phosphate (pH 7.0) with 300 mM NaCl, and 10 mM imidazole). The supernatant fractions of expressed proteins in the lysis buffer were loaded on the resin after shaking slowly for 1 h. Unbound proteins were washed with the washing buffer (50 mM sodium phosphate (pH 7.0) with 300 mM NaCl, and 20 mM imidazole). Target proteins were eluted with the elution buffer (50 mM sodium phosphate (pH 7.0) with 300 mM NaCl, and 250 mM imidazole), and the eluents were collected.

The purified protein was analyzed by SDS-PAGE. The fractions containing the purified protein were dialyzed and concentrated with 10 kDa Amicon filter (Millipore Co., MA, USA). Protein concentration was determined according to Bradford method (Kruger, 1994) using bovine serum albumin (BSA) as a standard.

2.2. Inhibitory characterization of SnuCalCpIs

2.2.1. Papain inhibition assay

The papain inhibition assay was performed as described by the method of Abe (Abe, Abe, Kuroda, & Arai, 1992). Papain (Sigma Co., Missouri, USA) activity was measured using 1 mM N-α-Benzoyl-DL-arginine-β-naphthylamide (BANA) as a substrate. Various concentrations of recombinant SnuCalCpI samples in 0.2 mL were mixed with 0.2 mL 0.5 M sodium phosphate (pH 7.0) with 10 mM EDTA , 0.2 mL of 50 mM 2-mercaptoethanol, and 0.2 mL of 25 μg/mL papain solution. The mixture was incubated at 37°C for 10 min. After pre-incubation, the reaction was started by addition of 0.4 mL of 1 mM BANA. After incubation at 37°C for 20 min, 150 μL of the mixture was added to 400 μL of 2% (v/v) HCl/ethanol to stop the enzyme reaction and then, 400 μL of 0.06% (w/v) *p*-dimethylaminocinnamaldehyde/ethanol was added to the mixture. The mixture was put at room temperature for 30 min for color development. After that, the absorbance of products was measured at 540 nm with a spectrophotometer. The relative residual activity of papain was calculated as follows,

$$\text{Relative residual activity } (\%) = \frac{T^*}{T} \times 100$$

where T is the absorbance at 540 nm in the absence of SnuCalCpI, and T* is the one in the presence of SnuCalCpI. The half-maximal inhibitory concentrations (IC_{50}) were determined from the fit curve equation using SigmaPlot 12.5 (Systat Software Co., CA, USA).

2.2.2. Kinetic analysis of SnuCalCpIs

Slow-binding inhibition kinetics were analyzed by papain inhibition assay. All kinetic measurements for papain were performed at 37°C for 18 min at 1.5 min time intervals in 0.5 M sodium phosphate buffer (pH 7.0) containing 10 mM EDTA. Various concentrations of SnuCalCpI and a substrate (BANA) concentration of 1.5 mM were used for slow-binding kinetic analysis ([S]≪Km). The reaction was started by adding substrate after 10 min of pre-incubation. Under these experimental conditions, progress curves for the inhibition of papain by SnuCalCpIs at pH 7.0 followed typical slow-binding kinetics as defined by the following equation,

$$[\mathrm{P}] = v_i t + \frac{(v_i - v_s)[1 - \exp(-k_{obs} t)]}{k_{obs}}$$

where [P] is the concentration of product formed, v_i and v_s are the initial and steady-state velocities of the reaction in the presence of inhibitor, t is the reaction time, and k_{obs} is the apparent first-order rate constant for the

interconversion between v_i and v_s. Nonlinear regression using the program SigmaPlot 12.5 (Systat Software Co.) provide the individual parameters (v_i, v_s, and, k_{obs}) for each progress curve.

2.2.3. Inhibition of bacterial growth

2.2.3.1. Micro-dilution test

The three Gram-positive strains (*Staphylococcus aureus* ATCC 49444, *Listeria welshimeri* ATCC 35897, and *Pseudomonas aeruginosa* ATCC 15692) and the three Gram-negative strains (*Salmonella typhimurium* ATCC 19585, *Escherichia coli* ATCC 43889, and *Bacillus cereus* ATCC 10361) were used. All bacterial strains were grown on tryptic soy agar (TSA) plates at 37°C for overnight and overnight cultures were prepared in TSB by inoculation of a single colony picked up from the agar plates and incubated at 37°C for overnight. Inocula of all bacteria strains were prepared by suspending overnight cultures to achieve approximately $1.5x10^8$ colony-forming units (CFU)/mL based on 0.5 McFarland standard and subsequently diluting them 1:150 in TSB.

Serial dilutions of each concentration of 50 µL SnuCalCpI were mixed with 50 µL TSB to the final volume of 100 µL in 96-well microplate. Then, each well was inoculated with 100 µL of suspension and incubated with 100 µL serially diluted SnuCalCpI at 37°C with low shaking. The growth rate was measured spectrophotometrically at 600 nm at 30 min time intervals for 16 h by using a Multiskan Go Microplate Spectrometer (Thermo Fisher Scientific Co., Vantaa, Finland).

2.2.3.2. Viable cell count

The overnight culture was subcultured using 1% inoculum to 50 mL of TSB supplemented with a range of concentrations (0, 10, 100 µg/mL) of SnuCalCpI17. During incubation at 37°C for 9 h with constant shaking at 220 rpm, aliquots (100 µL) were taken at 1.5 h time intervals from the culture medium and serially diluted in phosphate buffered saline (PBS). Aliquots of 50 µL from each of these dilutions were dropped on TSA plates and incubated at 37°C for overnight. The number of survivors (CFU/mL) was determined by counting the colonies and growth curves were constructed by plotting the log CFU/mL versus time.

3. Results and Discussion

3.1. Structural similarity

The propeptide inhibitors interact stereospecifically with target enzymes (Kwon, et al., 2018). Therefore, it was necessary to confirm that the selected eight kinds of SnuCalCpI were structurally similar to papain propeptide to determine whether they were able to inhibit papain activities. To analyze the structural similarity, homology modeling was performed between eight kinds of SnuCalCpIs and the propeptide domain of papain (Genbank: M15203. 1) obtained from *Carcia papaya*. All of SnuCalCpIs showed structural similarity of high identity of 39-60% to papain propeptide (Table 2). These results suggested that SnuCalCpIs had a possibility of interaction of SnuCalCpI with papain. For identification of motifs specific to papain propeptide, multiple alignments were performed between the SnuCalCpIs and the papain propeptide (Fig. 1). It was found that the amino acid sequences of SnuCalCpIs had the high identity with that of the papain propeptide. In particular, there were highly conserved ERFNIN (ExxxRxxxFxxNxxxIxxxN) and GNFD (GxNxFxD) motifs in SnuCalCpIs in common with the papain propeptide.

The ERFNIN motif is known as a motif existing in the α2 helix of a large number of cysteine protease propeptides (Wiederanders, Kaulmann, &

Schilling, 2003). Amino acids conserved in this motif presented on one side of the α2 helix interact with the consecutive β strand, contributing to the stabilization of the folding of propeptide (Coulombe, Grochulski, Sivaraman, Ménard, Mort, & Cygler, 1996). The GNFD motif is also present in most of the cysteine protease propeptides and located in the kink of the β-sheet that precedes the short α3 helix, which blocks opening and closing of the active site (Kwon, et al., 2015). This position of an α3 helix is stabilized by two salt bridges, one between the Asp residue of the GNFD motif and the Arg residue of the ERFNIN motif, and the other between Arg of the ERFNIN motif and Glu residue in the α3 helix (Wiederanders, Kaulmann, & Schilling, 2003). The presence of the predicted ERFNIN-GNFD motif in SnuCalCpIs indicated that SnuCalCpIs could inhibit papain activity (Kwon, et al., 2015; Martinez & Diaz, 2008; Roy, Choudhury, Aich, Dattagupta, & Biswas, 2012).

Table 2. Homology modeling of SnuCalCpIs with *Carcia papaya* papain propeptide

Unigene ID	Number of amino acid residues	Molecular weight (kDa)	Template protein	Identity (%)
SnuCalCpI02	113	13.3	*Carcia papaya* papain propeptide	60
SnuCalCpI03	106	12.5		55
SnuCalCpI08	124	14.1		39
SnuCalCpI12	107	12.8		52
SnuCalCpI14	110	12.3		46
SnuCalCpI15	117	14.0		58
SnuCalCpI16	120	14.0		47
SnuCalCpI17	119	13.7		54

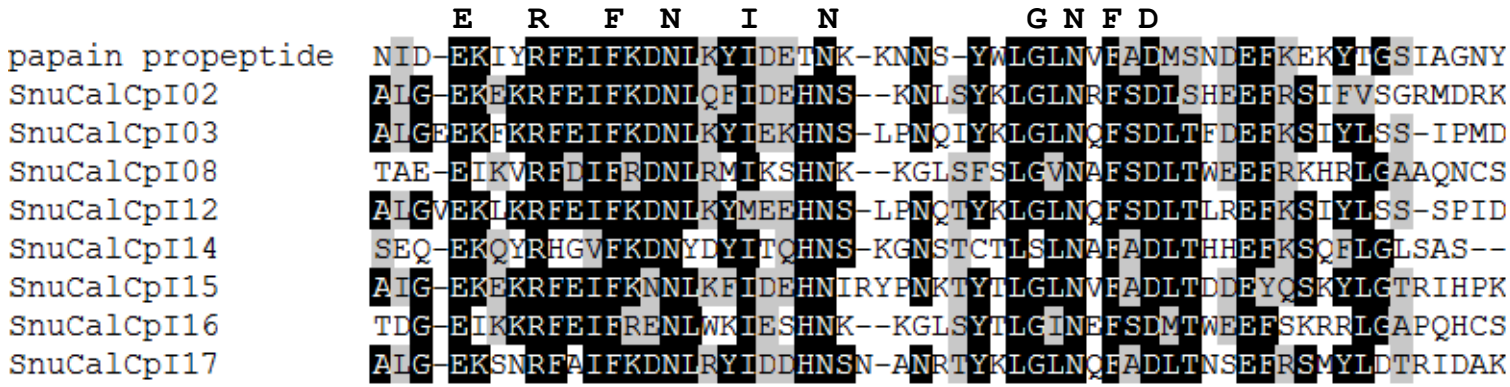

Fig. 1. Multiple protein sequence alignment of SnuCalCpIs with papain propeptide by ClustalW.

3.2. Molecular cloning of SnuCalCpI genes

Using cDNA extracted and synthesized from *Calotropis procera* R. Br., SnuCalCpI08, SnuCalCpI14, and SnuCalCpI17 were amplified by PCR, and then they were examined by electrophoresis whether there was gene amplification. The size of SnuCalCpI genes, which were designed to include the in-fusion cloning sites, was ranged from 360 to 400 bp, which could be found between 250 bp and 500 bp of the marker (Fig. 2). After the impurity removal and purification of the amplified gene, the pET29b(+)-SnuCalCpI-C6His plasmid was produced by cloning of DNA fragment into the pET29b(+) expression vector and successfully transforming it into *E. coli* DH5α (Fig. 3). The pET29b(+) expression vector belonged to the pET system with a strong T7 promoter and was used for the high expression rate of SnuCalCpI. A primer specific to pET29b(+) was made and screened by colony PCR, and the size of the product was observed by electrophoresis. Finally, the presence of SnuCalCpI gene in the recombinant plasmid was verified by DNA sequencing.

As for protein expression, most problems stem from the different codon usage between eukaryotic cell and prokaryotic cell proteins, and in many cases, rare codons can especially be a problem. Amino acids are encoded by more than one codon, and each organism is known to have codon usage bias in the

use of the 61 available amino acid codons (Terpe, 2006). The rare codon means a codon that is rarely used in an organism. Thus, when genes derived from other organisms are expressed in *E. coli*, the rarity or deficiency of codons may interfere with protein translation or expression due to the different codon usage bias among organisms (Kane, 1995). Based on this, codon of SnuCalCpI was examined (Table 3), and the frequency of rare codons for *E. coli* such as Gly codon GGA, Ile codon AUA, Leu codon CUA, Pro codon CCC, and Arg codon AGG, AGA, CGG, and CGA was converted to percentages. SnuCalCpI08 and SnuCalCpI17 among those, which were not expressed in the existing *E. coli* system, were found to hold 10% or more of rare codons, unlike other inhibitor genes. In addition to the overall rare codon ratio, SnuCalCpI08 and SnuCalCpI17 were found to have AGG and AGA ratio of 4.84% and 5.04%, respectively, which was relatively high compared to other inhibitors. In particular, AGG and AGA, which are rare Arg codons, have the lowest frequency of 0.14% or below and 0.21%, respectively, in *E. coli* (Kane, 1995). Especially, the presence of these codons is known to act sensitively to protein expression (Kane, 1995; Tegel, Tourle, Ottosson, & Persson, 2010). Therefore, *E. coli* system capable of controlling rare codons was used to express the proteins.

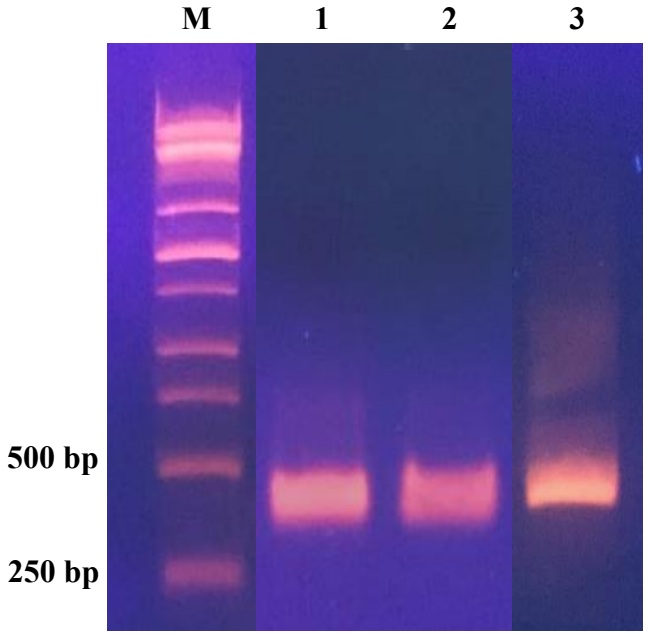

Fig. 2. Gel electrophoresis of amplified SnuCalCpI genes by PCR; lane M, DNA size marker; lane 1, SnuCalCpI08; lane 2, SnuCalCpI14; lane 3, SnuCalCpI17 (95°C, 30 s denaturation; 53°C, 30 s annealing; 72°C, 1 min elongation; 35 cycles, 2X TOPsimpleTM PreMIX-Forte).

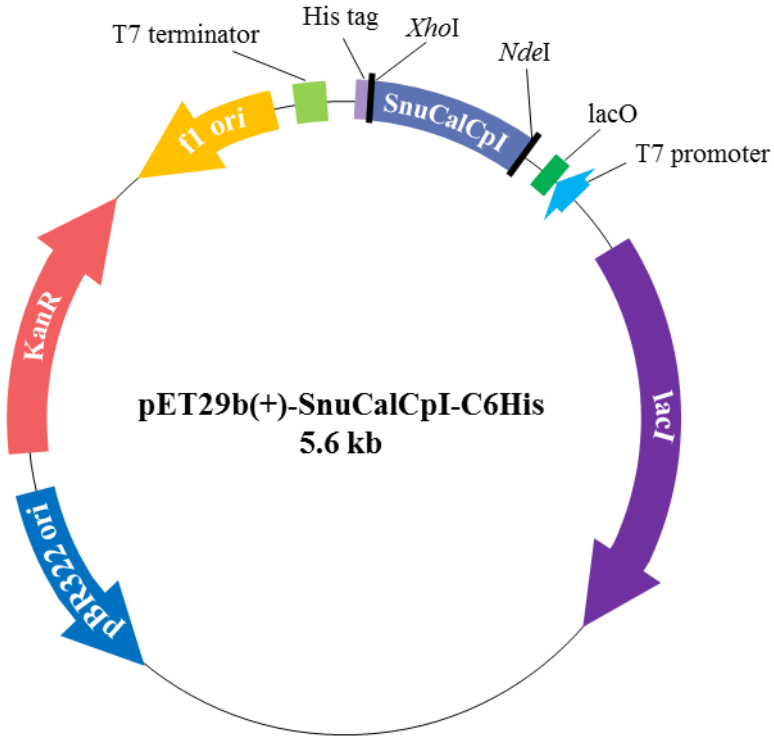

Fig. 3. Construction map of expression vector (pET29b(+)-SnuCalCpI-C6His).

Table 3. The ratio of rare codons for *E. coli* in the SnuCalCpI genes

Amino Acid	Rare codon	Percentage of rare codons for *E. coli* in SnuCalCpI genes (%)							
		02	**03**	**08**	**12**	**14**	**15**	**16**	**17**
Gly	**GGA**	0	0.94	4.84	0.94	1.82	1.71	3.33	1.68
Ile	**AUA**	0	3.77	1.61	1.87	0	0.86	0.83	1.68
Leu	**CUA**	0	0	0.81	0	0	0.86	1.67	0
Pro	**CCC**	0	1.89	0	1.87	0.91	0	0	0
Arg	**AGG**	3.54	0	4.84	0.94	0	0.86	1.67	3.36
	AGA	0	0.94	0	0.94	0.91	2.56	1.67	1.68
	CGG	1.77	0	0.81	0.94	0	0	0	1.68
	CGA	0.89	0.94	0	0.94	2.73	0	0	2.52
Total		6.20	8.50	**12.9**	8.41	6.36	6.84	9.17	**12.6**

3.3. Expression and purification of SnuCalCpIs

The recombinant plasmid pET29b(+)-SnuCalCpI-C6His for protein expression was transformed into the *E. coli* Rosetta(DE3). *E. coli* Rosetta(DE3) contains a plasmid of pRARE (Cm^R) encoding rare codon tRNAs for the AGG, AGA, AUA, CUA, CCC, and GGA codons that are used commonly in eukaryotes but rarely in *E. coli* (Fu, Lin, & Cen, 2007). Hence, this strain that can directly resolve protein non-expression resulting from codon bias by co-expression of the gene that encodes a rare tRNAs in *E. coli* (Fu, Lin, & Cen, 2007; Gustafsson, Govindarajan, & Minshull, 2004; Kane, 1995; Tegel, Tourle, Ottosson, & Persson, 2010). The results of expression at each temperature with 0.1 mM IPTG induction are shown in Table 4. SnuCalCpI08 and SnuCalCpI17 were successfully expressed in *E. coli* Rosetta(DE3); however, most of them were insoluble at 37°C. To increase solubility, the protein synthesis rate was controlled by lowering the temperature (to 18°C). As a result, while the expressed protein of SnuCalCpI08 was insoluble at 37°C, 30°C, and 18°C, the expressed protein of SnuCalCpI17 was soluble at 18°C in a large amount when it was induced by 0.1 mM IPTG for 16 h (Fig. 4 (A)). Based on these results, it could be found that rare codons significantly affected the expression of SnuCalCpI08 and SnuCalCpI17. On the other hand, SnuCalCpI14 was not expressed. Since SnuCalCpI14 has a low ratio of overall rare codons and its

ratio of AGA/AGG is also low at 0.91%, it could not be expressed using *E. coli* Rosetta(DE3) which is designed to control rare codons. SnuCalCpI17 that was expressed in the soluble form was purified by Ni-NTA column it was confirmed by SDS-PAGE whether the target protein was successfully purified (Fig. 4 (B)).

Table 4. Expression results of SnuCalCpI08, SnuCalCpI14, and SnuCalCpI17 in *E. coli* Rosetta(DE3)

Host	Plasmid name	37℃		30℃		18℃	
		Expression	**Solubility**	**Expression**	**Solubility**	**Expression**	**Solubility**
E. coli Rosetta (DE3)	pET29b(+)-SnuCalCpI08-C6His	++	I	+	I	N.E.	
	pET29b(+)-SnuCalCpI14-C6His	N.E.		N.E.		N.E.	
	pET29b(+)-SnuCalCpI17-C6His	++	I	++	S < I	+	S > I

1) ++, high level of expression; +, low level of expression; N.E., no expression.
2) S, soluble; I, insoluble.

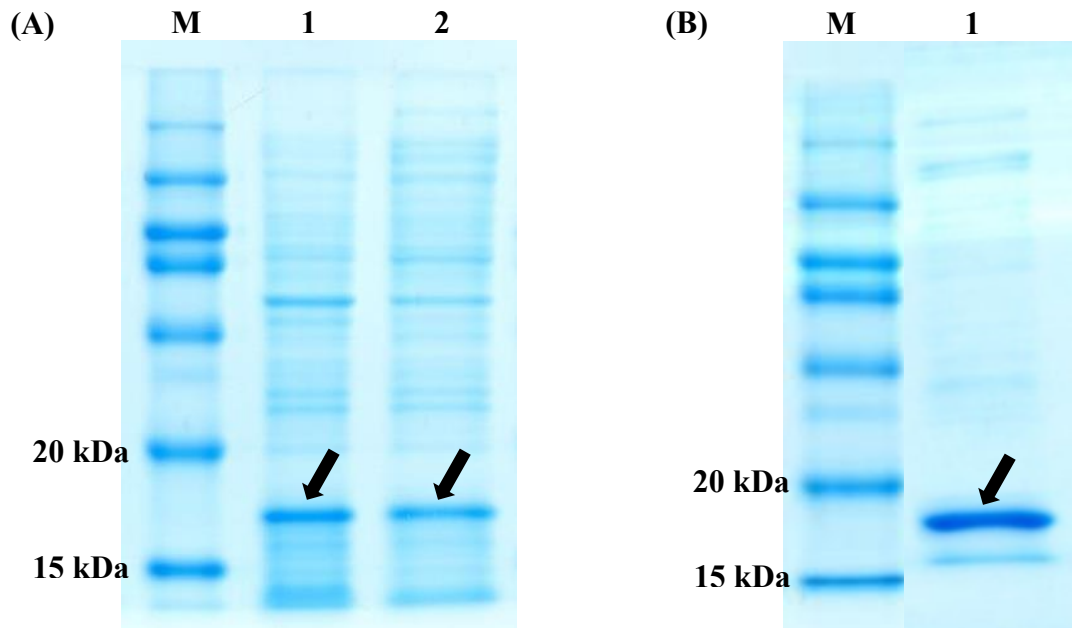

Fig. 4. SDS-PAGE analysis of SnuCalCpI17. (A) Expression in recombinant *E. coli* Rosetta(DE3); lane M, protein size marker; lane 1, cell lysate of *E. coli* Rosetta(DE3); lane 2, soluble cytoplasmic proteins and (B) Ni-NTA affinity-purified SnuCalCpI17; lane M : protein size marker; lane 1, SnuCalCpI17 after purification.

3.4. Inhibitory characterization

3.4.1. Inhibitory activity against papain

From the result of the previous experiment, it could be found that SnuCalCpI03 and SnuCalCpI15 have a similar structure to papain propeptide and are the same in SnuCalCpI17. To confirm the possibility of SnuCalCpI17 as a cysteine protease inhibitor like SnuCalCpI03 and SnuCalCpI15, it was necessary to find out whether it inhibit papain activity. Therefore, papain inhibition assay was performed to measure the inhibitory activities of the expressed SnuCalCpI17 and previously obtained SnuCalCpI03 and SnuCalCpI15 at pH 7.0 and 37°C, which is the optimal condition for papain. The half-maximal inhibitory concentrations (IC_{50}) were used to evaluate the inhibitory effect of the three recombinant propeptides (Fig. 5). As a result, the IC_{50} values of the SnuCalCpI03 and SnuCalCpI15 were 103.702±2.060 μg/mL and 23.900±0.654 μg/mL, respectively, and the value of SnuCalCpI17 was 105.671±9.857 μg/mL. This indicates that SnuCalCpI17 demonstrates an inhibitory activity against papain at pH 7.0. However, as shown in Fig. 5, SnuCalCpI17 showed different inhibitory patterns from SnuCalCpI03 and SnuCalCpI15. ERFNIN and GNFD motifs in SnuCalCpIs are mediators of propeptide inhibitory activity, which induce the appropriate protein folding to

interact with the catalytic site of cysteine protease (K. C. Pandey, Barkan, Sali, & Rosenthal, 2009). It is predicted that the different inhibitory patterns of SnuCalCpIs were caused by the differences of some residues that interact with residues around a catalytic site of cysteine protease mature domain.

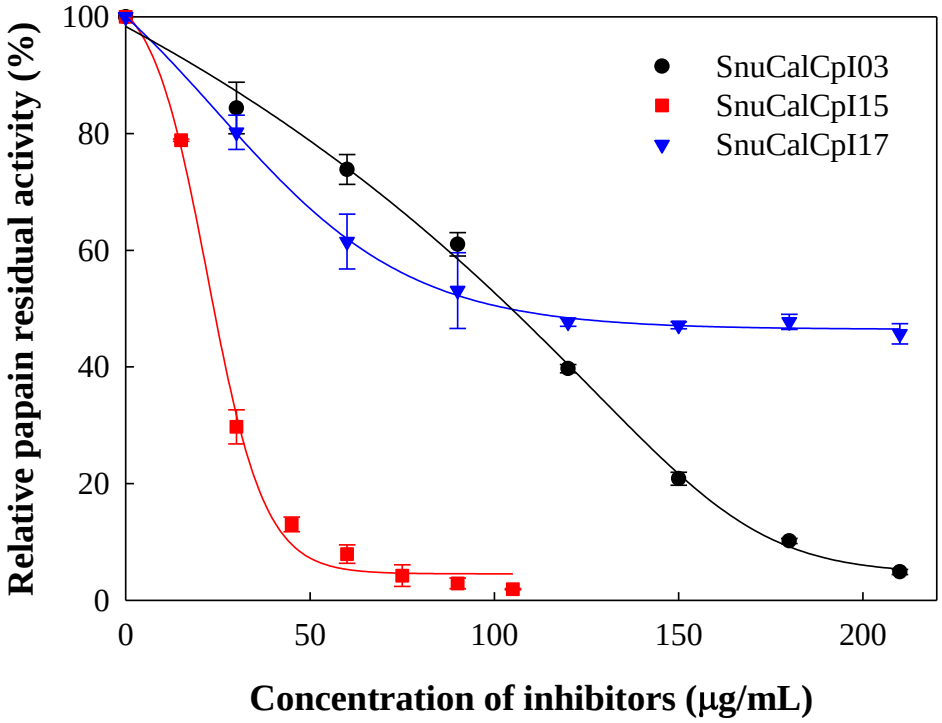

Fig. 5. Inhibitory activity of SnuCalCpIs against papain; ●: SnuCalCpI03, ■: SnuCalCpI15, ▼: SnuCalCpI17.

3.4.2. Slow-binding inhibition of SnuCalCpIs

Based on the IC_{50} values for papain, the difference in inhibition mechanism—the difference derived from kinetic mechanism—among SnuCalCpI03, SnuCalCpI15, and SnuCalCpI17 was examined. In the long-term enzymatic reaction, the three kinds of SnuCalCpI commonly showed the lag phase at the beginning of the reaction as the inhibitor concentration increased (Fig. 6). This was caused by slow-binding inhibition, and there were many examples of propeptide-type slow-binding inhibitor (Billington, Mason, Magny, & Mort, 2000; Copeland & Robert, 2002; Fox, De Miguel, Mort, & Storer, 1992; Kwon, et al., 2018). In the case of classical inhibition, while the equilibrium between the enzyme (E), inhibitor (I) and enzyme-inhibitor (EI) complex is established rapidly in the reaction of E with I, it takes from several seconds to minutes for the slow-binding inhibitor to establish the equilibrium between the E, I and EI complex. Thus, slow-binding inhibitors mean the compounds that inhibit the target enzyme in a time-dependent manner (Baici, 2015; Billington, Mason, Magny, & Mort, 2000; Copeland & Robert, 2002). Since EI complex is formed depending on time, it is possible to determine steady-state kinetic parameters (v_i, v_s, and k_{obs}) by using the progress curves of product formation.

Generally, the slow-binding inhibition follows two mechanisms (Cha, 1975; R. A. Copeland, Pompliano, & Meek, 2006; Fox, De Miguel, Mort, & Storer, 1992; Holdgate, Meek, & Grimley, 2017; Kwon, et al., 2018; Marko & Jure, 2004). Mechanism A (Fig. 7 (A)), which is the representative mechanism for the single-step process, is a direct-binding model in which the k_3[I] (association constant) and k_4 (dissociation constant) are both low due to the slow binding of the inhibitor to the enzyme active site. Mechanism B (Fig. 7 (B)), a two-step process, is an isomerization model in which the inhibitor binds rapidly to the enzyme first but is then a slowly converted to the higher affinity EI* complex.

The association and dissociation constants of the slow-binding inhibitor can be determined by plotting the apparent first-order rate constant (k_{obs}) in accordance with the inhibitor concentration. To calculate the respective parameters, the linear portion of the uninhibited reaction progress curve, where substrate depletion does not appear under the inhibitor-free papain reaction, was established and the papain inhibition assay was performed based on the reaction time of 18 min (Fig. 8). As for the papain inhibition pattern, the curve of its reaction velocity showed constant changes according to the time and inhibitor concentration, gradually reaching a steady state. Thus, it was possible to determine the steady-state kinetic parameters of the

progress curves of product formation using the following equation (Fig. 9).

$$[\mathrm{P}] = v_i t + \frac{(v_i - v_s)[1 - \exp(-k_{obs} t)]}{k_{obs}}$$

As the concentration of SnuCalCpI increase, k_{obs} remained linear, which indicated mechanism A, a single-step slow-binding inhibition. Therefore, the apparent dissociation constant (K_i^{app}) of the steady-state SnuCalCpIs were calculated using the following equation (Copeland & Robert, 2002).

$$k_{obs} = k_3[\mathrm{I}] + k_4$$

$$k_{obs} = k_4 + \frac{k_4}{K_i^{app}}[\mathrm{I}]$$

K_i^{app} of SnuCalCpI03, SnuCalCpI15, and SnuCalCpI17 against papain was calculated to be 28.70 μg/mL, 13.63 μg/mL, and 75.80 μg/mL, respectively. This result verified that the three kinds of SnuCalCpI are slow-binding inhibitors to papain.

Identification of the binding kinetics is important to enhance efficiency and effectiveness of drug discovery and development. From the current point of view, one of the most important factors for sustained drug efficacy *in vivo* is the residence time of the drug molecule to the target molecule, not the apparent affinity to the drug target itself (R. A. Copeland, Pompliano, & Meek, 2006;

A. R. Johnson, 2013; Lu & Tonge, 2010; Schindler, 2018). To secure a long residence time, there is a strategy to avoid reaching equilibrium for competition between the substrate and the inhibitor for the active site of the target. There are numerous ways to do that, and slow-binding inhibition is one of them (Swinney, 2004). Slow-binding inhibitor has a long residence time because they dissociate slowly from their targets. Slow-binding inhibitor will remain bound to the targets and effectively inhibit their target even when the concentration of the inhibitor is well below the IC_{50} of the inhibitor and allows for lower inhibitor doses to be used (A. R. Johnson, 2013). This implies that SnuCalCpI with a long residence time derived from slow-binding inhibition can be used as an effective inhibitor drug for diseases such as cancer metastasis, osteoporosis, and viral, bacterial, and parasitic infections in which papain-like cysteine protease plays crucial roles.

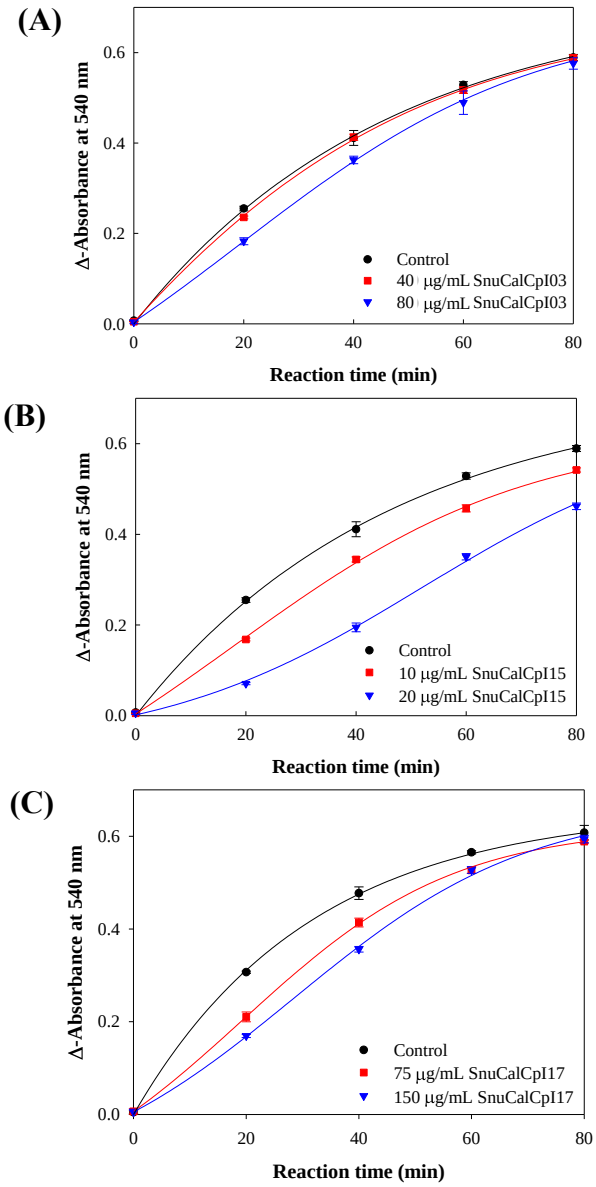

Fig. 6. Progress curves for inhibition of 25 μg/mL papain by SnuCalCpIs for 80 min. The reactions were started by adding 0.5 mM α-N-benzoyl-DL-arginine-3-naphthylamide (BANA) after 10 min of pre-incubation of papain with (A) SnuCalCpI03, (B) SnuCalCpI15, and (C) SnuCalCpI17.

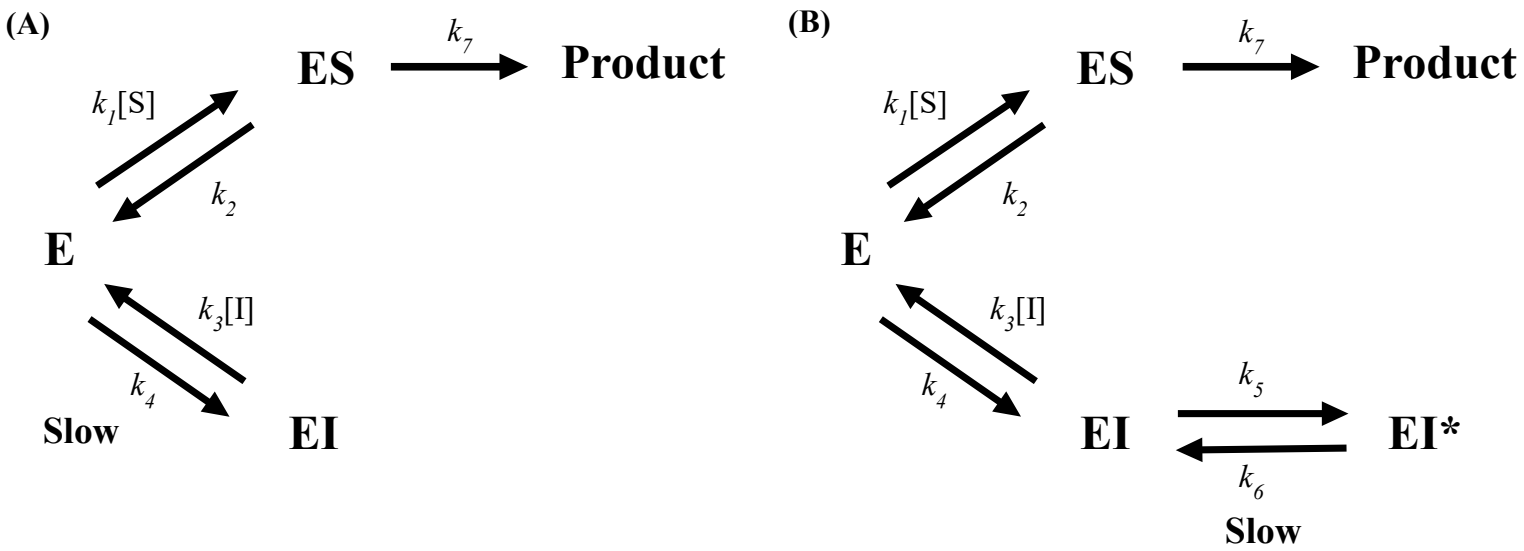

Fig. 7. Mechanism of slow-binding inhibition. (A) Single-step slow-binding mechanism, in which the binding of the inhibitor to enzyme (EI) complex is slow. (B) Two-step slow-binding mechanism, in which EI is formed rapidly then slow conformational change to the higher affinity EI* complex; ES, non-covalent enzyme–substrate complex; [I], inhibitor concentration; $k_{1–7}$, rate constant of the specific step; [S], substrate concentration.

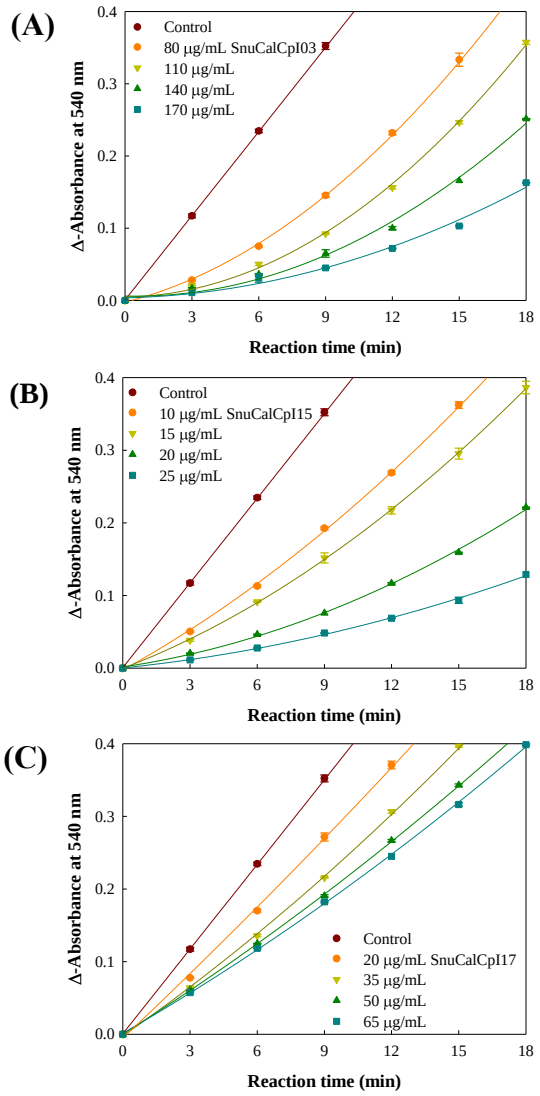

Fig. 8. Progress curves for inhibition of 25 µg/mL papain in the presence of varying concentrations of (A) SnuCalCpI03, (B) SnuCalCpI15, and (C) SnuCalCpI17 for 18 min. The reactions were started by adding 1.5 mM α-N-benzoyl-DL-arginine-3-naphthylamide (BANA) after 10 min of pre-incubation of papain and SnuCalCpIs. The uninhibited papain (control) displays a linear progress curve.

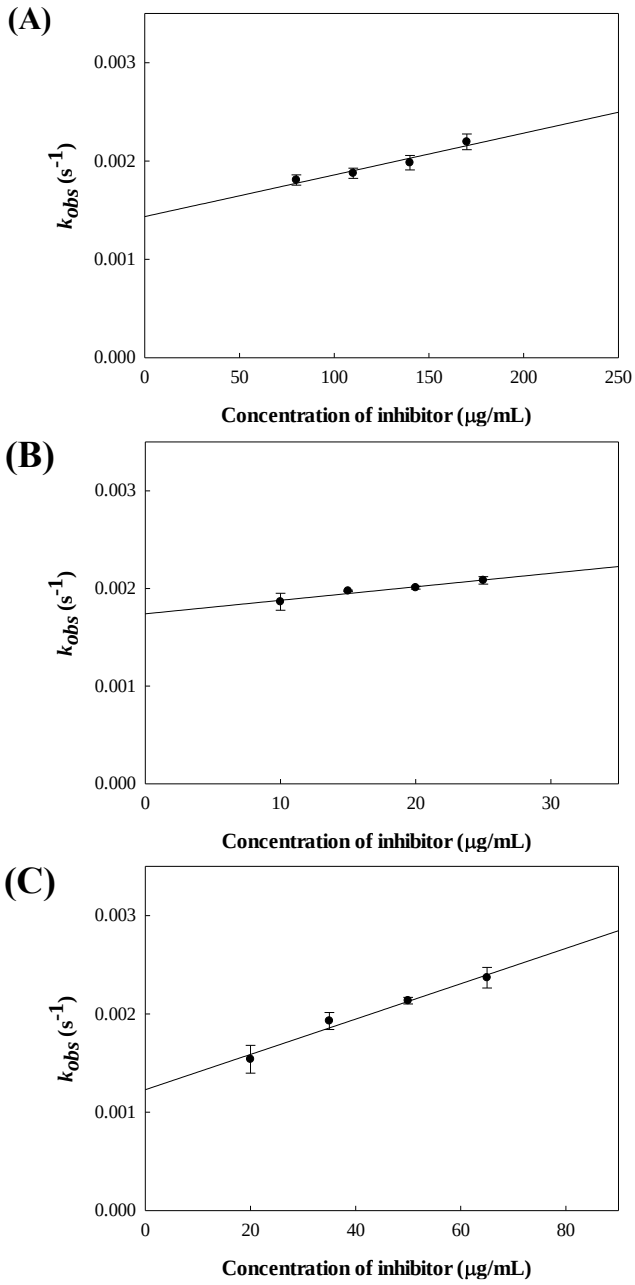

Fig. 9. Replot of the observed rate constant (k_{obs}) from the inhibition of papain by (A) SnuCalCpI03, (B) SnuCalCpI15, and (C) SnuCalCpI17.

3.4.3. Inhibition of bacterial growth

In order to identify the functionality of SnuCalCpI, it was studied to measure its inhibitory effect on bacterial cysteine protease activity. In fact, there were examples of antibacterial properties of cysteine protease inhibitors (Popovic, Andjelkovic, Grozdanovic, Aleksic, & Gavrovic-Jankulovic, 2013; Zindel, Kaman, Fröls, Pfeifer, Peters, Hays, et al., 2013). In this regard, three Gram-positive bacteria and three Gram-negative bacteria were selected, and a micro-dilution test was performed (Table 5). SnuCalCpI affected the growth curve of *L. welshimeri* only, Gram-positive bacteria (Fig. 10). In order to verify this, the viable cell count was performed with SnuCalCpI17 (Fig. 11), and the result showed a similar curve to that measured by absorbance. Specifically, the inhibited growth curve was identified during the logarithmic phase, and attention was drawn to the inhibitory activity of SnuCalCpI on bacterial cysteine protease.

The cell wall of the bacteria consists of peptidoglycan (PG), an exoskeleton, and is a physical support to protect bacteria from environmental attack. During the bacterial growth or cell division, the PG is steadily broken down by peptidoglycan-cleaving enzymes, and more than 50% of the PG is reconstituted. This phenomenon is called the cell wall turnover, and this

process represents a significant loss of resource if the liberated cell-wall fragments are not recovered and recycled (Carvalho, Sousa, & Cabanes, 2014; J. W. Johnson, Fisher, & Mobashery, 2013; Reith & Mayer, 2011). Briefly describing bacterial cell wall turnover, the PG precursor is biosynthesized in the bacterial cytoplasm, and LipidII, the final intermediate, is translocated into the periplasm for *de novo* PG polymer synthesis. Then, the PG polymer is reconstituted during bacterial growth and cell division (J. W. Johnson, Fisher, & Mobashery, 2013). In Gram-positive bacteria, as PG polymerization progresses, the cell wall material moves from inside to outside, making the outermost layer aged and maximally stretched. The outermost PG is then hydrolyzed by potentially autolytic and extracellular enzymes, the autolysin (Carvalho, Sousa, & Cabanes, 2014; J. W. Johnson, Fisher, & Mobashery, 2013).

There are various sorts of autolysin, and among them, NlpC/P60 domain related to the CHAP (cysteine, histidine-dependent amidohydrolase/peptidase) superfamily is known as cysteine protease having catalytic activity as a result of conserved Cys and His residues. It is known as a papain-like cysteine protease belonging to clan CA, structurally taking a papain-like fold form (Anantharaman & Aravind, 2003; Carvalho, Sousa, & Cabanes, 2014; Xu, Sudek, McMullan, Miller, Geierstanger, Jones, et al., 2009). Among the

proteins with the NlpC/P60 domain, p60 protein is found specifically in *Listeria* (Anantharaman & Aravind, 2003; Carvalho, Sousa, & Cabanes, 2014). In fact, the analysis of the sequence of the p60 protein (Genbank: GenBank: M80348.1) present in *L. welshimeri* revealed that p60 protein was specifically present in *Listeria* (data not shown).

The p60 protein, which is specifically present in *Listeria*, was inhibited due to structural similarity with SnuCalCpI, and it was expected that the inhibition of the p60 protein influenced cell wall turnover system, affecting morphology and growth of *Listeria* (Pilgrim, Kolb-Mäurer, Gentschev, Goebel, & Kuhn, 2003; Popovic, Andjelkovic, Grozdanovic, Aleksic, & Gavrovic-Jankulovic, 2013; Zindel, et al., 2013). Although additional experiments are required to confirm this hypothesis, this study showed the possibility of SnuCalCpI to affect the growth of *Listeria.*

Table 5. Inhibitory effects of SnuCalCpIs on bacterial growth

Inhibitors	Gram-positive			Gram-negative		
	S. aureus	*L. welshimeri*	*B. cereus*	*E. coli*	*S. thyphimurium*	*P. aeruginosa*
SnuCalCpI 03	-	+	-	-	-	-
SnuCalCpI 15	-	+	-	-	-	-
SnuCalCpI 17	-	+	-	-	-	-

1) +, inhibition; -, no inhibition

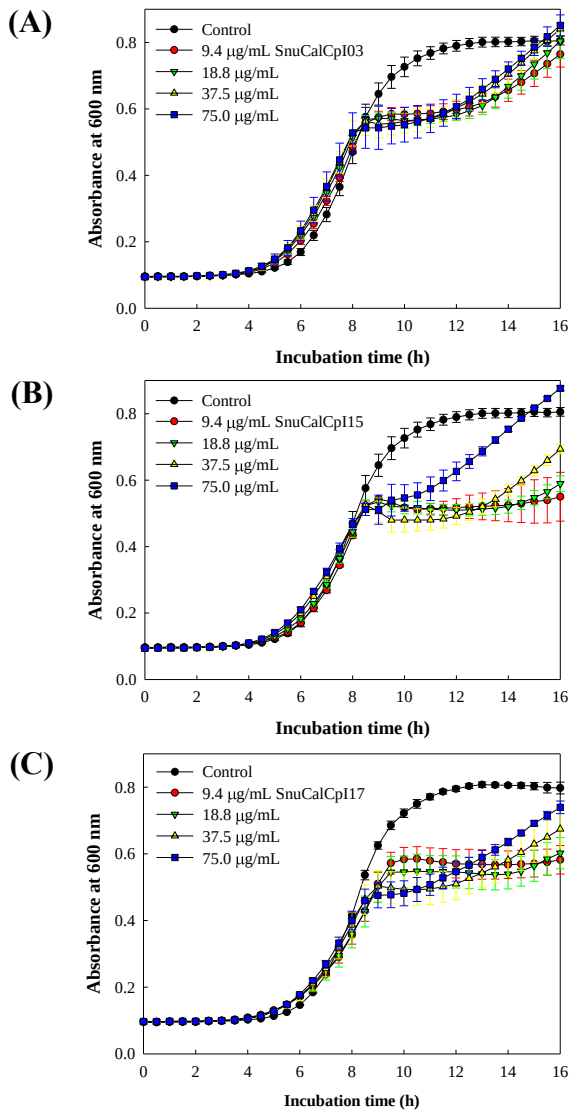

Fig. 10. Effect of SnuCalCpIs on the growth curve of *Listeria welshimeri* ATCC 35897 in TSB as determined by absorbance at 600 nm. *L. welshimeri* was cultured at 37°C for 16 h in the presence of (A) SnuCalCpI03, (B) SnuCalCpI15, and (C) SnuCalCpI17. Error bars represent the standard errors of the means from 3 independent experiments.

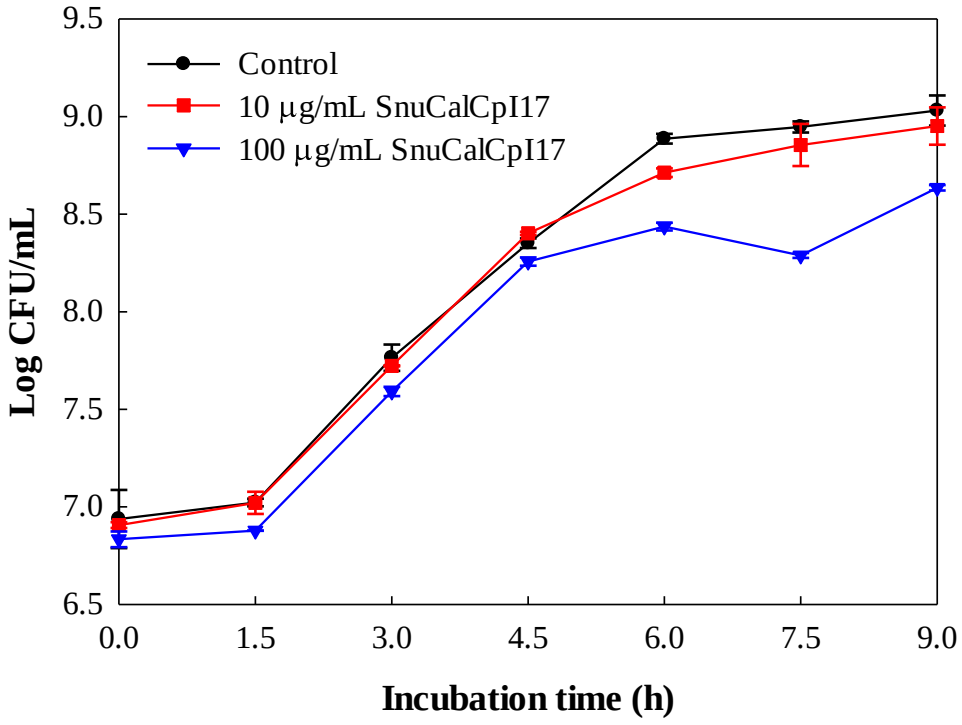

Fig. 11. Effect of SnuCalCpI17 on the growth curve of *L. welshimeri* ATCC 35897 in TSB as determined by colony counting.

4. Conclusion

In this study, a novel SnuCalCpI was cloned and expressed in a pET system and *E. coli* Rosetta(DE3) providing tRNAs for six rare codons. It has been identified that rare codons had a great influence on the expression of the protein. SnuCalCpI17 was possible to obtain as a soluble form, and inhibition assay against papain was carried out on SnuCalCpI17 and previously obtained SnuCalCpI03 and SnuCalCpI15. SnuCalCpI03, SnuCalCpI15, and SnuCalCpI17 commonly follow a slow-binding inhibition mechanism against papain at pH 7.0, and their K_i^{app} were 28.70 μg/mL, 13.63 μg/mL, and 75.80 μg/mL, respectively. Slow-binding inhibitor has the slow dissociation from the inhibitor-target complex, resulting in long residence time *in vivo*, and can effectively inhibit the target at doses far below the IC_{50} of the inhibitor. These results imply that SnuCalCpIs obtained from natural sources not only have the advantages of slow-binding inhibitions but also are effective inhibitors of papain-like cysteine protease in a broad spectrum of diseases as well as bacterial and parasitic infections.

5. References

Abe, M., Abe, K., Kuroda, M., & Arai, S. (1992). Corn kernel cysteine proteinase inhibitor as a novel cystatin superfamily member of plant origin. *European Journal of Biochemistry, 209*(3), 933-937.

Anantharaman, V., & Aravind, L. (2003). Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome Biology, 4*(2), R11-R11.

Baici, A. (2015). Slow-onset enzyme inhibition. In *Kinetics of Enzyme-Modifier Interactions: Selected Topics in the Theory and Diagnosis of Inhibition and Activation Mechanisms*, (pp. 367-444). Vienna: Springer Vienna.

Billington, C. J., Mason, P., Magny, M.-C., & Mort, J. S. (2000). The slow-binding inhibition of cathepsin K by its propeptide. *Biochemical and Biophysical Research Communications, 276*(3), 924-929.

Carvalho, F., Sousa, S., & Cabanes, D. (2014). How *Listeria monocytogene*s organizes its surface for virulence. *Frontiers in Cellular and Infection Microbiology, 4*, 48.

Cha, S. (1975). Tight-binding inhibitors—I: kinetic behavior. *Biochemical Pharmacology, 24*(23), 2177-2185.

Choedon, T., Mathan, G., Arya, S., Kumar, V. L., & Kumar, V. (2006). Anticancer and cytotoxic properties of the latex of *Calotropis procera* in a transgenic mouse model of hepatocellular carcinoma. *World Journal of Gastroenterology : WJG, 12*(16), 2517-2522.

Copeland, & Robert, A. (2002). Time-dependent inhibition. In *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, (pp. 318-349). New York: Wiley-VCH.

Copeland, R. A., Pompliano, D. L., & Meek, T. D. (2006). Drug–target residence time and its implications for lead optimization. *Nature Reviews Drug Discovery, 5*, 730.

Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J. S., & Cygler, M. (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *The EMBO Journal, 15*(20), 5492-5503.

Dewan, S., Sangraula, H., & Kumar, V. L. (2000). Preliminary studies on the analgesic activity of latex of *Calotropris procera*. *Journal of Ethnopharmacology, 73*(1), 307-311.

F., K. T., W., O. T. P., & C., C. J. (2005). SpeB–Spi: a novel protease–inhibitor pair from *Streptococcus pyogenes*. *Molecular Microbiology, 57*(3), 650-666.

Fox, T., De Miguel, E., Mort, J. S., & Storer, A. C. (1992). Potent slow-binding inhibition of cathepsin B by its propeptide. *Biochemistry, 31*(50), 12571-12576.

Fu, W., Lin, J., & Cen, P. (2007). 5-Aminolevulinate production with recombinant *Escherichia coli* using a rare codon optimizer host strain. *Applied Microbiology and Biotechnology, 75*(4), 777-782.

Gustafsson, C., Govindarajan, S., & Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology, 22*(7), 346-353.

Hayes, M. (2014). Bioactive peptides and their potential use for the prevention of diseases associated with Alzheimer's disease and mental health disorders. *Ann Psychiatry Ment Health, 2*(3), 1017.

Hayes, M., & Tiwari, B. K. (2015). Bioactive carbohydrates and peptides in foods: an overview of sources, downstream processing steps and associated bioactivities. *International Journal of Molecular Sciences, 16*(9), 22485-22508.

Holdgate, G. A., Meek, T. D., & Grimley, R. L. (2017). Mechanistic enzymology in drug discovery: a fresh perspective. *Nature Reviews Drug Discovery, 17*, 115.

Hummel, K. M., Petrow, P. K., Franz, J. K., Müller-Ladner, U., Aicher, W. K.,

Gay, R. E., Brömme, D., & Gay, S. (1998). Cysteine proteinase cathepsin K mRNA is expressed in synovium of patients with rheumatoid arthritis and is detected at sites of synovial bone destruction. *The Journal of Rheumatology, 25*(10), 1887-1894.

Johnson, A. R. (2013). *In vitro* and *in vivo* assays. In J. J. L. E. J. Corey (Ed.), *Drug Discovery: Practices, Processes, and Perspectives*, (pp. 67-98). Hoboken, N.J: John Wiley & Sons.

Johnson, J. W., Fisher, J. F., & Mobashery, S. (2013). Bacterial cell-wall recycling. *Annals of the New York Academy of Sciences, 1277*(1), 54-75.

Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology, 6*(5), 494-500.

Khan, A. R., & James, M. N. (1998). Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Science : A Publication of the Protein Society, 7*(4), 815-836.

Kruger, N. J. (1994). The bradford method for protein quantitation. In J. M. Walker (Ed.), *Basic Protein and Peptide Protocols*, (pp. 9-15). Totowa, NJ: Humana Press.

Kumar, V. L., & Basu, N. (1994). Anti-inflammatory activity of the latex of

Calotropis procera. *Journal of Ethnopharmacology, 44*(2), 123-125.

Kumar, V. L., & Roy, S. (2007). *Calotropis procera* latex extract affords protection against inflammation and oxidative stress in Freund's complete adjuvant-induced monoarthritis in rats. *Mediators of Inflammation, 2007*, 47523.

Kwon, C. W., Park, K.-M., Kang, B.-C., Kweon, D.-H., Kim, M.-D., Shin, S. W., Je, Y. H., & Chang, P.-S. (2015). Cysteine protease profiles of the medicinal plant *Calotropis procera* R. Br. revealed by *de novo* transcriptome analysis. *PLoS One, 10*(3), e0119328.

Kwon, C. W., Yang, H., Yeo, S., Park, K.-M., Jeong, A. J., Lee, K. W., Ye, S.-K., & Chang, P.-S. (2018). Molecular cloning and anti-invasive activity of cathepsin L propeptide-like protein from *Calotropis procera* R. Br. against cancer cells. *Journal of Enzyme Inhibition and Medicinal Chemistry, 33*(1), 657-664.

Lu, H., & Tonge, P. J. (2010). Drug-target residence time: critical information for lead optimization. *Current opinion in chemical biology, 14*(4), 467-474.

Lukomski, S., Montgomery, C. A., Rurangirwa, J., Geske, R. S., Barrish, J. P., Adams, G. J., & Musser, J. M. (1999). Extracellular cysteine protease produced by *Streptococcus pyogenes* participates in the pathogenesis

of invasive skin infection and dissemination in mice. *Infection and Immunity, 67*(4), 1779-1788.

Marko, G., & Jure, S. (2004). Slow-binding inhibition: a theoretical and practical course for students. *Biochemistry and Molecular Biology Education, 32*(4), 228-235.

Martinez, M., & Diaz, I. (2008). The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. *BMC Evolutionary Biology, 8*(1), 198.

McGrath, M. E. (1999). The lysosomal cysteine proteases. *Annual Review of Biophysics and Biomolecular Structure, 28*(1), 181-204.

Ovat, A., Muindi, F., Fagan, C., Brouner, M., Hansell, E., Dvořák, J., Sojka, D., Kopáček, P., McKerrow, J. H., Caffrey, C. R., & Powers, J. C. (2009). Aza-peptidyl Michael acceptor and epoxide inhibitors—potent and selective inhibitors of *Schistosoma mansoni* and *Ixodes ricinus* Legumains (asparaginyl endopeptidases). *Journal of Medicinal Chemistry, 52*(22), 7192-7210.

Pandey, K. C. (2013). Cysteine proteases of human malaria parasites. In S. Chakraborti, Dhalla, Naranjan S. (Ed.), *Proteases in Health and Disease*, (pp. 121-134): Springer, New York, NY.

Pandey, K. C., Barkan, D. T., Sali, A., & Rosenthal, P. J. (2009). Regulatory

elements within the prodomain of Falcipain-2, a cysteine protease of the malaria parasite *Plasmodium falciparum*. *PLoS One, 4*(5), e5694.

Pilgrim, S., Kolb-Mäurer, A., Gentschev, I., Goebel, W., & Kuhn, M. (2003). Deletion of the gene encoding p60 in *Listeria monocytogenes* leads to abnormal cell division and loss of actin-based motility. *Infection and Immunity, 71*(6), 3473-3484.

Popovic, M., Andjelkovic, U., Grozdanovic, M., Aleksic, I., & Gavrovic-Jankulovic, M. (2013). *In vitro* antibacterial activity of cysteine protease inhibitor from kiwifruit (*Actinidia deliciosa*). *Indian Journal of Microbiology, 53*(1), 100-105.

Radim, V., Matthias, B., Knut, B., & Tanja, S. (2006). Inhibitors of cysteine proteases. *Current Topics in Medicinal Chemistry, 6*(4), 331-353.

Reith, J., & Mayer, C. (2011). Peptidoglycan turnover and recycling in Gram-positive bacteria. *Applied Microbiology and Biotechnology, 92*(1), 1.

Rosenthal, P. J. (2004). Cysteine proteases of malaria parasites. *International Journal for Parasitology, 34*(13), 1489-1499.

Roy, S., Choudhury, D., Aich, P., Dattagupta, J. K., & Biswas, S. (2012). The structure of a thermostable mutant of pro-papain reveals its activation mechanism. *Acta Crystallographica, Section D: Biological Crystallography, 68*(Pt 12), 1591-1603.

Schindler, M. E. a. J. (2018). Impact and evolution of biophysics in medicinal chemistry. In A. Canales (Ed.), *Biophysical Techniques in Drug Discovery*, (pp. 1-22): Royal Society of Chemistry.

Scott, C. J., & Taggart, C. C. (2010). Biologic protease inhibitors as novel therapeutic agents. *Biochimie, 92*(11), 1681-1688.

Swinney, D. C. (2004). Biochemical mechanisms of drug action: what does it take for success? *Nature Reviews Drug Discovery, 3*, 801.

TAKAYUKI, S., & L., V. D. H. R. A. (2008). Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. *Molecular Plant Pathology, 9*(1), 119-125.

Tegel, H., Tourle, S., Ottosson, J., & Persson, A. (2010). Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). *Protein Expression and Purification, 69*(2), 159-167.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology, 72*(2), 211.

Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B., & Turk, D. (2012). Cysteine cathepsins: from structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta, Proteins and Proteomics,*

1824(1), 68-88.

Uhlig, T., Kyprianou, T., Martinelli, F. G., Oppici, C. A., Heiligers, D., Hills, D., Calvo, X. R., & Verhaert, P. (2014). The emergence of peptides in the pharmaceutical business: from exploration to exploitation. *EuPA Open Proteomics, 4*, 58-69.

Wiederanders, B., Kaulmann, G., & Schilling, K. (2003). Functions of propeptide parts in cysteine proteases. *Current Protein and Peptide Science, 4*(5), 309-326.

Xu, Q., Sudek, S., McMullan, D., Miller, M. D., Geierstanger, B., Jones, D. H., Krishna, S. S., Spraggon, G., Bursalay, B., Abdubek, P., Acosta, C., Ambing, E., Astakhova, T., Axelrod, H. L., Carlton, D., Caruthers, J., Chiu, H.-J., Clayton, T., Deller, M. C., Duan, L., Elias, Y., Elsliger, M.-A., Feuerhelm, J., Grzechnik, S. K., Hale, J., Won Han, G., Haugen, J., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Kozbial, P., Kumar, A., Marciano, D., Morse, A. T., Nigoghossian, E., Okach, L., Oommachen, S., Paulsen, J., Reyes, R., Rife, C. L., Trout, C. V., van den Bedem, H., Weekes, D., White, A., Wolf, G., Zubieta, C., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A., & Wilson, I. A. (2009). Structural basis of murein peptide specificity of a γ-D-glutamyl-L-diamino acid endopeptidase. *Structure, 17*(2), 303-313.

Zindel, S., Kaman, W. E., Fröls, S., Pfeifer, F., Peters, A., Hays, J. P., & Fuchsbauer, H.-L. (2013). The papain inhibitor (SPI) of *Streptomyces mobaraensis* inhibits bacterial cysteine proteases and is an antagonist of bacterial growth. *Antimicrobial Agents and Chemotherapy, 57*(7), 3388-3391.

국문초록

대부분의 cysteine protease는 clan CA, C1 family로 분류되는 papain-like cysteine protease에 속한다. Cysteine protease는 모든 유기체에 다양하게 존재하며 광범위한 질병과 세균, 기생충 감염의 원인으로 작용한다. 따라서 지금까지 cysteine protease에 대한 저해제의 개발이 계속해서 진행되어 왔다.

모든 papain-like cysteine protease는 공통적으로 signal peptide, propeptide, mature catalytic domain으로 이뤄진 zymogen 형태로 발현된다. Propeptide는 protease 저해제의 기능을 가지며 이를 이용한 저해제의 연구가 많이 진행되고 있다.

선행연구에서는 열대식물인 *Calotropis procera* R. Br.의 전사체 분석을 통해 20가지의 cysteine protease를 동정할 수 있었다. 20가지의 cysteine protease의 propeptide domain(SnuCapCpI) 중 선정된 8가지를 *Escherichia coli* 시스템에서 대량발현하여 얻은 SnuCalCpI03, SnuCalCpI15만이 cysteine protease 저해제로 쓰일 수 있었다. 따라서 발현되지 않은 후보를 대상으로 신규 천연 소재

저해제의 확보하고 cysteine protease 중 대표적인 papain에 대하여 저해 특성을 규명하는 연구를 진행하였다.

SnuCalCpI의 저해제로의 가능성은 서열분석을 이용하여 papain의 propeptide와의 높은 상동성 및 propeptide의 구조적으로 중요한 모티프의 존재를 통해 확인하였다. 또한 SnuCalCpI 유전자의 코돈 분석을 통해 선행연구에서 발현되지 않은 SnuCalCpI08, SnuCalCpI14, SnuCalCpI17 중 SnuCalCpI08과 SnuCalCpI17에서 *E. coli*에서 거의 사용되지 않는 코돈인 레어코돈 비율이 상대적으로 높음을 확인하였다.

Calotropis procera R. Br.로부터 total RNA 추출을 통해 cDNA를 합성하였으며, SnuCalCpI08, SnuCalCpI14, SnuCalCpI17 유전자를 PCR을 통해 증폭시켰다. 각각의 유전자는 pET29b(+) 벡터에 삽입하여 6가지의 레어코돈의 tRNA를 제공하는 *E. coli* Rosetta(DE3)에 발현하였다. SnuCalCpI08과 SnuCalCpI17이 발현되었으며, SnuCalCpI17만이 soluble하게 발현되었다.

선행연구로부터 확보한 SnuCalCpI03과 SnuCalCpI15를 포함하여 SnuCalCpI17의 cysteine protease에 대한 특성 분석을 진행하였다. Papain에 대한 저해 활성을 측정한 결과, IC_{50}은 각각

103.702±2.060 μg/mL, 23.900±0.654 μg/mL, 105.671±9.857 μg/mL이었다. 또한 SnuCalCpI는 공통적으로 papain에 대해 slow-binding 저해제로 작용하였으며, K_I^{app}는 각각 28.70 μg/mL, 13.63 μg/mL, 75.80 μg/mL이었다.

Slow-binding 저해제는 표적에서 해리되는 속도가 느려 생체 내에서 오랫동안 결합된 채로 유지되며, 저해제의 IC_{50}보다 훨씬 낮은 복용량에도 효과적으로 표적을 억제할 수 있다. 또한 펩타이드 형태의 저해제는 화학적으로 합성된 저해제와 달리 고분자이지만, 입체특이적으로 단백질-단백질 상호작용을 하기 때문에 활성부위에 특이적으로 잘 결합할 수 있다. 또한 천연자원에서 유래된 생리활성 펩타이드는 신체 조직에 축적되지 않으며, 부작용에 대한 사례가 거의 없다. 따라서 천연자원 유래 및 slow-binding 저해의 특성을 지닌 SnuCalCpI는 papain-like cysteine protease에 대한 효과적인 약물 제재로 사용할 수 있음을 시사한다.

주요어: *Calotropis procera* R. Br., cysteine protease 저해제, propeptide, 레어코돈, slow-binding 저해

학번: 2016-26913