Summer 2018

Biochemical Study of a Chymotrypsin-like Serine Protease, AaCHYMO, from the Female Aedes Aegypti Mosquito

Olive Escalicas Burata
San Jose State University

BIOCHEMICAL STUDY OF A CHYMOTRYPSIN-LIKE SERINE PROTEASE, AACHYMO, FROM THE FEMALE *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

Olive Burata

August 2018

The Designated Thesis Committee Approves the Thesis Titled

BIOCHEMICAL STUDY OF A CHYMOTRYPSIN-LIKE SERINE PROTEASE, AACHYMO, FROM THE FEMALE *AEDES AEGYPTI* MOSQUITO

by

Olive Burata

APPROVED FOR THE DEPARTMENT OF CHEMISTRY

SAN JOSÉ STATE UNIVERISTY

August 2018

Alberto A. Rascón, Jr., Ph.D	Department of Chemistry
Laura Miller Conrad, Ph.D.	Department of Chemistry
Elizabeth Migicovsky, Ph.D.	Department of Chemistry

ABSTRACT

BIOCHEMICAL STUDY OF A CHYMOTRYPSIN-LIKE SERINE PROTEASE, AACHYMO, FROM THE FEMALE *AEDES AEGYPTI* MOSQUITO

by Olive Burata

The *Aedes aegypti* female mosquito is a dominant vector of four bloodborne viruses: dengue, chikungunya, yellow fever, and Zika. Defenses against viral infection are reliant on human competence to limit interactions with the vector or through abatement using methods that are either too expensive or ecologically invasive upon other species. An alternative vector control approach focuses specifically on blood meal protein digestion, which is a required process to fuel the gonotrophic cycle. Reduction of eggs laid can be observed by inhibiting midgut serine proteases that are afterward released in the midgut in the presence of a blood meal. Trypsin-like serine proteases are the major enzymes involved in bloodmeal protein digestion and have been the focus of study in the 1990s. However, the study of chymotrypsin-like serine proteases has been neglected; therefore, the focus of this work will be on AaCHYMO, a chymotrypsin-like serine protease first identified in 1996. To understand the role of this protease in the blood meal digestion process, we must isolate and study the chymotrypsin *in vitro*. For the first time, recombinant AaCHYMO has been recombinantly expressed solubly in bacteria, and more surprisingly, the zymogen form of the enzyme is autocatalytic and has enough activity to cleave a well-known chymotrypsin-like substrate (Suc-Ala-Ala-Pro-Phe-pNA). This activity may be novel since autocatalytic behavior in chymotrypsin-like proteases is not commonly found in both vertebrate and invertebrate organisms.

ACKNOWLEDGMENTS

I dedicate this work to Prof. Alberto A. Rascón, Jr. for showing me that science was not only something that we learned in classrooms, but a way of life. He opened a door for me that I never thought I would be able to walk through. Through his mentorship, I not only renewed my passion for science, but also gained the confidence and skills to be an independent researcher. He showed me how to be strong in the face of challenges, that there's always a way whenever there's a will, and to be responsible in my decisions. I am very blessed to have both a mentor and a father-figure. ¿QUIERES QUESO!?

I would also like to thank Miss Flor, Celeste, and Buddy for welcoming Rachael and I to your family. Flor, thank you for the awesome Barbacoa that you always make. It's one of those things that really warms your heart and you just want everyone else to leave you alone while you enjoy it, which is surprising for me to say since I'm always talking to/annoying people.

I am very blessed to have gained the support of two families: the Buratas and the Luceros. They have seen me at my worse and my greatest throughout my journey here in SJSU. Picking up a career as a scientist is something extremely new to the Burata Family. It was not easy to transition to this field and my family tried their best to adjust and support me as much as they could. I thank them for always being there to support me and for listening/understanding my challenges/successes that were so foreign to them. To the Lucero Family, I thank you for being my awesome second family and constant cheerleaders in my work. I am very happy that you guys are consistently a part of not only Rachael's journey, but mine as well by always asking us about how we're doing and

getting us food at crazy hours while we're stuck in lab! It's very reassuring to be able to talk/vent/laugh to Mom, Dad and Grandma about our daily routines in lab and I hope you guys know how much of a big deal that is in helping us keep sane!

I would like to also thank Professor Miller Conrad, not just being a part of my thesis committee, but importantly for always being there for me whenever I randomly pop in your office. I'm sorry that over 90% of these visits are always unannounced and I think it may have something to do with how often I am in lab and passing by your office in the 6th floor while I do my experiments. It could also have to do with the fact that I feel very comfortable and relief about being able to talk you about anything science and life related concerns as well! I thank you for your advices and guidance, and in return I hope to make you proud as I go through PhD at UMich.

Of course, I am very happy to have met and endured the Master's journey with my fellow Masters! Diane, James, Beccaw!, Anthony, Victor, Lauren, Leroy, Jacob, Thanh, Kanishka, Alexis, Hassi, Shalaka, Angelina, Caroline, Regina, and Josh! It's so nice to be able to share laughs/jokes/complaints with you guys and it's just another way for me to keep sane! I am very excited to be able to see some of you in the East when I move to UMich!

A big shout out to the Rascón lab, especially Alexia, Kamille, the Fong brothers (Jon, Danieeeeeeeeeeeel, Timmeh!), Jamie, Ate Sairaaaa, and yes, even James! You guys are the only ones who understand what it means to be a part of the Rascón lab and I'm glad to be able to be able to share that with you!

Most importantly, I'm very lucky to have had my better half with me, my fiancée, Rachael, throughout my entire Master's career. A lot of tears have been shed, pulling late-nights, outbursts, and doubts we have experienced, but they've all been countered by hard work, jokes and laughter, pure happiness from successful experiments, trust, and communication. I would never trade what we have for anything better. I hope to consistently make you proud and happy always.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Aa = *Aedes aegypti*
cDNA = complementary Deoxyribonucleic acid
mRNA = messenger Ribonucleic acid
AaCHYMO = *Aedes aegypti* chymotrypsin
AaET = *Aedes aegypti* Early Trypsin
AaJHA15 = *Aedes aegypti* Juvenile Hormone
AaSPVI = *Aedes aegypti* Serine Protease VI
AaSVPII = *Aedes aegypti* Serine Protease VII
AaLT = *Aedes aegypti* Late Trypsin

CHAPTER 1

INTRODUCTION

1.1 The Female *Aedes aegypti* Mosquito

The female *Aedes aegypti* mosquito is an arthropod belonging to the Culicidae family and the subgenus, Stegomygia.[1] This breed originated from Africa and is currently prominent globally, thriving mostly in tropical and subtropical regions. The mosquito is highly characterized by its stripes, not only on its segmented abdomen, but also on its legs, as observed in Figure 1. They are anautogenous feeders, requiring blood meals obtained from vertebrae hosts to facilitate egg production.[1] Initially, the female *Ae. aegypti* mosquito preyed on animal hosts; however, its domestication through the expansion of human communities has affected not only its overall distribution toward urbanized regions, but has also evolved the mosquito to become an anthropophilic feeder.[1] Hence, the *Ae. aegypti* earned the nickname "The Urban Mosquito."

Figure 1. Imago form (final stage in the metamorphosis from eggs to adult) of the female *Aedes aegypti* mosquito with characteristic white stripes on the abdomen and legs, differentiating it from other mosquito species. (Image obtained with permission from University of Michigan Museum of Zoology).

The abundance of this mosquito has been the greatest factor in spreading arthropod-borne viruses (arboviruses) such as Dengue, Chikungunya, Yellow Fever, and Zika. This characteristic is greatly linked to the mosquito's life cycle (Figure 2). It is initiated by the laying of hundreds of robust eggs on the sides of water-filled containers. These eggs then form rafts if water levels are adequate; otherwise, the eggs are also capable of desiccation during water deprivation for approximately a year.[2] The hatching of the eggs releases the larvae, the second aquatic stage of the life cycle, which feed on microscopic organic

matter found within the water source. The larvae must feed to build up their nutritional reserves before reaching their fourth instar, thereby allowing the physical transition to a pupa.[3] The stationary, yet active pupa is the third aquatic stage. Sexual differentiation occurs during the pupae stage along with gradual metamorphosis which ends in eclosion, or emergence, by the imago form (final stage in the metamorphosis from eggs to adult) of the mosquito.[1,3] Immediately after eclosion, the female adult mosquito enters a teneral phase or maturation period that exhibits further differentiation of the mosquito's anatomy, development of its flight muscles, and establishment of both the hormonal and digestive systems.[3]

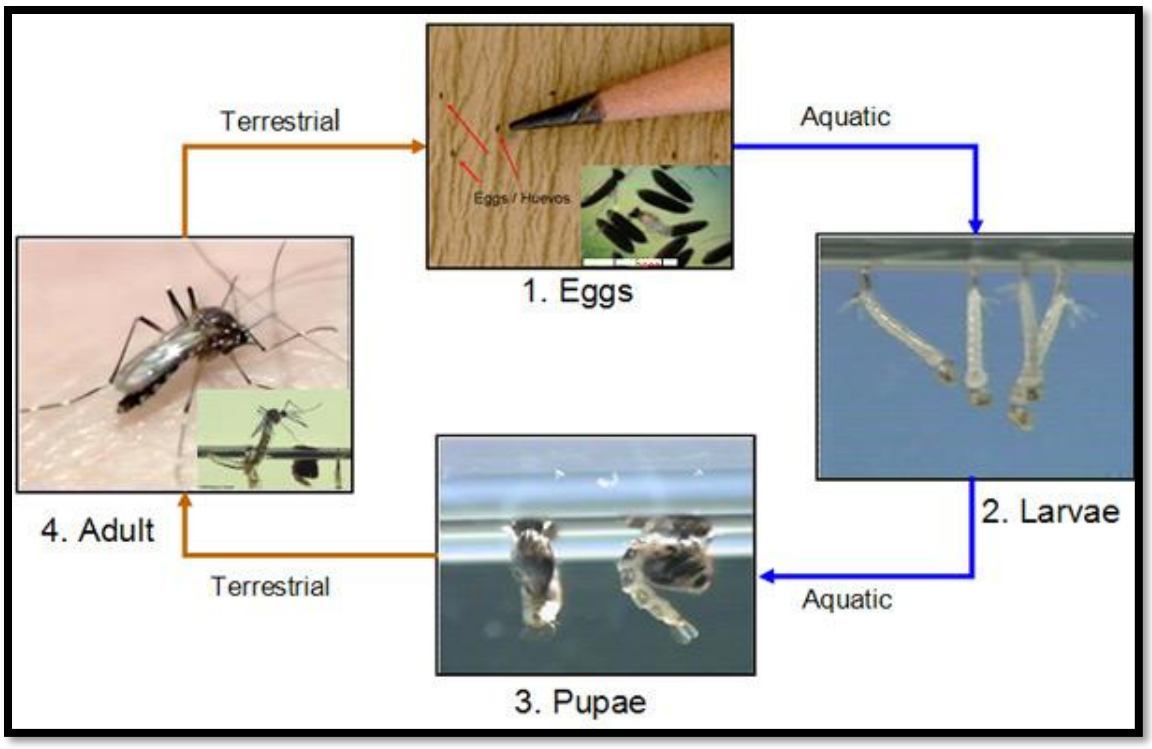


Figure 2. Life cycle of the *Ae. aegypti* mosquito with growth stages existing in both aquatic and terrestrial environments. Image obtained with permission from the CDC.

The duration of the life cycle is approximately 8-10 days depending on environmental factors (temperature, water level, food availability) and is entirely reliant on the gonotrophic cycle, a crucial step marked by the female *Ae. aegypti* mosquito acquiring the necessary nutrients from a blood meal.[4] The blood meal can be obtained from multiple human hosts until the female mosquito has fully engorged itself with approximately 2-5 µL of blood.[3] Two-thirds of the bloodmeal proteins are catabolized for lipid synthesis to sustain the energy requirements of the adult mosquito, while one-third of the bloodmeal is funneled into vitellogenesis, the process of yolk production, required for ovipositioning, the process of egg laying.[3] Typically, each female mosquito can produce up to 150 eggs per gonotrophic cycle followed by a 72-hour period of

oviposition.[2,5] The importance of blood meal protein digestion and egg laying was confirmed by Isoe and colleagues from the University of Arizona,[6] as shown in Figure 3. Ovaries were dissected from lab-grown Rockefeller strain of female *Ae. aegypti* mosquitos fed with a diluted blood meal, diluted with various percentages of feeding buffer (150 mM NaCl, 100 mM $NaHCO_3$ pH 7.0).[6] A clear decrease in the number of viable eggs can be observed beginning at 60% blood meal, indicating the dependence of egg production upon overall blood meal protein content and digestion.[6]

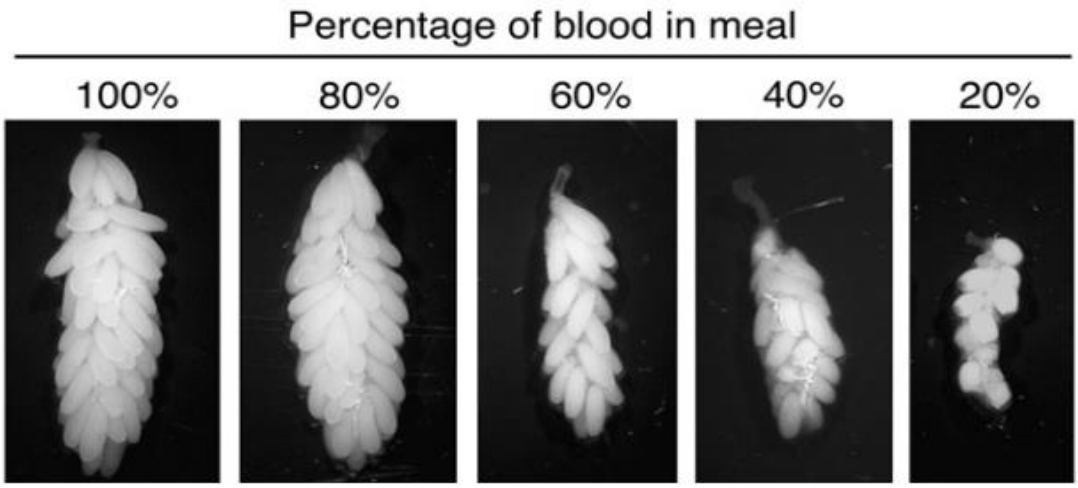


Figure 3. Ovaries dissected from female *Ae. aegypti* mosquitos fed with a diluted bloodmeal (diluted with feeding buffer containing 150 mM NaCl, 100 mM $NaHCO_3$ pH 7.0) to determine the relationship of blood meal to mosquito fecundity. As seen in the images, limiting the amount of nutrients leads to an overall effect on egg production. Image obtained, with permission, from Isoe *et al.*

1.2 The Transmission of Arthropod-borne Viruses

This anautogenous blood feeding behavior has facilitated the transmission of the Dengue virus, Yellow Fever virus, Chikungunya virus, and Zika virus. All these viruses originated in or near forested regions in Africa and have initially been found in simians.[7]

Sylvatic female *Ae. aegypti* mosquitos have allowed these zoonotic arboviruses to infect nearby human settlements or heavily urbanized regions. Thus, infected human hosts then became a viral source for the infection of peridomestic female *Ae. aegypti* mosquitos (Figure 4).[7]

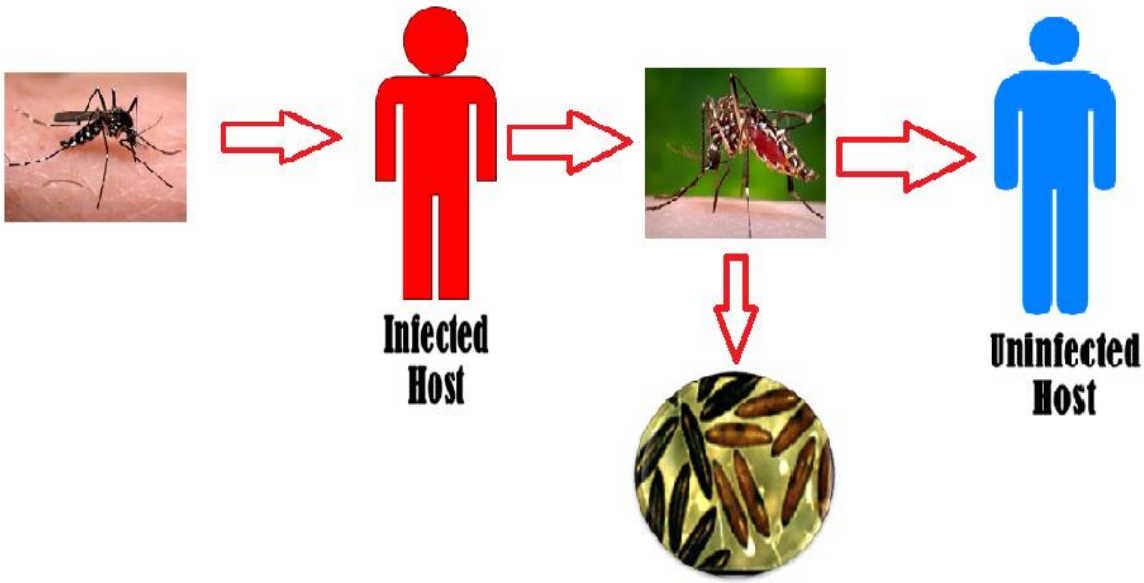


Figure 4. Horizontal transmission (from infected host to peridomestic female *Ae. aegypti* mosquito) and vertical transmission (from infected female *Ae. aegypti* mosquito to its offspring) of arboviruses. Images modified from CDC.

A vaccine is available for Yellow Fever, however, no vaccines have yet to be finalized for Zika or Chikungunya.[2,5] To combat the Dengue virus (DEN-V), only one semi-suitable vaccine is currently available called Dengvaxia.[2] This vaccine is only effective against four out of five serotypes of DEN-V, and a second heterotypic infection is usually accompanied with an increase in the severity of the disease, which limits the efficacy of the vaccine, and so is only available in highly endemic areas.[2] The low establishment of vaccines in combatting these diseases have turned defenses toward

either vector surveillance or preventive measures, which rely heavily on human compliance.[2] Other measures involve an aggressive use of organophosphate-based or pyrethroid-based insecticides or the introduction of biological predators, both of which have caused ecological effects for other organisms, especially pollinators.[1,5,8] Originally, the viral pathogens that the *Ae. aegypti* mosquito is capable of transmitting have only been endemic in the African and Asian continents. However, with climate change, urbanization, deforestation, increased human travels, and human overpopulation have facilitated the spread and outbreaks of these diseases in other continents.[1,9]

1.3 Potential Novel Vector Control Strategy

In the late 1940s, the foundation of a potential novel vector control strategy against *Ae. aegypti* was established upon studies of the mosquito's digestive system. A study by F.W. Fisk in 1950 established the foundation of this strategy by discovering how ingestion of the bloodmeal by the female mosquito stimulated the release of proteolytic enzymes or proteases into the midgut to facilitate efficient digestion of blood meal proteins.[10] In addition, male mosquitoes, sugar-fed mosquitoes and unfed female mosquitoes did not exhibit proteolytic activity higher than the residual amounts. Interestingly, the female mosquito can concentrate a blood meal by removal of water through excretion, which is important for a mosquito when a blood meal is limited and not enough nutrients are available. The complexity of a blood meal is typically low and dominantly consists of protein (~80%) in three forms: hemoglobin (~83%), serum albumin (~13%), and immunoglobulin (~4%).[6] The presence of amino acids derived from

the blood meal have been shown to induce the synthesis and secretion of these proteases in the midgut post-blood feeding.[11,12]

1.4 Biphasic Digestion Process of Blood Meal Proteins

Trypsin-like proteases were believed to be the major enzymes involved in midgut blood meal protein digestion, and focus on these protease types led to the discovery of four of the most abundant trypsin-like proteases during the 1990's.[13-15] These proteases were expressed and released at various times throughout the course of blood meal protein digestion, which led to the observation of two distinct phases: an early and a late phase.[16,17] Of the four abundant proteases, the major contributors of trypsin-like activity are Early trypsin (AaET) being a major trypsin-like protease in the early phase, while the other two are SPVI (AaSPVI) and SPVII (AaSPVII) in the late phase.[13-15] Felix *et al.* (1991) initially suggested that proteolytic activity peaks at approximately 18 hours post-feeding.[17] Isoe *et al.*, however, was able to specifically confirm the biphasic trypsin-like activity in individual midgut extracts dissected after different hours post blood meal (PBM) using a synthetic substrate, N-α-Benzoyl-DL-arginine-4-nitroanilide hydrochloride (BApNA), which releases a chromophore *p*-nitroaniline that can be measured spectrophotometrically at $OD_{405\ nm}$ (Figure 5).[6] The total activity was measured for a total of 48 hours PBM where two major increases in trypsin-like activity was observed, an initial increase between 0 to 15 h PBM, and a second increase past 18 h PBM was observed and peaked at approximately 30 h PBM. During the early phase, AaET is expressed immediately and peaks at approximately 6 h PBM.[6,18] A second

increase (late phase) is then observed from 15 h PBM where active AaSPVI and AaSPVII begin to be expressed at 18 h PBM with maximal expression peaking at 24 h PBM.[6]

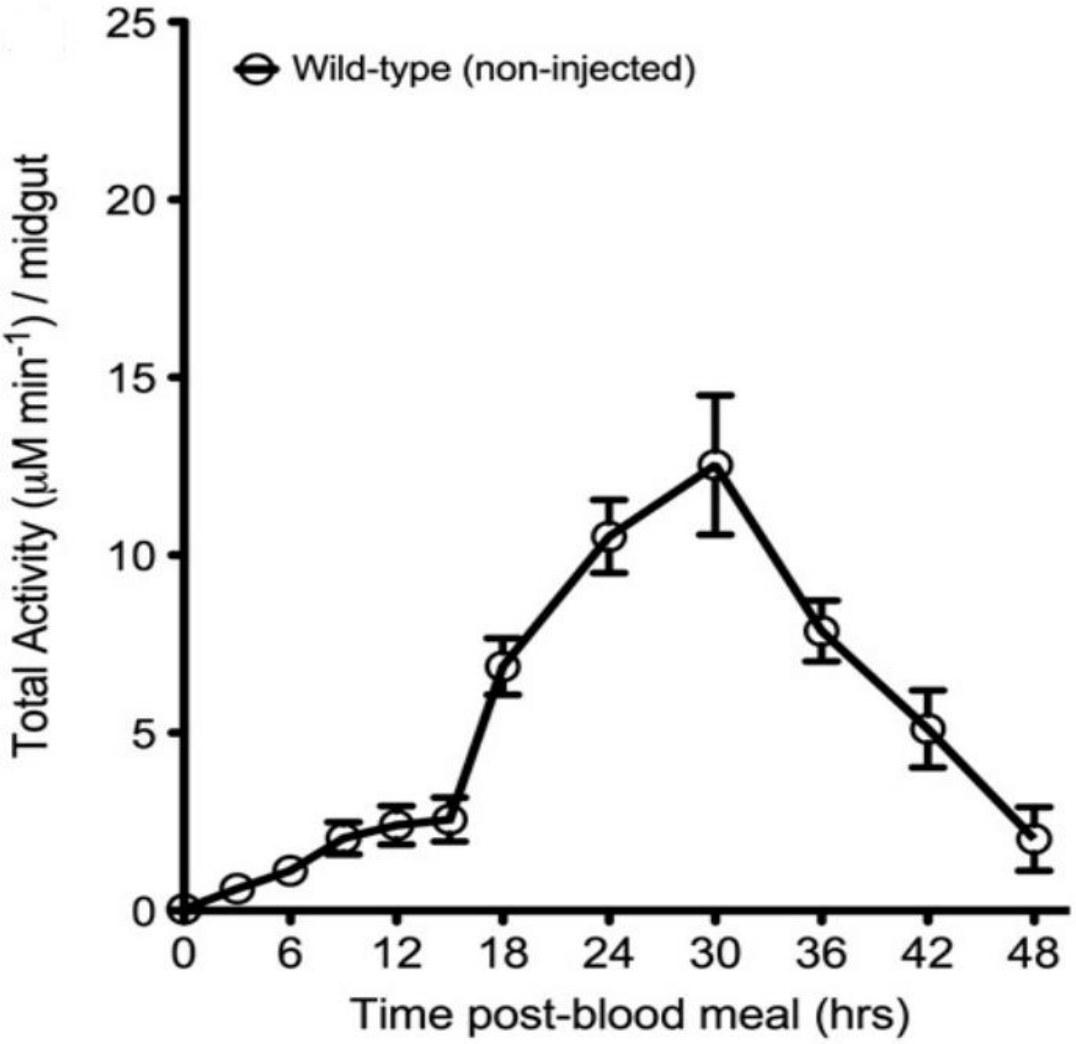


Figure 5. Using steady-state assumptions, enzyme kinetics of total trypsin-like activity was determined from cleavage of BApNA to *p*-nitroaniline, measured using UV-VIS at 405 nm. Each data point represents timepoints of when activity was measured from each dissected midgut for total of 48 h PBM. Two significant increases were observed: from 0 to 15 h PBM and 16 h PBM to 30 h PBM. Image obtained, with permission, from Isoe *et al.* (2009).[6]

Additionally, Isoe *et al.* investigated the phenotypic relationship between fecundity and the abundant proteases released during blood meal protein digestion (Figure 6).

RNAi technology was used to knockdown the transcripts of the three trypsin-like proteases, as well as a serine collagenase-like protease, AaLT.[6] Knockdown of AaET had no effect in overall fecundity. However, knockdown of the three late phase proteases, AaSPVI, AaSPVII, and AaLT resulted in a clear decrease in the number of viable oocytes laid by the mosquito.[6,12] Interestingly, a mixed knockdown of all three proteases at once only resulted in only ~31% reduction in the number of eggs oviposited. These findings suggest that other proteases could be compensating for the lack of trypsin-like activity in the knockdown of the abundant late phase proteases.[6]

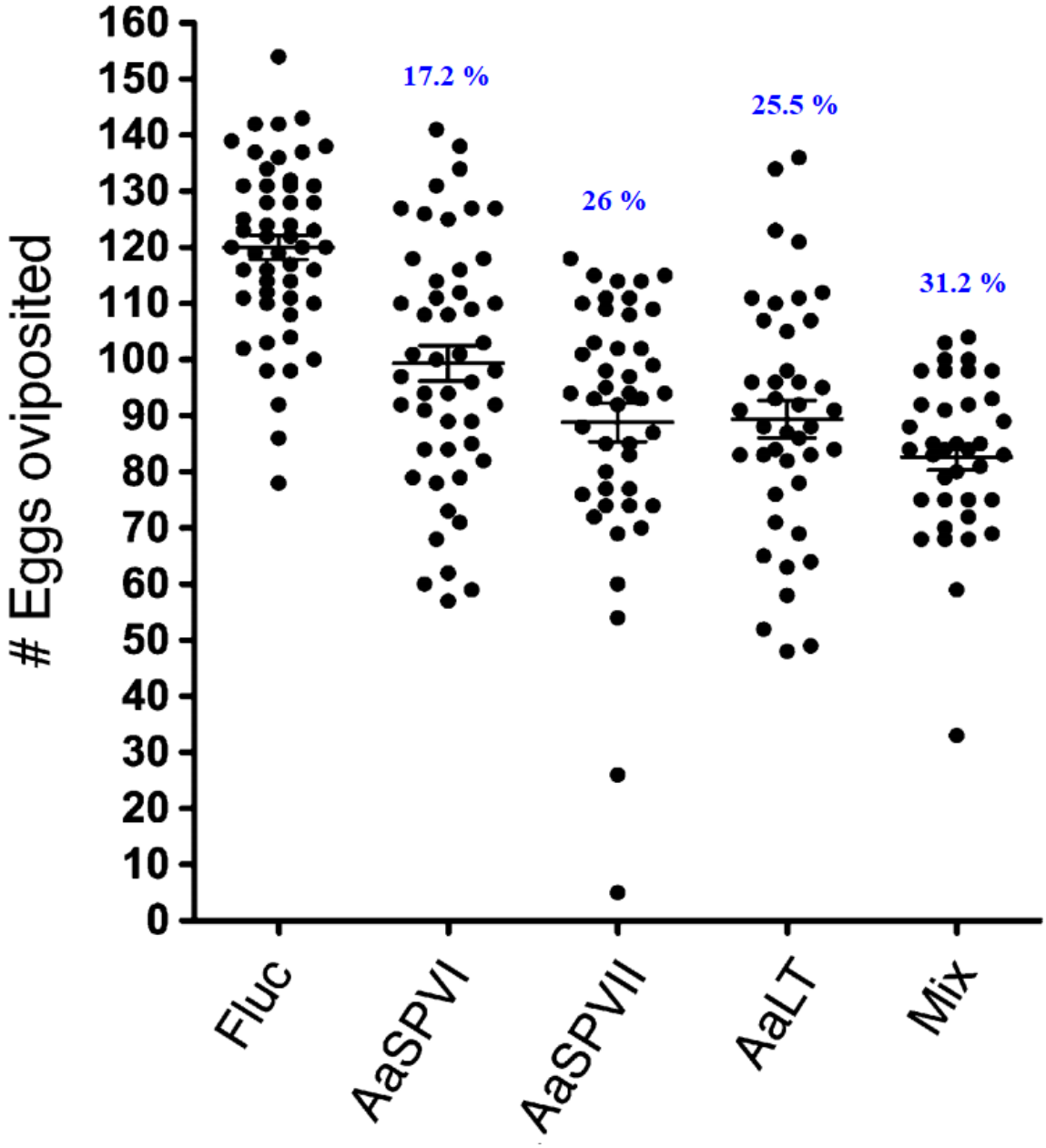


Figure 6. Knockdown of late phase mRNA transcripts and their effects on eggs laid using dsRNA injections specific to each protease (AaSPVI, AaSPVII, AaLT, and Mixed of all three). An average reduction of eggs oviposited are shown above each protease that are knockdown. Significant reductions in eggs oviposited are observed for individually knockdown protease, however Mix knockdown did not exhibit an additive reduction in eggs oviposited. Modified image from Isoe *et al.* (2009), obtained with permission.[6]

Other proteases that could possibly be involved are chymotrypsins[12,19,20] that may cleave intact globular blood meal proteins, and the exopeptidases such as aminopeptidases[21] and carboxypeptidases,[22] which only cleave smaller peptides (cleavage products from the endoproteases), as well as serine collagenases.[18] Of importance to this work, there are two midgut chymotrypsin-like proteases thought to be involved in the blood meal digestion process: JHA15[12] and AaCHYMO.[20] JHA15 is a juvenile-hormone regulated, chymotrypsin-like serine protease, which is expressed starting in the early phase, similar to the expression pattern of AaET,[12] but unlike AaET, JHA15 is constantly expressed throughout the whole blood meal digestion process. AaCHYMO is also expressed in the early phase of the blood meal protein digestion process along with AaET and JHA15.[12,13,20] However, prior to the referenced studies focused on chymotrypsins (see 12,13,20), their role in the female *Ae. aegypti* mosquito midgut digestion process had been minimally studied.

In 1997, Jiang *et al.* provided the first and only molecular analysis (both *in vivo* and *in vitro*) of the first identified chymotrypsin-like protease (AaCHYMO).[20] Through *in vivo* studies on adult female *Ae. aegypti* mosquitoes, Jiang and colleagues observed AaCHYMO transcripts in dissected midguts of unfed female mosquitoes starting at 24 hours post-eclosion, with a dramatic increase observed within the next 48 h.[20] Interestingly, when the mosquitoes were fed a ɣ-globulin protein meal the AaCHYMO transcripts did not decrease but were observed to remain constant for approximately 24 hours throughout the digestion process. Furthermore, using AaCHYMO-specific antibodies, Western Blot analysis of unfed and fed female mosquitoes led to the detection

of residual levels of AaCHYMO protease prior to feeding, and once a protein meal was given, a significant increase in AaCHYMO mRNA translation was observed 3 hours after feeding, peaking at approximately 24 hours.[20] This is analogous to the release of trypsin-like midgut proteases in response to an increased presence of amino acid concentration in the midgut.[11,12] The chymotrypsin-like activity of crude midgut extracts was also measured using a synthetic chymotrypsin-specific substrate N-Succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-pNA). Based on these experiments, there was no chymotrypsin-like activity in unfed female mosquitoes and activity was only observed in fed female mosquito crude extracts. This study was conducted in 1997, with no other chymotrypsin-like proteases identified at that point, and so the activity in the extracts were assumed to only come from AaCHYMO.

Interestingly, work since then has revealed that AaLT (named *Ae. aegypti* Late Trypsin) did not exhibit trypsin-like behavior, and instead could possibly be a serine-collagenase.[6,18] Most serine collagenases are capable of cleaving multiple substrates, including chymotrypsin-preferred substrates[18] indicating that the enzymatic assays performed by Jiang *et al.* (1997) did not take into account the possible activity on Suc-AAPF-pNA coming from AaLT. In addition, with the discovery of a second chymotrypsin-like protease (JHA15) in 2008, with the same expression pattern of both AaET and AaCHYMO,[12] may have also contributed to the chymotrypsin-like activity observed in the midgut extract assays. Second, Jiang *et.al.*'s 1997 study has determined that both the expression of midgut AaCHYMO and its respective proteolytic activity were initially detected at 3 hours post-feeding followed by a dramatic increase until

approximately 24 hours post-feeding.[20] However, the Ph.D. dissertation by Jiang presenting the same biochemical studies for AaCHYMO contradicted the former of a dramatic increase in both AaCHYMO protein concentration and activity, instead the dissertation reported a constant level of expression for AaCHYMO isolated from crude midgut extracts and its respective proteolytic activity throughout protein digestion.[23] As a result, it is unclear what the true expression pattern of AaCHYMO is in the female *Ae. aegypti* midgut. Third, Jiang *et al.* used recombinant AaCHYMO to help compare and detect the protease in mosquito midgut extracts, but it was ambiguous on whether the observed expression patterns in the Western Blot analysis is of the recombinant full length AaCHYMO zymogen with the leader sequence or without the leader sequence. Since no molecular weight ladders were used to compare the approximate molecular weights of the protein targets using SDS-PAGE, it is unclear whether the expression patterns for the crude midgut AaCHYMO is either the active mature form or the zymogen form without the leader sequence.[20,23]

Therefore, the work in this thesis will focus on clearing some of the discrepancies found in the initial study by Jiang *et al.* (1997) on AaCHYMO. AaCHYMO will be investigated *in vitro* in order to determine whether it is indeed a chymotrypsin-like protease, but also to determine its specific role in the blood meal digestion process. It is unknown whether the protease contributes to the blood meal digestion process, and so should help answer the question, does this protease cleave intact blood meal proteins or does it depend on the activity of trypsin-like proteases to first cleave the blood meal

proteins releasing products that can then be further digested by AaCHYMO. This work will help set-up future *in vivo* studies in the mosquito focusing on AaCHYMO.

CHAPTER 2

INITIAL BIOCHEMICAL STUDY OF N-TERMINALLY HIS_6-TAGGED AACHYMO

2.1 Introduction

The initial work by Jiang et al (1997) helped identify the first chymotrypsin-like protease, and more importantly, identified chymotrypsin-like activity in crude midgut extracts.[20] However, the role of this identified protease (AaCHYMO) was only hypothesized to be involved in the blood meal protein digestion process, and any *in vitro* activity using recombinant AaCHYMO was minimal and vague at best. Therefore, to aid in the *in vitro* biochemical study of AaCHYMO and its potential role in midgut blood meal protein digestion, the gene will be cloned and isolated using recombinant DNA technology, followed by over-expression using bacteria. Using bacteria as the host organism, copious quantities of recombinant protein can be obtained.[24] It is an alternative method to directly harvesting proteins from host tissues, a process commonly associated with extremely low protein yields that are inadequate for downstream biochemical studies.[24] This has been a challenging pattern observed in *in vivo* characterization of mosquito serine proteases dating back to the 1990s and has motivated the shift to *in vitro* studies using recombinant technology.[18,20]

To produce the recombinant gene of interest, the open reading frame of zymogen AaCHYMO full length with leader sequence was PCR amplified using gene specific primers that incorporate restriction sites to allow proper cloning into the pET28a expression vector to incorporate an N-terminal His_6-tag on the recombinant protein

target. The plasmid containing the amplified AaCHYMO gene sequence was initially expressed in BL21(DE3) *Escherichia coli*. Based on initial recombinant expression of the abundant full-length trypsin-like proteases by Rascon *et. al.*, total expression of recombinant AaCHYMO is predicted to be observed using BL21(DE3) and a derivative. However, the likelihood of producing soluble protein is not expected due to the reducing cytoplasm of these bacterial cells and the dependence on disulfide bridge formation for structural stability and function.[18] Therefore, a new AaCHYMO zymogen construct without the leader sequence was designed and expressed in SHuffle T7 *E. coli* competent cells, bacterial cells with a more oxidizing cytoplasm which will allow proper formation of disulfide bonds (Lobstein 2012). Once soluble protein expression was obtained, the expression conditions were optimized to obtain the highest zymogen yield possible to facilitate the purification, activation, and downstream biochemical studies of recombinant AaCHYMO. This chapter will focus on these strategies, as well as the initial enzymatic assays and determination of kinetic parameters using the synthetic chymotrypsin substrate, Suc-AAPF-pNA. This will lay the foundation to help clarify the discrepancies represented in the initial study by Jiang *et. al.* (1997).

2.2 Methods

2.2.1 Cloning of the Full Length AaCHYMO Zymogen with Leader Sequence in to the pET28a Vector

Midgut-specific cDNA (obtained from blood-fed, total RNA) using Rockefeller strain *Ae. aegypti* female mosquitoes was provided by the University of Arizona. Forward and reverse primers specific to the full-length AaCHYMO zymogen with the leader sequence (AaCHYMO-Z Full Length with Leader) were designed and are listed in Table 1. The forward primer was manually designed to incorporate the leader sequence and a recognition site for NdeI, while the reverse primer contains the 3' end of the gene of interest reverse complemented from the AaCHYMO ORF and a recognition site for HindIII and two stop codons (TAA). Both primers contain a $Poly_5(A)$ tail at the 5' ends. These designed primers were subsequently purchased from Elim Biopharmaceuticals, Inc. and obtained at stock concentrations of 200 μM (Table 1). Their respective melting temperatures were calculated using NetPrimer from Premier Biosoft. The AaCHYMO-Z Full-Length with Leader was amplified using three different reverse primers in separate PCR reactions. AaCHYMO-Zym-pET-Rev 1 primer was the original design for the reverse primer used for amplifying the AaCHYMO WT Full-Length with Leader Sequence construct. However, the primer contained a trinucleotide tandem repeat of *GCTGCTGCT* which can result in non-specific annealing[25]; therefore, Rev 2 and Rev 3 were subsequently designed as a precaution to disrupt the 3rd and 2nd tandem repeat.

Table 1. Primers designed for PCR-amplification of the AaCHYMO-Z Full Length with Leader sequence for cloning into the pET28a vector to produce the AaCHYMO-Z Full Length/pET28a pDNA construct. The respective melting temperatures (T_M) are respectively shown. The restriction sites are bolded and each primer contains a $Poly_5(A)$ tail at the 5' ends.

Gene Construct	Primer	Primer Sequence	T_M (°C)
AaCHYMO Zymogen (WT) Full-Length with Leader	AaCHYMO-Zym-pET-Fwd	5'-AAAAA**CATATG**GCTTTCAAACTCACGGTAGC-3'	62.61
	AaCHYMO -Zym-pET-Rev 1, 2, &3 (starting from the top)	5'-AAAAA**AAGCTT**ATTACGCACGGAGCTGCTGCT-3' 5'-AAAAA**AAGCTT**ATTACGCACGGAGCTGCTGCTC-3' 5'-AAAAA**AAGCTT**ATTACGCACGGAGCTGCTGC-3'	64.26 66.11 63.07

Prior to cloning, the AaCHYMO-Z Full Length with Leader open reading frame was PCR-amplified using the Promega GoTaq Green Master Mix (Promega; Cat# M712B), 10 μM of both the designed forward and reverse primers, and the amplicon source (*Ae. aegypti* midgut cDNA), following the manufacturer's protocol. Amplification was facilitated with the settings listed in Table 2 using the Eppendorf Mastercycler Gradient Thermocycler.

Table 2. PCR settings and amplicon source required for PCR-amplification of the AaCHYMO-Z Full Length with Leader gene using the Eppendorf Mastercycler Gradient Thermocycler.

Gene Construct	Amplicon Source	Amplicon Amounts (µL)	PCR Settings
AaCHYMO-Z (WT) Full-Length with Leader	*Ae. aegypti* midgut cDNA	2	*Lid Heating:* 95°C for 5 minutes *Denaturation:* 95°C for 20 seconds *Annealing*: 60°C for 20 seconds *Elongation*: 72°C for 20 seconds *Cycle*: 25 *Final Elongation*: 72°C for 5 minutes

The AaCHYMO-Z Full Length with Leader PCR products and 1 kB DNA Ladder (Thermofisher Scientific; Cat# SM0311) were loaded onto 1% agarose gels (212 mL 1X Tris-Acetate-EDTA (TAE) buffer, 2.12 g Agarose Type II Medium EEO (Sigma-Aldrich; Cat#A9918), and 21.2 µL Gel Red Nucleic Acid Stain (Biotin; Cat# 41002)) and run for approximately 30 minutes. The gels were observed and analyzed under ultraviolet (UV) light using the FluorChem 8980. The PCR products were extracted from the agarose gels using the Qiagen Miniprep Gel Extraction Kit (Qiagen; Cat# 28706) based on the provided protocol. Elution of the PCR products was facilitated via centrifugation from the purification columns provided by the kits after incubating with 25 µL Millipore water. Since the dry column automatically absorbs 2 µL Millipore water, only 23 µL of the amplified AaCHYMO-Z Full Length with Leader remains.

The AaCHYMO-Z Full Length with Leader PCR products and pET28a vectors (Novagen) were digested separately with NdeI (Thermofisher Scientific; Cat# FD0584)

and HindIII (Thermofisher Scientific; Cat# FD0505) restriction enzymes. Three digestion reactions were set up for the PCR products (Samples 1-1, 2-1, and 3-1, based on reverse primer numbering, see Table 1). Each reaction contained 1 µL of each restriction enzyme, the PCR products and 1X Fast Digest Buffer, and Millipore water. A similar reaction was set for the pET28a vectors, but only 1000 ng of pET28a vectors were used and obtained from three different concentrations. All digestion reactions (20 µL total volume) were incubated in a 37°C water bath for 1 hour. The vectors were then placed in a 90°C heat block for 10 minutes, incubated on ice for 2 minutes, and subsequently followed by incubation with 1 µL of Thermosensitive Alkaline Phosphatase (Promega; Cat#M9910) in a 37°C water bath for 1 hour. Loading Dye (6X) was added to all digested vectors (final 1X) and ran using 1% agarose gel electrophoresis. Each digested full length AaCHYMO-Z WT gene was co-purified with a digested vector product that was excised out of the 1% agarose gel using the Qiagen Miniprep Gel Extraction Kit. The co-purified products were eluted with 20 µL Millipore water and ligated using 1 µL T4 DNA Ligase (New England Biolabs; Cat# M0202S) with 1X T4 DNA Ligase Buffer incubated at room temperature for 20 minutes. The newly ligated AaCHYMO-Z (full-length) WT in pET28a plasmid DNA (pDNA) were all transformed using NEB 10ß cells (Cat# C9019H) following the manufacturer's protocol. The transformed cells were allowed to recover in 950 µL SOC media (Biolabs; Cat# B9020S) at 37°C for 1 h and plated on Luria-Broth (LB) agar (Fisher Scientific; Cat#BP9724-500) containing 30 µg/µL Kanamcyin (Kan) (Fisher Scientific; Cat#BP906-5) (plates were incubated overnight in a 37°C incubator). Colony PCR was used to verify the ligated plasmids by

choosing four colonies from each transformation plate and amplifying the gene in the colony using the respective designed primers (Table 1). The samples were analyzed via gel electrophoresis and visualized using the Fluorchem 8900 under UV light.

From the positive colony PCR hits, another set of four colonies were obtained from each transformation plate and transferred to 50 mL falcon tubes containing 25 mL LB media (MP Biomedicals; Cat# 3002-032) and 30 ug/mL kanamycin. A total of 16 overnight cultures were set and grown in a 37°C shaker (250 rpm) for approximately 16 hours. The overnight cultures were centrifuged for 12 minutes at 2500 rpm and the supernatant was discarded. The pDNA was isolated using the GeneJet Plasmid Miniprep Kit (Thermofisher Scientific; Cat# K0503) with the pellet initially solubilized in 250 µL Resuspension Solution. Elution of the pDNA from the purification columns was facilitated with 35 µL Millipore water. The concentrations of purified pDNA were estimated using the Nanodrop 1000 Spectrophotometer. PCR was used to verify the presence of the gene of interest in each of the 12 pDNA samples using the respective primers in Table 1, followed by agarose gel analysis. After isolation, all pDNA constructs were sent to ELIM Biopharmaceuticals, Inc. (Hayward, CA)) for DNA sequencing, which provided DNA sequencing chromatograms analyzed using Chromas (Technelysium; https://technelysium.com.au/wp/chromas/). The sequences were analyzed for the presence of the restriction sites CATATG (NdeI) and AAGCTT (HindIII), ribosomal bindings site (AAGGAG), the His_6-tag sequence (CATCATCATCATCATCAT) at the 5' end, and both the start (ATG) and stop (TAA) codons. In addition, the generated sequences were compared to check for any

misalignments between the T7 forward and reverse primer sequences, which could indicate a mistake in sequencing or introduction of an undesired mutation, using Clustal Omega <Multiple Sequence Alignment < EMBL-EBI (http://www.ebi.ac.uk/Tools/msa/clustalo/) to help with confirming the final gene sequence. The AaCHYMO ORF from the final gene sequence was then verified using the Basic Local Alignment Search Tool (BLAST) on the NIH's National Center for Biotechnology Information (NCBI) website to ensure correct annotation of the cloned gene with the *Ae. aegypti* genome. The predicted amino acid sequence of the gene target was obtained using the Translation tool from ExPASy, the Swiss Institute of Bioinformatics (SIB)) and subsequently verified using BLAST. The molecular weight (MW) and the pI of the final amino acid sequence for the AaCHYMO-Z Full Length with Leader protease was obtained using the "Compute pI/MW tool" from ExPASy. More importantly, amino acid sequence analysis was also used to determine the presence of cysteine residues involved in the formation of disulfide bridges and the presence of the catalytic triad.

2.2.2 Small Scale Bacterial Expression of AaCHYMO-Z Full Length with Leader/pET28a pDNA in BL21(DE3)

To determine if the newly constructed AaCHYMO-Z Full Length with Leader/pET28a pDNA can be expressed in a bacterial cell line, 50 ng of purified pDNA was transformed into 50 µL of BL21(DE3) Competent *E. coli* cells (Bioline; Cat# 55032) and plated on LB agar plates containing 30 ug/mL kanamycin, which were grown overnight in a 37°C incubator. The overnight cultures were set by choosing a colony from

the transformation plates and growing in 5 mL LB media with 30 ug/mL kanamycin. The overnights incubated for approximately 16 hours in a 37°C shaker (250 rpm). An optical density at 600 nm (OD_{600}) of each overnight culture was measured to determine the amount needed for dilution to start the growth at an initial OD_{600} of 0.05. They were grown in a 125 mL Erlenmeyer flask, containing 50 mL LB media and 30 ug/mL kanamycin, and incubated in a 37°C shaker (220 rpm) until an OD_{600} of 0.5-0.8 was achieved. A 1 mL timepoint (T0) was collected from the culture before induction with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) then placed back in the 37°C shaker (220 rpm). Various 1 mL timepoints were collected in 1.5 mL Eppendorf tubes for analysis and centrifuged for 2 minutes at 13,000 rcf. The supernatant was removed, and the pellets were stored in the -20°C freezer.

2.2.3 SDS-PAGE Analysis of Recombinant AaCHYMO-Z Full Length with Leader Expressed in BL21(DE3) Cells

The pellets collected above were solubilized in cold 20 mM Tris-HCl pH 7.2 and sonicated for 10 seconds at 20% intensity for a total of 3 cycles (kept on ice during each cycled interval). Total samples were prepared by obtaining 16.6 µL of the sonicated pellets mixed with 6X SDS Loading dye (in-lab preparation: 2% (w/v) SDS, 50 mM Tris-HCl pH 7.2, 0.2 mg/mL bromophenol blue, 0.1 M DTT, & 10% v/v glycerol) for a final 1X concentration. The sonicated samples were cold centrifuged (4°C) for 5 minutes at 13,000 rcf to collect the soluble sample (the supernatant). Afterwards, 16.6 µL of the soluble sample were transferred to 1.5 mL Eppendorf tubes mixed with 6X SDS Loading dye (final 1X). All total and soluble samples were vortexed and centrifuged for 10

seconds on a tabletop microcentrifuge before heating at 90°C for 4 minutes to fully denature the proteins. The samples were loaded on to 4-12% Bis-Tris SDS-PAGE gels (Life Technologies; Cat#NP0323BOX) and ran with a PAGE Ruler Pre-stained Protein Ladder (Thermofisher Scientific; Cat# 26616). The gels were stained with Simply Blue Safe Stain (Novex Life Technologies; Cat# LC6065) overnight or with InVision His-Tag In-Gel stain (Life Technologies; Cat# LC6060), based on the manufacturer's protocol.

2.2.4 Design and Cloning of the Full Length AaCHYMO Zymogen No Leader Sequence in to the pET28a Vector

Since insolubility was predicted to be an issue in the expression of AaCHYMO-Z Full Length with the Leader Sequence using BL21(DE3) cells, the removal of the leader sequence should help alleviate this issue. The exact amino acid sequence of the leader/signal peptide sequence was determined using SignalP 4.1. Since the removal of the leader sequence only changes the N-terminus of the new construct, a new forward primer was designed (Table 3), while only one of the reverse primers was utilized during PCR (Rev 1, Table 1 & 3). The new forward primer used for PCR amplification of AaCHYMO-Z Full Length without the Leader, incorporated only the pro-peptide sequence (ACCCACAAG, which translates to THK), a portion of the mature sequence, the Nde I recognition site, and a $Poly_5(A)$ tail at the 5' end (see Table 3).

Table 3. Primers designed for PCR-amplification of the AaCHYMO-Z Full Length No Leader sequence for cloning into the pET28a vector to produce a leaderless AaCHYMO-Z Full Length/pET28a pDNA construct. The respective melting temperatures (T_M) are respectively shown. The restriction sites are bolded, and each primer contains a $Poly_5(A)$ tail at the 5' ends.

Gene Construct	Primer	Primer Sequence	T_M (°C)
AaCHYMO Zymogen (WT) No Leader	AaCHYMO-Zym-pET-Fwd No Leader	5'-AAAAA**CATATG**ACCCACAAGATCGTCGGTGG-3'	67.68
	AaCHYMO-Zym-pET-Rev 1	5'-AAAAA**AAGCTT**ATTACGCACGGAGCTGCTGCT-3'	64.26

The DNA template for amplifying the AaCHYMO-Z Full Length No Leader sequence was the purified AaCHYMO-Z Full Length with Leader/pET28a pDNA construct (Table 4). In addition, the table also lists the thermocycler setting required for PCR. The following steps involving excision of PCR products from 1% agarose gel, digestion with NdeI and HindIII restriction enzymes, ligation of the gene target to pET28a vector, purification of the AaCHYMO-Z Full Length No Leader/pET28a pDNA, and sequencing analysis to confirm proper cloning were followed as described above for the cloning of AaCHYMO-Z Full Length with Leader/pET28a pDNA (Section 2.2.1).

Table 4. PCR settings and amplicon source required for PCR-amplification of the AaCHYMO-Z Full Length No Leader gene using the Eppendorf Mastercycler Gradient Thermocycler.

Gene Construct	Amplicon Source	Amplicon Amounts (µL)	PCR Settings
AaCHYMO-Z (WT) No Leader	AaCHYMO (WT) with Leader/pET28a Plasmid	1	*Lid Heating:* 95°C for 5 minutes *Denaturation:* 95°C for 20 seconds *Annealing*: 64°C for 20 seconds *Elongation*: 72°C for 20 seconds *Cycle*: 25 *Final Elongation*: 72°C for 5 minutes

2.2.5 Small Scale Bacterial Expression of AaCHYMO-Z Full Length No Leader/pET28a pDNA in SHuffle T7

Removal of the leader sequence should facilitate the soluble expression of AaCHYMO by preventing cytoplasmic aggregation due to the hydrophobicity of the leader sequence[26], but because of the reducing cytoplasm in BL21(DE3) cells, there is still a barrier to the formation of disulfide linkages required for tertiary stability and protein function. Therefore, the newly cloned and purified AaCHYMO-Z Full Length No Leader/pET28a pDNA was expressed in Shuffle T7 *E. coli* competent cells (New England Biolabs; Cat#C2566H). Transformation of 50 ng of the AaCHYMO No Leader pDNA into 25 µL of SHuffle T7 cells was followed as described in the manufacturer's protocol and plated on LB agar plates containing 30 ug/mL kanamycin. Plates were incubated overnight at 30°C. For small-scale growth experiments, overnight cultures and growth conditions were set as described above for AaCHYMO-Z Full Length with Leader/pET28a pDNA (see Section 2.2.2). However, SHuffle T7 cells have an optimum

growth temperature of 30°C, and so the initial growth and expression experiments post-induction were all conducted at this temperature. Multiple small scale bacterial expression experiments were done to determine the optimal conditions required to harvest the highest yield of full length AaCHYMO No Leader in zymogen form, which included lowering the temperature to 12°C after induction with IPTG. Timepoints were collected from the small-scale experiments and analyzed via SDS-PAGE and stained with SimplyBlue as explained in Section 2.2.3.

2.2.6 Large Scale Bacterial Expression of AaCHYMO-Z Full Length No Leader/pET28a pDNA in SHuffle T7

Once the optimal soluble expression of recombinant AaCHYMO was achieved, large scale expression of AaCHYMO-Z Full Length No Leader/pET28a in Shuffle T7 cells was set. The differences, however, involved using larger volumes of overnight cultures (scaled up to either 50 or 100 mL) and the actual growths were set in 6 x 1 L Erlenmeyer flasks containing 500 mL of desired media and 30 ug/mL kanamycin. The growth experiments were stopped at the time-points that gave the maximal soluble expression of AaCHYMO-Z Full Length No Leader (determined from the small-scale growth experiments) and the cells were subsequently harvested through cold (4°C) centrifugation at 6000 rpm for 10 minutes, with the pellets flash frozen in liquid N_2 and stored in the -80°C freezer.

2.2.7 Purification of Recombinant N-terminally His_6-tagged AaCHYMO-Z (WT) Full Length No Leader Protease

Cell paste containing AaCHYMO-Z Full Length No Leader recombinant protease collected from large-scale growths (Section 2.2.6) were solubilized (5 mL per gram of

cell paste) in lysis buffer (Buffer A: 10 mM imidazole, 250 mM NaCl, and 20 mM Tris-HCl pH 7.2). The solubilized cell paste was sonicated at 30% intensity with a pulse setting of 15 seconds on and 30 seconds off (while on ice), for a total sonication time of 9 minutes. A total sample of 0.5 µL was collected after sonication and diluted with 19.5 µL Millipore water and 6X SDS Loading Dye (final 1X). The sonicated lysate was then cold centrifuged (8°C) at 16,000 rpm for 30 minutes. A 1.0 µL volume of the supernatant (clear lysate) was collected (soluble sample) and prepared the same way as the total sample. The entire supernatant was loaded onto a 50 mL superloop and purified using a 5 mL Histrap FF Ni^{2+} Column (GE Healthcare; Cat# 17-5255-01) on the AKTA Pure L1 FPLC using lysis buffer and elution buffer (Buffer B: 500 mM imidazole, 250 mM NaCl, and 20 mM Tris-HCl pH 7.2). The fractions (20 µL), flowthrough (10 µL), total and soluble samples were analyzed via SDS-PAGE. The fractions containing the protein of interest were pooled together in Dialysis Membrane Tubing (10K NMWL Fisher Scientific, Cat# 08-667D) and dialyzed overnight at 4°C in a 2 L graduate cylinder with 20 mM Tris-HCl pH 7.2 and 10 mM $CaCl_2$. The buffer was replaced with fresh dialysis buffer and dialyzed for another 4-5 hours. The dialyzed protein was cold centrifuged (4°C) at 3500 rpm for 30 minutes to remove any precipitate. Supernatant was transferred to a 10,000 NMWL 15 mL Regenerated Cellulose Centrifugal Filter (Merck, Cat#UFC901024) and centrifuged at 3500 rpm (4°C) until a final volume between 500 µL and 1 mL was achieved. The concentration of the protease was estimated using the BCA Protein Assay Kit (Pierce; Cat# 23227). SDS-PAGE was used to analyze the purity

of the purified protein using various amounts of the diluted purified protease (starting at 5 mg/mL).

2.2.8 Activation of Purified Recombinant N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader

Purified AaCHYMO (WT) Zymogen with no leader was activated via dialysis using a 2 L graduated cylinder with 20 mM Tris-HCl, 10 mM $CaCl_2$, and 170 mM NaCl at room temperature. The protein samples were diluted to a total volume of 10 mL with 20 mM Tris-HCl pH 7.2 at a final concentration of 0.98 mg/mL. Timepoints were collected at various hours at volumes of 1 µL, 5 µL, and 10 µL and analyzed via SDS-PAGE gel until AaCHYMO zymogen was fully converted to the active form (no presence of the zymogen band).

2.2.9 *In vitro* Suc-AAPF-pNA Spectrophotometric Assays of Recombinant Fully Active AaCHYMO (Mature Protease)

The activated AaCHYMO protease was tested for chymotrypsin-like activity using the synthetic chymotrypsin substrate N-Succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanine 4-nitroanilide (Suc-AAPF-pNA) (Bachem Cat. # L1400.0050) dissolved in DMSO. Upon cleavage of the peptide bond of the C-terminal end of Phe on the substrate the *p*-nitroaniline chromophore is released (with a molar extincition coefficient (ε) of 8,800 $M^{-1}cm^{-1}$).[27] A concentration of 50 nM activated AaCHYMO was incubated in 20 mM Tris-HCl pH 7.2, 10 mM $CaCl_2$, Millipore water, and varying substrate concentrations of Suc-AAPF-pNA (from 1 µM to 2506 µM). The absorbance was measured at 410 nm as a function of time (3-10 minutes) using a Cary 60 UV-visible spectrophotometer. The measurements were conducted in

triplicate and the slopes per substrate concentration were determined using XCEL giving the initial velocity at the given substrate concentration. These were used to determine kinetics parameters using Graphpad Prism and reported as mean values with their respective standard error of the mean (± SEM).

2.2.10 Multiplex Substrate Profiling by Mass Spectrometry of Mature AaCHYMO

To determine the initial amino acid specificity of the AaCHYMO protease, activated AaCHYMO was sent to the University of California, San Diego (UCSD) for analysis using multiplex substrate profiling by mass spectrometry (MSP-MS).[28] For this experiment, 50 nM active/mature AaCHYMO was incubated with 10 μM peptides (14-mer sequences with different combinations of amino acids), and at different time-points post-incubation, samples were analyzed via liquid chromatography-tandem mass spectrometry system (LC-MS/MS) as described in O'Donoghue *et al*.[28] The approach was used to develop the initial substrate signature of active AaCHYMO.

2.3 Results

The open reading frame of the full length AaCHYMO zymogen with leader sequence was successfully amplified from *Ae. aegypti* midgut cDNA. Gene-specific forward and reverse primers were annealed at 60°C which incorporated both the NdeI and HindIII restriction sites into the amplified gene. In addition, the gene sequence also contained a poly(A) tail at each end to serve as an overhang for restriction enzyme attachment improving cleaving efficiency.[29] Sequencing analysis confirmed successful cloning of the 817 base pairs (bp) AaCHYMO PCR product. In addition, the translated amino acid sequence was used to confirm the presence of the N-terminal His_6-tag (Figure 7), but also the highly conserved catalytic triad (Asp-His-Ser) found in all serine proteases and the three-disulfide bridges required for structure and function. The predicted molecular weight of full-length N-terminally His_6-tagged AaCHYMO with Leader is approximately 31.1 kDa with a pI of 5.45.

RRYTMGSS**HHHHHH**SSGLVPRGSHMAFKLTVAFLLVASL
ALASSRA**THK***IVGGDEAEAHEFPYQISLQWNFNDGQTETMH*
FCGASVLNENFVLTAAHCKTAYSNTGFIEVVAAEHDVAVAEGS
EQRRLVAEFIVHEDYQGGVSPDDIAVIRVDKPFELNDKVKAVK
LPKQLEQFDGDVTLSGWGSVSTTVFPDYPDKLRKVVLPLVDY
EQCDTLWGNDSALAKSNVCAGPIDGSKSACSADSGGPLVKQS
GEEVIQVGVVSWGAVPCGSPRRPTVFAGVSHYVDWIEQQLRA

Figure 7. Predicted amino acid sequence of AaCHYMO-Z (WT) Full Length with the Leader Sequence obtained from DNA and amino acid sequencing analysis. The ribosomal binding site and N-terminal His_6-tag linker were identified. The mature sequence that will be released after proteolytic cleavage at the C-terminal end of the pro-peptide sequence (THK), is labelled in italics.

To determine if the AaCHYMO-Z Full Length with Leader/pET28a pDNA can be expressed in bacteria, the most common overexpression bacterial cells BL21(DE3) were

used.[30] The cells were grown in LB media at 30°C and expression was induced with IPTG, which successfully yielded total expression of the protein of interest. However, no soluble expression was observed (data not shown). This was not surprising, since recombinant expression of the most abundant mosquito trypsin-like proteases expressed in BL21(DE3) led to similar results.[18] However, the researchers indicated that the result is due to both the reducing cytoplasm in bacteria and the hydrophobic nature of the leader (signal) sequence. To overcome this, the leader sequence in AaCHYMO was removed and cloned in to the pET28a vector, producing the AaCHYMO-Z Full Length No Leader/pET28a pDNA construct. In addition, the leaderless construct was expressed in SHuffle T7 cells, bacterial cells with a more oxidizing cytoplasm, to ensure proper formation of the three disulfide linkages present in the native protein target[20]. Each variable was tested individually and failed to yield significant soluble expression (data not shown), so both variables were applied since the goal is to obtain abundant quantities of enzyme for downstream biochemical studies.

Small-scale growth in SHuffle T7 using LB media at 30°C yielded soluble forms of the leaderless AaCHYMO protease (28.8 kDa, pI 5.27) (Figure 8). A band slightly above the 30 kDa molecular weight marker was visually detected starting at 1-hour post-induction with IPTG, which is the predicted N-terminally His_6-tagged AaCHYMO No Leader Zymogen. The band intensity increases overtime until the 24th hour time-point, where the intensity of the band significantly decreases. A secondary band can be initially observed at S2 (2 hrs) and increasing over time (Figure 8). Initial activity assays of the growth timepoints with Suc-AAPF-pNA showed significant catalytic activity at

timepoints containing the secondary band (data not shown). This could indicate possible autocatalysis with the secondary band representing the active mature form of AaCHYMO, which is surprisingly unexpected, since chymotrypsins (in mammals) have always been activated through proteolytic cleavage by trypsins.[23,31]

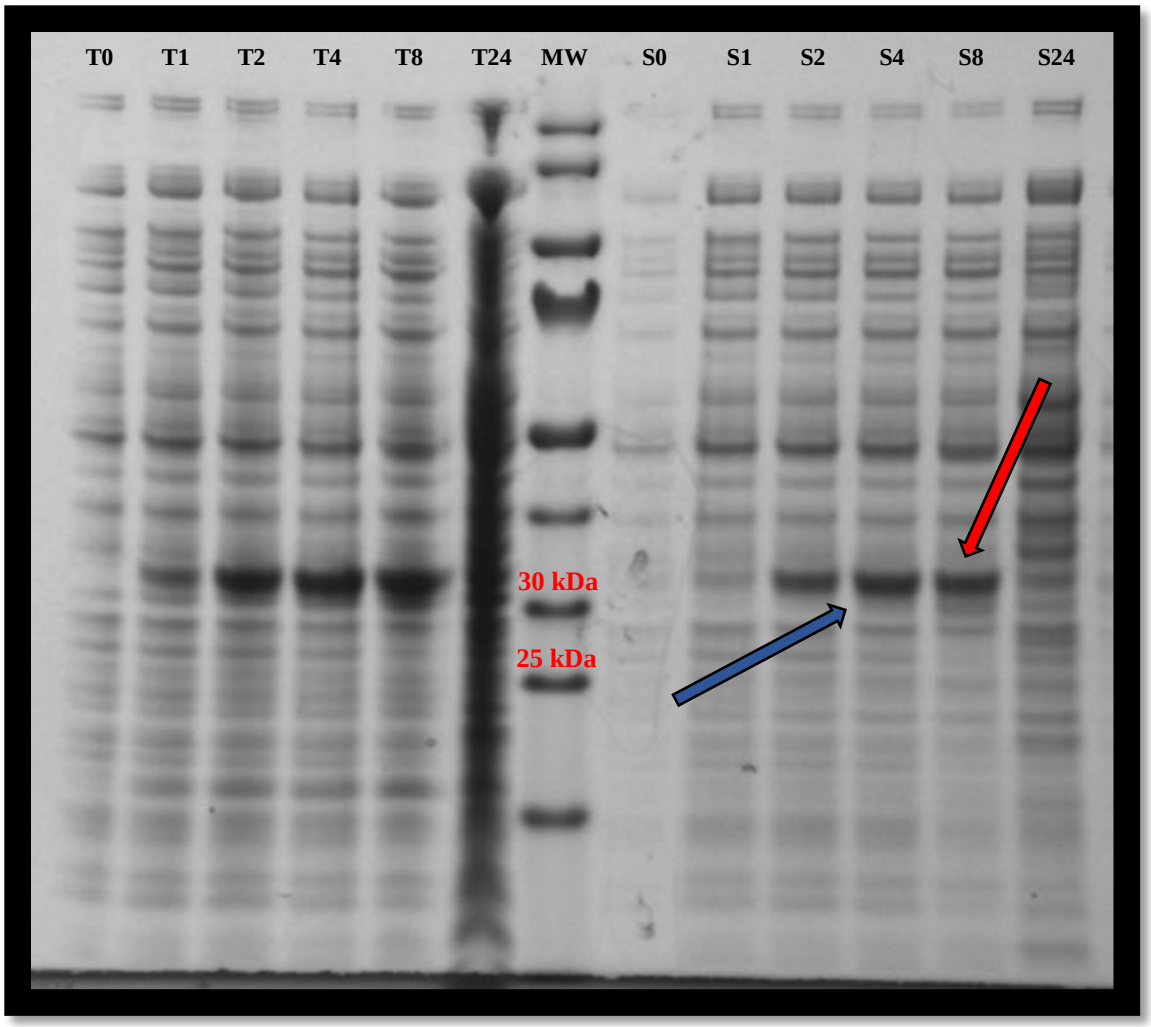


Figure 8. SDS-PAGE analysis of N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader/pET28a expressed in SHuffle T7 cells at 30°C in LB media induced with 0.1 mM IPTG at an OD_{600} of 0.5-0.8. The gel was stained with SimplyBlue. Total and soluble samples are indicated by (T) and (S), respectively. The numbers associated with the samples are the time-points (in hours) that were collected during the growth experiment. The predicted molecular weight of the N-terminally His_6-tagged AaCHYMO Zymogen without the leader sequence is approximately 28.9 kDa. A single band (red arrow) can be observed 1 h post-induction and is increasing over time until the 24^{th} h time-point. A secondary band (blue arrow) with a smaller molecular weight is initially observed at S2 beneath the 30 kDa marker, increasing over time.

If the secondary band is the active mature AaCHYMO, this means that the N-terminal His_6-tag upstream from the pro-peptide sequence is lost which will prevent the protein

target from being purified using a HisTrap Ni^{2+} FF Nickel column. Therefore, expression of the N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader/pET28a pDNA in SHuffle T7 cells was optimized through induction at 12°C (Figure 9). Expression at lower temperatures serves several purposes: it slows down bacterial growth and the translation process allowing the nascently expressed proteins enough time to fold properly.[30,32] It also decreases the activity of the protease preventing any potential autocatalysis. To compensate for the slow bacterial growth and expression, the growth experiment had to be extended beyond 24 hours, which also slowed the activation of the zymogen species to the active mature form, as seen in Figure 9. It can be observed from Figure 8 that the intensity of the zymogen band at S8 (8 hrs) when expressed at 30°C is very similar to that of S26 (26 hrs) or S30 (30 hrs) from Figure 9 when expressed at 12°C.

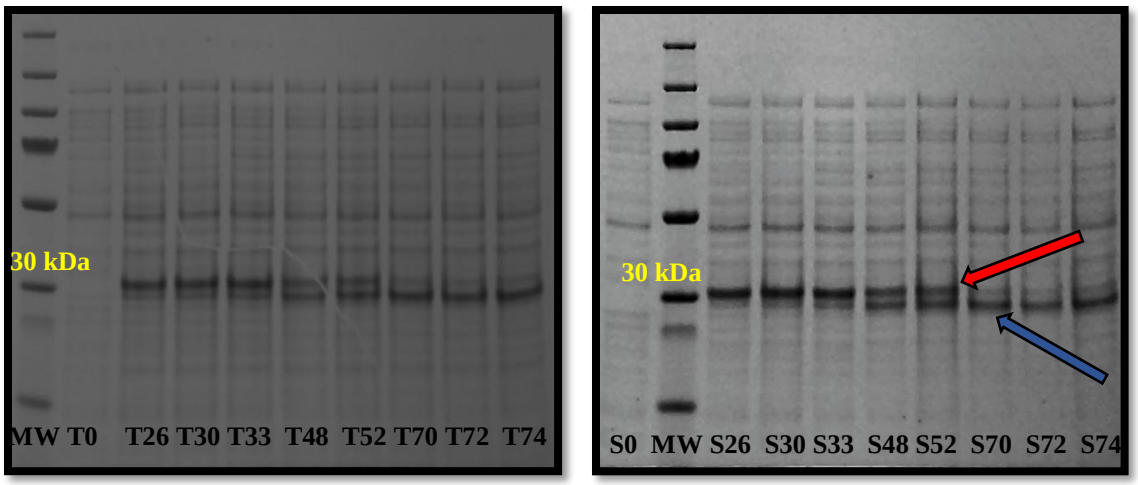


Figure 9. SDS-PAGE analysis of N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader/pET28a expressed in SHuffle T7 cells at 12 °C in LB media when induced with 0.1 mM IPTG at an OD_{600} of 0.5-0.8 and stained in SimplyBlue. Total and Soluble samples are indicated by (T) and (S), respectively. The values associated with the samples are hours upon when the 1mL timepoints were collected. The calculated molecular weight of the full-length zymogen AaCHYMO without the leader sequence is approximately 28.9 kDa. A single band (red arrow) can be initially observed at S26 at the 30 kDa marker and increasing over time until a gradual decrease at S48 and S52. A secondary band (blue arrow), with a smaller molecular weight is initially observed at S26 beneath the 30 kDa marker, increasing over time, until it becomes the dominant species approximately S70-S74.

Based on the optimized small-scale expression of N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader/pET28a in SHuffle T7 at 12 °C in LB media, the maximum expression of the predicted zymogen species occurred between 26 and 30 hours. Therefore, large scale expression of the protease was halted between those hours post-induction. Approximately 8-9 grams of cell paste were collected at this time range when grown at 12 °C in LB media and subsequently purified using a Histrap Ni^{2+} FF Column. Approximately 100-175 mM of imidazole was observed to initiate elution of the target protein from the column (Figure 10) which was then subject to buffer exchange using dialysis to remove imidazole and concentrated to approximately 500 µL-1 mL via

centrifugation. The approximate range of the concentration of purified total protein was 35 mg/mL to 45 mg/mL. A final SDS-PAGE gel was analyzed containing the purified N-terminally His_6-tagged AaCHYMO-Z No Leader at various protein amounts and estimated to be ~95% pure (Figure 11).

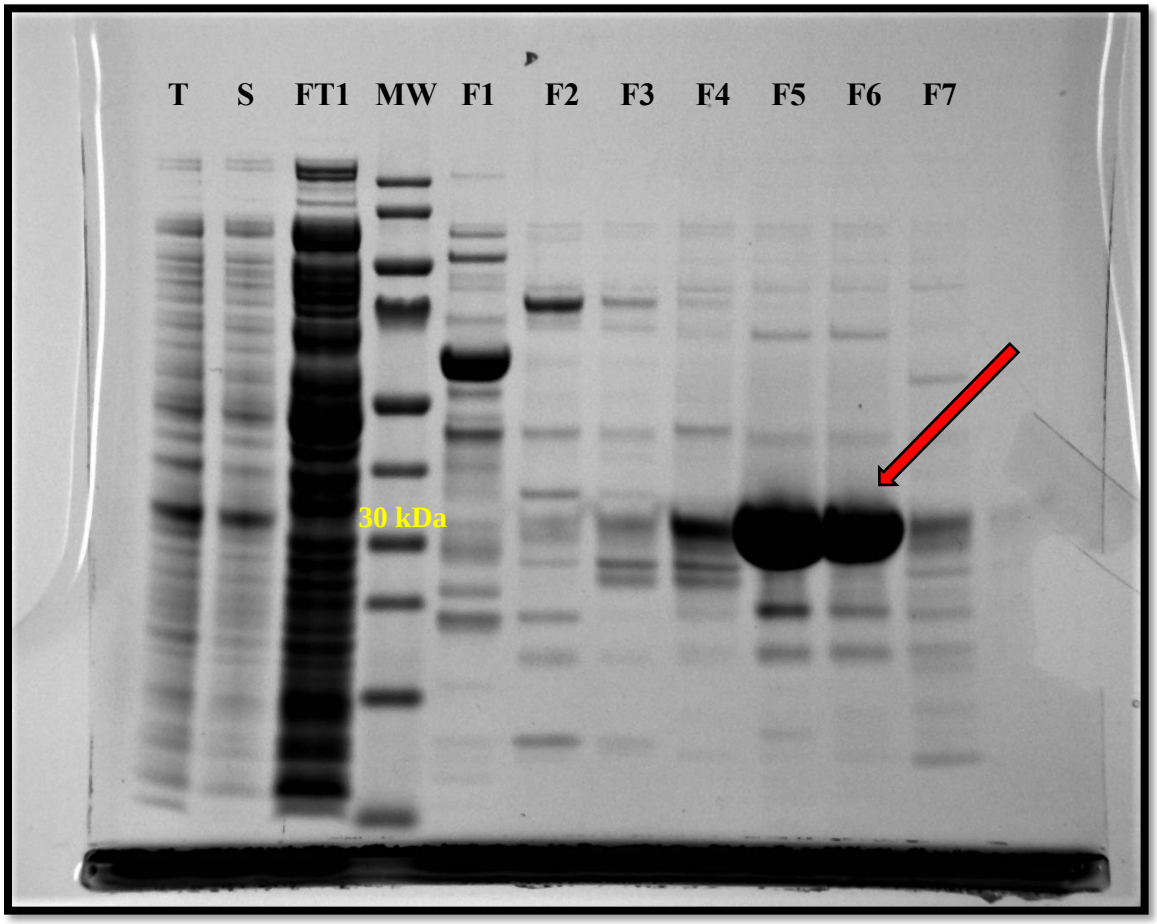


Figure 10. SDS-PAGE analysis containing the total (T), Soluble (S), Flowthrough (FT), and Fractions (F#) collected during Histrap Ni^{2+} Purification of the N-terminally His_6-tagged purified AaCHYMO-Z No Leader eluted with increasing concentrations of imidazole. The single band (red arrow) is the AaCHYMO protease of interest, which is the predicted no leader zymogen with a MW of approximately 28.9 kDa.

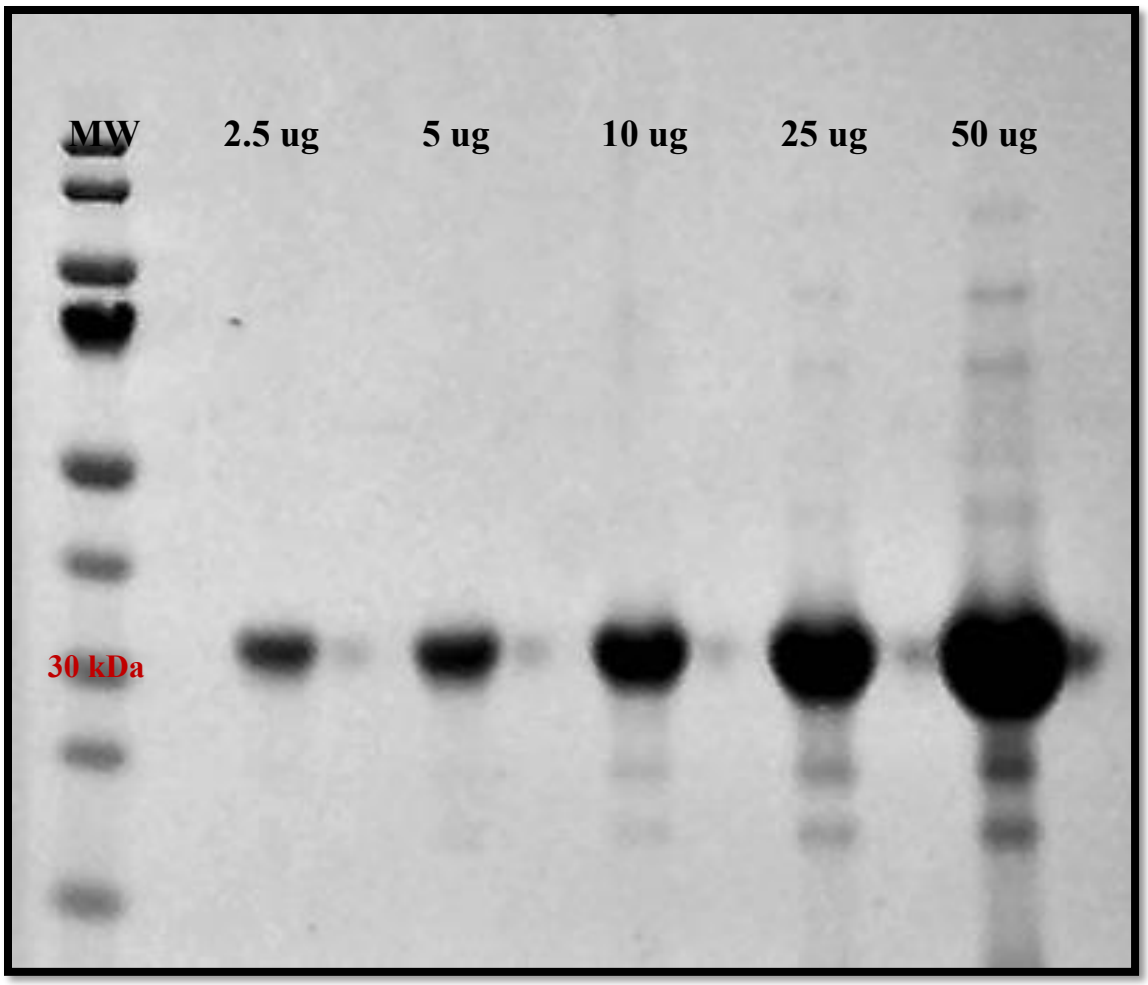


Figure 11. SDS-PAGE analysis of various dilutions of Histrap Ni^{2+} N-terminally His_6-tagged purified AaCHYMO-Z No Leader after dialysis in 2 L buffer containing 20 mM Tris-HCl pH 7.2 and10 mM $CaCl_2$. The dialyzed samples were concentrated to a final volume of 500 µL with a total protein concentration of 44.02 mg/mL. From this concentrated sample, the dilutions were prepared and run on the gel. The single band, which is the predicted no leader zymogen with a MW of approximately 28.9 kDa is observed slightly above the 30 kDa molecular marker.

The newly purified recombinant N-terminally His_6-tagged AaCHYMO-Z No Leader enzyme was spectrophotometrically tested for chymotrypsin-like activity using 1 mM Suc-AAPF-pNA (in 20 mM Tris-HCl pH 7.2 and 10 mM $CaCl_2$), at 410 nm for 10 minutes (Figure 14). Slight activity was observed from the post-dialyzed purified

AaCHYMO protease, although most of the purified species is in the zymogen form. This was not expected, especially since not only is there a clear presence of a single band above the 30 kDa line indicating the zymogen species, this is a chymotrypsin which should not exhibit auto-catalytic behavior. Since this was unexpected, we did not take proper precautions against autocatalysis of AaCHYMO during and after purification. For example, the purified protein was not aliquoted into smaller volumes, and so, during the freeze thawing process increased auto-catalysis of the purified zymogen may have occurred. This was immediately obvious during the activation of the zymogen via dialysis in 20 mM Tris-HCl pH 7.2, 10 mM $CaCl_2$, and 170 mM NaCl. As seen in Figure 15, a clear secondary band below the expected zymogen species band at T0 is observed, but after full activation (72 hours, ~3 days) only one single band is observed (the active mature AaCHYMO protease).

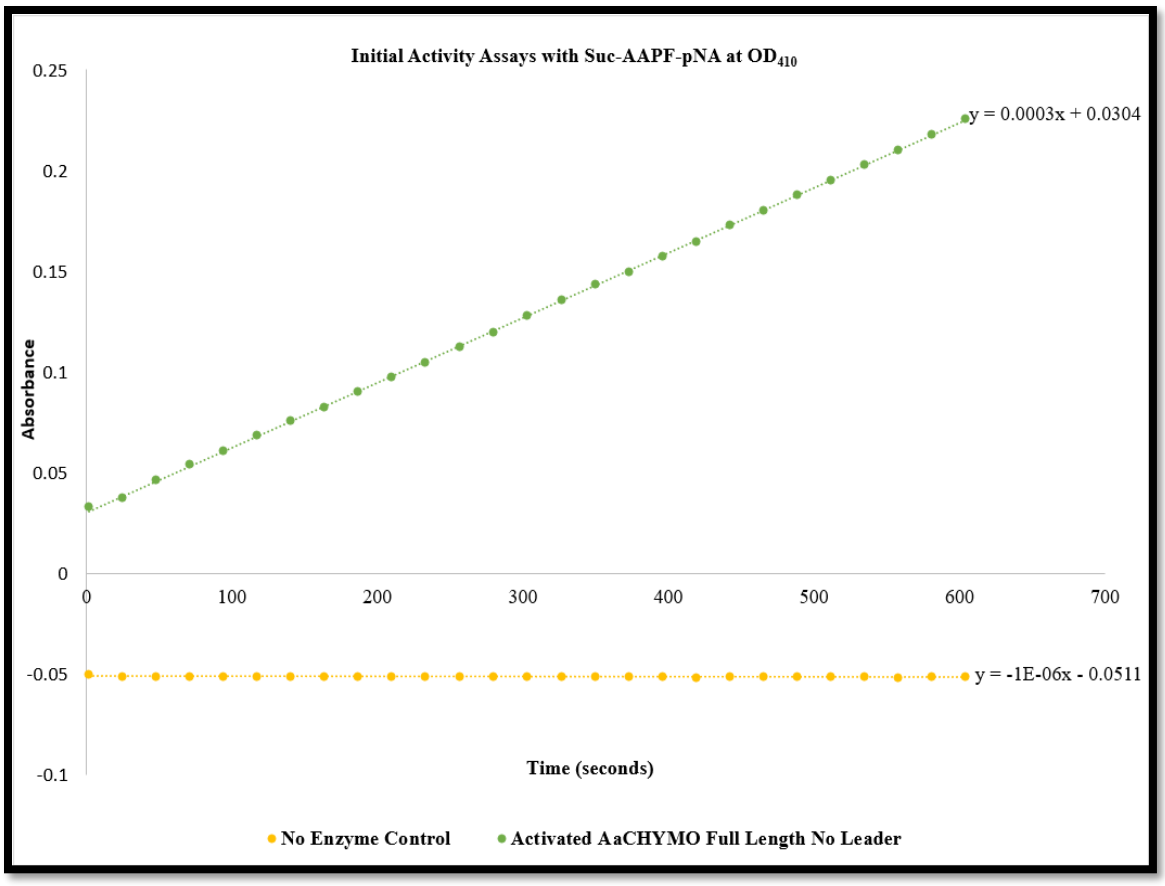


Figure 12. Initial activity assay for post-dialyzed zymogen AaCHYMO Full Length No Leader purified using the Histrap Ni^{2+} FF column. Chymotrypsin-like activity was tested using Suc-AAPF-pNA in 20 mM Tris-HCl pH 7.2 + 10 mM CaCl2. The observed activity (individually tested for each purified construct) for AaCHYMO was compared to a no enzyme control.

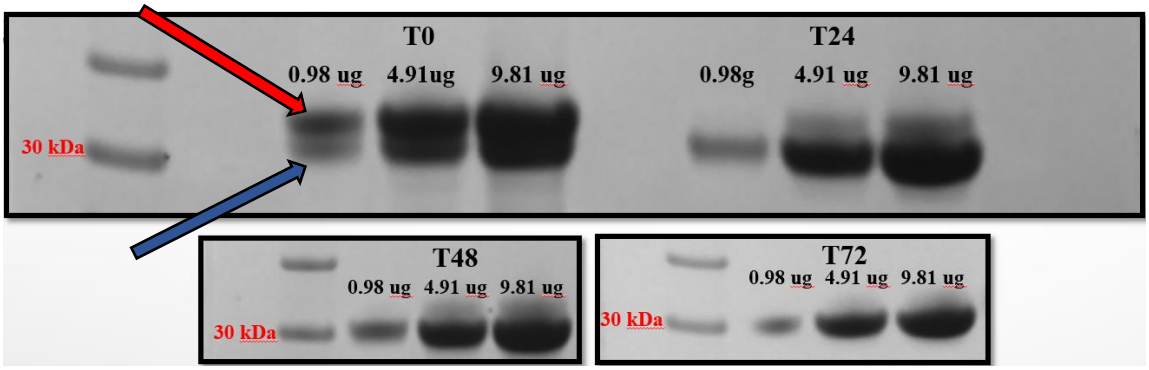
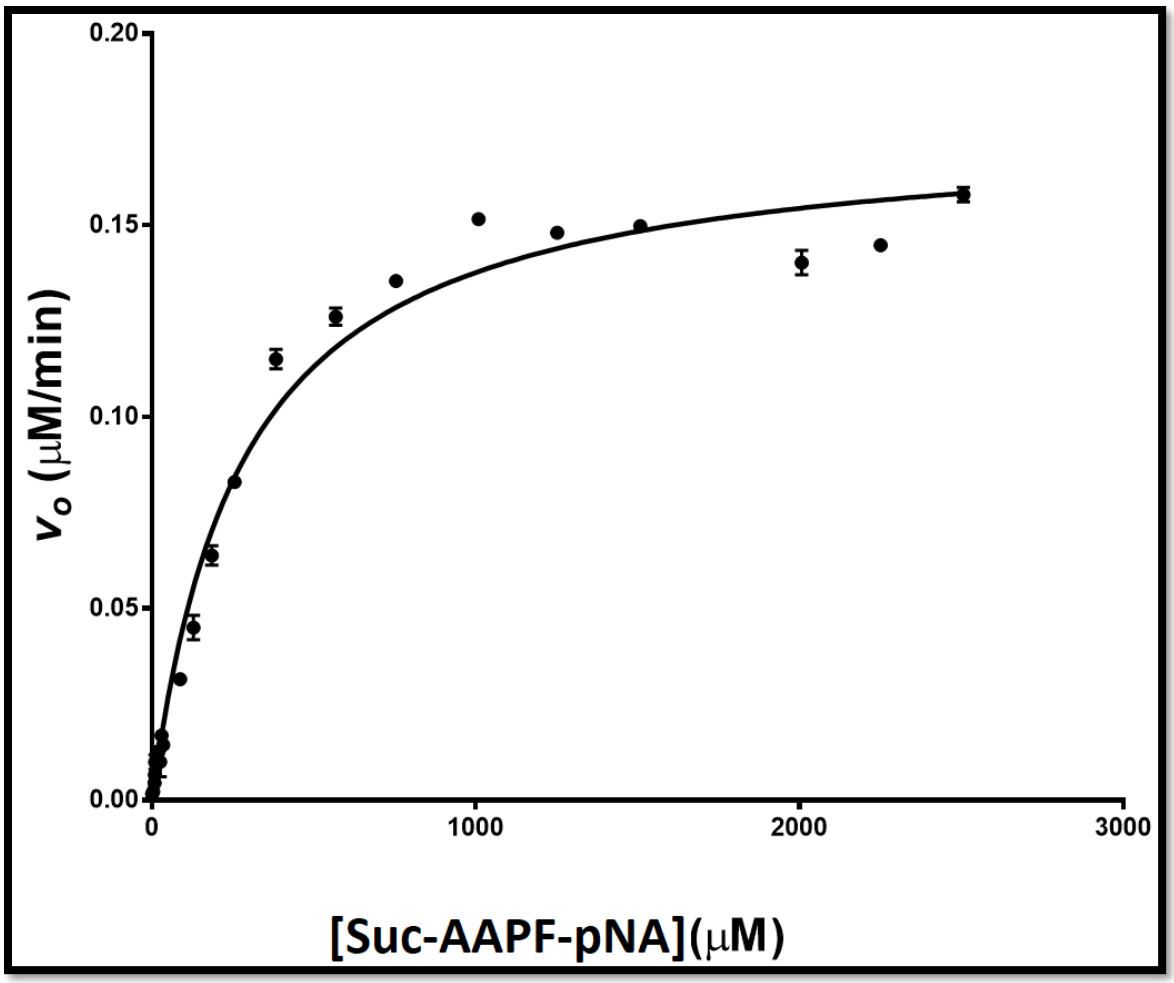


Figure 13. SDS-PAGE analysis of samples collected before (T0), during (T24, T48) and post (T72) sequential activation of N-terminally His_6-tagged purified AaCHYMO-Z No Leader. At T0, observed partial auto-activation was observed while incubated on ice. The activation was set at room temperature (mean temperature of 23°C) in 2L buffer with 170 mM NaCl, 20 mM Tris-HCl pH 7.2, and 10 mM $CaCl_2$.Timepoints were collected in triplicates (containing various amounts shown above) which showed full activation of the zymogen protease at approximately 3 days.

Once AaCHYMO was activated, initial kinetic parameters were determined using various concentrations of Suc-AAPF-pNA in order to create a Michaelis-Menten plot. The measurements were accomplished in triplicates at 410 nm using 50 nM of activated AaCHYMO. Graphpad Prism was then used to determine the mean and their respective ±SEM values for each triplicate reaction rate (initial velocity in μM/min) collected and fit into a Michaelis-Menten curve (Figure 14). The kinetic parameters were determined and are reported as a table in Figure 4. One unit of AaCHYMO enzyme is characterized as cleavage of 1.0 μmol Suc-AAPF-pNA to *p*-nitroaniline per mg of total purified protein.



Protease*	K_m (μM)	k_{cat}(s^{-1})	k_{cat}/ K_m (mM^{-1} s^{-1})	Specific Activity (μmol min^{-1} (mg of protein)$^{-1}$)
AaCHYMO	277.2±17.43	0.05863	0.2115	0.1337

Figure 14. Michaelis-Menten curve showing steady state parameters of 50 nM active, mature AaCHYMO using various concentrations of Suc-AAPF-pNA as a substrate. Graphpad Prism was used to determine the kinetic parameters as shown in the table below the Michaelis-Menten plot. Reaction conditions: 20 mM Tris-HCl at pH 7.2 and 10 mM $CaCl_2$, 170 mM NaCl at 23°C. SEM was determined from triplicate spectrophotometric measurements at each [Suc-AAPF-pNA]. Slopes from each measurement was fitted into the Michaelis-Mented plot with an R^2 value of 0.9856.

Once the kinetic parameters of active AaCHYMO were determined, the initial substrate signature of the protease was then determined through incubation of the enzyme with a 14-mer peptide library, which contain different combinations of amino acids allowing for specific determination of preferred amino acid residues in the P1 (preferred cleavage site), as well as P2-P4 sites (peptide sites on the N-terminal side of the P1 site) and P1'-P4' sites (peptide sites on the C-terminal side of the P1 site).[28] AaCHYMO was incubated with the substrates and cleavage products were determined via LC-MS/MS at different time-points after incubation. As shown in Figure 15, the signature of preferred amino acids is shown above in the positive percent difference scale, while the non-preferred amino acids are shown below the positive percent difference scale. Based on the signature, mature AaCHYMO prefers to cleave Tyr, Leu, Phe, and Met, which are amino acids commonly cleaved by chymotrypsin-like proteases[31]. Leu and Met residues are again preferred amino acid residues that are recognized at the P2 and P1' sites. Surprisingly, Trp, an amino acid residue usually preferentially recognized by mammalian chymotrypsins, is a non-preferred residue for this mosquito chymotrypsin-like protease.

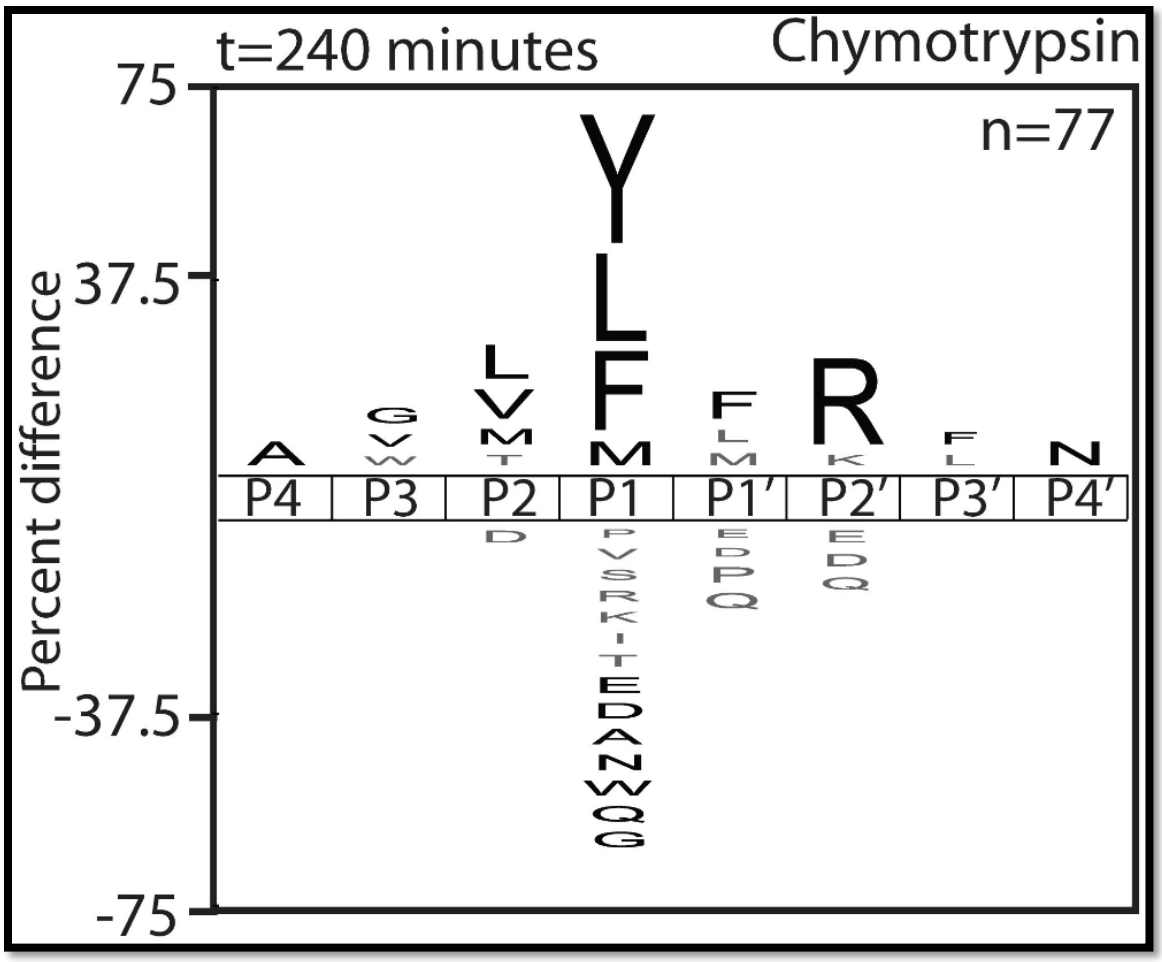


Figure 15. Substrate specificity signature of activated AaCHYMO using multiplex substrate profiling by mass spectrometry (MSP-MS). An enzyme concentration of 50 nM was incubated with 10 μM substrates from a 14-mer peptide library developed by Dr. Anthony J. O'Donoghue.[28] The enzyme was allowed to cleave the peptides while samples were collected at specific time points for a total of 240 minutes, then subsequently analyzed using LC-MS/MS. The preferred amino acid residues are shown above the individual recognition sites, while the non-preferred amino acids are depicted below them.

2.4 Discussion

Although Jiang *et. al.* provided initial *in vitro* work on the full-length zymogen AaCHYMO, they were unable to clearly report soluble expression of the recombinant protease nor produce enough soluble protein for full *in vitro* studies.[23] Thus, in my work, the full-length zymogen AaCHYMO with Leader Sequence was isolated through PCR-amplification from *Ae. aegypti* midgut cDNA, cloned into a bacterial expression vector, and was used to determine the conditions necessary for total protein expression in a bacterial cell line, but failed to produce soluble protein. Two inherent properties of the enzyme had to be addressed to successfully produce soluble protein in a bacterial cell line. First, the leader (signal) sequence from the full-length zymogen contains mostly hydrophobic, nonpolar residues.[26] Since the leader sequence is not recognized by the bacterial system's translocation machinery, protein targets are synthesized and remain in the cytoplasm where the high hydrophobicity of the leader sequence induces misfolding and aggregation of the native conformation of the protease upon expression.[24] Improved soluble expression levels have been observed in other studies involving recombinant expression in bacterial systems after the removal of the leader sequence from their protein targets.[33] Therefore, a similar approach was used for the zymogen AaCHYMO full length with leader sequence.

Second, AaCHYMO contains six cysteine residues that are involved in the formation of three disulfide linkages, which aids in the stability of the protein's tertiary structure and overall activity.[23,34] The cysteine residues must be oxidized prior to the formation of the linkage, which is impossible to accomplish in the naturally reducing cytoplasmic

environment within bacterial expression cells.[24,34] The reducing environment is maintained by the high presence of thiol reductases and glutathione, both of which prevent oxidation of cysteine residues, thereby disallowing formation of disulfide linkages.[35] These post-translational modifications can only be facilitated in highly regulated oxidized compartments, which in the gram-negative and gram-positive bacteria are found in the periplasmic space.[24,36] Unfortunately, the site of recombinant expression in bacteria is in the cytoplasm, which prohibits proper disulfide bond formation.[36] Therefore, a new *E. coli* strain was used, called SHuffle T7 competent cells, which contain mutations in both cytoplasmic reductases thereby creating a more oxidizing cytoplasm.[36] In addition, the mutated *E. coli* strain expresses a chaperone molecule, DsBc isomerase to help in proper disulfide bond formation in mis-oxidized cysteine residues.[37]

As shown in Figure 9, successful soluble expression of AaCHYMO-Z Full Length No Leader was observed when using the SHuffle T7 cells grown at 12°C. Surprisingly, however, a secondary band (not present at 26 h post-induction) became visible. Initial Suc-AAPF-pNA assays on crude lysates led to chymotrypsin-like activity with increasing activity as soon as the secondary species begins to appear (data not shown). Based on these surprising results, we predicted this second species to be the active, mature form of AaCHYMO. This suggests that the mosquito-origin zymogen AaCHYMO has enough activity to auto-catalyze and start activation of its own pro-peptide, a characteristic unheard of in mammalian and invertebrate chymotrypsins.[23,31] Typically, vertebrate chymotrypsinogen is activated by tryptic cleavage of its pro-peptide sequence thereby releasing the active chymotrypsin.[38,39] A similar mechanism is also observed in

invertebrate chymotrypsinogen.[40] Therefore, there was no precedent in anticipating the potential autocatalytic nature of zymogen AaCHYMO. Because of this, we were able to auto-activate the AaCHYMO Zymogen without the leader sequence through incubation in Tris-HCl pH 7.2 buffer, and for the first time, we were able to provide kinetic parameters of active mature AaCHYMO (Figure 14). This finding is important because Jiang *et al.* utilized trypsin to activate the recombinant zymogen AaCHYMO in their studies yielding the active form, but without a clear indication if they had purified away the active trypsin.[23]

It is important to note that Suc-AAPF-pNA was used to determine initial kinetics parameters, which unlike BApNA, a common synthetic substrate used to observe trypsin-like activity, is a tetrapeptide that leads to higher specificity for chymotrypsin-like proteases .[41] However, initial determination of the substrate signature of active AaCHYMO using MSP-MS has revealed the preferred P1 recognition site to be for Tyr, Leu, Phe, and Met residues (Figure 15) making the choice of Suc-AAPF-pNA not ideal for kinetic parameter determination. Therefore, a better substrate with a Tyr residue at the P1 site would be better to gain better kinetic parameters. Furthermore, these initial results do not provide enough information about the substrate specificity of the zymogen form, especially since the amino acid residue that must be recognized for pro-peptide cleavage activation is a Lysine (K) residue (see amino acid sequence on Figure 7). The Lys amino acid is missing from the substrate specificity profile for active AaCHYMO at the P1 site (Figure 15). The difference in substrate specificity between zymogen and active AaCHYMO is expected since the possibility exists that the three regions in the active site

required for efficient catalytic activity [S1 pocket for P1 recognition of the substrate, the localized order of the catalytic triad (Asp-His-Ser) needed to establish a catalytic charge relay system for substrate cleavage, and the oxyanion hole for substrate intermediate stability] are all not in the proper orientations needed for full catalytic activity in the zymogen.[42,43] It is not until activation of the zymogen are the regions in the proper orientation suitable for efficient catalytic activity. Therefore, there is a high possibility that zymogen AaCHYMO could have a different substrate profile than its active counterpart.

Based on the results observed with the N-terminally His_6-tag AaCHYMO Zymogen, it was quite surprising that a chymotrypsin-like zymogen may have enough activity to lead to auto-activation. However, after close examination of Jiang's dissertation work in 1996, they had initially proposed that the zymogen recombinant AaCHYMO produced may have exhibited proteolytic activity using Suc-AAPF-pNA.[23] However, it was unclear which zymogen form of AaCHYMO was enzymatically tested for activity, the full-length construct with the Leader sequence or No Leader sequence. In addition, no further kinetic parameters were provided for the active, mature form of AaCHYMO, which was activated with commercially available bovine trypsin.[23] Hence, the kinetic parameters obtained in this current thesis for the active, mature AaCHYMO cannot be compared to parameters provided in Jiang's dissertation. Therefore, if indeed the zymogen has activity, we must determine the kinetic parameters of the inactive zymogen and compare with the active AaCHYMO protease. This will confirm Jiang's initial observation, but also confirm the auto-catalysis nature of AaCHYMO.

CHAPTER 3

INITIAL BIOCHEMICAL STUDY OF N-TERMINALLY HIS_6-TAGGED AACHYMO MUTANTS (K4A and S197A)

3.1 Introduction

Although soluble expression of zymogen AaCHYMO was successfully facilitated (Chapter 2), the subsequent presence of a secondary band during initial growth experiments (Figure 9), and even the successful auto-activation of the zymogen form to the active mature form (Figure 15) raised several questions to the autocatalytic capability of AaCHYMO. These concerns were raised because chymotrypsinogen isolated from vertebrate and invertebrate species is not known to be autocatalytic, but rather is activated by trypsin.[23,38,44]

Mosquito chymotrypsins should not be an exception to this rule, but Jiang *et al.* reported potential zymogen activity that was not further elaborated or elucidated in their study.[23] We therefore set out to determine if the zymogen form of AaCHYMO does indeed have enough catalytic activity for autoactivation. To do this, using amino acid sequence analysis, we predicted the cleavage site (K4-I5 peptide bond) in the propeptide region that is required for activation using the ExPAsY ProSite protein domain determination tool. Once identified, we used site-directed mutagenesis to mutate the N-terminal amino acid Lys (K4) at the proposed cleavage site, to a small, inert amino acid residue (alanine, Ala). Theoretically, applying this change should prevent the zymogen from auto-cleaving after the propeptide region and maintain the enzyme in the zymogen form. Furthermore, a separate mutation was made for Ser195, the catalytic amino acid in

the catalytic triad, to ensure that the activity is indeed coming from the zymogen form of AaCHYMO and not anything else. This chapter will focus on site-directed mutagenesis used to create the mutant constructs, the optimal bacterial growth conditions used to solubly express the mutant AaCHYMO proteases, purification, and enzymatic assays.

3.2 Methods

3.2.1 Design and PCR-amplification of the Full Length AaCHYMO Zymogen No Leader (K4A)/pET28a and (S197A)/pET28a pDNA Mutant Constructs

The mutant constructs, AaCHYMO MU (K4A) and AaCHYMO MU (S197A) were both designed using gene-specific primers containing the desired mutations (Table 5). The Quickchange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Cat# 200522-5) was used to produce the mutation in each construct using AaCHYMO-Z Full Length No Leader as the amplicon source and amplified using the PCR settings shown in Table 6. The manufacturer's protocol was followed, and the only deviation was using 500 µL of SOC media and plating 100 µL of the resuspension in LB agar plates supplied with 30 ug/mL kanamycin.

Specific DNA cloning into the pET28a vector was not required for the mutant constructs, because site-directed mutagenesis amplifies the entire plasmid while incorporating the mutation using the designed primers (Table 5). After mutagenic PCR, the reaction was incubated with DpnI in order to digest the methylated parental pDNA without the wanted mutations and transformed into XL1-Blue supercompetent cells that accompanied the kit, following the manufacturer's protocol. It is important to note that according to the manufacturer, usage of other competent cells will lead to a decrease in transformation efficiency or low colony number. Therefore, XL1-Blue cells were used to transform the DpnI digested PCR products and plates were incubated overnight at 37°C, followed by plasmid isolation and subsequent sequencing analysis, which was conducted based on the steps already covered in Section 2.2.1.

Table 5. Primers designed for PCR-amplification of the full lengths AaCHYMO (K4A) and AaCHYMO (S197A) sequences for cloning into the pET28a vector to produce a leaderless AaCHYMO (K4A) Full Length/pET28a and AaCHYMO (S197A) Full Length pDNA constructs. The respective melting temperatures (T_M) are respectively shown. The restriction site for NdeI are bolded in the K4A forward and reverse primers, while the mutated regions are highlighted.

Gene Construct	Primer	Primer Sequence	T_M (°C)
AaCHYMO Zymogen MU (K4A)	AaCHYMO - Zym-pET-Fwd KtoA	5'-**CATATG**ACCCACGCGATCGTCGGTGGCG-3'	81.61
	AaCHYMO - Zym-pET-Rev KtoA	5'-CGCCACCGACGATCGCGTGGGT**CATATG**-3'	81.61
AaCHYMO Zymogen MU (S197A)	AaCHYMO - Zym-pET-Fwd StoA	5'-CTCGGCTGATGCGGGTGGCCCGTTG-3'	81.55
	AaCHYMO - Zym-pET-Rev StoA	5'-CAACGGGCCACCCGCATCAGCCGAG-3'	81.55

Table 6. PCR settings and amplicon source required for site-directed mutagenesis of the full length AaCHYMO-Z No Leader gene to produce the AaCHYMO (K4A) and AaCHYMO (S197A) mutations using the Eppendorf Mastercycler Gradient Thermocycler.

AaCHYMO-ZMU (K4A) No Leader	AaCHYMO (WT) No Leader/pET28a Plasmid (151.8 ng/µL)	1	*Lid Heating:* 95°C for 1 minutes *Denaturation:* 95°C for 50 seconds *Annealing*: 60°C for 50 seconds *Elongation*: 68°C for 7 minutes *Cycle*: 18 *Final Elongation*: 72°C for 5 minutes
AaCHYMO-ZMU (S197A) No Leader	AaCHYMO (WT) No Leader/pET28a Plasmid (ng/µL)	1	*Lid Heating:* 95°C for 5 minutes *Denaturation:* 95°C for 50 seconds *Annealing*: 60°C for 50 seconds *Elongation*: 72°C for 7 minutes *Cycle*: 18 *Final Elongation*: 72°C for 5 minutes

3.2.2 Small Scale Bacterial Expression of AaCHYMO-Z Full Length No Leader/pET28a (K4A) pDNA and AaCHYMO-Z Full Length No Leader/pET28a (S197A) pDNA in SHuffle T7

Both the AaCHYMO-Z Full Length No Leader (K4A)/pET28a and AaCHYMO-Z Full Length No Leader (S197A)/pET28a pDNA constructs were transformed and expressed in SHuffle T7 cells as described in Section 2.2.5. Several small-scale growth experiments were done to test various temperatures (30°C, 15°C, and 12°C) and media (LB or Terrific Broth (TB)) to determine the optimal soluble expression conditions for the constructs.

3.2.3 Large Scale Bacterial Expression of AaCHYMO-Z Full Length No Leader/pET28a (K4A) pDNA and AaCHYMO-Z Full Length No Leader/pET28a (S197A) pDNA in SHuffle T7

Upon determination of the optimal conditions for maximal soluble expression, large-scale expression was set as described in Section 2.2.6. The only exception was that both constructs were grown in TB media at 12°C and induced with 0.1 mM IPTG upon reaching an OD_{600} of 0.5-0.8. Growth was stopped at 27 hours for the K4A construct (~12 grams of cell paste) and 48 hours for S197A construct (~25 grams of cell paste), and the cells were subsequently harvested through centrifugation and immediately flash frozen in liquid N_2 before storing in the -80°C freezer.

3.2.4 Purification and "Activation" of Recombinant N-terminally AaCHYMO-Z Full Length No Leader (K4A) and AaCHYMO-Z Full Length No Leader (S197A) Proteases

The harvested cell paste of both constructs was nickel purified following the steps described in Section 2.2.7. It is important to note that both constructs were purified using brand new 5 mL Histrap Ni^{2+} FF Columns to prevent cross-contamination between protease constructs during purification. After subsequent buffer exchange, concentration, and protease storage, the concentrations for each mutant protease were estimated using the BCA assay kit, followed by "activation" using the same conditions and steps provided in Section 2.2.9.

3.2.5 *In vitro* Suc-AAPF-pNA Spectrophotometric Assays of Recombinant Fully Activated AaCHYMO K4A and AaCHYMO S197A

After both mutant recombinant proteases were "activated," chymotrypsin-like activity was tested using Suc-AAPF-pNA. The enzymatic assay and determination of the kinetic parameters for AaCHYMO Zymogen (K4A) were obtained as described in Section 2.2.9.

3.3 Results

Site-directed mutagenesis using PCR yielded the proper mutations, (K4A) and (S197A), which were confirmed through DNA analysis and the translated amino acid sequence was used to determine the molecular weights of the constructs. The calculated molecular weights of both mutant constructs were predicted to be 28.7 kDa (K4A) and 28.8 kDa (S197A). The newly isolated pDNA mutant constructs were then expressed in SHuffle T7 cells using TB media and grown at 12°C, following induction with 0.1 mM IPTG upon reaching an OD_{600} of 0.5-0.8. Surprisingly, in the bacterial expression of the K4A mutant, a second species with lower MW than 28.7 kDa was observed starting at the 48 h post-induction sample (Figure 16). These results are similar to the band observed during the small-scale expression of WT AaCHYMO-Z Full Length No Leader protease (Figure 9). Based on these results, we predicted that the secondary band might be the active species, but activity assays with the Suc-AAPF-pNA resulted in no significant detectable activity. As for the S197A mutant, only a single band was observed, which is expected since the mutation results in an inactive protease (Figure 17). Activity assays using the crude lysates of the S197A mutant growths using Suc-AAPF-pNA resulted in no observable *p*-nitroaniline release (data not shown).

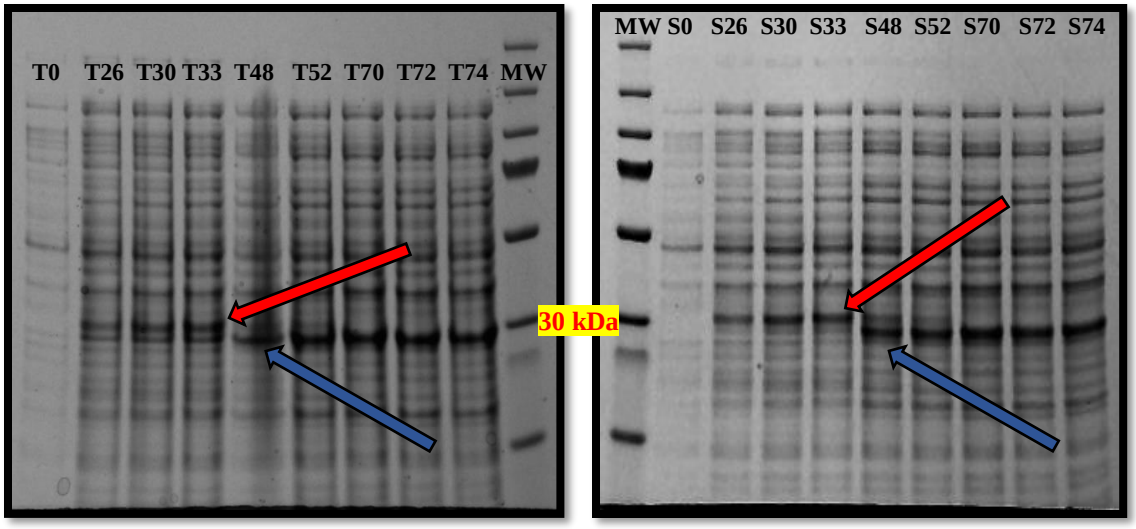


Figure 16. SDS-PAGE analysis of N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader (K4A)/pET28a expressed in SHuffle T7 cells at 12 °C in TB media when induced with 0.1 mM IPTG at an OD_{600} of 0.5-0.8 and stained in SimplyBlue. Total and Soluble samples are indicated by (T) and (S), respectively. The values associated with the samples are hours upon when the 1 mL timepoints were collected. The calculated molecular weight of the full-length zymogen AaCHYMO (K4A) without the leader sequence is approximately 28.7 kDa. A single band (red arrow) can be initially observed at S26 at the 30 kDa marker and increasing over time until a significant decrease at S48 and eventually increasing overtime. A secondary band (blue arrow), with a smaller molecular weight is initially observed at S26 beneath the 30 kDa marker and becomes the dominant species approximately S48-S74.

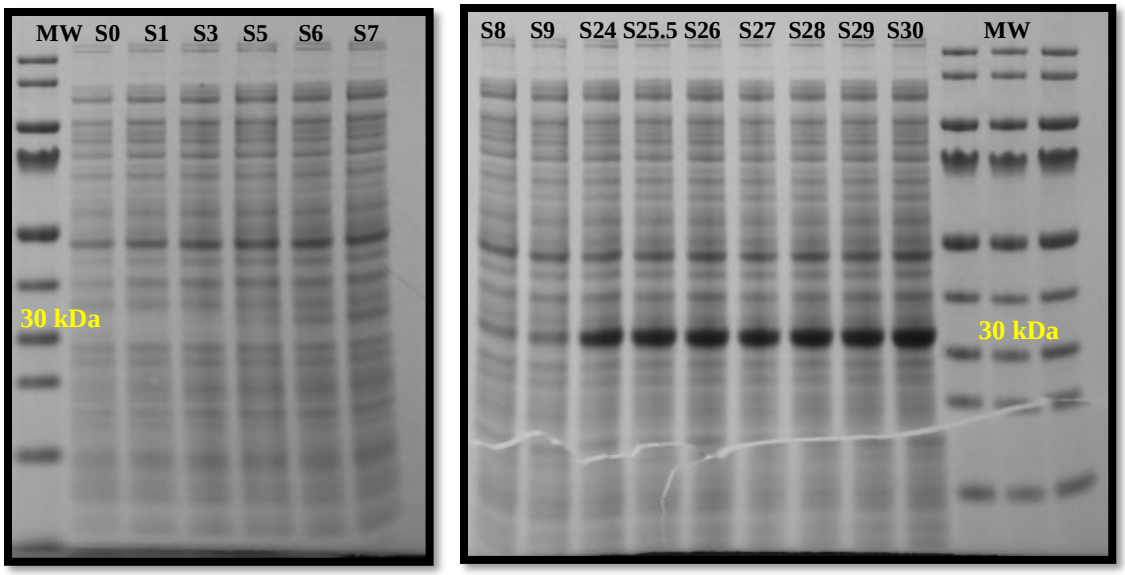


Figure 17. SDS-PAGE analysis of N-terminally His_6-tagged AaCHYMO-Z Full Length No Leader (S197A)/pET28a expressed in SHuffle T7 cells at 12 °C in TB media when induced with 0.1 mM IPTG at an OD_{600} of 0.5-0.8 and stained in SimplyBlue. Total and Soluble samples are indicated by (T) and (S), respectively. The values associated with the samples are hours upon when the 1 mL timepoints were collected. The calculated molecular weight of the full-length zymogen AaCHYMO (S197A) without the leader sequence is approximately 28.8 kDa. A single band (red arrow) can be initially observed at S26 at the 30 kDa marker and increasing over time until a significant decrease at S48 and eventually increasing overtime. A secondary band (blue arrow), with a smaller molecular weight is initially observed at S26 beneath the 30 kDa marker and becomes the dominant species approximately S48-S74.

Maximum expression of the predicted zymogen K4A mutant was observed between 30-33 hours, therefore large-scale expression was stopped between that time range and approximately 12-13 grams of cell paste were collected. Large-scale expression of the S197A mutant was halted at 48 hours, just to harvest as much of the cell paste as possible (~24 grams). Both were purified with new Histrap Ni^{2+} FF columns and a final SDS-PAGE gel post-dialysis was used to determine purity of the mutant proteases (Figure 18).

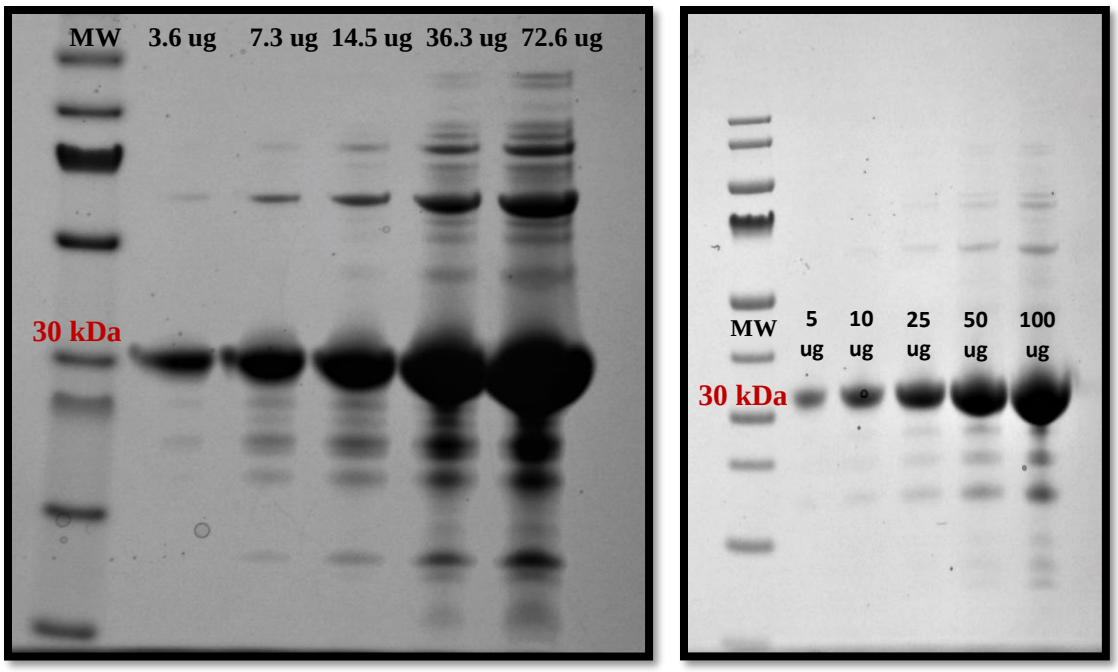


Figure 18. SDS-PAGE analysis of various dilutions of Histrap Ni^{2+} N-terminally His_6-tagged purified AaCHYMO-Z No Leader K4A (*left*) and S197A (*right*) after dialysis in 2 L buffer containing 20 mM Tris-HCl pH 7.2 and10 mM $CaCl_2$. The dialyzed samples were concentrated to a final volume of 500 uL with a total protein concentration of 24.4 mg/mL for the S197A mutant and 7.3 mg/mL for the K4A mutant. From this concentrated sample, the dilutions were prepared and ran on the gel. Single bands, which are the predicted leaderless zymogen species for both mutant proteases at approximately 28.8 kDa, are observed slightly above the 30 kDa molecular marker.

Based on the post-dialysis gels (Figure 18), single bands were observed for both the K4A and the S197A mutants and both constructs were determined to be at least 85% pure for the S197A mutant and 70% pure for the K4A mutant. Both purified mutants were then tested for chymotrypsin-like activity using Suc-AAPF-pNA and compared to that of active AaCHYMO (Figure 19). The S197A mutant did not exhibit any activity, but detectable activity was observed for the K4A mutant, albeit low compared to active AaCHYMO.

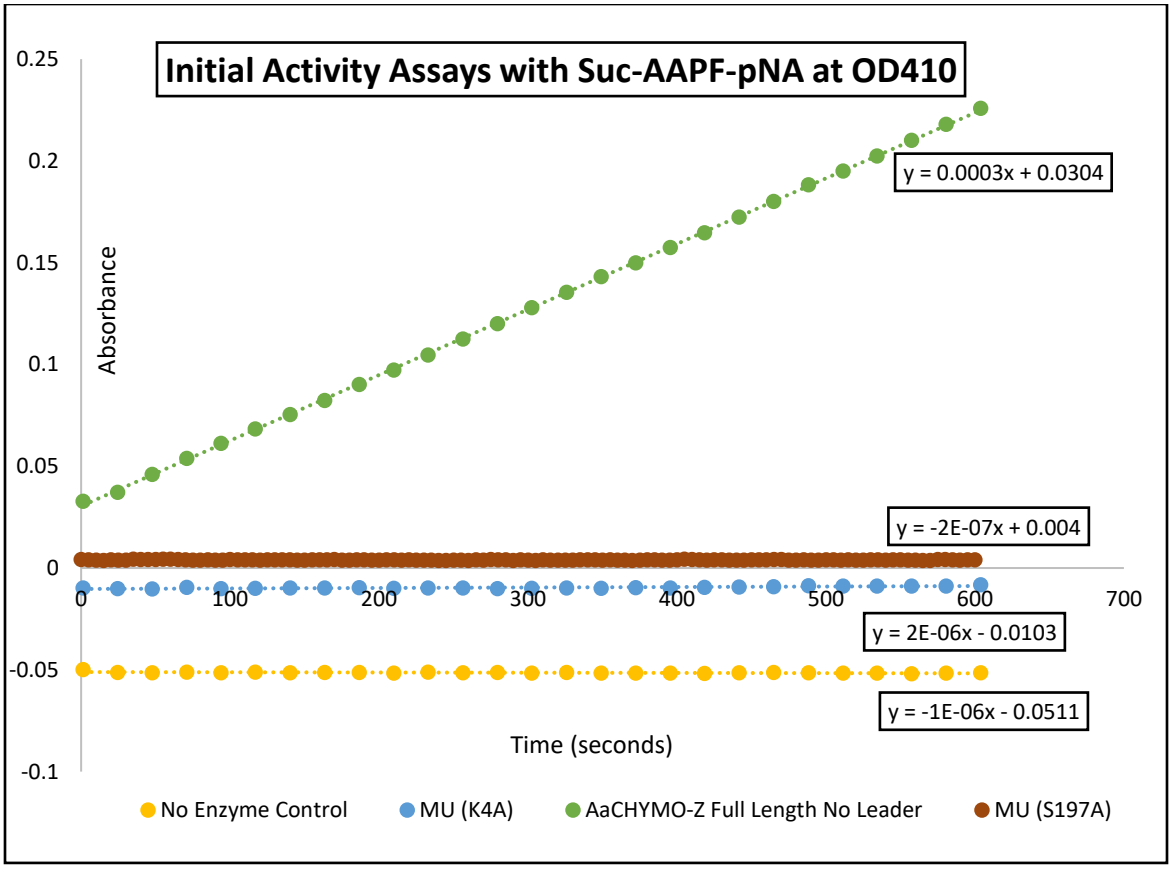


Figure 19. Initial activity assay for post-dialyzed zymogen AaCHYMO Full Length No Leader Mutants K4A and S197A purified using the Histrap Ni^{2+} FF column. The initial rates (Au/min) are compared to post-dialyzed zymogen AaCHYMO Full Length No Leader. Chymotrypsin-like activity was tested using Suc-AAPF-pNA in 20 mM Tris-HCl pH 7.2 + 10 mM CaCl2. The observed activity for AaCHYMO was compared to a no enzyme control.

Before full enzymatic assays were conducted on both the K4A and S197A mutants, the mutant proteases were run through the same activation process as the wildtype AaCHYMO-Z Full Length No Leader (see Section 2.2 and Figure 18). The S197A was diluted to a final concentration of 5.8 mg/mL in a total volume of 10 mL and placed in a 2L graduated cylinder with 20 mM Tris-HCl at pH 7.2, 10 mM $CaCl_2$, and 170 mM NaCl for approximately 60 hours at room temperature (Figure 20).

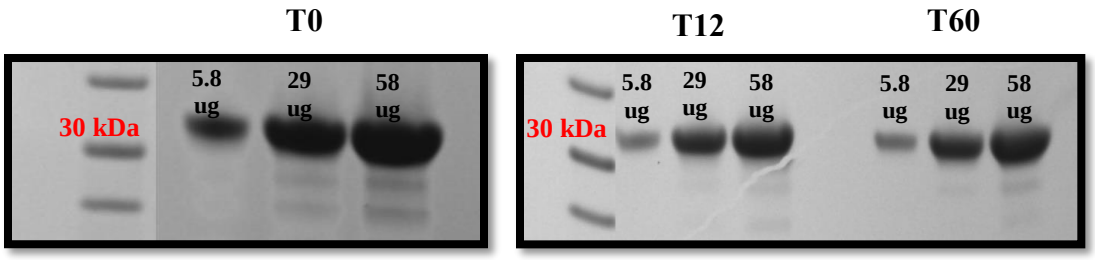


Figure 20. SDS-PAGE analysis of samples collected before (T0), during (T12) and post (T60) sequential activation of N-terminally His_6-tagged purified AaCHYMO-Z No Leader (S197A). The activation was set at room temperature (mean temperature of 23°C) in 2L buffer with 170 mM NaCl, 20 mM Tris-HCl pH 7.2, and 10 mM $CaCl_2$.Timepoints were collected in triplicates (containing various amounts shown above) and show no "activation" of the protease.

Only a single band above the 30 kDa molecular marker was observed for the S197A mutant, indicating the presence of only the full-length with the His_6-tag zymogen species. The K4A mutant was also subject to the same conditions for activation, diluted to a final concentration of 2.1 mg/mL with a total volume of 5 mL. The activation of K4A was set at room temperature using the same buffer components for the activation of the wildtype AaCHYMO-Z Full Length No Leader. It took approximately 9 days (~216 hours) to "activate" the zymogen K4A mutant species (Figure 21). Based on the gel, a single band was observed at T0 above the 30 kDa molecular marker (the zymogen species with the His_6-tag linker), but over time, the appearance of a second band was observed. By T216, a single band is observed at exactly the 30 kDA molecular marker. This shift in molecular weight was predicted to be the "activated" species. Therefore, kinetic parameters were subsequently determined for the K4A and S197A. Both mutant S197A and "activated" mutant K4A were used for enzymatic assays with Suc-AAPF-pNA, using the same process as wildtype active, mature AaCHYMO described in Section 2.2.9.

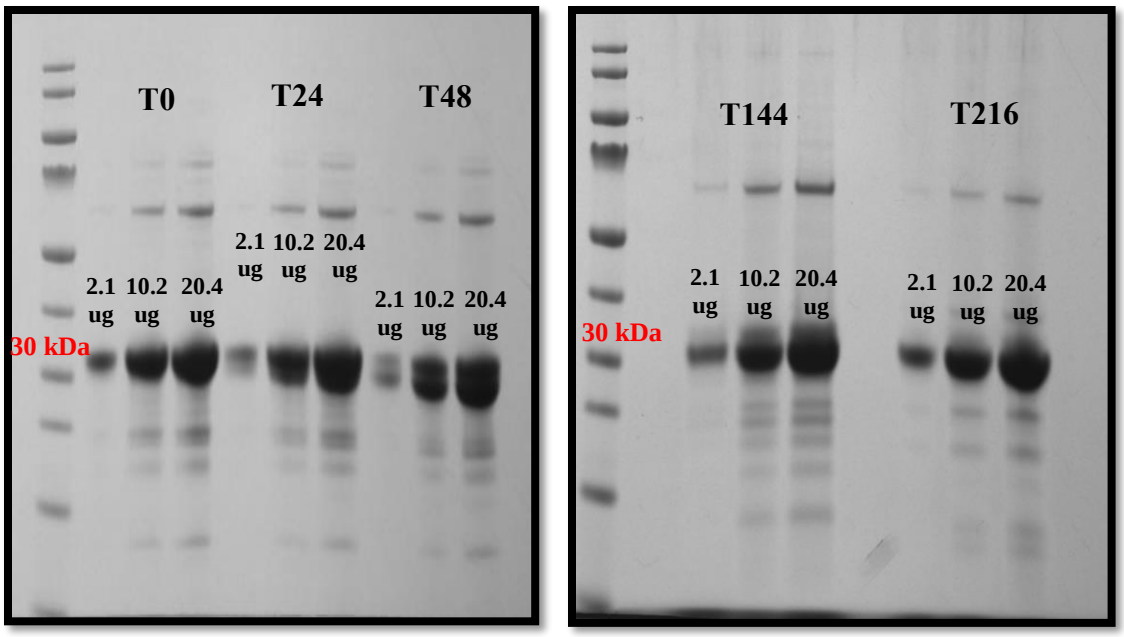
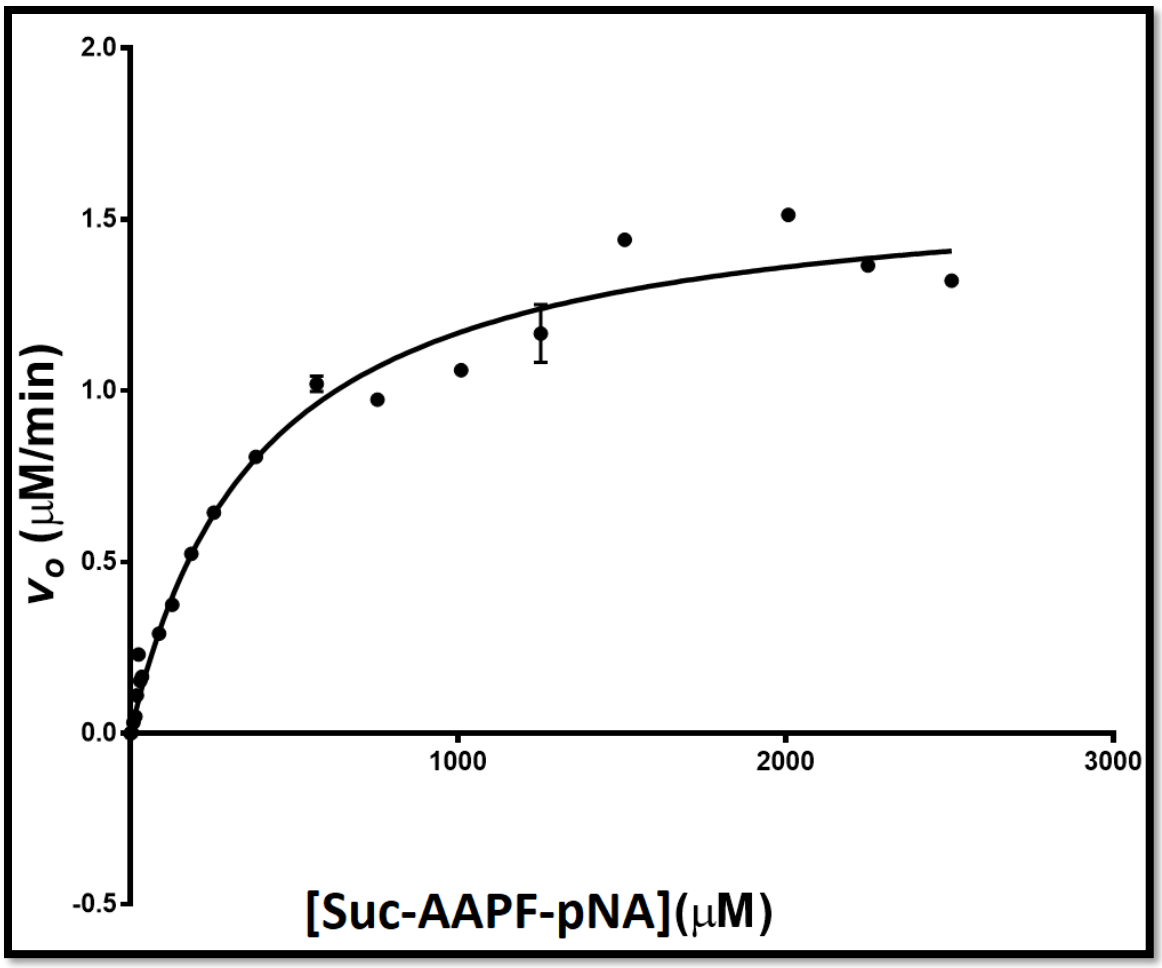


Figure 21. SDS-PAGE analysis of samples collected before (T0), during (T24, T48, T144) and post (T216) sequential activation of N-terminally His_6-tagged purified AaCHYMO-Z No Leader (K4A). The activation was set at room temperature (mean temperature of 23°C) in 2L buffer with 170 mM NaCl, 20 mM Tris-HCl pH 7.2, and 10 mM $CaCl_2$.Timepoints were collected in triplicates (containing various amounts shown above) which showed full "activation" of the zymogen protease at approximately 9 days.

The mutant S197A did not exhibit any activity at any concentration of Suc-AAPF-pNA. However, very surprisingly, we were able to detect activity at the different concentrations tested (1 μM to 2506 μM) using the mutant K4A. Therefore, using steady-steady conditions, the initial velocities at different substrate concentrations were used to create and fit into the Michaelis-Menten plot using GraphPad Prism as shown in Figure 22.



Protease*	**K_m (μM)**	**k_{cat}(s^{-1})**	**k_{cat}/ K_m (mM^{-1} s^{-1})**	**Specific Activity (μmol min^{-1} (mg of protein)$^{-1}$)**
AaCHYMO MU (K4A)	394.8±30.02	0.5433	1.376	1.109

Figure 22. Michaelis-Menten curve showing steady state parameters of 50 nM activated AaCHYMO K4A using various concentrations of Suc-AAPF-pNA as a substrate. Graphpad Prism was used to determine the kinetic parameters as shown in the table below the Michaelis-Menten plot. Reaction conditions: 20 mM Tris-HCl at pH 7.2 and 10 mM $CaCl_2$, 170 mM NaCl at 23°C.SEM was determined from triplicate spectrophotometric measurements at each [Suc-AAPF-pNA]. Slopes from each measurement was fitted into the Michaelis-Menten plot with an R^2 value of 0.9809.

3.4 Discussion

For the first time, large quantities of recombinant zymogen AaCHYMO was successfully expressed in soluble form. This was accomplished without having to rely on denaturing/refolding strategies commonly used to isolate mosquito proteases *in vitro* as was done for the most-abundant trypsin-like proteases.[18] However, the subsequent presence of a secondary band during the expression of the wild-type AaCHYMO No Leader Zymogen protease (Chapter 2) led to the possibility that the zymogen form may have autocatalytic properties. This is especially surprising since both vertebrate and invertebrate chymotrypsins are activated via trypsin cleavage of the pro-peptide region.[23,44] Hence, two questions had to be addressed to answer whether this behavior was possible: 1. Is AaCHYMO truly autocatalytic? 2. Can a chymotrypsin-like protease recognize a non-preferred amino acid (Lys) at the predicted cleavage site leading to the active species? To help answer these questions, two mutant constructs were designed in order to determine the possibility of chymotrypsin autocatalysis.

The first construct is the K4A mutant with the cleavage/activation Lys at the propeptide region mutated to Ala. Surprisingly, expression of the mutant K4A AaCHYMO protease (Figure 16) resulted in very similar results to the expression of the WT AaCHYMO zymogen (Figure 9). Both expression experiments led to the presence of double bands, one band slightly above the 30 kDa molecular marker predicted to be the zymogen species and a secondary band approximately at the 30 kDa molecular marker predicted to be the active form. Therefore, we expected the secondary band observed in the K4A mutant to be active, mature AaCHYMO. However, no observable

chymotrypsin-like activity was observed in the K4A crude lysates (data not shown). This indicates that the K4A mutant may be cleaving at another region upstream of the predicted cleavage site.

At the moment, we are unsure where the zymogen is cleaving, but we predict that it might be somewhere in the thrombin cleavage site (LVPRG) found in the His_6-tag linker region that is introduced when cloning into the pET28a vector (Figure 7). Although, it is difficult to predict at which amino acid in the thrombin cleavage site the K4A mutant is cleaving, based on the specificity profile of the active AaCHYMO form it is likely that it would be the Leu residue (Figure 15). However, the zymogen form of AaCHYMO K4A may have a different substrate specificity profile compared to its active, mature counterpart, which is indicated by the different kinetic parameters obtained. Active AaCHYMO has nearly a 2-fold lower Km value (Figure 14) compared to the K4A mutant (Figure 20), but the overall kinetic parameters (k_{cat}, $k_{cat/}$ K_m, and specific activity) for the K4A mutant are much better compared to the active mature form of AaCHYMO. This indicates that the Phe at the P1 position of Suc-AAPF-pNA is preferred for the mutant rather than the active form, which agrees well with the substrate specificity profile of the active form where Phe is the third preferred substrate after Tyr (Figure 15). Therefore, a better substrate for the mature form is needed to obtain better kinetic parameters and the substrate specificity profile of the K4A mutant needs to be determined to help identify the preferred amino acid residue at the P1 site.

There is a possibility that the K4A mutant (the zymogen form of AaCHYMO) may have an improperly formed active site. Recall, that the inactive zymogen forms of

proteases must be activated to the mature form through cleavage at the peptide bond, which induces an overall conformational change leading to proper formation of the active site and opening of the substrate-binding pocket.[42] Because of this, the specificity of each protease (the K4A mutant and the active AaCHYMO) cannot be directly compared since they are two different enzymes with potentially different substrate-binding pocket and active sites. Once the best substrates are identified for both active and K4A mutant AaCHYMO, and the substrate specificity profile for the K4A mutant identified can we determine the preferred P1 cleavage site and know more about the autocatalytic nature of the enzyme.

To further ensure that any autocatalytic activity was indeed coming from AaCHYMO, a second construct (the S197A mutant) where the nucleophilic Ser 197 at the active site was mutated to Ala was designed. This mutation yielded a "dead" enzyme, which upon soluble expression of the protease led to a single band above the 30 kDa molecular marker throughout the 30-hour growth experiment (Figure 17). This was compared to the expression of wildtype AaCHYMO-Z, which already showed the formation of the secondary band at 26 hours post-induction (Figure 9). Furthermore, chymotrypsin-like activity from the S197A crude lysates from Figure 17 and the purified form that was "activated" (Figure 18) similarly to the wild-type mature AaCHYMO failed to exhibit cleavage of Suc-AAPF-pNA to *p*-nitroaniline. This indicated that the autocatalytic activity observed is indeed coming from the zymogen form of AaCHYMO and not from any bacterial enzymes produced during the growth expression experiments.

Overall, the determined kinetic parameters provide a clear understanding that there is a distinction in the behavior of the active and mutant proteases towards the Suc-AAPF-pNA substrate. The S197A mutant did not result in detectable activity towards Suc-AAPF-pNA, but the K4A zymogen did result in enough detectable activity to determine kinetic parameters and to conclude that AaCHYMO zymogen is autocatalytic. With the determination of the substrate specificity signature for active AaCHYMO (Figure 15), especially the P1 substrate specificity, which is Tyr>Leu>Phe>Met, indicates that the zymogen may have a preferential difference for activation. The specificities for these amino acid residues by active AaCHYMO are very similar to a 1976 study that provided the k_{cat}/Km preference for these same amino acid residues for bovine chymotrypsin in the following order of preference, Tyr>Phe>Met>Leu.[46] Therefore, it not only confirmed the specificity of our chymotrypsin-like protease, but also exhibited similar P1 specificities to the well-studied chymotrypsin model. However, the substrate profile for mature AaCHYMO does not provide insight into the substrate preference for the AaCHYMO K4A zymogen. The P1 specificity for active AaCHYMO alone tell us that it has no specificity for the Lys residue (Figure 15) at the propeptide region, which is the amino acid residue that must be recognized prior to cleaving the scissile bond needed to release the active mature form of the enzyme (Figure 7). Therefore, only a substrate specificity profile for the K4A mutant can potentially provide a clear explanation to the autocatalytic nature of AaCHYMO.

CHAPTER 4

SUMMARY AND FUTURE STUDIES

The idea of inhibiting midgut serine proteases during bloodmeal digestion as a potential novel vector control strategy against *Ae. aegypti* mosquitoes has been proposed since the 1950s, but the majority of studies have focused solely on trypsin-like serine proteases.[6,10,18] Therefore, the work in this thesis focused on one of the minimally studied chymotrypsin-like serine proteases, AaCHYMO, and its *in vitro* isolation and initial biochemical study. Jiang *et al.* provided the only molecular study of AaCHYMO back in 1997, but failed to elaborate on its potential role or impact upon blood meal digestion. Before this work, it was suggested that AaCHYMO is constantly expressed throughout both early and late phases of the biphasic bloodmeal digestion process. However, several discrepancies found in Jiang *et al*.'s study made it difficult to determine the properties inherent to AaCHYMO, which resulted in the weak elucidation of its catalytic function.

In this work, AaCHYMO was successfully recombinantly expressed in soluble form using a bacterial cell line. Obtaining recombinant soluble protein was one of the greatest challenges researchers faced when trying to isolate mosquito proteases *in vitro*, since three disulfide bonds are required for the protein to be properly folded, and in turn relied on denaturation/refolding strategies that have usually resulted in low protein yields.[18,20] The insolubility problems were addressed and the expression of the protease was accomplished through usage of SHuffle T7 bacterial cells with a more oxidizing cytoplasm.[36] As a result of this successful soluble expression, autocatalytic behavior not commonly found in vertebrate and invertebrate chymotrypsins[23,38] was observed in

AaCHYMO. The mutant constructs designed helped to determine the possibility of this observed behavior and resulted in the "activated" AaCHYMO K4A zymogen having enough activity to catalyze Suc-AAPF-pNA and determine kinetic parameters. However, at the moment there is no evidence of the preferred amino acid of the AaCHYMO zymogen form. Therefore, the substrate specificity must be obtained for this enzyme to determine if it has a P1-site preference for Lys. This is important because the predicted amino acid at the propeptide cleavage site is Lys (see Figure 7). At the same time, we can also use N-terminal sequencing to determine the exact cleavage site of AaCHYMO K4A zymogen after the activation process. In addition, we plan to co-incubate the AaCHYMO K4A zymogen with inactive AaCHYMO S197A, and through N-terminal sequencing analysis determine the preferred site of cleavage in the N-terminal His_6-tag linker region leading to the identification of the preferred amino acid cleavage site. This will inform us on where the K4A mutant is cleaving to go from the full-length N-terminally His_6-tagged linker to the no His_6-tagged zymogen.

Although this study has laid the foundation for initial biochemical studies of AaCHYMO, it failed to investigate the specific role of AaCHYMO in the bloodmeal digestion process. Therefore, co-incubation with the natural bloodmeal proteins will tell whether or not the protease can cleave intact globular proteins or if other proteases present during digestion must first act on the bloodmeal proteins, followed by the activity of AaCHYMO. Additionally, we can also determine if there is a relationship between other proteases present during the initial release of AaCHYMO in the early phase. Thus far, the only other serine proteases known to be present during the early phase with

AaCHYMO is AaET (a trypsin-like protease) and AaJHA15 (another chymotrypsin-like protease). Although we have shown that AaCHYMO zymogen can activate to the mature form, AaET may speed-up the process since it does prefer either a Lys or Arg at the P1 position.[18] AaCHYMO has a Lys residue at the propeptide region, so a synergistic effect between the two proteases may exist. Only co-incubation between the two proteases can help with understanding this process. To do this, we can incubate the "dead" AaCHYMO S197A mutant with purified active AaET to determine if this protease is capable of cleaving at the propeptide region, suggesting a secondary mode of AaCHYMO zymogen activation. Overall, the work on AaCHYMO has laid the foundation to fully understanding the blood meal digestion process but may also further advance the vector control strategy of whether inhibiting midgut proteases in general can result in a novel vector control strategy against one of nature's most highly efficient disease vectors.

References

1. Christophers, S. R. (1960) *Aedes Aegypti (L.) The Yellow Fever Mosquito*. New York, NY: The Syndics of the Cambridge University Press: Bentley House.

2. World Health Organization. (2018). Dengue Control. Retrieved from *http://www.who.int/denguecontrol/en/*

3. Briegel, H. Physiological bases of mosquito ecology. *Journal of vector ecology: Journal of the Society for Vector Ecology* **2003**, *28*, 1.

4. Scott, T. W.; Amerasinghe, P. H.; Morrison, A. C.; Lorenz, L. H.; Clark, G. G.; Strickman, D.; Kittayapong, P.; Edman, J. D. Longitudinal Studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: Blood Feeding Frequency. *J. Med. Entomol.* **2000**, *37*, 89-101.

5. Centers for Disease Control and Prevention. (2018). Dengue. *https://www.cdc.gov/dengue/index.html*

6. 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.

7. Ross, R. W. The Newala epidemic: III. The virus: isolation, pathogenic properties and relationship to the epidemic. *J. Hyg.* **1956**, *54*, 177-191.

8. Kamgang, B.; Marcombe, S.; Chandre, F.; Nchoutpouen, E.; Nwane, P.; Etang, J.; Corbel, V.; Paupy, C. Insecticide susceptibility of Aedes aegypti and Aedes albopictus in Central Africa. *Parasites & Vectors* **2011**, *4*, 79-79.

9. Gould, E. A.; Higgs, S. Impact of climate change and other factors on emerging arbovirus diseases. *Trans. R. Soc. Trop. Med. Hyg.* **2009**, *103*, 109-121.

10. Fisk, F. W. Studies on proteolytic digestion in adult *Aedes aegypti* mosquitoes. *Annals of the Entomological Society of America*. **1950**, *43*, 555-572.

11. Noriega, F. G.; Colonna, A. E.; Wells, M. A. Increase in the size of the amino acid pool is sufficient to activate translation of early trypsin mRNA in Aedes aegypti midgut. *Insect Biochem. Mol. Biol.* **1999**, *29*, 243-247.

12. Bian, G.; Raikhel, A. S.; Zhu, J. Characterization of a juvenile hormone-regulated chymotrypsin-like serine protease gene in Aedes aegypti mosquito. *Insect Biochem. Mol. Biol.* **2008**, *38*, 190-200.

13. Noriega, F. G.; Wang, X.; Pennington, J. E.; Barillas-Mury, C.; Wells, M. A. Early trypsin, a female-specific midgut protease in Aedes aegypti: Isolation, amino-terminal sequence determination, and cloning and sequencing of the gene. *Insect Biochem. Mol. Biol.* **1996**, *26*, 119-126.

14. Kalhok., S. E.; Tabak, L. M.; Prosser, D. E.; Brook, W.; Downe, A. E. R.; White, B. N. Isolation, sequencing and characterization of two cDNA clones coding for trypsin-like enzymes from the midgut of Aedes aegypti. *Insect Mol. Biol.* **1993**, *2*, 71-79.

15. Barillas-Mury, C.; Graf, R.; Hagedorn, H. H.; Wells, M. A. cDNA and deduced amino acid sequence of a blood meal-induced trypsin from the mosquito, Aedes aegypti. *Insect Biochemistry.* **1991**, *21*, 825-831.

16. Graf, R.; Briegel, H. The synthetic pathway of trypsin in the mosquito Aedes aegypti L. (Diptera: Culicidae) and in vitro stimulation in isolated midguts. *Insect Biochemistry.* **1989**, *19*, 129-137.

17. Felix, C. R.; Betschart, B.; Billingsley, P. F.; Freyvogel, T. A. Post-feeding induction of trypsin in the midgut of Aedes aegypti L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochemistry.* **1991**, *21*, 197-203.

18. Rascón, A. A.; Gearin, J.; Isoe, J.; Miesfeld, R. L. In vitro activation and enzyme kinetic analysis of recombinant midgut serine proteases from the Dengue vector mosquito Aedes aegypti. *BMC biochemistry.* **2011**, *12*, 43.

19. Gooding, R. H. Studies on Proteinases from Some Blood-Sucking Insects. **1969**.

20. Jiang, Q.; Hall, M.; Noriega, F. G.; Wells, M. cDNA cloning and pattern of expression of an adult, female-specific chymotrypsin from Aedes aegypti midgut. *Insect Biochemistry and Molecular Biology.* **1997**, *27*, 283-289.

21. Graf, R. & Briegal, H. Comparison between aminopeptidase and trypsin activity in blood-fed females of *Aedes* aegypti. *Revue seuisse de zoologie.* **1982**, *89*, 845-850.

22. Isoe, J.; Zamora, J.; Miesfeld, R. L. Molecular analysis of the Aedes aegypti carboxypeptidase gene family. *Insect Biochem. Mol. Biol.* **2009**, *39*, 68-73.

23. Jiang, Q. Cloning and characterization of midgut-specific gene/gene products in the mosquito Aedes aegypti. **1996**.

24. Sorensen, H.; Mortensen, K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial Cell Factories.* **2005**, *4*.

25. Carl, M. H.; Frantzeskakis, L.; Huang, M.; Bekir Ülker PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Scientific Reports.* **2014**, *4*.

26. Singh, P.; Sharma, L.; Kulothungan, S. R.; Adkar, B. V.; Prajapati, R. S.; Ali, P. S.; Krishnan, B.; Varadarajan, R. Effect of Signal Peptide on Stability and Folding of Escherichia coli Thioredoxin.(Research Article). *PLoS ONE.* **2013**, *8*, e63442.

27. Barrett, A. J. Cathepsin G. *Meth. Enzymol.* **1981**, *80 Pt C*, 561.

28. Anthony, J. O.; A, A. E.; Giselle, M. K.; Ingram, J.; Zhou, M.; Jacob, B. S.; Alexander, L. G.; Daniel, R. H.; Qu, G.; David, A. M.; Marc, O. A.; Joseph, L. D.; James, H. M.; Alma, L. B.; Charles, S. C. Global identification of peptidase specificity by multiplex substrate profiling. *Nature Methods.* **2012**, *9*, 1095.

29. Gubler, U. Methods in Molecular Biology, Vol. 192, PCR Cloning Protocols, (2nd edition), B.Y. Chan and H.W. James. Human Press, Totowa, NJ, 2002 (Book Review). *Anal. Biochem.* **2003**, *312*, 262-262.

30. Rosano, G. L.; Ceccarelli, E. A. Recombinant protein expression in Escherichia coli : advances and challenges. *Frontiers in Microbiology.* **2014**, *5*.

31. Voet, D. & Voet, J. (2010) *Biochemistry*. Hoboken, NJ: John Wiley & Sons, Inc.

32. De Groot, N. S.; Ventura, S. Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Lett.* **2006**, *580*, 6471-6476.

33. Promdonkoy, B.; Promdonkoy, P.; Tanapongpipat, S.; Luxananil, P.; Chewawiwat, N.; Audtho, M.; Panyim, S. Cloning and Characterization of a Mosquito Larvicidal Toxin Produced During Vegetative Stage of Bacillus sphaericus 2297. *Curr. Microbiol.* **2004**, *49*, 84-88.

34. Bocian-Ostrzycka, K.; Grzeszczuk, M.; Dziewit, L.; Jagusztyn-Krynicka, E. Diversity of the Epsilonproteobacteria Dsb (disulfide bond) systems. *Frontiers in Microbiology.* **2015**, *6*.

35. Locker, J. K.; Griffiths, G. An Unconventional Role for Cytoplasmic Disulfide Bonds in Vaccinia Virus Proteins. *J. Cell Biol.* **1999**, *144*, 267-279.

36. Lobstein, J.; Emrich, C. A.; Jeans, C.; Faulkner, M.; Riggs, P.; Berkmen, M. SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories.* **2012**, *11*, 56-56.

37. Bessette, P. H.; Qiu, J.; Bardwell, J. C. A.; Swartz, J. R.; Georgiou, G. Effect of sequences of the active-site dipeptides of DsbA and DsbC on in vivo folding of multidisulfide proteins in Escherichia coli. *J. Bacteriol.* **2001**, *183*, 980.

38. Sharma, S. K.; Hopkins, T. R. Activation of bovine chymotrypsinogen A. Isolation and characterization of mu- and omega-chymotrypsin. *Biochemistry.* **1979**, *18*, 1008.

39. Miller, D. D.; Horbett, T. A.; Teller, D. C. Reevaluation of the activation of bovine chymotrypsinogen A. *Biochemistry.* **1971**, *10*, 4641.

40. Muller, H. M.; Catteruccia, F.; Vizioli, J.; Dellatorre, A.; Crisanti, A. Constitutive and Blood Meal-Induced Trypsin Genes in Anopheles gambiae. *Exp. Parasitol.* **1995**, *81*, 371-385.

41. Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. *Nat. Biotechnol.* **2000**, *18*, 187.

42. Hedstrom, L. Serine protease mechanism and specificity. *Chem. Rev.* **2002**, *102*, 4501.

43. Huber, R.; Bode, W. Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* **1978**, *11*, 114-122.

44. Dreyer, W. J.; Neurath, H. THE ACTIVATION OF CHYMOTRYPSINOGEN. *J. Am. Chem. Soc.* **1955**, *77*, 814-815.

45. Yamada, T. 2009. *Textbook of Gastroenterology.* West Sussex, United Kingdom: Wiley-Blackwell.

46. Kasafírek, E.; Fric, P.; Slabý, J.; Malis, F. p-Nitroanilides of 3-carboxypropionyl-peptides. Their cleavage by elastase, trypsin, and chymotrypsin. *European Journal of Biochemistry.* **1976**, *69*, 1.