

# Expression and characterization of a new serine protease inhibitory protein in *Escherichia coli*

Tran Thi Hong<sup>1,2,#</sup>, Ton That Huu Dat<sup>1,\*,#</sup>, Nguyen Phuong Hoa<sup>1</sup>, Tran Thi Kim Dung<sup>1</sup>, Vu Thi Thu Huyen<sup>3</sup>, Le Minh Bui<sup>4</sup>, Nguyen Thi Kim Cuc<sup>3</sup>, Pham Viet Cuong<sup>1,\*</sup>



Use your smartphone to scan this QR code and download this article

#equally contributed to this work

<sup>1</sup>Mien Trung Institute for Scientific Research, Vietnam Academy of Science and Technology, 321 Huynh Thuc Khang, Thua Thien Hue 531600, Vietnam

<sup>2</sup>Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Ha Noi 122300, Vietnam

<sup>3</sup>Institute of Marine Biochemistry, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Ha Noi 122300, Vietnam

<sup>4</sup>NNT Hi-Tech Institute, Nguyen Tat Thanh University, 300A Nguyen Tat Thanh, Ho Chi Minh city 748000, Vietnam

## Correspondence

**Ton That Huu Dat**, Mien Trung Institute for Scientific Research, Vietnam Academy of Science and Technology, 321 Huynh Thuc Khang, Thua Thien Hue 531600, Vietnam

Email: huudat96@gmail.com

## Correspondence

**Pham Viet Cuong**, Mien Trung Institute for Scientific Research, Vietnam Academy of Science and Technology, 321 Huynh Thuc Khang, Thua Thien Hue 531600, Vietnam

Email: phamvc1956@gmail.com

## History

- Received: Nov 12, 2019
- Accepted: Feb 04, 2020
- Published: Feb 29, 2020

DOI : 10.15419/bmrat.v7i2.590



## Copyright

© Biomedpress. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.



## ABSTRACT

**Introduction:** Proteases are enzymes that catalyze the hydrolysis of peptide bonds and play an important role in almost all biological processes. However, excessive protein proteolysis can be implicated in several diseases, such as cancer, as well as cardiovascular, inflammatory, neurodegenerative, bacterial, viral and parasitic diseases. In these cases, protease inhibitors can be used as one of versatile tools for regulating proteolytic activity of target proteases as well as therapeutic applications. In this study, we expressed and characterized a new serine protease inhibitory protein (PI-QT) from the metagenome of sponge-associated microorganisms in *Escherichia coli*. **Methods:** The gene PI-QT encoding for a new serine protease inhibitory protein was expressed in *E. coli* BL21(DE3). In addition, the expressed protein was purified and characterized. **Results:** Optimization of expression of the recombinant protein PI-QT in *E. coli* showed that suitable conditions for expression of the protein were pre-induction cell density (OD<sub>600</sub>) of 0.6 - 0.7, IPTG concentration of 1 mM and temperature of 25°C. The protease inhibitory protein was also purified and identified by mass spectrometry LC-MS/MS. The recombinant protein showed inhibitory activity against trypsin and  $\alpha$ -chymotrypsin with activity values of  $975 \pm 26$  U/mg and  $417 \pm 14$  U/mg, respectively. Maximum activity of the protease inhibitory protein was obtained at pH 7 and temperature 20-35°C. The inhibitor was stable over pH 4-9 and up to temperature 50°C. Addition of Zn<sup>2+</sup>, Mg<sup>2+</sup> and Ca<sup>2+</sup> enhanced inhibitory activity, whereas other metal ions, surfactants and oxidants reduced inhibitory activity of the protease inhibitor. **Conclusion:** The recombinant protein PI-QT is a potential protease inhibitor for therapeutic applications.

**Key words:** *Escherichia coli*, expression vector, protease inhibitor, recombinant protein, sponge-associated microorganisms

## INTRODUCTION

Proteases are enzymes that catalyze the hydrolysis of peptide bonds and play an important role in almost all biological processes; however, their uncontrolled activity often leads to diseases. Excessive protein proteolysis can relate to several diseases, including cancer, cardiovascular, inflammatory, neurodegenerative, bacterial, viral and parasitic diseases<sup>1</sup>. In these cases, protease inhibitors (PIs) can be used as one of the versatile tools for regulating proteolytic activity of target proteases<sup>2</sup>. To date, enzyme inhibitors have received increasing attention, not only for discovery of their structures and action mechanisms but also for potential application in different fields<sup>1,3</sup>.

Marine sponges are known to harbor diverse microbial communities<sup>4-6</sup> and represent a prolific source of natural products from their associated microorganisms<sup>7-10</sup>. Recent studies have shown that many potential protease inhibitors have been isolated from

sponge-associated microorganisms<sup>11-14</sup>; however, the exploration and exploitation potential PIs from microorganisms are still a big challenge because almost all microorganisms are resistant to cultivation in laboratory conditions, especially symbiotic microbes. Fortunately, new approaches (e.g., metagenomics) provide powerful tools for predicting, detecting and expressing novel bioactive genes from non-cultured microorganisms<sup>15-19</sup>. This opens new avenues for discovering new bioactive compounds, including PIs, in the future.

In order to meet the demands of finding novel and potential PIs from the marine environment, especially PIs from sponge-associated microorganisms, we optimized the expression and characterization of a new serine protease inhibitory protein PI-QT in an *E. coli* expression system.

## METHODS

**Cite this article :** Hong T T, Dat T T H, Hoa N P, Dung T T K, Huyen V T T, Bui L M, Cuc N T K, Cuong P V. **Expression and characterization of a new serine protease inhibitory protein in *Escherichia coli*.** *Biomed. Res. Ther.*; 7(2):3633-3644.

## Materials

The gene PI-QT is a new gene encoding a serine protease inhibitor from the metagenome of marine sponge QT collected from Quang Tri, Vietnam. It was synthesized and inserted into cloning vector pUC57 (GenScript, Piscataway, NJ, USA). The sequence of gene PI-QT was deposited in the National Center for Biotechnology Information (NCBI) database with the accession number MK359987.

## Sequence analysis, multiple sequence alignment, and phylogeny analysis

The PI-QT sequence was compared with other known sequences in the NCBI database using the Blast program. The open reading frame (ORF) of PI-QT was determined using ORF Finder program (<https://www.ncbi.nlm.nih.gov/orffinder/>). The molecular weight and pI of the deduced protein PI-QT were examined using the Compute pI/Mw tool of the ExPASy server ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). Multiple sequence alignments were performed using the ClustalW algorithm in the MEGA 7.0 software (<http://megasoftware.net>). A phylogenetic tree was constructed by the neighbor-joining (NJ) method and support of a bootstrap analysis with 1,000 replications implemented in the MEGA 7.0 software. Protein structure model of PI-QT was predicted using SWISS-MODEL program (<https://swissmodel.expasy.org/>) and (PS)<sup>2</sup>-v2: protein structure prediction server (<http://ps2.life.nctu.edu.tw/>).

## Construction of expression vector pET-32a(+)/PI-QT

The vector pUC57 (Thermo Fisher Scientific, Waltham, MA, USA) containing the gene PI-QT and expression vector pET-32a(+) (Invitrogen, Carlsbad, CA, USA) were digested with restriction enzymes *EcoRI* and *NotI* (Fermentas, Vilnius, Lithuania). The gene PI-QT was then purified and inserted into expression vector pET-32a(+) using enzyme T4 ligase (Thermo Fisher Scientific).

## Optimization of the expression of recombinant protein PI-QT

The recombinant vector pET-32a(+)/PI-QT was transformed into *E. coli* strain BL21(DE3) (Invitrogen) by the heat shock method. The transformed cells were plated on LB plates containing 50 µg/mL ampicillin (LB/amp) (Sigma-Aldrich, St. Louis, MO, USA) and incubated overnight at 37°C. Expression of the recombinant protein PI-QT was optimized using different conditions such as temperature (20,

25, 30, 37°C), IPTG (Sigma-Aldrich, St Louis, MO, USA) concentration (0.1, 0.5, 1.0, 1.5 mM), and pre-induction cell density (OD<sub>600</sub> = 0.4, 0.5, 0.6, 0.7).

## Determination of protein in soluble and insoluble fractions

Cell culture was centrifuged at 10,000 rpm for 15 min and then supernatant removed. The cells were resuspended in TE buffer (Tris 20mM, EDTA 10mM, PM F 0.05mM) and incubated at -75°C for 1 hr, then thawed at 50°C for 30 min, and sonicated on ice with a Misonix Ultrasonic Liquid Processors (company name & location). The solution was centrifuged at 13,000 rpm for 15 min, then the supernatant (soluble fraction) collected. The pellet was resuspended in TE buffer to the original volume (insoluble fraction). Next, SDS loading buffer was added to the pre-treatment culture (total protein), soluble protein fraction and insoluble protein fraction, and then the proteins were denatured at 100°C for 10 min. The expression of the protein was checked on 12.6% SDS-PAGE gel.

## Western blot analysis

After electrophoresis on SDS-PAGE gel, the recombinant protein was transferred to polyvinylidene difluoride (PVDF) membrane. The membrane was then blocked with TBS (containing 5% skimmed milk) and incubated with anti-TRx antibody (Sigma-Aldrich, St Louis, MO, USA) at a dilution of 1:1000 for overnight at 4 °C. The membrane was washed 4 times with TBS washing buffer and incubated with anti-mouse IgG secondary antibody at a dilution of 1:5000 for 1 hr at room temperature. The membrane was washed with TBS washing buffer and visualized by addition of *p*-nitro blue tetrazolium chloride 5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) solution.

## Purification of protein PI-QT

Purification of the recombinant protein PI-QT was performed by chromatography on Ni-NTA affinity chromatography column and then eluting with imidazole at concentrations of 100, 300, and 500 mM. The purified protein was removed from salts by dialysis tube (SnakeSkin™ Dialysis tubing) (Thermo Fisher Scientific) in TBS buffer (50 mM Tris HCl, 50 mM NaCl, pH 7.4) at 4 °C for 24 h. Next, the recombinant protein was removed from the TRx-tag fusion using Thrombin kit (Novagen, Darmstadt, Hesse, Germany).

## Protease inhibitory assay

### Trypsin inhibition assay

The purified protein (50  $\mu$ l of 0.05 mg/ml) was added to a mixture of 50  $\mu$ l of trypsin (Sigma-Aldrich) solution (0.05 mg trypsin/ml of 0.05 M Tris-HCl) and 100  $\mu$ l of 0.05 M Tris-HCl, pH 8.0, containing 0.03 M CaCl<sub>2</sub>. The mixture was incubated at 37°C for 10 min, then 1.0 ml of 0.8 mM BapNA (Sigma-Aldrich) solution was added and incubated at 37°C for 10 min. The reaction was stopped by adding 20  $\mu$ l of 30 % (v/v) glacial acetic acid. Subsequently, the solution was centrifuged at 10,000 rpm for 15 min and the absorbance were measured at 410 nm against appropriate blanks.

### $\alpha$ -Chymotrypsin inhibition assay

The purified protein (50  $\mu$ l of 0.05 mg/ml) was added to a mixture of 50  $\mu$ l of  $\alpha$ -chymotrypsin (Sigma-Aldrich) solution (0.05 mg  $\alpha$ -chymotrypsin/ml of 0.05 M Tris-HCl) and 100  $\mu$ l of 0.05 M Tris-HCl, pH 8.0, containing 0.03 M CaCl<sub>2</sub>. The mixture was incubated at 37°C for 10 min, then 1.0 ml of 0.86 mM BTPNA (Sigma-Aldrich) solution was added and incubated at 37°C for 10 min. The reaction was stopped by adding 20  $\mu$ l of 30 % (v/v) glacial acetic acid. Subsequently, the solution was centrifuged at 10,000 rpm for 15 min and the absorbance were measured at 410 nm against appropriate blanks.

Bowman-Birk Inhibitor (BBI) (Sigma-Aldrich) was used as a positive control for trypsin and  $\alpha$ -chymotrypsin inhibition assays.

## Identification of the recombinant protein

### Tryptic digestion in gel

The protein band from SDS-PAGE was excised, destained and dehydrated with 50% acetonitrile. The protein was reduced with DTT (65 mM) and subsequently alkylated with IAA (55 mM). The peptides were desalted using Zip-Tips (Millipore, Bedford, MA) according to the manufacturer's instructions. The purified peptides were collected for analysis by LC-MS/MS.

### Identification of the recombinant protein by shotgun proteomics

NanoLC-nanoESI-MS/MS analysis was performed on a nanoAcquity system (Waters, Milford, MA, USA) connected to an OrbitrapVelos hybrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a PicoViewnanospray interface (New Objective, Woburn, MA, USA). Peptide mixtures were loaded onto a 75  $\mu$ m ID, 25 cm length C18

BEH column (Waters, Milford, MA, USA), packed with 1.7  $\mu$ m particles with a pore size of 130 Å, and were then separated using a segmented gradient in 60 min from 5% to 40% solvent B (acetonitrile with 0.1% formic acid) at a flow rate of 300 nl/min and a column temperature of 35°C. Solvent A was 0.1% formic acid in water. The mass spectrometer was operated in the data-dependent mode. After acquisition of spectra, the proteins and peptides were identified by the PEAKS software (Bioinformatics Solutions Inc., Ontario, Canada).

## Characterization of protease inhibitor PI-QT

Effect of pH and temperature on activity of the protease inhibitor were determined by performing protease inhibitor assay at pH 7 after incubating the purified protease inhibitor in buffers of different pH (2-12) for 24 h and at different temperatures (30°C to 70°C) for 1 h.

Effect of metal ions, surfactants, and oxidizing agents on activity of the protease inhibitor were evaluated after incubating the protease inhibitor with 1 mM concentrations of metal salts (MgSO<sub>4</sub>, CuSO<sub>4</sub>, ZnSO<sub>4</sub>, CaCl<sub>2</sub>, MnCl<sub>2</sub>, FeCl<sub>2</sub>, NaCl), 1% (v/v) of surfactants (Tween 20, Tween 80 and Triton X 100), and oxidants (H<sub>2</sub>O<sub>2</sub> and DMSO) for 30 min at 37 °C.

## Statistical Analysis

The assays were performed in triplicate and expressed as the mean  $\pm$  standard error of the mean (SEM). The statistical analysis was performed by *t*-test and one-way analysis of variance (ANOVA) followed by Tukey's multiple comparison tests using the SPSS v.22 (SPSS Inc, Chicago, IL, USA). The results were considered to be significant at *P* < 0.05.

## RESULTS

### Amino acid sequence and phylogeny analysis of protein PI-QT

The gene PI-QT was 1,287 bp in length and had an open reading frame of 429 amino acid with a calculated molecular mass of about 50 kDa and a theoretical isoelectric point of 4.56. Comparison of the deduced amino acid sequence of PI-QT with the sequences in the NCBI database showed that protein PI-QT was most similar with serpins with similarities <55%. Multiple alignments of the deduced amino acids of PI-QT with the most homologous serpins in NCBI database showed that the deduced peptide PI-QT shared conserved active site residues with microbial serpin members (Figure 1). The phylogenetic tree

based on the neighbor-joining method (Figure 2A) located the protein PI-QT between two microbial serpin clades: one serpin clade from a candidate phylum of bacteria (Poribacteria) originally identified in the microbiome of marine sponges and another serpin clade from bacterial phyla Firmicutes and Cyanobacteria. Based on the database comparison, the protein PI-QT was considered as a new microbial serpin. The protein structure of PI-QT was predicted by SWISS-MODEL (Figure 2B) and (PS)<sup>2</sup>-v2 model (Figure 2C).

### Construction of the expression vector pET-32a(+)/PI-QT

The vectors pUC57 containing gene PI-QT and the expression vector pET-32a(+) were digested with *EcoRI* and *NotI* (Supplementary 1A). The gene PI-QT was then inserted into the opened expression vector pET-32a(+) using enzyme T4 ligase, transformed into *E. coli* strain TOP10F', and plated on LB plate supplemented with ampicillin (50 µg/mL). Agarose gel analysis of the plasmids isolated from colonies and cut with enzymes *EcoRI* and *NotI* (Supplementary 1B) showed bands of ~1.3 kb (corresponding to size of gene PI-QT) and bands of 5.9 kb (corresponding to size of vector pET-32a(+)). In addition, the sequencing results of the plasmid showed that the gene sequence inserted into the vector was identical to the sequence of the designed gene PI-QT (data not shown), thereby demonstrating that gene PI-QT was inserted successfully into expression vectors pET-32a(+) to form the recombinant vector pET-32a(+)/PI-QT.

### Expression of recombinant protein in *E. coli* strain BL21(DE3)

The recombinant vector pET-32a(+)/PI-QT was transformed into *E. coli* BL21(DE3) by heat shock and incubated on LB/amp medium overnight. Therefore, white colonies (Figure 3A) were incubated in LB/amp broth until OD<sub>600</sub> was about 0.8 - 1.0 IPTG was then added in culture and incubated at 37°C for 4 h. SDS-PAGE gel analysis of protein expression profile showed the presence of an overexpressed protein band of 64 kDa (Figure 3B, lane 3), corresponding to the size of recombinant protein PI-QT with Trx-tag fusion, whereas this foreign protein band was not present in negative control sample (Figure 3B, lane 1) containing only vector pET32a(+) and samples containing the recombinant vectors but not induced with IPTG (Figure 3B, lane 2). However, the recombinant protein was expressed mainly in the insoluble fraction (Figure 3C, lane 2). Therefore, optimization

of the expression conditions is necessary to increase the amount of the recombinant protein in the soluble fraction.

### Optimization of recombinant protein expression

#### Determination of suitable IPTG concentration for recombinant protein expression

Expression of the recombinant protein PI-QT was induced with different IPTG concentrations (0, 0.5, 1.0, and 1.5 mM). The SDS-PAGE gel analysis showed the recombinant protein was expressed at 1 mM IPTG, but not at low or high concentration of IPTG (0.5 mM and 1.5 mM). However, the expressed protein was still mainly present in the insoluble fraction (data not shown).

#### Determination of suitable temperature for recombinant protein expression

The expression of the recombinant protein was also investigated at different temperature (20, 25, 28, and 30 °C). The experimental results showed that the amount of recombinant protein in soluble fraction reached the highest value (423 mg/L with about 45% in the soluble fraction) when the recombinant protein was expressed at 25 °C (Figure 4), whereas the expression of recombinant protein was not observed at 20 °C (data not shown). At 28 °C and 30 °C, the amount of recombinant protein produced was higher; however, only a small amount of the recombinant protein was detected in the soluble fraction (only about 20% of total protein).

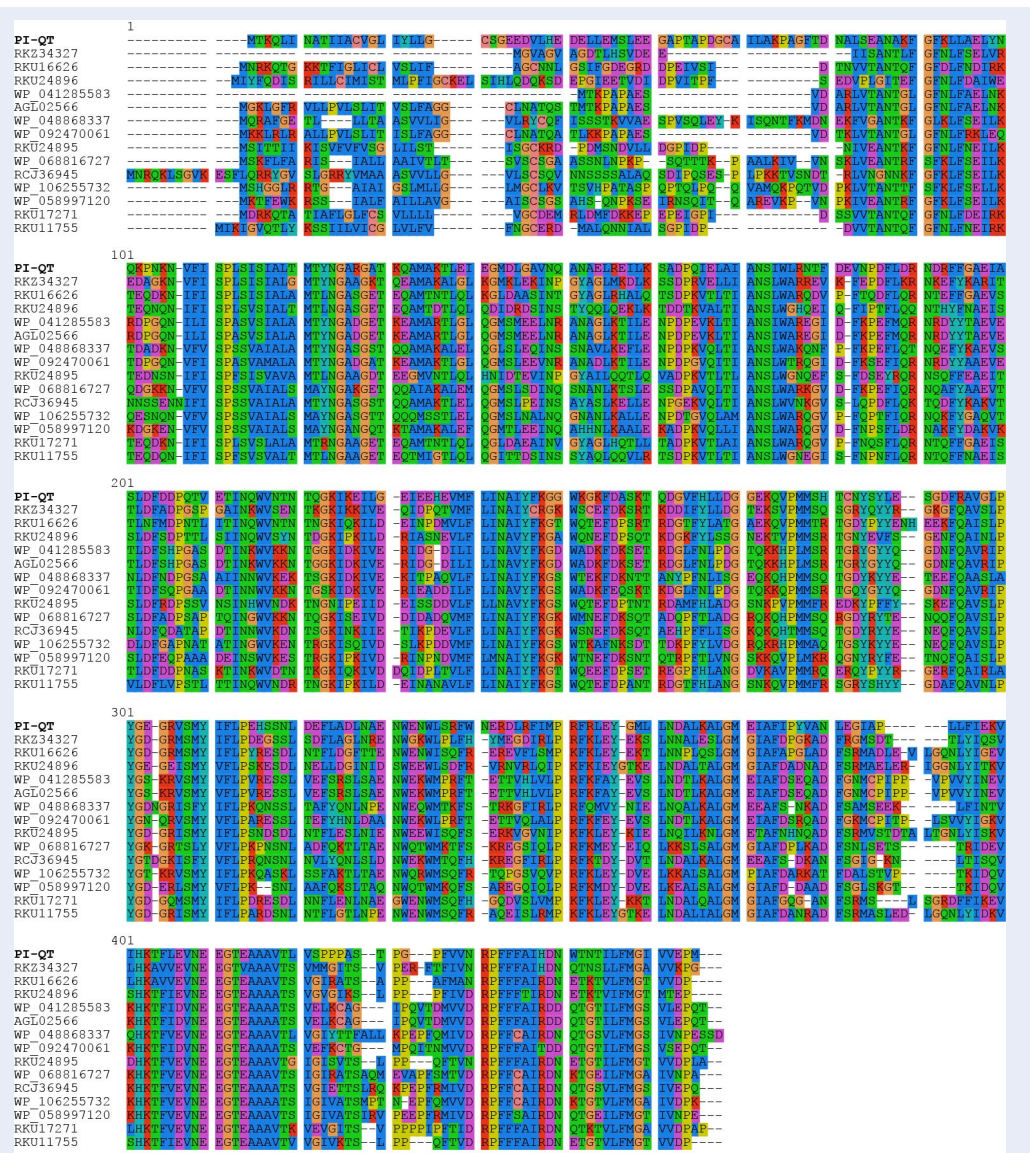
#### Determination of suitable cell density for recombinant protein expression

The expression of the recombinant protein at different pre-induction cell densities (OD<sub>600</sub> = 0.4, 0.5, 0.6, and 0.7) was investigated (Figure 5). The results showed that expression of the protein was not observed at OD<sub>600</sub> = 0.4. The total amount of produced protein and protein content in the soluble fraction were increased with increasing pre-induction cell density and reached the highest values at OD<sub>600</sub> = 0.6 - 0.7 (409 mg/L with 90% in the soluble fraction).

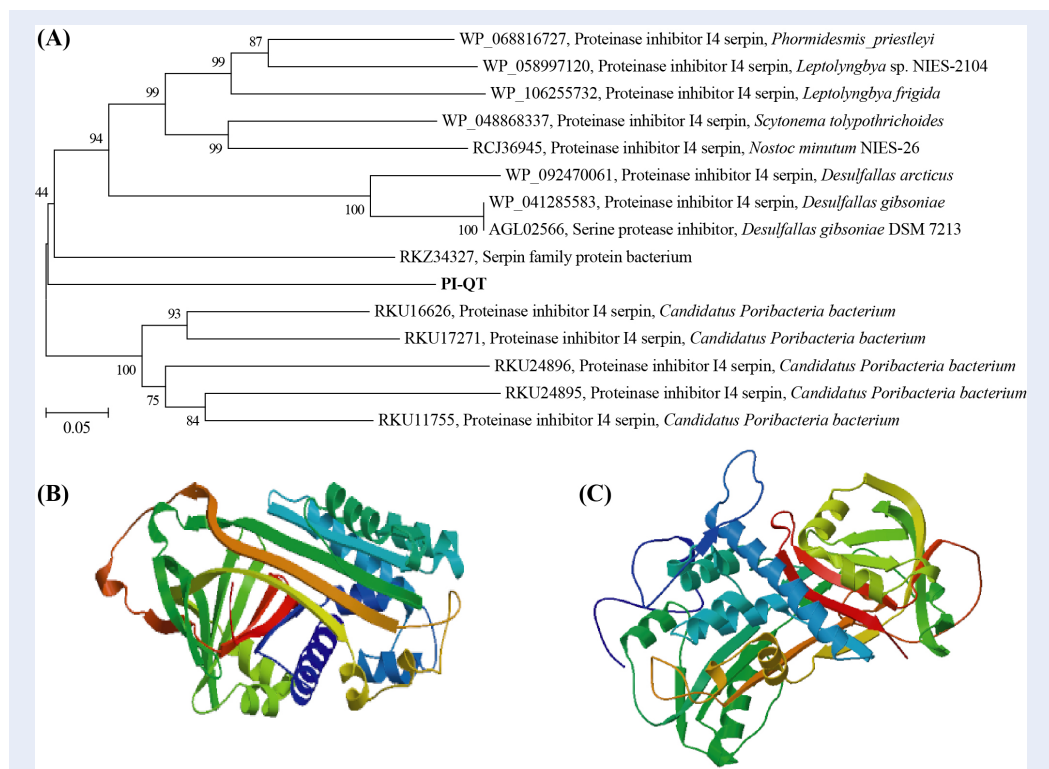
### Purification and identification of the recombinant protein

The purification of the recombinant protein by Ni-NTA affinity chromatography column showed that the recombinant protein was of high yield with imidazole concentrations of 100 and 300 mM; however,

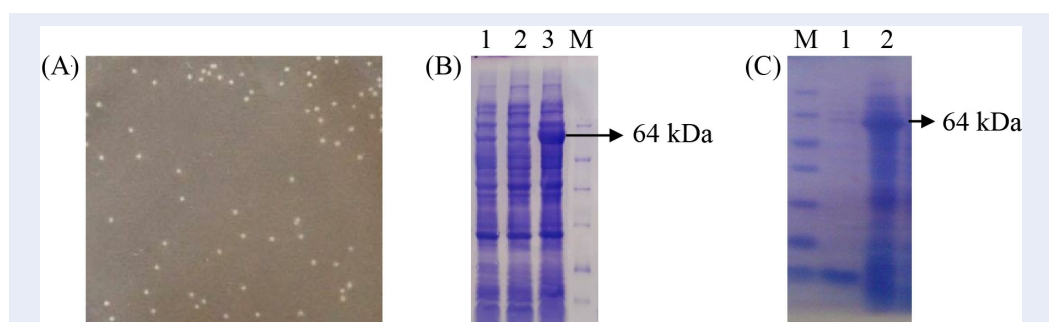




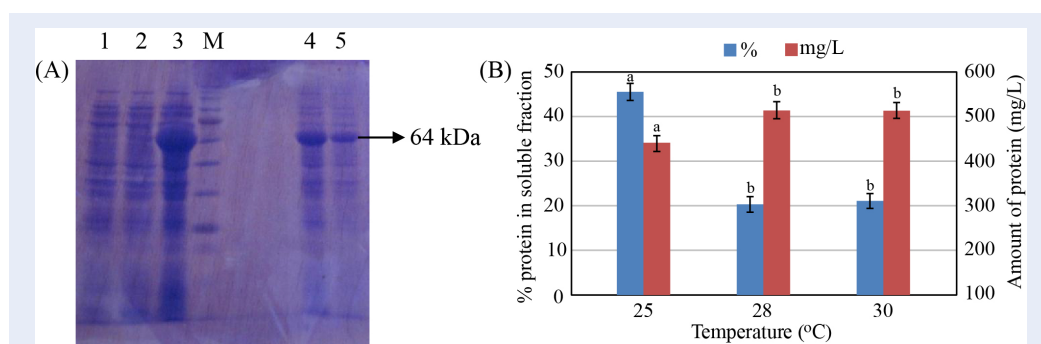
**Figure 1: Multiple alignments of the deduced amino acid sequence of PI-QT with other known serpins.** The multiple sequences were aligned by ClustalW algorithm. The known serpins from NCBI database are as follows: RKU17271, RKU16626, RKU24896, RKU24895, RKU11755 (proteinase inhibitor I4 serpin *Candidatus Poribacteria* bacterium), RKZ34327 (Serpin family protein bacterium), WP\_068816727 (proteinase inhibitor I4 serpin *Phormidesmis priestleyi*), WP\_058997120 (proteinase inhibitor I4 serpin *Leptolyngbya* sp. NIES-2104), WP\_106255732 (proteinase inhibitor I4 serpin, *Leptolyngbya frigida*), WP\_048868337 (proteinase inhibitor I4 serpin *Scytonema tolypothrichoides*), RCJ36945 (proteinase inhibitor I4 serpin *Nostoc minutum* NIES-26), WP\_092470061 (proteinase inhibitor I4 serpin *Desulfallas arctica*), WP\_041285583 (proteinase inhibitor I4 serpin *Desulfallas gibsoniae*), AGL02566 (serine protease inhibitor *Desulfallas gibsoniae* DSM 7213).



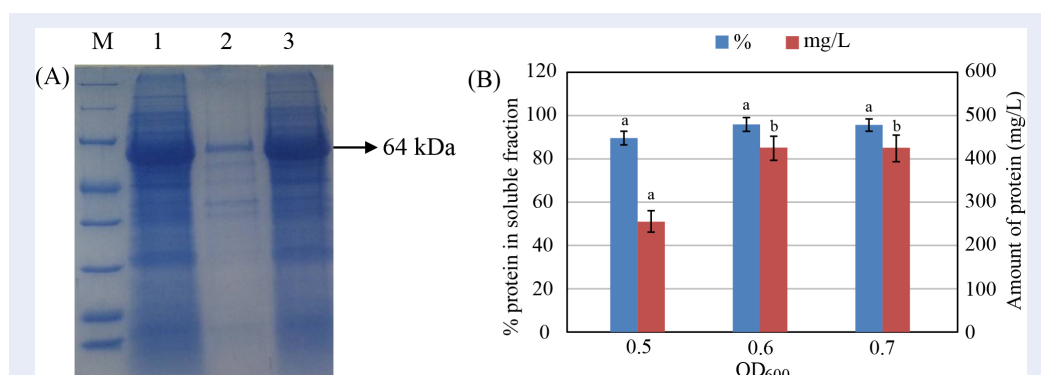
**Figure 2:** (A) Phylogenetic relationship of the Protein PI-QT with other serpins. The sequence alignment was performed using ClustalW algorithm. The phylogenetic tree was constructed by the neighbour-joining method using MEGA 7.0. Boot-strapping values were used to estimate the reliability of the phylogenetic reconstructions (1000 replicates). The numbers associated with the branches refer to bootstrap values (confidence limits) representing the substitution frequencies per amino acid residue. (B) Protein structure of PI-QT was predicted using SWISS-MODEL. (C) Protein structure of PI-QT was predicted using PS<sup>2</sup>-v2 server.



**Figure 3:** (A) White colonies containing the recombinant vector on agar plate. (B) SDS-PAGE gel of recombinant protein, lane M: marker protein (Novagen); lane 1: *E. coli* BL21(DE3) containing vector pET-32a(+); lane 2: *E. coli* BL21(DE3) containing the recombinant vector pET-32a(+)/PI-QT was not induced by IPTG; lane 3: *E. coli* BL21(DE3) containing the recombinant vector pET-32a(+)/PI-QT was induced by IPTG 1 mM at 37 °C. (C) SDS-PAGE gel of the recombinant protein, lane M: marker protein (iNtron); lane 1: total protein in soluble fraction; lane 2: total protein in insoluble fraction.



**Figure 4: Influence of temperature on expression of the recombinant protein.** (A) SDS-PAGE gel, lane M: marker protein (iNtRON), lane 1: total protein of *E. coli* BL21(DE3) containing vector pET-32a(+), lane 2: total protein of *E. coli* BL21(DE3) containing recombinant vector pET-32a(+)/PI-QT was not induced by IPTG; lane 3: total recombinant protein of *E. coli* BL21(DE3) containing recombinant vector pET-32a(+)/PI-QT was induced by IPTG 1mM at 25 °C, lane 4: total protein in insoluble fraction, lane 5: total protein in soluble fraction. (B) Amount of total protein and percentage of protein in soluble fraction at different temperatures. Different letters indicate significant differences between groups ( $p < 0.05$ ).



**Figure 5: Influence of pre-induction cell density on expression of the recombinant protein.** (A) SDS - PAGE gel of protein, lane M: marker protein (Thermo); lane 1: total protein of *E. coli* BL21(DE3) containing the recombinant vector pET-32a(+)/PI-QT; lane 2: total protein of *E. coli* BL21(DE3) containing the recombinant vector pET-32a(+)/PI-QT in insoluble fraction; lane 3: total protein of *E. coli* BL21(DE3) containing the recombinant vector pET-32a(+)/PI-QT in soluble fraction. (B) Amount of total protein and percentage of protein in soluble fraction at different pre-induction cell density. Different letters indicate significant differences between groups ( $p < 0.05$ ).

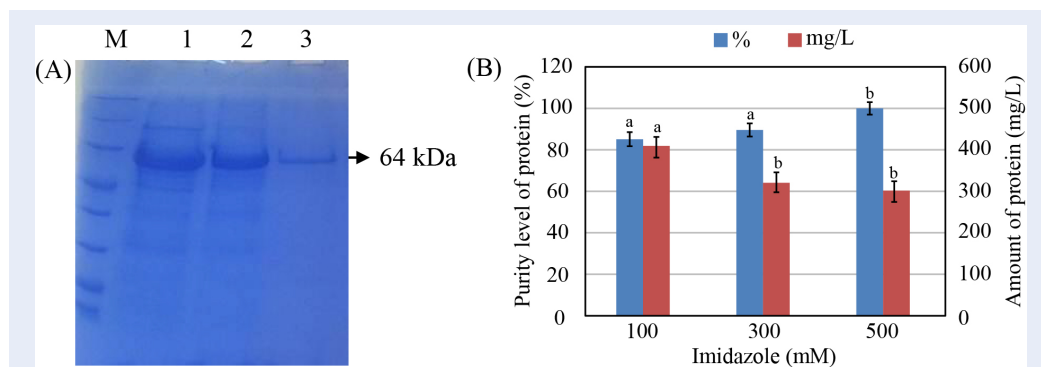
protein purity was low. In the case of imidazole concentration of 500 mM, the amount of the obtained protein was lower but had higher purity than those of the two above cases (Figure 6).

To confirm that the purified protein was the protein PI-QT, we performed Western blot assay using anti-Trx antibody. The Western blot analysis showed the presence of a protein band of 64 kDa on the hybrid membrane that was similar to the size of protein PI-QT with Trx-tag fusion (Figure 7A). In addition, in order to remove the Trx-tag fusion from the recombinant protein PI-QT, the protein was cut off from Trx-tag by thrombin. SDS-PAGE gel analysis of the protein treated with thrombin showed 2 bands of 50 kDa

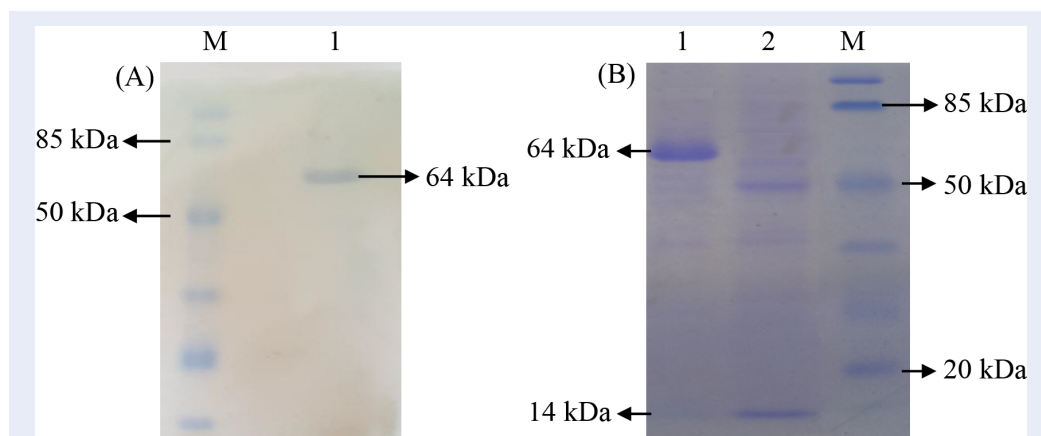
and 14 kDa, which corresponded to the size of protein PI-QT and Trx-tag, respectively (Figure 7B). This revealed that the Trx-tag fusion was removed successfully from the recombinant protein.

In order to identify the recombinant protein by mass spectrometry, the recombinant protein was cut, hydrolyzed, chromatographed, and analyzed for spectroscopic data (see methods). The LC-MS/MS spectroscopic data analysis of the recombinant protein (Supplementary 2 & 3) revealed that the sequence of the peptide fragments hydrolyzed and extracted from the recombinant protein was identical to the designed sequence. Since the protein was hydrolyzed into small peptides and only about 30% of peptides were recov-





**Figure 6: Protein fractions after chromatography on Ni-NTA affinity chromatography column.** (A) SDS-PAGE gel, lane M: marker protein (Thermo); lane 1, 2, 3: total protein eluted with imidazole of 100 mM, 300 mM, 500 mM, respectively. (B) Purity level and amount of protein after chromatography. Different letters indicate significant differences between groups ( $p < 0.05$ ).



**Figure 7: (A)** Western blot analysis, lane M: marker protein (Bio-Basic); lane 1: the recombinant protein containing Trx-tag tail was hybrid to antibody anti-Trx. **(B)** SDS-PAGE gel of recombinant protein was removed Trx-tag by thrombin, lane M: marker protein (Bio-Basic); lane 1: the recombinant protein was not removed Trx-tag; lane 2: the recombinant protein was removed Trx-tag.

ered (*Supplementary 3*), it is therefore not possible to find similar polypeptides on available databases. However, analyses based on the protein database built from the designed sequence by the software PEAKS and of the recovered peptide sequences confirmed that the expressed and purified recombinant protein was the desired protein.

### Activity and characterization of the recombinant protease inhibitor

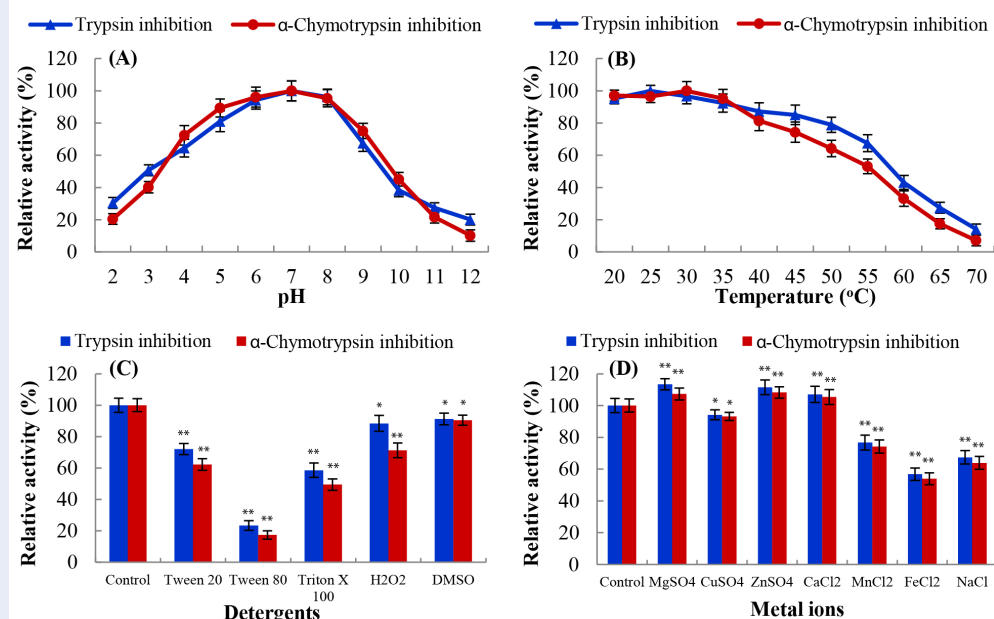
The recombinant protein was evaluated for its protease inhibitory activity against trypsin and  $\alpha$ -chymotrypsin. Protease inhibitory assay showed that the recombinant protein exhibited inhibitory effects against trypsin and  $\alpha$ -chymotrypsin with specific activities of  $975 \pm 26$  U/mg and  $417 \pm 14$  U/mg, respec-

tively. Compared to protein PI-QT, BBI (positive control) showed better inhibitory effects against trypsin and  $\alpha$ -chymotrypsin with specific activities of  $3303 \pm 66$  U/mg and  $1340 \pm 58$  U/mg, respectively.

The activity of the protease inhibitor peaked at pH 7 and still maintained more than 60% of its activity within pH 4-9. The activity of the protease inhibitor sharply declined at highly acidic (pH 3) and alkaline (pH 10) conditions (**Figure 8A**). The obtained results also showed that the protease inhibitor was most active at temperatures of 20-35 °C, and still maintained more than 60% of its activity up to 50 °C. The protease inhibitor activity decreased rapidly at temperature >60 °C (**Figure 8B**).

The experimental results showed that the presence of surfactants (Tween 20, Tween 80, and Triton X 100)





**Figure 8: Characterization of protease inhibitor PI-QT.** Effects of pH (A), temperature (B), surfactants and oxidants (C), and metal ions (D) on protease inhibitory activity of the protein PI-QT. Significant differences are indicated by asterisks. \*:  $p < 0.05$ ; \*\*:  $p < 0.01$ .

led to decreases of activity of the protease inhibitor compared to the control (Figure 8C). Furthermore, the activity reduction of the protease inhibitor was observed with the presence of oxidizing agents, H<sub>2</sub>O<sub>2</sub> and DMSO (Figure 8 C).

Effects of metal ions on the protease inhibitor were observed in this study (Figure 8 D). The obtained results showed that the presence of Zn<sup>2+</sup>, Mg<sup>2+</sup>, and Ca<sup>2+</sup> ions enhanced the activity of the protease inhibitor, whereas the presence of Cu<sup>2+</sup>, Mn<sup>2+</sup>, Fe<sup>2+</sup>, and Na<sup>+</sup> ions had negative effects and did not support protease inhibitor activity compared to the control.

## DISCUSSION

Protease inhibitors play an important role in the regulation of protease activity and have been used as a potential tool in different fields. Indeed, the discovery and exploitation of novel PIs from different sources have garnered increased attention across scientific research areas. In the present study, we have expressed and characterized a new serine protease inhibitor protein (PI-QT) from the metagenome of sponge-associated microorganisms in *E. coli*. The homogenous serpins in the NCBI database with the protein PI-QT are from microorganisms, suggesting that the protein PI-QT is a microbial serpin. Interestingly, several homogeneous serpins with the protein

PI-QT have been detected from sponge-associated bacteria (candidate phylum *Poribacteria*) based on metagenome data.

The protein PI-QT had an open reading frame of 429 amino acid with a calculated molecular mass of about 50 kDa. These values were consistent with the observed average molecular mass of most serpin proteins, ranging from 40 - 60 kDa with 330 - 500 amino acids<sup>20</sup>. The protein PI-QT exhibited inhibitory effects against trypsin and  $\alpha$ -chymotrypsin with specific activities of  $975 \pm 26$  and  $417 \pm 14$  U/mg, respectively. The specific activity of the protein PI-QT was comparable with that of novel serpins reported in recent studies<sup>21-24</sup>. For example, Jiang *et al.*<sup>21</sup> have cloned and expressed a novel serpin (Spi1C) from metagenomic library of uncultured marine microorganisms; the Spi1C protein exhibited inhibitory effects against trypsin and  $\alpha$ -chymotrypsin with specific values of 6940 and 3640 U/mg, respectively. Chan *et al.*<sup>22</sup> have purified a thermostable trypsin inhibitor from small pinto beans and reported that the purified protein showed inhibitory activity against trypsin with specific value of 2398 U/mg. Shamsi *et al.*<sup>23</sup> have isolated and purified a novel Kunitz trypsin inhibitor (ASPI) from garlic *Allium sativum*; the ASPI protein showed inhibitory activity against

trypsin with specific activity of 30376 U/mg. In another study, Mohan *et al.*<sup>24</sup> purified and characterized a protease inhibitor from *Capsicum frutescens* with specific activity against trypsin of 6749 U/mg.

The effects of IPTG concentration, pre-induction cell density, and temperature on the expression of the protease inhibitor PI-QT were evaluated to investigate the suitable and optimal conditions for expression of the protease inhibitor PI-QT in *E. coli*. Previous studies have reported that the synthesis of recombinant protein by expression vector pET-32a(+) is controlled by the T7 promoter and that this promoter is induced by the presence of IPTG in the culture medium. The concentration of IPTG in culture medium can, therefore, influence the expression of recombinant protein. The transcription may be hindered and the amount of foreign protein produced may be reduced due to low concentrations of IPTG in medium, whereas high concentrations of IPTG can cause cytotoxicity and inhibit cell growth<sup>25</sup>. In addition, it has also been reported that the optimal temperature for growth of *E. coli* is 37 °C; however, this temperature is not suitable for producing foreign protein. The high temperature may result in eliminating plasmids. In addition, expression of protein at high temperatures can produce proteins with a tertiary structure, resulting in the loss of biological activity of proteins, whereas low expression temperature may reduce the cleavage of target protein by intracellular protease and significantly increase the amount of recombinant protein<sup>26,27</sup>.

The experiments also showed that the protease inhibitor PI-QT were stable and maintained its activity within a wide range of pH and temperature. The stability of the protease inhibitor PI-QT within a wide range of pH and temperature is one of promising and interesting characteristic for its applications in biotechnological and pharmaceutical industries. Similar results were observed for many protease inhibitors of the Kunitz family in previous studies. The protease inhibitors in the Kunitz family are stable in wide pH ranges (pH 4-10), but sensitive to extreme pH conditions<sup>21,28-31</sup>. Under strongly acidic or alkaline conditions, the proteinaceous inhibitors are denatured and then lose their activity partially or completely. In addition, high temperature can affect the intramolecular disulfide bridges which are presumably responsible for the functional stability of Kunitz type protease inhibitors<sup>32</sup>.

The negative effects of surfactant and oxidizing agents on the activity of the protease inhibitor PI-QT were also observed in this study. The effects of surfactants on the activity of protease inhibitors could be attributed to a reduction of hydrophobic interactions,

whereas effects of oxidizing agents on the activity of protease inhibitors may be attributed to the probable oxidation of the amino acid methionine at the reactive site of the inhibitors<sup>33</sup>. Furthermore, metal ions can reduce the activity of the protease inhibitor PI-QT, with the exception of some metal ions, such as Mg<sup>2+</sup>, Ca<sup>2+</sup> and Zn<sup>2+</sup>.

Our study, together with previous studies, reveal that exploration of new bioactive compounds by expression of metagenome-based screening genes can be a versatile and potential strategy for mining new natural products from the marine environment as well as other environments<sup>15-17</sup>.

## CONCLUSIONS

A new gene encoding a serine protease inhibitor PI-QT screened from the metagenome of sponge-associated microorganisms was optimized and expressed successfully in *E. coli*. The recombinant protein had a molecular mass of about 50 kDa. The Western blot assay and spectroscopic data analyses confirmed expression of the recombinant protein PI-QT. The recombinant protein exhibited inhibitory activity against trypsin and  $\alpha$ -chymotrypsin with 975  $\pm$  26 U/mg and 417  $\pm$  14 U/mg, respectively. The protease inhibitor was stable within a wide range of pH and temperature. This study has shown that expression of recombinant protein from metagenome-based screening genes is a promising tool for the discovery and potential use of new bioactive compounds.

## ABBREVIATIONS

**Amp:** Ampicillin  
**BApNA :** N $\alpha$ -Benzoyl-L-arginine 4-nitroanilide  
**BTpNA :** N-Benzoyl-L-tyrosine p-nitroanilide  
**DMSO:** Dimethyl sulfoxide  
**DTT:** Dithiothreitol  
**EDTA:** Ethylenediaminetetraacetic acid  
**IAA:** Iodoacetamide  
**IPTG:** Isopropyl  $\beta$ -D-1-thiogalactopyranoside  
**LB:** Luria-Bertani broth  
**LC-MS/MS:** Liquid chromatography-tandem mass spectrometry  
**NCBI:** National Center for Biotechnology Information  
**OD:** Optical density  
**PI:** Protease inhibitor  
**PMSF:** Phenylmethylsulfonyl fluoride  
**PVDF :** polyvinylidene difluoride membrane  
**SDS:** Sodium dodecyl sulfate  
**SDS-PAGE :** Sodium dodecyl sulfate-Polyacrylamide gel electrophoresis  
**TBS:** Tris-buffered saline  
**TE:** Tris-EDTA

## COMPETING INTERESTS

The authors declare no conflict of interest.

## AUTHORS' CONTRIBUTIONS

Conceived and designed the experiments: Tran Thi Hong, Ton That Huu Dat, Nguyen Thi Kim Cuc, Pham Viet Cuong, Le Minh Bui. Performed the experiments: Tran Thi Hong, Ton That Huu Dat, Nguyen Phuong Hoa, Tran Thi Kim Dung, Vu Thi Thu Huyen. Analyzed the data: Tran Thi Hong, Ton That Huu Dat. Contributed reagents/materials/analysis tools: Le Minh Bui, Nguyen Thi Kim Cuc, Pham Viet Cuong. Wrote and revised the manuscript: all authors. All authors read and approved the final manuscript.

## ACKNOWLEDGEMENTS

LTQ-Orbitrap data and additional technical assistance were supported by the Academia Sinica Common Mass Spectrometry Facilities located at the Institute of Biological Chemistry, Taiwan.

## FUNDING

This work was financially supported by Vietnamese Ministry of Science and Technology under grant number: ĐTĐLCN.17/14.

## REFERENCES

- Sabotij, Kos J. Microbial and fungal protease inhibitor-sand potential applications. *Appl Microbiol Biotechnol*. 2012;93(4):1351–75. PMID: 22218770. Available from: 10.1007/s00253-011-3834-x.
- Shamsi TN, Parveen R, Fatima S. Characterization, biomedical and agricultural applications of protease inhibitors: A review. *Int J Biol Macromol*. 2016;91:1120–33. PMID: 26955746. Available from: 10.1016/j.ijbiomac.2016.02.069.
- Ryan BJ, Henahan GT. Overview of Approaches to Preventing and Avoiding Proteolysis During Expression and Purification of Proteins. *Curr Protoc Protein Sci*. 2013;71(1):5.25.1–5.25.7. Available from: 10.1002/0471140864.ps0525s71.
- Cuc NTK, Dat TTH, Hong TT, Cuong PV. Phylogenetic diversity of microorganisms associated with three marine sponges from Mien Trung sea of Vietnam. *Vietnam Journal of Science and Technology*. 2017;55(2):168. Available from: 10.15625/0866-708X/55/2/8577.
- Dat TTH, Cuong PV, Dung TTK, Huyen VTT, Cuc NTK. Bio-diversity assessment of microorganisms associated with two marine sponges (*Haliclona oculata* and *Amphius huxleyi*) collected at the Lang Co bay of Vietnam. *J Mar Sci Technol*. 2018;18(3):286–95. Available from: 10.15625/1859-3097/18/3/10642.
- Dat TTH, Steinert G, Cuc NTK, Smidt H, Sipkema D. Archaeal and bacterial diversity and community composition from 18 phylogenetically divergent sponge species in Vietnam. *PeerJ*. 2018;6:e4970. PMID: 29900079. Available from: 10.7717/peerj.4970.
- Brinkmann MC, Marker A, Kurtböke ID. An Overview on Marine Sponge-Symbiotic Bacteria as Unexhausted Sources for Natural Product Discovery. *Diversity (Basel)*. 2017;9(4):40. Available from: 10.3390/d9040040.
- Dat TTH, Hong TT, Dung TTK, Hoa NP, Cuong PV, Binh PT, Dat NT, Cuc NTK. Screening of antimicrobial producing bacteria associated with sponge and isolation of secondary metabolites from selected strain. *Proceedings of National Conference on Biotechnology 2018*; 2018 Oct 26; Hanoi, Vietnam. Hanoi, Vietnam: Publishing House for Science and Technology; 2018. p. 774-9.
- Bibi F, Faheem M, Azhar El, Yasir M, Alvi SA, Kamal MA. Bacteria From Marine Sponges: A Source of New Drugs. *Curr Drug Metab*. 2017;18(1):11–5. PMID: 27739371. Available from: 10.2174/1389200217666161013090610.
- Zhang H, Zhao Z, Wang H. Cytotoxic Natural Products from Marine Sponge-Derived Microorganisms. *Mar Drugs*. 2017;15(3):68. PMID: 28287431. Available from: 10.3390/md15030068.
- Hong TT, Cuc NTK, Dat TTH, Cuong PV. Screening of anti-protease producing bacteria associated with marine sponges from Da Nang sea of Vietnam. *Proceedings of National Conference on Biotechnology 2018*; 2018 Oct 26; Hanoi, Vietnam. Hanoi, Vietnam: Publishing House for Science and Technology; 2018. p. 743-8.
- Hong TT, Dat TTH, Cuc NTK, Cuong PV. Protease inhibitors from marine sponge and sponge-associated microorganisms. *Vietnam Journal of Science and Technology*. 2018;56(4):405–23. Available from: 10.15625/2525-2518/56/4/10911.
- Pimentel-Elardo SM, Buback V, Gulder TA, Bugni TS, Reppart J, Bringmann G. New tetracycline derivatives with anti-trypansomal and protease inhibitory activities. *Mar Drugs*. 2011;9(10):1682–97. PMID: 22072992. Available from: 10.3390/md9101682.
- Wahyudi AT, Qatrunnada, Mubarik NR. Qatrunnada, Mubarik NR. Screening and Characterization of Protease Inhibitors from Marine Bacteria Associated with Sponge *Jaspis* sp. *Hayati J Biosci*. 2010;17(4):173–8. Available from: 10.4308/hjb.17.4.173.
- Zhang T, Han WJ, Liu ZP. Gene cloning and characterization of a novel esterase from activated sludge metagenome. *Microb Cell Fact*. 2009;8(1):67. PMID: 20028524. Available from: 10.1186/1475-2859-8-67.
- Verma D, Kawarabayasi Y, Miyazaki K, Satyanarayana T. Cloning, expression and characteristics of a novel alkalistable and thermostable xylanase encoding gene (*Mxyl*) retrieved from compost-soil metagenome. *PLoS One*. 2013;8(1):e52459. PMID: 23382818. Available from: 10.1371/journal.pone.0052459.
- Zhang H, Li F, Chen H, Zhao J, Yan J, Jiang P. Cloning, expression and characterization of a novel esterase from a South China Sea sediment metagenome. *Chin J Oceanology Limnol*. 2015;33(4):819–27. Available from: 10.1007/s00343-015-4170-2.
- Singh R, Dhawan S, Singh K, Kaur J. Cloning, expression and characterization of a metagenome derived thermoactive/thermostable pectinase. *Mol Biol Rep*. 2012;39(8):8353–61. PMID: 22711301. Available from: 10.1007/s11033-012-1685-x.
- Maruthamuthu M, van Elsas JD. Molecular cloning, expression, and characterization of four novel thermo-alkaliphilic enzymes retrieved from a metagenomic library. *Biotechnol Biofuels*. 2017;10(1):142. PMID: 28588643. Available from: 10.1186/s13068-017-0808-y.
- van Gent D, Sharp P, Morgan K, Kalsheker N. Serpins: structure, function and molecular evolution. *Int J Biochem Cell Biol*. 2003;35(11):1536–47. PMID: 12824063. Available from: 10.1016/S1357-2725(03)00134-1.
- Jiang CJ, Hao ZY, Zeng R, Shen PH, Li JF, Wu B. Characterization of a novel serine protease inhibitor gene from a marine metagenome. *Mar Drugs*. 2011;9(9):1487–501. PMID: 22131953. Available from: 10.3390/md9091487.
- Chan YS, Zhang Y, Sze SC, Ng TB. A thermostable trypsin inhibitor with antiproliferative activity from small pinto beans. *J Enzyme Inhib Med Chem*. 2014;29(4):485–90. PMID: 23859150. Available from: 10.3109/14756366.2013.805756.

23. Shamsi TN, Parveen R, Amir M, Baig MA, Qureshi MI, Ali S. *Allium sativum* Protease Inhibitor: A Novel Kunitz Trypsin Inhibitor from Garlic Is a New Comrade of the Serpin Family. *PLoS One*. 2016;11(11):e0165572. PMID: 27846232. Available from: [10.1371/journal.pone.0165572](https://doi.org/10.1371/journal.pone.0165572).
24. Mohan M, Kozhithodi S, Nayarisseri A, Elyas KK. Screening, Purification and Characterization of Protease Inhibitor from *Capricum frutescens*. *Bioinformation*. 2018;14(6):285–93. PMID: 30237674. Available from: [10.6026/97320630014285](https://doi.org/10.6026/97320630014285).
25. Shin CS, Hong MS, Bae CS, Lee J. Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21(DE3)[pET-3aT2M2]. *Biotechnol Prog*. 1997;13(3):249–57. PMID: 9190075. Available from: [10.1021/bp970018m](https://doi.org/10.1021/bp970018m).
26. Piserchio A, Ghose R, Cowburn D. Optimized bacterial expression and purification of the c-Src catalytic domain for solution NMR studies. *J Biomol NMR*. 2009;44(2):87–93. PMID: 19399371. Available from: [10.1007/s10858-009-9318-y](https://doi.org/10.1007/s10858-009-9318-y).
27. Tomala M, Lavrentieva A, Moretti P, Rinas U, Kasper C, Stahl F. Preparation of bioactive soluble human leukemia inhibitory factor from recombinant *Escherichia coli* using thioredoxin as fusion partner. *Protein Expr Purif*. 2010;73(1):51–7. PMID: 20381622. Available from: [10.1016/j.pep.2010.04.002](https://doi.org/10.1016/j.pep.2010.04.002).
28. Bacha AB, Jemel I, Moubayed NMS, Abdelmalek IB. Purification and characterization of a newly serine protease inhibitor from *Rhamnus frangula* with potential for use as therapeutic drug. *3 Biotech*. 2017;7(2):148. PMID: 28597160. Available from: [10.1007/s13205-017-0764-z](https://doi.org/10.1007/s13205-017-0764-z).
29. Bijina B, Chellappan S, Basheer SM, Elyas KK, Bahkali AH, Chandrasekaran M. Protease inhibitor from *Moringa oleifera* leaves: Isolation, purification, and characterization. *Process Biochem*. 2011;46(12):2291–300. Available from: [10.1016/j.procbio.2011.09.008](https://doi.org/10.1016/j.procbio.2011.09.008).
30. Cruz AC, Massena FS, Migliolo L, Macedo LL, Monteiro NK, Oliveira AS. Bioinsecticidal activity of a novel Kunitz trypsin inhibitor from *Catanduva* (*Piptadenia moniliformis*) seeds. *Plant Physiol Biochem*. 2013;70:61–8. PMID: 23770595. Available from: [10.1016/j.plaphy.2013.04.023](https://doi.org/10.1016/j.plaphy.2013.04.023).
31. Mello GC, Oliva ML, Sumikawa JT, Machado OL, Marangoni S, Novello JC. Purification and characterization of a new trypsin inhibitor from *Dimorphandra mollis* seeds. *J Protein Chem*. 2001;20(8):625–32. PMID: 11890203. Available from: [10.1023/A:1013764118579](https://doi.org/10.1023/A:1013764118579).
32. KidriM, Fabian H, Brzin J, PopoviT, Pain RH. Folding, stability, and secondary structure of a new dimeric cysteine proteinase inhibitor. *Biochem Biophys Res Commun*. 2002;297(4):962–7. PMID: 12359248. Available from: [10.1016/S0006-291X\(02\)02328-8](https://doi.org/10.1016/S0006-291X(02)02328-8).
33. Johnson D, Travis J. The oxidative inactivation of human alpha-1-proteinase inhibitor. Further evidence for methionine at the reactive center. *J Biol Chem*. 1979;254(10):4022–6. PMID: 312289.