Spring 2017

Biochemical and Enzymatic Studies of the Midgut Serine Protease AaSPVII from the Aedes aegypti Mosquito

James ThienToan Nguyen
San Jose State University

BIOCHEMICAL AND ENZYMATIC STUDIES OF THE MIDGUT SERINE
PROTEASE AASPVII FROM THE *AEDES AEGYPTI* MOSQUITO

A Thesis

Presented to

The Faculty of the Department of Chemistry

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

James Nguyen

May 2017

The Designated Thesis Committee Approves the Thesis Titled

BIOCHEMICAL AND ENZYMATIC STUDIES OF THE MIDGUT SERINE
PROTEASE AASPVII FROM THE *AEDES AEGYPTI* MOSQUITO

by

James Nguyen

APPROVED FOR THE DEPARTMENT OF CHEMISTRY

SAN JOSÉ STATE UNIVERSITY

May 2017

Prof. Alberto A. Rascón, Jr.	Department of Chemistry
Prof. Laura Miller Conrad	Department of Chemistry
Prof. Daryl Eggers	Department of Chemistry

ABSTRACT

BIOCHEMICAL AND ENZYMATIC STUDIES OF THE MIDGUT SERINE PROTEASE AASPVII FROM THE *AEDES AEGYPTI* MOSQUITO

by James Nguyen

The *Aedes aegpyti* mosquito is an efficient biological vector of four known neglected tropical arboviruses: the yellow fever, dengue, chikungunya, and Zika viruses. Since the female mosquito can acquire multiple blood meals to complete the gonotrophic cycle, she can potentially infect multiple hosts. With the reemergence of these viruses, especially due to global warming, it is crucial to slow the spread of the mosquito and the blood-borne diseases she transmits. The *Ae. aegypti* mosquito relies heavily on midgut serine proteases that digest blood meal proteins to provide nutrients required for the egg laying process. The focus of this work is on a late phase midgut serine protease, AaSPVII. AaSPVII is of interest due to its potential difference in specificity and functional role in the digestion process, despite its sequence similarity to another late phase protease known as AaSPVI. By understanding the important role these midgut serine proteases play, we could potentially inhibit these proteases and disrupt fecundity. Currently, two different expression constructs were produced, one with an N-terminal His_6-tag and the other with a C-terminal His_6-tag. Both constructs have been recombinantly expressed using bacteria, but only the C-terminally His_6-tagged AaSPVII has been purified and activated successfully. This enzyme has comparable steady-state kinetic parameters to previously published results, suggesting that we have optimized soluble bacterial expression.

ACKNOWLEDGMENTS

I would also like to thank Department Chair Karen Singmaster, Dr. Elizabeth Pollum, Dr. Roy Okuda, Dr. Bradley Stone, Dr. Maria Matyska, and Associate Dean Marc d'Alarcao for all being part of my journey by supporting and cracking jokes with me.

Big shout out to Brian Huynh and Dan Dang for always keeping in touch and trying to set up hangouts! Special thanks to Jimmy Nguyen, Eliza Vien, Frank Nguyen, Kevin "The Flash" Tran, Tan Tran, Josh Sun, Simon Du, Alice Nguyen, Mandeep Banwait, Chi Nghiem, and Jeff Nguyen for just being awesome! Also, the crazy SJSU graduate students I got to befriend during my time here! Special thanks to Anthony Balistreri and Victor Sosa for always counting down the hours we have before critical deadlines, Alexis Sarabia and Joshua Topete for looking out for my well-being, M.A.L.L. for being Leroy, and B.O.B for being Rebecca! Cannot forget Lauren Blankenship and Diane Eilert's get-togethers and cooking!

And finally, a very special thanks to my family. Thank you to my brother George and sister Ngoc for being curious about my research and giving me their support; despite our harmless arguments. My mom and dad for their many sacrifices, unyielding support, and for putting up with my frequent moodiness. Without their love and support, I would not be the person I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

INTRODUCTION

Domestication and Spread of the *Aedes aegypti* Mosquito

The biological vector *Aedes aegypti* mosquito is classified with the subfamily

Culicinae of the order Diptera.[1,2] From a genetic standpoint, *Ae. aegypti* is the most

widely studied and medically important species of mosquitoes because of the ease of

transition from the field to a laboratory setting and the feasibility of controlled mating.[2] It

has become almost certain that the predominant subspecies of *Ae. aegypti* in Africa was

the *Ae. aegypti formosus*, while the subspecies that originated in Africa and dominated

the subtropical and tropical regions outside of Africa are the domesticated forms known

as *Ae. aegypti*.[3,4] *Ae. aegypti* spread to the New World between the fifteenth and

eighteenth centuries due to the increase in global trade[3,5], and their preference for human

domestication was most certainly due to human modifications of global landscapes

through movement and trade.[5] What may also have contributed to the domesticated

behavior of these mosquitoes was the extreme harsh landscape of the Sahara 4,000-6,000

years ago[6] and the fact that the only available source of fresh water that can maintain

breeding colonies was found in human settlements and on ships.[3,7] Based on the historical

data of the emergence of urbanized dengue (a viral infection transmitted by *Ae. aegypti*),

the mosquito may have spread to Asia during the late nineteenth century via cargo ships

and trade.[3]

Like all mosquitoes in the Culicidae family, the *Ae. aegypti* life cycle consists of a

metamorphic stage from larvae, pupae, to the adult mosquito. The interesting

characteristic about these mosquitoes is their eggs, which can be highly resilient and can survive through desiccation for several days to up to a year.[1,8] Their pupa can also survive through desiccation for several hours.[1] This particular attribute may have allowed for the mosquitoes to be transported to other parts of the world until they were able to find an adequate water source to hatch and, as expected, allowed for further spread of specific viruses and increased population due to their close proximity to human hosts.[1,7]

Blood Feeding Behavior – Spread of Vector-Borne Diseases

Typically, both male and female mosquitoes require carbohydrates, commonly from nectar, as the main energy source to maintain flight and other basic functions. Eventually, the female mosquito would be required to imbibe a blood meal to complete the gonotrophic cycle.[1,9] Unfortunately, because of the *Ae. aegypti* mosquito's close association to human settlements, this "new" domestic habitat led to a heavy reliance on human blood as their new primary energy source for egg production.[9] The proximity between human and mosquito allowed the mosquitoes to adapt and evolve, thus giving them a fitness advantage to respond to certain host cues (such as CO_2 release and host odors)[10] to detect the presence of humans, and as such locate blood and its nutritional quality.[9,11] This adaptation has also affected the female's sugar-feeding behavior, which unexpectedly reveals that the female mosquitoes seldom feed on nectar in their domestic environments, demonstrating that sugar-feeding was facultative.[9] This lack of carbohydrate feeding was observed as the mosquitoes would mate near humans and be found indoors more often, thereby reducing flight demands.[9] The anthropophilic behavior of the female mosquito has also suggested that she may survive solely on a blood meal by

increasing the amount she intakes.[12] This will allow for multiple events of blood feeding

and host-seeking, thereby increasing fecundity and population.[9,13] However, because of

this blood feeding behavior, the *Ae. aegypti* is an efficient biological vector for the yellow

fever (YFV), dengue (DENV), chikungunya (CHIKV), and Zika viruses (ZIKV), having

the ability to carry more than one virus and being able to infect a human host with

multiple viruses from a single bite.[14] Thus, the presence of the *Ae. aegypti* mosquito

becomes a critical issue for humid and urbanized human environments.[9,13]

One of most recent global concerns around the world is the mosquito-borne flavivirus

ZIKV.[15] ZIKV is the latest culprit on the long list of arboviruses originating in Africa.[15]

Like YFV and DENV in the family *Flaviviridae*, ZIKV are enveloped, single-stranded,

positive-sense, obligate intracellular RNA viruses. Although ZIKV, YFV, and DENV are

similar in genomic organizations, ZIKV is fundamentally different in host infectivity and

how the viruses are spread.[15,16] The virus was initially isolated in the Zika forest in 1947

from the serum of the rhesus monkey.[17] The virus was not detected and isolated from

humans until the 1950s, being reported in Uganda, Tanzania, and Nigeria.[18,19] Between

the 1950s and 1980s, ZIKV was predominately localized in Asian and African countries

and did not migrate to the Pacific Islands until the early 2000s. During this time, ZIKV

infection became more rampant, spreading to the Americas and now observed in

U.S.territories and the U.S.[15,20] There have been over 36,000 reported cases in the U.S.

territories and 200 locally acquired cases in the U.S.[20] Thus far, ZIKV has been the only

flavivirus to be sexually transmitted with 41 confirmed cases in the U.S. (as of February

2017), which was reported by the Centers for Disease Control and Prevention (CDC).[20]

One major issue with ZIKV is that infected people will typically not have any immediate onset of symptoms or will not show any symptoms at all. If symptoms do arise, they may last anywhere from several days to a week, with a likelihood of being protected from subsequent ZFV infections. However, the duration of the immunity is still unknown.[15] Common symptoms include fever, rash, joint pain, conjunctivitis, muscle pain, and headaches. The most striking congenital abnormality occurs within pregnant women and has been shown to have a direct association with microcephaly of the fetus.[15,21] This type of manifestation was linked to a 2015 Zika outbreak in Brazil where microcephaly was observed 20 times more than usual.[15] There have been rare cases of adults dying with ZIKV, but the exact contributions from the virus, if any, to these deaths are still relatively unknown.[22]

Another important and dangerous arthropod-borne viral infection that affects humans is DENV. It has been estimated that 2.5 billion people around the world are currently at risk of contracting dengue, with approximately 40% of those people living in tropical and sub-tropical urbanized areas, such as the Pacific and the Americas.[23,24] Unlike the other flaviviruses that are carried by the *Ae. aegypti* mosquito, DENV is known to have four serotypes (DENV-1, DENV-2, DENV-3, DENV-4)[23,24], and recently a fifth serotype (DENV-5) has been identified.[25] Each serotype shares approximately 65% genetic similarity and produces nearly undistinguishable syndromes and circulates in the same ecological niche.[23] Infection against one serotype will provide a long-lasting immunity against only that serotype, as a secondary heterotypic infection has been shown to lead to more severe symptoms.[23,24] The onset of this disease can lead from minor to severe fever,

headaches, muscle and joint pains, nausea, vomiting, swollen glands, and rashes. Secondary infection can increase the risk of dengue hemorrhagic fever, which can cause severe abdominal pains, bleeding gums, and blood in vomit.[23,24] Recent published data have also hypothesized that DENV may drive antibody-dependent enhancement (ADE) of infection with ZIKV, as DENV specific antibodies could bind to the ZIKV but are unable to neutralize and defend against the virus.[26] This is of concern because the results reported by Dejnirattisai *et al.* (2016) and Priyamvada *et al.* (2016) implicated that antibodies that were present to neutralize DENV were cross-reacting and promoting ADE of ZIKV infection.[26,27] This type of ADE of infection in children was also hypothesized with the recently licensed CYC-TDV vaccine[26], also known as Dengvaxia as created by Sanofi Pasteur which is currently in use in Brazil, Costa Rica, El Salvador, Mexico, Paraguay, and the Philippines.[28] Vaccination of DENV naïve patients experienced a greater rate of hospitalization than patients that were initially infected. This case observation led to the hypothesis that vaccination of seronegative patients was poorly protected from the pathogen despite developing enhanced antibodies for DENV1-4 and had a higher risk of infection.[26,28] Other vaccines are being developed to combat DENV serotypes 1-4, with the company Takeda currently in phase II and Butantan in phase III clinical trials.[29]

Yellow fever is a flavivirus infection caused by YFV and found mainly in sub-Sahara Africa and tropical South America where one billion people in 46 countries are at risk of infection[30] The disease was greatly feared until the development of a live attenuated vaccine by Max Theiler and associates in the 1930s.[30] Although, the vaccine helped

combat the virus considerably in the following decades, reports for yellow fever have been substantially underreported due to nonspecific symptoms between febrile illnesses and the limitations in surveillance and healthcare systems in the affected areas.[31] About 180,000 people are infected and 78,000 die every year despite the availability of a working vaccine.[30,31] This fact is attributed to the decline of mass vaccination after the 1960s, which may have allowed for the resurgence of the YFV.[30,31] Yellow fever symptoms are separated by two phases: the acute and toxic phases. A period of 3-6 days after infection, fever, muscle pain, headache, shivers, loss of appetite, nausea, and vomiting can occur. After an additional 3-4 days, many people tend to improve and the illness dissipates. However, approximately 15% of the infected people can enter the toxic phase, which can lead to stronger fevers, jaundice, and bleeding.[24]

The CHIKV is a re-emerging alphavirus that belongs to the Togaviridae family, which was observed in 2004 in Kenya.[32] After this, the virus began to spread to other novel locations, such as the South-East Asia region, including India where over one million cases were reported in 2006.[24,33] CHIKV is comprised of an enveloped, positive single-stranded RNA virus, which upon infection leads to a rapid onset of symptoms. Common symptoms are very similar to the previously described Zika, dengue, and Yellow fever viruses, except that, in severe cases, chikungunya fever can cause long-lasting joint pains that may last for several weeks or years.[24,33] There is currently no treatment available for chikungunya which continues to be a growing problem along with the other viruses.

The Nature of the Problem

There has been an increasing presence of arboviruses in the United States, mainly due to the appearance of the *Ae. aegypti* and *Ae. albopictus* mosquitoes, which are the principal vectors of the tropical diseases associated with YFV, DENV, CHIKV, and ZIKV viral infections.[34] The viruses are transmitted through the saliva when the female mosquito acquires a blood meal. The blood meal is needed for the female mosquito to obtain the necessary nutrients and energy to complete the gonotrophic cycle.[35,36] Global warming and the blood feeding behavior of these mosquitoes have contributed to facilitating the transmission and resurgence of the DENV and YFV in tropical and subtropical regions[31,37-39] and the recent rapid spread of CHIKV and ZIKV across the globe.[40] As of now, a vaccine is only available to combat the yellow fever virus[30] with a potential vaccine against dengue remaining elusive due to disappointing results from clinical trials of a leading candidate known as Dengvaxia.[28] Therefore, the only option to combat these diseases and the mosquito is through vector control. However, mosquito control methods have specific flaws that lead to increasing resistance due to overuse of insecticides and have very limited specificity, killing off other important insects (such as butterflies and bees). Also, removal of standing water that provides ideal breeding sites for the mosquito may be impractical for regions with rainy seasons.[41,42] Currently, a new specific vector control strategy against the *Ae. aegypti* mosquito utilizes male transgenic mosquitoes that will pass on a lethal gene to the offspring, killing the progeny before they can reach adulthood.[43] However, due to the varying degree of arbovirus density in different regions of the world and public resistance to field trials, it becomes extremely

difficult to release these engineered mosquitoes into the wild globally.[44-46] All of these factors make it extremely crucial to further study the *Ae. aegypti* mosquito and its vector-borne diseases and to investigate potential novel techniques to slow or reduce the spread of the mosquitoes and pathogens.

Digestion of Blood Meal by Regulated Midgut Proteases

The midgut of the mosquito is the main site for blood meal protein digestion and subsequent extraction of any essential nutrients and adsorption. The midgut also acts as the site where viruses can infect the vector, and once infected the mosquito can transmit the virus to vertebrate human hosts.[47] Once the female mosquito imbibes the blood meal, an extracellular layer produced by midgut epithelial cells is formed around the ingested blood. This peritrophic matrix allows for the retention of the blood meal until it can be fully degraded for its nutrients.[48] The peritrophic matrix is composed of chitin fibrils and proteoglycans that act to control the traffic of molecules between the midgut epithelial and lumen and to protect against any type of abrasion and toxic products that may derive from blood digestion.[48] The typical body weight of the mosquito is around two milligrams, and she can acquire more than her own body weight in blood.[47] This can result from a single blood meal or from multiple blood feedings.[9,13,47] For this reason, she is under immense stress, and the midgut must respond accordingly by releasing proteolytic enzymes to quickly digest and detoxify the blood meal.[47] The digestive enzymes are produced in the midgut and have evolved to be a specialized system to efficiently cleave the protein-rich blood meal[47] within two days.[49]

When the female *Ae. aegypti* mosquito obtains a blood meal, this initiates the release

of proteolytic enzymes in the midgut aiding in the degradation of blood meal proteins

into peptides and amino acids that can be used for carbohydrate, lipid, and yolk protein

synthesis, fueling mosquito functionality and egg production.[36,50] The expressed digestive

enzymes consist of endopeptidases, which are comprised of trypsin-[51,52] and

chymotrypsin-like[53,54] serine proteases, and exopeptidases, which are comprised of

aminopeptidases[55] and carboxypetidases.[56,57] Although many different proteolytic

enzymes are present, published work by Felix *et al.* in 1991 revealed a two-phase process

of midgut trypsin-like activity in the *Ae. aegypti* mosquito[58], suggesting that trypsin may

be the main proteolytic enzyme responsible for blood meal digestion.[58,59] As seen in

Figure 1, biphasic trypsin-like activity begins with an initial early phase almost

immediately after a blood meal to 6 hours post blood meal (PBM), followed by the late

phase starting at 12-18 hours PBM, reaching maximal activity at the 30th hour PBM.[59]
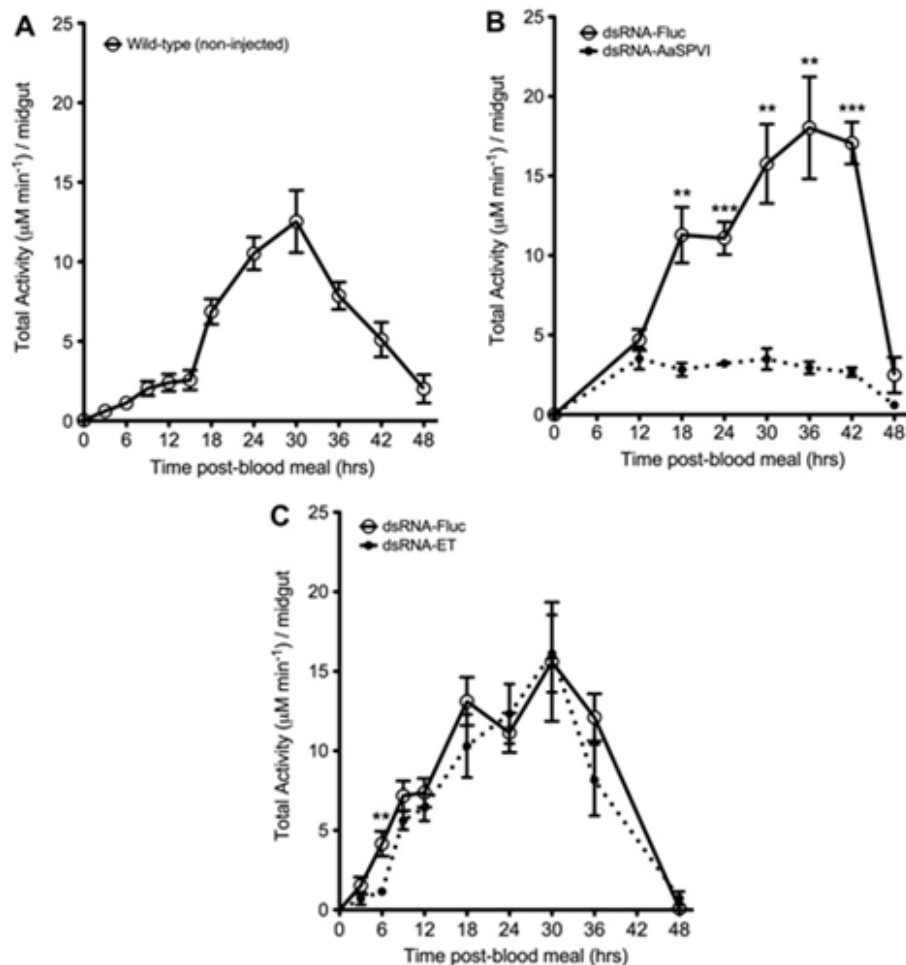
Figure 1. Total BApNA activity in individual mosquito midguts as determined by quantitative enzyme kinetic analysis using steady-state assumptions. A) Total midgut extract. B) Suppression of AaSPVI only. C) Suppression of AaET only. Reproduced from *Insect Biochemistry and Molecular*, 39/12, Isoe, J., Rascón, A.A., Kunz, S., Miesfeld, R.L., Molecular genetic analysis of midgut serine proteases in *Aedes aegpyti*, 903-912, 2009, with permission from Elsevier.

The observed biphasic manner of midgut trypsin-like activity was suggested to have specific proteases induced at specific hours.[58] During the first 6 hours PBM, a protease known as early trypsin (AaET) was found to be abundantly expressed.[51,60] Interestingly, high levels of un-translated AaET mRNA transcripts are already present in the midgut epithelial cells of the adult female mosquito before a blood meal and are not translated to

protein until a blood meal is ingested.[61] Two late serine endopeptidases known as late

trypsin (AaLT) and 5GI (later renamed AaSPVI) were also characterized around the same

time as early trypsin. They were shown to be strongly expressed at the 24th hour PBM,

insinuating that they may be essential during late phase digestion.[51,52,62] In the 1990s,

AaET was thought to activate AaLT through an unknown signal transduction system. For

this experiment, soybean trypsin inhibitor (STI) was added to an artificial blood meal and

fed to mosquitoes resulting in inhibition of AaET protease activity, which led to the loss

of AaLT protease activity.[63] This was reevaluated in 2006 when Lu and associates

discovered that the original STI used in their experiments was contaminated with

potential toxins leading to the aforementioned results. Therefore, Lu *et al.* (2006) then

repeated the feeding experiments using a variety of specific trypsin inhibitors that did

inhibit early trypsin activity, but had no effect on late trypsin activity. These results were

unexpected as it contradicted the original results obtained with the original STI.[64] To

ensure that the inhibitors were not affecting other metabolic and transcriptional processes,

the researchers suppressed AaET mRNA translation using RNA interference (RNAi)

technology. At the time, RNAi was successfully used in initial studies by Blandin and

associates in 2002 studying the function of the *Defensin* gene in *Anopheles gambiae*[65].

The method was optimized for *Ae. aegypti* to reduce AaET levels, which led to no

significant effects on late trypsin activity.[64] This result, combined with the inhibition

study provided evidence that AaET activity was not necessary for late phase enzyme

activity.[64]

Considering the high abundance of AaET[51,60,61], AaLT[52,62], and AaSPVI[51], and the

discovery of an additional highly induced late phase protease known as AaSPVII[59,66],

Isoe and associates investigated the important impacts these four highly expressed midgut

serine proteases have on blood meal protein digestion and egg development.[59] However,

before these studies were conducted, trypsin-like activity was measured in individual *Ae.*

aegypti mosquito midguts using Nα-benzoyl-L-arginine-4-nitroanilide (BApNA), a

common chromogenic trypsin substrate analog. As seen in figure 1C, AaET knockdown

confirmed the results from Lu *et al.* (2006) (described above), which did not affect late

phase midgut trypsin-like activity.[59] In fact, the majority of midgut late phase BApNA

activity was shown to be from AaSPVI (Figure 1B), while knockdown of AaSPVII and

AaLT had no significant effect on overall activity.[59] However, as shown in Table 1, the

suppression of AaSPVII and AaLT had a greater effect on the amount of eggs laid,

despite AaSPVI's prominent late phase BApNA activity.[59] These results are interesting

because AaSPVII is shown to be homologous to AaSPVI[66] (68% sequence identity,

National Center for Biotechnology Information (NCBI) database), and so similar activity

should have been observed, but these studies suggest that the enzymes may be playing

different functional roles in the blood meal digestion process.[59] In addition, research

studies by Rascón and associates have shown that AaSPVII has double the turnover

number (k_{cat}) and higher specific activity towards BApNA than AaSPVI *in vitro*[67], which

may indicate differences in substrate specificity *in vivo*. However, as seen in Figure 2 and

Table 1, by inhibiting AaSPVII and the other midgut serine proteases, midgut blood meal

protein digestion may be prevented, and thus eliminate the necessary nutrients needed for

egg development.[59] Therefore, it is essential to further characterize these proteolytic

enzymes and verify if targeting them can aid in reducing the mosquito population as a

potential vector control strategy.
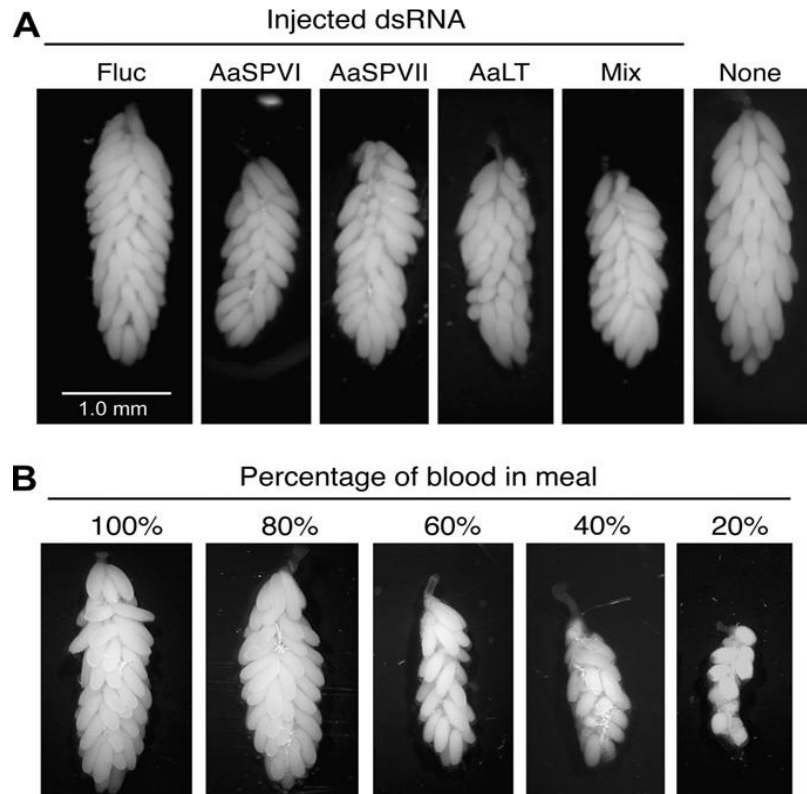
Figure 2. The effects of dsRNA injection or diluted protein bloodmeal on ovary development. A) dsRNA for each protease were designed, synthesized, and injected into female mosquitoes prior to blood feeding. Mix represented the combination of AaSPVI, AaSPVII, and AaLT. B) Blood was diluted from 20 to 80% with feeding buffer and fed to uninjected mosquitoes. Reprinted from *Insect Biochemistry and Molecular Biology*, 39/12, Isoe, J., Rascón, A.A., Kunz, S., Miesfeld, R.L., Molecular genetic analysis of midgut serine proteases in *Aedes aegpyti*, 903-912, 2009, with permission from Elsevier.

Table 1. Summary of Fecundity Data from Individual dsRNA-Injected Mosquitoes.

First gonotrophic cycle					
dsRNA	N	Eggs laid Mean ± SEM	Reduction (%)	N	Viability (%)
Fluc	51	120.0 ± 2.2	—	10	94.0
AaSPVI	48	99.3 ± 3.2***	17.2	9	92.2
AaSPVII	44	88.8 ± 3.5***	26.0	9	86.7
AaLT	41	89.4 ± 3.3***	25.5	10	91.3
Mix	38	82.6 ± 2.3***	31.2	9	91.9
Second gonotrophic cycle					
dsRNA	N	Eggs laid Mean ± SEM	Reduction (%)	N	Viability (%)
Fluc	19	99.0 ± 2.6	—	6	97.2
AaSPVI	21	72.3 ± 5.7***	27.0	6	98.3
AaSPVII	20	72.1 ± 3.8***	27.2	6	96.7
AaLT	17	63.4 ± 5.8***	36.0	6	88.9
Mix	27	72.4 ± 4.1***	26.9	6	95.6

N: Number of mosquitoes tested.
***Indicates egg numbers significantly different from dsRNA-Fluc injected mosquitoes ($P < 0.001$ in all cases).
Mix refers to injection with a mixture of AaSPVI, AaSPVII, and AaLT dsRNA using 400 ng f each.

Chapter II

Initial Biochemical Characterization of N-terminally His$_6$-tagged AaSPVII

Introduction

The ingestion of a blood meal by the female *Ae. aegypti* mosquito releases digestive enzymes known as exo- and endopeptidases.[66,68] The abundantly expressed midgut serine proteases released PBM are endopeptidases, more specifically trypsin-like serine proteases, which account for the majority of the digestion of blood meal proteins.[58] Trypsin, in general, has been known to be one of the most important, extensively studied enzymes for its role in food protein hydrolysis and its ability to autocatalyze its zymogen form (trypsinogen) to the active form and activate other digestive proteolytic enzymes in both vertebrates and invertebrates.[68-70]

Serine proteases are so named because proteolytic activity depends on a catalytic serine residue.[69] When catalysis begins, the serine at the active site becomes deprotonated by histidine, becoming a nucleophile and attacking the carbonyl of a peptide substrate forming a tetrahedral intermediate.[69,70] This serine deprotonation by histidine and histidine stabilization by aspartate are part of a charge relay system known as the catalytic triad (Ser-His-Asp).[69,70] The specificity of each enzyme is due to the amino acid in the substrate pocket. With trypsin, there is an aspartate residue in the substrate pocket, which stabilizes the positively charged lysine and arginine amino acid residues, leading to the specificity and cleavage of the peptide bonds on the carboxyl side of these amino acids. Chymotrypsin prefers peptide bonds associated with the carboxyl side of the large hydrophobic, aromatic amino acids, tyrosine, tryptophan, and phenylalanine.[69,70] In

contrast, serine collagenases have a serine residue in the binding pocket, therefore preferring different specificities than that of trypsin and chymotrypsin.[67]

Initial biochemical studies of recombinant AaET, AaSPVI, and AaSPVII show that they exhibit trypsin-like activity when using BApNA as a substrate (Table 2).[67] AaET was shown to have the highest specific activity towards BApNA, having 5-fold more activity than the other three abundant midgut proteases, and AaET was about 1.5-fold higher than the bovine trypsin control. AaET, AaSPVI, and AaSPVII have been classified as classical trypsin enzymes[59,66,67]; however, their exact specificity and catalytic mechanism are still unknown. As mentioned above, AaSPVI and AaSPVII are homologous to each other (68% sequence identity), but AaSPVII was shown to cleave BApNA more efficiently than AaSPVI despite their amino acid similarities.[67] This was interesting because, as seen in Figure 1, it was previously shown that knockdown of AaSPVI expression reduced late-phase trypsin-like activity by ~78%[59] despite its low specific activity and turnover number.[67] This raises the question as to why AaSPVII was shown to have higher specific activity than AaSPVI, but did not contribute to much of the late phase BApNA cleavage and contributed significantly to the overall fecundity in blood fed mosquitoes.[59,67] To attempt to answer this question, further biochemical studies using bacterial recombinant and protein expression techniques must be done on the protease to investigate its structure and function.

Bacterial recombinant technology has revolutionized biochemistry because, by applying the right combination of bacterial expression vectors and bacterial expression strains, one can achieve high-level production of recombinant proteins that are needed for

in vitro biochemical characterization.[71] Rascón and associates attempted to first purify the full-length zymogen forms of AaET, AaSPVI, AaSPVII, and AaLT, including the signal peptide (leader) sequence.[67] Unfortunately, the low levels of bacterial expression, insoluble protein, and purification of recombinant forms of these proteases did not yield enough products to perform biochemical studies.[67] There may have been two main reasons for the high level of insoluble expression of the proteins. First, the presence of the leader sequence may have affected the overall efficiency of the expression. In Figure 3, the full-length gene with the leader sequence is present before the propeptide region, the region that needs to be cleaved to form the mature protease.[67,69]
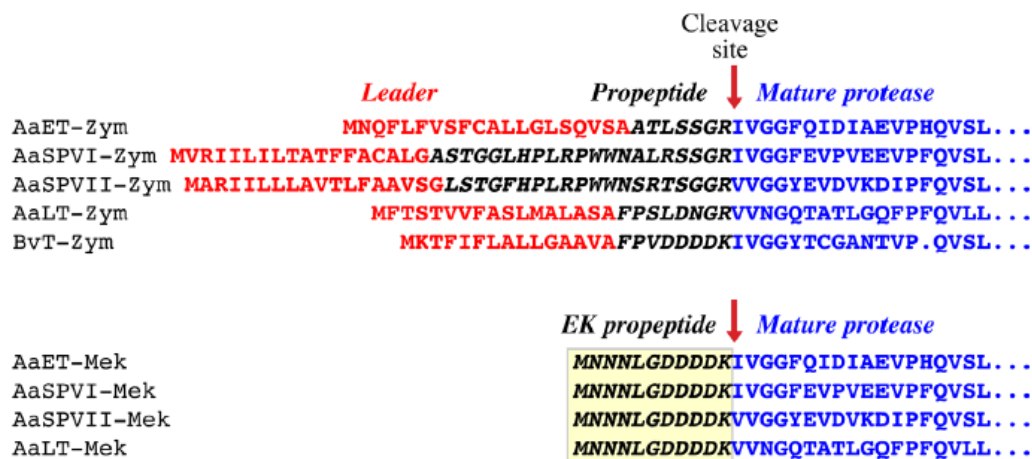
Figure 3. Amino acid sequences of the wild-type zymogen and the engineered heterologous enterokinase artificial propeptide region (Mek) forms of the enzymes. Reprinted from *Biomed Central Biochemistry*, 12/43, Rascón, A.A., Gearin, J., Isoe, J., Miesfeld, R.L., *In vitro* activation and enzyme kinetics analysis of recombinant midgut serine proteases from dengue vector mosquito *Aedes aegypti*, 2011, with permission from BioMed Central Ltd.

Although the leader sequence plays a key role in targeting and membrane insertion of secretory and membrane proteins in both prokaryotes and eukaryotes, the presence of the

17

signal peptide during recombinant protein expression has been shown to lead to thermodynamic destabilization, which reduces the activity of the desired protein and increases aggregation propensity.[72] In addition, results published from Promdonkoy *et al.* (2004) showed improved expression levels of mosquitocidal toxin 1 after removal of the putative leader sequence.[73] For the Rascón *et al.* (2011) studies, the leader sequence was kept for AaET, AaSPVI, AaSPVII, and AaLT, which may have contributed to the low expression levels.[67] As an alternative, Rascón and associates engineered a heterologous enterokinase cleavage site, which has the designated DDDDK sequence. As seen in Figure 3, the enterokinase (EK) cleavage was engineered to allow for the activation of the desired protease by using commercially available enterokinase to cleave the EK propeptide site, forming the mature protease, but also to help improve expression levels by removing the leader sequence.[67] However, this approach still resulted in insoluble expression, leading to the second main reason why high levels of insoluble protein expression were observed. In the paper, Rascón *et al.* described using Novagen Rosetta 2 (DE3) and BL21 (DE3) *E. coli* bacterial expression systems. The BL21 (DE3) *E. coli* strain is a common bacterial T7 expression system that works by using a synthetic inducer isopropyl-β-D-thiogalactopyranoside (IPTG) to activate the expression of genes cloned into pET vectors, while the Rosetta 2 (DE3) is a BL21 (DE3) derivative that aims to increase eukaryotic protein production by supplying specific rare codons that are not commonly found in bacteria[71]. Although this may be the case, Rosetta (DE3) strains in general have been shown to both improve levels of protein production and protein insolubility.[71] The reason for expression of insoluble protein using these cells has to do

with the fact that the bacterial cytoplasm is a reducing environment, which is the site

where recombinant protein expression takes place.[71] This is important to understand

because proteases have been shown to rely on proper disulfide bond formation to

properly fold and be biologically active.[74] By using these cells, proper disulfide formation

is not possible due to the reducing cytoplasm of wild-type *E. coli*, which can lead to the

expression of insoluble, unfolded proteins and aggregates known as inclusion bodies.[74]

With high levels of inclusion bodies, denaturing/refolding strategies must be considered

to obtain soluble proteins.[75] Refolding is not straightforward and may require an

extensive trial-and-error approach with various reagents needed for proper solubilization

and refolding. Denaturation occurs by using specific agents (such as guanidine HCl or

urea) that can solubilize the desired inclusion bodies efficiently. The rate at which the

protein can refold depends on the method to reduce the denaturant concentration and the

solvent conditions used.[75] During this rigorous process, misfolding and aggregation may

also occur, which can lead to varying yields of soluble protein.[75]

Due to the challenges presented, several changes can be made. To start, primer design

and DNA cloning can be utilized to remove the hydrophobic leader sequence from the

desired protease gene. By analyzing the full-length AaSPVII amino acid sequence using

the SignalP 4.1 Server, we can identify the location of the signal peptide cleavage site

sequence and develop specific primers to remove the leader sequence. The leaderless

protease gene can then be cloned into the pET28a vector expression system to utilize the

poly His_6-tag feature for purification. Once cloned, AaSPVII can be expressed in BL21

(DE3) derivatives known as T7 SHuffle competent cells (New England BioLabs). These

E. coli strains carry mutations to favor an oxidizing cytoplasm and disulfide bond isomerase, DsbC, that assist in disulfide bond formation.[76] This chapter will discuss the efforts to apply these conditions to express the late-phase midgut serine protease AaSPVII in bacteria and the approach taken for obtaining purified, soluble protein. In addition, site-directed mutagenesis will also be discussed as it has been hypothesized that the *Ae. aegpyti* midgut serine proteases may be autocatalytic in the absence of the enterokinase cleavage site.[51] Enterokinase serves to regulate the activation of trypsin[69], but the putative insect trypsinogens lack this region, which may also explain why enteropeptidases have not been found in mosquitoes.[51] We aimed to perform point mutations at the arginine residue of the predicted propeptide cleavage site (see Figure 3) and at the serine of the active site of AaSPVII to further elucidate the activation process of this protease if it is indeed autocatalytic.

Table 2. Steady-State Kinetic Parameters of Mosquito Proteases and Bovine Trypsin using BApNA as a Substrate.

Protease[a]	$K_{M(BApNA)}$ (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (mM^{-1} s^{-1})	Specific Activity (μmol min^{-1} (mg of protein)$^{-1}$)
AaET	63.0 ± 1.3	2.03	32.3	5.08
AaSPVI	11.8 ± 0.3	0.16	13.5	0.40
AaSPVII	19.2 ± 1.0	0.35	18.5	0.88
AaLT	ND	ND	ND	ND
Bovine Trypsin	1051 ± 54	1.30	1.2	3.28

[a]Reaction conditions: 20 mM TRIS-HCl+10 mM CaCl$_2$ pH 8.0 and 24°C.
ND: Not detected, exceeded lower level of detection for p-Nitroaniline (cleavage product of BApNA) >0.0125 abs units.
Values were determined from a mean of three determinations.

Reprinted from *Biomed Central Biochemistry*, 12/43, Rascón, A.A., Gearin, J., Isoe, J., Miesfeld, R.L., *In vitro* activation and enzyme kinetics analysis of recombinant midgut serine proteases from dengue vector mosquito *Aedes aegypti*, 2011, with permission from BioMed Central Ltd.

Methods

Preparation of Late-Phase Midgut Serine Protease AaSPVII Plasmid Construct

The leader (signal peptide) sequence cleavage site was identified using SignalP 4.1 Server and ExPASy – Translate Tool. The leader sequence was removed, and identification of the proper zymogen sequence was determined to design the correct 5'-end (forward) primer and 3'-end (reverse) primer. Restriction cleavage sites of NdeI (CATATG) and HindIII (AAGCTT) were included to form complementary sticky ends for proper cloning into the pET28a vector. A poly-A tail was also added for restriction enzyme cleaving efficiency. A double stop codon (TAATAA) was introduced as part of the HindIII restriction site for the reverse primer. The melting temperatures (T_m) of the designed primers were determined using NetPrimer from Premier Biosoft (http://www.premierbiosoft.com/netprimer/), as listed in Table 3.

Table 3. Primers Designed and used for PCR Amplification and Cloning of No Leader Wild-Type AaSPVII Zymogen. Primers were purchased from Elim Biopharmaceuticals, Inc.

Protease	Primer	Primer Sequence	T_m (°C)
AaSPVII - Zymogen	AaSPVII-Zym-pET-Fwd	5' – **AAAAA**<u>CAT*ATG*</u>CTATCAACCGGATTCCATCCGC – 3'	65.4
AaSPVII - Zymogen	AaSPVII-Zym-pET-Rev	5'– **AAAAA**<u>AAGC*TTATTA*</u>AACTCCACTGACTTCGGCCA– 3'	63.7

Poly-A tail used is in **BOLD**.
Restriction sites used for cloning of the genes of interest are <u>underlined</u>.
Start and double stop codons are *italicized*.
Annealing temperature calculated using NetPrimer.

PCR amplification was performed as described in the *Promega GoTaq Green Master Mix* (Cat. # M7122) protocol using the designed forward and reverse primers purchased from Elim Biopharmaceuticals, Inc. The DNA template used to amplify the No Leader AaSPVII gene of interest was that of the full-length AaSPVII gene cloned into the pET29a vector (AaSPVII Z/pET29a). No leader AaSPVII Zymogen was amplified using the Eppendorf Mastercycler Gradient thermocycler. The thermocycler was set at 95 °C for 5 minutes (initial denaturation), 25 cycles at 95 °C for 20 seconds (denaturation), 60°C for 20 seconds (annealing), 72 °C for 20 seconds (elongation), and then a final elongation at 72 °C for 5 minutes. An amount of 1000 ng of pET28a vector (Novagen Cat. #6986-3) and the amplicon were digested separately in 1X FastDigest buffer (ThermoScientific), 1 µl NdeI (ThermoScientific Cat. #FD0584), and 1uL HindIII (ThermoScientific Cat. #FD0505) restriction enzymes at 37 °C for 60 minutes. The enzymes were inactivated by incubating the reaction tubes at 90 °C for 10 minutes, followed by a 5-minute ice incubation. The digested vector was then dephosphorylated with thermosensitive alkaline phosphatase (TSAP) (Promega Cat. # M9910) to prevent the vector from ligating. PCR product, digested/dephosphorylated vector, and GeneRuler 1 kb DNA Ladder (ThermoScientific, #SM0311) were run on a 1% agarose gel using 1X TAE buffer. The desired bands were excised and extracted by following the protocol described in the *QiaQuick Gel Extraction Kit* (Qiagen Cat. # 28707). Both the digested amplicon and the dephosphorylated vector were concentrated by ethanol precipitation and ligated in 1X T4 DNA ligase buffer (New England BioLabs Cat. #M0202), 18 µl Milli-Q water, and 1 µl T4 DNA ligase (New England BioLabs Cat. #M0202). The ligation

reaction was incubated at room temperature for 30 min. Ligated AaSPVII (No Leader) zymogen/pET28a was transformed into NEB 10-β competent *E. coli* cells (New England BioLabs # C3019**)** as described in the *High Efficiency Transformation Protocol* with the following modifications. From the AaSPVII (No Leader) zymogen/pET28a ligation reaction, 10 μl was added into separate transformation tubes containing the bacterial cells, and transformation protocol followed. For the recovery phase, SOC media was added and the reaction was incubated in a Lab-line Orbit Envirion Shaker at 37 °C (220 rpm) for 60 minutes. The transformation tubes were centrifuged for one minute at full speed and then 950 μl of the supernatant was discarded. The pelleted cells were re-suspended with the remaining 50 μl and the whole volume was plated on to pre-warmed luria broth (LB) agar plates supplemented with kanamycin (30 μg/ml). Once colonies were observed, overnight cultures were set and placed in 37 °C shaker overnight (220 rpm). Isolation of the plasmid was done by following the *QIAprep Spin Miniprep Kit* (Qiagen, Cat # 27106) protocol, with the exception of the elution step. Plasmid DNA was eluted using 55 μl MilliQ water. The concentration of the AaSPVII (No Leader) zymogen pET28a plasmid was determined using the ND 1000 Spectrophotometer Nanodrop. All plasmid samples were sent to Elim Biopharmaceuticals, Inc. for DNA sequencing and then analyzed and confirmed by using the Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi), Clustal Omega <Multiple Sequence Alignment < EMBL-EBI (http://www.ebi.ac.uk/Tools/msa/clustalo/), and ExPASy – Translate Tool (http://web.expasy.org/translate/).

Small-Scale Bacterial Overexpression Using BL21 (DE3) and T7 Shuffle Competent Cells and *In Vitro* Spectrophotometric BApNA Assay

The AaSPVII (No Leader) Z/pET28a plasmid construct was transformed into the BL21 (DE3) bacterial expression system as described in the *BL21 Competent Cells* (BioLine Cat. # BIO-85032) protocol. However, rather than plating the entire volume of transformant (as described above for NEB 10-β), only 10 μl of the resuspended pellet was used. The overexpression of the protein of interest was done by selecting a single colony from the transformation plate and by growing overnight in 5 mL LB medium supplemented with 30 μg/ml kanamycin (37 °C, 220 rpm). From the overnight culture, the initial volume of cell culture was calculated to initially begin at an OD_{600} of 0.05 in an autoclaved 125 mL Erlenmeyer flask with 50 mL LB + kanamycin (30 μg/ml) and set in a 37 °C shaker at 220 rpm. At OD_{600} ~0.5-0.8, the cell culture was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown for 4 hours and 1 mL samples were collected at different time points.

The AaSPVII Z/pET28a plasmid construct was also transformed into *SHuffle T7 Express Competent E. Coli High Efficiency Transformation* cells (New England BioLabs Cat. #C3029), as described above for BL21 (DE3) cells. Overnight cultures were prepared in LB with kanamycin and set in a 30 °C shaker (250 rpm) overnight. The bacterial growth experiment was repeated as above, except the cells were grown in terrific broth (TB) at 30 °C (250 rpm) until an OD_{600} ~0.5-0.8, and before inducing with 0.1 mM IPTG, the temperature was lowered to 23 °C. Cells were grown for 28 hours and 1 mL samples collected at different time points.

For SDS-PAGE analysis, all time points were resuspended in 20mM Tris-HCl, pH 7.2, and sonicated at 20% amplification for 10 seconds each (total of three cycles) on ice. A sample was collected (total protein sample) before centrifugation and after centrifugation at 13,000 rpm (4 °C) a soluble protein sample was collected. All protein samples were prepared for electrophoresis with 6x SDS-PAGE sample buffer and denatured at 95°C for 4 minutes. SDS-PAGE was done using Novex NuPAGE SDS-PAGE 4-12% Bis-Tris gels (Invitrogen, Cat. #NP0321BOX), prestained protein ladder (ThermoScientific PageRuler Cat. # 26616), and 1X NuPAGE MES running buffer (Invitrogen Cat. #NP0002). After running the gels, they were stained with InVision His-Tag In-Gel stain (Life Technologies Cat. #LC6030) and SimplyBlue Stain (Life Technologies, Cat. # LC6060). The BL21 (DE3) expressed proteins were only blue stained as per the manufacturer's protocol.

The synthetic chromogenic substrate BApNA (Acros Organics, Cat. # AC227740010) was used to test for trypsin-like activity. The time point samples collected from the AaSPVII expressed in T7 Shuffle cells were used. A typical BApNA master mix reaction can be seen in Table 4. Reactions were initiated with 20 μL of the AaSPVII crude lysate, and the absorbance at 405 nm was monitored using the Hewlett Packard (HP) 5452A Diode Array spectrophotometer. Initial reaction rates of each time point were monitored for 10 minutes.

Table 4. Typical BApNA Master and Reaction Mixture used for Activity Assays. BApNA was dissolved in DMSO.

BApNA Activity					
	20 mM	**10 mM**	**1 mM**		
Sample (µl)	**1 M Tris-HCl, pH 7.2 (µl)**	**1 M CaCl$_2$ (µl)**	**138 mM BApNA (µl)**	**ddH$_2$O (µl)**	**Total Volume (mL)**
20	20	10	7.2	942.8	1
400	400	200	144.9	18855.1	20
For Master Mix, combine water, Tris, and CaCl$_2$ only					
For a single reaction:	1 mL				
Master Mix (µl)	**Sample (µl)**	**Total Volume (mL)**			
980	20	1			

Large-scale Overexpression in T7 SHuffle and Purification

The AaSPVII (No Leader) Z/pET28a construct was transformed in T7 SHuffle bacterial cells (New England BioLabs) and overnight cultures set, as above. However, a larger volume was used to set the overnight cultures (75 mL LB + kanamycin (30 µg/ml)). The overnight cultures were incubated in a ThermoScientific MaxQ 4000 shaker at 30 °C and set at 250 rpm overnight. The bacterial growth expression was done as described above, but rather than a single 50 mL growth, 6 x 0.5 L of TB + kanamycin (30 µg/ml) in autoclaved 1 L Erlenmeyer flasks were used. The cultures were induced and grown for 5 hours. Bacterial cells were harvested by centrifugation at 6000 rpm (8 °C) for 10 minutes using the IEC B-22M Programmable Centrifuge. Pellets were then flash frozen in liquid nitrogen and stored at -80 °C until needed.

The AaSPVII (No Leader) Zymogen enzyme was highly expressed in soluble form in the conditions described (see Results section for more information). For purification, 7 to 10 g of the collected cell paste was used. Cell paste was lysed by pulse sonication at 30%

amplification for 15 minutes on ice in buffer containing 20 mM Tris-HCl, pH 7.2 + 250 mM NaCl + 10 mM imidazole + 2 mM dithiothreitol (DTT) (Buffer A). The lysed cells were then centrifuged at 16,000 rpm at 4°C for 30 minutes to obtain a clear crude lysate. The protein sample was loaded on to a Nickel Column 5 ml HisTrap FF (GE Healthcare Cat no. 17-5255-01), which had been equilibrated with Buffer A using the AKTA FPLC Amersham Pharmacia Biotech instrument. The column was then washed with 10 column volumes (CV) of Buffer A. The sample was then eluted with a three-step linear gradient using 20 mM Tris-HCl, pH 7.2 + 250 mM NaCl + 500 mM imidazole + 2 mM DTT (Buffer B): 10%B for 3 CV, 30% B for 6 CV, 50% B for 5CV, and 100% B for 20 CV. Fractions of 1.5 mL were collected and analyzed using SDS-PAGE. All fractions that contained the desired protein were pooled together into a dialysis membrane tubing (FisherScientific, Cat. #08-6670) and dialyzed in 50 mM sodium acetate, pH 5.2 + 2 mM DTT (2 L x 2) at 4 °C to remove the imidazole and excess salt. The dialyzed protein was then centrifuged at 4 °C (3500 rpm) for 10 minutes using the Beckman GPR Centrifuge to remove any precipitated products. The protein was then concentrated using the Amicon Ultra-15 Centrifugal Filter Units Regenerated Cellulose 10,000 NMWL Ultracel-10k (Millipore, REF UFC901024). The concentrated protein was analyzed using SDS-PAGE, and concentration was estimated using the BCA Protein assay kit (Pierce, Cat. #23227) and BioTek ELx808 plate reader. The final concentrated protein was then aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C.

Activation of Nickel Purified AaSPVII Z/pET28a

A separate nickel purification of AaSPVII (No Leader) Zymogen enzyme was done with same methods as described above (all steps, dialysis, and concentrating to 7 mL). Approximately half of the nickel purified AaSPVII (No Leader) Zymogen concentrate (in 50 mM sodium acetate, pH 5.2 + 2 mM DTT) was transferred into dialysis membrane tubing and allowed to dialyze in 2 L 20 mM Tris-HCl, pH 7.2 + 10 mM $CaCl_2$ with stirring using a magnetic stirrer in the cold room (4 °C). Samples of 20 μL volumes were collected for both SDS-PAGE analysis and BApNA assays every 15 minutes for a total of 135 minutes. SDS-PAGE analysis was done with the collected time points and the BApNA assay was monitored for color change but not analyzed on the spectrophotometer for trypsin-like activity (qualitative assay).

In Vitro **BApNA Spectrophotometric Assay of Recombinant AaSPVII (No Leader) Zymogen**

The synthetic chromogenic substrate BApNA (Acros Organics, Cat. # AC227740010) was used to test for trypsin-like activity. Different concentrations of BApNA and a constant concentration of AaSPVII (No Leader) Zymogen enzyme were used to obtain a Michaelis-Menten saturation curve.[67] Reactions were initiated with 50 or 100 nM of the nickel purified AaSPVII Z (No Leader) Zymogen enzyme into the master mix and the absorbance at 405 nm was monitored using the Hewlett Packard (HP) 5452A Diode Array spectrophotometer as a function of time. Reaction rates were determined from the linear portions of the absorbance at 405 (A_{405}) versus time plots using the extinction coefficient (ε_{405nm}) 8.80 $mM^{-1}*cm^{-1}$. Each concentration was done in triplicate. Steady-state conditions were set to help determine the kinetic parameters of the wild-type

AaSPVII enzyme. The results obtained from the spectrophotometer for each

concentration was fit into the Michaelis-Menten equation and statistically analyzed using

GraphPad Prism statistical software package (Version 5.0b for MAC OS X, GraphPad

Software Inc).

Site-Directed Mutagenesis of the Propeptide Cleavage Site and Active Site of AaSPVII

Specific primers were designed with point mutations, in order to mutate the arginine

at the propeptide cleavage site (see Figure 3) and the serine at the active site to alanine,

following the guidelines from the QuikChange XL Site-Directed Mutagenesis Kit

(Agilent Technologies, Cat. #200516). The arginine cleavage site was predicted to be the

site of activation[67] and the serine residue at the active site was predicted using ExPASy

Prosite – Scan Prosite Tool (http://prosite.expasy.org/scanprosite/). The codon coding for

arginine was changed from CGT to GCT to code for alanine. The codon coding for serine

was changed from TCC to GCG to code for alanine (Table 5). Mutagenesis of the

AaSPVII Z/pET28a plasmid was done with the purchased primers and following the

methods described in the QuikChange XL Site-Directed Mutagenesis Kit, with the

following changes. The annealing temperature was set to 60 °C and the elongation time

was set to 7 minutes for each cycle for a total of 18 cycles. Super optimal broth (SOC)

outgrowth medium (New England BioLabs, Cat. #B9020S) was used instead of NZY⁺

broth. After the one hour incubation, the 500 µL sample was pelleted and then 400 µL of

the supernatant was removed. The pellet was resuspended with the remaining 100 µL and

then plated onto an LB + kanamycin (30 µg/ml) agar plate. Overnight culture and

isolation were done as described above. Mutations at these sites were confirmed through

DNA sequencing. The mutants were transformed into T7 SHuffle, and small-scale

bacterial growth experiments were conducted as described above. The cells were induced

and grown for a total of 28 hours, and 1 mL samples were collected at different time

points.

Table 5. Mutagenic Primers of AaSPVII. Point mutations were introduced at the arginine
propeptide cleaving site and the active site serine.

Protease	Primer	Primer Sequence	T_m (°C)
AaSPVII – Zymogen	AaSPVII-Mut-R to A-Fwd	5'-AAT AGC CGC ACT TCC GGC GGT **GCT** GTC GTC GGC GGT TAT GAA GT-3'	91.82
AaSPVII - Zymogen	AaSPVII-Mut-R to A-Rev	5'-ACT TCA TAA CCG CCG ACG AC**A GC**A CCG CCG GAA GTG CGG CTA TT-3'	91.82
AaSPVII - Zymogen	AaSPVII-Mut StoA-Fwd New	5' –TGC CAG GGT GAT **GCG** GGT GGT CCA TTG– 3'	80.79
AaSPVII - Zymogen	AaSPVII-Mut StoA-Rev New	5' –CAA TGG ACC ACC **CGC** ATC ACC CTG GCA– 3'	80.79

Bold and underlined are the mutated codons of the selected sites.
Melting temperature calculated using NetPrimer.

Results

The presence of the leader sequence was hypothesized to lower expression levels of

AaSPVII, as this was seen in other published articles as well.[72,73] The designed primers

were made specifically to remove the leader sequence and to facilitate cloning of the gene

into the pET28a vector system to exploit the N-terminal His$_6$-tag feature. The poly-A tail

was added at the 5'-ends of each primer to improve cleaving efficiency of the restriction

enzymes as it has been known that restriction enzymes have difficulty cleaving the

recognition sites if they are located near or extremely near the 5'-ends.[77] The primers

were purchased from Elim Biopharmaceuticals, Inc. and successfully worked in DNA cloning. The gene of interest was calculated to have 764 base pairs (bp) and, after PCR amplification, was shown to align with the DNA molecular weight ladder.

The expression of the wild-type AaSPVII Zymogen in pET28a was carried out in both BL21 (DE3) and T7 SHuffle competent cells to confirm improved protein expression. The BL21 (DE3) cells were grown at 37°C and found to have improved expression without the leader sequence[67] but failed to yield soluble protein (Figure 4). This was expected, as the disulfide linkage is crucial for proper protein folding and activity, which is impossible to form in bacterial systems due to their reducing environment[74,76,78]; this becomes arduous if soluble active enzyme is needed.
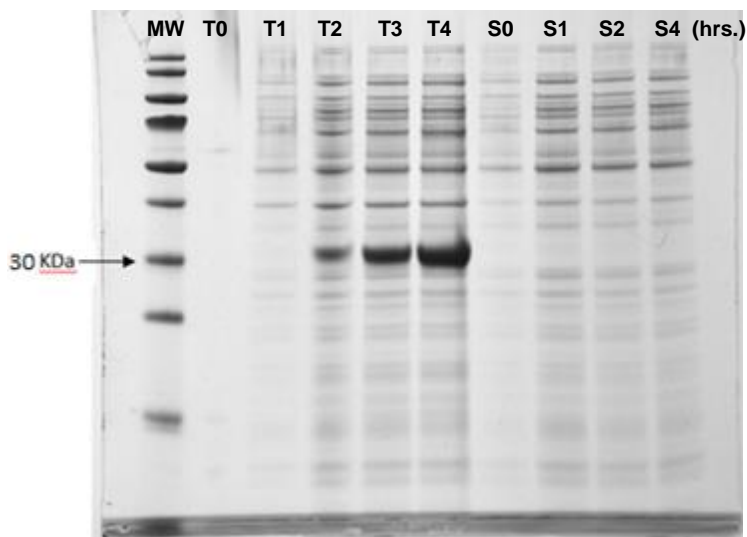
Figure 4. SDS-PAGE analysis of AaSPVII Z/pET28a transformed in BL21 (DE3) competent cells stained in SimplyBlue. Molecular weight of the AaSPVII (No Leader) Zymogen protein was calculated to be 28.7 kDa, running near the 30 kDa band on the prestained protein ladder (ThermoScientific, Cat. #PI-26616). Total protein (T) was shown to be the highly expressed over 4 hours after 0.1 mM IPTG induction. Soluble protein (S) was not observed, revealing that the desired gene was expressed in insoluble form. Lane numbers represent time in hours.

We therefore turned to the BL21 (DE3) derivative known as T7 SHuffle, which contains

mutations that create a more oxidizing environment in the cytoplasm. These competent

cells also contain DsbC, a protein that helps with proper folding of disulfide bonds in the

recombinant enzyme.[76] Induction by IPTG in these bacterial strains led to soluble

expression of the desired AaSPVII (No Leader) Zymogen of interest (Figure 5). When

compared to the insoluble expression in BL21 (DE3), the T7 SHuffle expression resulted

in soluble, properly folded AaSPVII, but also initial evidence of possible autocatalysis.

The His-tag in gel stain (which specifically binds to His_6-tagged enzymes) revealed a

decrease in intensity over time (after induction) of AaSPVII grown in T7 SHuffle cells.

This may suggest cleavage of the N-terminal His_6-tag (Figure 5A). Once visualized with

the His-tag stain, the same gel was stained with SimplyBlue to allow visualization under

white light and comparison of expressed AaSPVII between the two stains. From both

stained gels, the presence of three total bands appear with a band near the 30 kDa marker

(AaSPVII No Leader Zymogen with the His_6-tag approximately 28.7 kDa), and

interestingly, the molecular weight of each species decreases, but the intensity compared

to the rest of the proteins is higher. To test if any trypsin-like activity was present, a

preliminary BApNA assay was set using the individual crude lysates obtained from the

time point samples collected from the T7 SHuffle growth experiments. Surprisingly,

trypsin-like activity increased over time (after induction), with maximal activity observed

at the 18$^{\text{th}}$ hour of expression and gradually decreasing afterwards (Figure 6).
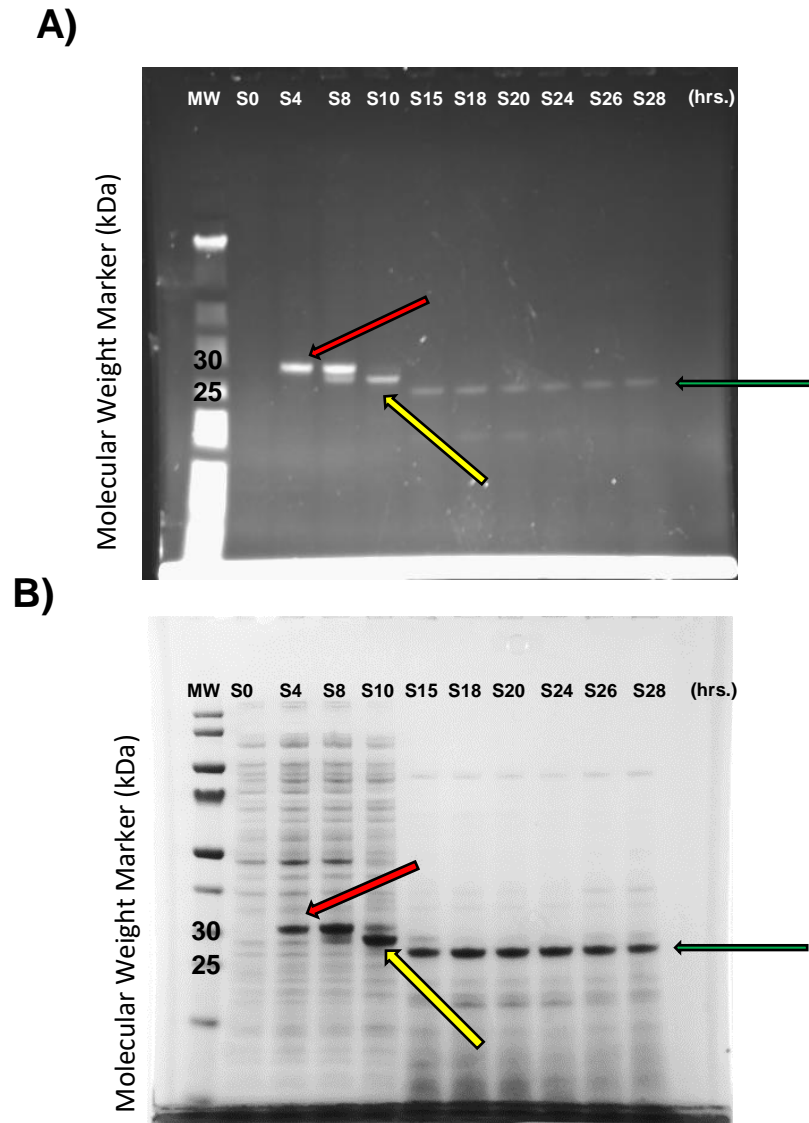
Figure 5. SDS-PAGE analysis of solubly expressed AaSPVII (No Leader) Zymogen expressed in T7 SHuffle competent cells, induced with 0.1 mM IPTG, grown for 28 hours at 23 °C. Multiple species can be observed in both gels and may be suggesting auto-activation of the protease. Molecular weight of the AaSPVII Zymogen was calculated to be 28.7 kDa (red arrow). The no His$_6$-tagged "Inactive" protease was calculated to be 26.8 kDa (yellow arrow). The mature protease was calculated to be 24.1 kDa (green arrow). All predicted molecular weight calculations were done using ExPASy. A) Gel visualized with His-tag in gel stain. The intensity of the AaSPVII (No Leader) Zymogen enzyme can be seen fading as expression time increases, which may suggest that the enzyme is losing the N-terminal His$_6$-tag. B) The same gel used in A was stained with SimplyBlue. The dominant expressed AaSPVII (No Leader) Zymogen species is over-expressed in the T = 4 and 8 hour samples, but as time increases, two new over-expressed species can be clearly seen. These results indicate the possibility of auto-activation. Lane numbers represent time in hours.
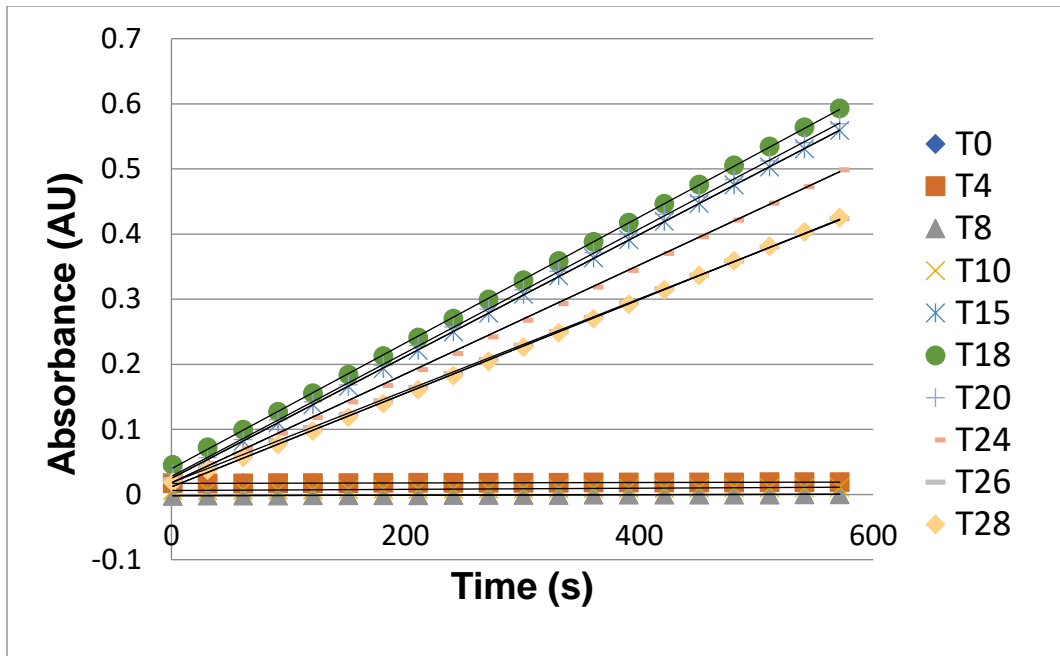
Figure 6. BApNA activity of AaSPVII Z/pET28a crude lysate. BApNA was used as a substrate to detect trypsin-like activity. From the collected time points, peak activity is observed at the 18th hour of expression and begins to decline post-18 hours. No significant activity is observed until after 10 hours of expression and growth.

After several small-scale bacterial growth experiments, we discovered that there are certain time points where the 28.7 kDa band was present before the other lower molecular weight species appeared. The maximal expression of the putative zymogen form was discovered to be intact until the 6th hour of expression and then visible auto-cleavage occurs (data not shown). Because we are hypothesizing that the N-terminal His6-tag is being cleaved (Figure 5A) and there is no observed trypsin-like activity between 4-8 hours using the BApNA assay (Figure 6), we decided to collect cell paste for purification from the large-scale bacterial growth at 5 hours post-induction (Figure 7).
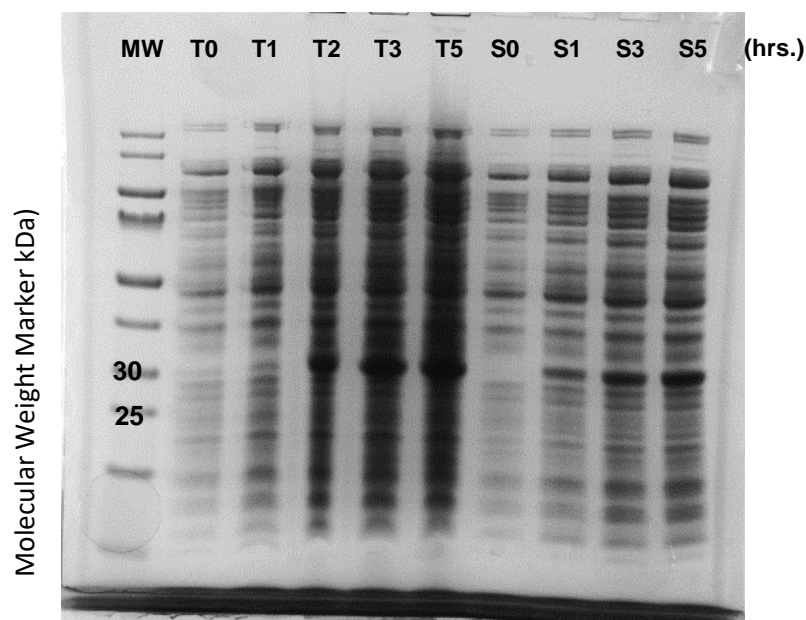
Figure 7. Large-scale bacterial expression of AaSPVII Z/pET28a transformed in T7 SHuffle *E. coli* competent cells. Calculated molecular weight of AaSPVII (No Leader) Zymogen is 28.7 kDa. The cells were grown at 30 °C until OD_{600} ~0.5-0.8, induced with 0.1 mM IPTG, and grown for 5 hours at 23 °C. A single band near the 30 kDa marker is increasing over time in both the total and soluble samples. Lane numbers represent time in hours.

To purify N-terminally His_6-tagged AaSPVII (No Leader) Zymogen enzyme, multiple HisTrap FF Nickel purifications were attempted. As seen in Figure 5, autocatalyzed AaSPVII would pose a problem and not bind to the Nickel column due to the loss of the desired N-terminal His_6-tag. Therefore, we hypothesized that by stopping the expression at the 5th hour and collecting the cell paste, we would be able to purify intact AaSPVII Zymogen, only if proper conditions are set. Unfortunately, trypsins have been shown to have broad maximal pH ranging from 7 to 9 for activity[79]. Regardless, the initial attempt was done with buffers at physiological pH (20 mM Tris-HCl, pH 7.2 + 10mM $CaCl_2$) to observe if it was possible to purify intact zymogen without changing the buffer conditions. As seen in Figure 8A, for purified AaSPVII, after dialysis and

concentration, a small band near the 25 kDa marker is observed, suggesting that the

mature protease (24.1 kDa) is present and that it may be auto-degrading due to the

presence of many other lower molecular weight protein species. In order to slow or

inhibit this process, we attempted to add 2 mM dithriothreitol (DTT) to the purification

and dialysis buffers to partially reduce disulfide bonds and minimize auto-degradation of

the protease.[67] As seen in Figure 8B, auto-digestion and full activation of the protease

was slowed as the presence of the 26.8 kDa (no His$_6$-tagged zymogen AaSPVII) and 24.1

kDa (mature AaSPVII) bands were more pronounced compared to Figure 8A, and the

concentration of degraded products were not as prominent. Because of this, and the fact

that we were able to purify intact AaSPVII Zymogen (Figure 8C), we determined that

simply adding DTT to the dialysis buffer was not enough to prevent activity and

degradation. Therefore, we decided to alter the pH of the dialysis buffer. The purification

buffers still contained 2 mM DTT, however, the dialysis buffer was changed to 50 mM

sodium acetate, pH 5.2 + 2 mM DTT. As seen in Figure 8C, auto-activation and auto-

digestion was prevented, thus obtaining intact AaSPVII Zymogen.
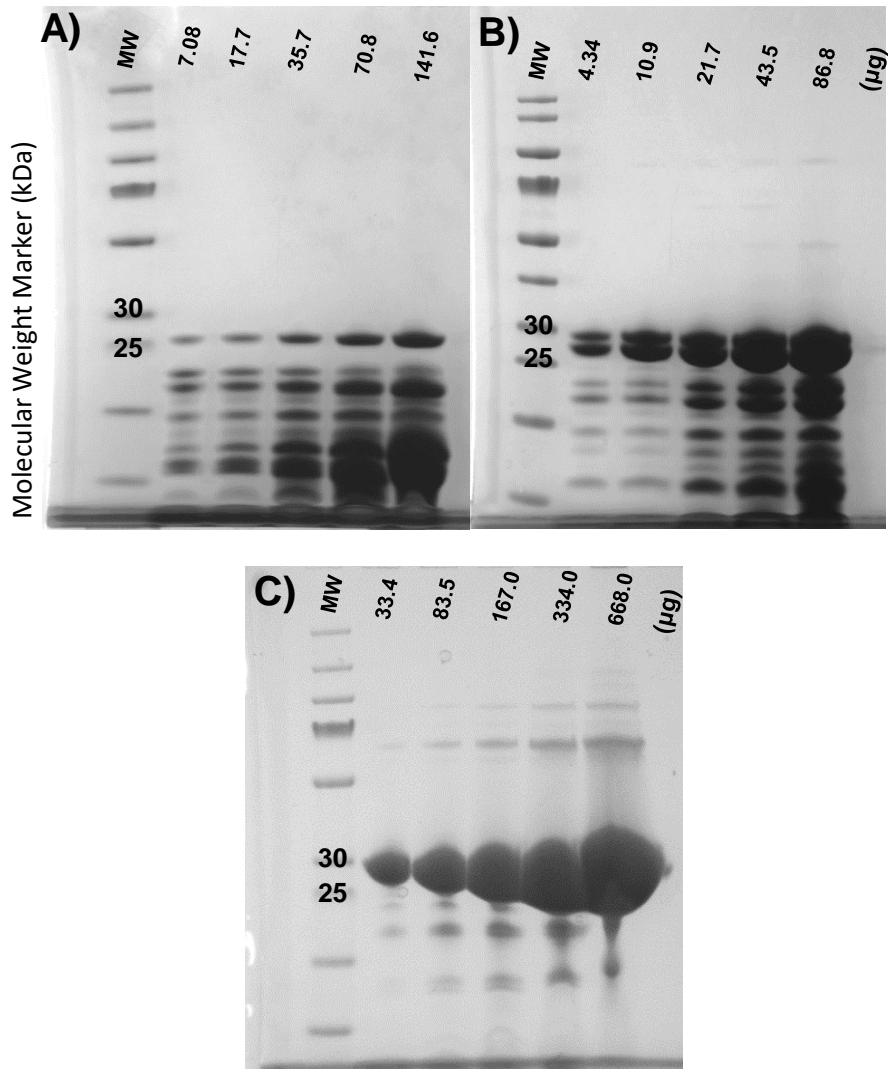
Figure 8. Post-dialysis concentrate of three separate nickel purified AaSPVII Zymogen preparations. A) Dialysis was done in 2 L 20 mM Tris-HCl, pH 7.2 + 10 mM $CaCl_2$ (buffer exchanged twice and set at 4 °C). B) Dialysis was done in 20mM Tris-HCl, pH 7.2 + 10 mM $CaCl_2$ + 2 mM DTT (buffer exchanged twice and set at 4 °C). C) Dialysis was done in 50 mM sodium acetate, pH 5.2 + 2 mM DTT (buffer exchanged twice and set at 4 °C). All samples were concentrated after dialysis and analyzed using SDS-PAGE and Simple blue stained. All concentrations above are in micrograms, µg.

To determine the activation of AaSPVII Zymogen, nickel purification of the AaSPVII was repeated, but after dialysis in sodium acetate buffer, 3.5 mL of the concentrated protein was dialyzed in 20 mM Tris-HCl, pH 7.2 + 10 mM $CaCl_2$. Samples were taken at

different time-points and analyzed via SDS-PAGE and BApNA to determine if the

protease was active prior to buffer exchange (Figure 9). SDS-PAGE showed consistent

intensity of the 24.1 kDa "mature" protease band. This was also supported by the

observed auto-activation and auto-degradation of the protease (Figure 9A), with no

observable differences in BApNA cleavage over time (Figure 9B). It may be possible that

there was not enough BApNA substrate compared to the amount of AaSPVII suggesting

that the protease may prefer to cleave itself and not BApNA.
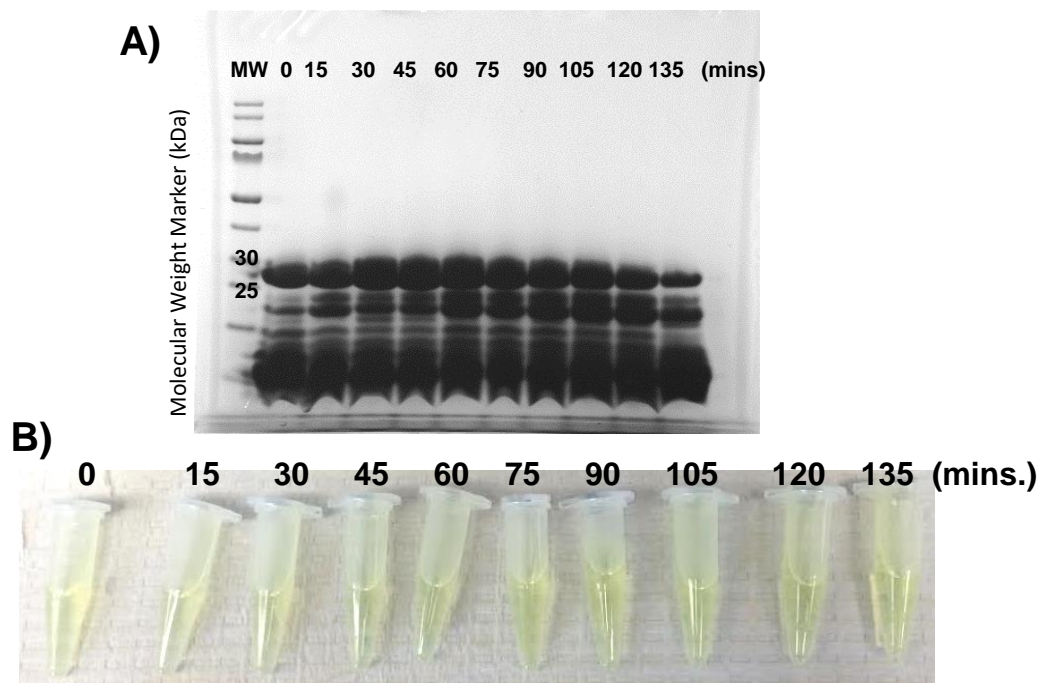
Figure 9. Activation attempt of nickel purified AaSPVII Zymogen by buffer exchange
from sodium acetate pH 5.2 to Tris-HCl pH 7.2. Activation and auto-degradation of the
protease was achieved prior to buffer exchange. A) Band(s) can be seen between the 30
or 25 kDa markers, however, due to the thickness of the band, it may be either the 26.8
no His$_6$-tagged "inactive" or the 24.1 kDa "mature" protein. B) Qualitative BApNA assay
to check trypsin-like activity, which is nearly constant in all samples. Lane numbers
represent time in minutes.

Rascón *et al.* measured trypsin-like activity of AaSPVII and determined initial kinetic parameters using BApNA as the substrate.[67] We wanted to replicate these steady-state kinetic parameters for the nickel purified AaSPVII Zymogen (Figure 8C). We hypothesized that the activation of the protease should be possible when transferring the zymogen into a buffer environment with a physiological pH. Therefore, we set reactions assuming steady state conditions, using 20 mM Tris-HCl, pH 7.2 + 10 mM $CaCl_2$ and BApNA as a substrate, as described in Rascón *et al.* (2011).[67] However, despite adding excess amount of BApNA, the reaction rates were zero (with no detectable activity) and very similar across all concentrations (data not shown).

Due to the autocatalytic nature of AaSPVII, site-directed mutagenesis was performed to introduce point mutations to the wild-type AaSPVII enzyme at the arginine in the propeptide region (Figure 3) and the serine at the active site in order to alter this observed activity. After predicting the position of the arginine (R38) in the propeptide region[67] and the serine (S219) in the active site using the ExPASy Prosite – Scan Prosite Tool, mutations were made to alter the selected amino acids residues into an inert alanine[80] in order to give insight into autocatalysis and prevent auto-activation of AaSPVII. The mutants were induced and grown like wild-type, and as seen in Figure 10, the AaSPVII R38A mutant expression resulted in possible loss of the N-terminal His_6-tag, despite the propeptide region (the region that needs to be cleaved for full activation) being mutated. Figure 10A is the His-stain gel of samples collected from the small-scale expression experiment, which shows the disappearance of the N-terminally His_6-tagged AaSPVII R38A mutant over time and appearance of a second band. It is important to note that

there was no third band present as seen in Figure 5 for wild-type AaSPVII (the third band

being the active mature form). A preliminary BApNA assay repeated for these mutant

samples revealed that there was no trypsin-like activity (data not shown). We

hypothesized that the second species observed in Figure 10A is the no His$_6$-tagged

AaSPVII R38A zymogen mutant. As for the AaSPVII S219A mutant, expression resulted

in a single band (Figure 11). This reveals the importance of the serine residue at the

active site for AaSPVII function, as auto-activation and auto-digestion was prevented.
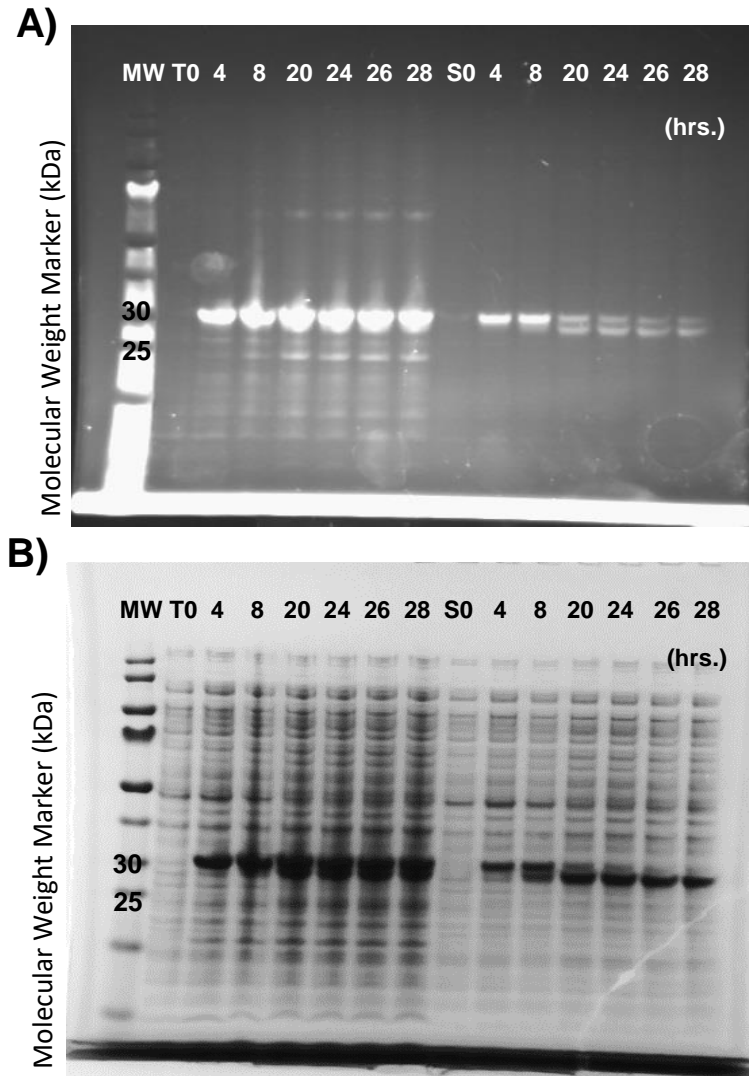
Figure 10. SDS-PAGE analysis of solubly expressed AaSPVII (No Leader) R38A Z/pET28a transformed in T7 SHuffle competent cells. Bacteria were grown at 23 °C, induced with 0.1 mM IPTG and allowed to grow for 28 hours. Two prominent species can be observed in both gels and may be suggesting auto-activation of the protease. Molecular weight of the AaSPVII mutant was calculated to be 28.7 kDa. The no His$_6$-tag "Inactive" protease was calculated to be 26.8 kDa. A) Gel visualized with His-tag in gel stain. The intensity of the N-terminally His$_6$-tagged AaSPVII mutant can be seen fading as expression time increases, which may suggest that the enzyme is losing the N-terminal His$_6$-tag. B) The same gel used in A was stained with SimplyBlue. The dominant expressed AaSPVII mutant species is over-expressed in the T = 4 and 8 hour samples, but as time increases, a new over-expressed species can be clearly seen. These results indicate the possibility of auto-activation, but not into its fully active mature state as seen in the wild-type AaSPVII (Figure 5).
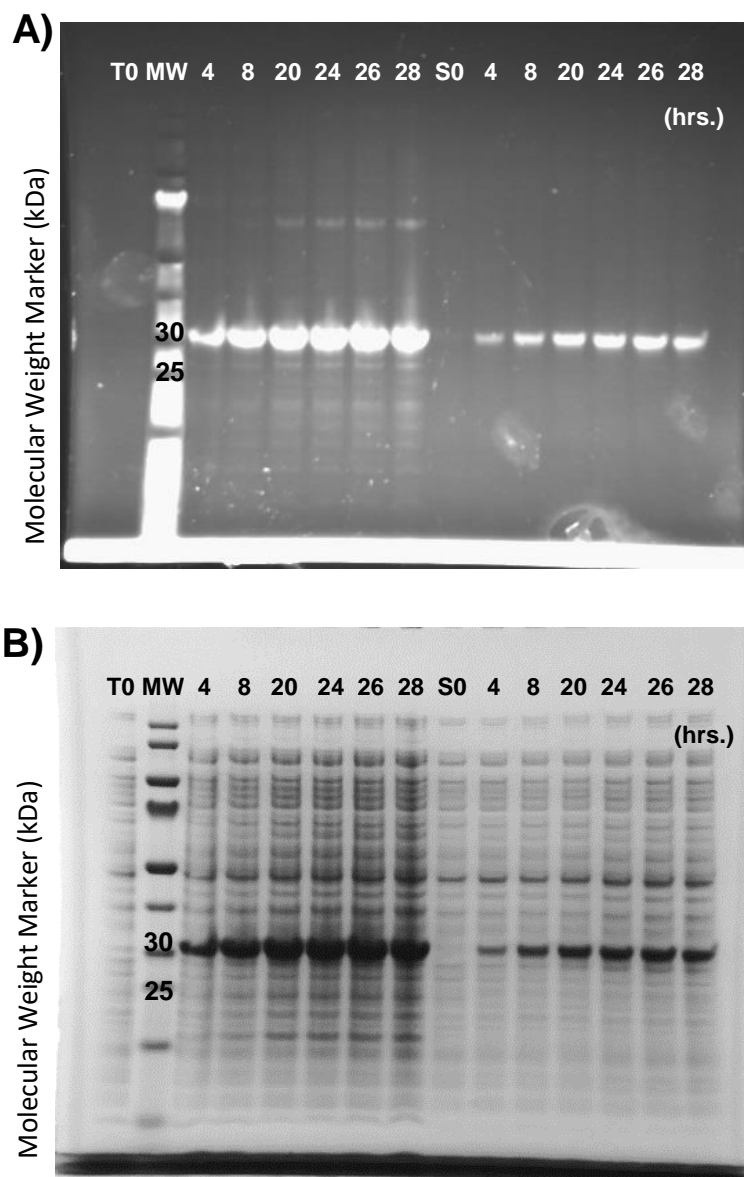
Figure 11. SDS-PAGE analysis of solubly expressed AaSPVII (No Leader) S219A Z/pET28a transformed in T7 SHuffle competent cells. Bacteria were grown at 23 °C, induced with 0.1 mM IPTG and allowed to grow for 28 hours. Only one prominent species can be observed in both gels suggesting no auto-activation of the protease. Molecular weight of the AaSPVII mutant was calculated to be 28.7 kDa. A) Gel visualized with His-tag in gel stain. The intensity of the AaSPVII mutant band can be seen increasing as expression time increase, suggesting no loss of the N-terminal His_6-tag. B) The same gel used in A was stained with SimplyBlue. The expressed AaSPVII mutant species increases over time. These results indicate the presence of the mutant AaSPVII Zymogen that is unable to undergo autocatalysis.

Discussion

Previous expression experiments failed to yield sufficient soluble wild-type zymogen and mature AaSPVII protein for biochemical characterization.[67] One of the major issues was the improper formation of the disulfide bonds due to the reducing environment of the BL21 (DE3) and Rosetta 2 (DE3) *E. coli* cells.[67,71,74,76] Despite the fact that these cells are designed with the T7 expression system in order to enhance protein expression using pET vectors, the environmental cellular conditions are not optimal for proper expression and folding of midgut serine proteases. By switching to SHuffle T7 competent cells, which are derivatives of BL21 (DE3), the cells are more compatible for expression of mosquito proteases due to a more oxidizing environment.[76] With these cells, we were able to successfully express soluble inactive and active forms of AaSPVII (Figure 4 and 5). In addition, the leader sequence (Figure 3) was removed from the gene, as it has been shown that the deletion of the putative signal sequence can aid in increasing protein expression by promoting thermodynamic stability and decreasing propensity for aggregation.[72,73]

The T7 SHuffle *E. coli* cells were able to provide the necessary environment and the DsbC protein to aid in disulfide bond formation and proper protein folding. This change led to the unexpected auto-activation of AaSPVII (Figure 5 and 6), which was not reported by Rascón *et al.* (2011).[67] For their studies, an AaSPVII construct was engineered with a pseudo peptide region containing an enterokinase cleavage site (AaSPVII-Mek).[67] According to the authors, this construct was designed to improve expression and to allow activation *in vitro* by using commercially available enterokinase (Figure 3)*,* as it is unknown whether this enzyme and other midgut proteases are

autocatalytic.[67] From the growth experiments of AaSPVII (No Leader) Zymogen/pET28a

in SHuffle T7 cells, three distinct species can be observed. A visible band near the 30

kDa MW marker (indicative of the zymogen AaSPVII species ~28.7 kDa) is initially

present, followed by two other protein species appearing over time after induction. At the

8th hour after induction, the presence of a second species at around 26.8 kDa might be

indicative of the no His_6-tag "inactive" AaSPVII Zymogen. Several supportive

experiments were carefully analyzed to arrive at this hypothesis. As seen in Figure 5B,

the His-stain clearly shows the disappearance of the His_6-tagged AaSPVII Zymogen

suggesting that the N-terminal poly His_6-tag may be getting cleaved through possible

autocatalysis of AaSPVII. However, full activation is not immediately observed. A

second protein species in Figure 5B, which seems to be around 26.8 kDa, might be the

zymogen without the N-terminal His_6-tag. This is based on the presence of the thrombin

cleavage site that is located between the N-terminal poly His_6-tag and the NdeI restriction

site on the pET28a vector.[81,82] The thrombin site was included in the vector because the

thrombin enzyme is known for its high proteolytic specificity, and therefore is a valuable

biochemical tool to be exploited for cleaving off the fusion/purification tags that are

normally included in linker regions of recombinant fusion protein constructs.[83] Thrombin

recognizes the cleavage site Leu-Val-Pro-Arg-Gly-Ser where it selectively cleaves

between the Arg and Gly residues, even with the proline residue next to the Arg site.[83]

Because AaSPVII has been identified as a classical trypsin[66,67], and because, trypsins do

not readily cleave peptides if a proline precedes an Arg residue[69,70], we were surprised

that this second species might be the product of cleavage at the thrombin cleavage site.

Subsequently, in collaboration with Prof. Anthony O'Donoghue at the University of California of San Diego, nickel purified mature AaSPVII was tested against 7-amino-4-methylcoumarin (AMC)-based peptide substrates to gain initial insight into AaSPVII substrate specificity. These AMC-based substrates are important fluorogenic compounds used to probe proteolytic activity and specificity of proteases with unknown functions.[84] For these experiments, 2 nM AaSPVII was incubated with 20 μM of each substrate (in 20 mM Tris-HCl. pH 7.4 + 100 mM NaCl + 10 mM $CaCl_2$ + 0.01% Tween-20) and the maximal velocity (relative fluorescence units per second) measured and reported (Figure 12). Surprisingly, using one of the AMC-based substrates (Boc-Ala-Gly-Pro-Arg-AMC) that is similar to the thrombin cleavage site, AaSPVII showed higher activity than both AaSPVI (another midgut late phase protease) and the porcine trypsin control (Figure 12). Additionally, alanine-scanning mutagenesis was used to probe its catalytic role and function[80] by mutating R38 in the propeptide cleavage site and S219 in the active site to an inert alanine. The auto-activation of AaSPVII was completely prevented due to the S219A mutation (Figure 11) and, from preliminary BApNA assays, no trypsin-like activity was observed (data not shown). Further evidence that the inactive wild-type AaSPVII Zymogen is cleaving the N-terminal His_6-tag at the thrombin cleavage site comes from the R38A point mutation. This mutation was used for the expression of both the zymogen with the His_6-tag and zymogen without the His_6-tag, inhibiting full auto-activation as no trypsin-like activity was observed in the *in vitro* spectrophotometric BApNA assays (Figure 10). Initial experiments have established AaSPVII to have trypsin-like activity, but unlike classical trypsins, this enzyme is able to cleave at an Arg

residue that is preceded by a Pro. In addition, there is a possibility that the "inactive

zymogen" may not truly be inactive if the R38A mutant resulted in loss of the His$_6$-tag.

The S219A mutant did not lose the tag indicating that the bacterial cells are not

producing an enzyme with this type of specificity.
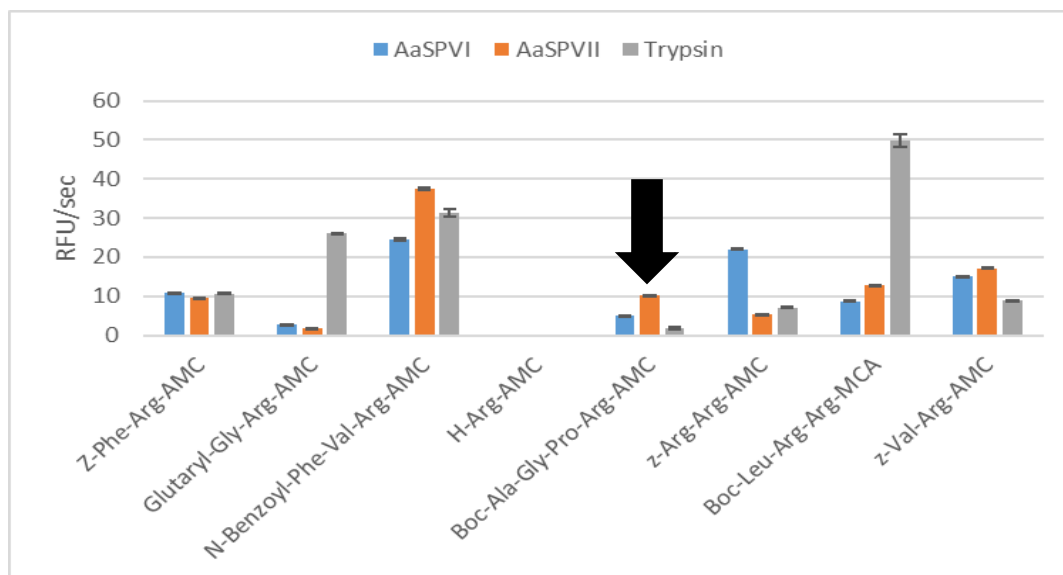
Figure 12. Fluorogenic assay of 7-amino-4-methylcoumarin (AMC)-based substrates
with nickel purified mature AaSPVI and AaSPVII. A variety of AMC-based substrates
were used to randomly test for activity and specificities of each protease. A working
stock of 10 μM protease was made in 20 mM citrate phosphate, pH 5.0 + 100 mM NaCl.
The 10 μM was diluted to 2 nM in 20 mM Tris-HCl, pH 7.4 + 100 mM NaCl + 10 mM
CaCl$_2$ + 0.01% Tween-20 assay buffer. A 25 μL volume of the solution reacted with 25
μL of each AMC-based substrate in triplicate. Porcine trypsin was used as a control
trypsin. A 1 nM solution of AaSPVI and AaSPVII had approximately the same cleaving
efficiency as 40 nM solution of porcine trypsin when Z-Phe-Arg-AMC was used as the
substrate. The data is reported as the maximal velocity (RFU/sec) measured over 20
points in 60 seconds for each reaction.

The third band observed in Figure 5 is hypothesized to be the active "mature"

AaSPVII, as it correlated with the calculated 24.1 kDa molecular weight of the active

species. In addition, the presence of the band (after induced expression) correlated in time

with BApNA activity, especially reaching maximal activity at the 18th hour post-

induction (Figure 5 and 6). These results were not too surprising, especially since

trypsinogens in vertebrates have an enteropeptidase site (a region of repeating aspartic

acid residues, followed by a lysine) at the propeptide region that allows enteropeptidases

to regulate the activation of trypsin.[51] However, trypsinogens in *Ae. aegypti* lack these

repeated residues before the arginine, and therefore active trypsin may be produced from

auto-activation. This is also supported by the fact that there has been no confirmed

enteropeptidases found in the mosquito.[51]

Multiple purification attempts of N-terminally His_6-tagged AaSPVII Zymogen were

conducted to determine the effect of buffer conditions on the auto-activation of AaSPVII.

Previously, only the artificially engineered Mek form of the midgut serine proteases was

expressed in insoluble form and purified using denaturation/refolding techniques.[67] For

this thesis work, cell paste containing solubly expressed N-terminally His_6-tagged

AaSPVII Zymogen was collected at the 5th hour post-induction to obtain only the intact

zymogen. A HisTrap FF nickel column was used due to histidine's strong interaction

with the immobilized transition metal ion matrices where histidine from the recombinant

enzyme can readily form coordination bonds.[85] This is beneficial since the N-terminal

poly-His_6-tag in the pET28a vector can bind to the nickel matrix and then be eluted using

an increasing concentration of imidazole.[85] Initial attempts used 20 mM Tris-HCl, pH 7.2

+ 10 mM $CaCl_2$ as the dialysis buffer after nickel purification, despite the fact the

optimum pH for trypsin activity is near physiological pH.[69,70,79] After concentrating and

SDS-PAGE analysis, the protein was revealed to be heavily auto-degraded and still active

(Figure 8A). Addition of 2 mM DTT slowed down autocatalysis and auto-degradation

(Figure 8B), as a mild concentration of DTT has shown to help create a mild reducing environment to partially misfold the protein of interest and minimize protease activity.[67] Auto-activation and auto-digestion of the protein was not observed after dialysis of the pooled fractions in 50 mM sodium acetate, pH 5.2 + 2 mM DTT. This is thought to be due to the pH change, going from 7.2 to 5.2 (Figure 8C). It is important to note that regardless of DTT being added to the purification buffers, the isolated fractions prior to dialysis were all observed to be intact, based on SDS-PAGE analysis.

The auto-activation of AaSPVII may be concentration-dependent because fractions collected from nickel purification shows intact AaSPVII Zymogen, but only after we concentrate the pooled fractions do we see auto-activation and digestion (Figure 8A). This is not too surprising because trypsin activation and auto-degradation is very sensitive to the concentration of the active form of trypsin present.[86] After activation of trypsin, the active form can immediately activate all other trypsinogens, and therefore create an escalating chain reaction.[86] This concentration-dependent auto-activation, may explain why the *in vitro* spectrophotometric assay using BApNA and the AaSPVII Zymogen form resulted in insignificant activity of the trypsin (data not shown). The nanomolar concentration of enzyme was too low and failed to undergo autocatalysis.

Since we hypothesized that AaSPVII was concentration-dependent, we decided to concentrate the isolated zymogen in 50 mM sodium acetate, pH 5.2 + 2 mM DTT and attempt to activate it by buffer exchanging into 20 mM Tris-HCl, pH 7.2 + 10 mM $CaCl_2$. Unfortunately, as seen in Figure 9, partial auto-activation occurred prior to buffer exchange. The auto-activation was most likely initiated during the concentration step, but

it was discovered that the sodium acetate stock used for this experiment was at pH 9.4 and not pH 5.2. The pH value is important because trypsins have broad maximal activity at pH ranges from 7 to 9[79], therefore, this allowed for auto-activation and auto-degradation to occur.

The N-terminal His$_6$-tagged AaSPVII protein was solubly expressed successfully after using the SHuffle T7 bacterial cells as confirmed by SDS-PAGE. However, the overexpression of this protein in bacteria led to auto-activation and auto-degradation over time after induction as observed from both SDS-PAGE and BApNA activity assays (Figure 5 and 6). Even after expressing and nickel purifying the N-terminally His$_6$-tagged AaSPVII Zymogen, the enzyme auto-activated after dialyzing and concentrating in the optimal pH 7.2 buffer conditions. To prevent auto-activation of the enzyme, mutation at the Arg propeptide region (Figure 3) to an Ala resulted in the loss of the N-terminal His$_6$-tag, which indicates the possibility that the "inactive zymogen" may not be truly inactive. The bacterial cells used to express the wild-type and mutant forms of AaSPVII do not express an enzyme that cleaves at the thrombin cleavage site otherwise the S219A mutant would have more than one band observed during expression (Figure 11). Based on initial pepride fluorgenic AMC substrate activity assays, AaSPVII was shown to cleave arginine despite having a proline preceding it. In fact, the Pro-Arg-AMC substrate (Figure 12) is similar to the thrombin cleavage site on the N-terminal His$_6$-tag linker on all of the AaSPVII constructs, indicating a preference for this type of sequence. Despite auto-activation and loss of the N-terminal His$_6$-tag of wild-type and R38A AaSPVII, we successfully isolated intact N-terminally His$_6$-tagged AaSPVII (No Leader) Zymogen

using 50 nM sodium acetate, pH 5.2 + 2 mM DTT. Once the optimal activation

conditions are found, the N-terminal His$_6$-tagged AaSPVII (No Leader) Zymogen will be

activated, purified, and used to obtain kinetic parameters using BApNA as the substrate

and compared to previously published results.[67]

Chapter III

Initial Biochemical Characterization of C-terminally His$_6$-tagged AaSPVII

Introduction

An unexpected and ultimately great result came from the expression and

purification of wild-type AaSPVII Zymogen; under the right buffer conditions, the

enzyme is autocatalytic (Chapter II). We were concerned that purification of the N-

terminally His$_6$-tagged AaSPVII Zymogen would be deemed unsuccessful due to loss of

the His$_6$-tag. A possible solution to this problem is to move the N-terminal His$_6$-tag to the

C-terminus[82,87,88] thereby allowing us to purify the mature (activated) form of the

protease if isolation and activation of the zymogen is not possible. Placement of these

poly-histidine affinity tags may alter the expression or solubility of the recombinantly

expressed enzyme.[89,90] Since not all recombinant enzymes behave the same upon addition

of the His$_6$-tag (at either termini), optimal placement of the affinity tag may be protein-

specific.[89-93] For example, when the cleavable poly His$_6$-tag was placed on the C-terminal

end of the human serum transferrin enzyme, the enzyme's ability to release iron at

various pH conditions changed dramatically, while attachment at the N-terminus had no

effect on the release rate.[89] Another example of His$_6$-tag having an effect on the

recombinantly expressed enzyme was a study on the petrobactin biosynthesis gene *asbF*

from *Bacillus anthracis*, which was shown to have the N-terminally His$_6$-tag partially

buried. Thus, the His$_6$-tag may have interfered with protein folding and was not optimal

for crystallization.[93] Therefore, the pursuit of a C-terminal His$_6$-tagged AaSPVII

Zymogen as a novel construct was initiated, and we set out to investigate if this

attachment would alter the activity and/or expression of the protease and to compare the results with the solubly expressed N-terminal His$_6$-tagged AaSPVII Zymogen.

In this chapter, the cloning, recombinant expression, purification, and the initial biochemical characterization of the C-terminal His$_6$-tagged AaSPVII Zymogen construct using the pET29b vector[82] will be discussed and its contrasted to AaSPVII in the pET28a vector.

Methods

Preparation of C-terminal His$_6$-tagged AaSPVII Construct

The engineering of the C-terminal poly His$_6$-tag construct was made by designing primers to facilitate the cloning of the AaSPVII Zymogen gene into the pET29b vector. The previous forward primer for wild-type AaSPVII was used (Table 6). The reverse primer was designed to remove the stop codons at the end of the gene and include the XhoI (CTCGAG) restriction cleavage site and a poly A-tail (Table 6). Cloning of AaSPVII Zymogen into the pET29b was done as described in Chapter II with the following modifications: 1) Instead of HindIII, used 1 μL of XhoI (ThermoScientific, Cat. #ER0691) and 2) Ethanol precipitation was not done, instead, the digested AaSPVII PCR product and the digested/desphosphorylated pET29b vector were co-purified using the QIAquick Gel Extraction Kit and then ligated together overnight at 4 °C.

Large-Scale Bacterial Overexpression and Purification

The AaSPVII (No Leader) Z/pET29b plasmid construct was expressed as described previously in Chapter II with the following modifications: 1) The cells were grown to

OD$_{600}$ ~0.5-0.8 at 30 °C, then were induced with 0.1 mM IPTG and grown at 12 °C for

44 hours, and 2) 5 mM DTT was added to Buffer A and B for FPLC purification.

In Vitro BApNA Spectrophotometric Assay of Recombinant AaSPVII Z/pET29b

Nickel purified mature AaSPVII was used for a Michaelis-Menten steady-state

kinetic experiment using the synthetic chromogenic substrate BApNA as described in

Chapter II.

Table 6. Designed Primers used for PCR Amplification and Cloning of C-terminal His$_6$-tagged Wild-Type AaSPVII Zymogen. Primers were purchased from Elim Biopharmaceuticals, Inc.

Protease	Primer	Primer Sequence	T$_m$ (°C)
AaSPVII - Zymogen	AaSPVII-Zym-pET-Fwd	5'–**AAAAA**<u>CAT</u>*ATG*CTATCAACCGGATTCCATCCGC– 3'	65.4
AaSPVII - Zymogen	AaSPVII-C term His-Rev	5'–**AAAAA**<u>CTCGAG</u>AACTCCACTGACTTCGGCCACC–3'	65.0

Poly-A tail used is in **BOLD**.
Restriction sites used for cloning of the genes of interest are <u>underlined</u>.
Start codons is *italicized*.
Annealing temperature calculated using NetPrimer.

Results

The cloning of AaSPVII into the pET29b vector system was done not only to purify

the mature protease in the event of autocatalysis, but to also investigate if affinity tag

placement may affect recombinant protein expression. The soluble expression of N-

terminally His$_6$-tagged AaSPVII (No Leader) Zymogen was done at 23 °C (Chapter I).

However, recombinant expression of the C-terminally His$_6$-tagged AaSPVII (No Leader)

Zymogen was done at 12 °C to allow for proper folding since we are unsure if changing

His$_6$-tag placement will alter the properties of the enzyme.[94] This lower growth

temperature should inhibit or slow down the auto-activation process, as was seen for the

AaSPVII Z/pET28a construct. As seen in Figure 13, after induction with 0.1 mM IPTG,

the AaSPVII Zymogen was slowly expressed over 44 hours. The molecular weight of

AaSPVII Z/pET29b was calculated to be 27.6 kDa and the "active" no propeptide

protease is 25.2 kDa. And based on the figure, the majority of the expressed species is the

zymogen form (27.6 kDa), running near the 30 kDa marker, as desired. With the intact C-

terminal poly His-tag on AaSPVII, nickel purification was pursued.
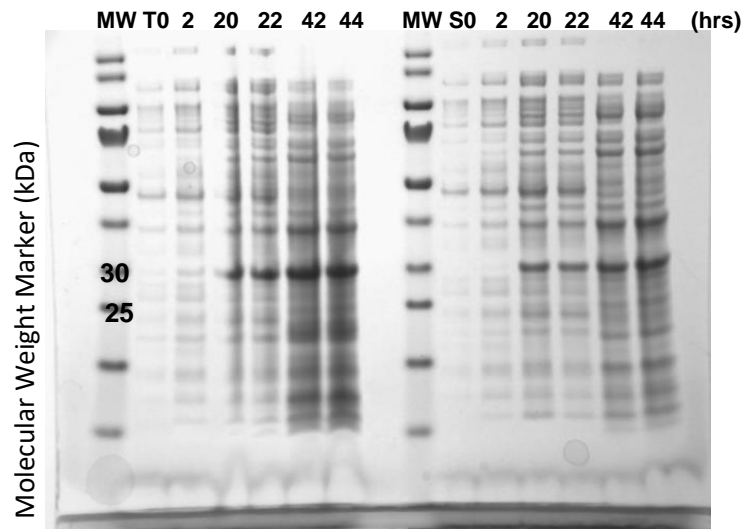
Figure 13. SDS-PAGE analysis of solubly expressed AaSPVII (No Leader) Z/pET29b
transformed in T7 SHuffle competent cells. Bacteria were grown at 12 °C, induced with
0.1 mM IPTG, and allowed to grow for 44 hours. Only one prominent species can be
observed in the gel, which may suggest no auto-activation of the protease. Molecular
weight of the AaSPVII Zymogen protein was calculated to be 27.6 kDa. The 'active'
AaSPVII was calculated to be 25.2 kDa. Gel visualization with SimplyBlue stain.

Purification of the C-terminally His$_6$-tagged AaSPVII (No Leader) Zymogen was

repeated as performed with the wild-type N-terminally His$_6$-tagged AaSPVII Zymogen

described in Chapter II. Surprisingly, after dialysis in 50 mM sodium acetate, pH 5.2 + 2

mM DTT and concentration, the mature form of the protease was isolated. This was based on SDS-PAGE analysis and the presence of what seems to be auto-degradation products below the MW of the "active" AaSPVII protease (25.2 kDa) (Figures 14).
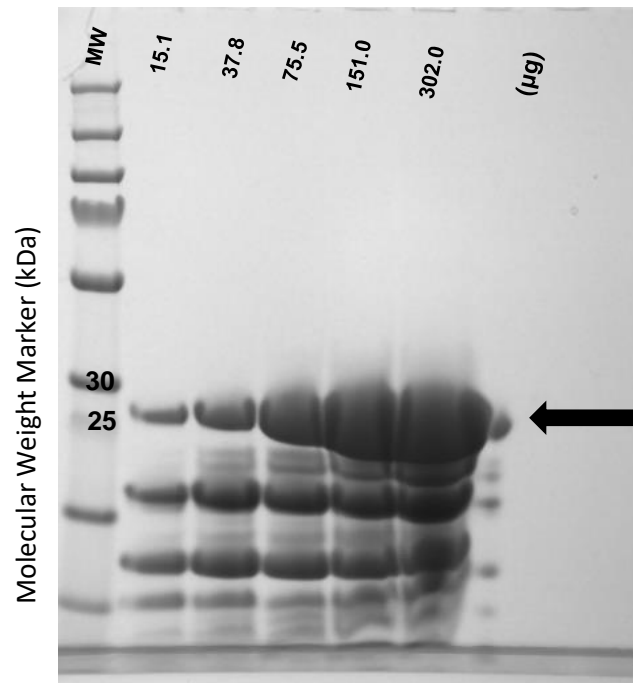
Figure 14. SDS-PAGE analysis of nickel purified "active" mature AaSPVII. The mature protease was the final product observed after dialysis and concentration. The expected molecular weight of mature AaSPVII is 25.2 kDa. Several auto-degradation products are observed.

To determine if this purified AaSPVII is indeed active, trypsin-like activity was tested using BApNA as the substrate.[67] The concentration of enzyme was selected to be 50 nM for the lower BApNA concentrations and 100 nM for the higher concentrations[67] in order to maintain a 10-fold lower enzyme concentration compared to the lowest substrate concentration used[95]. Therefore, to obtain kinetic parameters, 50 nM active AaSPVII was used at each BapNA concentration in order to produce the Michaelis-Menten plot in

Figure 15. The kinetic parameters obtained were comparable to the results published by

Rascón *et al.*[67] (Table 2 and Figure 15).
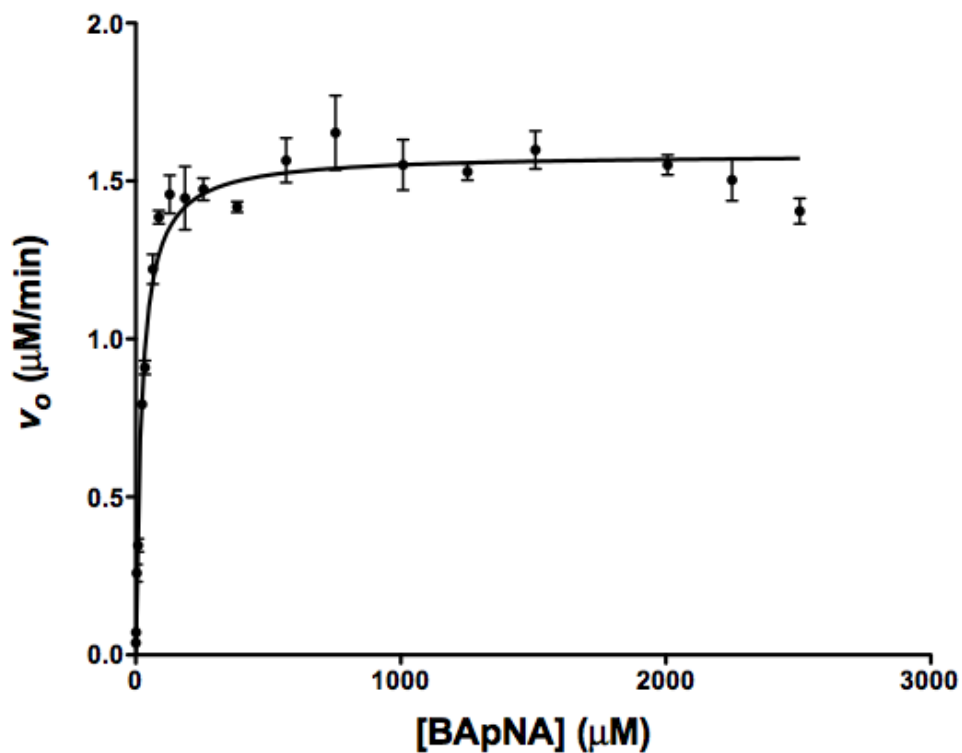
Protease*	$K_{M (BApNA)}$ (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (mM^{-1} s^{-1})	Specific Activity (μmol min^{-1} (mg of protein)$^{-1}$)
AaSPVII	22.0±1.9	0.53	24.1	1.31

*Reaction conditions: 20 mM Tris-HCl, pH 7.2 + 10 mM CaCl$_2$ at 23 °C.
All values were calculated from the mean of three reactions.
The produced Michaelis-Menten plot had an R^2 value of 0.9639.

Figure 15. Steady-state kinetic parameters of mature AaSPVII using BApNA as a substrate. AaSPVII at 50nM was reacted with an excess amount of BApNA to create the Michaelis-Menten plot. The kinetic parameters were determined using GraphPad prism.

Discussion

The C-terminally His$_6$-tagged protease (AaSPVII Z/pET29b) construct was prepared to help alleviate the purification problem, and to investigate if the placement of the affinity tag would alter expression, folding, and/or activity.[89,90,93,96]

Proper folding and soluble recombinant protease expression in bacteria can vary due to the placement of the poly-histidine affinity tag (as mentioned in the Introduction), but also with growth temperature.[85,94] As seen in Figures 5 and 13, the C-terminally His$_6$-tagged AaSPVII Zymogen was fully expressed until the 44th hour (post-induction) at 12 °C, while the N-terminally His$_6$-tagged AaSPVII Zymogen was expressed earlier at the 6th hour (post-induction) at 23 °C. By lowering the growth temperature of the bacterial cells, the metabolism and protein expression machinery of the host cell is slowed to allow the protease to fold properly[74], but to also slow the auto-activation activity of the enzyme. Mammalian proteases in general have optimal activity at 37 °C[97], and similarly these mosquito proteases are expected to function at the elevated temperature of the blood meal. However, since blood meal digestion is a relatively long process (between 30 to 40 hours), the temperature of the mosquito and the blood meal decreases to ambient levels[98] and protease activity is unaffected. So, we hypothesized that if the temperature is reduced below the optimal ambient temperature (typically 23 °C), proteolytic activity should be minimized. Therefore, for the C-terminally His$_6$-tagged AaSPVII Zymogen recombinant expression, the temperature was dropped to 12 °C. The lower temperature approach led to successful overexpression, but more importantly, slowed down auto-activation of the C-terminally His$_6$-tagged AaSPVII for purification

(Figure 13). Comparing this expression with AaSPVII Zymogen/pET28a construct grown at 23 °C, the ideal proteolytic temperature of the midgut proteases, auto-autoactivation was clearly observed (Figure 5). Once the C-terminally His$_6$-tagged AaSPVII Zymogen was expressed, the enzyme was nickel purified and dialyzed in sodium acetate buffer as was done with the N-terminally His$_6$-tagged AaSPVII Zymogen. Unfortunately, after concentrating the dialyzed C-terminally His$_6$-tagged AaSPVII Zymogen, the enzyme had autocatalyzed to the "mature" protease. This was attributed to the sodium acetate used during dialysis. The sodium acetate buffer was found to have a pH of 9.4, the pH of the salt dissolved in water before adjusting to pH 5.2 with acetic acid. Proper preparation of the buffer was not followed, leading to the result observed.

As mentioned, Rascón and associates determined steady-state kinetic parameters for active AaSPVII isolated from inclusion bodies (Table 2).[67] This *in vitro* spectrophotometric assay using BApNA was repeated with purified N-terminally His$_6$-tagged AaSPVII Zymogen, but the assay failed to show any significant activity (data not shown). However, with the nickel purified C-terminally His$_6$-tagged AaSPVII mature protease (Figure 14), the assays were repeated and are comparable to the steady-state kinetic parameters published.[67] These results suggest that the AaSPVII enzyme in these experiments was solubly expressed and purified successfully, which had never been repeated since the AaSPVII Rascón *et al.* (2011) work.[67] More importantly, however, placement of the His$_6$-tag on the C-terminus did not have an effect on AaSPVII expression or activity, which will allow for further protease purification optimization and enzyme specificity assays.

Chapter IV

Summary and Future Directions

The *Ae. aegypti* mosquito is an efficient biological vector, transmitting several viruses when the female mosquito acquires a blood meal.[34,36,47] Digestion of blood meal proteins is required to complete the gonotrophic cycle[35,36], and is highly dependent on midgut proteases. Therefore, it may be possible to inhibit these digestive enzymes with small molecule inhibitors to prevent the completion of the gonotrophic cycle, and therefore, potentially reduce the mosquito population and viral transmission.[59,67] As seen in Figure 2A, specific digestive enzymes were suppressed and were shown to have varying effects on the overall fecundity of the female *Ae. aegypti* mosquito. Additionally, diluted blood meals were fed to mosquitoes, and a significant reduction of eggs was observed (Figure 2B), thus demonstrating how crucial protein blood meal digestion is for the egg laying process.[59] With the reemergence of the YFV[30,31], DENV[38,40], CHIKV[32,40], and ZIKV[15,40], and the need for better control strategies, investigating the midgut serine proteases involved in the digestive process may prove to be a potential vector control strategy.[59,67]

The work presented in this thesis serves to re-evaluate the potential role of the late phase midgut serine protease AaSPVII that was previously characterized by Rascón *et al.*[67]. In this study, soluble expression of the AaSPVII Zymogen was not sufficient and had to be engineered with an artificial propeptide region containing an enterokinase site. Even then, however, expression was insoluble and denaturation/refolding techniques were utilized to obtain purified soluble protein.[67] Interestingly, the isolated AaSPVII was shown to be biochemically different from AaSPVI despite being homologous to each

other.[59,66] AaSPVI was shown to contribute to the majority of late-phase trypsin-like activity in the midgut, while AaSPVII had no significant effect[59], despite both proteases cleaving the chromogenic trypsin-like substrate BApNA *in vitro*[67] (Table 2). This raises the interesting question as to the specificity and functionality of AaSPVII, a question that warrants further studies to elucidate the role of AaSPVII in blood meal protein digestion.

Although Rascón and associates were successful at obtaining soluble AaSPVII using denaturation/refolding techniques, the refolding process is rigorous and has the potential to reduce protein yield.[67,75] Two of the main issues regarding low protein yield are the presence of the leader (signal peptide) sequence (Figure 3) and the reducing environment of the competent bacterial cells used for protein expression, which prevents proper disulfide formation.[76] Removal of the leader sequence has been shown to increase the yield of desired proteins.[72,73] By designing specific leaderless primers for AaSPVII to facilitate cloning into the pET28a and pET29b vectors[82], and by using bacterial cells with a more oxidizing cytoplasm[76], we were able to successfully express and purify the zymogen and active mature forms of AaSPVII in soluble form (Chapters II and III).

The first construct (AaSPVII (No Leader) Z/pET28a) led to AaSPVII autocatalysis during expression with maximal trypsin-like activity detected at the 18th hour (post-induction) (Figures 5 and 6). However, we did manage to purify the enzyme using a nickel column by collecting the cell paste at 5-6 hours post-induction (Figure 5) and altering the buffer conditions needed to prevent auto-activation and auto-digestion (Figure 8). AaSPVII has been classified as a classical trypsin enzyme[59,66,67], but trypsins cannot cleave at an Arg residue with a Pro residue preceding it.[69,70] However, AaSPVII

was shown to cleave an AMC fluorogenic substrate that contains a Pro-Arg sequence (Figure 12 and Chapter II). This proved to be slightly problematic since a thrombin cleavage site (Leu-Val-Pro-Arg-Gly) is engineered into the pET28a vector to facilitate removal of the His_6-tag. AaSPVII was shown to prefer the thrombin cleavage site (Figure 5), and during expression the enzyme was cleaving the His_6-tag, which would not allow for nickel purification.

As an alternative, AaSPVII was cloned into the pET29b vector to yield a C-terminally His_6-tagged enzyme in order to purify either the zymogen or mature form of the protease. Purification of the C-terminally His_6-tagged AaSPVII resulted in isolation of the mature "active" form of AaSPVII (data not shown), even in the presence of sodium acetate, pH 5.2 and the reducing agent DTT. However, it was later discovered that the sodium acetate buffer was at a pH of 9.4 because acetic acid was not added to lower the pH. We hypothesized that the N-terminally His_6-tagged enzyme autocatalyzes only after removal of imidazole, NaCl, and DTT (dialysis step) and concentrated at the optimal pH conditions for activity (pH 7.2). The autocatalytic nature of the C-terminally His_6-tagged AaSPVII has not been fully investigated. At this point, we are unsure if the enzyme is expressed as the active form or the zymogen form, but the final isolated product (after cell lysis, centrifugation, purification, dialysis, and concentration) is fully active. This conclusion is based on the steady-state kinetic parameters obtained from solubly purified AaSPVII, which is comparable to published results[67] (Table 2 and Figure 15). In addition, the expression and purification approach described in this thesis has led to a much more efficient process of obtaining active AaSPVII enzyme.

To better understand the autocatalytic mechanism of AaSPVII, we can focus on the N-terminally His$_6$-tagged AaSPVII Zymogen, which was successfully purified in its inactive form (Figure 8C). A pH profile using various pH buffers (ranging from pH 4 to 10) can be set to determine the best conditions for auto-activation and optimal protease activity. Once the mechanism of activation is determined, we can then compare the kinetic parameters between this construct and the active C-terminally His$_6$-tagged AaSPVII protease, as well as to the Rascón et al. published data shown in Table 2.[67]

The work described in this thesis has laid the foundation to further investigate and biochemically study not only AaSPVII, but other midgut serine proteases as well. The specificity of AaSPVII is still unknown. However, in collaboration with Prof. Anthony O'Donoghue from UCSD, we hope to utilize the Multiple Substrate Profiling by Mass Spectroscopy technique to determine the proteolytic signature and specificity of AaSPVII. This technique is suitable for the identification of substrate specificity of exo- and endopeptidases.[99] Therefore, as an immediate goal, we need to optimize the purification scheme to obtain homogeneous AaSPVII Zymogen and the mature active form for these studies. In addition, to further characterize AaSPVII, x-ray crystallography can be used to further understand the relationship between the three-dimensional structure of the enzyme and its biological activity. More importantly, these structural studies can then serve to help design inhibitors if these proteases prove to be suitable targets for vector control. Lastly, once the roles of AaSPVII and other midgut proteases in the blood meal protein digestion process are fully understood, it might be worth investigating the interaction of the enzymes with the viruses transmitted by the *Ae.*

aegypti mosquito. The first site of viral infection and replication in the mosquito is the midgut.[100] Brackney and associates examined a knockdown of AaSPVI *in vivo* which resulted in an increase in infectivity of DENV-2 in the mosquito, suggesting that certain midgut serine proteases may play a role in regulating viral infectivity.[101] Before mosquito-selective inhibitors can be formulated, further biochemical studies on these midgut serine proteases have to be completed in order to elucidate their specific roles in blood meal protein digestion, but also to determine if there is a host protease interaction network that increases viral infectivity in the mosquito.

References

(1) Christophers, S. R. Cambridge University Press: Great Britain, **1960**.

(2) Severson, D. W.; Brown, S. E.; Knudson, D. L. *Annu. Rev. Entomol.* **2001**, *46*, 183-219.

(3) Tabachnick, W. J. *Am. Entomol.* **1991**, *37*, 14-26.

(4) Lounibos, L. P. *Ecol. Entomol.* **1981**, *6*, 129-154.

(5) Lounibos, L. P. *Annu. Rev. Entomol.* **2002**, *47*, 233-266.

(6) Kröpelin, S.; Verschuren, D.; Lézine, A.-M.; Eggermont, H.; Cocquyt, C.; Francus, P.; Cazet, J.-P.; Fagot, M.; Rumes, B.; Russell, J. M.; Darius, F.; Conley, D. J.; Schuster, M.; von Suchodoletz, H.; Engstrom, D. R. **2008**, *320*, 765-768.

(7) Brown, J. E.; Evans, B. R.; Zheng, W.; Obas, V.; Barrera-Martinez, L.; Egizi, A.; Zhao, H.; Caccone, A.; Powell, J. R. *Evolution.* **2014**, *68*, 514-525.

(8) Faull, K. J.; Williams, C. R. *J. Vector Ecol.* **2015**, *40*, 292-300.

(9) Edman, J. D. Takken, W., Scott, T. W., Eds.; Kluwer Academic: Wageningen UR frontis series, Dordrecht, 2003; pp 63-74.

(10) Cardé, R. *Curr. Biol.* **2015**, *25*, R793-R795.

(11) Mccall, P. J.; Eaton, G. *Med. Vet. Entomol.* **2001**, *15*, 197-203.

(12) Braks, M. A. H.; Juliano, S. A.; Lounibos, L. P. *Med. Vet. Entomol.* **2006**, *20*, 53-59.

(13) Farjana, T.; Tuno, N. *J. Med. Entomol.* **2013**, *50*, 838-846.

(14) Nuckols, J. T.; Huang, Y. S.; Higgs, S.; Miller, A. L.; Pyles, R. B.; Spratt, H. M.; Horne, K. M.; Vanlandingham, D. L. *J. Med. Entomol.* **2015**, *52*, 447-451.

(15) Wong, S. S.; Poon, R. W.; Wong, S. C. *J. Formosan Med. Assoc.* **2016**, *115*, 226-242.

(16) Blitvich, B. J.; Firth, A. E. *Viruses* **2015**, *7*, 1927-1959.

(17) Dick, G. W. A.; Kitchen, S. F.; Haddow, A. J. *Trans. R. Soc. Trop. Med. Hyg.* **1952**, *46*, 509-520.

(18) MacNamara, F. N. *Trans. R. Soc. Trop. Med. Hyg.* **1954**, *48*, 139-145.

(19) Simpson, D. I. H. *Trans. R. Soc. Trop. Med. Hyg.* **1964**, *58*, 335-337.

(20) Areas with Zika: Case Counts in the US. *https://www.cdc.gov/zika/geo/united-states.html* (accessed February 15, 2017, 2017).

(21) Lazear, H. M.; Diamond, M. S. *J. Virol.* **2016**, *90*, 4864-4875.

(22) Epidemiological Update: Zika virus infection. *http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&Itemid=270&gid=33296&lang=en* (accessed February 15, 2017).

(23) Guzman, M. G.; Halstead, S. B.; Artsob, H.; Buchy, P.; Farrar, J.; Gubler, D. J.; Hunsperger, E.; Kroeger, A.; Margolis, H. S.; Martinez, E.; Nathan, M. B.; Pelegrino, J. L.; Simmons, C.; Yoksan, S.; Peeling, R. W. *Nat Rev Micro* **2010**, *90*, 4864-4875.

(24) A Global Brief on Vector-Borne Diseases. *World Health Organization.* **2014**, 1-56.

(25) Mustafa, M. S.; Rasotgi, V.; Jain, S.; Gupta, V. *Med. J. Armed Forces India.* **2015**; *71*, 67-70.

(26) Dejnirattisai, W.; Supasa, P.; Wongwiwat, W.; Rouvinski, A.; Barba-Spaeth, G.; Duangchinda, T.; Sakuntabhai, A.; Cao-Lormeau, V.; Malasit, P.; Rey, F. A.; Mongkolsapaya, J.; Screaton, G. R. *Nat. Immunol.* **2016**, *17*, 1102-1108.

(27) Priyamvada, L.; Quicke, K. M.; Hudson, W. H.; Onlamoon, N.; Sewatanon, J.; Edupuganti, S.; Pattanapanyasat, K.; Chokephaibulkit, K.; Mulligan, M. J.; Wilson, P. C.; Ahmed, R.; Suthar, M. S.; Wrammert, J. *Proc. Natl. Acad. Sci.* **2016**, *113*, 7852-7857.

(28) Halstead, S. B. *J. Infect. Dis.* **2016**, *214*, 1793-1795.

(29) Dengue Vaccine Candidates in Clinical Development. *http://www.denguevaccine.org/vaccine-development/* (accessed February 15, 2017).

(30) Barrett, A. D. T. *N. Engl. J. Med.* **2016**, *375*, 301-303.

(31) Garske, T.; Van Kerkhove, M. D.; Yactayo, S.; Ronveaux, O.; Lewis, R. F.; Staples, J. E.; Perea, W.; Ferguson, N. M. *PLoS Med.* **2014**, *11*, e1001638.

(32) Staples, J. E.; Breiman, R. F.; Powers, A. M. *Clin. Infect. Dis.* **2009**, *49*, 942-948.

(33) Schwartz, O.; Albert, M. L. *Nat. Rev. Micro.* **2010**, *8*, 491-500.

(34) Surveillance and Control of *Aedes aegypti* and *Aedes albopictus* in the United States. *https://www.cdc.gov/chikungunya/resources/vector-control.html* (accessed February 9, 2017).

(35) Surtees, G. *Bull. World Health Organ.* **1967**, *36*, 594-596.

(36) Briegel, H. *J. Vector Ecol.* **2003**, *28*, 1-11.

(37) Butler, D. *Nature.* **2016**, *532*, 155-156.

(38) Mackenzie, J. S.; Gubler, D. J.; Petersen, L. R. *Nat. Med.* **2004**, 10, S98-S109.

(39) Githeko, A. K.; Lindsay, S. W.; Confalonieri, U. E.; Patz, J. A. *Bull. World Health Organ.* **2000**, *78*, 1136-1147.

(40) Patterson, J.; Sammon, M.; Garg, M. *West. J. Emerg. Med.* **2016**, *17*, 671-679.

(41) Alphey, L.; McKemey, A.; Nimmo, D.; Neira Oviedo, M.; Lacroix, R.; Matzen, K.; Beech, C. *Pathog. Glob. Health.* **2013**, *107*, 170-179.

(42) Ramirez, J. L.; Garver, L. S.; Dimopoulos, G. *Curr. Mol. Med.* **2009**, *9*, 116-130.

(43) Paes, d. A.; Aragão, F. J. L.; Colli, W.; Dellagostin, O. A.; Finardi-Filho, F.; Hirata, M. H.; Lira-Neto, A.; Almeida, d. M.; Nepomuceno, A. L.; Gorgônio, d. N.; Delfino, d. S.; Valicente, F. H.; Zanettini, M. H. B. *Bull. World Health Organ.* **2016**, *94*, 766-771.

(44) Resnik, D. B. *Dev. World Bioeth.* **2012**, *14*, 37-46.

(45) Amin, L.; Hashim, H. *Sci. Eng. Ethics.* **2015**, *21*, 655-681.

(46) Adalja, A.; Sell, T. K.; McGinty, M.; Boddie, C. *PLoS Currents* **2016**, *8*, ecurrents.outbreaks.1c39ec05a743d41ee39391ed0f2ed8d3.

(47) Marquardt, W. C. 2nd ed.; Elsevier: Oxford, **2005**.

(48) Magalhaes, T. *Parasit. Vectors.* **2014**, *7*, 362.

(49) O'Gower, A. K. *Aust. J. Biol. Sci.* **1956**, *9*, 125-129.

(50) Zhou, G.; Flowers, M.; Friedrich, K.; Horton, J.; Pennington, J.; Wells, M. A. *J. Insect Physiol.* **2004**, *50*, 337-349.

(51) Kalhok., S. E.; Tabak, L. M.; Prosser, D. E.; Brook, W.; Downe, A. E. R.; White, B. N. *Insect Mol. Biol.* **1993***, 2*, 71-79.

(52) Barillas-Mury, C.; Wells, M. A. *Insect Mol. Biol.* **1993***, 2*, 7-12.

(53) Jiang, Q.; Hall, M.; Noriega, F. G.; Wells, M. *Insect Biochem. Mol. Biol.* **1997***, 27*, 283-289.

(54) Bian, G.; Raikhel, A. S.; Zhu, J. *Insect Biochem. Mol. Biol.* **2008***, 38*, 190-200.

(55) Noriega, F. G.; Edgar, K. A.; Bechet, R.; Wells, M. A. *J. Insect Physiol.* **2002***, 48*, 205-212.

(56) Edwards, M. J.; Moskalyk, L. A.; Donelly-Doman, M.; Vlaskova, M.; Noriega, F. G.; Walker, V. K.; Jacobs-Lorena, M. *Insect Mol. Biol.* **2000***, 9*, 33-38.

(57) Isoe, J.; Zamora, J.; Miesfeld, R. L. *Insect Biochem. Mol. Biol.* **2009***, 39*, 68-73.

(58) Felix, C. R.; Betschart, B.; Billingsley, P. F.; Freyvogel, T. A. *Insect Biochem. Mol. Biol.* **1991***, 21*, 197-203.

(59) Isoe, J.; Rascón, A. A.; Kunz, S.; Miesfeld, R. L. Molecular Genetic Analysis of Midgut Serine Proteases in *Aedes aegypti* Mosquitoes. *Insect Biochem. Mol. Biol.* **2009***, 39*, 903-912.

(60) Noriega, F. G.; Wang, X.; Pennington, J. E.; Barillas-Mury, C. V.; Wells, M. A. *Insect Biochem. Mol. Biol.* **1996***, 26*, 119-126.

(61) Noriega, F. G.; Pennington, J. E.; Barillas-Mury, C.; Wang, X. Y.; Wells, M. A. *Insect Mol. Biol.* **1996***, 5*, 25-29.

(62) Barillas-Mury, C.; Graf, R.; Hagedorn, H. H.; Wells, M. A. *Insect Biochem. Mol. Biol.* **1991***, 21*, 825-831.

(63) Barillas-Mury, C. V.; Noriega, F. G.; Wells, M. A. *Insect Biochem. Mol. Biol.* **1995***, 25*, 241-246.

(64) Lu, S. J.; Pennington, J. E.; Stonehouse, A. R.; Mobula, M. M.; Wells, M. A. *Insect Biochem. Mol. Biol.* **2006***, 36*, 336-343.

(65) Blandin, S.; Moita, L. F.; Köcher, T.; Wilm, M.; Kafatos, F. C.; Levashina, E. A. *EMBO Rep.* **2002***, 3*, 852-856.

(66) Brackney, D. E.; Isoe, J.; Black, W. C.; Zamora, J.; Foy, B. D.; Miesfeld, R. L.; Olson, K. E. *J. Insect Physiol.* **2010**, *56*, 736-744.

(67) Rascón, A. A.; Gearin, J.; Isoe, J.; Miesfeld, R. L. *BMC Biochem.* **2011**, *12*, 43.

(68) Muhlia-Almazán, A.; Sánchez-Paz, A.; García-Carreño, F. L. *J. Comp. Physiol. B.* **2008**, *178*, 655-672.

(69) Voet, D.; Voet, J. G. 4th ed.; J. Wiley & Sons: New York, 2010.

(70) Nelson, D.; Cox, M. 5th ed.*;* W.H. Freeman and Company: New York, 2008.

(71) Rosano, G. L.; Ceccarelli, E. A. *Front. in Microbiol.* **2014**, *5*, 172.

(72) Singh, P.; Sharma, L.; Kulothungan, S. R.; Adkar, B. V.; Prajapati, R. S.; Ali, P. S.; Krishnan, B.; Varadarajan, R. *PLoS One* **2013**, 8, e63442.

(73) Promdonkoy, B.; Promdonkoy, P.; Tanapongpipat, S.; Luxananil, P.; Chewawiwat, N.; Audtho, M.; Panyim, S. *Curr. Microbiol.* **2004**, *49*, 84-88.

(74) Baneyx, F.; Mujacic, M. *Nat. Biotech.* **2004**, *22*, 1399-1408.

(75) Tsumoto, K.; Ejima, D.; Kumagai, I.; Arakawa, T. *Protein Expr. Purif.* **2003**, *28*, 1-8.

(76) Lobstein, J.; Emrich, C. A.; Jeans, C.; Faulkner, M.; Riggs, P.; Berkmen, M. *Microb. Cell Fact.* **2012**, *11*, 56-56.

(77) Chen, B.; Janes, H. W. Humana Press Inc.: New Jersey, 2002.

(78) Kardos, J.; Bódi, A.; Závodszky, P.; Venekei, I.; Gráf, L. *Biochemistry.* **1999**, *38*, 12248-12257.

(79) Sipos, T.; Merkel, J. R. *Biochemistry.* **1970**, *9*, 2766-2775.

(80) Lefèvre, F.; Rémy, M. H.; Masson, J. M. *Nucleic Acids Res.* **1997**, *25*, 447-448.

(81) TB074VM pET28a-c(+) Vector Map. *http://www.emdmillipore.com/US/en/product/pET-28a(%2B)-DNA---Novagen,EMD_BIO-69864#documentation*. (accessed February 28, 2017).

(82) *pET System Manual.* Novagen: Darmstadt, Germany, 2005.

(83) Amersham Pharmacia Biotech*;* AB ed.; Amersham Pharmacia Biotech: Amersham, United Kingdom, 2001.

(84) Gray, C. J.; Sullivan, J. M. *J. Chem. Technol. Biotechnol.* **1989***, 46*, 11-26.

(85) Bornhorst, J. A.; Falke, J. J. *Meth. Enzymol.* **2000***, 326*, 245-254.

(86) Yamada, T. 5th ed.*;* Wiley-Blackwell: West Sussex, United Kingdom, 2009.

(87) Ghasemi, Y.; Dabbagh, F.; Ghasemian, A. *Mol. Biotechnol.* **2012***, 52*, 1-7.

(88) Bjerga, G. E. K.; Arsın, H.; Larsen, Ø; Puntervoll, P.; Kleivdal, H. T. *J. Biotechnol.* **2016***, 222*, 38-46.

(89) Mason, A. B.; He, Q.; Halbrooks, P. J.; Everse, S. J.; Gumerov, D. R.; Kaltashov, I. A.; Smith, V. C.; Hewitt, J.; MacGillivray, R. T. A. *Biochemistry* **2002***, 41*, 9448-9454.

(90) Ledent, P.; Duez, C.; Vanhove, M.; Lejeune, A.; Fonzé, E.; Charlier, P.; Rhazi-Filali, F.; Thamm, I.; Guillaume, G.; Samyn, B.; Devreese, B.; Van Beeumen, J.; Lamotte-Brasseur, J.; Frère, J. *FEBS Lett.* **1997***, 413*, 194-196.

(91) Terpe, K. *Appl. Microbiol. Biotechnol.* **2003***, 60*, 523-533.

(92) Sabaty, M.; Avazeri, C.; Pignol, D.; Vermeglio, A. *Appl. Environ. Microbiol.* **2001***, 67*.

(93) Eschenfeldt, W. H.; Maltseva, N.; Stols, L.; Donnelly, M. I.; Gu, M.; Nocek, B.; Tan, K.; Kim, Y.; Joachimiak, A. *J. Struct. Funct. Genomics.* **2010***, 11*, 31-39.

(94) Vera, A.; González-Montalbán, N.; Arís, A.; Villaverde, A. *Biotechnol. Bioeng.* **2007***, 96*, 1101-1106.

(95) Brooks, H. B.; Geeganage, S.; Kahl, S. D.; Montrose, C.; Sittampalam, S.; Smith, M. C.; Weidner, J. R. Sittampalam, G. S., Coussens, N. P., Brimacombe, K., Grossman, A., Arkin, M., Auld, D., Austin, C., Baell, J., Bejcek, B., Chun, T. D. Y., Dahlin, J. L., Devanaryan, V., Foley, T. L., Glicksman, M., Hall, M. D., Hass, J. V., Inglese, J., Iversen, P. W., Lag-Nag, M., Li, Z., Mcgee, J., McManus, O., Riss, T., Trask, O. J., Weidner, J. R., Xia, M. and Xu, X.; Eli Lilly & Company: Bethesda, Maryland, **2012**; pp 63-76.

(96) Sabaty, M.; Grosse, S.; Adryanczyk, G.; Boiry, S.; Biaso, F.; Arnoux, P.; Pignol, D. *BMC Biochem.* **2013***, 14*, 28.

(97) Chelulei Cheison, S.; Brand, J.; Leeb, E.; Kulozik, U. *J. Agric. Food Chem.* **2011***, 59*, 1572-1581.

(98) Benoit, J. B.; Lopez-Martinez, G.; Patrick, K. R.; Phillips, Z. P.; Krause, T. B.; Denlinger, D. L. *Proc. Natl. Acad. Sci. U. S. A.* **2011***, 108*, 8026-8029.

(99) O'Donoghue, A. J.; Eroy-Reveles, A.; Knudsen, G. M.; Ingram, J.; Zhou, M.; Statnekov, J. B.; Greninger, A. L.; Hostetter, D. R.; Qu, G.; Maltby, D. A.; Anderson, M. O.; DeRisi, J. L.; McKerrow, J. H.; Burlingame, A. L.; Craik, C. S. *Nat. Meth.* **2012***, 9*, 1095-1100.

(100) Franz, A. W. E.; Kantor, A. M.; Passarelli, A. L.; Clem, R. J. *Viruses.* **2015***, 7*, 3741-3767.

(101) Brackney, D. E.; Foy, B. D.; Olson, K. E. *Am. J. Trop. Med. Hyg.* **2008***, 79*, 267-274.