# EXPRESSION OF RECOMBINANT *MANDUCA SEXTA* PROPHENOLOXIDASE ACTIVATING PROTEINASE-1 IN *BACILLUS SUBTILIS*

by

WENJING WANG

B.S., Shan Dong Agricultural University, 2006
M.S., Institute of Botany, Chinese Academy of Sciences, 2010

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Graduate Biochemistry Group

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Approved by:

Major Professor
Michael R. Kanost

# Abstract

Prophenoloxidase-activating proteinase (proPAP) activates prophenoloxidase when bacteria or fungi invade *Manduca sexta*. Upon activation, phenoloxidase initiates synthesis of melaninin, which can encapsulate the invaders and kill them. *M. sexta* contains three proteases that can activate prophenoloxidase, proPAP1, proPAP2, and proPAP3. The study of proPAP function has been slowed by the difficulty of expressing the proteins in recombinant systems. ProPAP1 contains one clip domain and one serine proteinase domain, a simpler structure than proPAP2 and proPAP3, which have two clip domains. For this reason, proPAP1 was selected for this investigation, to develop an improved system for expression of recombinant proPAP zymogens. In past experiments proPAP1 had a low expression level in insect cells using a baculovirus vector. In *Escherichia coli*, proPAP1 was expressed as an insoluble protein that could not be refolded successfully. The *Bacillus subtilis* expression system offers a potential improvement for expression of recombinant clip domain proteases because it can secrete recombinant proteins into the medium, it is a Biosafety Level 1 organism that is easy to handle, and it is less expensive to culture than insect cells. Four constructs for expression of proPAP1 and proPAP1 mutants were produced in the plasmid shuttle vector pHT43, which is compatible with both *E. coli* and *B. subtilis*. Experiments were carried out to test and optimize expression and purification of proPAP1 in *B. subtilis*. Conditions were optimized for IPTG (isopropyl β-D-1-thiogalactopyranoside) concentration, IPTG induction time, growth medium and induction temperature. Results showed that 0.5mM IPTG with 20 hours induction at 37°C in 2xYT medium was the optimum condition for proPAP1 production in the *B. subtilis* system. The recombinant proPAP1 was precipitated from the medium in 50% saturated ammonium sulfate and partially purified by nickel affinity chromatography. In addition to the full length proPAP1 protein, degradation of proPAP1 was also observed. Further experiments should be done to try to solve this problem. With purified protein, future work can be aimed at study of the structure and function of proPAP1.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| dNTP | deoxyribonucleotide triphosphate |
| EDTA | ethylenediaminetetraacetic acid |
| EGTA | ethylene glycol tetraacetic acid |
| h | hour |
| IPTG | isopropyl β-D-1-thiogalactopyranoside |
| LB | lysogeny broth |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| rpm | revolutions per minute |
| RT | room temperature |
| SDS | sodium dodecyl sulfate |
| U | unit |

# Acknowledgements

I would like to thank my major professor Dr. Michael R. Kanost for all the encouragement, support and kind advice he offered me during my study in his laboratory as well as the good laboratory environment he provided for all of us.

My thanks also go to my other two committee members, Dr. Subbaratnam Muthukrishnan and Dr. Mark Weiss, for their guidance and valuable suggestions on my research project and thesis.

I also want to thank Dr. Daisuke Takahashi for all his help and suggestions on my experimental design and analysis of results. I want to thank Dr. Di Wu, for providing me the experiment materials including the vector and mutants, as well as all her help when I just joined the laboratory.

I would like to thank Zhen Li, Fabrice Ngansop, Sue Huang and Yan Shipelskiy, for their support and kind help to fulfill my Teaching Assistant job. I also want to thank Dr. Phillip E. Klebba for offering me this great opportunity to work in his department as a graduate Teaching Assistant. I want to thank Xiangming Li, Ying Ding, Lorne Jordan and Qasim Al Souhail for all the help and encouragement they provided through my hardest time here.

I want to thank my parents for all their support and encouragement.

# Dedication

This thesis is dedicated to my dear parents who always encourage me and help me to pursue my dreams.

# Chapter 1 - Introduction

## The importance to study *Manduca sexta*'s innate immune system

*Manduca sexta*, also known as the tobacco hornworm or Carolina sphinx moth, can be found in most of the states in the US (Lange, et al., 1981). Often considered as agricultural pests since they feed on tobacco, tomato and other plants (Lange, et al., 1981), *M. sexta* has a long history as a model system for scientific research due to its short life span, easy to raise and propagate in laboratory settings and its large, easily accessible and dissectible organs. Two of the areas that benefit the most from research performed with *M. sexta* are neurobiology and immunology. *M. sexta* has been a model system for neurodevelopment, memory formation (Weeks et al, 1997), serotonergic and dopaminergic neurotransmission (Ellen and Mercer, 2012) and olfaction (Kloppenburg and Mercer, 2008). Together with other insect model systems including the malaria mosquito, *Anopheles gambiae*, the red flour beetle, *Tribolium castaneum* and fruit fly, *Drosophila melanogaster, M. sexta* has helped in the understanding of many immune processes that are applicable to human. For example, the Toll pathway was discovered both in *Drosophila* embryonic development (Halfon et al., 1998) and in immunity (Lemaitre et al., 1996). These discoveries led to the elucidation of Toll-like receptor in mammals (Medzhitov et al, 1997) and also helped Dr. Hoffmann to win the 2011 Nobel Prize for physiology or medicine (Hoffmann, 2011).

*Manduca sexta* usually has two generations per year in nature (Lange et al., 1981). Its life stages can be divided into egg, larva, pupa and adult. During its life cycle, *M. sexta* encounters invaders, including bacteria, fungi and other pathogens. Unlike mammals that have both the innate and adaptive immune response, insects such as *M. sexta* lack an adaptive immune response. Although they are not able to produce antibodies after exposure to antigens, insects have a powerful innate immune system (Dunn, 1990). This non-specific immune response provides immediate defense for the insects against invading pathogens and parasites, therefore allowing them to adapt to the severe environment (Dunn, 1990; Gillespie et al., 1997; Lavine, 2002; Jiang et al., 2010). For these reasons, *M. sexta* has been an excellent model system to study innate immune response upon microbe infection. Especially, due to its larger larvae size in

comparison to other insects, *M. sexta* is more amendable to *in vitro* biochemical studies and has become a system of choice for dissecting molecular pathways involved in the immune response (Kanost et al., 2004). In addition, knowledge obtained in the innate immune response in *M. sexta* may be applicable to other organisms including mammals.

The *M. sexta* genome consists of 28 chromosomes and contains about 500 MB of information (Hanrahan and Johnson, 2011; Yasukochi et al., 2009). The whole genomic sequence is in progress and will be released soon (M. Kanost, personal communication). The genomic information will be a tremendous boost to the current and future studies in many disciplines, including innate immunity.

The innate immune response in *M. sexta* is a relatively simple but efficient self-defense system that involves several interconnected steps (Kanost et al., 2004; An et al., 2010; Jiang 2008). The presence of invading pathogens is first recognized by a group of proteins in the hemolymph. These so-called surveillance proteins, including hemolin, peptidoglycan recognition proteins, β-1,3-glucan recognition proteins (βGRPs), and C-type lectins, microbe binding protein (MBP) recognize the invading pathogens by binding to their surface molecules (Kanost et al., 2004; Dai et al., 2013; Wang et al., 2011; Jiang et al., 2004; Yu et al., 2002). Binding of the surveillance proteins to pathogens triggers a number of responses that consist of the innate immune response to fight off the pathogens (Kanost et al., 2004). Components in the hemolymph such as anti-microbial peptides and macrophage-like hemocytes are effective in defense against bacteria, parasites and fungi. These proteins or peptides are secreted into hemolymph during the innate immune response. Their expression level increased dramatically in hemolymph during the innate immune response (Kanost et al., 2004). Hemocytes phagocytose the pathogens or trap them in hemocyte aggregates called nodules (Horohov et al., 1983). Nodules and other sites of infection can be melanized through initiation of a serine proteinase cascade that activates prophenoloxidase (Kanost et al., 2004). Recent studies have used RNA-seq to identify the genes that participating in *M. sexta* innate immune response (Gunaratna and Jiang, 2013; Zhang et al., 2011).

Although the whole immune response pathway is not clear yet, more than 20 serine proteinases in hemolymph have been identified in *M. sexta* (Jiang et al., 2005; An et al., 2009). However, the functions for most of the proteinases are not well defined (An et al., 2009). In hemolymph, pathogen recognition triggers the activation of the cascades of serine proteinases (Kanost et al., 2004). At the end of the cascade, a proteinase cleaves prophenoloxidase (proPO) to generate phenoloxidase (PO) and activate it, and then the active PO oxidizes the catechols, leading to melanin formation (Cerenius et al., 2004; Kanost and Gorman, 2008; Jiang et al., 1997; Jiang and Kanost, 2000).

## Function of PAP family and its importance in the innate immune response

One of the most important molecular pathways that are involved in the innate immune response in *M. sexta* is the prophenoloxidase (proPO) activation pathway (An and Kanost, 2010). Although there are still many unknowns, the proPO activation pathway is relatively well understood compared with other pathways (Jiang et al., 2005). Activation of proPO pathway can lead to melanin synthesis (Gorman et al., 2007; Liu et al., 2007). Upon activation, PO initiates the melanin synthesis process. In addition to immune defense, melanin is also believed to be involved in wound healing (Sugumaran, 2002). Phenolic compounds are the substrates of PO. When oxidized they are converted to quinones. Quinones can assist in killing the invading pathogens and can also be used to synthesize melanin for wound healing and encapsulation of parasites (Kanost and Gorman, 2008; Zhao et al., 2007). 5,6-dihydroxyindole (DHI), a substrate of proPO, is very toxic to both bacteria and the host cells (Zhao et al., 2007).

POs belong to the tyrosinase type of enzyme family (E.C. 1.14.18.1). Tyrosinases hydroxylate monophenols such as tyrosine and oxidize o-diphenols such as DOPA (Korner and Pawelek, 1982). PO is synthesized in the form of zymogen, proPO (Kan et al., 2008). ProPO is expressed constitutively in certain hemocytes (Gorman et al., 2007; Charalambidis et al., 1996). There are several types of hemocytes in *M. sexta*, such as oenocytoid, plasmatocytes, granulocyte and spherule cells (Willott et al., 1994). Oenocytoid is the only type of hemocytes that can synthesize PO (Jiang et al., 1997) The zymogen form of proPO is activated by a family of

proteinases named proPO activating proteinases (PAPs) (Gupta et al., 2005a). PAPs cleave proPO in the hemolymph and activate the enzyme.

Three proPO activating proteinases (PAPs) have been discovered so far in *M. sexta*. They are named PAP1, PAP2, and PAP3 (Jiang et al., 1998, 2003a and 2003b). All of the three PAPs themselves are synthesized as zymogens (proPAPs), and they only become active by specific proteolytic cleavage by an upstream proteinase. PAP1, PAP2 and PAP3all cleave proPO at $Arg^{51}$ (Gupta et al., 2005b), and the reaction is mediated by serine proteinase homologs (SPH) (Wang and Jiang, 2004b). The structure of proPAP1 is the simplest among the three proPAPs. It contains a clip domain in the amino-terminus of the protein, and a serine proteinase domain in its carboxyl-terminus. Besides proPAP1, SPH-1 and SPH-2 together with immulectin-2 also helps to activate proPO (Yu et al., 2002). ProPAP2 and proPAP3 each contain two amino-terminal clip domains and one serine proteinase domain in the carboxyl-terminus. The proteinases that cleave and activate proPAP zymogens are called hemolymph proteinases (HPs). ProPAP1 can be recognized and activated by HP6 through the cleavage after Arg-127 (Jiang et al., 1998; An et al., 2009). ProPAP2 and proPAP3, however, are activated by HP21 through the cleavage after Lys-153 and Lys-146 respectively (Jiang et al., 2003a and 2003b; Gorman et al., 2007; Wang and Jiang, 2007). In addition to the activation cleavages by the HPs, all three PAPs need the presence of two noncatalytic serine proteinase homologs (SPHs), SPH1 and SPH2, to activate proPO efficiently (Jiang et al., 2003a; Wang and Jiang, 2004b; Wang and Jiang, 2008).

ProPAP1 was the first PAP identified in *M. sexta* through a function - guided protein purification protocol (Jiang et al, 1998). PAP1 has the apparent molecular weight, as measured in the SDS PAGE under non-reducing conditions, of 44 kDa and is composed of two disulfide-linked polypeptide chains – the 31 kDa catalytic domain and 13 kDa clip domain. The cDNA was isolated from a hemocyte cDNA library (Jiang et al., 1998). The full length cDNA contains an open reading frame of 383 amino acid residues, with a calculated mass 39.5 kDa. The native mRNA for proPAP1 is about 1.5 kb in length, and the expression level of proPAP1 mRNA is increased dramatically in fat body upon injection of bacteria, suggesting that proPAP1 is involved in the innate immune response of *M. sexta* (Jiang et al., 1998; Zou et al., 2005).

Serine proteinases such as PAPs are regulated by a family of inhibitors called serpins. In *M. sexta*, the primary inhibitor of PAPs appears to be serpin-3 (Zhu et al, 2003; Christen et al., 2012). Serpin-3 shares sequence homology with the *Drosophila* serpin-27A, also a regulator of melanization (De Gregorio et al., 2002). In vitro studies with purified proteins showed that serpin-3 associates with both PAP1 and PAP3 to inhibit their proteinase activity (Zhu et al, 2003). These results suggest that serpin-3 regulates the melanization process by blocking the activation of proPO. *M. sexta* maintains a low but constitutive expression of serpin-3 in its hemolymph. Upon microbial infection, the expression level of serpin-3 increases by about 6 fold (Zhu et al, 2003), indicating that serpin-3 expression increased as part of the insect's immune response. Recently study shows that, besides inhibits PAPs, serpin-3 can also inhibit HP8 (Christen et al., 2012). Serpin-6 also inhibits PAP3 (Wang and Jiang 2004a; Zou and Jiang 2005)

In summary, there appears to be both positive and negative regulatory mechanisms for the innate immune response in the hemolymph of *M. sexta*, and that the PAP family of serine proteinases is a center player in the process of proPO activation.

## Expression of recombinant serine proteinases

Proteinases are one of the most abundant families of enzymes and are involved in many physiological processes. Serine proteinases belong to the endopeptidase family which catalyses the hydrolysis of peptide bonds. Endopeptidases can be further divided into four subtypes based on their catalytic mechanisms. They are called serine, cysteine, aspartic and metallo proteinases. In general, serine proteinases are first synthesized as zymogens which can be activated through self-cleavage or by other proteinases (Khan and James, 1998).

In vitro recombinant systems are the most straight forward system for understanding the structure and function of serine proteinases. Various host organisms including bacteria, fungi, insects, plants and animals have been used for this purpose. *Escherichia coli* is often employed for its simplicity, low cost, fast high-density cultivation, well-known genetics and the large number of available tools (Sørensen and Mortensen, 2005).

One of the biggest challenges for heterologous recombinant protein expression is their folding issues. There are at least three reasons why misfolding frequently happens. First, the host for recombinant expression may lack the specific interaction partners (e.g., chaperones) and post-translational modification mechanisms that ensure correct folding in the native environment of the protein (Sørensen and Mortensen, 2005). Second, the highly crowded environment in the cytoplasm (e.g. protein concentration at 200–300 mg/ml in *E. coli*) is also unfavorable for the correct folding of recombinant proteins (van den Berg et al., 2000). Third, the potential toxicity to the host by overexpressing the serine proteinases could also be a factor that affects the expression level and downstream processing (Georgiou and Segatori, 2005). Therefore, it is advisable that the host organisms and the location of the recombinant protein products should be carefully selected depending on biochemical properties of the proteins to be expressed. For example, the toxicity to the host might be alleviated if the toxic protein is secreted into the medium.

Baculovirus expression system had been used to express *Haemonchus contortus* recombinant cysteine protease (CP) (Estefan et al., 2008). Insect Sf9 cells were used to express CP, and a 6x His tag was used to purify active CP through affinity purification (Estefan et al., 2008). *E. coli* TB1 was a good system for the recombinant murine coronavirus (MHVJHM) 3C-like proteinase expression (Seybert, et al., 1997). The factor Xa recognition site was used to link the 5' *E. coli* maltose-binding protein and the 3' MHV-JHM 3C-like proteinase domain (Seybert, et al., 1997). After the protein expression and purification, factor Xa was used to cleave and release the MHV-JHM 3C-like proteinase domain for further functional study (Seybert, et al., 1997).

*E. coli* strain pLysS had been used as a host strain to express human mannan-binding lectin-associated serine protease (SP) (MASP)-1 and MASP-2 (Ambrus et al., 2003). The recombinant proteins were expressed under the isopropyl-D-thiogalactoside induction. The recombinant proteins could form inclusion body, which were then solubilized and refolded successfully (Ambrus et al., 2003). Another example was to use baculovirus to express fragments from a human modular serine protease C1r, a protease that mediates activation of the C1 complex, which is comprised of two complement control protein (CCP) modules followed by a serine protease (SP) domain. Baculovirus-mediated expression was employed to produce fragments

containing the SP domain and both CCP modules (CCP1/2-SP) and one containing only the second CCP module (CCP2-SP). (Lacroix et al., 2001).

## The *Bacillus subtilis* expression system

Although most commonly used, heterologous expression of recombinant proteins in *E. coli* often results in the formation of insoluble protein aggregates packaged in inclusion bodies. These protein aggregates are inactive, although can be easily isolated to relatively high purity (Niwa et al., 2009). To acquire the active form of the protein, these aggregates are often completely denatured then re-natured in hope the protein will fold into the active form (Vallejo et al., 2004). Enzymes refolded using such procedure are often of low activity (Sørensen and Mortensen, 2005). Alternatively a secretion vector can be used, which may eliminate the possibility for inclusion body formation and secrete the native form of the protein into the medium. However, when trying to express the recombinant proPAP1 serine proteinase, neither of the systems gave satisfactory outcome. Under the optimal condition, proPAP1 was secreted into the medium at a low concentration in the baculovirous expression system (~ 0.37 mg/L, Wang et al. 2001). The *E. coli* system was also used to express proPAP1, but the protein was insoluble and could not be refolded (Wu and Kanost, unpublished results). To improve the expression level of proPAP1, the *Bacillus subtilis* expression system was explored in this project.

*Bacillus subtilis* is an attractive host for expression of proPAP1 for the following reasons: (1) *B. subtilis* is generally considered safe and non-pathogenic (Sorokulova et al,. 2008); (ii) it has no significant bias in codon usage (Moszer et al., 1999), which will enable efficient translation of *M. sexta* proPAP1; (iii) it can secrets proteins directly into the culture medium (Tjalsma et al., 2004), therefore may alleviate the potential toxicity of proPAP1, a serine proteinase; (iv) since *B. subtilis* has been widely for commercial production of many enzymes (Schallmey et al., 2004), there is a large body of information concerning the expression and large scale production.

An example to use *B. subtilis* to express secreted proteins is to express heterologous human interferon- $\alpha$ (hIFN-α) with the α-amylase (AmyE) secretion peptide. The results shown that the expression level of hIFN-α with the α-amylase (AmyE) secretion peptide was 15mg/L which was increased by more than three fold compared with the expression of hIFN-α without the α-

amylase (AmyE) secretion peptide. Besides that, *B. subtilis* had also been used to express a heterologous enzyme cutinase from the fungus *F. solani pisi*. Cutinase was constructed with a 6xhis tag for further purification. The signal peptide LipA helped cutinase to completely secreted into the medium with an active form. A strong promoter $P_{59}$ helped to increase the secreted amount to 60 mg/L detected from the cultured supernatant (Brockmeier et al., 2006).

The new vector pHT43 (Molecular Biotechnology, GmbH) allows high level expression of recombinant proteins and directs the recombinant proteins into the medium using its built-in signal peptide. This vector contains the strong $\sigma^A$-dependent promoter in front of the *groE* operon of *B. subtilis*, which has been converted into an IPTG-inducible promoter by adding the *lac* operator. An efficient Shine-Dalgarno (SD) sequence as well as a multiple cloning site (*Bam*H I, *Xba* I, *Aat* II, *Sma*I) were also inserted for strong translation and convenient subcloning (*Bacillus subtilis* Expression Vectors manual, Molecular Biotechnology).

Based on the information available, *B. subtilis* expression system (Molecular Biotechnology, GmbH) may hold the promise to achieve high level expression of *M. sexta* proPAP1 serine proteinase where baculovirus and *E. coli* expression systems failed to deliver. The newly constructed vector pHT43 is a shuttle vector which offers convenience in subcloning and production of large quantity of plasmid DNA in *E. coli*. The IPTG inducible operon and the ability to secrete proPAP1 into the medium may minimize the potential toxicity of the proteinase to the host.

## The importance and goal of my project

According to our previous results, proPAP1 had a low expression level in either baculovirus or *E. coli* expression system. Another issue needs to be considered is the cost of production. *Baculovirus* expression system is much more expensive than the *B. subtilis* system. In addition, proPAP1 can be secreted into medium under the *B. subtilis* expression system. The goal of my project is to establish a new expression system in *B. subtilis* to express high level of proPAP1 and its mutants for future structural and functional relationship studies.

# Chapter 2 - Materials and Methods

## Subcloning of proPAP1 WT and mutants into pHT43 vector

The coding sequences of the four constructs (proPAP1$_{WT}$, proPAP1$_{I}$, proPAP1$_{Xa}$, proPAP1$_{I,Xa}$) in the pET28a vector were provided by Di Wu. PCR reactions were performed to amplify the inserts using primers pHTPAP1-F and pHTPAP1-R (Table 1). The amino-terminal signal peptide was excluded from the constructs, and a 5' *Bam*HI recognition site in primer pHTPAP1-F was added to the 5' end through the PCR reactions. A sequence encoding a 6-histine tag was also added to the 3' of the constructs as well as an *Aat*II digestion site. The PCR reaction was carried out in 50 μL with all the reagents from Invitrogen: 0.5 μL template DNA (200 ng/ μL), 1 uL forward primer pHTPAP1-F (200 nmol/L), 1 uL reverse primer pHTPAP1-R (200 nmol/L), 5 uL 10X high fidelity PCR buffer, 1 μL 10 mM dNTPs, 2 μL 50mM $MgSO_4$, 0.2 μL High fidelity polymerase (5 U/ μL), 39.3 μL nuclease-free water. The cycling parameters were as follows: 94°C for 4 s, followed by 30 cycles: 30 s at 94°C, 30 s at 55°C, and 1.5 min at 68°C for each cycle. After cycling, the PCR reactions were further elongated for 10 min at 68°C. After PCR, 45 μL reaction products together with 8 μL 6x loading buffer were loaded onto a 0.8% agarose gel containing ethidium bromide. Gel electrophoresis was performed in 1xTAE buffer. PCR products were separated through the electrophoresis under 150V for 30 min and then the DNA fragments of interest were cut from the gel under the UV light. The fragments were then purified using a Gel Extraction Kit (Qiagen). After purification, the concentration was determined by gel electrophoresis and comparison with known amounts of DNA in bands of a low DNA mass ladder (Invitrogen), and then ligated into the TOPO vector (TA cloning kit, Invitrogen) at 4°C overnight. After ligation, the vectors with inserts were used to transform *E. coli* strain DH5α (Invitrogen), which were then plated on LB plates with ampicillin (50 μg /mL). The plates were left in a 37°C incubator overnight. To purify the plasmids, four colonies for each construct were used to inoculate 4 mL LB medium with ampicillin (50 μg /mL) in a shaking incubator (300 rpm) overnight at 37°C. The plasmid DNAs were then isolated using a High-Speed Plasmid Mini Kit (IBI scientific). Because the fragments contained 5' *Bam*HI site and 3' *Aat*II site, a double digestion was performed using these two enzymes to identify the colonies containing the desired plasmids. The double digestion was performed in 20 μL with the following: Buffer 4 (NEB) 2 μL, *Bam*HI 0.5 μL, *Aat*II 0.5 μL, bovine serum albumin (10 mg/ml) (100x) 0.2 μL, DNA sample

2 μL, water 14.8 μL. After overnight incubation at 37°C, the digestion fragments were resolved by electrophoresis in an 0.8% agarose gel. Clones containing expected size bands were further verified by DNA sequencing.

To subclone into the pHT43 expression vector, the coding sequences of the four constructs were released from TOPO vector with *Bam*HI and *Aat*II double digestion. The fragments were purified as described above and ligated into the pHT43 vector that had been digested with the same two enzymes. The ligation reactions were performed as follows: pHT43 (22 ng/μL) 10 μL, PCR fragments (140 ng/μL) 1.5 μL, T4 ligase (NEB) 1 μL, T4 ligase buffer 2 μL, water 5.5 μL. After overnight ligation at 4°C, the ligated vectors with inserts were used to transform *E. coli* strain DH5α (Invitrogen). The colonies were then screened on SOC plates with ampicillin (50 μg/L) in a 37°C incubator overnight. After incubation, the positive clones were picked and used to inoculate 13 mL tubes containing 4 mL SOC medium and ampicillin (50 μg/mL) in a shaking incubator at 300 rpm, 37°C overnight. DNA extraction was performed to isolate plasmids using High-Speed Plasmid Mini Kit (IBI scientific). *Bam*HI and *Aat*II double digestion was performed, and the fragments were resolved by electrophoresis in a 0.8% agarose gel to find the right clones. The sequences of the clones were further verified by DNA sequencing. To verify the sequence, two pairs of primers were designed to cover full length of all the fragments (Table 1).

## Preparation of WB800N *Bacillus subtilis* competent cells and transformation

*Bacillus subtilis* strain WB800N was acquired from MoBiTec GmbH (Germany), and the preparation protocol was provided by the MoBiTec GmbH as follows: An overnight culture of the *B. subtilis* strain WB800N was grown in 5 mL HS medium at 37°C. 0.5 mL of the overnight culture was used to inoculate 50 mL HS medium and incubated under vigorous shaking at 37°C. OD600 of the cultural was measured at time intervals to draw the growth curve. Samples of 10 mL were removed at 15 min intervals when cells reached the stationary phase. One mL of sterile glycerol (87% saturation) was added to the bacteria, mixed and left for 15 min on ice. The cells were fractionated into 1 mL aliquots, frozen in liquid nitrogen, and stored at -80°C.

The transformation process was carried out as follows: An aliquot of cells was thawed at 37°C. These cells were used to inoculate 20 mL LS medium and shake slowly in a 30°C water bath for 2 h to obtain maximal competence. One mL aliquots were added to a 2 mL Eppendorf tube, add 10 μL of 0.1 M EGTA, and incubated for 5 min at room temperature. Added 1 μg plasmid DNA and incubated for 2 h at 37°C, with the incubator shaking at 300 rpm. Plated 500 μL bacteria on SOC medium with chloramphenicol (5 μg/mL) and ampicillin (50 μg/mL). Incubated at 37°C overnight to select the positive colonies.

## Protein expression

To test IPTG induction, the appropriate *B. subtilis* colonies were used to inoculate in 5 mL fresh SOC medium (chloramphenicol (5 μg/mL)) in a shaking 37°C incubator (300 rpm) at 37°C overnight. The next day, the overnight culture was diluted 1:10 with fresh SOC medium and then incubated in a shaking incubator until the OD600 reached 1.0. Then the culture was split into two halves, with one being induced with 1 mM IPTG and the other one without IPTG. Samples (200 μL) were collected at different time points from both halves for analysis (t = 0, 2, 4, 6, 20 h). To analyze the protein induction, 100 μL of culture was mixed with 20 μL 6 x SDS loading buffer, and heated at 95°C for 5 min to lyse the cells and denature the proteins. After denaturation, the samples were centrifuged at 12,000 rpm for 2 min to remove the precipitate, and the supernatants were analyzed by SDS PAGE.

## Immunoblot analysis

Protein samples were analyzed on a 12-well, gradient (4-12%) NuPAGE Bis-Tris gel (Invitrogen Life Technologies). The gel was electrophoresed in the MES SDS running buffer (Invitrogen Life Technologies) at 200V for 50 min. After electrophoresis, proteins were transferred to a nitrocellulose membrane. The membrane was rinsed briefly in 1x TTBS (137 mM NaCl, 2.7 mM KCl, 25 mM Tris (pH7.4), 0.05% (v/v) Tween20), and the marker lanes were cut and stained with amido black for 1 min, followed by 2x 1 min destaining and 1x 5 min wash with water.

The rest of the membrane was used for immune detection. The immunoblotting procedures were as follows. Membranes were blocked in 3% dry milk in TTBS (mixture of Tris-Buffered Saline

and Tween 20) buffer for 30 min. Then, membranes were incubated with proPAP1 antibody (made by Haobo Jiang) (rabbit serum) at 1:2000 dilution in 3% dry milk overnight (Kanost lab protocol). The membranes were washed with TTBS buffer for 3 x 5 min. Then the membranes were incubated in goat Anti-Rabbit IgG (H+L)-AP conjugate antibody (Bio-Rad) (1:3000 diluted in 3% dry milk in TTBS) for 1 h. After incubation with the secondary antibody, the membranes were washed with TTBS buffer 2 x 5 min, followed by a wash with TBS (137 mM NaCl, 2.7 mM KCl, 25 mM Tris pH 9.2) for 5 min. The bands were detected using AP conjugate substrate kit (Bio-Rad).

## SDS-PAGE and Coomassie staining

Samples were loaded to a 10-well NuPAGE gradient (4-12%) Bis-Tris gel (Invitrogen Life Technologies), marker 12 (Invitrogen) was used as a molecular weight standard. The gel was electrophoresed in 1xMES SDS running buffer (Invitrogen Life Technologies) at 200V for 50 min. After electrophoresis, the gel was stained with Coomassie Staining Solution (12% (v/v) glacial acetic acid, 50% (v/v) methanol, 0.2% (w/v) Coomassie Brilliant Blue) for 40 min, then changed to destaining solution (10% (v/v) glacial acetic acid, 30% (v/v) methanol) until the background became clear.

## Optimum IPTG concentration for protein induction

Because IPTG concentration may affect the protein expression level, an IPTG concentration test was performed using proPAP1$_{I,Xa}$ and proPAP1$_{WT}$ as examples. The appropriate *B. subtilis* colonies were used to inoculate 5mL fresh SOC medium (chloramphenicol (5 μg/mL)) in a shaking incubator (300 rpm) at 37°C overnight. The overnight culture was diluted 1:10 with fresh SOC medium and then incubated in a shaking incubator until the OD600 reached 1.0. Then the culture was split into five aliquots, and IPTG was added at 0.5 mM, 1 mM, 2 mM, 4 mM respectively and one without IPTG as control. 200 μL of each sample was collected at 4 h after continued culture with shaking at 37°C. To analyze protein induction, 100 μL of culture was mixed with 20 μL 6 x SDS loading buffer, and heated at 95°C for 5 min to lyse the cells and denature the proteins. After denaturation, the samples were centrifuged at 12,000 rpm for 2 min to remove the precipitates, and 20 μL each of the supernatants were analyzed by immunoblot.

## Test of secretion of the proPAP1 proteins

Since pHT43 contains a secretion signal peptide, the recombinant proPAP was expected to be secreted into the medium. To test the amount of secretion, the following study was performed. Overnight cultured *B. subtilis* colonies were diluted 1:10 into 5mL of fresh SOC medium (chloramphenicol (5 μg/mL)) and inoculated in a shaking incubator (300 rpm) at 37°C until the OD600 reached 1.0. Then all the cultures were split into two halves, with one induced with 1 mM IPTG and the other one without IPTG. 200 μL of each sample was collected and centrifuged at 4°C, 12,000 rpm for 10 min to separate the supernatants and pellets, then 100 μL supernatant or 100 μL medium were mixed with 20 μL 6 x SDS loading buffer separately. The pellets were suspended in 120 μL 1x SDS loading buffer and heated all of these at 95°C for 5 min to lyse the cells and denature the proteins. After denaturation, all the samples were centrifuged at 12,000 rpm for 2 min to remove the precipitates, and 20 μL each supernatant was analyzed with immunoblot.

## Comparison of growth media for IPTG induction

An overnight culture of *B. subtilis* containing pTH43 constructs from a single colony in SOC medium was diluted (1:10) separately into a 50 mL flask with 5 mL fresh 2xYT medium (chloramphenicol (5 μg/mL)) and another 50 mL flask with 5mL fresh SOC medium (chloramphenicol (5 μg/mL)). Cultures were inoculated in a shaking 37°C incubator (300 rpm) until the OD600 reached 1.0. Then the cultures were split into two halves, with one induced with 1 mM IPTG and the other without IPTG. 200 μL of samples were collected at 0, 1, 2, 4, 6, 20 h induction in 2xYT medium or at 0, 4, 6 hours induction in SOC medium. To analyze proPAP1 expression, 100 μL each of these samples were mixed with 20 μL 6 x SDS loading buffer, and heated at 95°C for 5 min to lyse the cells and denature the proteins. After denaturation, the samples were centrifuged at 12,000 rpm for 2 min to remove the precipitate, and 20 μL each of the supernatants were analyzed by immunoblots.

## Comparison of induction temperature for IPTG induction

An overnight culture of *B. subtilis* from a single colony containing pTH43 constructs was diluted 1:10 into two 50 mL flasks with 5 mL fresh 2xYT medium each (chloramphenicol (5 μg/mL)) and inoculated in a shaking 37°C incubator (300 rpm) until the OD600 reached 1.0. Then the cultures were split into two halves, and incubated at 16°C with 1 mM IPTG or without IPTG induction. Samples (200 μL) were collected at 0, 2, 4, 6, 20 h induction at 16°C or 37°C. For the analysis, 100 μL of culture was mixed with 20 μL 6 x SDS loading buffer, and heated at 95°C for 5 min to lyse the cells and denature the proteins. After denaturation, the samples were centrifuged at 12,000 rpm for 2 min to remove the precipitate, and 20 μL each of the supernatants were analyzed by immunoblotting.

## Ammonium sulfate precipitation and purification the proPAP1$_I$ protein

To establish the purification protocol, proPAP1$_I$ was used to test the best ammonium sulfate precipitation concentration. A *B. subtilis* strain harboring a plasmid for expression of proPAP1$_I$ was cultured in 5 mL SOC medium overnight, then expanded into 50 mL 2xYT medium (16 g tryptone, 10 g casamino acids, 5 g sodium chloride, add distilled water to 1000 mL, autoclave 15 mins) with chloramphenicol (5 μg/mL), and inoculated in a shaking incubator at 37°C until OD600 reached 0.8. IPTG was added to 1 mM, and the induction was carried out for 20 h. After induction, bacteria pellet and supernatant were separated by centrifugation at 4°C at 12,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) for 15 min. The supernatant was then transferred to a 1L beaker, and 50 mL of the supernatant was subjected to ammonium sulfate precipitation by sequentially adding 15.05 g, 9.49 g and 11.30 g of ammonium sulfate until the concentration of ammonium sulfate reached 50%, 75%, 100% saturation accordingly. After stirring for 1 h at 4°C for each ammonium sulfate concentration, the mixtures were centrifuged at 12,000 rpm for 15 min at 4°C. The supernatant was discarded and the protein pellets were dissolved in 1mL 1xSDS gel loading buffer, then heated at 95°C for 3 min before protein gel analysis.

## Ni-NTA affinity chromatography

The supernatant from 1 L proPAP1$_I$ culture following IPTG induction was prepared as described above, and subjected to precipitation at 50% saturated ammonium sulfate. After precipitation for 1 h at 4°C, the protein pellet was collected by centrifugation at 15,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) at 4°C for 20 min. The pellet was dissolved in 25 mL of starting buffer (299 mM NaCl, 10mM imidazole, 0.5 M sodium phosphate buffer, pH 8.0), then centrifuged again at 15,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) at 4°C for 15 min. The supernatant was then dialyzed against 4 L of starting buffer overnight to remove ammonium sulfate. After dialysis, the supernatant was centrifuged again at 15,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) at 4°C for 15 min. Then the supernatant was collected, and 100 μL of the supernatant was mixed with 20 μL 6x SDS gel loading buffer, heated at 95°C for 5 min and saved for further analysis.

Four mL of Ni-NTA (Qiagen) beads were centrifuged at 12,000 rpm for 10 min at 4°C. After centrifugation, the supernatant was discarded, and the beads were washed again with water. After washing, the beads were incubated with the protein in 25 mL starting buffer placed on a rotator and mixed for 1 h at 4°C (at a speed of 12). After 1 h of binding, the entire contents were poured into a 1.5 cm diameter Econo column. After collecting the flow through, the column was washed twice with 10 mL wash buffer (299 mM NaCl, 20mM imidazole, 0.5 M sodium phosphate buffer, pH 8.0. The column was then washed with additional 5 mL of wash buffer and collected. After the last wash, the bound proteins were eluted with elution buffer (299 mM NaCl, 250 mM imidazole, 0.5 M sodium phosphate buffer pH 8.0). Ten 200 μL elution fractions were collected, followed by four 500 μL elution fractions in 1.5 mL microfuge tubes. After collection, $A_{280}$ value of flow through (FT), Wash1-3 (W1, W2, W3) and Elution 1-14 (E1-E14) were analyzed with a Nanodrop spectrophotometer. Samples of each fraction (100 μL) were transferred to 1.5 ml centrifuge tubes, mixed with 20 μL of 6x SDS loading buffer, and heated at 95°C for 5 min. All fractions were then analyzed by SDS PAGE followed by western blot or Coomassie blue staining.

## Protease and Phosphatase Inhibitor experiment

The supernatant from 1 L proPAP1$_I$ culture following IPTG induction was prepared as described above, two tablets protease and phosphatase inhibitor-EDTA free (Thermo Scientific, 88669) was added during IPTG induction. Another 1L proPAP1$_I$ culture was also done following IPTG induction but without any inhibitor. One ml of each supernatant and whole medium were collected for further analysis. After that, both supernatants were subjected to precipitation at 50% saturated ammonium sulfate. After precipitation for 1 h at 4°C, the protein pellets were collected by centrifugation at 15,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) at 4°C for 20 min, 1 mL supernatants were collected separately for further analysis. Then each pellet was dissolved in 30 mL of starting buffer (299 mM NaCl, 10mM imidazole, 0.5 M sodium phosphate buffer, pH 8.0), and then centrifuged again at 15,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) at 4°C for 15 min. 100 µL each of the supernatants were collected for further analysis. The supernatants were then dialyzed against 8 L of starting buffer for two days to remove ammonium sulfate. After dialysis, the supernatants were centrifuged again at 15,000 rpm (Sorvall RC-5C Plus Superspeed Centrifuge) at 4°C for 15 min. Then the supernatant were collected, and 100 µL of the supernatant was saved at -20°C for further analysis.

Four mL of Ni-NTA (Qiagen) beads were centrifuged at 12,000 rpm for 10 min at 4°C. After centrifugation, the supernatant was discarded and the beads were washed again with water. After washing, the beads were incubated with the protein in 25 mL starting buffer placed on a rotator and mixed for 1 h at 4°C (at a speed of 12). After 1 h of binding, the entire contents were poured into a 1.5 cm diameter Econo column. After collecting the flow through, the column was washed twice with 10mL wash buffer (299 mM NaCl, 20mM imidazole, 0.5 M sodium phosphate buffer, pH 8.0. After the second wash, the bound proteins were eluted with elution buffer (299 mM NaCl, 250 mM imidazole, 0.5 M sodium phosphate buffer pH 8.0). Eight 500 µL elution fractions were collected in 1.5 mL microfuge tubes. After collection, $A_{280}$ value of flow through (FT), Wash1-3 (W1, W2) and Elution 1-14 (E1-E8) were analyzed with a Nanodrop spectrophotometer. Samples of each fraction (100 µL) were transferred to 1.5 ml centrifuge tubes, mixed with 20 µL of 6x SDS loading buffer, and heated at 95°C for 5 min. All fractions were then analyzed by SDS PAGE followed by western blot or Coomassie blue staining.

# Chapter 3 - Results

## Subcloning of proPAP1 proteins into pHT43 shuttle vector

ProPAP1 has a clip domain near its amino-terminal and a serine proteinase domain near its carboxyl-terminal (Jiang et al., 1998) (Fig1A). To further study the function of proPAP1, three mutants generated by Di Wu through point mutagenesis were named proPAP1$_{I,}$ proPAP1$_{I,Xa}$ and proPAP1$_{Xa}$ (Fig 1B). All four constructs contain a six histidine tag at the carboxyl end. To attempt to achieve higher expression of proPAP1$_{WT}$ and mutant constructs, a *B. subtilis* expression system was used. The pHT43 shuttle vector (Fig 1C) was selected as the expression vector, which includes a signal peptide that enables secretion of the proteins of interest into the medium. Since pHT43 is a shuttle vector, which can exist both in *E. coli* and *B. subtilis*, subcloning could be carried out within *E. coli*, then the plasmids used to transform *B. subtilis*. In order to do this, three propap1 mutants 6xHis propap1$_{I}$ (S311A), 6xHis propap1$_{I,Xa}$ (N105I, G106E, D107G, S311A), 6xHis propap1$_{Xa}$ (N105I, G106E, D107G) generated by Di Wu in our laboratory were used as templates (Fig1B). Propap1$_{I}$'s catalytic serine was replaced with alanine, propap1$_{Xa}$'s activation site was changed to permit activation by bovine factor Xa, and propap1$_{I,Xa}$ was a double mutant. Two restriction sites -- *AatII* and *BamHI* were introduced into the 3' end and 5' ends of proPAP1 and its mutants using PCR method, then subcloned into pHT43. Plasmids were first screened using digestion with *AatII* and *BamHI*, and then tested on agarose gel electrophoresis to identify the expected 8 Kb and 1095 bp fragments (Fig 2A). The sequences were further verified by DNA sequencing (Iowa State University DNA facility). The protein alignment is shown in Figure 3.

ProPAP1$_{WT}$ and mutants were subcloned into the *BamHI* and *AatII* sites, and the positive clones were identified by double digestion with the same restriction enzymes (Fig 2). When first screening using plates made with LB medium, the resulting clones all contained frame-shift mutations at the 5' end. I hypothesized that the expression of these constructs may be toxic to the bacteria, and that the small amount of lactose contamination in casein could increase the expression of these constructs beyond the level that could be tolerated by the bacteria, therefore resulting in the survival of only the clones with the frame-shift mutations. To avoid this issue,

SOC medium, which lacks casein, was used for all the subcloning procedures (Sun et al., 2009) and the frame-shift was eliminated in the subsequent subcloning step.

Because pHT43 plasmid DNA needs to form linear plasmid multimers before successful transformation into *B. subtilis* competent cells (Bron et al., 1985; Gruss et al., 1988), the plasmids were used to transform *E. coli* strain JM101, which is RecA (+), and can help the formation of multimers. All constructs were first used to transform *E. coli* strain JM101 competent cells, then pHT43 plasmid DNA was extracted from JM101 and stored at -20°C before being used to transform *B. subtilis*.

## Preparing *B. subtilis* WB800N competent cells

*B. subtilis* cells are naturally competent (Singh et al., 2012), *B. subtilis* strain WB800N was prepared according to the protocol provided by the Molecular Biotechnology Company. The main growth phases of *B. subtilis* undergoes the lag phase, the exponential phase, the stationary phase and the death phase. To get the maximum competence, the *B. subtilis* competent cells were collected at three consecutive time points near the stationary phase (Fig 4). *B. subtilis* competent cells were divided into several aliquots and stored at -80°C until use.

## Optimizing conditions for proPAP1 expression and secretion

Protein induction may be affected by IPTG concentration, induction time, growth medium, and induction temperature. I evaluated the effect of IPTG concentration on proPAP1 expression. Expression clones of $proPAP1_{WT}$ (Fig 5A) and $proPAP1_{I,Xa}$ (Fig 5B) were induced with final IPTG concentration of 0.5 mM, 1 mM, 2 mM or 4 mM for 4 h in SOC medium. Supernatants were all collected at different IPTG concentration and western blot with proPAP1 antibody was employed to detect the 44 KDa band with the size similar to the calculated molecular weight of $proPAP1_{WT}$ and $proPAP1_{I,Xa}$ (Fig 5). The proPAP1 expression reached a plateau at 0.5 mM IPTG and did not increase further at higher IPTG concentrations. To test the best IPTG concentration for producing secreted proPAP1, another experiment was done using 6 h induction and 20 h induction, with the concentration of IPTG from 0.5 mM, 1 mM, 2 mM to 4 mM in 2xYT medium. Supernatants were collected, and western blot with proPAP1 antibody and

coomassie staining were done to detect the best IPTG concentration (Fig 5C). An IPTG concentration of 0.5 mM with 20 h induction time could induce the highest protein expression level, although protein degradation was observed. With no IPTG, there was a low proPAP1 protein expression level. The difference between different IPTG concentration greater than 0.5mM was negligible. For the following experiments, 1 mM IPTG was used for induction.

## IPTG induction time course

Since induction time is also a very important factor to affect protein expression level, a time course experiment from 2 h, 4 h, 6 h to 20 h induction time was performed with all constructs (Fig 6 A, B, C, D) in SOC medium. Pellets and supernatants were collected together at each time point, and immunoblotting with proPAP1 antibody was used to detect the recombinant protein expression level at around 44 KDa. Expression of proPAP1 was greatest at 20 h for each construct. To test the best induction time for production of secreted proPAP1, another time course experiment was done using with induction in 2xYT medium. Supernatants were collected at different time points, western blot with proPAP1 antibody and coomassie staining were done to detect the best induction time (Fig 6E). Twenty hour induction resulted in the highest proPAP1 protein expression level. Degradation of proPAP1 was apparent at all induction times.

## Secretion of proPAP1 proteins

Because pHT43 vector contains a signal peptide, proPAP1 was expected to be secreted into medium. To test this, bacteria were collected after 6 h induction with 1mM IPTG and with no IPTG as a control. The supernatant and pellet were separated, and the appearance of the proPAP protein was monitored by western blot (Fig 7). More than half of the $proPAP1_{I,Xa}$ was secreted into supernatant (Fig 7A). This phenomenon was also observed with $proPAP1_{WT}$ (Fig 7B) and $proPAP1_{I}$ (Fig 7C). However, with $proPAP1_{Xa}$, it seemed most protein remained in the cell pellet (Fig 7D). Also from the comparison of these proteins expression level it appeared that $proPAP1_{I}$ had the highest protein expression, while $proPAP1_{Xa}$ had the lowest protein expression.

## Comparison of growth media for IPTG induction

Because proPAP1$_I$ appeared to have the highest level of protein expression, proPAP1$_I$ was chosen for the next set of optimization experiments. To know if 2xYT medium, which is recommended by MoBiTec, is better than SOC medium, a comparison of the two media was performed (Fig 8). A time course from 1 h, 2 h, 4 h, 6 h induction with proPAP1$_I$ was done using 2xYT medium. Supernatants were collected at each time point and western blot with proPAP1 antibody was applied to compare the expression level at different induction times. The time course induction from 4 h and 6 h induction in SOC medium was also used as a control. The proPAP1$_I$ expression level appeared to be significantly greater in 2xYT medium compared with SOC medium.

## Comparison of induction temperature

ProPAP1 and proPAP1 mutants appeared to become degraded during IPTG induction. To try to solve this problem, a lower temperature 16°C induction was tested (Fig 9). ProPAP1$_I$ was induced at 16°C for 0, 2, 4, 6 and 20 h and compared with a 6 h induction at 37 °C. Low temperature did not help to prevent protein degradation, while the protein expression level was dramatically less at 16°C. In this case, low temperature induction did not improve yield of proPAP1.

Ni-NTA affinity purification of proPAP1$_I$ proteinThe 6xHis tag at the C-terminal of the four constructs was designed for use in Ni-NTA affinity chromatography. To test the procedure, supernatant from 1 L proPAP1$_I$ culture induced with 1mM IPTG for 20 h at 37°C was subjected to 50% saturated ammonium sulfate to precipitate proPAP1$_I$. The pellet was dissolved in buffer (see Material and Methods) and then incubated with Ni-NTA beads. Washing and elution with imidazole fraction was analyzed by $A_{280}$ and immunoblot or Commassie blue detection after SDS-PAGE (Fig 10). Most proPAP1$_I$ protein bound to the column, and then eluted in buffer containing 250 mM imidazole (Fig 10 B, C). Contaminating protein was removed by flow through and washing steps. However, protein degradation was detected as low molecular weight bands detected by the proPAP1 antibody. Measurement of the $A_{280}$ of the eluted fractions containing proPAP1$_I$ was used to estimate the protein concentration. In this experiment approximately 0.06 mg of proPAP1$_I$ was isolated, starting from one L of medium,

## Protease Inhibitors for proPAP1$_I$ expression

Degradation of proPAP1 occurred during expression and purification. To test if protease inhibitors could help prevent protein degradation, a mixture of protease inhibitors was added to 1 L 2xYT medium during the IPTG induction step and compared to a culture without protease inhibitors during the induction. The protease inhibitor mixture contains aprotinin, which inhibits serine proteases, bestatin which inhibits aminopeptidase proteases, E-64 which inhibits cysteine proteases, and leupeptin which inhibits serine and cysteine proteases. In order to be able to make direct comparison, the supernatants from the induction with or without protease inhibitors were treated equally and subjected to the same purification protocol as outlined in Figure 10 (see Material and Methods). The total input, flow through, washing and elution fractions were analyzed by $A_{280}$ absorbance, Western blotting and Commassie blue staining (Fig. 11). The previously purified proPAP1$_{I,Xa}$ from sf9 cells (purified by Di Wu and Daisuke Takahashi) was used as a standard.. From Fig 11 B&D, it seems that the serine proteases inhibitors did not effectively prevent proPAP1 protein degradation. Protein degradation could also be seen from Fig 11 D from several elution fractions, but the degradation may have occurred during the purification process, when the inhibitors were not present.

# Chapter 4 - Discussion

## Recombinant expression and purification of clip domain proteases

Clip-domain serine proteases are very important for innate immunity and embryonic development of invertebrates, and the clip domains are always located at the N-terminal of serine proteases (Jiang et al., 1997; Jiang et al., 2003a; Jiang et al., 2003b). Due to the functional importance of clip-domain serine proteases, a number of biochemistry studies had been done to try to understand the function of clip-domain serine proteases. The crucial first step for the functional study is to express and purify these proteases.

Several systems have been used to obtain purified clip-domain serine proteases of insects. The first is direct purification from the hemolymph. BAEEase, a serine protease which hydrolyzes the synthetic substrate $N^{\alpha}$-benzoyl-L-arginine ethyl ester BAEE in *Bombyx mori*, was purified from *Bombyx mori* hemolymph as a zymogen which contains an amino-terminal CLIP domain and a carboxyl-terminal serine protease catalytic domain (Jang et al., 2006). *M. sexta* hemolymph is also a good system for direct purification of the clip-domain serine proteases due to the relative large volume (1-2ml) can be harvested from one insect (Kanost et al, 2004) . The *M. sexta* hemolymph had been used as a source to purify serine protease HP21 which recognizes and activates proPAP3 (Gorman et al., 2007). *M. sexta* hemolymph had also been used to purify other serine proteases such as proPAP1, proPAP2 and proPAP3 (Jiang et al., 1997; Jiang et al., 2003a; Jiang et al., 2003b). The hemolymph of a large beetle, *Tenebrio molitor* larvae had also been used to purify several serine protease, such as SAE (Spatzle processing enzyme-activating enzyme), GNBP3 (Gram-negative-binding protein 3), MSP (modular serine protease), SPE (Spatzle processing enzyme), (Kim et al., 2008; Roh et al., 2009).A second method is recombinant expression in *Drosophila* S2 cells. Spatzle processing enzyme which was a clip domain protease identified in *Drosophila* as a BAEEase homolog protein participated in the activation of spatzle and toll signaling pathway. To express this protein, the pMT/V5-His vector containing SPE was used to transfect *Drosophila* S2 cells (Jang et al., 2006). *Drosophila* S2 cell line was also used to express two other clip domain serine proteases from *Drosophila* snake and easter (Tian et al., 2008). *Drosophila* S2 cells were also used to express *M. sexta* proHP6 and proHP8.

A third expression system is the baculovirus expression system. In the expression of proPAP2, the proPAP-2 cDNA was inserted into pFastBac1with a hexa-histidine tag fused at the carboxyl terminus. The virus was then used to infect *Spodoptera frugiperda* Sf21 cells, and the recombinant proPAP-2 was efficiently secreted into the media (Ji et al., 2003). The protein was purified by affinity chromatography on Ni-NTA column. Another example to employ baculovirus as an overexpression system was the expression of prophenoloxidase activating factor-II (PPAF-II) protein, a clip-domain serine proteases which was derived from the beetle *Holotrichia diomphalia* larvae. The recombinant virus was amplified using *Spodoptera frugiperda* 9 (SF-9) cells and a modified vector of pFASTBAC-HTa was used with the mellitin signal sequence inserted before the hexahistidine coding region to secrete the resulting proteins into culture media (Piao et al., 2005). Baculovirus expression system was also used to express *M. sexta* HP2, proPAP1, proPAP2, and proPAP3 recombinant proteins 1 (Wang et al., 2001; (Ji et al., 2003; Wang and Jiang, 2007; Gorman et al, 2007). .

The last expression system I found is to use *Escherichia coli* to express and purify clip-domain serine proteases. The *M. sexta* proPAP2 clip 1 and clip 2 domain had been constructed within *E. coli* strain M15 (Huang et al., 2007). A hexahistidine tag was attached to the amino-terminal. NMR was performed to dissect the structure of these two domains. Another example was to use *E. coli* strain BL21 to purify the clip-domian seine protease PtsSP from *Portunus trituberculatus*. IPTG induction was added to the cultured medium when $OD_{600}$ reached to 0.5-0.8. The cobalt affinity chromatography was performed to purify the PtsSP and then refolded through a linear urea gradient (Song et al., 2013).

## Development of a new system for proPAP1 expression

ProPAP1 is an important serine proteinase involved in the innate immune response of insect *M. sexta*. To conduct biochemical studies to better understand its function, large quantity of purified proPAP1 is required. However, heterologous expression of recombinant serine proteinases such as proPAP1 is always challenging. Two expression systems – the baculovirus and the *E.coli* expression system had been tried to express proPAP1 protein. In the baculovirus expression

system, proPAP1 had a low expression level – around 0.37 mg/L (Wang et al., 2001). In the *E. coli* expression system, over-expression of proPAP1 resulted in inclusion body formation, making purification of active correctly folded proPAP1 difficult. Low temperature induction was tried to decrease the inclusion body formation in order to increase the yield of correctly folded proPAP1, which was also failed to eliminate the inclusion body formation (M. Kanost, personal communication).

To conquer these problems, *B. subtilis* was employed as a new expression system. *B. subtilis* had been engineered and was widely used as a host for large scale non-recombinant protein expression (Priest, 1977; Liu et al., 2013; Van and Hecker, 2013; Harwood and Cranenburgh, 2008). Although subcloning in the *B. subtilis* system was not as straightforward as in the *E. coli* system, a shuttle vector pHT43 has been created to facilitate the subcloning step. The secretion system in *B. subtilis* may help to maintain the biological activity of the expressed protein and to simplify the further purification step (Wu et al., 1991). The other advantages for *B. subtilis* are very safe, easy to handle, easy to culture and adapt to large fermentation and no significant bias in codon usage (Moszer et al., 1999). Because *B. subtilis* is a gram-positive soil bacterium which does not have an outer membrane, compared with gram-negative bacteria, it secretes and absorbs substances more efficiently (Jäger et al., 1992). Furthermore, there is neither endotoxin nor pathogenicity with *B. subtilis* (Wenzel et al., 2011). All of these characteristics make *B. subtilis* a good expression system for recombinant proteins.

To get stable and high yield of protein, a WB800N strain of *B. subtilis* was chosen as the host strain. WB800N strain was engineered to be deficient in eight extracellular proteinases, which will reduce the degradation of the recombinant protein (Zhang et al., 2005). In addition, pHT43 vector containing an amyQ signal peptide gene was constructed in front of proPAP1 to help proPAP1 protein secrete into the medium to keep it in its native conformation.

When I used this system for protein expression, I noticed a lot of degradation during proPAP1 expression and purification, which decreased the amount of protein that could be obtained. Although WB800N is an engineered strain which lacked extracellular serine proteinases, the proPAP1 protein degradation was still very severe. To solve this problem, the protease inhibitors

were used to try to inhibit the protease degradation. From the experimental results, it shows that the protease inhibitors did not prevent protein degradation of proPAP1.

## Compare costs per mg PAP1 of expression in baculovirus and *B. subtilis*

Previous work in our lab used recombinant baculovirus to infect *Spodoptera frugiperda* (Sf21) cells to purify proPAP1$_{WT}$ protein. ProPAP1 was secreted into the medium at a concentration around 0.37 mg/L (Wang et al., 2001). To culture the Sf21 cell, the Sf-900 II medium (Gibco BRL) was used, and the cost for 1 Liter medium is $73. To obtain 1 mg of proPAP1 protein, 2.7 Liter Sf-900 II medium should be used which roughly costs $197. Also each cell culture flask (Corning® 1L Polycarbonate Erlenmeyer Flask with Flat Cap, Product #431146) costs $25 and usually 3 flasks are needed for 1 L culture, so the total cost for making 2.7 L medium should be $398. The cost for virus infection, virus culture, and maintaining the cell line are not included in my calculation. To generate 1 mg of proPAP1 from *B. subtliis*, 16.6 Liters 2xYT medium and 1 mM IPTG will be needed. And the costs for making 16.6 Liters 2xYT is, $48.7 for tryptone (fisher, BP 1421-500), $2.25 for sodium chloride (fisher, BP 358-212), $114.3 for casamino acid (fisher, BP 1424-100) and $22 for IPTG (sigma 16758-1G). The total costs was $187. From the above data, it is evident that to produce 1 mg of secreted protein proPAP1, *B. subtilis* system will only need less than half of the money spent on the baculovirus expression system.

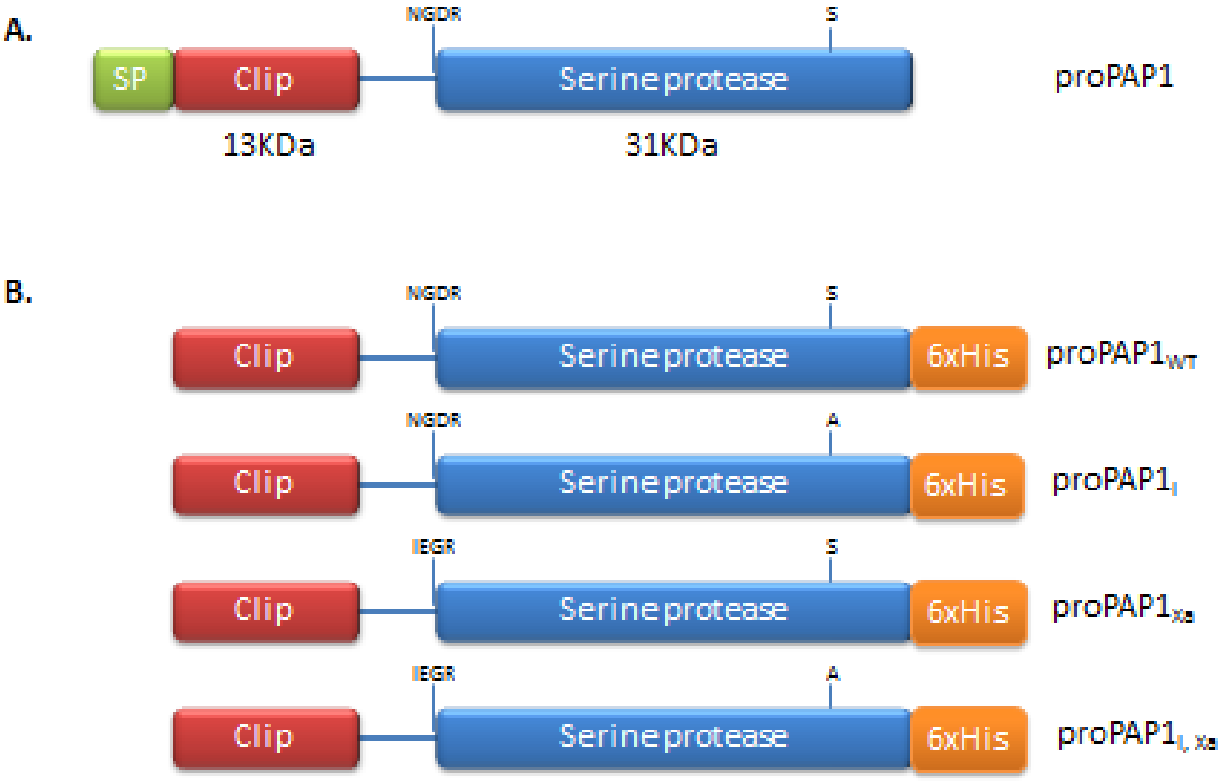
From the above data, it is evident that to produce 1 mg of secreted protein proPAP1, *B. subtilis* system will only need nearly 1/2 money spent on the baculovirus expression system.

The *B. subtilis* expression system seems to be a reasonably good system for proPAP1 expression, but there is still some further work to be done to improve the expression level and downstream protein purification steps. The use of protease inhibitors needs to be optimized to decrease degradation and improve the protein expression level. A fermenter (> 10 liter) will be a good system for large scale culture. Ammonium sulfate precipitation and nickel affinity chromatography are useful as initial purification steps. Since proPAP1 is negatively charged at near neutral pH, an anion exchange chromatography could be employed for the next purification step, following the Ni-NTA affinity chromatography. Also from the results, it seems that there

may still be some protein degradation during the Ni-NTA purification step and more work needs to be done to resolve this issue.

## Future work

ProPAP1 was identified in 1998 by Jiang et al. as a clip-domain serine protease (Jiang et al., 1998). With the discovery of other proPAP family members, a lot of work had been done to try to elucidate the function of PAP1, PAP2 and PAP3. Untill now, we know that proPO is the substrate of PAP family. However, it is unclear if there are some other substrates of PAPs. We would like to use proPAP1 to answer these questions, since it possesses a simpler structure compared to other proPAPs with one clip domain located near the amino-terminal of the protein (Jiang et al., 1998; Jiang et al., 2003a; Jiang et al., 2003b). The first question is if there are some other proteins in hemolymph that are substrates of PAP1. There are several traditional methods to screen for new substrates. The protease degradome could be used as a new tool to understand the question on the system-wide scale (Lopez-Otin and Overall, 2002; Doucet and Overall, 2008). We would like to use the in vitro activated pure $proPAP1_{Xa}$ to mix into hemolymph and perform the degradome experiment to try to answer this question. $ProPAP1_{I}$ and $proPAP1_{I,Xa}$ will be used as negative controls. The mass spectrometry experiments will be used to discover the other protease substrates. Another question we are interested in is the X-ray structure of proPAP1. The solution structure of clip 1 and clip 2 domain in proPAP2 had been dissected by Huang et al., 2007, but the crystal structure of proPAP2 has not been dissected yet. Also the crystal structure of proPAP1 is not available, due to the low expression level of proPAP1 in both *E. coli* and baculovirus system. We tried to build a new expression system of proPAP1 in *B. subtilis* to help us further understand the structure and function of proPAP1. Future improvements may yield $proPAP1_{WT}$ that could be used to generate the crystal structure of proPAP1 and help us further understand the function of clip domain and the serine protease domain.


A.
NGDR
S
SP
Clip
Serine protease
proPAP1
13KDa
31KDa
B.
NGDR
S
Clip
Serine protease
6xHis
proPAP1WT
NGDR
A
Clip
Serine protease
6xHis
proPAP1I
IEGR
S
Clip
Serine protease
6xHis
proPAP1Xa
IEGR
A
Clip
Serine protease
6xHis
proPAP1I, Xa

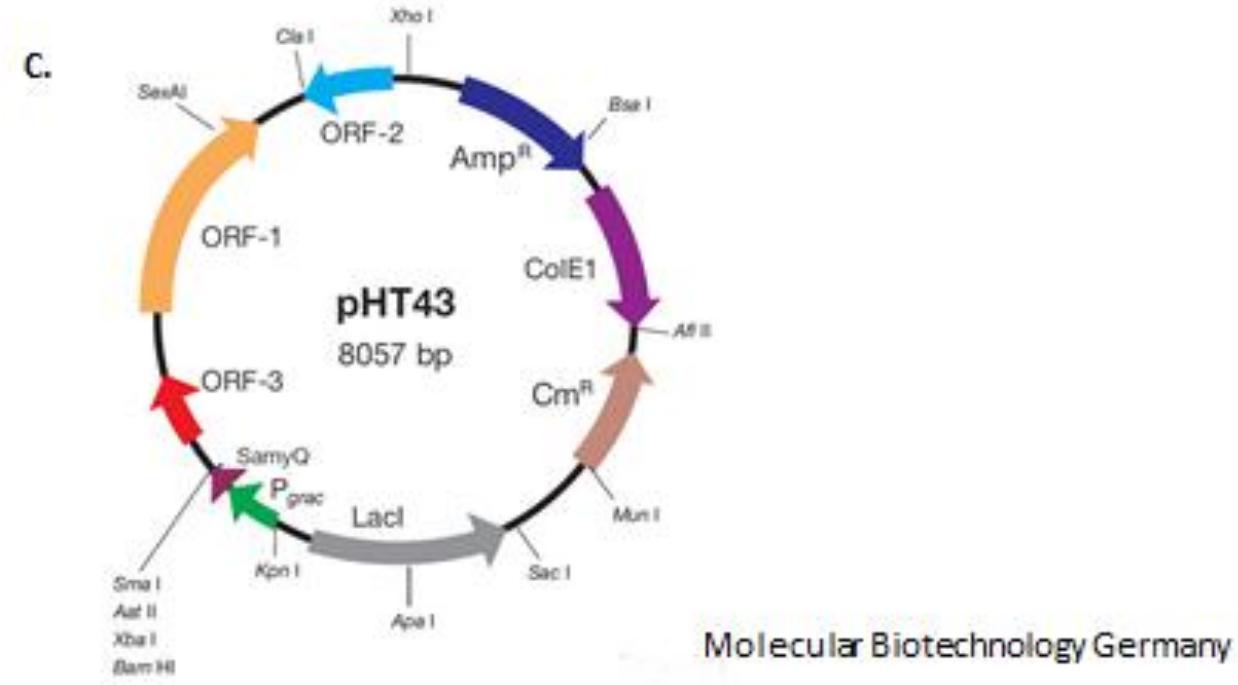



**Figure 1** ***Manduca*** **proPAP1 structure, constructs and expression system**

A. The domain structure of *Manduca sexta* proPAP1. The signal peptide (SP), clip domain, and serine protease domain are highlighted. The HP-6 protease's activation cleavage site (NDGR) in proPAP1 and the serine residue (S) in the active site are also indicated. B. Four molecular constructs used in the study. proPAP1 wild type; proPAP1$_I$: S to A mutation, protease inactive construct; proPAP1$_{Xa}$: HP-6 cleavage site replaced by factor Xa cleavage site; proPAP1$_{I, Xa}$: S to A mutation in combination with factor Xa cleavage site. C. The plasmid map of *Bacillus subtilis* expression vector pHT43 (figure from Molecular Biotechnology, Germany). pHT43 is a shuttle vector, it can be used both in *Bacillus subtills* and *E. coli*. The four molecular constructs listed in Figure 1B were subcloned into the *Bam* HI and *Aat* II site of the vector.

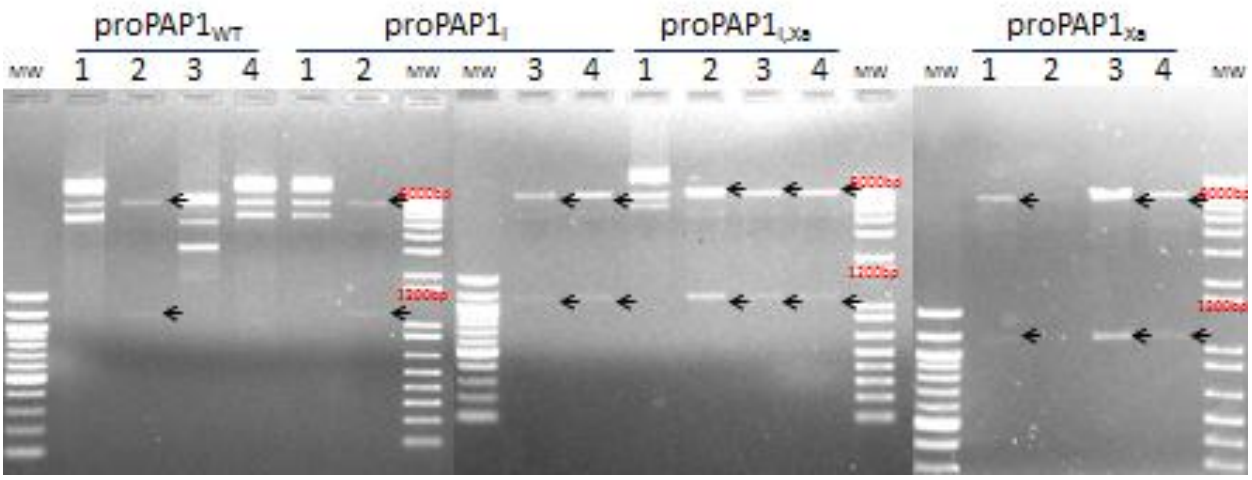



**Figure 2 Subcloning of four molecular constructs into the pHT43 expression vector**

Four colonies (labeled 1-4) from each ligation were picked for restriction digestion analysis with *Bam*HI and *Aat*II. Correct constructs should release a 1095 bp insert fragment and a 9 Kb vector fragment upon completion of digestion. (Arrows pointed to the correct digestion fragments).

```
MIQKRKRTVSFRLVLMCTLLFVSLPITKTSAVGSQSCTTPQGADSNCISLYECPQLLSAF proPAP1WT
MIQKRKRTVSFRLVLMCTLLFVSLPITKTSAVGSQSCTTPQGADSNCISLYECPQLLSAF proPAP1I
MIQKRKRTVSFRLVLMCTLLFVSLPITKTSAVGSQSCTTPQGADSNCISLYECPQLLSAF proPAP1Xa
MIQKRKRTVSFRLVLMCTLLFVSLPITKTSAVGSQSCTTPQGADSNCISLYECPQLLSAF proPAP1I,Xa

EQRPLPSPVVNYLRKSQCGFDGYTPRVCCGPLPQQASRPQPTPAPVPTRAPPVNPGGVDP
EQRPLPSPVVNYLRKSQCGFDGYTPRVCCGPLPQQASRPQPTPAPVPTRAPPVNPGGVDP
EQRPLPSPVVNYLRKSQCGFDGYTPRVCCGPLPQQASRPQPTPAPVPTRAPPVNPGGVDP
EQRPLPSPVVNYLRKSQCGFDGYTPRVCCGPLPQQASRPQPTPAPVPTRAPPVNPGGVDP

TYDEDSSPAPRNQCGVDMNGDRIYGGQITDLDEFPWMALLGYLTRTGSTTYQCGGVLINQ
TYDEDSSPAPRNQCGVDMNGDRIYGGQITDLDEFPWMALLGYLTRTGSTTYQCGGVLINQ
TYDEDSSPAPRNQCGVDMIEGRIYGGQITDLDEFPWMALLGYLTRTGSTTYQCGGVLINQ
TYDEDSSPAPRNQCGVDMIEGRIYGGQITDLDEFPWMALLGYLTRTGSTTYQCGGVLINQ

RYVLTAAHCTIGAVEREVGKLITVRLGEYDTQNSVDCVDDVCADPPQNIPIEAAYPHSGY
RYVLTAAHCTIGAVEREVGKLITVRLGEYDTQNSVDCVDDVCADPPQNIPIEAAYPHSGY
RYVLTAAHCTIGAVEREVGKLITVRLGEYDTQNSVDCVDDVCADPPQNIPIEAAYPHSGY
RYVLTAAHCTIGAVEREVGKLITVRLGEYDTQNSVDCVDDVCADPPQNIPIEAAYPHSGY

SDNNKNRKDDIALVRLTRRAQYTYYVKPICLAHNNERLATGNDVFVAGWGKTLSGKSSPI
SDNNKNRKDDIALVRLTRRAQYTYYVKPICLAHNNERLATGNDVFVAGWGKTLSGKSSPI
SDNNKNRKDDIALVRLTRRAQYTYYVKPICLAHNNERLATGNDVFVAGWGKTLSGKSSPI
SDNNKNRKDDIALVRLTRRAQYTYYVKPICLAHNNERLATGNDVFVAGWGKTLSGKSSPI

KLKLGMPIFDKSDCASKYRNLGAELTDKQICAGGVFAKDTCRGDSGGPLMQRRPEGIWEV
KLKLGMPIFDKSDCASKYRNLGAELTDKQICAGGVFAKDTCRGDAGGPLMQRRPEGIWEV
KLKLGMPIFDKSDCASKYRNLGAELTDKQICAGGVFAKDTCRGDSGGPLMQRRPEGIWEV
KLKLGMPIFDKSDCASKYRNLGAELTDKQICAGGVFAKDTCRGDAGGPLMQRRPEGIWEV

VGIVSFGNRCGLDGWPGVYSSVAGYSDWILSTLRSTNVHHHHHH
VGIVSFGNRCGLDGWPGVYSSVAGYSDWILSTLRSTNVHHHHHH
VGIVSFGNRCGLDGWPGVYSSVAGYSDWILSTLRSTNVHHHHHH
VGIVSFGNRCGLDGWPGVYSSVAGYSDWILSTLRSTNVHHHHHH
```

**Figure 3 Protein sequence alignment of the four constructs**

Protein sequence alignment of the four constructs. All clones contain the 6xHis tag at the C-terminus. The mutations were highlighted.

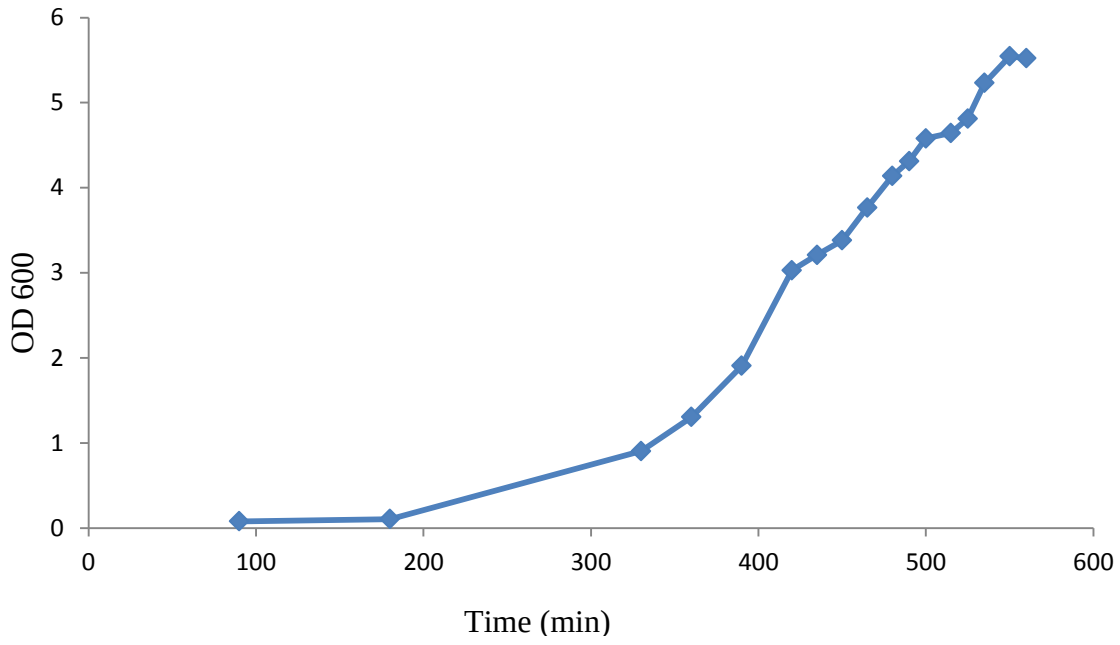



**Figure 4 Growth curve of *B. subtilis* WB800N**

*B. Subtilis* was cultured as described in methods. The OD600 values were plotted against time to generate to the growth curve. Competent cells were collected at OD 5.23, 5.54, 5.52.

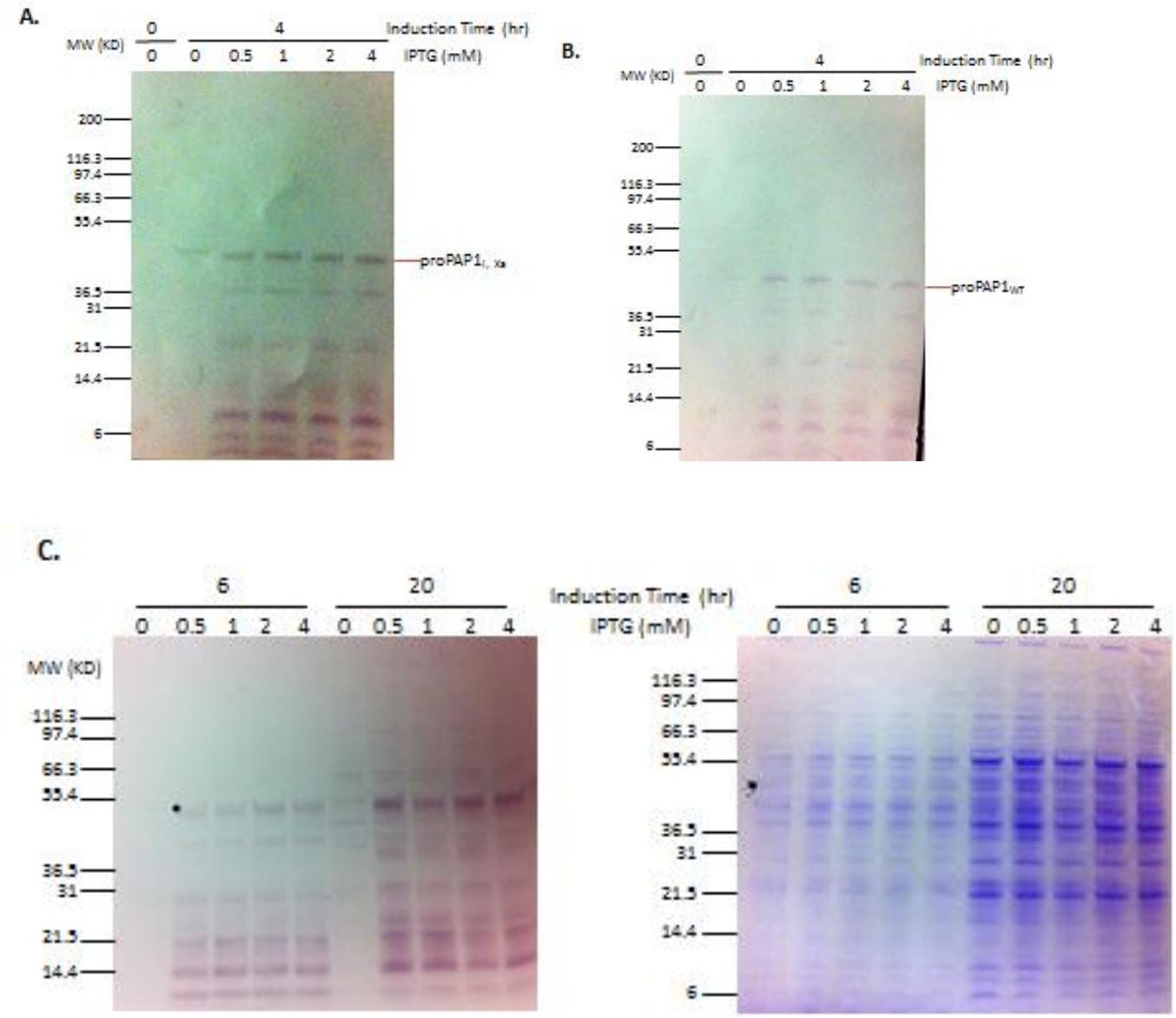



**Figure 5 Effect of IPTG concentration on proPAP1 expression**

To find the optimum concentration for induction, a series of IPTG concentrations was tested with $proPAP1_{I,\ Xa}$ (A) and $proPAP1_{WT}$ (B) constructs. 20 µL supernatant from medium after removal of bacteria was loaded into each well The presence of the proPAP1 protein was confirmed with immunoblot analysis using antibody to proPAP1. (C) Two induction times were chosen to test the expression level of secreted $proPAP1_I$. Immunoblot (left) and coomassie staining (right) were employed. ***** Indicates the position of $proPAP1_I$.

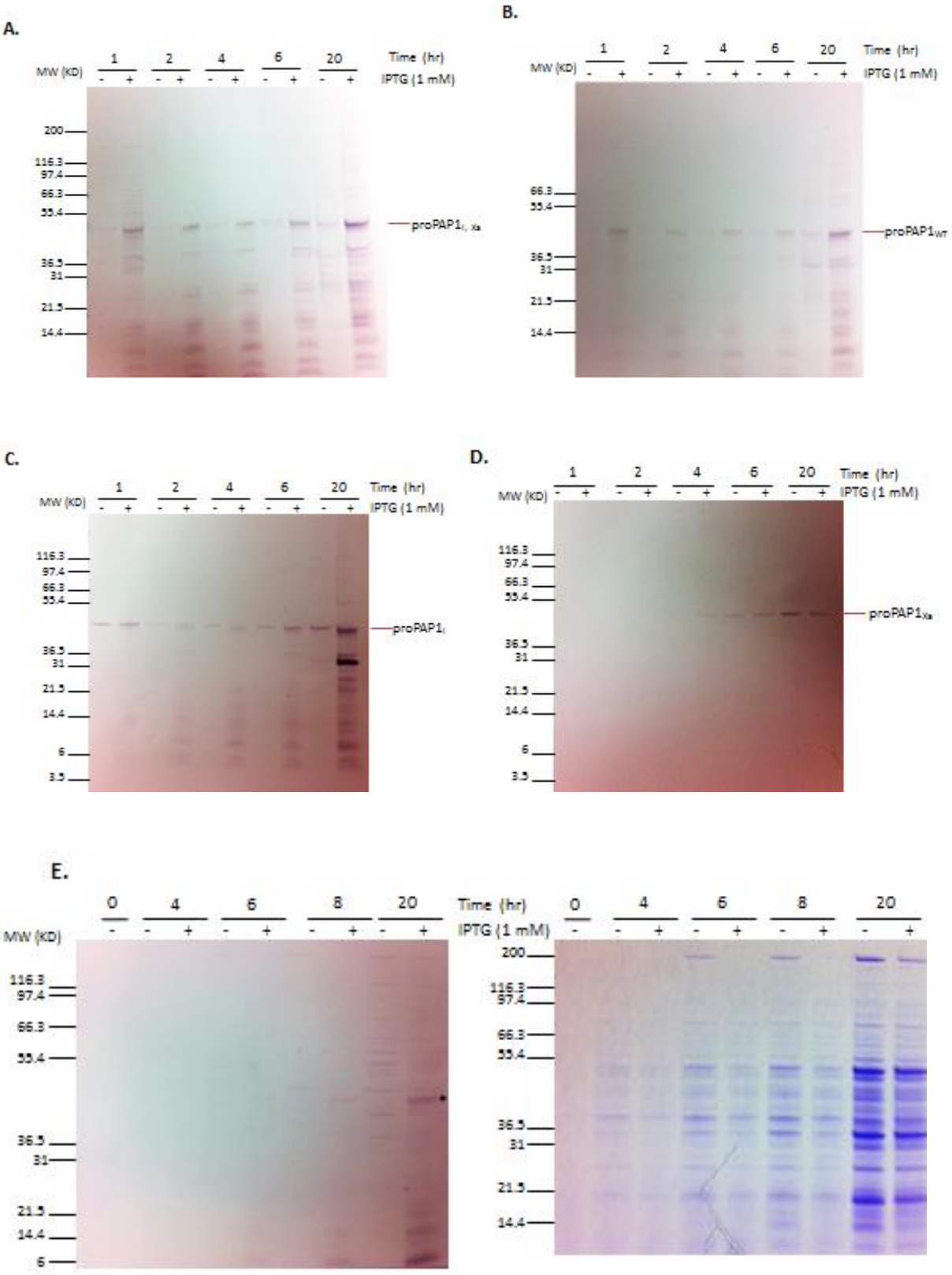

A.
MW (KD)
1
2
4
6
20
Time (hr)
IPTG (1 mM)
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
proPAP1i, Xa
B.
MW (KD)
1
2
4
6
20
Time (hr)
IPTG (1 mM)
66.3
55.4
36.5
31
21.5
14.4
proPAP1WT
C.
MW (KD)
1
2
4
6
20
Time (hr)
IPTG (1 mM)
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
3.5
proPAP1i
D.
MW (KD)
1
2
4
6
20
Time (hr)
IPTG (1 mM)
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
3.5
proPAP1Xa
E.
MW (KD)
0
4
6
8
20
Time (hr)
IPTG (1 mM)
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
200

**Figure 6 IPTG induction time course**

To find the optimum time for induction, an induction time course (0, 2, 4, 6 h and 20h) was performed with proPAP1$_{I, Xa}$ (A), proPAP1$_{WT}$ (B), proPAP1$_{I}$ (C), proPAP1$_{Xa}$ (D) using 1 mM IPTG. The induction of proPAP1 proteins was confirmed by immunoblot analysis, and effect of IPTG induction was tested by comparing plus or minus IPTG at the same time point. (E) A time course with secreted proPAP1$_{I}$ was performed and the supernatant was analyzed by immunoblot (left) analysis and coomassie staining (right).

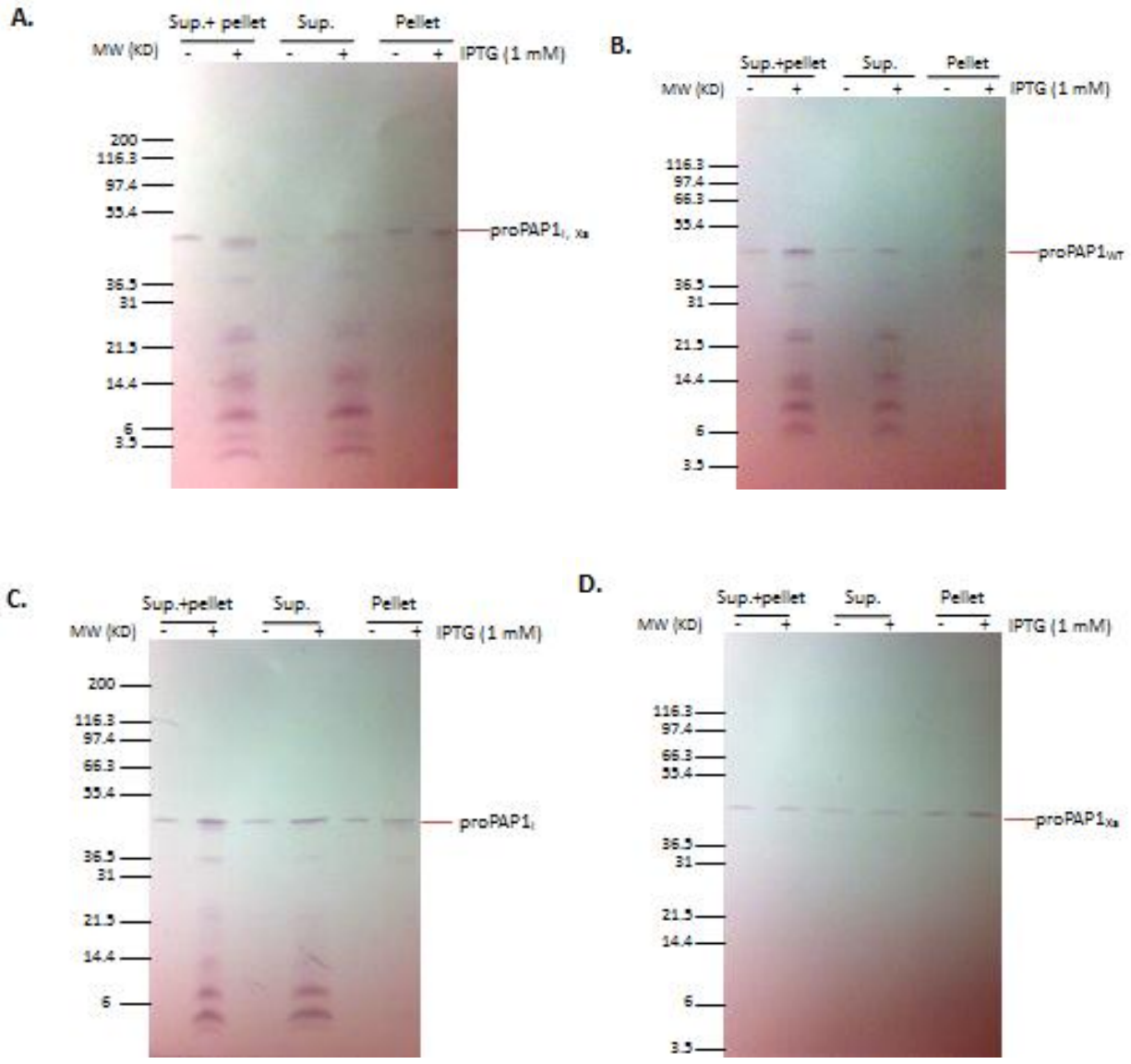



**Figure 7 Secretion of proPAP1 proteins**

The appearance of the proteins in the whole cell, supernatant, and cell lysate pellet were compared at 6h induction for $proPAP1_{I,\ Xa}$ (A), $proPAP1_{WT}$ (B), $proPAP1_{I}$ (C), $proPAP1_{Xa}$ (D), proPAP1 was detected by immunoblotting.

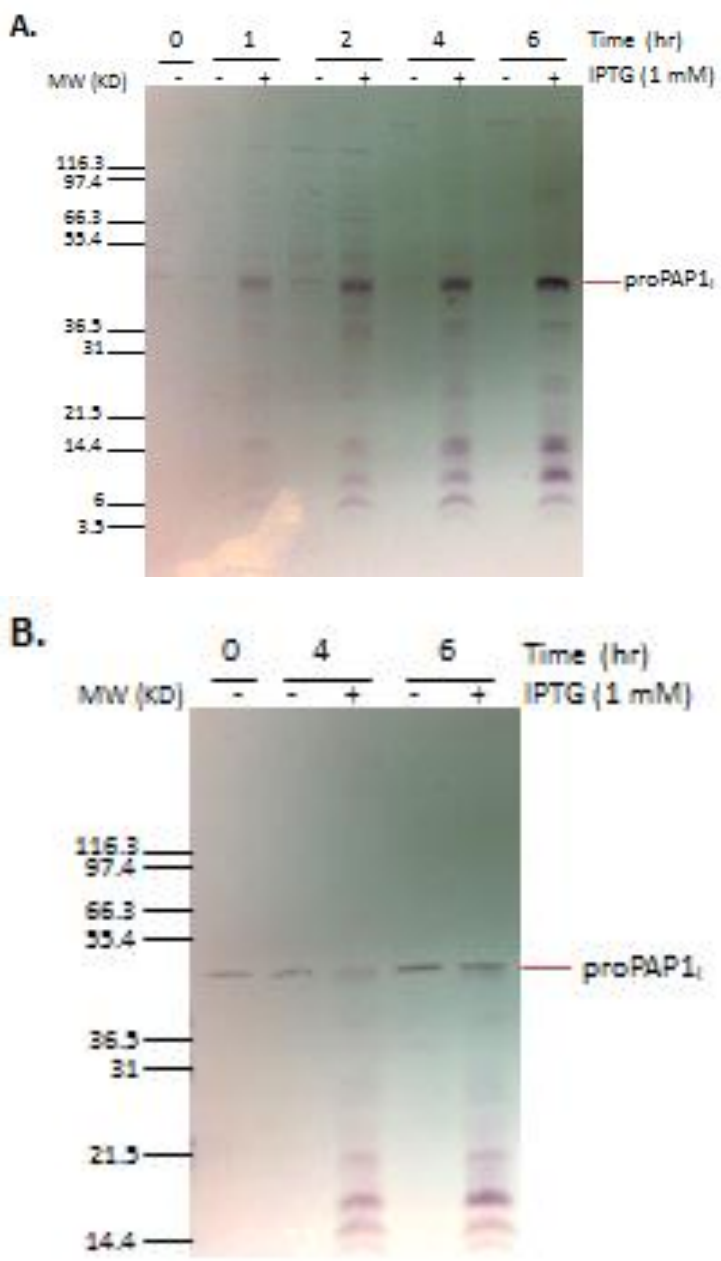



**Figure 8 Comparison of growth media for IPTG induction**

The expression level of proPAP1$_I$ was compared in 50 ml 2xYT(A) and 50 ml SOC(B) medium. ProPAP1$_I$ was induced by IPTG. proPAP1$_I$ was detected by immunoblotting using proPAP1 antibody.

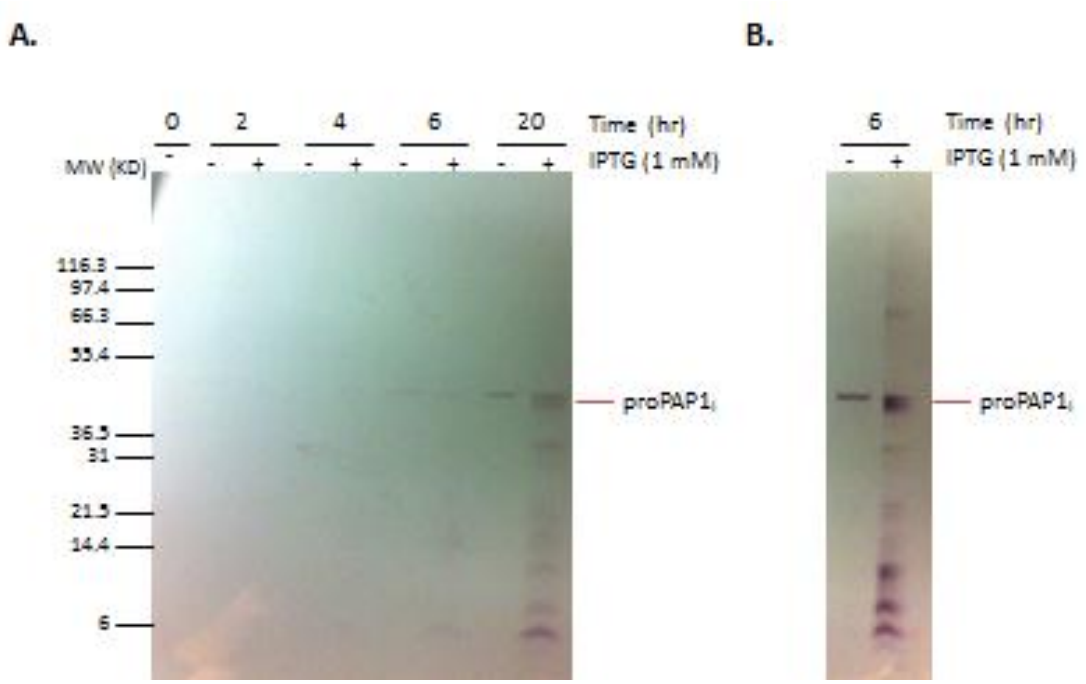



**Figure 9 Comparison of temperature for IPTG induction**

The expression level of proPAP1$_I$ was compared at 16°C (A) and 37°C (B). ProPAP1$_I$ was induced by IPTG in 2xYT at 16°C or 37°C. proPAP1$_I$ was detected by immunoblotting using proPAP1 antibody.

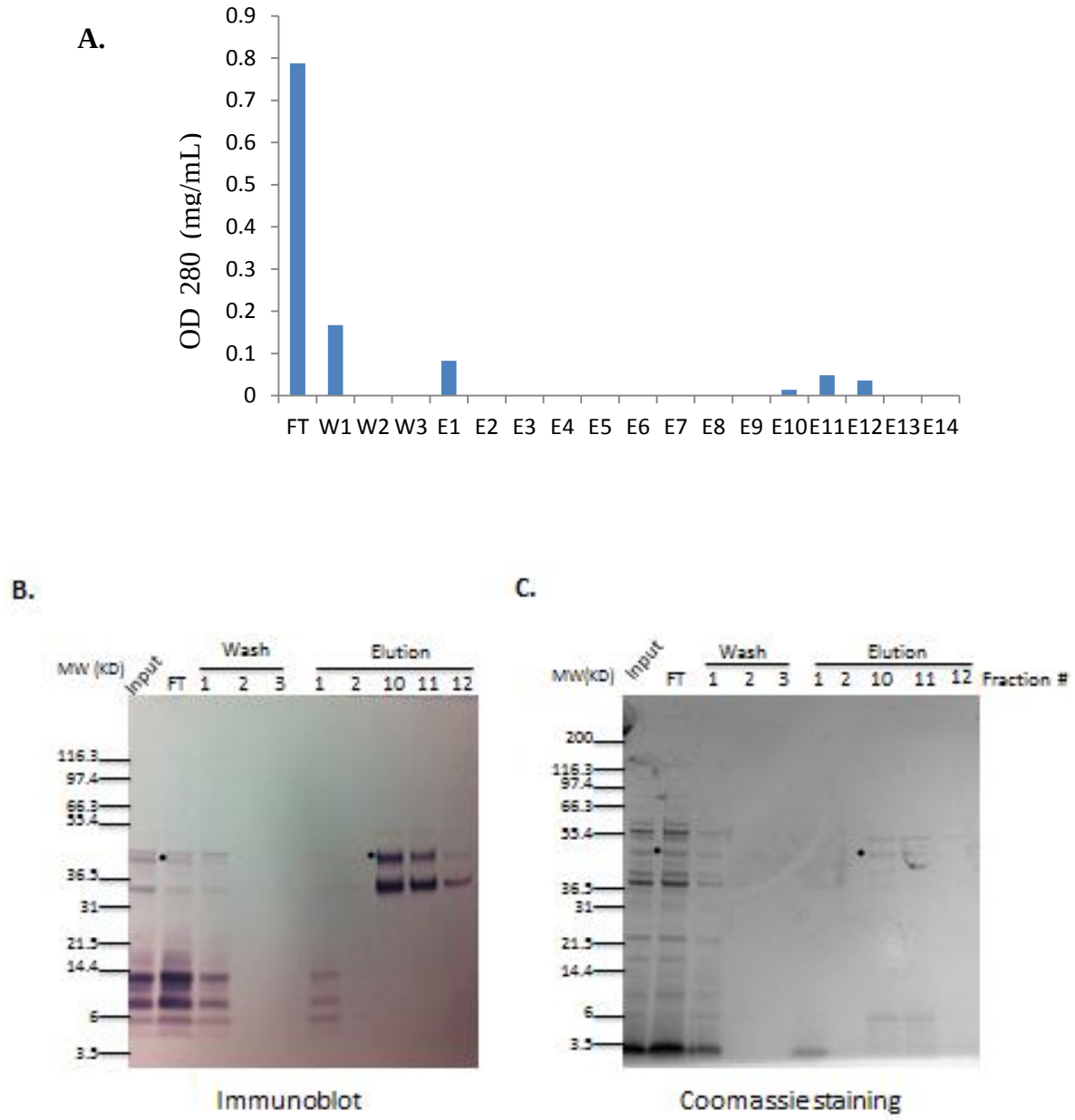
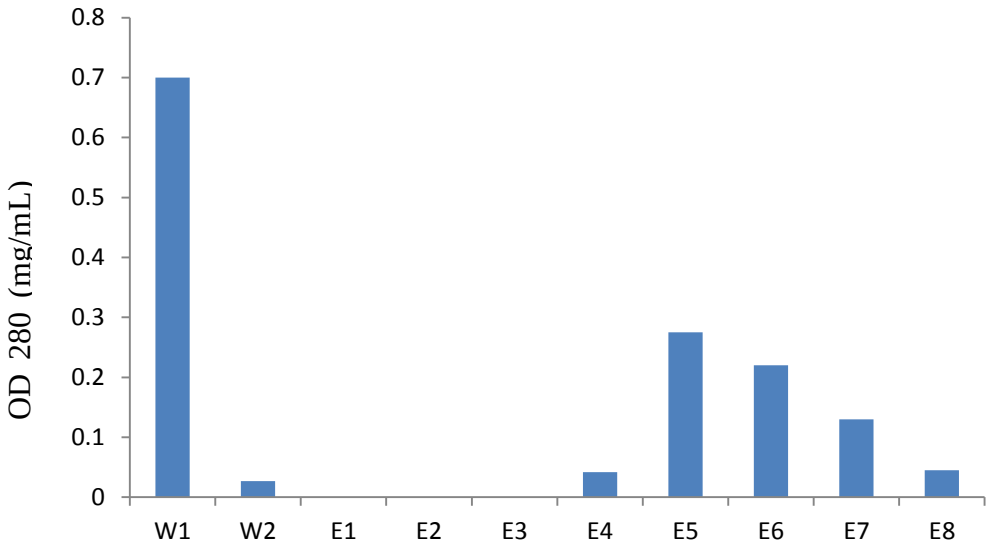
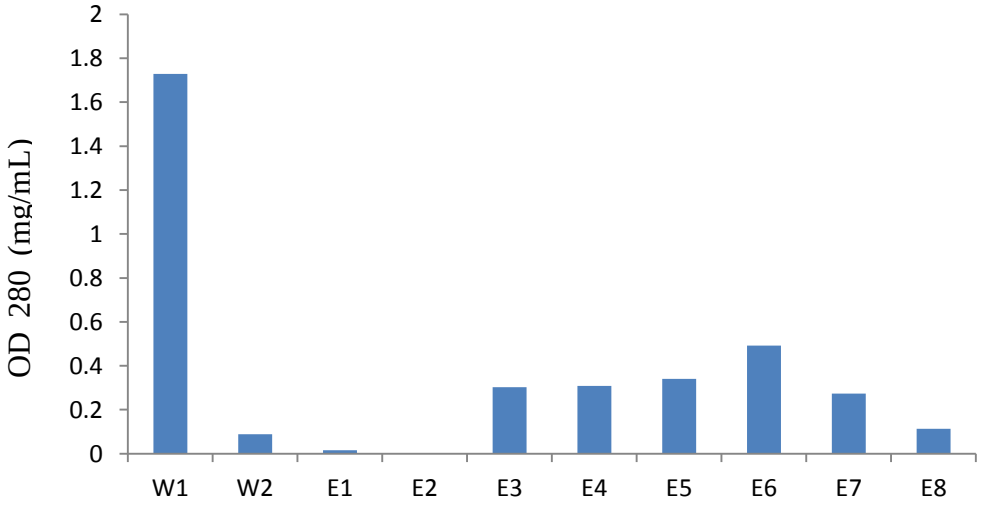



**Figure 10 Purification of proPAP1$_I$ protein**

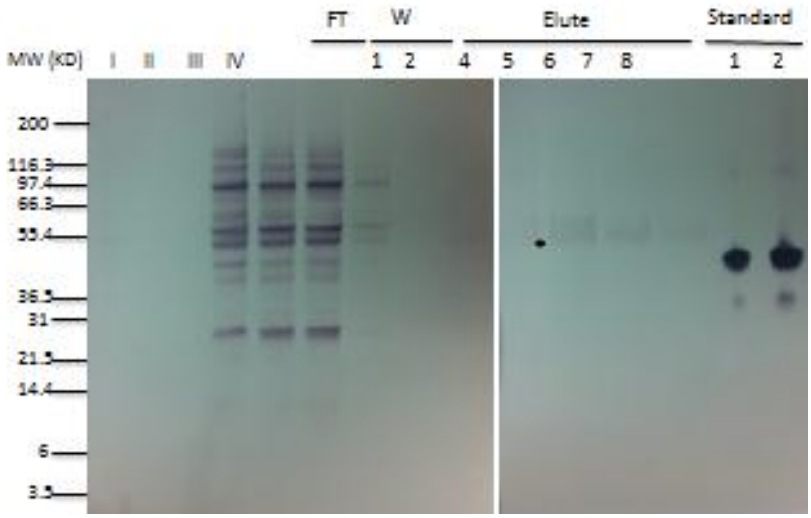
After dialysis, the proteins from 50% $(NH4)_2SO_4$ precipitation was subjected to Ni-NTA affinity chromatography. (A) $OD_{280}$ value of flow through (FT), wash (W), elution fractions (E), Elution was performed with 200 µL each from 1-10 fractions, 500 µL each from 11-14 fractions. (B &C) Analysis of proPAP1I during Ni-NTA affinity purification by immunoblot (B) and Coomassie blue staining (C) after SDS-PAGE separation of fractions. The full length of proPAP1$_I$ protein was indicated by *.

**A.**

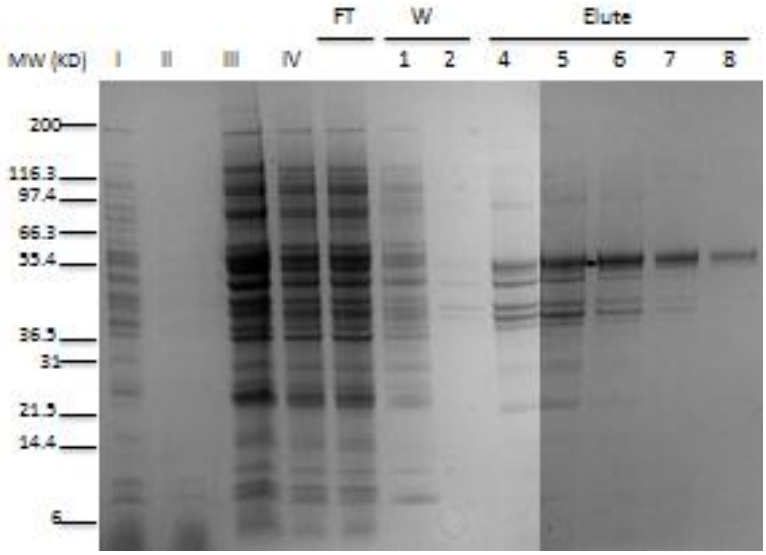

OD 280 (mg/mL)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
W1
W2
E1
E2
E3
E4
E5
E6
E7
E8


OD 280 (mg/mL)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
W1
W2
E1
E2
E3
E4
E5
E6
E7
E8

B.


FT
W
Elute
Standard
MW (KD)
I
II
III
IV
1
2
4
5
6
7
8
1
2
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
3.5

C.


FT
W
Elute
MW (KD)
I
II
III
IV
1
2
4
5
6
7
8
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6

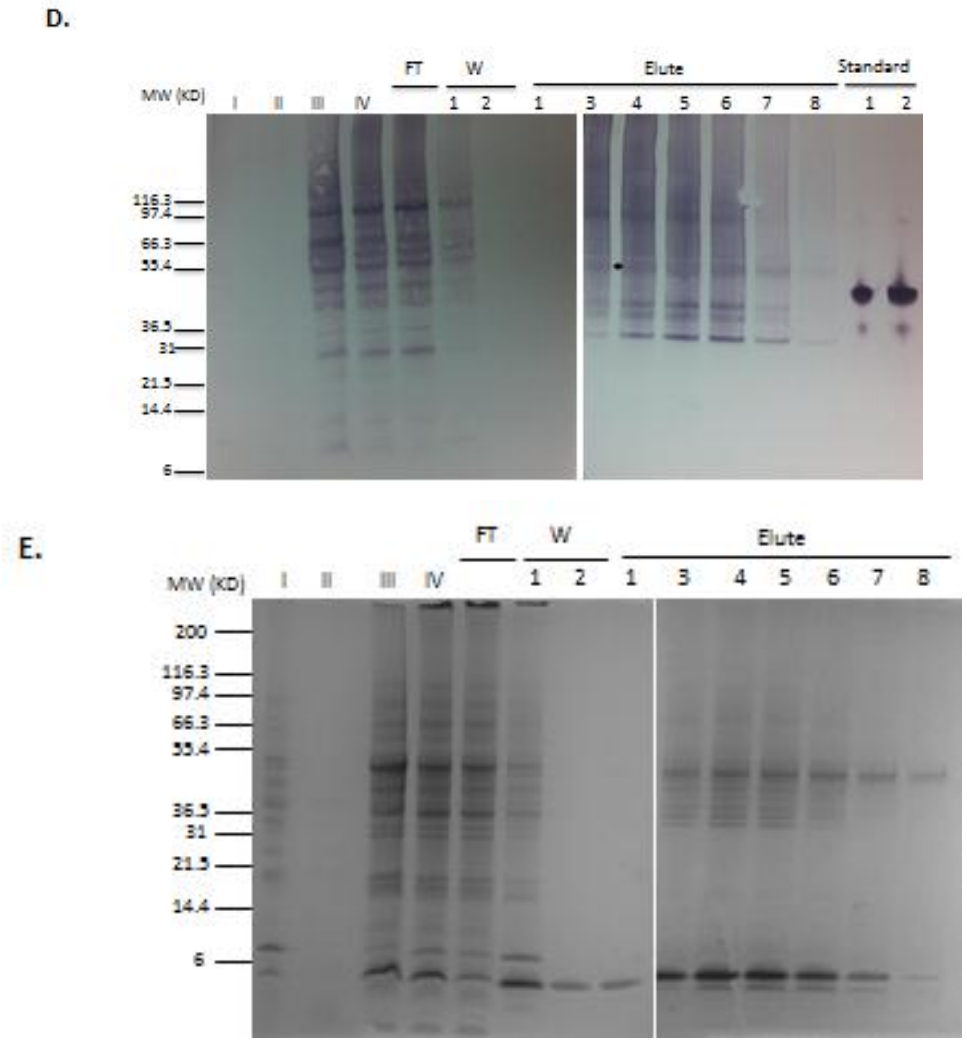



**Figure 11 Induction of proPAP1$_I$ protein in the presence of protease inhibitors**

IPTG induction was performed either in the absence (B&C) or presence (D&E) of protease inhibitors for 20 hr. After induction, the proteins in the supernatant were subjected to 50% saturated $(NH4)_2SO_4$ precipitation than dialyzed. After dialysis, the proPAP1$_I$ was purified by Ni-NTA affinity chromatography. (A) $OD_{280}$ value of flow through (FT), wash, elution fractions. Elution was performed with 500 μL each. (B) Western blot of induction in the absence of protease inhibitors. (C) Coomassie staining of induction in the absence of protease inhibitors. (D ) Western blot of induction in the presence of protease inhibitors. (E) Coomassie staining of induction in the presence of protease inhibitors. I: supernatant of culture; II: supernatant after 50% $(NH4)_2SO_4$ precipitation; III: pellet of 50% $(NH4)_2SO_4$ precipitation; IV: pellet of 50% $(NH4)_2SO_4$ precipitation after dialysis. 50 ng or 100 ng proPAP1$_{I,Xa}$ were used as standard. The full length of proPAP1$_I$ protein was indicated by *.

**Table 1 Summary of PCR primers and sequencing primers**

Subcloning primers:

| Primer name | Sequence |
|---|---|
| pHTPAP1-F | 5’- GGATCCATGGGCCAGTCATGTACAACACCA -3’ |
| pHTPAP1-R | 5’-GACGTCTCAATGGTGATGGTGATGATGCACATT GGTTGATCTGAG-3’ |

Sequencing vector primers:

| Primer name | Sequence |
|---|---|
| pHT43-VF | 5'GCTAACGGAAAAGGGAGC 3' |
| pHT43-VR | 5' CGTATGTTTCAACCATTTG3' |

Sequencing primers in the middle of proPAP1 CDS:

| Primer name | Sequence |
|---|---|
| proPAP1 CDS-F | 5' ACGATGTTTGTGCTGACCC3' |
| proPAP1 CDS-R | 5' GGGTCAGCACAAACATCGT3' |

**Table 2 Protocol for all reagents:**

1xTAE buffer

| |
|---|
| 40 mM Tris |
| 20 mM acetic acid |
| 1 mM EDTA |
| pH 8.0 |

SOC medium

| |
|---|
| 20 mM glucose |
| 20 g/L tryptone |
| 5 g/L yeast extract |
| 0.5 g/L NaCl |
| 2.5 mM KCl |
| 5 mM MgCl2 |
| 5 mM MgSO4 |
| pH 7.0 |

100 mL 10x S-base (Spizizen's salt)

| |
|---|
| 2 g $(NH4)_2SO_4$ |
| 14 g $K_2HPO_4$ |
| 6 g $KH_2PO_4$ |
| 1 g sodium citrate |
| After autoclave, 0.1 mL 1 M $MgSO_4$ |

100 mL HS medium

| |
|---|
| 10 mL 10x S-base |
| 2.5 mL 20% (w/v) glucose |
| 5 mL 0.1% (w/v) L-tryptophan |
| 1 mL 2% (w/v) casein |

| |
|---|
| 5 mL 10% (w/v) yeast extract |
| 10 mL 8% (w/v) arginine |
| 0.4% histidine |

Autoclave all components separately. Tryptophan solution: sterile filtration.

100 mL LS medium

| |
|---|
| 10 mL 10x S-base |
| 2.5 mL 20% (w/v) glucose |
| 0.5 mL 0.1% (w/v) L-tryptophan |
| 0.5 ml 2% (w/v) casein |
| 5 mL 2% (w/v) yeast extract (Difco) |
| 0.25 mL 1 M $MgCl_2$ |
| 0.05 mL 1 M $CaCl_2$ |

Autoclave all components separately. Tryptophan solution: sterile filtration.

0.1 M EGTA (100 mL):

| |
|---|
| 3.8 g EGTA |
| adjust pH to 7.2 using 10 N NaOH |
| add water to 100 mL, autoclave |

# References


Ambrus, G., Gál, P., Kojima, M., Szilágyi, K., Balczer, J., Antal, J., Gráf, L., Laich, A., Moffatt, B.E., Schwaeble, W., Sim, R.B., Závodszky, P. 2003. Natural substrates and inhibitors of mannan- binding lectin-associated serine protease-1 and -2: a study on recombinant catalytic fragments. J Immunol. 170, 1374-82.

An, C., Ishibashi, J., Ragan, E., Jiang, H.B., Kanost, M.R. 2009. Functions of *Manduca sexta* hemolymph proteinases HP6 and HP8 in Two Innate Immune Pathways. J. Biol. Chem. 284, 19716-19726.

An, C., Kanost, M.R. 2010. *Manduca sexta* serpin-5 regulates prophenoloxidase activation and the Toll signaling pathway by inhibiting hemolymph proteinase HP6. Insect Biochem. Mol. Biol. 40, 683-689.

An, C.J., Jiang, H.B., Kanost, M.R. 2010. Proteolytic activation and function of the cytokine Spätzle in the innate immune response of a lepidopteran insect, *Manduca sexta.* FEBS J. **277**, 148-162.

Brockmeier U, Wendorff M, Eggert T. 2006. Versatile expression and secretion vectors for *Bacillus subtilis*. Curr Microbiol. 52143-8.

Bron, S., Luxen, E. 1985. Segregational instability of pUB110-derived recombinant plasmids in *Bacillus subtilis*. Plasmid. 14, 235-244.

Cerenius, L., Söderhäll, K. 2004. The prophenoloxidase-activating system in invertebrates. Immunol. Rev. 198, 116-126.

Charalambidis, N.D., Foukas, L.C., Zervas, C.G., Marmaras, V.J. 1996. Hemocyte surface phenoloxidase (PO) and immune response to lipopolysaccharide (LPS) in *Ceratitis capitata*. Insect Biochem. Mol. Biol. 26, 867-74.

Christen, J.M., Hiromasa, Y., An, C.J., Kanost, M.R. 2012. Identification of plasma proteinase complexes with serpin-3 in *Manduca sexta* Insect Biochem. Mol. Biol. 42, 946-955.

Dai, H., Hiromasa, Y., Takahashi, D., Vander Velde, D., Fabrick, J.A., Kanost, M.R., Krishnamoorthi, R. 2013. An initial event in the insect innate immune response: structural and biological studies of interactions between β-1,3-glucan and the N-terminal domain of β-1,3-glucan recognition protein. Biochemistry. 52, 161-70.

De Gregorio, E., Han, S.J., Lee, W.J., Baek, M.J., Osaki, T., Kawabata, S., Lee, B.L., Iwanaga. S., Lemaitre, B., Brey, P.T. 2002. An immune-responsive Serpin regulates the melanization cascade in Drosophila. Dev. Cell. 3, 581-92.

Dunn, P.E. 1990. Humoral Immunity in Insects. Bioscience. 40, 738-744.

Doucet, A., Overall, C.M. 2008. Protease proteomics: revealing protease in vivo functions using systems biology approaches. Mol Aspects Med. 29, 339-58.

Ellen, C.W., Mercer, A.R. 2012. Modulatory actions of dopamine and serotonin on insect antennal lobe neurons: insights from studies in vitro. J. Mol. Histol. 43, 401-4.

Estefan, M.M., María, H. M. S., Raquel, C. B. 2008. Expression of a *Haemonchus contortus* cysteine protease in the baculovirus system. Electronic J of Biotechnol. 11, 1-7.

Georgiou, G., Segatori, L. 2005. Preparative expression of secreted proteins in bacteria: status report and future prospects. Curr. Opin. Biotechnol. 16, 538-545.

Gillespie, J.P., Kanost, M.R., Trenczek, T. 1997. Biological mediators of insect immunity. Annu. Rev. Entomol. 42, 611-43.

Gorman, M.J., Wang, Y., Jiang, H.B., Kanost, M.R. 2007. *Manduca sexta* hemolymph proteinase 21 activates prophenoloxidase-activating proteinase 3 in an insect innate immune response proteinase cascade. J. Biol. Chem. 282, 11742-11749.

Gruss, A., Ehrlich, S.D. 1988. Insertion of foreign DNA into plasmids from gram-positive bacteria induces formation of high-molecular-weight plasmid multimers. J. Bacteriol. 107, 1183-1190.

Gupta, S., Wang, Y., Jiang, H.B. 2005a. *Manduca sexta* prophenoloxidase (proPO) activation requires proPO-activating proteinase (PAP) and serine proteinase homologs (SPHs) simultaneously. Insect Biochem. Mol. Biol. 35, 241-8.

Gupta, S., Wang, Y., Jiang, H.B. 2005b. Purification and characterization of *Manduca sexta*prophenoloxidase-activating proteinase-1, an enzyme involved in insect immune responses. Protein Expr and Purif. 39, 261-268.

Halfon, M.S., Keshishian, H. 1998. The Toll pathway is required in the epidermis for muscle development in the *Drosophila* embryo. Dev. Biol. 199, 164-74.

Harwood, C, R., Cranenburgh, R. 2008. *Bacillus* protein secretion: an unfolding story. Trends Microbiol. 16, 73-9.

Hoffmann, J. 2011. The host defense of insects: A paradigm for innate immunity. Nobel Lecture. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2011/hoffman-lecture.pdf

Horohov, D.W., Dunn, P.E. 1983. Phagocytosis and nodule formation by hemocytes of *Manduca sexta* larvae following injection of *Pseudomonas aeruginosa.* J. Invertebr. Pathol. 41, 203-213.

Huang, R.D., Lu, Z.L., Dai, H.E., Velde, D.V., Prakash, O., Jiang, H.B. 2007. The solution structure of clip domains from *Manduca sexta* prophenoloxidase activating proteinase-2. Biochemistry. 46, 11431-11439.

Jäger, W., Schäfer, A., Pühler, A., Labes, G., Wohlleben, W. 1992. Expression of the Bacillus subtilis sacB gene leads to sucrose sensitivity in the gram-positive bacterium Corynebacterium glutamicum but not in Streptomyces lividans. J. Bacteriol.174, 5462-5465.

Jang, I.H., Chosa, N., Kim, S.H., Nam, H.J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., Brey, P.T., Lee. W.J. 2006. A Spatzle-Processing Enzyme Required for Toll Signaling Activation in *Drosophila* Innate Immunity. Dev Cell. 10, 45-55.

Ji, C.Y., Wang, Y., Ross, J., Jiang, H.B. 2003. Expression and in vitro activation of *Manduca sexta* prophenoloxidase-activating proteinase-2 precursor (proPAP-2) from baculovirus-infected insect cells. Protein Expr. Purif. 29, 235-243.

Jiang, H.B. 2008. The biochemical basis of antimicrobial responses in *Manduca sexta*. Insect Science. 15, 53-66.

Jiang, H.B., Kanost, M.R. 2000. The clip-domain family of serine proteinases in arthropods. Insect Biochem. Mol. Biol. 30, 95-105.

Jiang, H. B., Ma, C., Lu, Z.Q., Kanost, M.R. 2004. Beta-1,3-glucan recognition protein-2 (betaGRP-2) from *Manduca sexta*; an acute-phase protein that binds beta-1,3-glucan and lipoteichoic acid to aggregate fungi and bacteria and stimulate prophenoloxidase activation. Insect Biochem. Mol. Biol. 34, 89-100.

Jiang, H.B., Vilcinskas, A., Kanost, M.R. 2010. Immunity in Lepidopteran Insects. Adv. Exp. Med. Biol. 708, 181-204.

Jiang, H.B., Wang, Y., Gu, Y., Guo, X., Zou, Z., Scholz, F., Trenczek, T.E., Kanost, M.R. 2005. Molecular identification of a bevy of serine proteinases in *Manduca sexta* hemolymph. Insect Biochem. Mol. Biol. 35, 931-43.

Jiang, H.B., Wang, Y., Kanost, M.R. 1998. Pro-phenol oxidase activating proteinase from an insect, *Manduca sexta:* A bacteria-inducible protein similar to *Drosophila* easter. Proc. Natl. Acad. Sci. 95, 12220-12225.

Jiang, H., Wang, Y., Ma, C., Kanost, M.R. 1997. Subunit composition of pro-phenol oxidase from *Manduca sexta*: molecular cloning of subunit ProPO-P1. Insect Biochem. Mol. Biol. 27, 835-50.

Jiang, H.B., Wang, Y., Yu, X.Q., Kanost, M.R. 2003a. Prophenoloxidase-activating proteinase-2 from hemolymph of *Manduca sexta.* A bacteria- inducible serine proteinase containing two clip domains. J. Biol. Chem. 278, 3552-3561.

Jiang, H.B., Wang, Y., Yu, X.Q., Zhu, Y., Kanost, M.R. 2003b. Prophenoloxidase-activating proteinase-3 (PAP-3) from *Manduca sexta* hemolymph: a clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs. Insect Biochem. Mol. Biol. 33, 1049-60.

Kan, H., Kim, C.H., Kwon, H.M., Park, J.W., Roh, K.B., Lee, H., Park, B.J., Zhang, R., Zhang, J., Söderhäll, K., Ha, N.C., Lee, B.L. 2008. Molecular control of phenoloxidase-induced melanin synthesis in an insect. J. Biol. Chem. 283, 25316-25323.

Kanost, M.R., Gorman, M.J. 2008. Phenoloxidases in insect immunity. In Insect Immunology (N. Beckage, ed), Academic Press/Elsevier, San Diego, pp. 69-96.

Kanost, M.R., Jiang, H.B., Yu, X. 2004. Innate immune responses of a lepidopteran insect, *Manduca sexta*. Immunol Rev. 198, 97-105.

Khan, A.R., James, M.N. 1998. Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. Protein Sci. 7, 815-36.

Kim, C.H., Kim, S.J., Kan H, Kwon, H.M., Roh, K.B., Jiang, R., Yang, Y., Park, J.W., Lee, H.H., Ha, N.C., Kang, H.J., Nonaka, M., Söderhäll, K., Lee, B.L. 2008. A three-step proteolytic cascade mediates the activation of the peptidoglycan-induced toll pathway in an insect. J Biol Chem. 283,7599-607.

Kloppenburg, P., Mercer, A.R. 2008. Serotonin modulation of moth central olfactory neurons. Annu. Rev. Entomol. 53, 179-90.

Korner, A., Pawelek, J. 1982. Mammalian tyrosinase catalyzes three reactions in the biosynthesis of melanin. Science. 217, 1163-1165.

Lacroix, M., Ebel, C., Kardos, J., Dobó, J., Gál, P., Závodszky, P., Arlaud, G.J., Thielens, N.M. 2001. Assembly and enzymatic properties of the catalytic domain of human complement protease C1r. J Biol Chem. 276, 36233-40.

Lange, H., Bronson, L. 1981. Insect pests of tomatoes. Annu. Rev. Entomol. 263, 45-371.

Lavine, M.D., Strand, M.R. 2002. Insect hemocytes and their role in immunity. Insect Biochem. Mol. Biol. 32, 1295-1309.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J., Hoffmann, J.A. 1996. The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. Cell. 86, 973-983.

Liu, H., Jiravanichpaisal, P., Cerenius, L., Lee, B.L., Söderhäll, I., Söderhäll, K. 2007. Phenoloxidase is an important component of the defense against *Aeromonas hydrophila* infection in a crustacean, *Pacifastacus leniusculus*. J. Biol. Chem. 282, 33593-33598.

Liu, L., Liu, Y., Shin, H.D., Chen, R.R., Wang, N.S., Li, J., Du, G., Chen, J. 2013. Developing *Bacillus* spp. as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology. Appl. Microbiol. Biotechnol. 97, 6113-27.

Lopez, O.C., Overall, C.M. 2002. Protease degradomics: a new challenge for proteomics. Nat. Rev. Mol. Cell Biol. 3, 509-519.

Medzhitov, R., Brenton-Hurlburt, P., Janeway, C.A. 1997. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. Nature. 388, 394-396.

Moszer, I., Rocha, E.P., Danchin, A. 1999. Codon usage and lateral gene transfer in *Bacillus subtilis*. Curr. Opin. Microbiol. 2, 524-8.

Niwa, T., Ying, B.W., Saito, K., Jin, W.Z., Takada, S., Ueda, T., Taguchi, H. 2009. Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. Proc. Natl. Acad. Sci. 106, 4201-4206.

Piao, S.F., Kim, D., Park, J.W., Lee, B.L., Ha, N.C. 2005. Overexpression and preliminary X-ray crystallographic analysis of prophenoloxidase activating factor II, a clip domain family of serine proteases. Biochimica et Biophysica Acta. 1752, 103-106.

Priest, F.G. 1977. Extracellular enzyme synthesis in the genus *Bacillus*. Bacteriol. Rev. 41, 711-753.

Roh, K. B., Kim, C. H., Lee, H., Kwon, H. M., Park, J. W., Ryu, J. H., Kurokawa, K., Ha, N. C., Lee, W. J., Lemaitre, B., Söderhäll, K., Lee, B. L. 2009. Proteolytic cascade for the activation of the insect toll pathway induced by the fungal cell wall component. J. Biol. Chem. 284,19474-81.

Schallmey, M., Singh, A., Ward, O.P. 2004. Developments in the use of Bacillus species for industrial production. Can. J. Microbiol. 50,1-17.

Seybert, A., Ziebuhr, J., Siddell, S. G. 1997. Expression and characterization of a recombinant murine coronavirus 3C-like proteinase. J. Gen. Virol. 78, 71–75.

Singh, P. K., Ramachandran, G., Durán-Alcalde, L., Alonso, C., Wu, L. J., Meijer, W. J. 2012. Inhibition of *Bacillus subtilis* natural competence by a native, conjugative plasmid-encoded comK repressor protein. Environ. Microbiol. 14, 2812-25.

Song, C. W., Cui, Z. X., Liu, Y., Li, Q.Q., Li, X. H., Shi, G. H., Wang, C. L. 2013. Characterization and functional analysis of serine proteinase and serine proteinase homologue from the swimming crab *Portunus trituberculatus*. Fish. Shellfish. Immunol. 35, 231-239.

Sørensen, H. P., Mortensen, K. K. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J. Biotechnol. 115, 11-128.

Sorokulova, I. B., Pinchuk, I. V., Denayrolles, M., Osipova, I. G., Huang, J. M., Cutting, S. M., Urdaci, M. C. 2008. The safety of two *Bacillus* probiotic strains for human use. Dig. Dis. Sci. 53, 954-63.

Sugumaran, M. 2002. Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. Pigment. Cell. Res.15, 2-9.

Sun, Q.Y., Ding, L.W., He, L. L., Sun, Y. B., Shao, J. L., Luo, M., Xu, Z. F. 2009. Culture of *Escherichia coli* in SOC medium improves the cloning efficiency of toxic protein genes. Anal. Biochem. 394, 144-146.

Tian, S. F., LeMosy, E. K. 2008. Mutagenesis of the cysteine-rich clip domain in the *Drosophila* patterning protease, Snake. Arch. Biochem. Biophys. 475, 169-174.

Tjalsma, H., Antelmann, H., Jongbloed, J. D., Braun, P. G., Darmon, E., Dorenbos, R., Dubois, J. Y., Westers, H., Zanen, G., Quax, W. J., Kuipers, O. P., Bron, S., Hecker, M., Van, Dijl. J. M. 2004. Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. Microbiol. Mol. Biol. Rev. 68, 207-33.

Vallejo, L. F., Rinas, U. 2004. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb Cell. Fact. 3, 3-11.

Van den Berg, B., Wain, R., Dobson, C. M., Ellis, R. J. 2000. Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell. EMBO. J. 19, 3870-3875.

Van, Dijl, J,M., Hecker., M. 2013. *Bacillus subtilis*: from soil bacterium to super secreting cell factory. Microb Cell Fact. 14, 1-6

Wang, Y., Jiang, H. B. 2004a. Purification and characterization of *Manduca sexta* serpin-6: a serine proteinase inhibitor that selectively inhibits prophenoloxidase-activating proteinase-3. Insect Biochem. Mol. Biol. 34, 387-395.

Wang, Y., Jiang, H. B. 2004b. Prophenoloxidase (proPO) activation in Manduca sexta: an analysis of molecular interactions among proPO, proPO-activating proteinase-3, and a cofactor. Insect Biochem. Mol. Biol. 34, 731-742.

Wang, Y., Jiang, H. 2007. Reconstitution of a branch of the *Manduca sexta* prophenoloxidase activation cascade in vitro: snake-like hemolymph proteinase 21 (HP21) cleaved by HP14 activates prophenoloxidase-activating proteinase-2 precursor. Insect Biochem. Mol. Biol. 37, 1015-25.

Wang, Y., Jiang, H. 2008. A positive feedback mechanism in the *Manduca sexta* prophenoloxidase activation system. Insect Biochem Mol Biol. 38, 763-9.

Wang, Y., Jiang, H.B., Kanost, M.R. 2001. Expression and purification of *Manduca sexta* prophenoloxidase-activating proteinase precursor (proPAP) from baculovirus-infected insect cells. Protein Expr. Purif. 23, 328-337.

Wang, Y., Sumathipala, N., Rayaprolu, S., Jiang, H. 2011. Recognition of microbial molecular patterns and stimulation of prophenoloxidase activation by a β-1,3-glucanase-related protein in *Manduca sexta* larval plasma. Insect Biochem. Mol. Biol. 41, 322-31.

Weeks, J. C., Jacobs, G. A., Pierce, J. T., Sandstrom, D. J., Streichert, L. C., Trimmer, B. A., Wiel, D. E., Wood, E. R. 1997. Neural mechanisms of behavioral plasticity: metamorphosis and learning in *Manduca sexta*. Brain Behav. Evol. 50 Suppl. 1, 69-80.

Wenzel, M., Müller, A., Siemann-Herzberg, M., Altenbuchner, J. 2011. Self-inducible *Bacillus subtilis* expression system for reliable and inexpensive protein production by high-cell-density fermentation. Appl. Environ. Microbiol. 77, 6419-25.

Willott, E., Trenczek, T., Thrower, L.W., Kanost, M.R. 1994. Immunochemical identification of insect hemocyte populations: monoclonal antibodies distinguish four major hemocyte types in *Manduca sexta*. Eur. J. Cell. Biol. 65, 417-423.

Wu, X.C., Lee, W., Tran, L., Wong, S.L. 1991. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. J. Bacteriol. 173, 4952-4958.

Yasukochi, Y., Tanaka-Okuyama, M., Shibata, F., Yoshido, A., Marec, F., Wu, C., Zhang, H., Goldsmith, M.R., Sahara, K.. 2009. Extensive conserved synteny of genes between the karyotypes of *Manduca sexta* and *Bombyx mori* revealed byBAC-FISH mapping. PLoS One. 4, 7465-7487.

Yu, X.Q., Jiang, H. B., Wang, Y., Kanost, M.R. 2003. Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. Insect Biochem. Mol. Biol. 33, 197-208.

Yu, X. Q., Zhu, Y. F., Ma, C., Fabrick, J. A., Kanost, M. R. 2002. Pattern recognition proteins in *Manduca sexta* plasma. Insect Biochem. Mol. Biol. 32, 1287-93.

Zhang, S., Gunaratna, RT., Zhang, X., Najar, F., Wang, Y., Roe, B., Jiang, H. 2011. Pyrosequencing-based expression profiling and identification of differentially regulated genes from *Manduca sexta*, a lepidopteran model insect. Insect Biochem Mol Biol. 41, 733-46.

Zhang, X. Z., Cui, Z. L., Hong, Q., Li, S. P. 2005. High-Level expression and secretion of methyl parathion hydrolase in *Bacillus subtilis* WB800. Appl. Environ. Microbiol. 71, 4101-4103.

Zhao, P., Li, J., Wang, Y., Jiang, H. 2007. Broad-spectrum antimicrobial activity of the reactive compounds generated in vitro by *Manduca sexta* phenoloxidase. Insect Biochem. Mol. Biol. 37, 952-9.

Zhu, Y. F., Wang, Y., Gorman, M., Jiang, H. B., Kanost, M. R. 2003. *Manduca sexta* serpin-3 regulates prophenoloxidase activation in response to infection by inhibiting prophenoloxidase-activating proteinases. J. Biol. Chem. 278, 46556-64.

Zou, Z., Jiang, H.B. 2005. *Manduca sexta* serpin-6 regulates immune serine proteinases PAP-3 and HP8: cDNA cloning, protein expression, inhibition kinetics, and function elucidation. J. Biol. Chem. 280, 14341-14348.

Zou, Z., Wang, Y., Jiang, H. B. 2005. *Manduca sexta* prophenoloxidase activating proteinase-1 (PAP-1) gene: Organization, expression, and regulation by immune and hormonal signals. Insect Biochem Mol Biol. 35,627-636.