Summer 2023

Optimizing Soluble Recombinant Expression of the Aedes Aegypti Early Trypsin Protease

Neomi Millan
San Jose State University

OPTIMIZING SOLUBLE RECOMBINANT EXPRESSION OF THE *AEDES AEGYPTI* EARLY TRYPSIN PROTEASE

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

Neomi Millan

August 2023

The Designated Thesis Committee Approves the Thesis Titled

OPTIMIZING SOLUBLE RECOMBINANT EXPRESSION OF THE *AEDES AEGYPTI* EARLY TRYPSIN PROTEASE

by

Neomi Millan

APPROVED FOR THE DEPARTMENT OF CHEMISTRY

SAN JOSÉ STATE UNIVERSITY

August 2023

Alberto A. Rascón, Jr., Ph.D.	Department of Chemistry
Ningkun Wang, Ph.D.	Department of Chemistry
Gianmarc Grazioli, Ph.D.	Department of Chemistry

ABSTRACT

OPTIMIZING SOLUBLE RECOMBINANT EXPRESSION OF THE *AEDES AEGYPTI* EARLY TRYPSIN PROTEASE

by Neomi Millan

The female Aedes aegypti mosquito is a primary vector for the transmission of blood–borne viral pathogens such as Chikungunya, Zika, Yellow Fever, and Dengue. The female mosquito will become infected with these pathogens through the uptake of a blood meal from an infected host. The midgut digestive enzymes are required to digest the blood meal to obtain necessary nutrients needed for the completion of the gonotrophic cycle. Blood meal digestion is biphasic and digestive enzymes are expressed at different time points after uptake of a blood meal. The focus of this study is on an early phase protease known as Aedes aegypti Early Trypsin (AaET). Initial attempts at obtaining soluble inactive zymogen led to autoactivation and low abundance of recombinantly expressed protease, making it difficult to isolate enough enzyme for activity assays and structural studies. Thus, an optimized protocol was developed to improve proper folding of the protease as well as to produce an abundance of the inactive zymogen form. The AaET gene was cloned into the pET28a expression vector and transformed into E. coli T7 SHuffle cells. Several growth conditions were tested, such as varying growth temperatures, concentration of the inducer (IPTG), and the presence or absence of a small molecule osmolyte (betaine). From these studies, bacterial growth in TB media at 10℃ with 0.025 mM IPTG and the addition of 1 mM betaine resulted in the best and most successful production of soluble recombinant AaET in the inactive form. Future research will focus on protein purification, enzyme kinetics, proteolytic specificity, and cleavage of blood meal protein substrates.

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Rascón for literally everything! Thank you for mentoring me throughout my undergrad and graduate years. I was lost, especially in my undergrad, and I didn't have anyone back home who had already gone through the system. I am very grateful for all the opportunities you have given me throughout the years. I especially want to thank you for always encouraging me whether I went to industry, PhD, or any other kind of schooling, I know you would always be happy for me. It's crazy how when we first met, I was just a quiet girl who was afraid to ask a lot of questions or talk to about my day but now I can talk to you about literally anything. Thank you for taking the time to make me feel safe enough to confide with you about anything and everything. Lastly, I want to thank you for helping me realize my own potential (as a person and a scientist) and for helping me break out of my shell. You really pushed and encouraged me when I was feeling doubtful about myself and my future, and for that I will be forever grateful. You are the best mentor I could've ever had and without you I would not have come this far!

I would also like to thank Flor for always supporting me throughout my journey. Thank you so much for any time that you sent Dr. R with some extra food, the extra food really helped me get through those long days in lab, especially when I didn't have a lot of food at home. I will miss your wonderful cooking and I am really going to miss picking on Dr. R together.

Next, I would like to thank my committee members, Dr. Wang and Dr. Grazioli. Thank you both for taking the time to listen to my presentations and for reading my thesis and giving any feedback or advice throughout the Master's program. Especially Dr. Grazioli,

thank you for your advice on PhD programs, applications, interviews, and how to navigate life during a PhD program. I will make sure to put your advice to good use in Madison!

A special shoutout to my most current lab mates Kenia, Liana, and Daisy! Being in the lab for about 5 years, I saw many people come and go but these girls really made lab fun and inviting these last two years. We really have some talented girls in the lab and I am so proud of all of them for pursuing what they want. I wish them all the best of luck in their future endeavors.

I would also like to thank my mom (Adriana), my dad (Carlos) and my brothers (Tony and Arrow). While there were many personal issues going on throughout my college career thus far, they have always encouraged me and have been proud of my accomplishments. A huge thanks to my dad for working his butt off to help pay for my education. Without all his hard work, I would not have been able to attend college comfortably. Thank you so much for getting me through school and allowing me to study something that I am passionate about. I will forever be grateful to you dad.

Next, I would like to thank my good friend and roommate Christina. Thank you for letting me use you as an audience to practice my presentations. I know we're both students and both very busy with assignment, but you were still able to make some time for me to practice. I also wanted to thank you for helping bounce ideas about sentence structure or certain words when I couldn't think of the right thing to say in this thesis, without you it would have taken me forever to get this all together. I am going to miss going on our random food craving runs and binge-watching dramas afterwards. I am lucky to have gotten you as a roommate all those years ago and I am grateful that I have gotten such a wonderful friend out

of it (friend ??). I'm going to miss you when I leave, but maybe you'll consider coming with me when you're ready for grad school.

Finally, I would like to thank my boyfriend Nathan and the Luc family. Thank you for inviting me to your holiday gatherings when I was alone up here. Thank you for all the food you sent over with Nathan, your food allowed me to try new things and kept me from being hungry when I ran out of food. And thank you, Nathan, for always being there for me. While we had our ups and downs, we always pulled through and were able to encourage and uplift one another. Thank you for all your love and support throughout my journey. Especially now, when we spend most of our night together watching Bones in the background while I'm writing this thesis. Thank you for your patience and understanding. Words can't describe how grateful I am to you for always being by my side and for your love and understanding. The future will always be uncertain, but I hope we continue to love and make one another happy in the future.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

1.1 The *Aedes aegypti* Mosquito and Life Cycle

The female *Aedes aegypti* mosquito is one of the fiercest creatures found around the world. She is a primary vector for the transmission of blood–borne viruses such as Zika, Chikungunya, Dengue, and Yellow Fever.[1] The *Ae. aegypti* mosquito is a tiny insect from the *Culicidae* family, deriving from the word culex, which is the Latin word for gnat.[2] The most prominent feature of the *Ae. aegypti* mosquito are the white stripes found along the abdomen and legs.[3] Another prominent feature of mosquitoes are their mouthparts. Both male and female *Ae. aegypti* mosquito's feed on sugar solutions such as nectar, however, the mouthparts of the female mosquito are modified to facilitate blood feeding.[4] This blood feeding behavior is essential for the female mosquito because the uptake of the blood meal will initiate the release and activation of midgut proteases that will break down blood meal proteins to provide the necessary nutrients needed to complete the gonotrophic cycle.[5] The gonotrophic cycle is the most important process in the mosquito's life cycle because at this stage the female mosquito is primed to lay eggs and produce future offspring. Once blood meal nutrients have been acquired to complete the gonotrophic cycle, she will lay her eggs near water, and eventually the eggs will hatch into larvae. At this stage, the mosquito larvae need a lot of energy to grow, being submerged in water will allow the larvae to eat algae, plankton, fungi, and other microorganisms found in the water.[6] Once the larvae have gained enough nutrients, they will develop into pupae. Once pupation occurs, they will continue to

develop beneath the water but will no longer feed. Finally, the mosquito will emerge from the pupal case and will look for a blood meal to restart the cycle.

1.2 *Aedes aegypti* as a Vector of Blood–Borne Pathogens

Unfortunately, the female mosquito can become infected with several viruses stemming from feeding from an infected host. When she takes in a blood–meal from an infected host, the virus will travel through her digestive system and end up in the salivary glands, and when she takes the next blood–meal from a different host, she will transmit the virus, now secreted in the saliva, and infect the new host.[7] This is concerning because the *Ae. aegypti* mosquito is a carrier of many potentially deadly viruses such as Zika, Dengue, Yellow Fever, and Chikungunya. Zika, Dengue, and Yellow Fever are part of the *Flaviviridae* family, a family of single–stranded, RNA viruses primarily found in arthropods such as ticks and mosquitoes.[8] Chikungunya is part of the *Togaviridae* family, a family of small, single–stranded RNA viruses that are found in mosquitoes similar to the *Flaviviridae* family.[9]

Infections with these viruses in mild cases can lead to mild flu–like symptoms, but in more severe cases, can cause long term effects or even death. For example, Chikungunya initially causes headaches and muscle and joint pains but in severe cases, the muscle and joint pains persist and my lead to chronic arthritis.[10] Similarly, initial symptoms for Yellow Fever causes flu like symptoms and muscle aches but symptoms can develop into serious illnesses such as jaundice, serious bleeding, organ failure and death.[11] As for Dengue, severe symptoms may cause a sudden drop in blood pressure, serious bleeding, and even death.[12] As for Zika, Zika viral infection was found to be linked to microcephaly and Guillain–Barré syndrome.[13] The Zika virus was given the name “the silent one” because symptoms usually

go unnoticed until a child is born. Microcephaly is a condition in which an infant's brain is reduced in size which leads to intellectual and motor disabilities, and behavioral issues.[14] Guillain–Barré syndrome is an autoimmune disease in which the immune system attacks the nervous system causing muscle weakness and even paralysis.[15] Unfortunately, only the Yellow Fever virus has an effective vaccine to combat infections whereas the other viruses have ineffective or unreliable treatments.[16,17] For example, to combat the Dengue virus, a vaccine called Dengvaxia was produced. Initially, safety data of Dengvaxia did not indicate any safety concerns in ages nine and above, but studies have shown that the vaccine causes hospitalization in those who have not yet contracted the virus due to the antibody dependent enhancement effect in which people became sensitized by the vaccine.[18,19]

The *Aedes aegypti* mosquito can be found all around the world but prefers to live in tropical, subtropical, and temperate climates.[20] In fact, recent studies have found that many countries such as the USA, UK, Australia, and Brazil have reported an increase in *Ae. aegypti* populations due to climate change.[14,20,21] In 2017 the CDC predicted how likely the *Ae. aegypti* mosquito will be able to live and reproduce in the United States (Figure 1).[20] While this map is just a prediction of where *Ae. aegypti* may be able to reproduce, there have been reports of Zika and Dengue infection found in Florida and Texas within the last seven years.[22] Controlling the *Aedes aegypti* mosquito as a vector is essential in preventing the spread of viral blood–borne pathogens; however, since most vector control strategies are either ineffective or insufficient, a new vector control strategy is needed.[16,17]

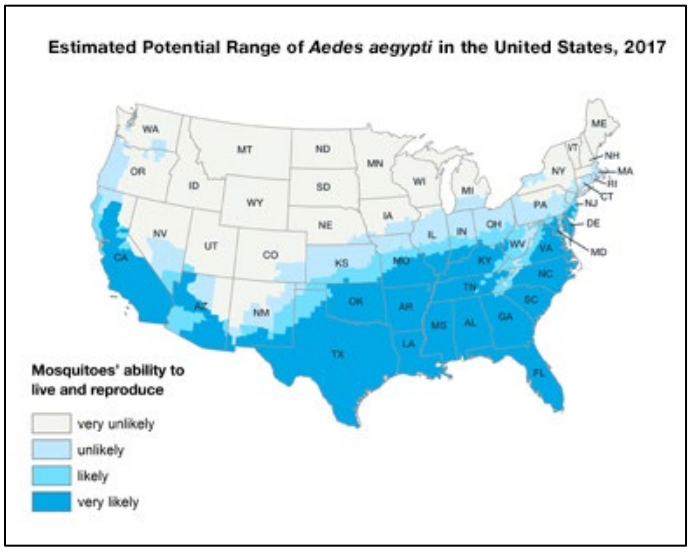


Figure 1. Prediction map of *Aedes aegypti* estimated population in the United States. The map represents an estimated population range of *Aedes aegypti* but does not represent spread of disease risk. Image obtained from the CDC.

1.3 *AEDES AEGYPTI* MIDGUT ENZYMES

Since modern medicine has unreliable and ineffective treatments for most of the viruses that *Ae. aegypti* mosquito carries, a new vector control strategy is needed to prevent the spread of mosquito–borne viruses. Vector control strategies such as insecticides, adulticides or eliminating larval habitats may cause harmful environmental and health effects.[23] For example, insecticides may be beneficial for eliminating mosquitoes, but insecticides can pollute the soil leading to pollution of ponds, lakes and drinking water and can become toxic to humans and livestock. Thus, a new vector control strategy will focus on the midgut digestive enzymes involved in producing nutrients for the egg laying process in the female *Aedes aegypti* mosquito. The idea stems from RNAi knockdown studies of mosquito digestive enzymes that led to a significant decrease in reproduction, as well as a direct correlation with the amount of nutrients available for the production of eggs (Figure 2).[5]

When a female mosquito is fully engorged with a blood meal she will produce large viable eggs, but when the blood meal is diluted the amount of nutrients decreases and thus causes a reduction in egg viability.[5] Based on these studies, it can be stated that the blood meal is essential for not only the gonotrophic cycle but also the mosquito's entire life cycle, and by inhibiting digestive enzymes, it would essentially inhibit the production of oligopeptides and amino acids needed for the mosquito to lay eggs. With the population of *Aedes aegypti* mosquitoes and number of viral infections increasing all over the world, it is important to gain an understanding of midgut enzymes and how they play a role in blood meal digestion, and possibly viral transmission.[24]

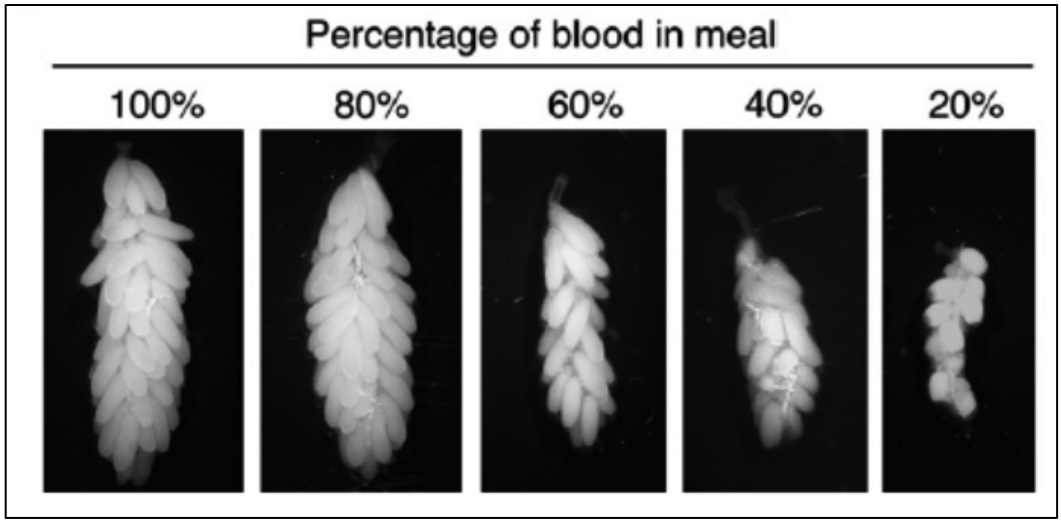


Figure 2. Representation of the relationship between the blood meal and mosquito fecundity. Dissected ovaries from the *Ae. aegypti* mosquito, fed with a diluted bloodmeal (100 mM $NaHCO_3$, 150 mM NaCl, pH 7.0) show that limiting nutrients in a blood meal affect overall egg production. Image obtained from Isoe *et al.* (2009).

Many digestive enzymes have been identified but only a few have been isolated and minimally studied. Two classes of proteolytic enzymes have been described, which are released and expressed at different points after a blood meal has been ingested.[5] The first

class of enzymes are endo–proteolytic which are known to be trypsin or chymotrypsin–like and cleave peptide bonds of non–terminal amino acids whereas the second class of enzymes are exo–proteolytic which are known to be aminopeptidases and carboxypeptidases and break peptide bonds of terminal amino acids.[25] Protease expression of midgut enzymes are found to be biphasic, being expressed in either the "Early phase" or the "Late phase".[5] The "Early phase" midgut enzymes are observed and expressed in the first 16–hours post blood meal. The *Aedes aegypti* midgut proteases that are expressed in the "Early phase" are Early Trypsin (AaET), Chymotrypsin (AaCHYMO), juvenile hormone–regulated serine protease (JHA15), and Serine Protease I (AaSPI).[5,26] The "Late phase" midgut enzymes are observed and expressed in the next 18–42 hours post blood meal. The *Aedes aegypti* midgut proteases that are expressed in the "Late phase" are Late Trypsin (AaLT), Serine Protease VI (AaSPVI), and Serine Protease VII (AaSPVII).[5,27] Interestingly, other midgut enzymes, Serine Protease II (AaSPII), Serine Protease IV (AaSPIV), and Serine Protease V (AaSPV) were found to be constantly expressed throughout both the Early and Late phases of blood meal digestion.[26] Although a few of these enzymes have been isolated, the role each enzyme plays during blood meal digestion is still unknown. Isoe *et al.* describes knockdown studies of a few of the isolated midgut digestive enzymes and reveal that AaSPVI, AaSPVII, and AaLT lead to a decrease in overall egg production.[5] Therefore, it is important to study and understand the role each enzyme plays during blood meal digestion, so we can better understand blood meal digestion and possibly inhibit the completion of the gonotrophic cycle or the transmission of viral pathogens. Furthermore, there might a correlation between midgut protease activity and viral replication that may be explored once the roles of these

midgut enzymes are fully understood. Thus, the focus of this study is of an early phase midgut enzyme known as *Aedes aegypti* Early Trypsin (AaET).

1.4 *AEDES AEGYPTI* EARLY TRYPSIN PROTEASE

1.4.1 Initial Recombinant Expression Studies

The *Aedes aegypti* Early Trypsin Protease is an early phase midgut digestive enzyme that aids in breaking down the blood meal in the first 16 hours after the blood meal has been ingested. Initial expression studies suggested that AaET may be required for gene expression in the late phase, however, later studies found that AaET is not needed for gene expression in the late phase.[5,28,29] The RNAi knockdown studies (Figure 3) also revealed that Early Trypsin led to a significant decrease in BApNA cleavage activity at 6–hours post blood meal in midgut extracts indicating that Early Trypsin is a major trypsin–like enzyme found in the early phase of blood meal digestion, but BApNA cleavage activity decrease in overall egg production compared to the late phase proteases.[5] Therefore, it is important to determine and understand what role Early Trypsin plays in the overall blood meal digestion process. To begin this endeavor, the enzyme must be recombinantly expressed and purified to be able to its activity *in vitro*.

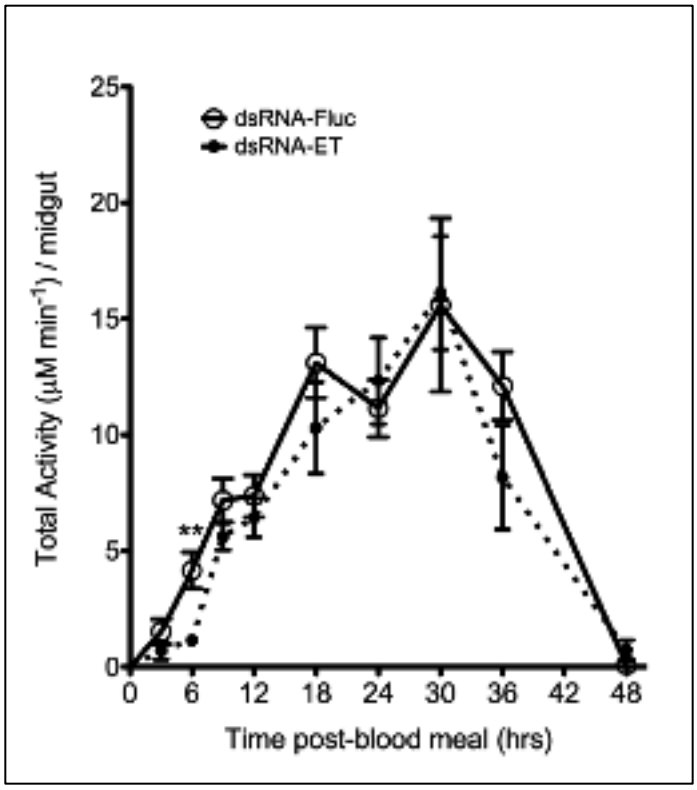


Figure 3. RNAi knockdown studies using BApNA to detect enzyme activity. Knockdown of AaET greatly decreases BApNA activity 6 hours post blood meal. Image obtained from Isoe *et al.* (2009).

1.4.2 AaET Autocatalysis

Multiple sequence alignment studies found that *Ae. aegypti* serine proteases depend on the catalytic triad composed of aspartic acid, serine, and histidine for proteolytic activity.[30] The catalytic triad at the active site, creates a charge relay system in which aspartate forms a hydrogen bond with histidine making histidine a powerful base to deprotonate the hydroxyl group of serine. Serine will then act as a nucleophile and attack the carbonyl of a peptide substrate. A hydrolysis reaction will occur to then liberate the cleaved substrate, while aspartate stabilizes the newly protonated histidine. As the name implies, AaET is a trypsin–like serine protease. In the *Ae.* mosquito, Early Trypsin is expressed in the inactive from,

known as a zymogen or trypsinogen. The zymogen of the protease, will prevent the cleavage of peptide bonds of any protein substrate, but once the propeptide region is cleaved from the zymogen, it will undergo a conformational change that leads to the active mature enzyme which facilitates substrate binding and increased proteolytic activity.[31]

Recombinant Early Trypsin has been shown to be autocatalytic and will catalyze the inactive trypsinogen to the active mature form.[27] Due to Early Trypsin's autocatalytic nature, there were difficulties when trying to express and isolate the protease. Initial attempts at recombinantly expressing and purifying AaET led to difficulties in producing soluble enzyme. To solve this issue, a refolding scheme using a denaturation/renaturation protocol of AaET expressed using BL21–DE3 cells with guanidine–hydrochloride (Gu–HCl) was used. The Gu–HCl chaotropic buffer fully denatured the protein followed by dialysis to slowly to remove the Gu–HCl to refold the protein. As the protein slowly folded, autoactivation of the protease was observed and a low abundance of protein was isolated.[27] Later, Nguyen *et al*. re–coned the gene into the pET28a vector and expressed N–terminally his_6–tagged AaET using *E. coli* T7 SHuffle competent cells, which are cells that are engineered to correctly produce proteins that contain disulfide bonds with the help of a chaperone that helps mis–oxidized proteins to fold properly.[27,32] Unfortunately, while they were able to improve and visualize soluble expression at low temperatures, expression in these cells still led to autocatalysis and a low abundance of the protease, making it difficult to isolate, purify and further analyze the protease.[33]

1.5 Optimizing Soluble Expression of Inactive AaET

To obtain an abundant amount of the inactive AaET protease, the recombinant expression protocol described in Nguyen *et al*. was optimized by manipulating culture conditions. Modifying factors such as temperature, inducer concentrations, and culture additives can influence the rate of synthesis and may aid in proper folding of the protein.[34]

For this study, Early Trypsin will be induced with varying concentrations of the inducer isopropyl β–d–1–thiogalactopyranoside (IPTG), going lower than the recommended 0.1 mM.[33] IPTG is a synthetic inducer that will not get metabolized and using low concentrations of the inducer should allow slow production of the protease and prevent autocatalysis. This is important because protease autoactivation is dependent on the concentration of the produced enzyme.[33,35] The higher the accumulation of the protease in the cell, the faster the protease can autoactivate. In addition, a small molecule osmolyte called betaine will be incorporated in the culture. This small molecule osmolyte is believed to assist in protein folding by interacting with the protein backbone, resulting in a decrease in the accumulation of aggregation.[36,37] Thus, possibly increasing recombinant soluble protein. Each manipulation factor will be tested at different temperatures ranging from 30°C to 10°C. The reason for this temperature difference is due to the fact that Nguyen *et al.* (2018) were able to determine that the production of soluble midgut proteases in T7 SHuffle cells is temperature dependent, and depending on the protease, temperature may be a contributing factor in recombinant soluble expression of the inactive protease.

Chapter 2

Recombinant Expression and Data Analysis Methods

2.1 Background: Initial Recombinant Expression Studies

Previous studies found that AaET is autocatalytic which ultimately led to difficulties in isolating and purifying an abundant amount of the protease. As a graduate student, Dr. Alberto A. Rascón Jr. produced AaET in bacterial cells, obtained activity data, and implemented a denaturation/renaturation protocol to obtain recombinant soluble AaET for initial studies.[27] The protocol described in Rascón *et al.* utilized BL21(DE3) cells with guanidine–hydrochloride to denature the protein followed by refolding using a dialysis buffer to remove the Gu–HCl chaotropic reagent.[27] In figure 1, the final isolated AaET protease is observed, with 10 micrograms (μg) of AaET loaded onto a gel to visualize the protease, also showing that AaET had autocatalyzed into the active mature form (Figure 4).[27] Unfortunately, this protocol still led to inadequate amounts of the protein for further activity assays and structural studies.

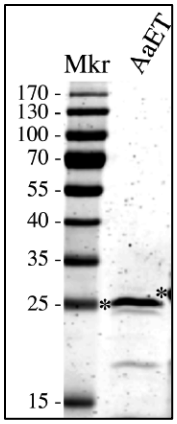


Figure 4. SDS–PAGE results of purified active mature Early Trypsin protease. Insoluble AaET was refolded using a denaturation/renaturation protocol to isolate the enzyme from inclusion bodies which led to autoactivation and low yields. The AaET protease was loaded with 10 μg of protein. Image obtained from Rascón *et al.* (2011).[27]

In 2018, Nguyen *et al.* then expressed Early Trypsin using SHuffle T7 *E. coli* competent cells instead of BL21(DE3) cells.[33] Previous studies have found that the BL21(DE3) bacterial strain leads to inclusion bodies of enzymes dependent on disulfide bond formation for folding and structural stability due to the highly reducing environment of the strain.[33,38] The SHuffle T7 *E. coli* competent cells, however, have a more oxidizing cytoplasm that helps support the formation of disulfide bonds.[32] As it turns out, AaET is dependent on three disulfide bonds for structural stability and activity. In addition to the more oxidizing cytoplasm, Lobstein *et al*. also engineered the production of the DsbC chaperone in the cytoplasm, which helps mis–oxidized proteins to fold properly.[32] This combination helped Nguyen *et al.* produce inactive (zymogen) AaET, which was observed via SDS–PAGE analysis near the 30 kDa molecular weight (MW) ladder mark, and unfortunately, recombinant soluble expression was also unsuccessful at 30°C (Figure 5). The recombinant expression protocol was repeated at 15°C and recombinant soluble expression was achieved

(Figure 6A). However, at 5–hours post induction the expression band at 30 kDa begins to disappear and a more pronounced band appears right below the 25 kDa marker, indicating the possibility of early trypsin autoactivating into the active mature form (Figure 6A).[33] The active mature form of Early Trypsin was confirmed by testing activity of the protease with the use of BApNA (benzoyl arginine p-nitroaniline) as a substrate.[33] This substrate is used to detect trypsin–like activity due to the preferred arginine amino acid at the pseudo peptide bond cleavage site between arginine and p–nitroaniline. Cleavage at this site releases the p-nitroaniline chromophore, leading to a color change from colorless to yellow.[39,40] At 5–hours post induction (as shown in Figure 6B), BApNA activity increases which correlates with the appearance of the active mature band seen in Figure 3A. While Nguyen *et al.* was able to achieve recombinant soluble expression, they were not able to maintain the inactive form and could not isolate an abundant amount of the protease for further *in vitro* activity assays, thus making further studies harder to achieve.

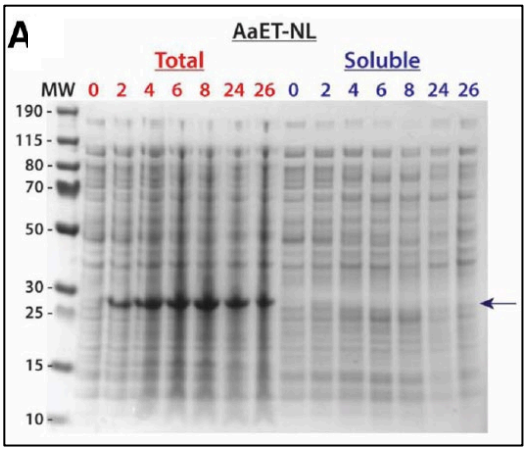


Figure 5. SDS–PAGE results of initial attempt at AaET recombinant soluble expression in T7 SHuffle *E. coli* competent cells. The cells were grown in TB media at 30°C and induced with 0.1 mM IPTG. Numbered bands represent the protein ladder associated with the molecular weight in kDa (MW lane). Expected soluble protease is indicated by the arrow, unfortunately, the protease was expressed insolubly and is only observed in the total samples. Image obtained from Nguyen *et al*. (2018).[33]

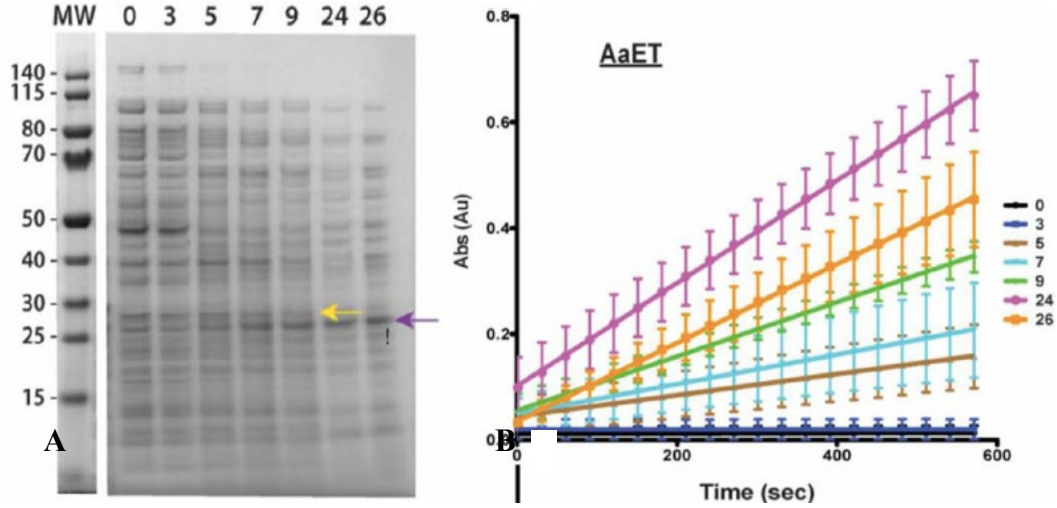


Figure 6. SDS–PAGE analysis and BApNA activity assays of samples collected from a small-scale growth experiment producing soluble AaET protease. A. The gel represents soluble expression of inactive (zymogen) AaET, indicated by the yellow arrow, and the active mature form, indicated by the purple arrow. B. The presence of active mature AaET starting at 5 hours post induction correlates with an increase in BApNA activity. Image obtained from Nguyen *et al*. (2018).[33]

Early Trypsin had previously been cloned using the pET vector system.[33] The pET vector system was developed for the cloning and expression of recombinant proteins in *E. coli*.[41] The pET28a vector was used to add an N–terminal his_6–tag to the protease to aid in one step purification using a Nickel column.[33,42] The his_6–tag on AaET will chelate to the Ni^{2+} metal on the column and will allow other non–his tagged proteins to elute out. However, the issue is that Early Trypsin is autocatalytic and will cleave off the N–terminal his_6–tag which makes it harder to purify and isolate the protease. Thus, to be able to purify and isolate the AaET protease in abundance using a Nickel column, it is important to produce recombinant soluble expression of the inactive form of AaET, the zymogen, to maintain the his_6-tag.

2.2 Recombinant Bacterial Expression

To recombinantly produce AaET, a bacterial system was used because of the feasibility and low cost compared to other expression systems.[27,33,43] A typical bacterial growth is composed of a lag phase, log phase, stationary phase, and a logarithmic decline phase, with the lag and log phase being the most important phases needed for the successful production of soluble proteins.[44] First, bacterial cells will adapt to the growth conditions (media and antibiotic added), the lag phase, and once the cells are acclimated, will then continue to grow and double until nutrients become depleted (the log phase). In the stationary phase, the rate of growth and the rate of death are equal, linearizing until the cells begin to die due to the lack of nutrients available for continued growth. Due to the abundance of nutrients still available, while the cells are doubling and growing, the log phase is where induction of protein expression is most ideal.[45] Therefore, all of the growth experiments conducted in this

project were induced at an optical density at 600 nm (OD_{600nm} ~ 0.6), the middle portion of the log phase, using SHuffle T7 cells.

As previously mentioned, SHuffle T7 *E. coli* competent cells have a more oxidizing environment than wildtype *E. coli* cells. Lobstein *et al*. engineered the SHuffle expression strain to correctly produce proteins that contain disulfide bonds.[32] Previous studies found that Early Trypsin contains three disulfide bonds, making the SHuffle expression strain useful during recombinant bacterial expression.[27,33] Furthermore, Lobstein *et al*. also ensured that the SHuffle expression strain was engineered with a DsbC chaperone which helps mis–oxidized proteins to fold properly.[32] Additionally, Nguyen *et al*. identified that terrific broth (TB) media is the best nutritious broth needed to provide the SHuffle cells with necessary nutrients for better overall expression.[33] Unlike other media, such as Lysogeny Broth (LB), TB media is a buffered, nutritious media that will lead to denser cell growths due to increased amounts of tryptone and yeast extract, as well as the addition of glycerol.[46] The extra amounts of tryptone and yeast extract will act as the main sources of carbon, nitrogen, vitamins, and amino acids needed for bacterial growth.[46,47] The buffering system comes from the added potassium phosphate buffer which provides potassium to help prevent cell death and phosphate to maintain a neutral pH.[46,47] The addition of glycerol acts as an extra source of carbon, which is an extra feeding source for the bacterial cells.[46,47]

Moreover, the synthetic inducer (isopropyl β–d–1–thiogalactopyranoside, IPTG), which is a mimic of allolactose (the natural transcription activator), triggers transcription without being metabolized.[48] However, it is important to note that because IPTG does not get metabolized, high concentrations of the inducer will increase the rate of transcription while

decreasing the rate of RNA synthesis, and at high concentrations, may be toxic to the cells leading to insoluble expression of the enzyme and cell death.[49] Similarly, if low concentrations of the inducer are utilized then there might not be enough inducer to initiate transcription within the cells.[50] Previous studies have concluded that the optimal concentration of IPTG can range between 0.05 and 0.1 mM.[50] For this project, 0.025, 0.0125, and 0.00625 mM concentrations of IPTG were used to determine which concentration will help improve soluble expression. By manipulating the concentration of IPTG to as low as 0.00625 mM, should slow the rate of transcription, producing fewer AaET protease molecules to not only prevent aggregation, but allow more time for the protease to fold properly.

Additionally, a small molecule osmolyte was incorporated into the protocol. An osmolyte can be described as a chemical chaperone that may help increase thermodynamic stability of folded proteins without stressing other cellular processes and maintain osmotic pressure.[37] Because betaine is a polar molecule, betaine can stabilize proteins by interacting with the protein backbone and essentially protect the hydrophobic regions from interacting with other hydrophobic regions of other simultaneously expressed proteins.[37] Thus, the small molecule osmolyte may be able to aid in protein folding, resulting in increased protein solubility.[51,52] However, it is possible that high salt concentrations in the media in the presence of an osmolyte may cause aggregation to occur and decrease protein solubility, so proper care is needed when choosing the correct concentration of osmolyte used.[53] Nonetheless, betaine was chosen as the osmolyte for these bacterial growth experiment studies at the recommended 1 mM concentration.[37]

In bacteria, transcription and translation happen simultaneously which leads proteins to fold too fast, sometimes not allowing for foreign recombinant protein to fold properly.[54] Inclusion bodies form as a result of intracellular accumulation of partially folded proteins, resulting in aggregation.[55] Inclusion bodies can be formed in both high and low temperatures but can be solubilized faster at lower temperatures.[56] Therefore, Lobstein *et al.* used a variety of temperatures to optimize protein expression but found that the effect of temperature is protein specific, with a decrease in temperature usually resulting in improved yields of solubly expressed protein.[32] Five out of the seven proteins that were tested in their study resulted in improved yields with temperatures of 25°C and below.[32] Decreasing temperature may decrease the rate of synthesis and metabolism, allowing more time for the protein to fold properly without aggregating.[57] Thus, for this study 30°C, 23°C, and 10°C temperatures were analyzed to determine which temperature will produce an abundance of soluble inactive AaET protein without autocatalysis occurring.

All bacterial growths were started at 30°C to approach the log phase within four hours, but once the cells reached the optical density at 600 nm (OD_{600nm}) at approximately 0.6, the temperatures were dropped, and culture additives were added (inducer and osmolyte). These conditions were modified depending on the experiment. For example, a growth, set at 15°C induced with 0.025 mM IPTG in the presence of 1 mM betaine, would start at 30°C then upon reaching OD_{600nm} ~0.6, the cells would be treated with IPTG and betaine then the growth temperature would be dropped to 15°C.

2.3 SDS–PAGE GEL ANALYSIS

For all growth experiments, various cell samples were collected over a 55–hour period and each sample was diluted with 450 μL of 20 mM Tris-HCl pH 7.2 buffer and cells were then sonicated to disrupt and break open the cells. Sonication uses high frequency sound waves that will agitate and lyse the cells.[58] Cells were sonicated using the Qsonica sonicator Q125 for 10 seconds with 25% amplification, all while being kept on ice to prevent any enzymatic activity from occurring, samples were then treated with a sodium dodecyl sulfate (SDS) dye. SDS is a detergent that will partially denature and bind to the protein backbone, proportional to the molecular weight of the protein, which contributes to a large negative charge.[59,60] Thus, electrophoresis in the presence of SDS will separate proteins based on molecular weight. Additionally, a reducing agent such as dithiothreitol is used to reduce disulfide bonds that are critical for proper protein folding.[59,60] The SDS dye was prepared by combining SDS (purchased from Pierce Chemical Company), glycerol, bromophenol blue, 1M Tris-HCl pH 7.2, and dithiothreitol to make the 6x SDS dye stock. However, only 1x SDS dye is needed to partially denature the proteins in each sample. To obtain the total samples, 20 μL of the freshly sonicated samples were incubated with the SDS dye on a 90°C dry heat block. Sonicated samples were then centrifuged (at 13,000 rcf and 4°C) to separate insoluble protein from the soluble protein. Soluble samples were then obtained by taking 20 μL of the supernatant and incubating samples with the SDS dye on a 90°C dry heat block. Samples were then loaded onto a pre-casted NuPAGE 4-12% Bis-Tris gel (Invitrogen). PAGE represents polyacrylamide gel electrophoresis, and the polyacrylamide forms a mesh–like matrix and will act as a molecular sieve that will slow the migration of proteins in

proportion to their charge to mass ratio.[60,61] The polyacrylamide gel was loaded onto the XCELL II running chamber with 1x MES buffer at 180 volts. The MES buffer is formulated for running PAGE Bis-Tris gels and is recommended for separating small and medium sized proteins.[62] To visualize proteins, the gel needs to be incubated in a dye or stain. The InVision his–tag stain is specific for his–tagged proteins because it contains a proprietary fluorescent dye that is conjugated to the Nickel (II) ion (Ni^{2+}).[63] The Ni^{2+} will chelate to the poly–histidine domain, which will allow for specific binding of his–tagged proteins. Additionally, after his-tag staining, the gels were treated with SimplyBlue Safe Stain (Invitrogen). This Safe Stain is an eco–friendly stain that is formulated for fast, sensitive, and specific detection of all proteins run via SDS-PAGE.[64]

After running the growth experiment samples through gel electrophoresis, the polyacrylamide gels were washed with a fixing solution (10% glacial acetic acid, 50% ethanol) to remove excess MES buffer and prevent dispersion of proteins to ensure protein bands remain sharp and are correctly resolved in the downstream staining steps.[65] After microwaving for 45 sec at full power and 10 min incubation in fixing solution, the gels were then rinsed with high purity Milli Q water (5 min water wash) for a total of two times. After the final water wash, the gels were incubated with 30 mL of InVision his-stain for 40 min and was then washed (4x 5 min) with 100 mL of 20 mM Sodium Phosphate buffer pH 7.8 to remove background noise. Since the InVision his-tag stain has a fluorescent tag, the gel needs to be excited using a UV transilluminator to be fully visualized. All gels were imaged using the GelDoc Go imaging system (BioRad), with his-stained gels imaged using the ethidium bromide setting under the nucleic acid gels application category. Using this setting

led to gels with a dark background and white detected protein bands. After the gels were imaged, the sodium phosphate buffer was discarded, and the gel was incubated in about 100 mL of the SimplyBlue SafeStain overnight. The blue-stained gels were then washed with high purity Milli Q water for 5 to 6 hours, then imaged using the Coomassie blue setting under the protein gels application category on the GelDoc Go imaging system leading to clear background and dark blue bands.

2.4 Gel Band Quantification and Analysis

To determine statistical differences between protein expression in Total and Soluble bands of all growth experiments, with the different temperature conditions, concentrations of IPTG, and samples containing the betaine osmolyte, gel bands were quantified. The ImageJ program was used to measure gel band intensity. ImageJ can calculate the area and pixel value statistics of a given selection, this case being the expression of AaET.[66] For the analyses, his–stained gels were utilized due to the stain only binding to the AaET his_6-tagged proteins, minimizing the background binding of other non-tagged proteins, which may interfere with the ImageJ gel analysis. Simply Blue stain binds to all proteins, and if any other proteins overlap with the molecular weight of AaET, then pixel area count will be incorrectly determined. In using the his-stained gels, the images were inverted to produce a white background and dark bands so that the program can analyze the number of pixels each band exhibits (Figure 7A). This is important because ImageJ only works for dark black bands in a white background.[66] Each gel figure was set to 32-bit to remove any RGB coloring and ensure that the figure could be inverted. Using the rectangle tool, the largest expressed band was measured first and a box was created and shaped using this band, then moved over to the

time zero band (Figure 7). Using the select first lane tool, under the analyze gels tab, allows for the program to compare the other selections to the first lane and remove any background pixels. By selecting the select next lane tool, all other bands are selected and can be analyzed for pixels when compared to the time zero background (Figure 7A). It should be noted that ImageJ requires each box within a selection to be the same size in order to perform pixel detection properly. Once all bands are boxed, ImageJ uses a plot lanes tool which analyzes the pixels in each box and plots the intensity of each band against the background band (Figure 7B). The line drawing tool was used to then draw a line from one base of the peak to another, defining each peak so that ImageJ can automatically produce the area of the peak, once the inside of each peak has been clicked. The area of each peak was used to obtain a number measurement for the intensity of each band to then be able to run statistical analysis tests. Furthermore, it was assumed that time zero produced no recombinant expression (pre-induction sample), thus, each time zero (background) was subtracted from all other bands to eliminate any background noise.

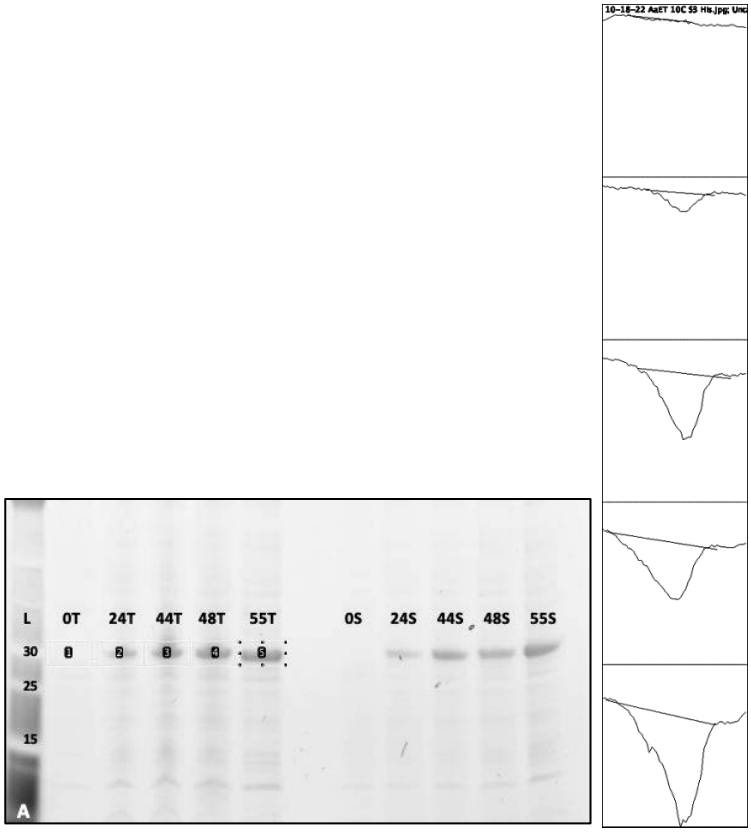


Figure 7. Gel band quantification using the ImageJ program. A. The gel presented is soluble expression of AaET using TB media and grown at 10°C with 0.0125 mM IPTG and stained with his-stain (more gel information detail is presented in Figure 22B). The figures represent the method used to measure band intensity of the gel. The yellow boxes around the gel bands represent the area taken to measure pixels and create the plots shown in B. B. These are the plots that were generated by ImageJ to help measure the intensity (pixels) of each band. The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences.

Once numerical data was obtained from ImageJ, each band measurement was used for either a T-test or an ANOVA test. An unpaired T-test is a statistical test that compares the means of two unrelated or independent samples.[67] The T-test uses the t-statistic, the t-distribution values, and the degrees of freedom to determine statistical significance and the

null hypothesis is rejected or accepted accordingly.[68] If the null hypothesis is rejected, it indicates that the data readings are strong and are probably not due to chance.[68] The unpaired T-test assumes a gaussian distribution and equal variance and standard deviation between groups, however, in order to obtain the best statistical results, a larger sample size is required.[67,68] The unpaired T-test was used to compare the means of the total and soluble samples. ANOVA, commonly referred to as an F test, is a statistical test that compares the means of three or more unmatched groups.[67] The F test is used to compare the variance in each group mean from the overall group variance.[67,69] If the variance within groups is smaller than the variance between the groups, it will result in a higher likelihood that the difference observed is not due to chance.[69] One-way ANOVA uses one independent variable, in this case the variable would be the different concentrations of the inducer. Each statistical test was performed using GraphPad Prism. GraphPad Prism can use multiple statistical tests on a given data set. To compare the means of the total and soluble samples for each condition at 10°C, data obtained from ImageJ was loaded into the XY presets with the X-value identified as the time-point samples that were collected (0-, 24-, 44-, and 48-hours) and the Y-value set to have three replicate values (the pixel intensities for each sample determined using ImageJ) in side-by-side sub-columns. Total and soluble samples were separated by groups, total being group A and soluble being group B. Once all data is loaded onto the GraphPad sheet, the analyze button allows us to choose which statistical test to use to analyze the data. When comparing the means of total and soluble samples, the T-tests (and nonparametric tests) settings under column analyses were selected. A pop-up screen will then appear to select specific parameters for the T-test. Selecting the unpaired test under experimental design and

ensuring two-tailed with 95% confidence level under options are the only parameters that were changed, all other settings were left as is. By selecting OK, GraphPad Prism automatically generates the results from the statistical analysis test as well as a graph that shows the comparisons from the data set. Under the change settings, the grouped graph family and the first graph representation under summary data was selected to ensure graphs can be visualized easily. A similar approach was taken when comparing the different IPTG concentrations and samples containing betaine using ANOVA. Only the soluble expression data obtained from ImageJ was loaded into the XY presets with the X-value identified as the time-point samples that were collected (0-, 24-, 44-, and 48-hours) and the Y-value set to have three replicate values (the pixel intensities for each sample determined using ImageJ) in side-by-side sub-columns. In one sheet, the different IPTG concentrations were separated by groups with 0.00625 mM being group A, 0.0125 mM being group B, and 0.025 mM being group C. In another sheet, samples induced with IPTG and 1 mM betaine were separated by groups described in the previous statement. Once all data is loaded onto the sheet, the One-way ANOVA (and nonparametric or mixed) settings under column analyses were selected. A pop-up screen will then appear to select specific parameters for the ANOVA test. Selecting 'compare the mean of each column with the mean of every other column' under multiple comparisons is the only parameter that was changed, all other settings were left as is. By selecting OK, GraphPad Prism automatically generates the results from the statistical analysis test as well as a graph that shows the comparisons from the data set. Under the change settings, the grouped graph family and the first graph representation under summary data was selected to ensure graphs can be visualized easily.

Chapter 3

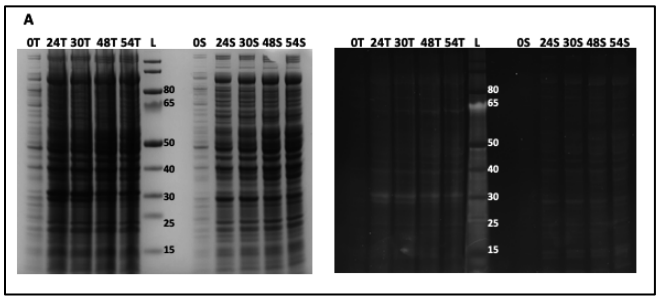
Results

3.1 Recombinant Bacterial Expression and SDS-PAGE Analysis

The AaET/pET28a plasmid was transformed into T7 SHuffle cells to form colonies that were used to cultivate large quantities of cells that contain plasmid DNA. Growth experiments were set in Terrific Broth (TB) nutritious media and incubated at 30°C to ensure the lag and log phases are reached, which were then induced with either 0.025 mM, 0.0125 mM, or 0.00625 mM IPTG and the addition or absence of 0.1 mM betaine when cells reached an optical density (OD_{600nm}) of about 0.6. After induction, temperature settings varied based on the experiment and cells were incubated at either 30°C, 23°C, or 10°C. Various samples were collected over an estimated 68-hour period and all samples were then diluted with 450 µL of 20 mM Tris-HCl pH 7.2 buffer in preparation for sonication. All samples were sonicated for 10 sec at 25% amplitude (Qsonica sonicator Q125) with 25 sec ice resting for a total of three times. After sonication, 20 µL samples were taken and treated with 6x SDS dye (final concentration of 1x), then the sonicated samples were centrifuged to remove the insoluble material and protein from properly folded protein, as described in Chapter 2 section 2. Once the samples were heated, they were loaded on to SDS-PAGE gels for analysis. All gels were initially stained with the InVision his–tag stain and then stained with SimplyBlue Safe Stain.

3.1.1 30°C Growth Experiments

At 30°C, samples induced with 0.025 mM IPTG only produced insoluble protein (Figure 8). In fact, no soluble expression was observed at each condition tested at 30°C (Figures 8-13), a result not too surprising since the rate of bacterial protein expression is too fast to allow disulfide bonds in AaET to fold properly.[32,33,54] In some of the his-stained gels a dim soluble band right below the 30 kDa marker can be observed (Figures 8A, 9A, 10A, 11A, 12A,13A). However, the soluble bands do not align with the insoluble protein bands and the soluble bands are not as bright, indicating that the soluble bands observed are not the AaET protein of interest. Unfortunately, multiple blue stained and his-stained gels have been ripped since his-stained gels are more sensitive to handle, but results were still observed in the his-stained gels and blue stained gels show similarities to the his-stained gels (Figures 8B, 9C, 12A, 13B, 13C).



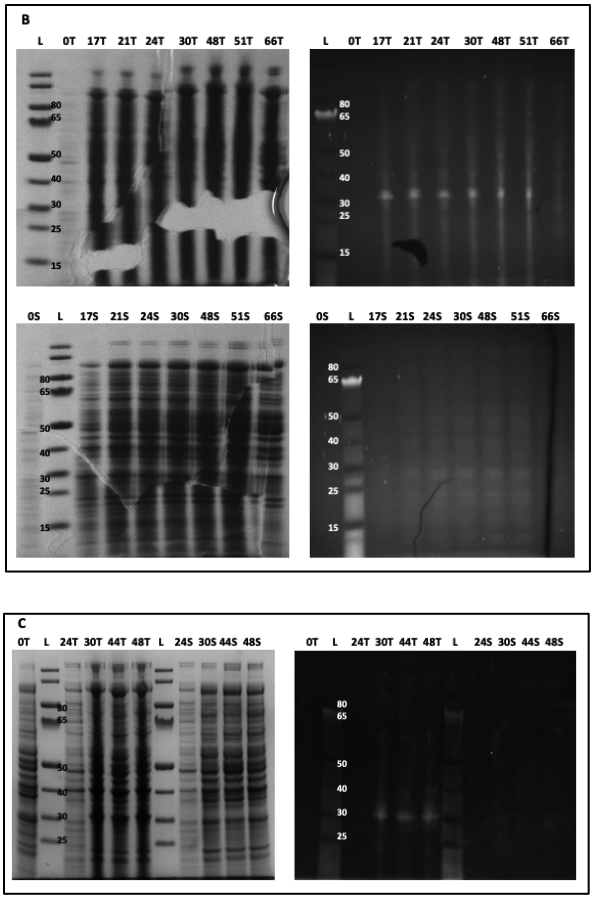
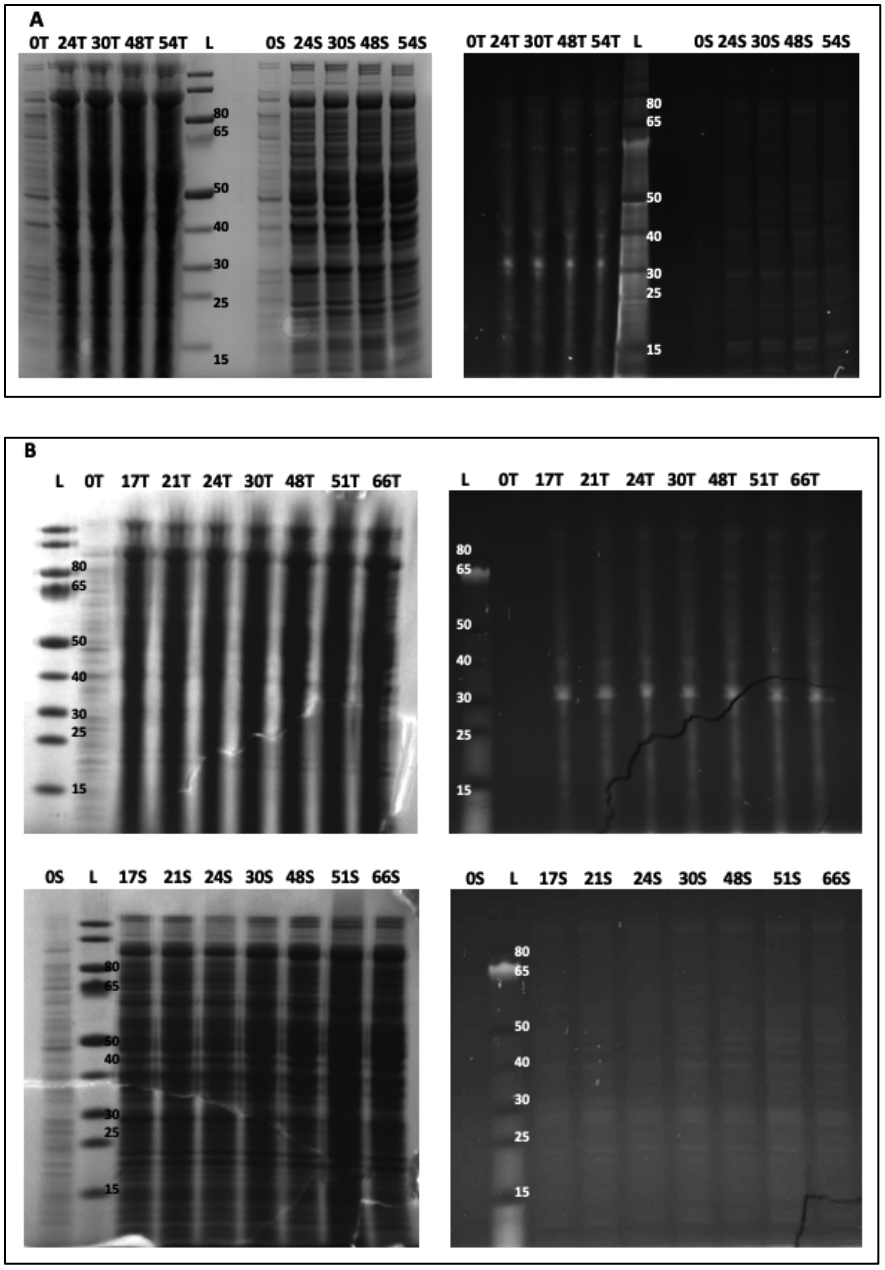


Figure 8. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 30°C in TB media and induced with 0.025 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 66 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gels have been ripped but results were still observed and show similarities to the his-stained gel. C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
0T 24T 30T 48T 54T L 0S 24S 30S 48S 54S
80
65
50
40
30
25
15
0T 24T 30T 48T 54T L 0S 24S 30S 48S 54S
80
65
50
40
30
25
15
B
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15

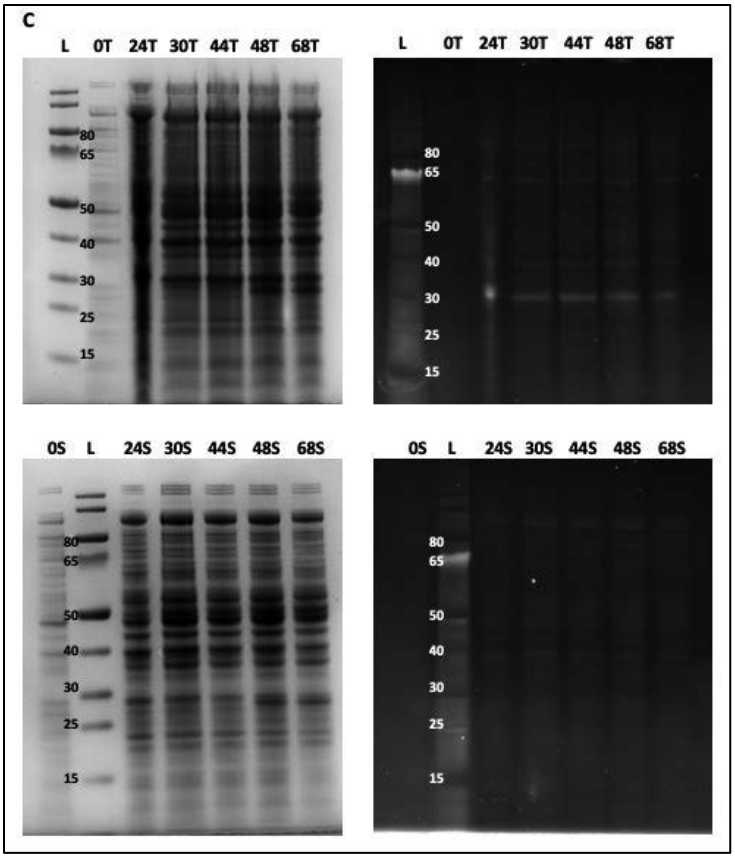
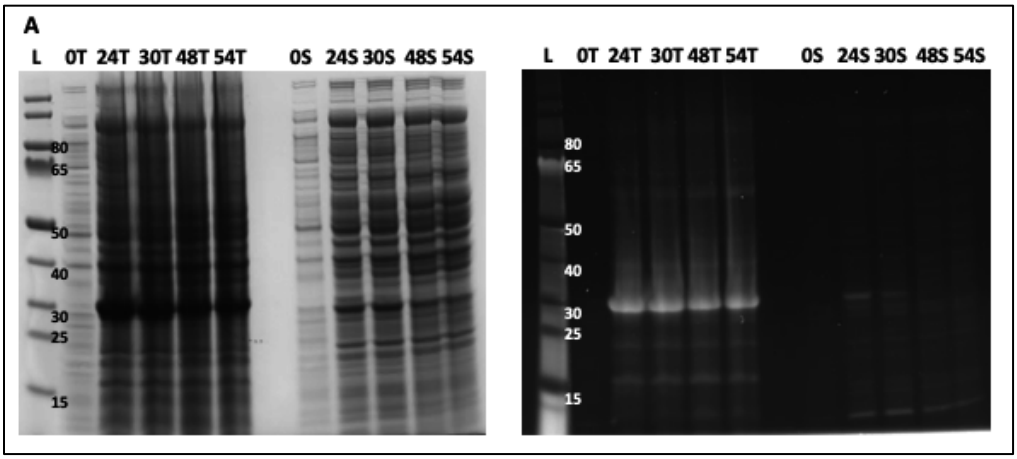
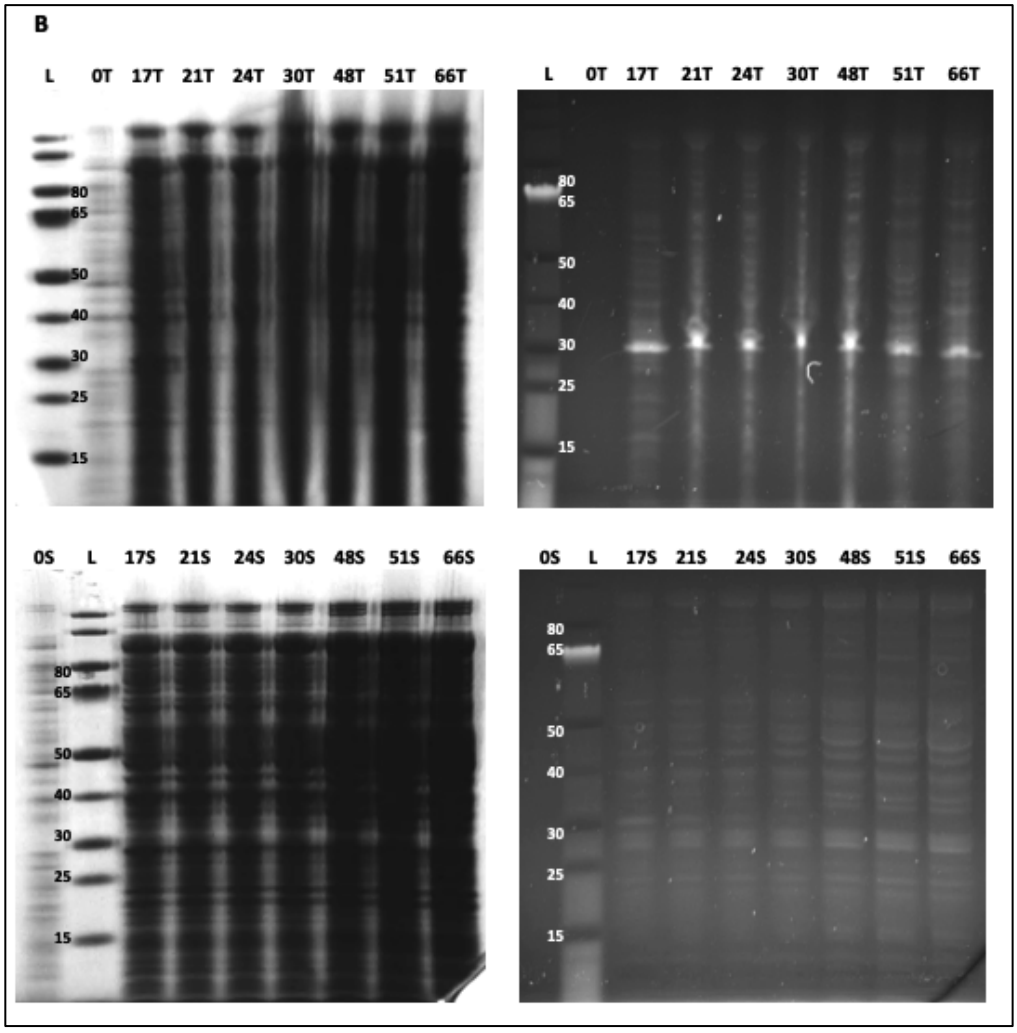


Figure 9. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 30°C in TB media and induced with 0.025 mM IPTG, and in the presence of 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 68 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). Gels have been ripped but results were still observed and show similarities to one another. C. Third triplicate set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right).


A
L 0T 24T 30T 48T 54T 0S 24S 30S 48S 54S
80
65
50
40
30
25
15
L 0T 24T 30T 48T 54T 0S 24S 30S 48S 54S
80
65
50
40
30
25
15


B
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15

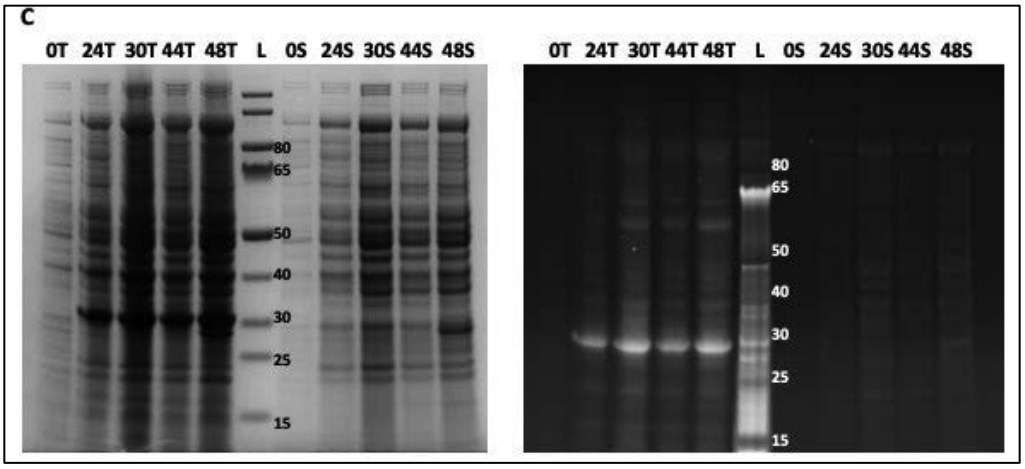
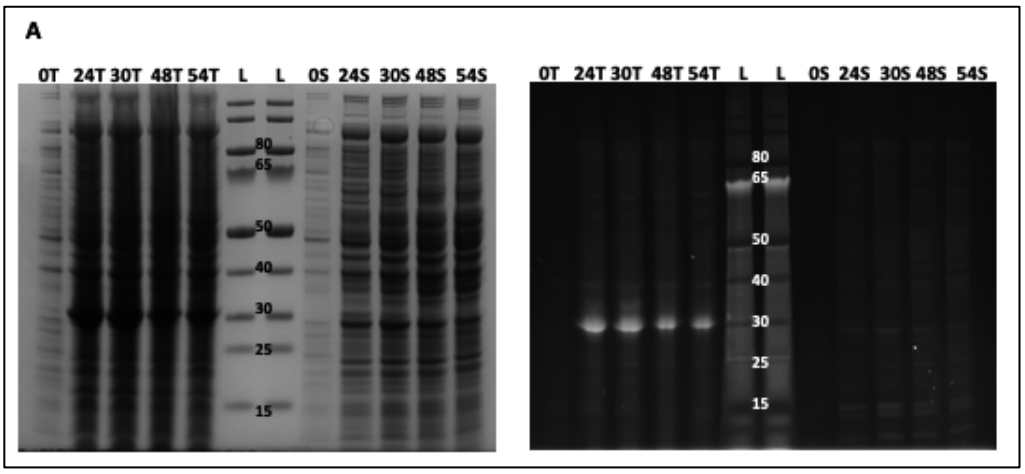


Figure 10. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 30°C in TB media and induced with 0.0125 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 66 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).



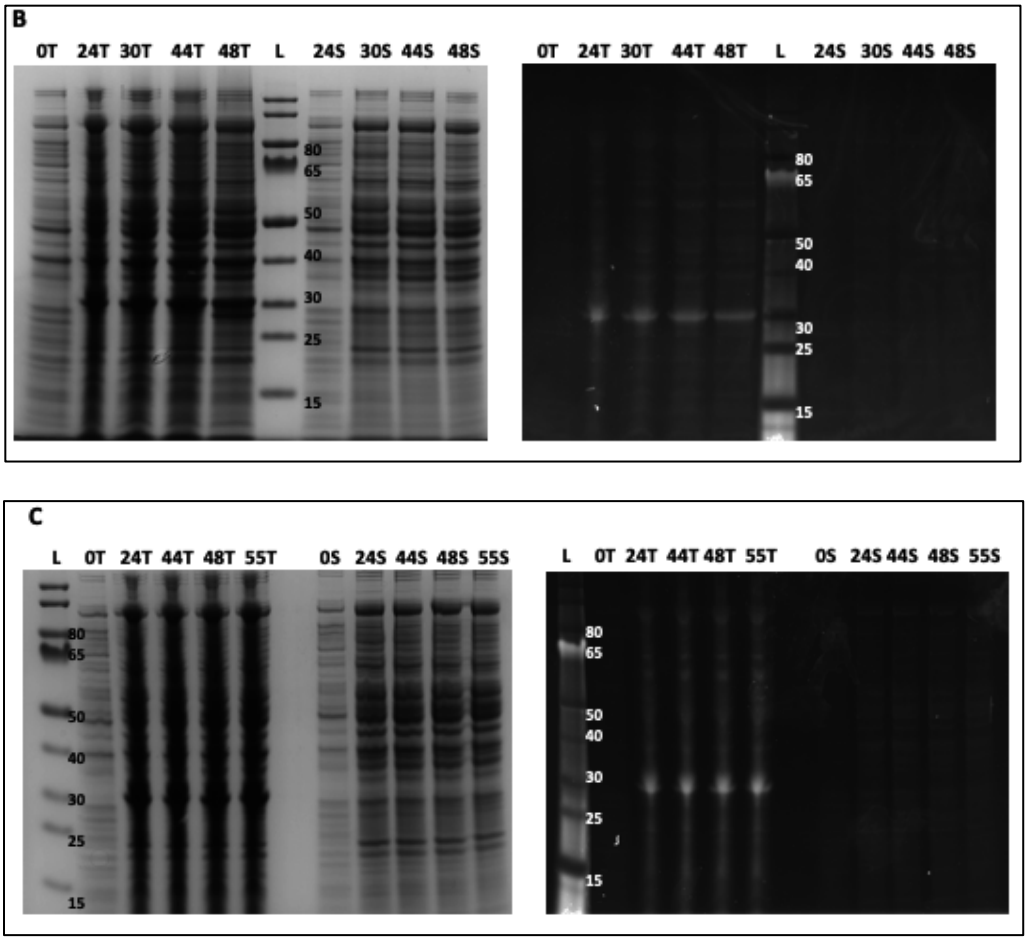
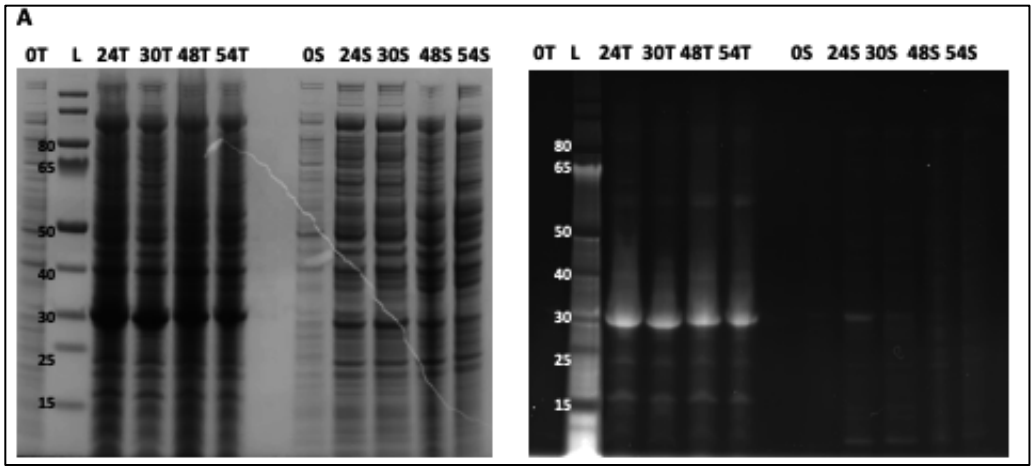
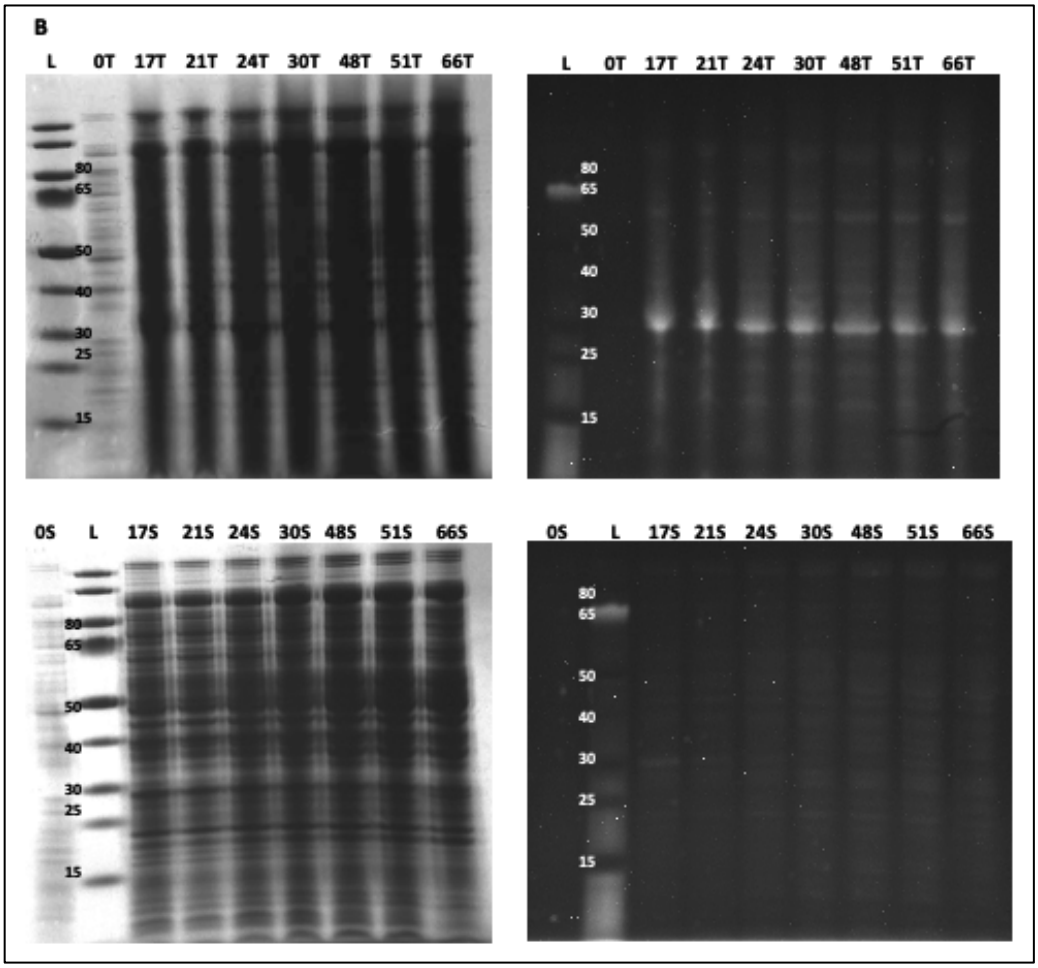


Figure 11. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 30°C in TB media and induced with 0.0125 mM IPTG, and in the presence of 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 55 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
0T L 24T 30T 48T 54T 0S 24S 30S 48S 54S
80
65
50
40
30
25
15
0T L 24T 30T 48T 54T 0S 24S 30S 48S 54S
80
65
50
40
30
25
15


B
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15

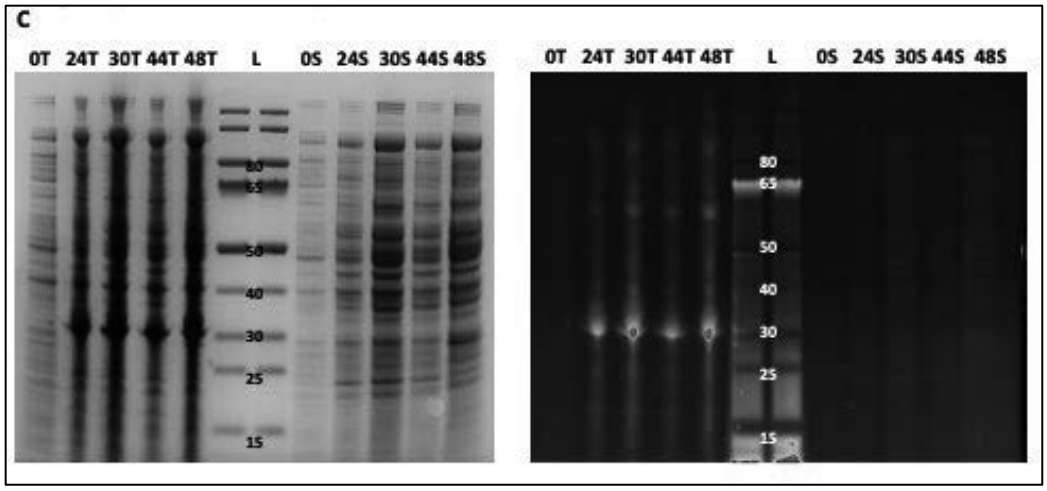
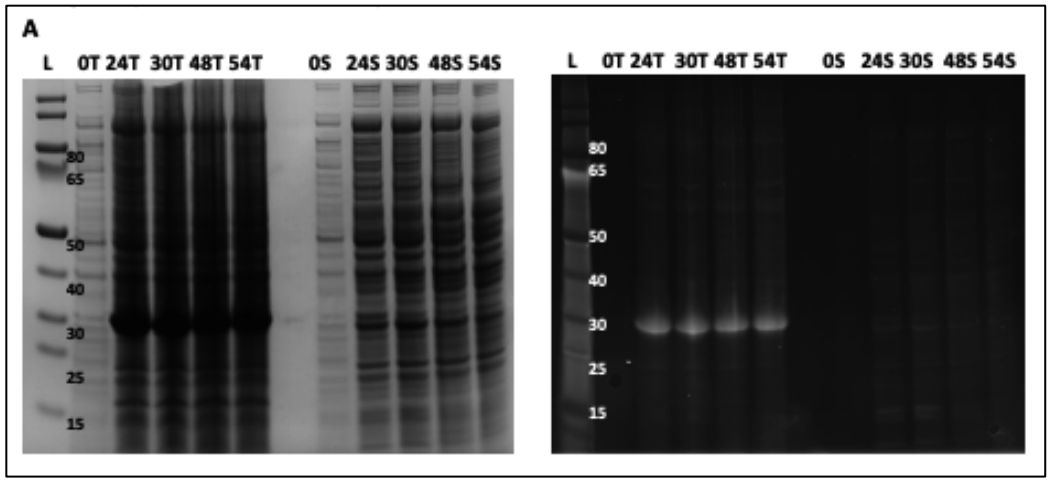
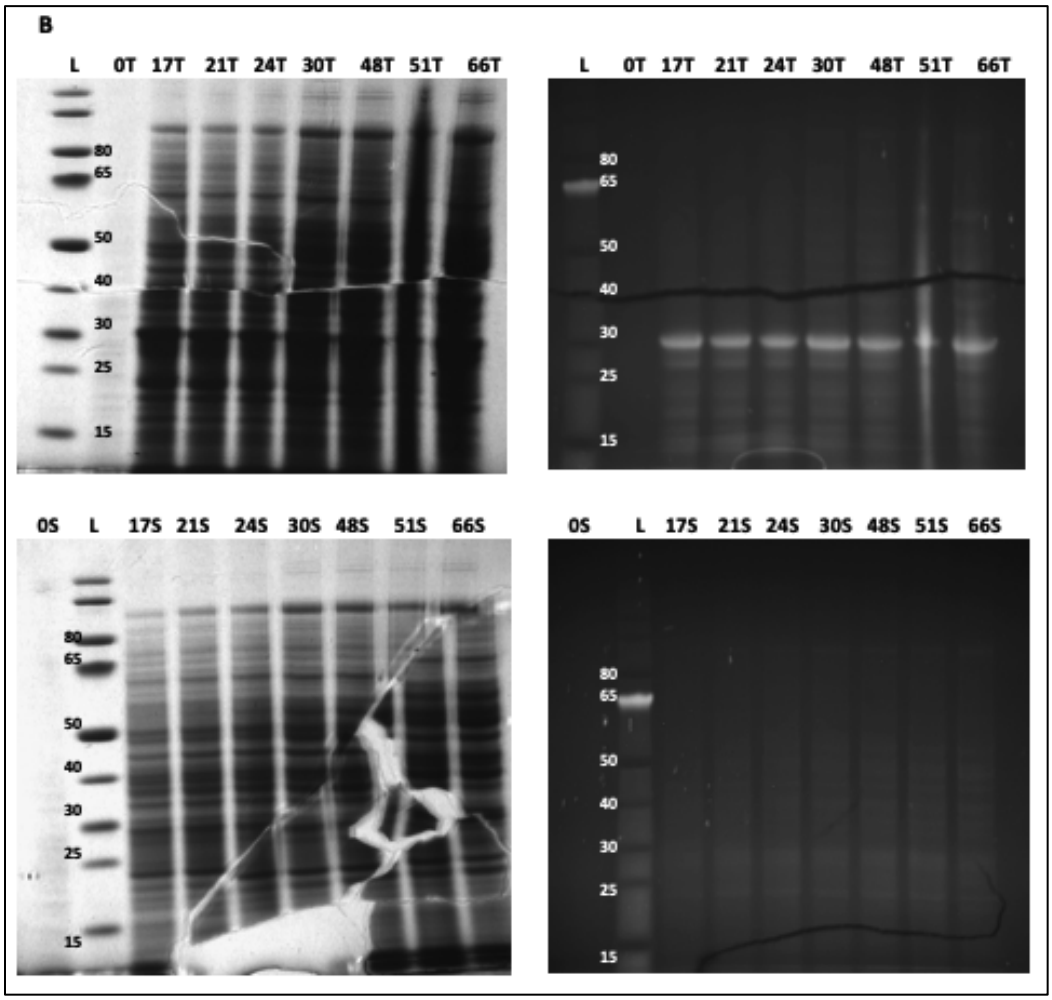


Figure 12. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 30°C in TB media and induced with 0.00625 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 66 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel. B. Repeat experiment set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
L 0T 24T 30T 48T 54T 0S 24S 30S 48S 54S
80
65
50
40
30
25
15
L 0T 24T 30T 48T 54T 0S 24S 30S 48S 54S
80
65
50
40
30
25
15


B
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
L 0T 17T 21T 24T 30T 48T 51T 66T
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15
0S L 17S 21S 24S 30S 48S 51S 66S
80
65
50
40
30
25
15

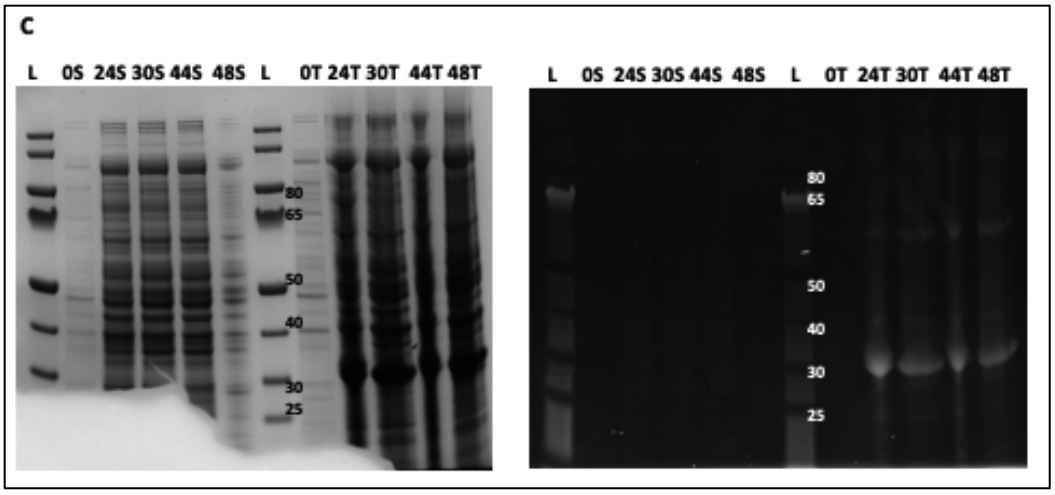
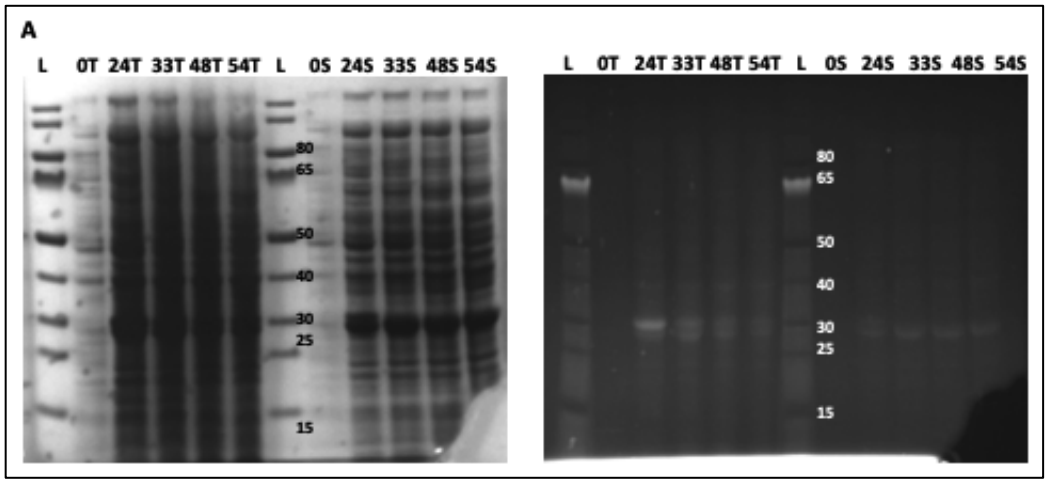


Figure 13. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 30°C in TB media and induced with 0.0025 mM IPTG, and in the presence of 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 66 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). Gels have been ripped but results were still observed and show similarities to one another. C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel.

3.1.2 23°C Growth Experiments

At 23°C, samples induced with 0.025 mM IPTG produced soluble expression (Figure 14). As a matter of fact, soluble expression can be observed at each condition tested at 23°C but unfortunately each sample ultimately led to possible autoactivation of Early Trypsin into the active mature form (Figures 14-19). To fully confirm this, activity assays in concert with the gel analysis need to be run to ensure this indeed is the case. Nonetheless, as observed in

the his-stained gels, the intensity of the soluble bands at ~27 kDa decreases over time which is likely due to the enzyme autocatalyzing and cleaving the N-terminal his_6-tag resulting in a lower molecular weight shift change to ~25 kDa. It is important to note that the his-tag stain only works with proteins that have a his_6-tag, so when the protease cleaves off the N-terminus (where the his_6-tag is located) the signal is lost which can clearly be observed in Figures 14-19. Furthermore, samples at each condition tested seem to follow the same pattern, a shift in MW of AaET from 27 kDa to 25 kDa at around 44- to 48-hours post induction. Interestingly, for the second set of replicates, the shift in AaET protease MW was observed around 54-hours post induction (Figure 14B), yet the third replicate shift is observed around 44-hours post induction (Figure 14C). This is an important reason why different experimental conditions were tested, to find the condition that will lead to the production of the inactive form of AaET, the species observed around 27 kDa.



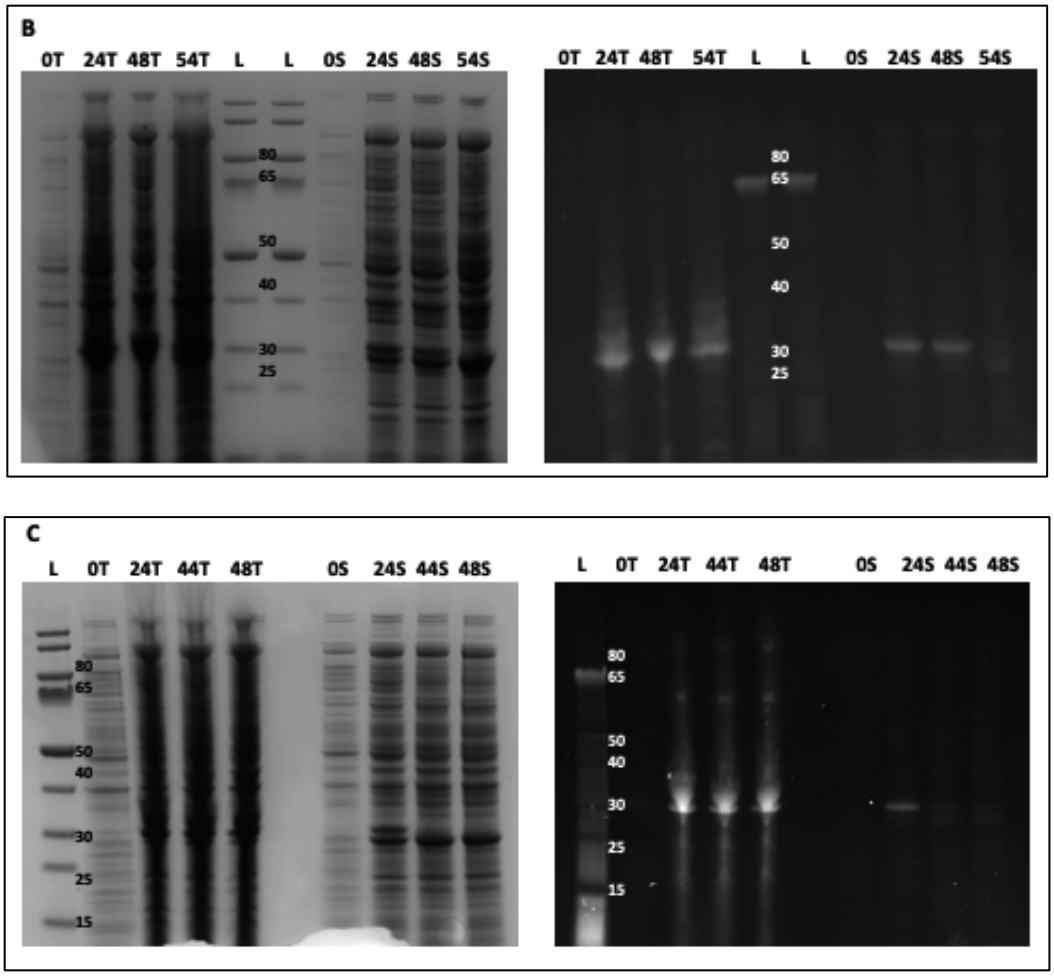
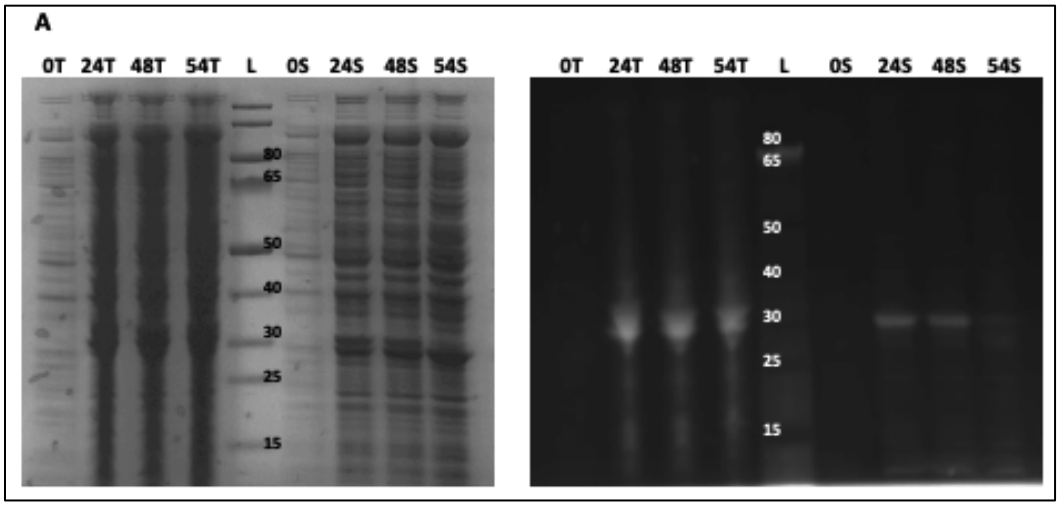
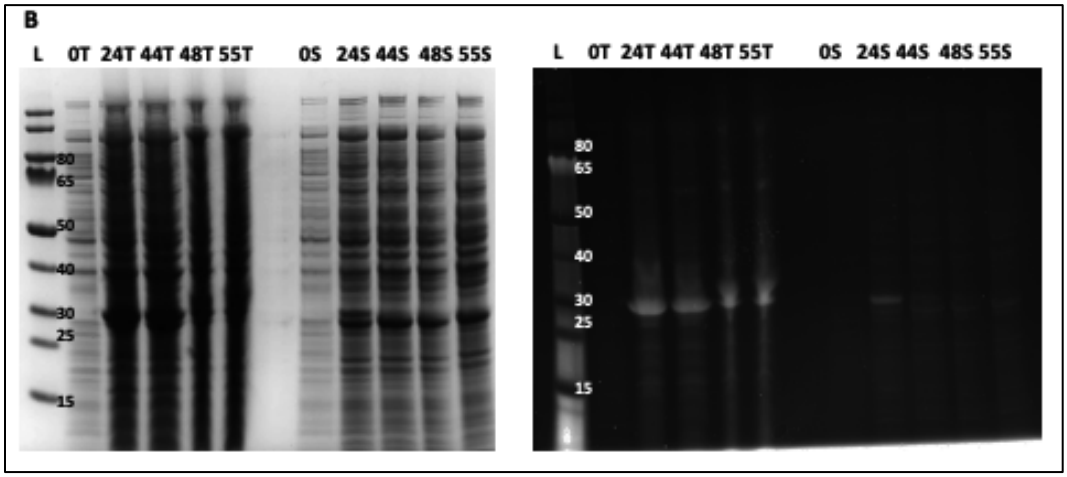


Figure 14. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 23°C in TB media and induced with 0.025 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 54 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
0T 24T 48T 54T L 0S 24S 48S 54S
80
65
50
40
30
25
15
0T 24T 48T 54T L 0S 24S 48S 54S
80
65
50
40
30
25
15


B
L 0T 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15
L 0T 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15

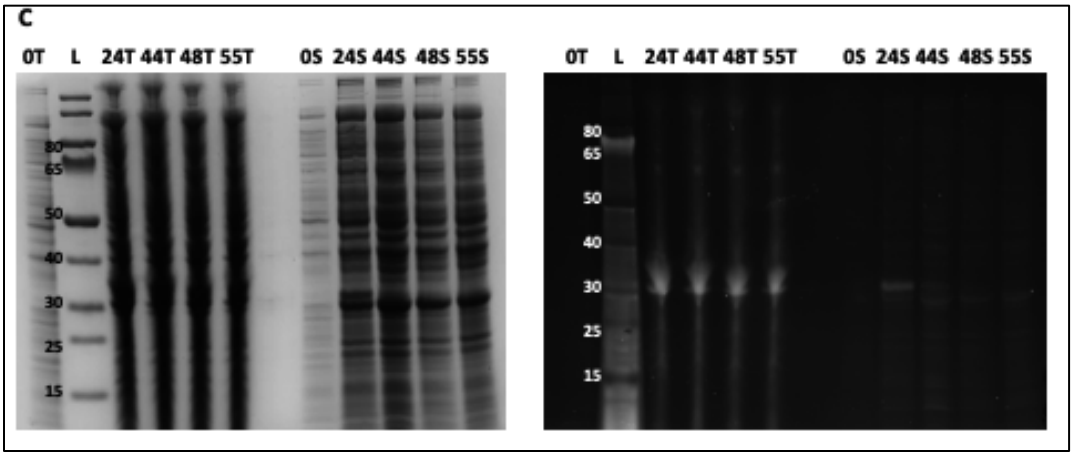
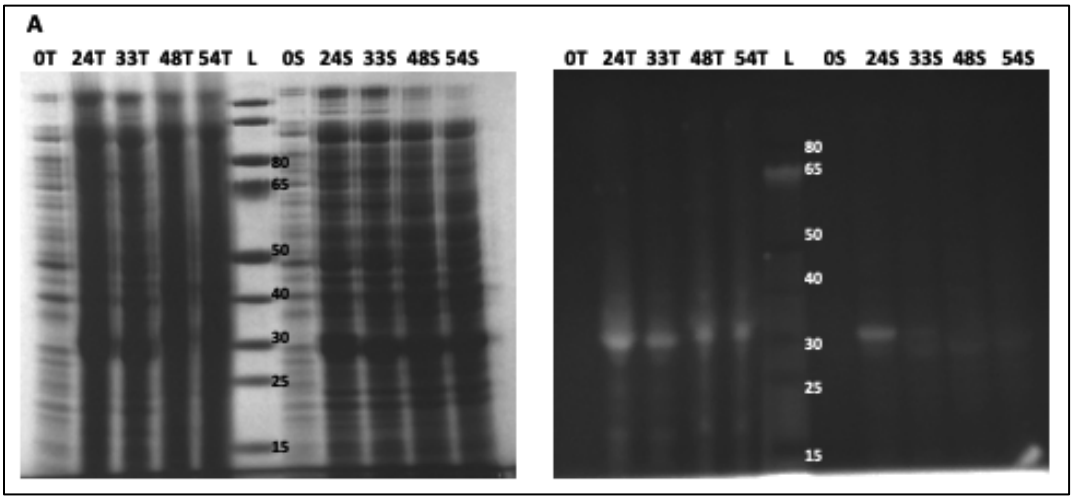


Figure 15. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 23°C in TB media and induced with 0.0125 mM IPTG, and 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 55 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).



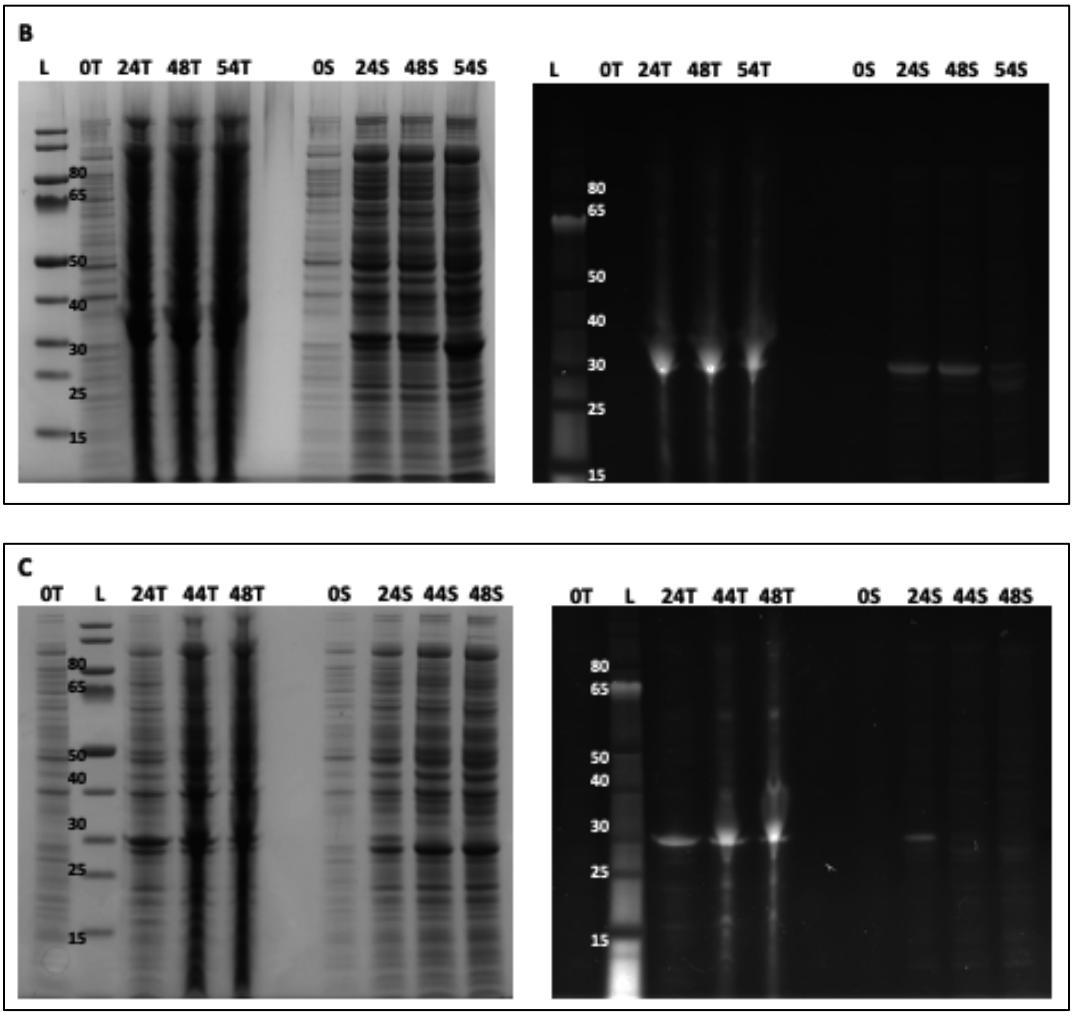
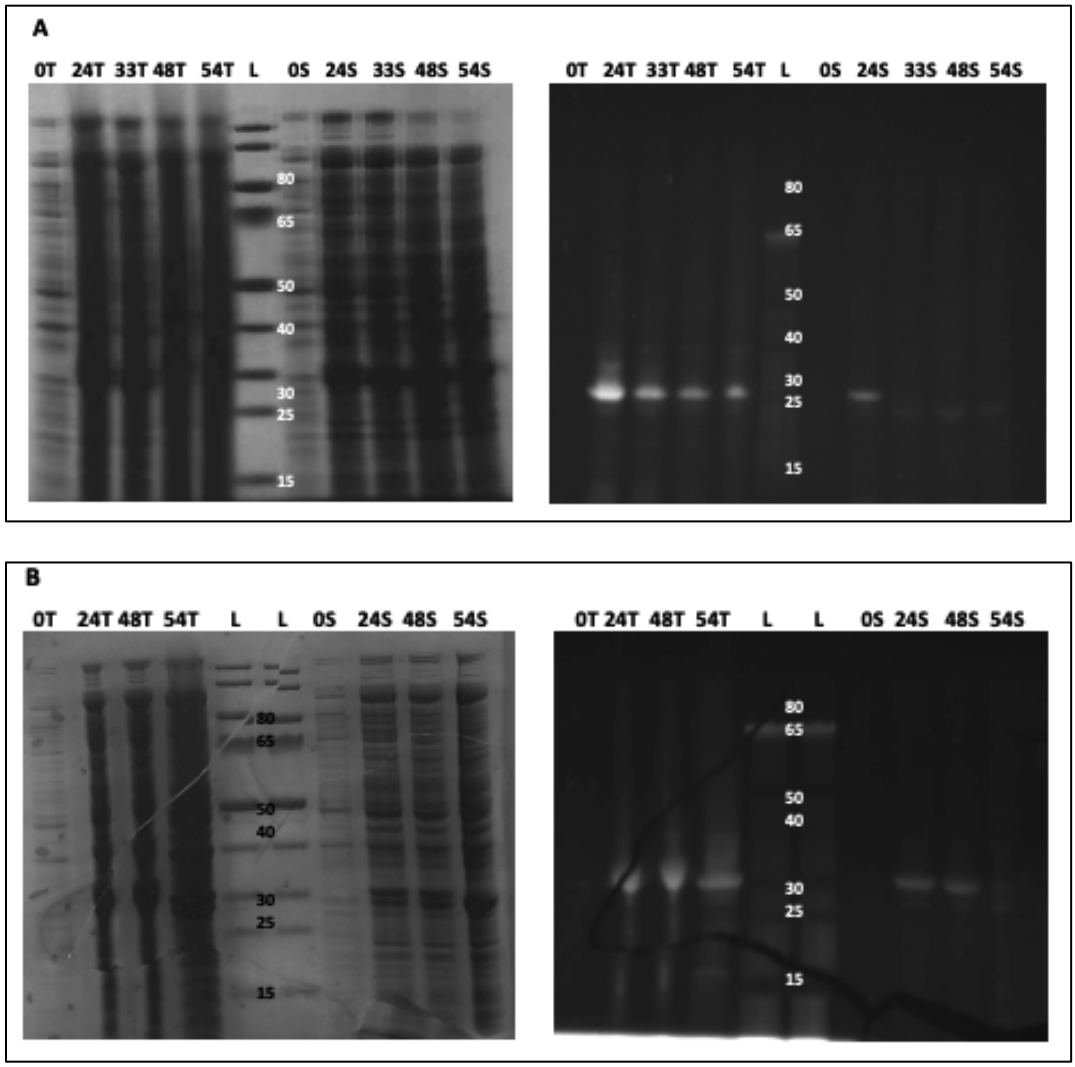


Figure 16. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 23°C in TB media and induced with 0.0125 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 54 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
0T 24T 33T 48T 54T L 0S 24S 33S 48S 54S
80
65
50
40
30
25
15
0T 24T 33T 48T 54T L 0S 24S 33S 48S 54S
80
65
50
40
30
25
15
B
0T 24T 48T 54T L L 0S 24S 48S 54S
80
65
50
40
30
25
15
0T 24T 48T 54T L L 0S 24S 48S 54S
80
65
50
40
30
25
15

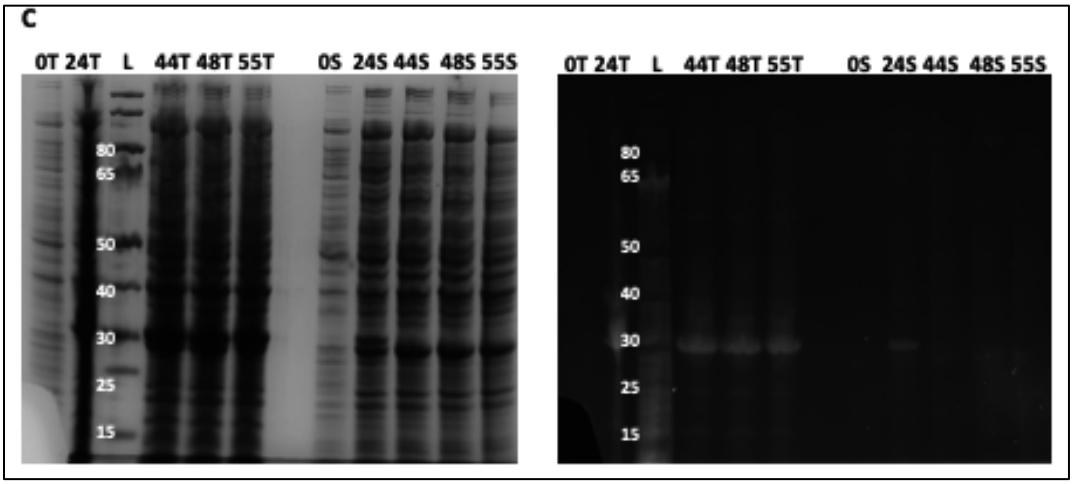
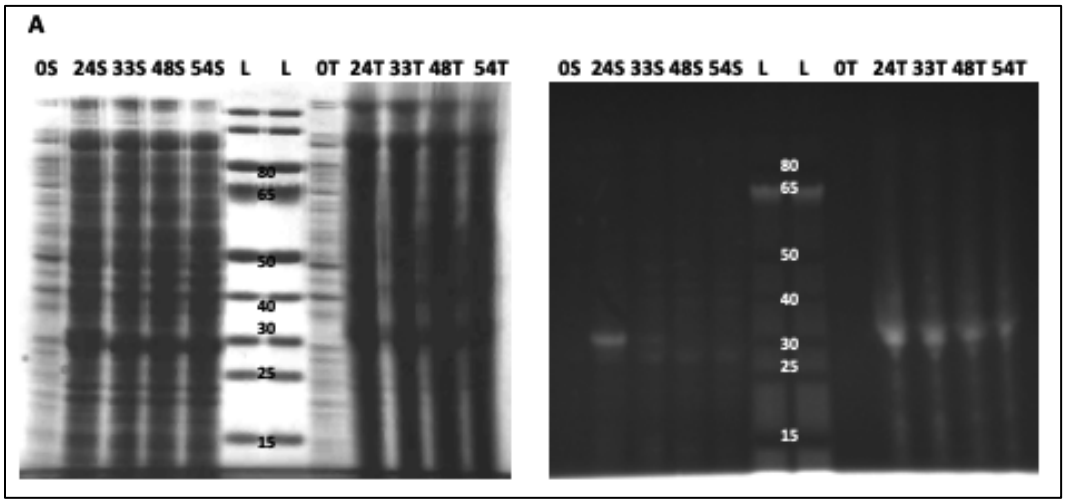


Figure 17. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 23°C in TB media and induced with 0.0125 mM IPTG, and 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 55 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gels have been ripped but results were still observed and show similarities to the his-stained gel. C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).



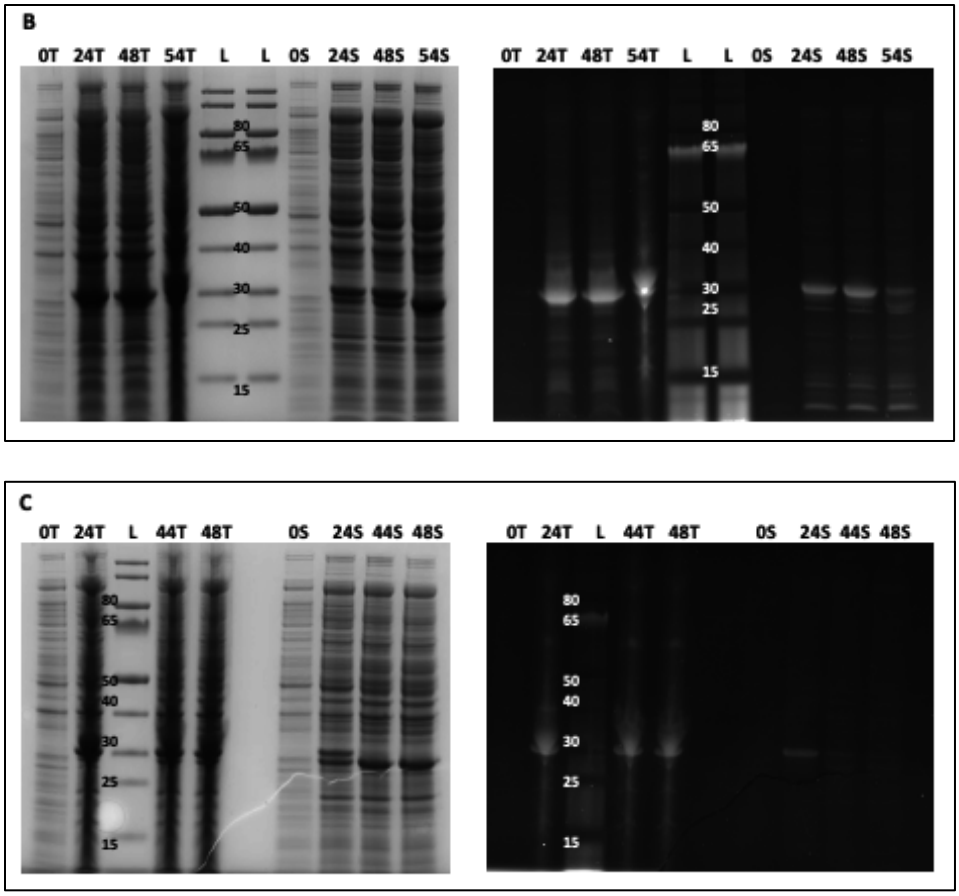
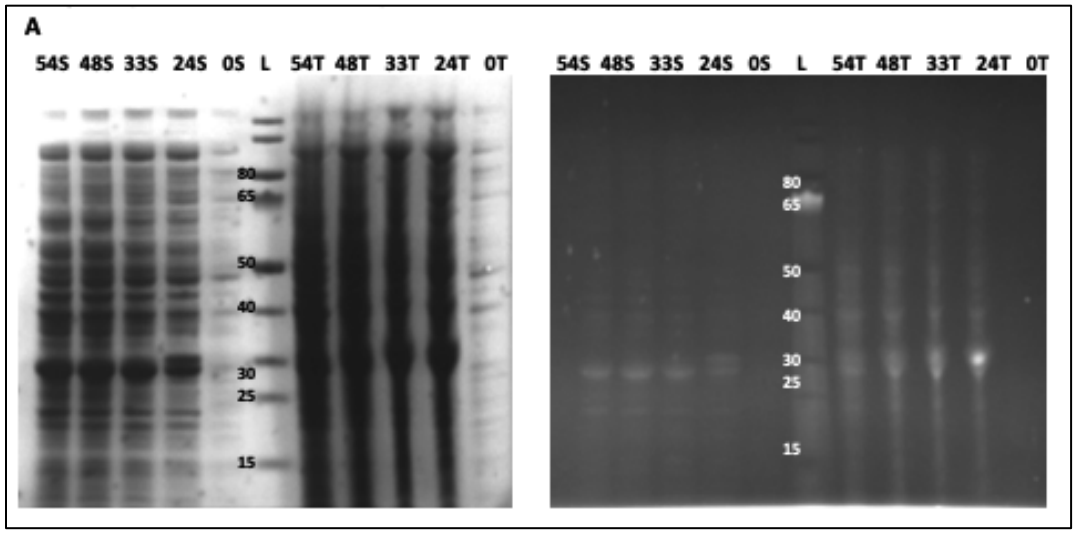
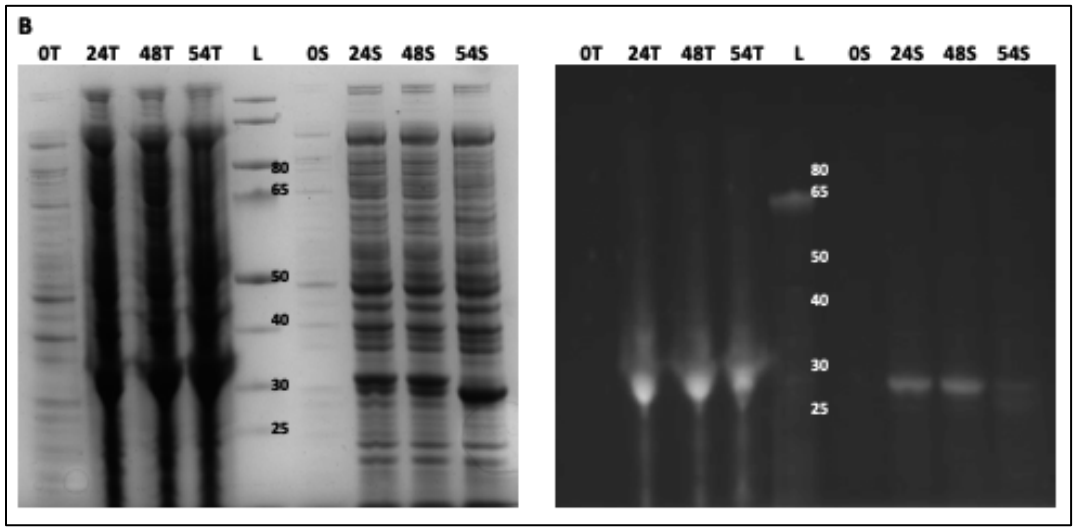


Figure 18. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 23°C in TB media and induced with 0.00625 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 54 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel.


A
54S 48S 33S 24S 0S L 54T 48T 33T 24T 0T
80
65
50
40
30
25
15
54S 48S 33S 24S 0S L 54T 48T 33T 24T 0T
80
65
50
40
30
25
15


B
0T 24T 48T 54T L 0S 24S 48S 54S
80
65
50
40
30
25
0T 24T 48T 54T L 0S 24S 48S 54S
80
65
50
40
30
25

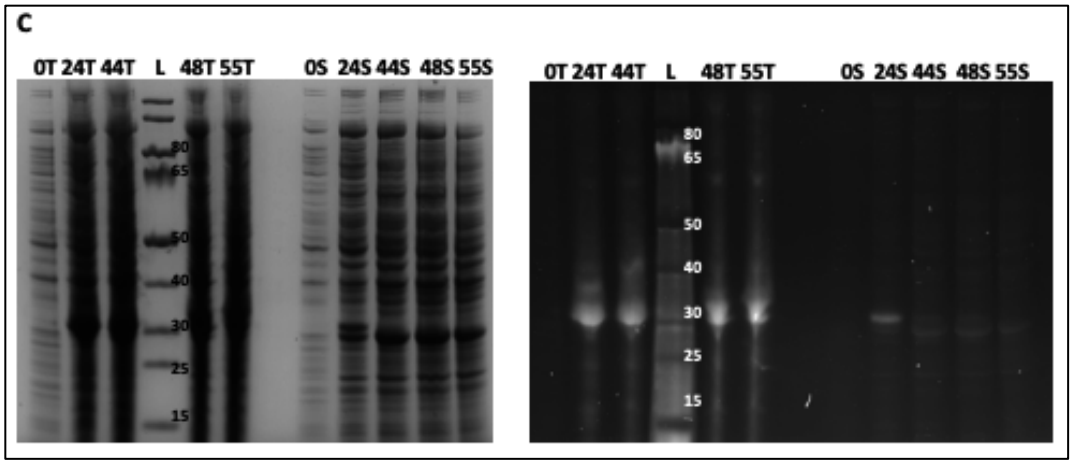
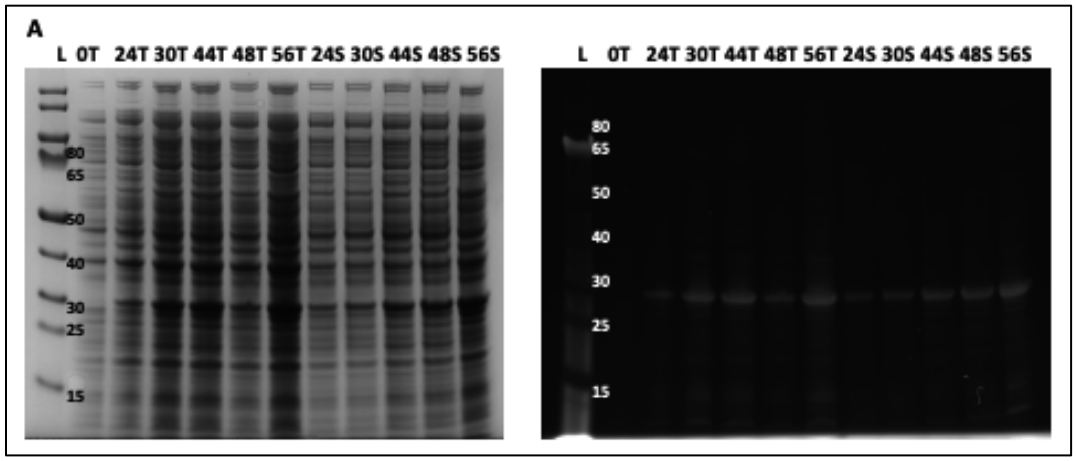


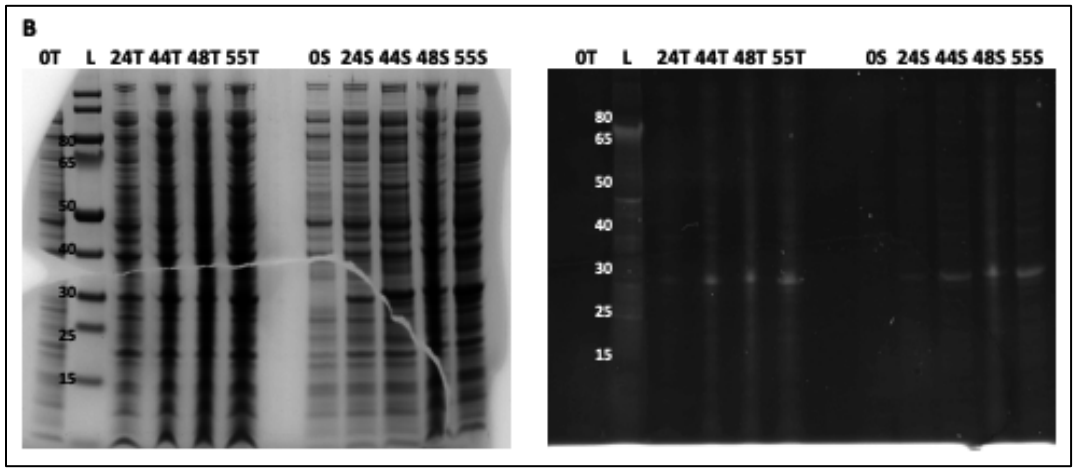
Figure 19. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 23°C in TB media and induced with 0.00625 mM IPTG and 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 55 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).

3.1.3 10°C Growth Experiments

All conditions tested at 10°C produced the most soluble recombinant expression of the inactive form of Early Trypsin (Figures 20-25). In both the his-stained and blue-stained gels for all conditions at 10°C, soluble expression near the 30 kDa marker (AaET MW ~27 kDa) increases in intensity over time (Figures 20-25). More importantly, samples that were induced with 0.025 mM IPTG, either with or without betaine, produced the most soluble expression leading to the most intense soluble bands when compared visually to other IPTG concentrations used for induction. And in comparison, between samples induced with 0.025

mM IPTG, in the presence of 1 mM betaine have slightly more intense bands (Figure 20) than samples induced with only 0.025 mM IPTG (Figure 21). As for the samples induced with 0.0125 mM IPTG (Figure 22) or 0.0125 mM IPTG with 1 mM betaine (Figure 23), each of the two conditions produce equally intense soluble bands. Thus, it is difficult to visually distinguish any differences between the two conditions at 0.0125 mM IPTG, in both blue-stained and his-stained gels, intensity of soluble expression bands increase over time. Lastly, for samples induced with 0.00625 mM IPTG, only the samples that did not contain betaine produced intense soluble expression bands (Figure 24) compared to cells induced with 0.00625 mM IPTG in the presence of 1 mM betaine, which produced dim soluble expression bands (Figure 25). This is surprising given that betaine improved solubility in all growth experiments when used (this will fully be discussed in the discussion, see Chapter 4 Section 1). Furthermore, similar to the previous conditions mentioned, intensity of soluble bands increase over time as observed in both his-stained and blue-stained gels (Figures 24 and 25). It should be noted that at 48-hours post induction in figure 24A, the soluble band is dimmer than the soluble band at 44-hours post induction which was likely due to the loss of protein when samples were being loaded onto the gel. This was a human error mistake on my part, but consistency was still maintained for all samples and more replicated further needed for better statistical analysis (see section 3.2).


A
L 0T 24T 30T 44T 48T 56T 24S 30S 44S 48S 56S
80
65
50
40
30
25
15
L 0T 24T 30T 44T 48T 56T 24S 30S 44S 48S 56S
80
65
50
40
30
25
15


B
0T L 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15
0T L 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15

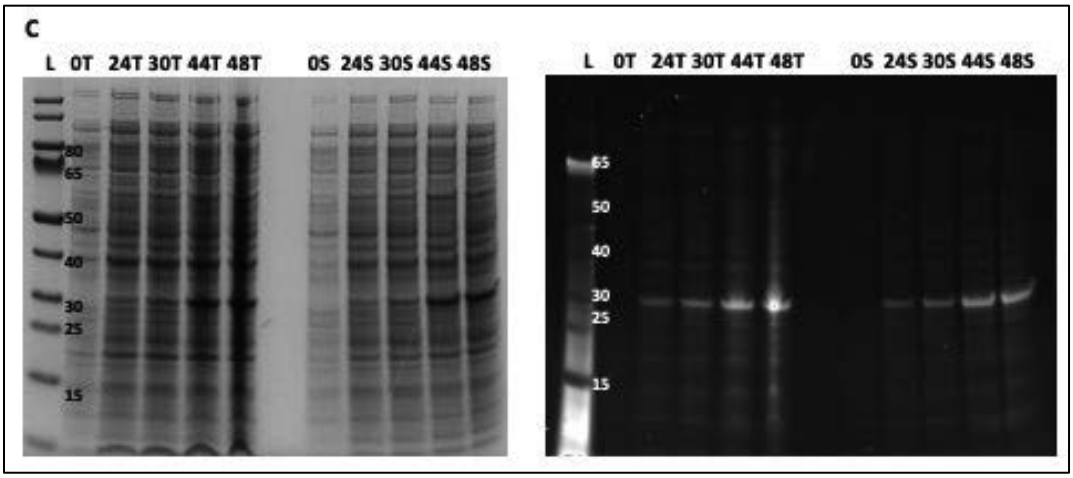
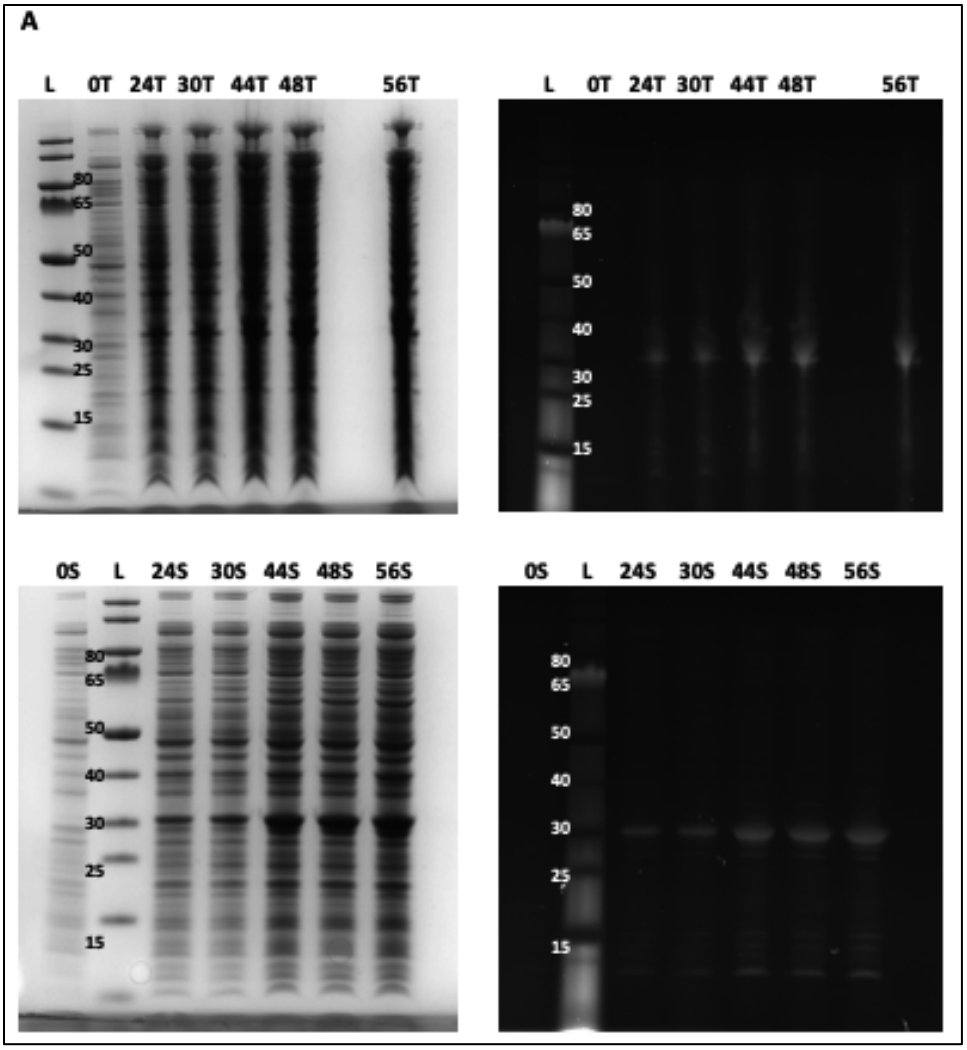
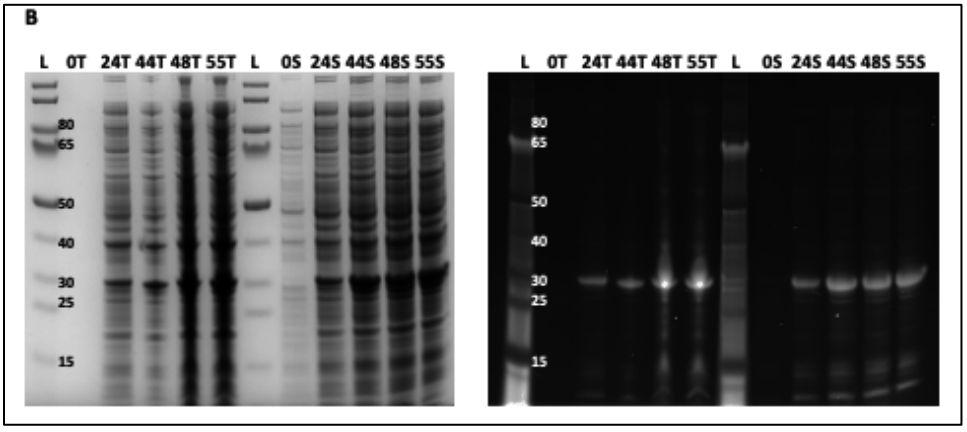


Figure 20. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 10°C in TB media and induced with 0.025 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 56 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel. C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
L 0T 24T 30T 44T 48T 56T
80
65
50
40
30
25
15
L 0T 24T 30T 44T 48T 56T
80
65
50
40
30
25
15
0S L 24S 30S 44S 48S 56S
80
65
50
40
30
25
15
0S L 24S 30S 44S 48S 56S
80
65
50
40
30
25
15


B
L 0T 24T 44T 48T 55T L 0S 24S 44S 48S 55S
80
65
50
40
30
25
15
L 0T 24T 44T 48T 55T L 0S 24S 44S 48S 55S
80
65
50
40
30
25
15

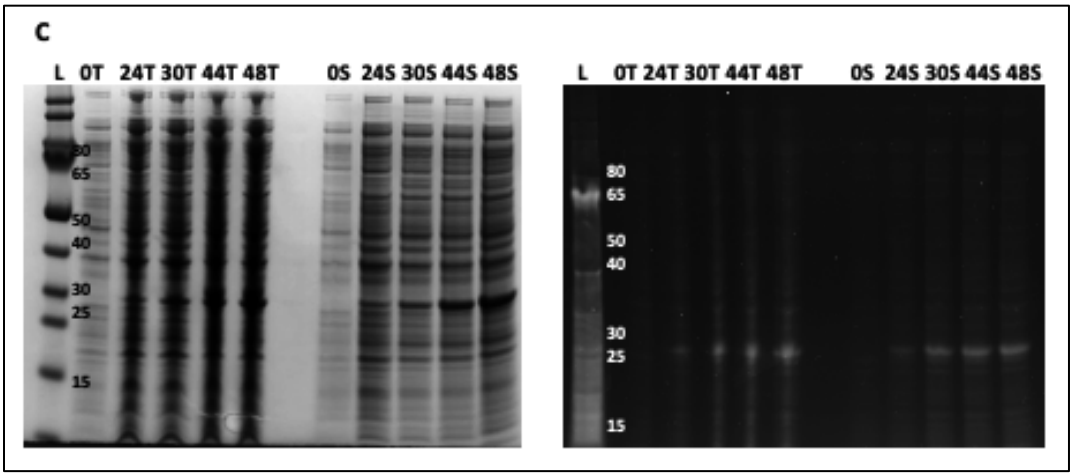
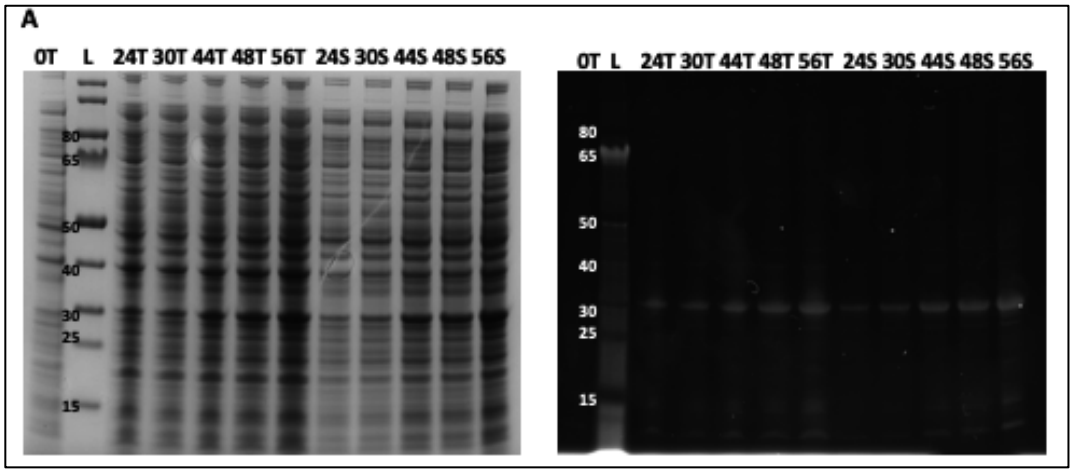


Figure 21. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 10°C in TB media and induced with 0.025 mM IPTG, and 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 56 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).



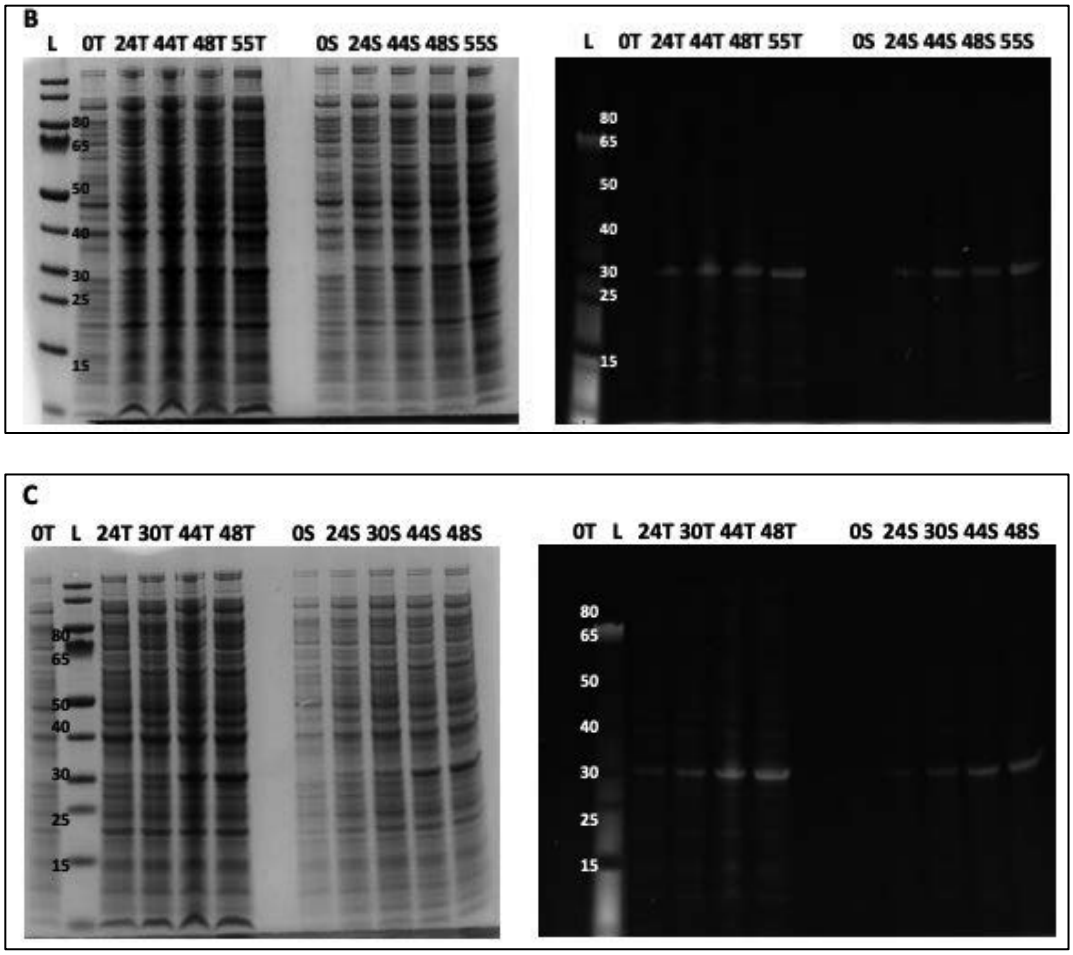
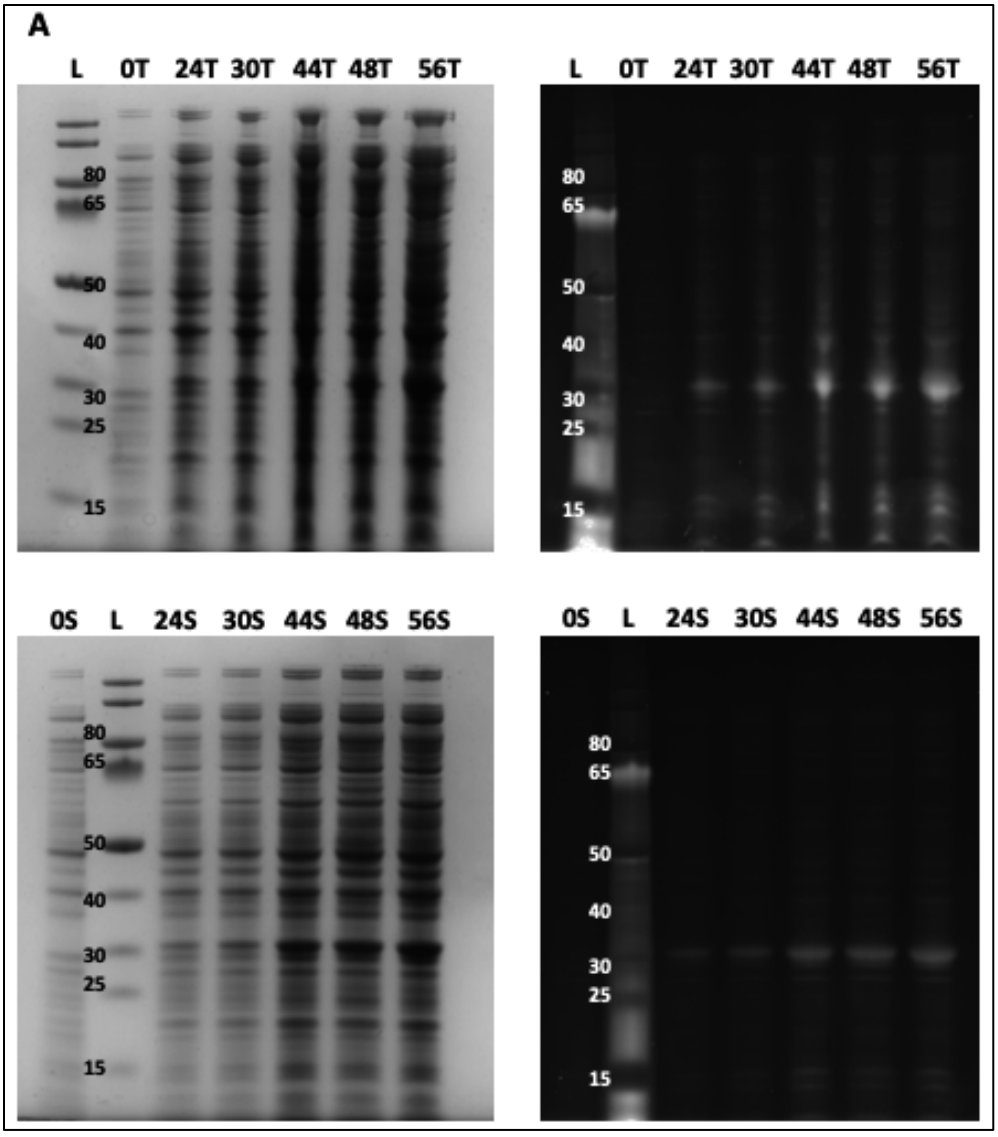
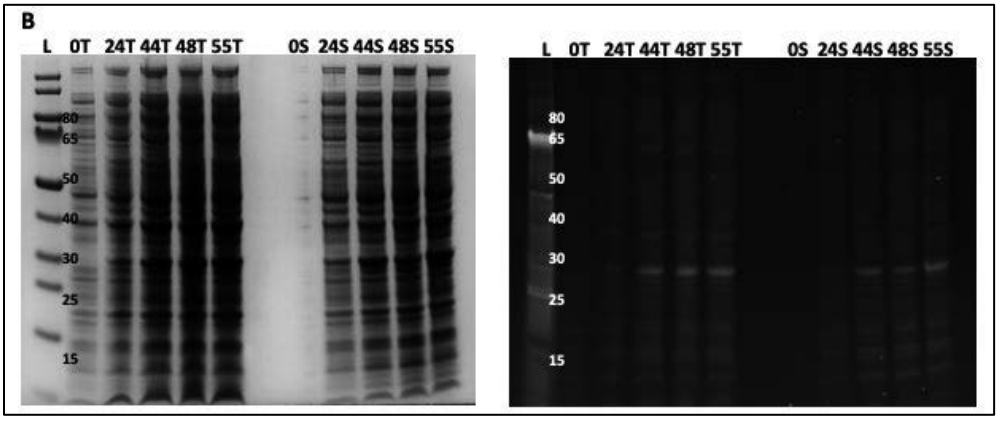


Figure 22. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 10°C in TB media and induced with 0.0125 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 56 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel. B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
L 0T 24T 30T 44T 48T 56T
80
65
50
40
30
25
15
L 0T 24T 30T 44T 48T 56T
80
65
50
40
30
25
15
0S L 24S 30S 44S 48S 56S
80
65
50
40
30
25
15
0S L 24S 30S 44S 48S 56S
80
65
50
40
30
25
15


B
L 0T 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15
L 0T 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15

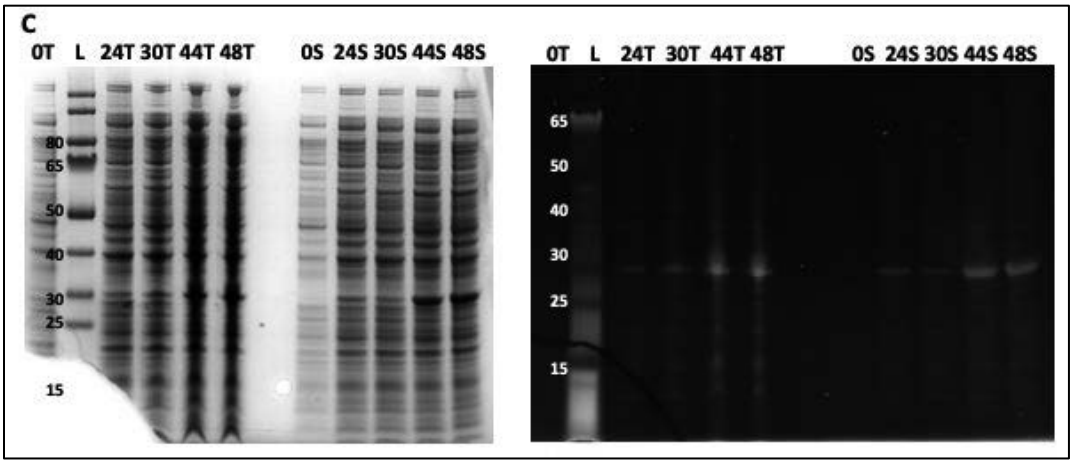
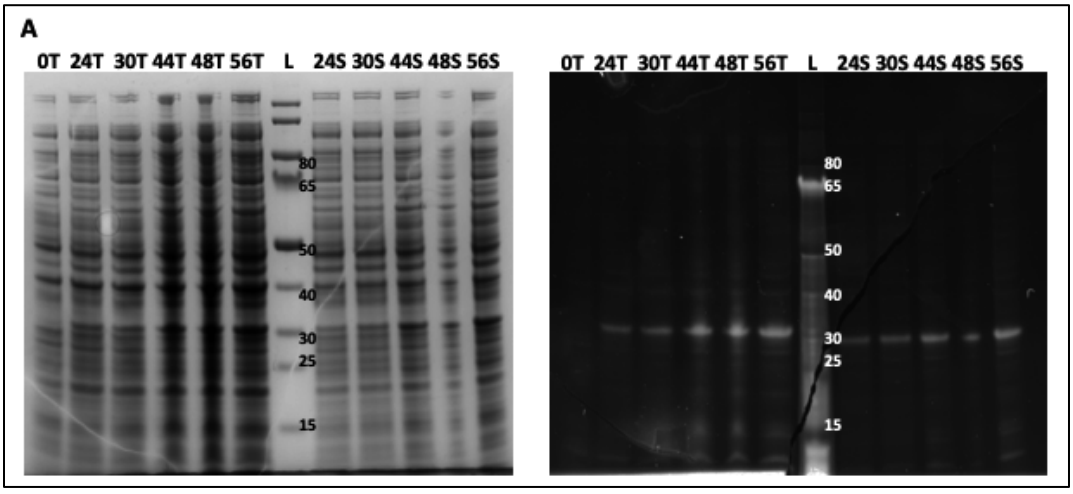


Figure 23. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 10°C in TB media and induced with 0.0125 mM IPTG, and 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 56 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel.



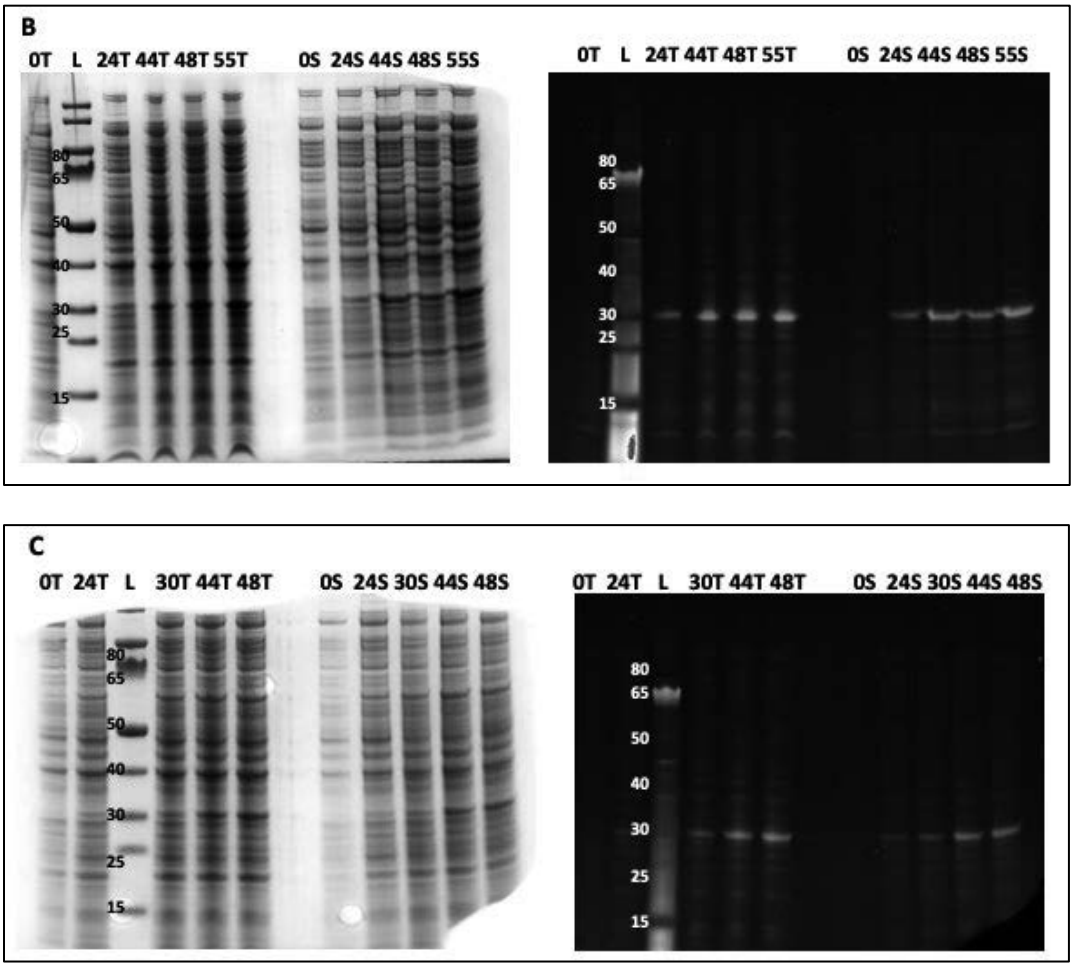
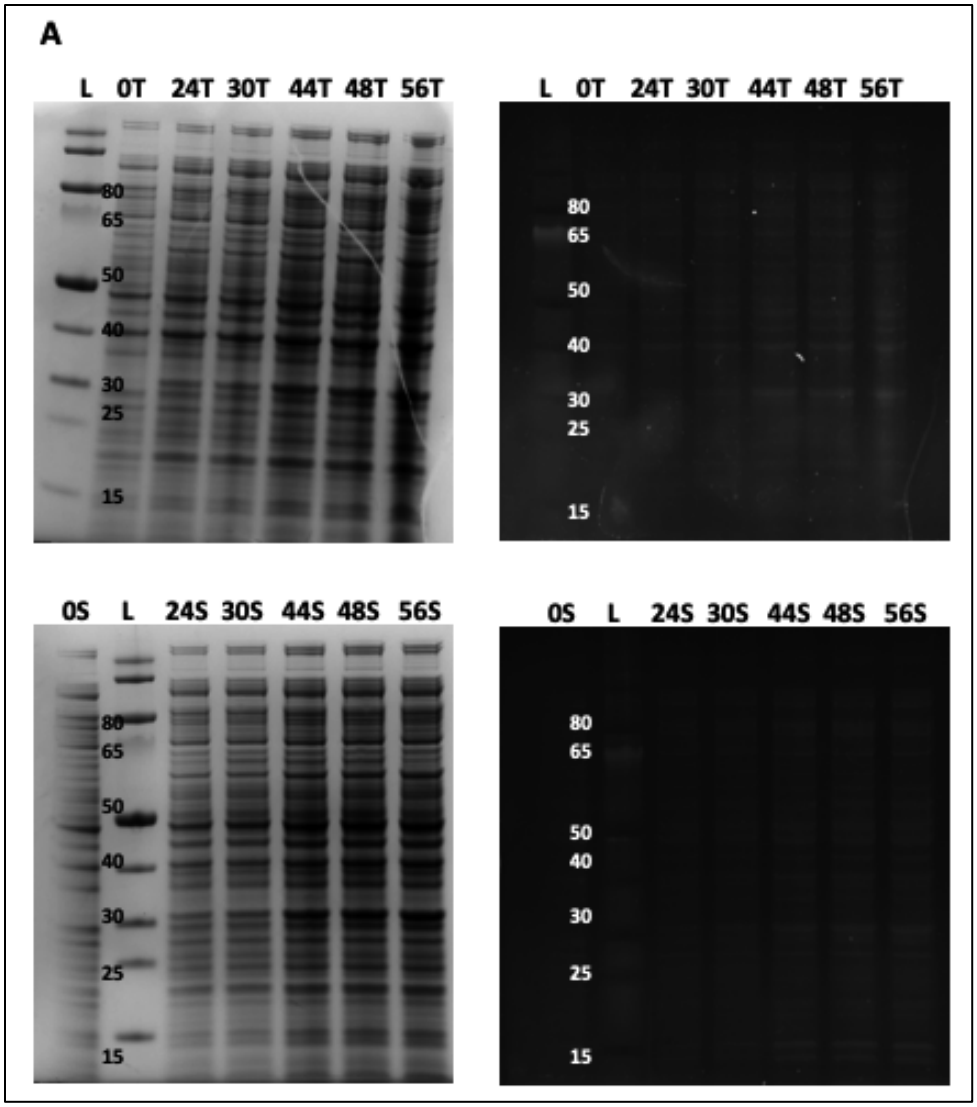
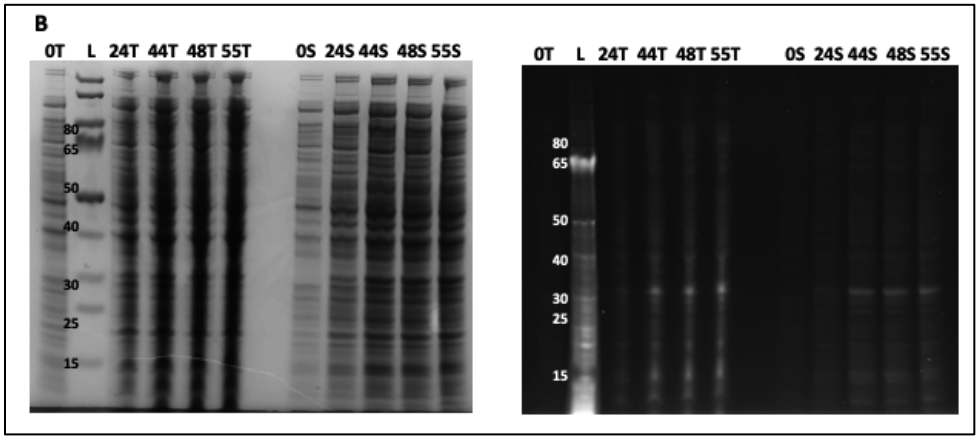


Figure 24. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 10°C in TB media and induced with 0.00625 mM IPTG. Various time points were collected starting at 0 hours (pre-induction) going up to 55 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gel has been ripped but results were still observed and show similarities to the his-stained gel. B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).


A
L 0T 24T 30T 44T 48T 56T
80
65
50
40
30
25
15
L 0T 24T 30T 44T 48T 56T
80
65
50
40
30
25
15
0S L 24S 30S 44S 48S 56S
80
65
50
40
30
25
15
0S L 24S 30S 44S 48S 56S
80
65
50
40
30
25
15


B
0T L 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15
0T L 24T 44T 48T 55T 0S 24S 44S 48S 55S
80
65
50
40
30
25
15

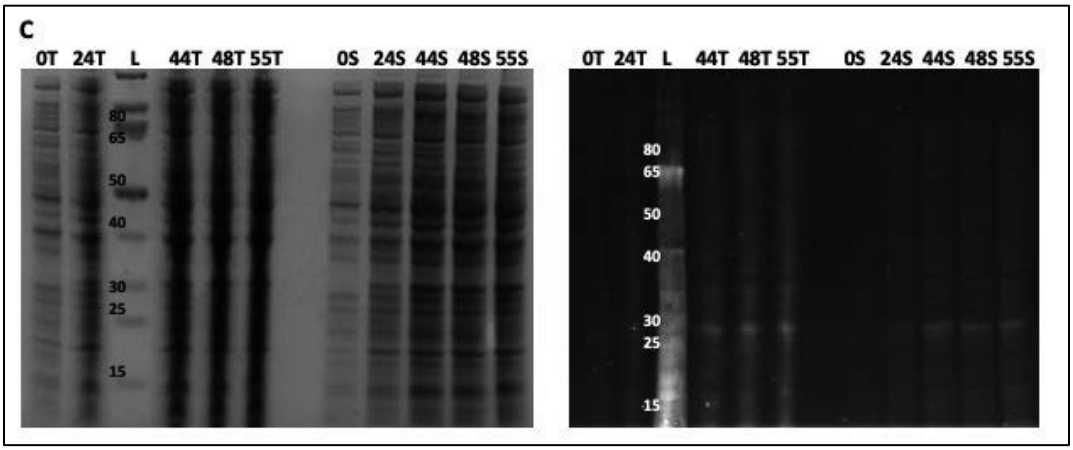


Figure 25. SDS–PAGE results of the recombinant expression of the AaET-WT-NL protease using T7 SHuffle Competent Cells grown at 10°C in TB media and induced with 0.00625 mM IPTG, and 1 mM Betaine. Various time points were collected starting at 0 hours (pre-induction) going up to 56 hours (post-induction). The samples were run on 4-12% Bis-Tris gels with 1x MES buffer and stained with his-stain (dark background and white bands) and simply safe blue stain (grey background with dark bands). The ladder (L) contains protein band standards with the corresponding molecular weight in kDa. AaET recombinant protein should have been observed at around 27.0 kDa (red arrow). A. First set of total (top half) and soluble (bottom half) samples stained with simply safe blue stain (left) and with his-stain (right). Blue stained gels have been ripped but results were still observed and show similarities to the his-stained gel. B. Repeat experiment set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right). C. Third triplicate set of total and soluble samples stained with simply safe blue stain (left) and with his-stain (right).

3.2 Gel Band Quantification

With the preparation and gel analysis of all growth experiment samples, band intensities of his-stained gels were measured to determine statistical differences between protein expression in Total and Soluble bands, with the different temperature conditions, concentrations of IPTG, and samples containing the betaine osmolyte. The ImageJ program was used to measure gel band intensity by specifically focusing on the inactive form of AaET at the ~27 kDa marker and calculating the area and pixel values of each selection (see

Chapter 2 section 3 for detailed information). It is important to note that only the inverted his-stained gels were utilized since this will produce dark protein bands of the his_6-tagged AaET protease with a light background, a condition needed by ImageJ to perform pixel detection properly. Using ImageJ, the pixels in each selection were used to generate plots (Figures 26-31). Each peak represents the number of pixels in each band compared to the time zero background, and the area of each peak was used for statistical analysis studies using GraphPad Prism. Since soluble expression is not observed in any samples at 30°C, band intensity was not measured for statistical analysis data. Similarly, since the inactive form of the protease seems to have disappeared due to possible autoactivation at 23°C it was difficult to measure band intensity of the inactive form of the enzyme, thus, band intensity was not measured for statistical analysis data. Since all samples at 10°C produce soluble recombinant expression of the inactive form, band intensity was measured for each condition. Only the 0-, 24-, 44-, and 48-hour time points were considered when obtaining peak area data as they were the most consistent time points across all the 10°C conditions.

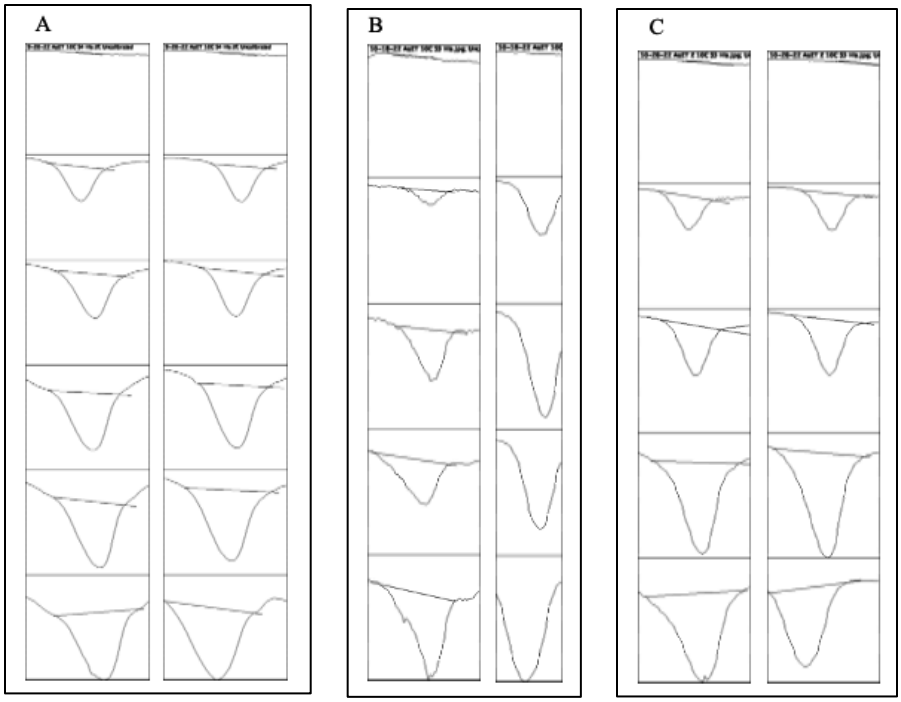


Figure 26. ImageJ plots generated from gel band selections of AaET-WT-NL recombinant protease expression grown at 10°C in TB media and induced with 0.025 mM IPTG. Plots that were generated by ImageJ to help measure the intensity (pixels) of each band (Figure 20 gels used for this analysis). The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences. Each peak represents a time point in which samples were collected, starting at 0 hours (post induction) going to 48 hours (post induction). A. First set of total (left) and soluble (right) samples plotted. B. Repeat experiment set of total (left) and soluble (right) samples plotted. C. Third triplicate set of total (left) and soluble (right) samples plotted.

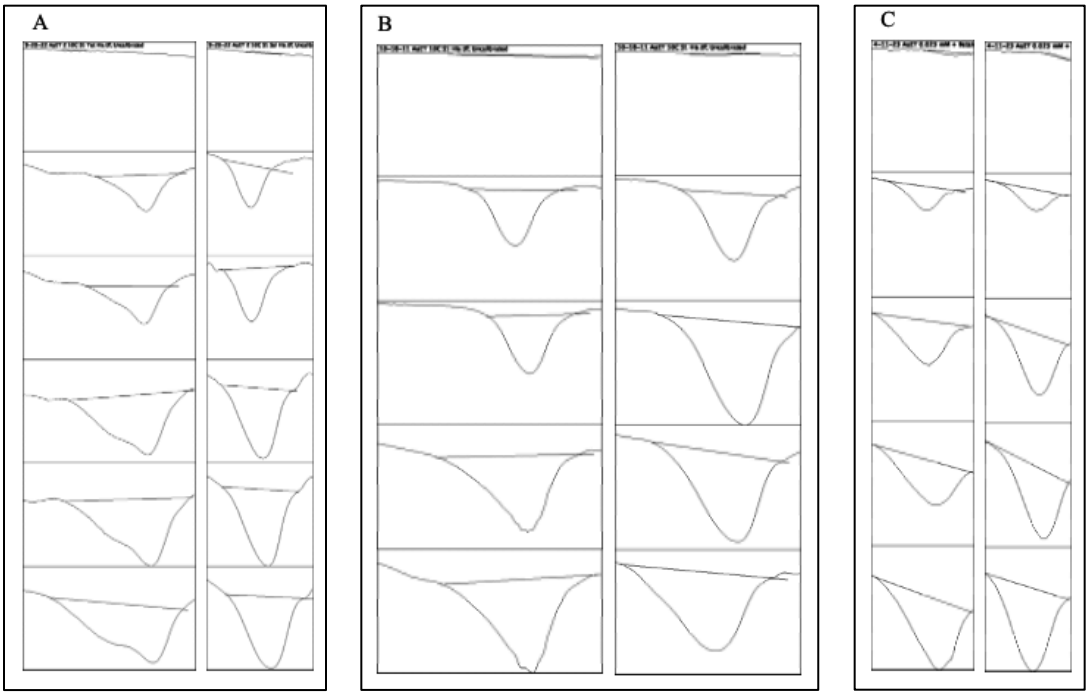


Figure 27. ImageJ plots generated from gel band selections of AaET-WT-NL recombinant protease expression grown at 10°C in TB media and induced with 0.025 mM IPTG and 1 mM Betaine. Plots that were generated by ImageJ to help measure the intensity (pixels) of each band (Figure 21 gels used for this analysis). The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences. Each peak represents a time point in which samples were collected, starting at 0 hours (post induction) going to 48 hours (post induction). A. First set of total (left) and soluble (right) samples plotted. B. Repeat experiment set of total (left) and soluble (right) samples plotted. C. Third triplicate set of total (left) and soluble (right) samples plotted.

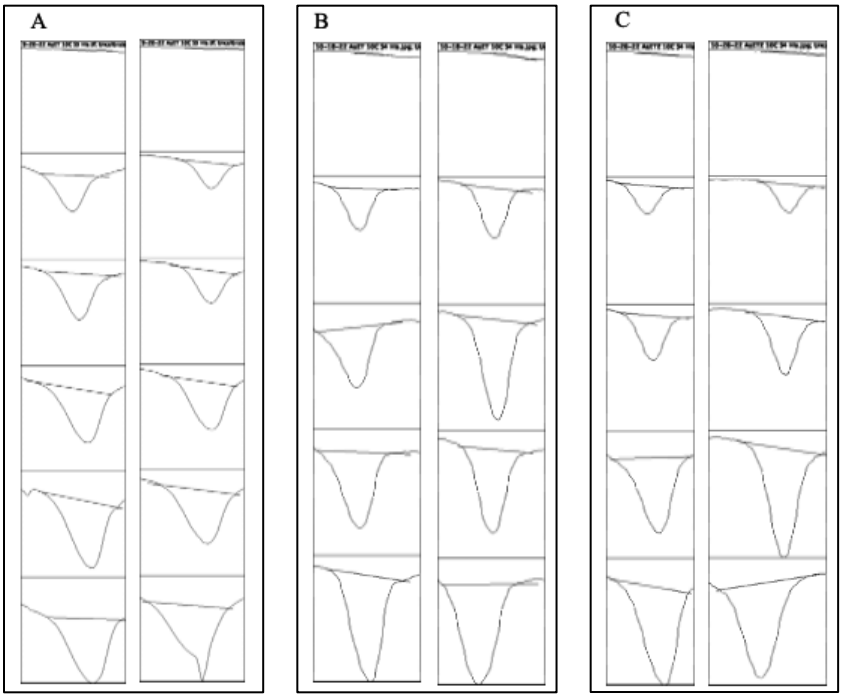


Figure 28. ImageJ plots generated from gel band selections of AaET-WT-NL recombinant protease expression grown at 10°C in TB media and induced with 0.0125 mM IPTG. Plots that were generated by ImageJ to help measure the intensity (pixels) of each band (Figure 22 gels used for this analysis). The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences. Each peak represents a time point in which samples were collected, starting at 0 hours (post induction) going to 48 hours (post induction). A. First set of total (left) and soluble (right) samples plotted. B. Repeat experiment set of total (left) and soluble (right) samples plotted. C. Third triplicate set of total (left) and soluble (right) samples plotted.

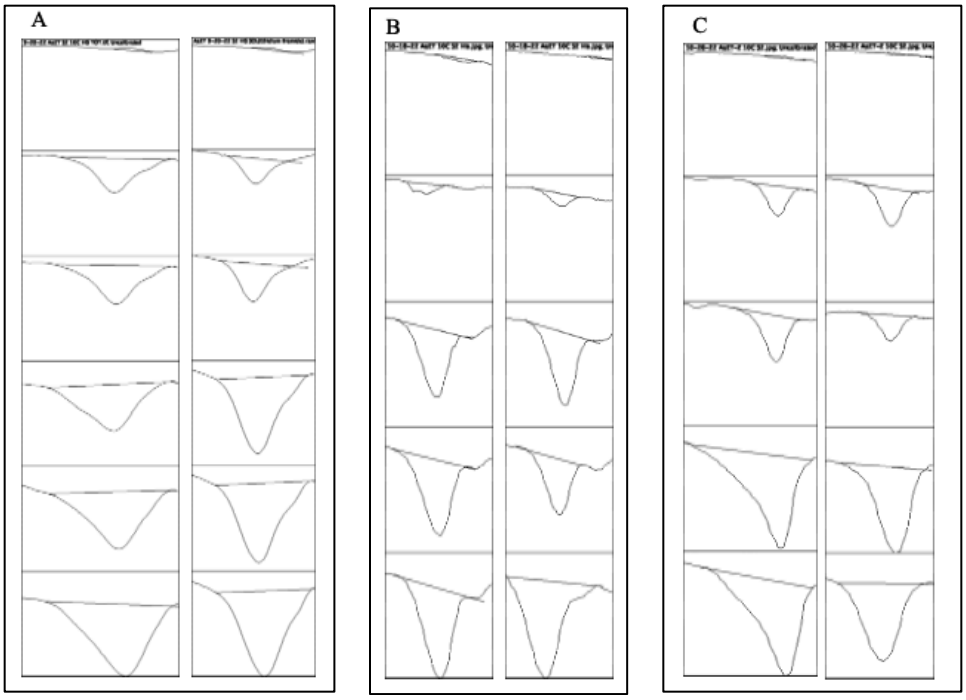


Figure 29. ImageJ plots generated from gel band selections of AaET-WT-NL recombinant protease expression grown at 10°C in TB media and induced with 0.0125 mM IPTG and 1 mM Betaine. Plots that were generated by ImageJ to help measure the intensity (pixels) of each band (Figure 23 gels used for this analysis). The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences. Each peak represents a time point in which samples were collected, starting at 0 hours (post induction) going to 48 hours (post induction). A. First set of total (left) and soluble (right) samples plotted. B. Repeat experiment set of total (left) and soluble (right) samples plotted. C. Third triplicate set of total (left) and soluble (right) samples plotted.

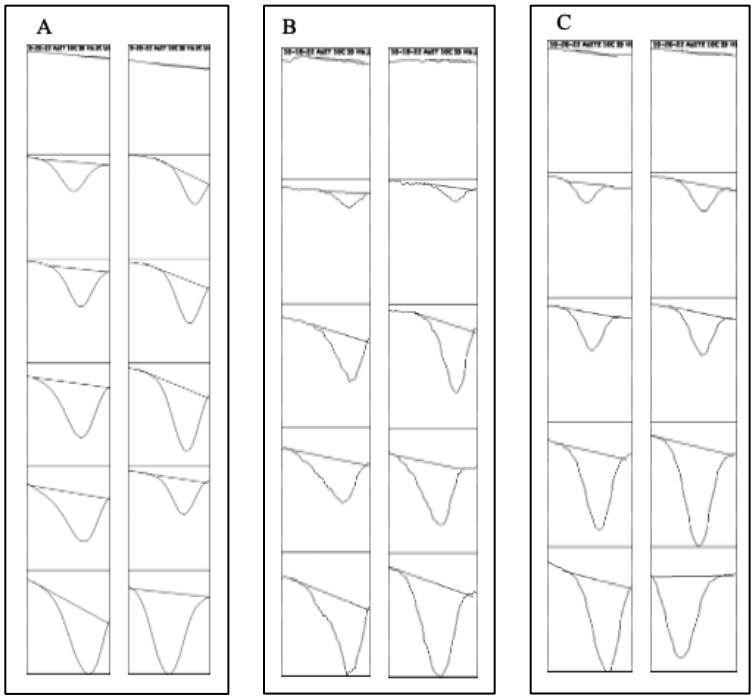


Figure 30. ImageJ plots generated from gel band selections of AaET-WT-NL recombinant protease expression grown at 10°C in TB media and induced with 0.00625 mM IPTG. Plots that were generated by ImageJ to help measure the intensity (pixels) of each band (Figure 24 gels used for this analysis). The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences. Each peak represents a time point in which samples were collected, starting at 0 hours (post induction) going to 48 hours (post induction). A. First set of total (left) and soluble (right) samples plotted. B. Repeat experiment set of total (left) and soluble (right) samples plotted. C. Third triplicate set of total (left) and soluble (right) samples plotted.

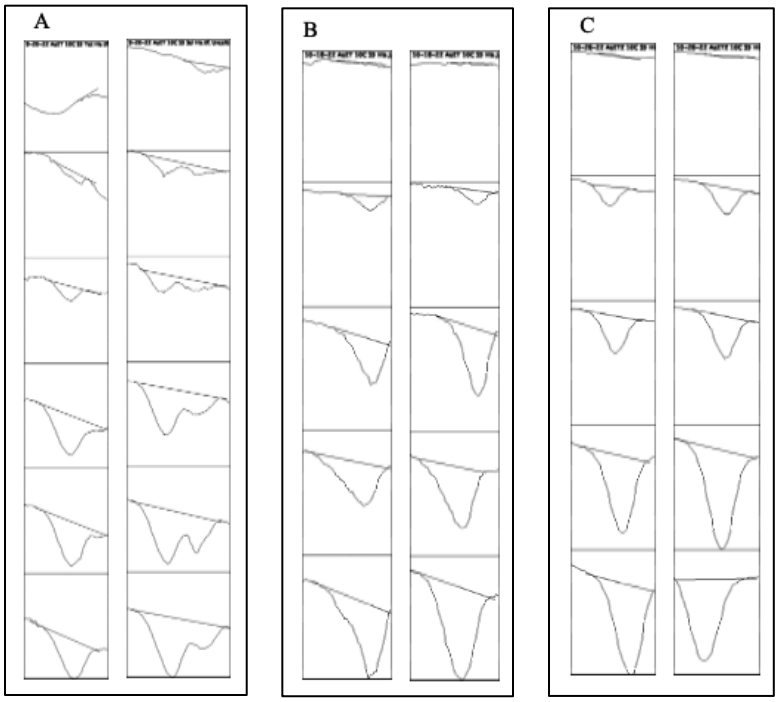


Figure 31. ImageJ plots generated from gel band selections of AaET-WT-NL recombinant protease expression grown at 10°C in TB media and induced with 0.025 mM IPTG and 1 mM Betaine. Plots that were generated by ImageJ to help measure the intensity (pixels) of each band (Figure 25 gels used for this analysis). The area within each peak was used to produce a numerical value (pixel intensity) and then used to determine statistical differences. Each peak represents a time point in which samples were collected, starting at 0 hours (post induction) going to 48 hours (post induction). A. First set of total (left) and soluble (right) samples plotted. B. Repeat experiment set of total (left) and soluble (right) samples plotted. C. Third triplicate set of total (left) and soluble (right) samples plotted.

Total and soluble measurements were compared with each other to determine if the samples were statistically different from one another using the student's T-test (Figure 32-34). The samples induced with 0.00625 mM IPTG show that there are no significant differences in expression between the total and soluble samples (Figure 32). However, it can

be noted that soluble expression seems to exhibit more intense bands than the total expression bands at 44- and 48-hours post induction as observed in figure 32A and at 24- and 44-hours post induction in figure 32B. Samples induced with 0.0125 mM IPTG show that there are differences in expression between total and soluble samples (Figure 33). For the most part, total expression measurements show higher intensity values when compared to soluble expression measurements. Interestingly, soluble expression intensity is much higher than the total expression intensity at 48 hours post induction as observed in figure 33A. Again, it should be noted that the soluble samples led to more intense bands than the total expression bands. This could be due to limitations in ImageJ with the box selection described in Chapter 2 Section 3, and experimental limitations in the number of total replicates (this will fully be discussed in the discussion, see Chapter 4 Section 1). At 44- and 48-hours post induction (Figure 33B), total expression intensity measurements exhibit a greater number of pixels when compared to the soluble expression intensity. The samples induced with 0.025 mM IPTG show that there are no statistical differences in expression between the total and soluble samples (Figure 34), which indicate that all protein produced is solubly expressed. It should be noted that at 44- and 48-hours post induction in figure 34A and at 44-hours post induction in figure 34B, soluble samples exhibit more intense bands than the total expression bands because the pixel measurements of the soluble samples obtained from ImageJ are greater than total sample pixel measurements. Unfortunately, statistical analysis of all growth experiments seems to have no significant statistical differences due to limitations in the number of total replicates (this will be fully discussed in the discussion, see Chapter 4 Section 1).

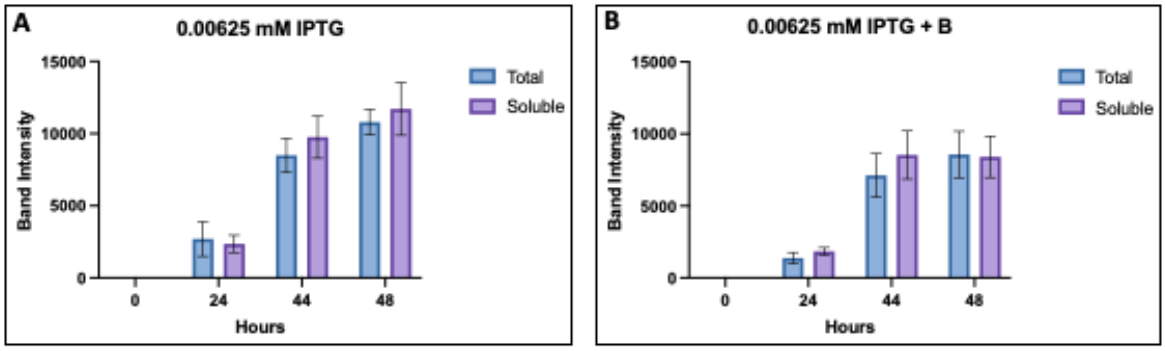


Figure 32. Graphical comparison of total and soluble measurements of AaET-WT-NL protease expression grown at 10°C using the student's T-test (GraphPad Prism). The means of the pixel intensity measurements were used to create a graphical representation of the band intensities at each timepoint collected during the growth experiment (0-, 24-, 44-, and 48-hours). Cells were induced with either A. 0.00625 mM IPTG or B. 0.00625 mM IPTG in the presence of 1 mM Betaine.

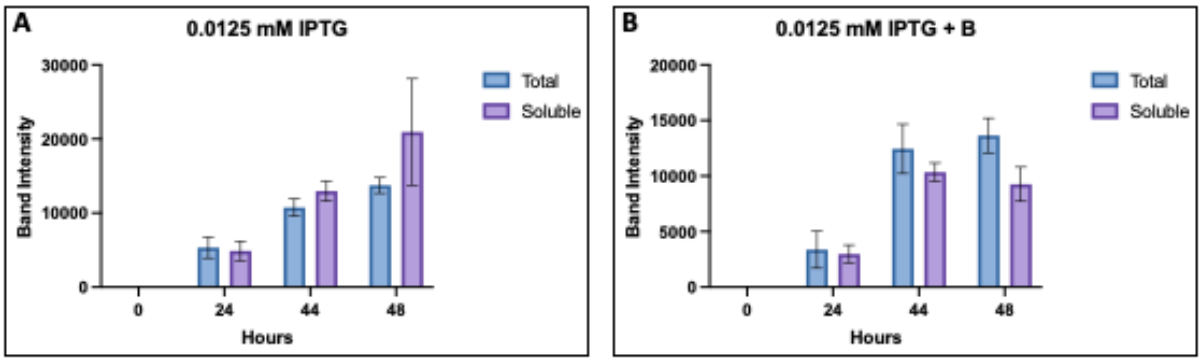


Figure 33. Graphical comparison of total and soluble measurements of AaET-WT-NL protease expression grown at 10°C using the student's T-test (GraphPad Prism). The means of the pixel intensity measurements were used to create a graphical representation of the band intensities at each timepoint collected during the growth experiment (0-, 24-, 44-, and 48-hours). Cells were induced with either A. 0.0125 mM IPTG or B. 0.0125 mM IPTG in the presence of 1 mM Betaine.

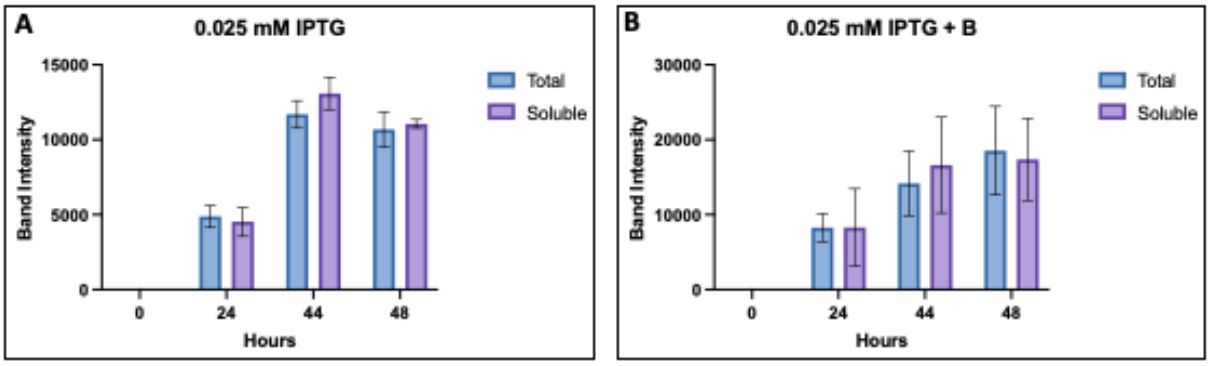


Figure 34. Graphical comparison of total and soluble measurements of AaET-WT-NL protease expression grown at 10°C using the student's T-test (GraphPad Prism). The means of the pixel intensity measurements were used to create a graphical representation of the band intensities at each timepoint collected during the growth experiment (0-, 24-, 44-, and 48-hours). Cells were induced with either A. 0.025 mM IPTG or B. 0.025 mM IPTG in the presence of 1 mM Betaine.

The three different concentrations of IPTG (0.025 mM, 0.0125 mM, and 0.00625 mM) at 10°C were then compared to determine if samples show any statistical significance between one another using ANOVA (Figure 35). Since the goal is to solubly express and maintain the inactive form of Early Trypsin only soluble expression intensity measurements were utilized to compare IPTG concentrations. In samples that did not contain betaine, a linear trend in which soluble expression increases over time is evident for samples induced with 0.0125 mM IPTG (Figure 35A). Samples induced with 0.0125 mM IPTG produced much higher amounts of soluble expression, especially at 48-hours post induction (Figure 35A). For samples induced with 0.00625 mM or 0.025 mM IPTG, soluble expression reaches its peak at around 44-hours post induction then expression decreases by 48-hours post induction (Figure 35A). In samples that did contain betaine, a linear trend in which soluble expression increases over time, also increases when inducer concentration increases, as seen in Figure 35B. Samples induced with 0.025 mM IPTG and 1 mM betaine produced much higher amounts of soluble

expression at every time point analyzed when compared to the other conditions that contained betaine (Figure 35B).

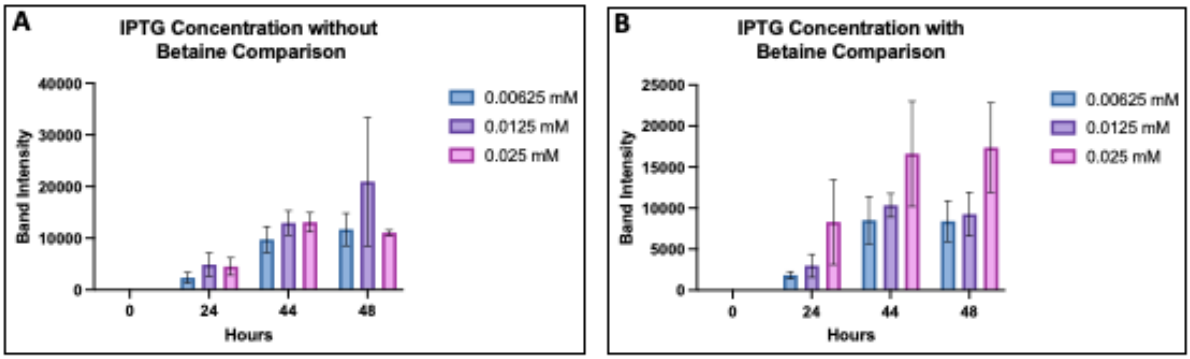


Figure 35. Graphical comparison of the different IPTG concentration (0.025 mM, 0.0125 mM, and 0.00625 mM) measurements of AaET-WT-NL protease expression grown at 10°C using an ANOVA test (GraphPad Prism). The means of the pixel intensity measurements were used to create a graphical representation of the band intensities at each timepoint collected during the growth experiment (0-, 24-, 44-, and 48-hours). A. Comparison of samples only induced with IPTG (0.025 mM, 0.0125 mM, and 0.00625 mM). B. Comparison of samples induced with IPTG (0.025 mM, 0.0125 mM, and 0.00625 mM) in the presence of 1 mM Betaine.

In general, the samples that did not contain betaine produced more soluble expression than samples that did not contain betaine, apart from 0.025 mM in the presence of 1 mM betaine (Figure 36). In samples that did not contain betaine, cells that were induced with 0.0125 mM IPTG produced the most soluble expression and cells induced with 0.00625 mM IPTG produced the least amount of soluble expression (Figure 36). In samples that did contain betaine, cells that were induced with 0.025 mM IPTG and 1 mM betaine produced the most soluble expression while cells induced with 0.00625 mM IPTG and 1 mM betaine produced the least amount of soluble expression (Figure 36). Samples induced with 0.0125 mM IPTG and 0.025 mM IPTG with 1 mM betaine show similar amounts of soluble

expression, however, the conditions tested at 10°C with 0.025 mM IPTG and 1 mM betaine produced the most amount of soluble expression overall.

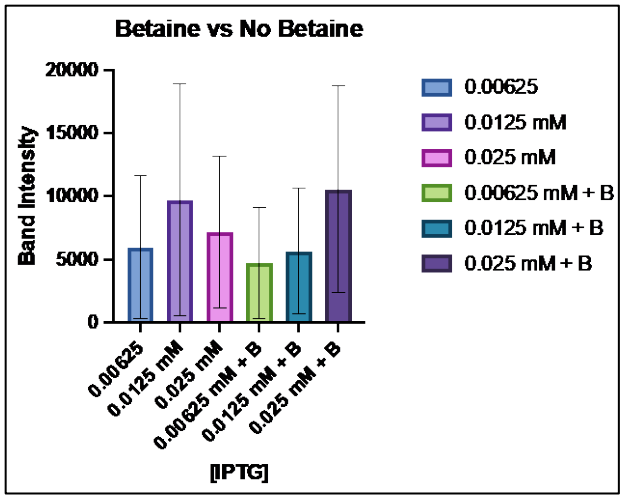


Figure 36. Graphical representation of the different IPTG concentrations (0.025 mM, 0.0125 mM, and 0.00625 mM) measurements and IPTG concentration measurements in the presence of 1 mM Betaine of AaET-WT-NL protease expression grown at 10°C using an ANOVA test (GraphPad Prism). The means of the pixel intensity measurements were used to create the graphical representation of the band intensities.

Chapter 4

Discussion and Future Directions

4.1 Discussion

While the blood feeding behavior of the female *Aedes aegypti* mosquito is essential for the completion of the gonotrophic cycle, this allows for the transmission of the Zika, Chikungunya, Dengue, and Yellow Fever blood-borne viruses.[1,5] Since many vector control strategies are ineffective or unreliable, new vector control strategies need to be explored to prevent the spread of these viral pathogens.[70] *Ae. aegypti* midgut digestive enzymes play an essential role in the gonotrophic cycle. These midgut enzymes will breakdown the blood meal to provide nutrients needed for the egg laying process.[5] Inhibiting the midgut proteases could potentially inhibit the production of amino acids and oligopeptides needed for egg production. While each midgut protease may play a role in digestion, it is still unknown how these proteases digest blood meal proteins individually. Thus, it is important to gain an understanding on the individual functions of each midgut enzyme to be able to better develop a much more efficient vector control strategy that inhibits these specific midgut proteases. Therefore, the focus of this work is specifically on the *Aedes aegypti* Early Trypsin (AaET) protease. AaET is a female specific protease that is expressed in the first 12-hours after ingestion of a blood meal.[5,27] Initial attempts at recombinantly expressing the native AaET zymogen form led to auto-activation and difficulties in purifying the enzyme.[33] In order to obtain an abundant amount of the inactive form of AaET, the recombinant expression protocol described in Nguyen *et al.* was optimized for this study.[33] Factors such as

temperature, inducer concentration and culture additives were modified to influence the rate of protein synthesis to aid in proper protein folding.

Growth experiments were set in Terrific Broth (TB) nutritious media and incubated at 30°C to ensure the log phase reached an optical density (OD_{600nm}) of about 0.6, and once reached, cells were then induced with either 0.025 mM, 0.0125 mM, or 0.00625 mM IPTG and the addition or absence of 0.1 mM betaine. After induction, temperature settings varied, based on the experiment, and cells were incubated at either 30°C, 23°C, or 10°C. Ultimately, the goal of these experiments was to determine the best bacterial growth condition to produce the recombinant and inactive form of AaET in abundance. In the literature, it is recommended to induce cells with 0.1 mM IPTG, which was thought to be optimal for expressing soluble *Aedes aegypti* midgut proteases.[33,50] The synthetic IPTG inducer is a mimic of allolactose and does not get metabolized, so levels of IPTG will remain constant and the lac repressor protein will bind to the IPTG molecule, inducing a conformational change in the protein structure that will reduce binding affinity for DNA, specifically the operator region of the lac operon.[49,71] Using higher levels of the inducer will prevent the lac repressor from binding at the operator region which will then allow RNA polymerase to bind to and transcribe the lac operon.[48] However, if lower levels of the inducer are used then there will not be enough of the inducer to inhibit the lac repressor from binding, thus preventing RNA polymerase from binding to the lac operator, and inhibiting enzyme production.[48] It is important to note that the *lacI* gene in bacteria is constitutive, so production of the lac repressor protein is always on to ensure tight regulation of transcription and translation.[48] Thus, the levels of IPTG present can affect gene transcription, and by decreasing the inducer

concentration, rate of transcription should also decrease which should allow more time for the protease to fold properly. However, as noted, too low may inhibit gene transcription, so care needs to be taken to determine the proper lower levels of IPTG that can be used.

Additionally, since the protein of interest is expressed in a prokaryotic environment (T7 SHuffle *E. coli* competent cells) the rate of synthesis and protein folding can be about 10 times faster than in a eukaryotic environment.[72,73] Unfortunately, the consequences of expressing eukaryotic genes in a prokaryotic environment is that the rate of protein synthesis is occurring so much faster that the eukaryotic proteins (especially those dependent on disulfide bonds for protein structure and stability) cannot fold properly.[54,73] Thus, at optimal temperatures in bacteria (30°C-37°C), transcription and translation function simultaneously, exposing the hydrophobic regions of the newly formed protein, which leads to improper protein folding and aggregation, and by decreasing the temperature the rate of synthesis will be much slower and can decrease the formation of inclusion bodies by allowing more time for the protein to fold properly.[49,54] This can be seen in the results presented in 3.1, samples induced at 30°C only produced insoluble protein (Figures 8-13), samples induced at 23°C ultimately produced soluble expression of the active mature enzyme (Figures 14-19), and samples induced at 10°C produced soluble recombinant expression of the inactive form of Early Trypsin (Figures 20-25). The levels of soluble recombinant expression varied with concentration of the inducer (Figures32-36). For example, the expression levels in samples induced with 0.00625 mM IPTG are much lower than samples induced with either 0.0125 mM or 0.025 mM IPTG (Figure 35 and 36). This may be due to 0.00625 mM IPTG being too low of a concentration, and at such a low concentration, that there is not enough of the

inducer to inhibit the lac repressor protein and initiate transcription within all the cells. However, this could be improved by increasing incubation time to allow for production of more protein over time. At lower temperatures (lower than the optimal bacterial growth temperature, 37°C or 30°C), the rate of protein synthesis and metabolism is much slower, due to a decrease in the kinetic energy of protein molecules.[74] Studies have found that decreased temperatures greatly decreases substrate affinity, and more substrate is needed to achieve increased rates of synthesis and metabolism at lower than optimum temperatures.[75] And thus, as temperature decreases, enzyme binding collisions decrease which hinder binding of RNA polymerase to the promoter region.[74,75] In addition, the low temperature affects other important enzymes and proteins involved in bacterial metabolism. Production of these enzymes is therefore slowed, using less resources, so there is still sufficient amount of nutrients in the media available to allow for the bacteria to continue to grow longer than 24 hours.[54,76] This allows for the log phase to be extended beyond this time point, which will allow for the production of more protein over time.[50] Even more interesting is that samples induced with 0.0625 mM IPTG produced intense soluble expression bands when compared to samples induced with 0.00625 mM IPTG in the presence of 1 mM betaine, which produced lower protein amounts leading to much more dim bands (Figures 24, 25, and 32). Betaine is a small polar molecule osmolyte that acts as a chaperone and is thought to improve thermodynamic stability of folded proteins by interacting with the protein backbone, while also maintaining osmotic pressure in the bacterial cell.[37] Since betaine is a polar molecule, betaine can prevent hydrophobic interactions from occurring as the protein of interest is folded and being expressed.[37] Furthermore, osmotic pressure within the cell can change,

especially when a cell's solute concentration exceeds that of the surrounding environment (changes in osmolality).[37,77,78] During recombinant protein production, the osmolality of the bacterial cells changes due to the release of water, and so to help balance changes in osmolality, the polar osmolyte will prevent osmotic stress by driving a water uptake gradient to offset changes in the cell.[37,77–79] Osmolality levels determine how many 'particles' are in a solvent and increased osmolality levels indicate that there are high amounts of a particle in a solvent.[79] Interestingly, previous studies state that increased osmolality levels in the cell producing recombinant protein result in decreased levels of expression.[80] In the case of 0.00625 mM IPTG induction, AaET expression levels were so low that no osmolality changes within the cells were observed, and so with the addition of 1 mM betaine, the cells may have been oversaturated with the osmolyte ultimately resulting in lower amounts of recombinant soluble expression.[80] Therefore, in this case, it may be that the addition of betaine may not be needed for cells induced with very low concentrations of IPTG. Unfortunately, fundamental research on betaine has not been fully studied, limiting our ability to understand the osmolyte to its full capacity.

To analyze the difference in expression levels between samples, the ImageJ program was used to measure gel band intensities of both total and soluble expression samples.[66] ImageJ can measure gel band intensities by calculating the area and pixel value statistics of a given selection.[66] The pixel analysis will provide plots that will give a numerical value for each band intensity that was then used for statistical analysis tests.[66] However, since all conditions tested at 30°C led to insoluble protein (Figures 8-13), there was no way to obtain soluble gel band measurements let alone compare total and soluble measurements. Thus, no ImageJ band

intensity measurements were performed for any conditions tested at 30°C, especially since the goal of this project is to produce the most abundant soluble expression of inactive AaET. Similarly, since all conditions tested at 23°C ultimately led to possible activation of the mature enzyme (Figures 14-19), due to the inactive form of the protease disappearing over time, it was difficult to measure band intensity of the inactive form of AaET, and thus, band intensity measurements were not obtained for any conditions tested at 23°C. Fortunately, all conditions tested at 10°C produced and maintained the inactive form of the soluble recombinant protease (Figures 20-25). Thus, only samples that were tested at 10°C were analyzed for band intensity measurements. Interestingly, soluble expression bands were more intense than total expression bands at multiple conditions tested at 10°C (Figures 32, 33A, and 34). This is most likely due to limitations with the ImageJ box selection described in Chapter 2 Section 3, and experimental limitations in the number of total replicates. For example, the limitations came from the ImageJ box selection tool, each box selection must be the same size and aligned horizontally with all other boxes within a selection to perform pixel detection properly.[66] Ensuring each protein band was selected and isolated properly within each box, and without any other protein background signals interfering with the analysis process, was difficult. This is due to multiple gels looking 'streaky' (Figures 20-25), meaning that the loaded samples may have been too concentrated causing the protein in the samples to not separate properly leading to the 'streaky' look. This can be prevented by diluting the protein sample with more of the 20 mM Tris-HCl pH 7.2 buffer as described in Chapter 2 Section 3.

Furthermore, many GraphPad Prism analysis plots show a clear linear trend in which soluble expression increases over time (Figure 35A) and soluble expression increases as inducer concentration increases (Figure 35B). Unfortunately, the data revealed that there were no significant statistical differences due to the lack in number of total replicates. The amount of variance was too little, resulting in high standard deviation values that ultimately led to insignificant statistical data.[81] The literature states that the greater the sample size the greater the statistical value, and a low sample size will ultimately lead to insignificant data and a lower statistical value.[81] And since there were only three replicates at each condition tested at the different temperatures, as described in Chapter 3, each statistical test ultimately led to the conclusion that the data is insignificant (Figures 32-36). Thus, a total of ten replicates should be conducted at minimum to obtain more variance data that will lead to better standard deviation results, and ultimately, match the linear trend leading to statistical significance between the samples.[81]

4.2 Future Directions

Based on these studies, the condition that produced the best recombinant soluble expression of the inactive form of AaET is induction at 10°C with 0.025 mM IPTG in the presence of 1 mM betaine (Figure 36). Since the statistical analysis data obtained was not statistically significant, further statistical studies need to be conducted to further validate the induction results at 10°C with 0.025 mM IPTG in the presence of 1 mM betaine. This can be accomplished simply by conducting more replicates of the growth experiments at 10°C (see Chapter 4 Section 1). As indicated above, seven more replicates (ten total) would be ideal. Additionally, a new project could focus on recombinant protein expression folding and how

betaine aids in this process. A better understanding of the osmolyte may be needed to better understand its role, and how it may benefit recombinant bacterial expression of midgut proteases and other enzymes from the *Aedes aegypti* mosquito.

Although determining the statistical significance of the observed linear trends (Figures 32-36), is important to produce a high impact manuscript, a large-scale growth of the best conditions (induction at 10°C with 0.025 mM IPTG in the presence of 1 mM betaine) should be performed to purify and isolate the zymogen form of AaET. Moving to 3 to 6 L growths should provide enough protein for activation and enzymatic assays. As stated in Chapter 2 Section 1, cloning the gene into the pET28a vector adds an N-terminal his_6-tag to the protease which aids with purification and isolation when using a Nickel column.[33,42] Once purified, activation and enzymatic activity assays will help determine enzymatic parameters of active AaET and compared to that published in the literature.[27] Overall, producing an abundance of recombinant AaET protease will facilitate biochemical studies needed to gain information regarding protease function in blood meal digestion, and help investigate its role in viral pathogen interaction or vector control.

REFERENCES

(1) Powell, J. R. Mosquito–Borne Human Viral Diseases: Why Aedes Aegypti? *Am. J. Trop. Med.* [Online] **2018**, *98*, 1563–1565. https://doi.org/10.4269/ajtmh.17-0866.

(2) Foster, W. A.; Walker, E. D. Chapter 15-Mosquitoes (Culicidae). In *Medical and Veterinary Entomology,* 3rd ed.; Mullen, G. R., Durden, L. A., Eds.; Academic Press: Cambridge, MA, 2019; pp 261–325.

(3) Christophers, R. *Aedes Aegypti (L.), The Yellow Fever Mosquito.*; The Syndics of the Cambridge University Press: Cambridge, UK, 1960.

(4) Wahid, I.; Sunahara, T.; Mogi, M. Mouthparts of Male Aedes (Stegomyia) Mosquitoes. *J. Am. Mosq. Control Assoc.* [Online] **2002**, *18*, 73–80.

(5) Isoe, J.; Rascón, A. A.; Kunz, S.; Miesfeld, R. L. Molecular Genetic Analysis of Midgut Serine Proteases in Aeded Aegypti Mosquitos. *Insect Biochem. Mol. Biol.* [Online] **2009**, *39*, 903–912. https://doi.org/10.1016/j.ibmb.2009.10.008.

(6) Souza, R. S.; Virginio, F.; Riback, T. I. S.; Suesdek, L.; Barufi, J. B.; Genta, F. A. Microorganism-Based Larval Diets Affect Mosquito Development, Size and Nutritional Reserves in the Yellow Fever Mosquito Aedes Aegypti (Diptera: Culicidae). *Front Physiol.* [Online] **2019**, *10*. https://doi.org/10.3389/fphys.2019. 00152.

(7) McCracken, M. K.; Christofferson, R. C.; Grasperge, B. J.; Calvo, E.; Chisenhall, D. M.; Mores, C. N. Aedes Aegypti Salivary Protein "Aegyptin" Co-Inoculation Modulates Dengue Virus Infection in the Vertebrate Host. *Virology* [Online] **2014**, *468–470*, 133–139. https://doi.org/10.1016/j.virol.2014.07.019.

(8) Viral Hemorrhagic Fevers (VHFs). Centers for Disease Control and Prevention. https://www.cdc.gov/vhf/virus-families/flaviviridae.html (last reviewed March 23, 2022).

(9) Chen, R.; Mukhopadhyay, S.; Merits, A.; Bolling, B.; Nasar, F.; Coffey, L. L.; Powers, A.; Weaver, S. C. ICTV Virus Taxonomy Profile: Togaviridae. *J. Gen. Virol*. [Online] **2018**, *99*, 761–762. https://doi.org/10.1099/jgv.0.001072.

(10) Baxter, V. K.; Heise, M. T. Genetic Control of Alphavirus Pathogenesis. *Mamm. Genome* [Online] **2018**, *29*, 408–424. https://doi.org/10.1007/s00335-018-9776-1.

(11) Yellow Fever. Centers for Disease Control and Prevention. https://wwwnc.cdc.gov/ travel/diseases/yellow-fever#:~:text=People%20who%20do%20get%20sick,organ%20 failure%2C%20and%20sometimes%20death (last reviewed July 06, 2023).

(12) Dengue and Dengue Hemorrhagic Fever. Centers for Disease Control and Prevention. https://www.cdc.gov/dengue/resources/denguedhf-information-for-health-care-practitioners_2009.pdf (accessed January 3, 2023).

(13) Basarab, M.; Bowman, C.; Aarons, E. J.; Cropley, I. Zika Virus. *BMJ* [Online] **2016**, *352*, i1049. https://doi.org/10.1136/bmj.i1049.

(14) Petersen, L.; Jamieson, D. J.; Powers, A. M.; Honein, M. A. Zika Virus. *N. Engl. J. Med.* [Online] **2016**, *374*, 1552–1563. https://doi.org/10.1056/NEJMra1602113.

(15) Guillain-Barré Syndrome. Centers for Disease Control and Prevention. https://www.cdc.gov/campylobacter/guillain-barre.html (last reviewed June 27, 2022).

(16) Yellow Fever Virus. Centers for Disease Control and Prevention. https://www.cdc.gov/yellowfever/vaccine/index.html (last reviewed April 21, 2021).

(17) Soares, T. S.; Gonzalez, B. L. R.; Torquato, R. J. S.; Lemos, F. J. A.; Costa-da-Silva, A. L.; de Lara Capurro Guimarães, M.; Tanaka, A. S. Functional Characterization of a Serine Protease Inhibitor Modulated in the Infection of the Aedes Aegypti with Dengue Virus. *Biochimie* [Online] **2018**, *144*, 160–168. https://doi.org/10.1016/j.biochi.2017.11.005.

(18) Halstead, S. B. Dengvaxia Sensitizes Seronegatives to Vaccine Enhanced Disease Regardless of Age. *Vaccine* [Online] **2017**, *35*, 6355–6358.

(19) Gessner, B. D.; Halsey, N. Dengue Vaccine Safety Signal: Immune Enhancement, Waning Immunity, or Chance Occurrence? *Vaccine* [Online] **2017**, *35*, 3452–3456. https://doi.org/10.1016/j.vaccine.2017.05.003.

(20) Potential Range of Aedes Mosquitoes. Centers for Disease Control and Prevention. https://www.cdc.gov/mosquitoes/mosquito-control/professionals/range.html#:~:text=These%20mosquitoes%20live%20in%20tropical,than%20other%20types%20of%20mosquitoes (last reviewed June 27, 2023).

(21) Piovezan-Borges, A. C.; Valente-Neto, F.; Urbieta, G. L.; Laurence, S. G. W.; de Oliveira Roque, F. Global Trends in Research on the Effects of Climate Change on Aedes Aegypti: International Collaboration Has Increased, but Some Critical Countries Lag Behind. *Parasites Vectors* [Online] **2022**, *15*. https://doi.org/10.1186/s13071-022-05473-7.

(22) Zika Cases in the United States. Centers for Disease Control and Prevention. https://www.cdc.gov/zika/reporting/index.html (last reviewed January 13, 2023).

(23) Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control: New Edition. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/44188/9789241547871_eng.pdf?sequence=1&isAllowed=y (accessed January 5, 2023).

(24) Brackney, D. E.; Beane, J. E.; Ebel, G. D. RNAi Targeting of West Nile Virus in Mosquito Midguts Promotes Virus Diversification. *PLoS Pathog.* [Online] **2009**, *5*. https://doi.org/10.1371/journal.ppat.1000502.

(25) Dana, A. N.; Hong, Y. S.; Kern, M. K.; Hillenmeyer, M. E.; Harker, B. W.; Lobo, N. F.; Hogan, J. R.; Romans, P.; Collins, F. H. Gene Expression Patterns Associated with Blood-Feeding in the Malaria Mosquito Anopheles Gambiae. *BMC Genomics* [Online] **2005**, *6*. https://doi.org/10.1186/1471-2164-6-5.

(26) Brackney, D. E.; Isoe, J.; C 4th, B. W.; Zamora, J.; Foy, B. D.; Miesfeld, R. L.; Olsom, K. E. Expression Profiling and Comparative Analyses of Seven Midgut Serine Proteases from the Yellow Fever Mosquito, Aedes Aegypti. *J. Insect Physiol.* [Online] **2010**, *56*, 736–744. https://doi.org/10.1016/j.jinsphys.2010.01.003.

(27) 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 Biochem.* [Online] **2011**, *12*. https://doi.org/10.1186/1471-2091-12-43.

(28) Barillas-Mury, C. V.; Noriega, F. G.; Wells, M. A. Early Trypsin Activity Is Part of the Signal Transduction System That Activates Transcription of the Late Trypsin Gene in the Midgut of the Mosquito, Aedes Aegypti. *Insect Biochem. Mol. Biol.* [Online] **1995**, *25*, 241–246. https://doi.org/10.1016/0965-1748(94)00061-l.

(29) Noriega, F. G.; Pennington, J. E.; Barillas-Mury, C. V.; Wang, X. Y.; Wells, M. A. Aedes Aegypti Midgut Early Trypsin Is Post-Transcriptionally Regulated by Blood Feeding. *Insect Mol. Biol.* [Online] **1996**, *5*, 25–29. https://doi.org/10.1111/j.1365-2583.1996.tb00037.x.

(30) 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 Biochem. Mol. Biol.* [Online] **1997**, *27*, 283–289. https://doi.org/10.1016/S0965-1748 (97)00001-5.

(31) Pasternak, A.; Liu, X.; Lin, T.-Y.; Hedstrom, L. Activating a Zymogen without Proteolytic Processing: Mutation of Lys15 and Asn194 Activates Trypsinogen. *Biochem.* [Online] **1998**, *37*, 16201–16210.

(32) 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. *Microb. Cell Fact.* [Online] **2012**, *11*. https://doi.org/10.1186/1475-2859-11-56.

(33) Nguyen, J. T.; Fong, J.; Fong, D.; Fong, T.; Lucero, R. M.; Gallimore, J. M.; Burata, O. E.; Parungao, K.; Rascón, A. A. Soluble Expression of Recombinant Midgut Zymogen (Native Propeptide) Proteases from the Aedes Aegypti Mosquito Utilizing E. Coli as a Host. *BMC Biochem.* [Online] **2018**, *19*. https://doi.org/10.1186/s12858-018-0101-0.

(34) Bhatwa, A.; Wang, W.; Hassab, Y. I.; Abraham, N.; Li, X.-Z.; Zhou, T. Challenges Associated with the Formation of Recombinant Protein Inclusion Bodies in Escherichia Coli and Strategies to Address Them for Industrial Applications. *Front. Bioeng. Biotech.* [Online] **2021**, *9*. https://doi.org/10.3389/fbioe.2021.630551.

(35) Abita, J. P.; Delaage, M.; Lazdunski, M. The Mechanism of Activation of Trypsinogen. The Role of the Four N-Terminal Aspartyl Residues. *Eur. J. Biochem.* [Online] **1969**, *8*, 314–324. https://doi.org/10.1111/j.1432-1033.1969.tb00530.x.

(36) Kumar, R. Role of Naturally Occurring Osmolytes in Protein Folding and Stability. *Arch Biochem. Biophys.* [Online] **2009**, *491*, 1–6. https://doi.org/10.1016/j.abb.2009.09.007.

(37) Oganesyan, N.; Ankoudinova, I.; Kim, S.-H.; Kim, R. Effect of Osmotic Stress and Heat Shock in Recombinant Protein Overexpression and Crystallization. *Protein Expr. Purif.* [Online] **2008**, *52*, 280–285.

(38) Arnér, E. S.; Holmgren, A. Physiological Functions of Thioredoxin and Thioredoxin Reductase. *Eur. J. Biochem.* **2000**, *267*, 6102–6109. https://doi.org/10.1046/j.1432-1327.2000.01701.x.

(39) 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 Biochem.* [Online] **1991**, *21*, 197–203.

(40) Erlanger, B. F.; Kokowsky, N.; Cohen, W. The Preparation and Properties of Two New Chromogenic Substrates of Trypsin. *Arch. Biochem. Biophys.* [Online] **1961**, *95*, 271–278. https://doi.org/10.1016/0003-9861(61)90145-x.

(41) Mierendorf, R. C.; Morris, B. B.; Hammer, B.; Novy, R. E. Expression and Purification of Recombinant Proteins Using the PET System. *Methods Mol. Med.* [Online] **1998**, *13*, 257–292. https://doi.org/10.1385/0-89603-485-2:257.

(42) Choi, T.-J.; Geletu, T. T. High Level Expression and Purification of Recombinant Flounder Growth Hormone in E. Coli. *J. Genet. Eng. Biotechnol.* [Online] **2018**, *16*, 347–355. https://doi.org/10.1016/j.jgeb.2018.03.006.

(43) Bacterial Expression Systems (E. Coli / Bacillus). Creative Biomart Inc. https://www.creativebiomart.net/bacterial-expression-systems.htm (accessed May 25, 2023).

(44) Kacena, M. A.; Merrell, G. A.; Manfredi, B.; Smith, E. E.; Klaus, D. M.; Todd, P. Bacterial Growth in Space Flight: Logistic Growth Curve Parameters for Escherichia Coli and Bacillus Subtilis. *Appl. Microbiol. Biotechnol.* [Online] **1999**, *51*, 229–234. https://doi.org/10.1007/s002530051386.

(45) Kadner, R. J.; Rogers, K. Bacteria. *Encyclopedia Britannica*; Encyclopædia Britannica, Inc.: Edinburgh, 2023.

(46) Kram, K. E.; Finkel, S. E. Rich Medium Composition Affects Escherichia Coli Survival, Glycation, and Mutation Frequency during Long-Term Batch Culture. *Appl. Environ. Microbiol.* [Online] **2015**, *81*, 4442–4450. https://doi.org/10.1128/AEM.00722-15.

(47) Tartoff, K. D. Improved Media for Growing Plasmid and Cosmid Clones. *Bethesda Res. Lab Focus* [Online] **1987**, *9*, 12.

(48) Clark, D. P.; Pazdernik, N. J. Chapter 2 - DNA, RNA, and Protein. Biotechnology (2nd ed.); Elsevier: Amsterdam, 2016; pp 33–61.

(49) Rizkia, P. R.; Silaban, S.; Hasan, K.; Kamara, D. S.; Subroto, T.; Soemitro, S.; Maksum, I. P. Effect of Isopropyl-β-D-Thiogalactopyranoside Concentration on Prethrombin-2 Recombinan Gene Expression in Escherichia Coli ER2566. *Procedia Chem.* [Online] **2015**, *17*, 118–124. https://doi.org/10.1016/j.proche.2015.12.121.

(50) Mühlmann, M.; Forsten, E.; Noack, S.; Büchs, J. Optimizing Recombinant Protein Expression via Automated Induction Profiling in Microtiter Plates at Different Temperatures. *Microb. Cell Fact.* [Online] **2017**, *16*. https://doi.org/10.1186/s12934-017-0832-4.

(51) Blackwell, J. R.; Horgan, R. A Novel Strategy for Production of a Highly Expressed Recombinant Protein in an Active Form. *FEBS Lett.* [Online] **1991**, *295*, 10–12. https://doi.org/ 10.1016/0014-5793(91)81372-f.

(52) de Marco, A.; Vigh, L.; Diamant, S.; Goloubinoff, P. Native Folding of Aggregation-Prone Recombinant Proteins in Escherichia Coli by Osmolytes, Plasmid- or Benzyl Alcohol-Overexpressed Molecular Chaperones. *Cell Stress Chaperones* [Online] **2005**, *10*, 329–339. https://doi.org/10.1379/csc-139r.1.

(53) Diamant, S.; Eliahu, N.; Rosenthal, D.; Goloubinoff, P. Chemical Chaperones Regulate Molecular Chaperones in Vitro and in Cells under Combined Salt and Heat Stresses. *JBC* [Online] **2001**, *276*, 39586–39591. https://doi.org/10.1074/jbc.M103081200.

(54) Ralston, A. Simultaneous Gene Transcription and Translation in Bacteria. *Nature Education* [Online] **2008**, *1*, 4.

(55) Rajan, R. S.; Illing, M. E.; Bence, N. F.; Kopito, R. R. Specificity in Intracellular Protein Aggregation and Inclusion Body Formation. *PNAS* [Online] **2001**, *19*, 13060–13065. https://doi.org/10.1073/pnas.1814797.

(56) Kaur, J.; Kumar, A.; Kaur, J. Strategies for Optimization of Heterologous Protein Expression in E. Coli: Roadblocks and Reinforcements. *Int. J. Biol. Macromol.* [Online] **2018**, *106*, 803–822.

(57) Song, J. M.; An, Y. J.; Kang, M. H.; Lee, Y.-H.; Cha, S.-S. Cultivation at 6–10 °C is an Effective Strategy to Overcome the Insolubility of Recombinant Proteins in Escherichia Coli. [Online] **2012**, *82*, 297–301. https://doi.org/10.1016/j.pep.2012.01.020.

(58) Borthwick, K. A. J.; Coakley, W. T.; McDonnell, M. B.; Nowotny, H.; Benes, E. Development of a Novel Compact Sonicator for Cell Disruption. *J. Microbiol. Methods* [Online] **2005**, *60*, 207–216. https://doi.org/10.1016/j.mimet.2004.09.012.

(59) Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* [Online] **1970**, *227*, 680–685. https://doi.org/10.1038/227 680a0.

(60) Al-Tubuly, A. A. SDS-PAGE and Western Blotting. *Diagnostic and Therapeutic Antibodies* [Online] **2000**, *40*, 391–405. https://doi.org/10.1385/1-59259-076-4:391.

(61) Saraswathy, N.; Ramalingam, P. 10 - Introduction to Proteomics. In *Concepts and Techniques in Genomics and Proteomics*; Woodhead Publishing: Sawston, Cambridge, 2011; pp 147–158.

(62) Ferris, C.; Casas, M.; Lucero, M. J.; de Paz, M. V.; Jiménez-Castellanos, M. R. Synthesis and Characterization of a Novel Chitosan-N-Acetyl-Homocysteine Thiolactone Polymer Using MES Buffer. *Carbohydrate Polymers* [Online] **2014**, *111*, 125–132. https://doi.org/10.1016/ j.carbpol.2014.03.078.

(63) Bittner, C.; Grassau, B.; Frenzel, K.; Baur, X. Identification of Wheat Gliadins as an Allergen Family Related to Baker's Asthma. *J. Allergy Clin. Immunol.* [Online] **2008**, *121*, 744–749. https://doi.org/10.1016/j.jaci.2007.09.051.

(64) Hu, X.; Beeton, C. Detection of Functional Matrix Metalloproteinases by Zymography. *J. Vis. Exp.* [Online] **2010**, *45*, e2445. https://doi.org/10.3791/2445.

(65) Goldring, J. P. D. The Roles of Acetic Acid and Methanol During Fixing and Staining Proteins in an SDS-Polyacrylamide Electrophoresis Gel. *Methods Mol. Biol.* [Online] **2018**, *1853*, 15–18.

(66) Rasband, W. *ImageJ*; National Institutes of Health: Maryland, 1997.

(67) Motulsky, H. J. *Analyzing Data with GraphPad Prism*; GraphPad Software Inc.: San Diego, CA, 1999.

(68) Gerald, B. A. Brief Review of Independent, Dependent and One Sample t-Test. [Online] **2018**, *4*, 50–54. https://doi.org/10.11648/j.ijamtp.20180402.13.

(69) Kim, T. K. Understanding One-Way ANOVA Using Conceptual Figures. *Korean Journal of Anesthesiology* [Online] **2017**, *70*, 22–26.

(70) 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. *Parasit. Vectors* [Online] **2011**, *4*. https://doi.org/10.1186/1756-3305-4-79.

(71) Donovan, R. S.; Robinson, C. W.; Glick, B. R. Review: Optimizing Inducer and Culture Conditions for Expression of Foreign Proteins under the Control of the Lac Promoter. *J. Ind. Microbiol.* [Online] **1996**, *16*, 145–154.

(72) Wingfield, P. T. Overview of the Purification of Recombinant Proteins. *Curr. Protoc. Protein Sci.* [Online] **2015**, *80*, 6.1.1-6.1.35.

(73) Widmann, M.; Christen, P. Comparison of Folding Rates of Homologous Prokaryotic and Eukaryotic Proteins. *JBC* [Online] **2000**, *275*, 18619–18622.

(74) Roe, J.-H.; Burgess, R. R.; Record, M. T. Temperature Dependence of the Rate Constants of the Escherichia Coli RNA Polymerase-ΛPR Promoter Interaction: Assignment of the Kinetic Steps Corresponding to Protein Conformational Change and DNA Opening. *JMB* [Online] **1985**, *184*, 441–453.

(75) Nedwell, D. B. Effect of Low Temperature on Microbial Growth: Lowered Affinity for Substrates Limits Growth at Low Temperature. *FEMS Microbiol. Ecol.* [Online] **1999**, *30*, 101–111.

(76) Clarke, A.; Fraser, K. P. P. Why Does Metabolism Scale with Temperature? *Funct. Ecol.* [Online] **2004**, *18*, 243–251.

(77) Ghosh, U. K.; Islam, Md. N.; Siddiqui, Md. N.; Khan, Md. A. R. Understanding the Roles of Osmolytes for Acclimatizing Plants to Changing Environment: A Review of Potential Mechanism. *Plant Signal Behav.* [Online] **2021**, *16*, 1913306.

(78) Finan, J. D.; Guilak, F. The Effects of Osmotic Stress on the Structure and Function of the Cell Nucleus. *J. Cell Biochem.* [Online] **2010**, *109*, 460–467.

(79) Zhu, M. M.; Goyal, A.; Rank, D. L.; Gupta, S. K.; Boom, T. V. Effects of Elevated PCO2 and Osmolality on Growth of CHO Cells and Production of Antibody-Fusion Protein B1: A Case Study. *Biotechnol Prog.* [Online] **2008**, *21*, 70–77.

(80) Ryu, J. S.; Kim, T. K.; Chung, J. Y.; Lee, G. M. Osmoprotective Effect of Glycine Betaine on Foreign Protein Production in Hyperosmotic Recombinant Chinese Hamster Ovary Cell Cultures Differs Among Cell Lines. *Biotechnol. Bioeng.* [Online] **2000**, *70*, 167–175.

(81) Wilson, C. R.; Voorhis, V.; Morgan, B. L. Understanding Power and Rules of Thumb for Determining Sample Sizes. *TQMP* [Online] **2007**, *3*, 43–50.