Limited Proteolysis of the AAA+ Protein Rubisco Activase from *Nicotiana tabacum*

by

Alan Conrad

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2012 by the
Graduate Supervisory Committee:

Rebekka Wachter, Chair
Thomas Moore
Kevin Redding

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is widely accepted as the world's most abundant enzyme and represents the primary entry point for inorganic carbon into the biosphere. Rubisco's slow carboxylation rate of ribulose-1,5-bisphosphate (RuBP) and its susceptibility to inhibition has led some to term it the "bottle neck" of photosynthesis. In order to ensure that Rubisco remains uninhibited, plants require the catalytic chaperone Rubisco activase. Activase is a member of the AAA+ superfamily, ATPases associated with various cellular activities, and uses ATP hydrolysis as the driving force behind a conformational movement that returns activity to inhibited Rubisco active sites. A high resolution activase structure will be an essential tool for examining Rubisco/activase interactions as well as understanding the activase self-association phenomenon. Rubisco activase has long eluded crystallization, likely due to its infamous self-association (polydispersity). Therefore, a limited proteolysis approach was taken to identify soluble activase subdomains as potential crystallization targets. This process involves using proteolytic enzymes to cleave a protein into a few pieces and has previously proven successful in identifying crystallizable protein fragments. Limited proteolysis, utilizing two different proteolytic enzymes (α-chymotrypsin and trypsin), identified two tobacco activase products. The fragments that were identified appear to represent most of what is considered to be the AAA+ C-terminal all α-domain and some of the AAA+ N-terminal αβα-domain. Identified fragments were cloned using the pET151/dTOPO. The project then moved towards cloning and recombinant

protein expression in E. coli. NtA$\beta_{(248-383)}$ and NtA$\beta_{(253-354)}$ were successfully cloned, expressed, purified, and characterized through various biophysical techniques. A thermofluor assay of NtA$\beta_{(248-383)}$ revealed a melting temperature of about 30°C, indicating lower thermal stability compared with full-length activase at 43°C. Size exclusion chromatography suggested that NtA$\beta_{(248-383)}$ is monomeric. Circular dichroism was used to identify the secondary structure; a plurality of α-helices. NtA$\beta_{(248-383)}$ and NtA$\beta_{(253-354)}$ were subjected to crystallization trials.

DEDICATION

This thesis is dedicated to my loving wife Melissa Conrad, my family, and my friends. I could never have done this without your support.

ACKNOWLEDGMENTS

I would like to thank my advisor, Rebekka Wachter, for giving me the opportunity to work for her and for all of the knowledge she has shared with me over the years. I would also like to thank the members of my supervisory committee: Tom Moore and Kevin Redding. Thank you to the Wachter lab; each and every one of you have helped me with my research at one point or another and that has meant a lot to me. I would also like to give special thanks to Nathan Henderson, Jim Zook, Alison Dunkle, and Chris Madden for their endless help with research techniques and motivation over the years.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Photosynthesis

Through a process called photosynthesis, plants along with photosynthetic microorganisms utilize an electron donor, such as water, along with atmospheric carbon dioxide (CO_2) to produce carbohydrates and oxygen (O_2), using photons from sunlight as the energy source for the conversion. This highly complex natural event is what drives life on earth. Photosynthesis is undoubtedly influential in our daily lives as nearly every aspect of life as we know it can be related in some way or another to photosynthesis or its products. The products of photosynthesis are a large part of why there is human life on the planet.

$$6CO_2 + 6H_2O \xrightarrow{Light} C_6H_{12}O_6 + 6O_2$$

The oxygen product of photosynthesis is the key to aerobic life. Along with the different forms of energy that photosynthetic organisms provide for themselves, some photosynthetic organisms such as plants that produce fruits and vegetables are sources of energy not only for humans, but also a source of energy for many animals and other organisms. Photosynthesis has evolved over millions of years to what it is today. While current global food production is sufficient, there will be more people on Earth in the future, which will require more food, and an increase in global temperatures, which has been shown to decrease the capacity of photosynthesis.[1]

1.2 Rubisco

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is the most abundant protein on earth and the main carbon fixing enzyme active during photosynthesis.[2] It is responsible for adding inorganic CO_2 to the sugar phosphate, ribulose-1,5-bisphosphate (RuBP). Rubisco's catalytic reaction turns one molecule of RuBP into a product of two molecules of 3-phosphoglycerate (3PGA), which can be reformed into biomass by metabolism in the organism.

The catalytic efficiency of Rubisco is dependent upon several factors. First, Rubisco requires ordered carbamylation of its large subunit lysine 201 residue in conjunction with coordination of a Mg^{2+} ion for activation of its active site.[3] The CO_2 added to Rubisco's lysine 201 is not the same CO_2 that is added to the RuBP during catalysis, it requires CO_2 as a cofactor and a substrate. Without the carbamylated lysine residue in Rubisco's active site, RuBP can bind tightly completely halting enzymatic activity in that particular site.[4] Another factor affecting the catalytic efficiency of Rubisco is the competition of the oxygenation reaction of RuBP by Rubisco. Rubisco is capable of performing either carboxylation or oxygenation of its substrate RuBP.

$$CO_2 + RuBP \xrightarrow{Rubisco} (2)\ 3phosphoglycerate$$

$$O_2 + RuBP \xrightarrow{Rubisco} (1)3phosphoglycerate\ (1)2phosphoglycolate$$

Rubisco's carboxylation reaction is favored to oxygenation despite O_2 being 25 times more prevalent in solution than CO_2 at 25 C.[5] The carboxylation product of RuBP is 2 molecules of 3PGA. The oxygenation of RuBP effectively reduces the amount of 3PGA produced during each reaction to half. This is

because the products of the oxygenation of RuBP are 1 molecule of 3PGA and 1 molecule of 2-phosphoglycolate. Production of 2-phosphoglycolate does not require CO_2 from the atmosphere and 2-phosphoglycolate is not used by plants to create biomass. Instead, 2-phosphoglycolate must be converted into a usable molecule through a process that requires energy in the form of adenosine triphosphate (ATP).

Rubisco has been identified as the bottleneck of photosynthesis with a turnover rate of about 5 s^{-1}.[6,7] For this reason, Rubisco has long been investigated in the study of carbon fixation and crop improvement.[9]

1.3 Rubisco Activase

In the earlier work of Salvucci *et al.*[8] a protein was discovered that is required to keep the active sites of Rubisco free from inhibitors. This protein uses energy from ATP hydrolysis as the driving force behind a relative movement, which is responsible for removing tightly bound inhibitors such as 2-carboxyarabinatol-1,5-bisphosphate from the active site of Rubisco.[10,11] This protein was termed ribulose-1,5-bisphosphate carboxylase/oxygenase activase or more simply, activase.[11,12] In this sense, activase is indirectly responsible for photosynthesis regulation. Once activase was discovered to be essential to Rubisco's activity, it became important to investigate activase as a means of increasing the efficiency of carbon fixation in plants.

Amino acid sequence comparisons lead to the discovery that activase was part of a family of P-loop NTPases and belongs to the subclass of proteins called AAA+, which is short for ATPases associated with a variety of cellular

activities.[13, 14] This family of P-loop NTPases has two highly conserved sequence motifs known as the Walker A and the Walker B motifs. The Walker A motif consists of GXXXXGKT/S and it coordinates the β-phosphate on the NTP. The Walker B motif consists of four hydrophobic residues followed by an aspartic acid or glutamic acid, which coordinates the γ-phosphate on the NTP.[15] Other important features present on AAA+ proteins are slightly conserved Box regions and Sensor regions. The Box and Sensor regions are less conserved between proteins than the Walker motifs; however, the Sensor 2 region is exceptionally important because it contains two specific amino acids shown to correlate to the species specificity recognition site of activase.[16]

When this project was started, crystallization of full-length activase had been unsuccessful, likely due to its propensity for polydispersity. This is one reason that limited proteolysis was an appealing technique for the investigation of the structure of activase. It is possible that truncated variants of activase are less likely to oligomerize if they are missing the specific amino acids needed for intermolecular interactions. Limited proteolysis has been used previously to cut proteins into fragments of proteins and to identify domain structure.[17, 18] Limited proteolysis has also been used to make crystallizable fragments from full-length proteins with success by Botos *et al.*[19] They were able to adjust the limited proteolysis conditions to isolate soluble stable fragments of the *E. coli* Lon protease. Then, they separated and crystallized the fragments, obtaining crystals that were able to diffract to a resolution of 1.9 Å and ultimately solved the structure of that fragment.

At the beginning of this project there were no structures in the protein data bank (PDB) that had a high sequence homology to activase, this is not an uncommon occurrence with AAA+ proteins. For example, the AAA+ proteins HsLU and FtsH have low sequence homology; however, both proteins' structures have been solved by X-ray crystallography, revealing a very similar fold. They both have an αβα region, connected by a short linker region to an all α region. It is for this reason that a homology model of activase, based on the structure of HsLU, was suspected to be fairly accurate. Since activase is a AAA+ protein, it is thought to have a similar structure to other AAA+ proteins even if the ultimate function differs.[20] Due to the aforementioned reasons, a homology model of activase, provided by Archie Portis (USDA-ARS, U. of Illinois, Emeritus), is occasionally referred to when designing recombinant proteins.

Since the initiation of this project, two important structural papers of higher plant Rubisco activase were published. The first paper, published by *Henderson et al.*[21] was the X-ray structure of a recombinant protein fragment of activase from the creosote plant. This fragment consists of 97 amino acid residues that represent the majority of the all α-domain. The largely C-terminus structure consists of four helices, as expected, from a protein in the AAA+ superfamily; however, it contains a unique elongated helix number three. This long helix contains two amino acids that are responsible for the species specificity of activase.[16,22] Species specificity is crucial because activase from the family of non-Solanaceae plants such as spinach is not capable of activating Rubisco from the family of Solanaceae such as tobacco and vice versa. The unique long helix

structure could be an artifact from crystallization or possibly an important feature for in vivo interactions with Rubisco.

A German group also published a structural paper of higher plant Rubisco activase. *Stotz et al.*[23] built a 293 amino acid long model of activase representing the majority of the activase protein. Stotz's model has the C-terminus truncated by 23 amino acids and the N-terminus missing 67 amino acids. Their larger, yet still truncated, tobacco activase structure shows a C-terminal α domain that differs from the classical AAA+ fold. Their C-terminal domain contains a small, fifth helix about 1.5 turns in length. Stotz showed by mutagenesis that this small helix contains the amino acids responsible for species specificity.[23]

1.4 Significance

Currently, there are at least 1 billion chronically malnourished people around the world. With the world population increasing as it is, it is expected that by 2030 food production will need to be increased by 50%, and by 2050, food production will need to be twice as high as it is now.[24] This creates pressure on the agricultural community to rapidly increase their production of food goods. In order to increase the amount of food from crops, there needs to be either more crops planted and harvested or the yield of crop per plant needs to be increased. Selecting for high yielding crops by planting seeds from the most productive plants has been successful in the past[25]; however, to increase crop yield by 100% of the current yield within 40 years is a challenge that will require the help of science. The challenge is not only that there will need to be such an incredibly large increase of available food goods, but also that over the next century the

average global temperature is expected to increase 2-4 °C. In certain areas such as Canada, the temperature is expected to increase 3-10 °C due to green house gases in the atmosphere.[26] Global temperature increases can cause the entire world crops' photosynthesis to be operating in temperatures that are, more frequently, above optimal for photosynthesis. It is possible that plants will be able to acclimate to the increase in temperatures naturally; however, it is also possible that some species will not be able to adapt to the new temperatures causing widespread die offs.[26]

Since Rubisco has been identified as the bottle neck of photosynthesis, it follows that it should be investigated as a means of increasing future crop yields. In order to make Rubisco a more efficient enzyme, all of its activities will need to be investigated thoroughly. Rubisco activase interacts directly with Rubisco. At the start of this thesis project the structure of activase was unknown as well as the specific physical interactions between Rubisco and activase. Identifying the structure will allow the specific interactions to be investigated further.

CHAPTER 2

EXPERIMENTAL SECTION

2.1 Recombinant Protein ($NtA\beta_{(248-383)}$), Design Based on Homology Model

2.1.1 $NtA\beta_{(248-383)}$ Recombinant Protein Cloning

Polymerase chain reaction (PCR) primers were ordered from Integrated DNA Technologies to amplify part of a DNA sequence. The amplified sequence coded for a 136 amino acid long protein, creating an N-terminal truncated variant of *Nicotiana tabacum* (tobacco) Rubisco activase. The forward primer utilized for this experiment was 5'-CAC CTT CTA CTG GGC ACC AAC-3'. The reverse primer used was 5'-CTA ACT AGC AAA GAA GGA TCC ATT G-3'. The PCR was performed using the previously mentioned primers along with a pET23a(+) plasmid containing a gene that coded for full-length tobacco activase protein. The PCR product was confirmed to be present and purified, using agarose gel electrophoresis. The PCR product was removed from agarose gel, using a Qiagen gel extraction kit. TOPO cloning of the PCR product into the pET151/dTOPO vector was performed according to the directions of Invitrogen's Champion pET Directional TOPO Expression Kit. TOP10 cells were then transformed with the pET151/dTOPO vector using heat shock and spread onto a carbenicillin supplemented (100 mg/mL carbenicillin) agar plate and allowed to grow at 37°C for 12 hours. 6 colonies were subsequently selected and allowed to grow in 25 ml of Lysogeny Broth (LB) media. The resulting plasmids were then extracted from each of the 6 samples of cells using a Qiagen spin miniprep kit. Samples of each plasmid were taken to the ASU DNA facility to be sequenced and the

concentrations of plasmid were determined using a NanoDrop 1000 spectrophotometer.

2.1.2 NtA$\beta_{(248\text{-}383)}$ Recombinant Protein Expression

The initial step of protein expression involves transformation of BL21*(DE3) cells. For the transformation, 50 µl of BL21*(DE3) cells were put into an Eppendorf tube on ice. Then, 0.5 µl of plasmid, identified by the ASU sequencing facility to contain the gene of interest, was added to the cells, mixed by stirring with a pipette tip, and let sit on ice for 30 minutes. The cells were subsequently heat shocked for 45 seconds at 42°C, then placed on ice for 2 minutes. Next, 0.5 ml of SOC was added to the cells in the Eppendorf tube and the cells were placed into an incubator/shaker for 1 hour at 37°C with a rotation rate of 250 rpm. After 1 hour, 100 µl of cell growth was plated onto a carbenicillin supplemented (100 mg/mL carbenicillin) LB agar plate and incubated at 37°C overnight. The following day, using sterile technique, 4 of the colonies were picked and put into 4 separate culture tubes, containing 25 ml of LB supplemented with carbenicillin (100 mg/mL carbenicillin) and allowed to grow overnight in an incubator/shaker at 37°C and 250 rpm. Following this, four 1 L LB cultures were created by inoculating each 1 L of LB supplemented with carbenicillin (100 mg/mL carbenicillin) using the 4 previously created 25 ml "overnight" cultures. The 1 L cultures were allowed to grow in the incubator/shaker at 37°C and 250 rpm until an optical density (OD) of 0.6 at 600 nm was reached. The cultures were then cooled to 18°C and protein production was induced by adding isopropyl β-D-1-thiogalactopyranoside to a 1 mM final

concentration. Cultures were placed back into the incubator/shaker with the temperature set to 18°C and rotations set at 200 rpm for 16 hours.

2.1.3 NtAβ$_{(248-383)}$ Recombinant Protein Purification

Cells were harvested by centrifugation and frozen at -80°C. To purify the protein, the cell pellet was suspended in lysis buffer consisting of 50 mM HEPES pH 8.0, 0.3 M KCL,10% glycerol, 1 mM DTT, 20 mM imidazole at pH 8.0 and 0.1% Triton X-100 and sonicated. The mixture was then centrifuged and the supernatant was harvested. The supernatant was then filtered through a 0.8 μm syringe filter. The filtered supernatant was loaded onto a 5 ml total volume nickel column (GE Healthcare) and eluted fractions were collected. The first wash was with 20 mM imidazole. The protein was then eluted with a 200 mM imidazole wash. The final wash was with 500 mM imidazole to clean any residual material off the column. A quick Bradford assay of all eluent was performed to confirm the presence of protein in the fractions where the protein was expected to elute. The proper protein factions were pooled, and the histidine tag was removed using a his-tagged tobacco etch virus (TEV) protease while being dialyzed overnight in order to remove the imidazole to allow for further purification. Next, the mixture was loaded onto a 5 ml volume nickel (GE Healthcare) column. The protein eluted in the flow through and with the first 20 mM imidazole wash. The protein was then concentrated to 18.6 mg ml^{-1} using a Millipore amicon ultra, ultracel 10 kDa centrifugal concentrator. A Bradford assay was performed in triplicate to determine the concentration of protein. The protein was then flash frozen in liquid nitrogen and stored at -80°C for future use.

2.1.4 $NtA\beta_{(248-383)}$ Recombinant Protein Characterization

Size exclusion chromatography was used to investigate the polydispersity of the Recombinant Protein $NtA\beta_{(248-383)}$. Size exclusion chromatography experiments were performed using a Waters 600S controller paired with a Waters 996 detector HPLC and a Phenomonex Biosep 4000 column.

A thermal denaturation thermofluor assay was performed using SYPRO Orange as the fluorescent dye. The assay was performed in an Applied Biosystems real time PCR machine. The secondary structure of this fragment was investigated by circular dichroism (CD) using a Jasco 710 Spectropolarimeter. Data from the CD experiment was analyzed to investigate secondary structure using an unpublished deconvolution program created by James Zook from Dr. Petra Fromme's lab.

Crystallization screens were setup in a cold room at 10°C using the hanging drop vapor diffusion method.[27] Both Hampton's Index Screen and Crystal Screen Cryo were used. Each screen was setup using two different concentrations of protein—the first 18.6 mg ml^{-1} and the second 6.2 mg ml^{-1}.

2.2 Recombinant Protein ($NtA\beta_{(253-354)}$), Design Based on Crystallized Creosote Activase

2.2.1 $NtA\beta_{(253-354)}$ Recombinant Protein Cloning

Polymerase chain reaction (PCR) primers were ordered from Integrated DNA Technologies to amplify part of a DNA sequence. The amplified sequence coded for a 102 amino acid long protein, creating an N-terminal and C-terminal truncated variant of *Nicotiana tabacum* (tobacco) Rubisco activase. The forward primer utilized for this experiment was 5'-CAC CAG GGA AGA CAG AAT TGG-3'. The reverse primer used was 5'-TTA TCT CTT CAC ATT TTC TTG C-3'. The PCR was performed using the previously mentioned primers along with a pET23a(+) plasmid containing a gene that coded for full-length tobacco activase protein. The PCR product was confirmed to be present using agarose gel electrophoresis. TOPO cloning of the PCR product into the pET151/dTOPO vector was performed according to the directions of the Invitrogen's Champion pET Directional TOPO Expression Kit. TOP10 cells were then transformed with the pET151/dTOPO vector using heat shock and spread onto a carbenicillin supplemented (100 mg/mL carbenicillin) agar plate and allowed to grow at 37°C for 12 hours. Multiple colonies were subsequently selected and allowed to grow in 25 ml of LB media. The resulting plasmids were then extracted from the media using a Qiagen spin miniprep kit. Samples of each plasmid were taken to the ASU DNA facility to be sequenced and the concentrations of plasmid were determined using a NanoDrop 1000 spectrophotometer.

2.2.2 NtAβ$_{(253-354)}$ Recombinant Protein Expression

The initial step involves transformation of BL21*(DE3) cells. For the transformation, 50 µl of BL21*(DE3) cells were put into an Eppendorf tube on ice. Then, 0.5 µl of plasmid, identified by the ASU sequencing facility to contain the gene of interest, was added to the cells, mixed by stirring with a pipette tip, and let sit on ice for 30 minutes. The cells were heat shocked for 45 seconds at 42°C then placed on ice for 2 minutes. Next, 0.5 ml of SOC was added to the cells in the Eppendorf tube and the cells were placed into an incubator/shaker for 1 hour at 37°C with a rotation rate of 250 rpm. After 1 hour, 100 µl of cell growth was plated onto a carbenicillin supplemented (100 mg/mL carbenicillin) LB agar plate and incubated at 37°C overnight. The following day, using sterile technique, 4 of the colonies were picked and put into 4 separate culture tubes, containing 25 ml of LB supplemented with carbenicillin (100 mg/mL carbenicillin) and allowed to grow overnight in an incubator/shaker at 37°C and 250 rpm. Following this, four 1 L LB cultures were created by inoculating each 1 L of LB supplemented with carbenicillin (100 mg/mL carbenicillin) using the 4 previously created 25 ml "overnight" cultures. The 1 L cultures were allowed to grow in the incubator/shaker at 37°C and 250 rpm until an optical density (OD) of 0.6 at 600 nm was reached. Next, protein production was induced by adding isopropyl β-D-1-thiogalactopyranoside to a 1 mM final concentration. Cultures were placed back into the incubator/shaker with the temperature set to 30°C and rotations set at 200 rpm for 14 hours.

2.2.3 NtAβ$_{(253-354)}$ Recombinant Protein Purification

Cells were harvested by centrifugation and frozen at -80°C. To purify the protein, the cell pellet was suspended in lysis buffer consisting of 50 mM HEPES pH 8.0, 0.3 M KCL,10% glycerol, 1 mM DTT, 20 mM imidazole at pH 8.0, and 0.1% Triton X-100 and sonicated. The mixture was then centrifuged and the supernatant was harvested. The supernatant was then filtered through a 0.8 μm syringe filter. The filtered supernatant was loaded onto a 5 ml nickel column (Qiagen Ni-NTA superflow) and eluted fractions were collected. The first wash was with 50 mM imidazole. The next wash was with 100 mM imidazole where a large amount of the protein was eluted. The final wash was with 500 mM imidazole to clean any residual material off the column. A quick Bradford assay was performed to confirm the presence of protein in the fractions where the protein was expected to elute. The proper protein factions were pooled, and the histidine tag was then removed by a his-tagged tobacco etch virus (TEV) protease while being dialyzed overnight to remove the imidazole to allow for further purification. Next, the mixture was loaded onto a 5 ml nickel column (Qiagen Ni-NTA superflow). The protein eluted in the flow through and with a 20 mM imidazole wash. The protein was then concentrated and buffer exchanged into 10 mM HEPES pH 7.0, 100 mM NaCl, and 0.04% NaN_3 using a Millipore amicon ultra, ultracel 10 kDa centrifugal concentrator. Protein concentration was quantified using a NanoDrop 1000 spectrophotometer.

2.2.4 NtAβ$_{(253-354)}$ Recombinant Protein Characterization

Using a Wyatt DynaPro NanoStar, a Dynamic light scattering experiment was performed on a 0.1 μm filtered, 5 mg/ml sample of NtAβ$_{(253-354)}$.

96 well sitting-drop trays were set up using a Rigaku Phoenix-HT crystallization robot to screen NtAβ$_{(253-354)}$ with the protein concentration at 60 mg/mL. Tested commercial screens from Hampton Research included: Cryo screens 1 and 2 (96 conditions), Index screen (96 conditions) and Crystal screen (96 conditions). Upon set up, experimental trays were sealed and transferred to a 4°C cold room.

2.3 Limited Proteolysis and Design of Genetic Constructs

2.3.1 Limited Proteolysis

Limited proteolysis was performed on a 51 amino acid N-terminally truncated variant of tobacco Rubisco activase provided by Dr. Michael Salvucci from the US Department of Agriculture. The technique utilized two different proteolytic enzymes: α-chymotrypsin (Sigma Aldrich, C4129) and trypsin (Sigma Aldrich, T8003). Reaction conditions were optimized first for α-chymotrypsin, then for trypsin. The conditions that were optimized are as follows: the molar ratio of activase to proteolytic enzyme, time allowed for proteolytic enzyme activity, and the concentration of adenosine diphosphate (ADP). Conditions were optimized by observing products of the proteolysis, using 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and adjusting the conditions as necessary.

The first conditions optimized were the molar ratio of activase to proteolytic enzyme and the time allowed for proteolysis. The limited proteolysis reactions (henceforth, referred to as digestions) were performed in 0.5M HEPES pH 7.9, at 26°C, with a constant concentration of 16.8 μM activase and 18.3 μM ADP in a total volume of 1 ml. The concentration of α-chymotrypsin was varied. Experimentation commenced with separate digestions utilizing molar ratios of activase to α-chymotrypsin of 233:1, 624:1, 6,247:1, 62,475:1 and 624,750:1. These ratios were created by adding specific volumes of a 0.018 mg ml^{-1} stock solution of α-chymotrypsin. The digestions were allowed to proceed for 2 hours (120 min) with samples of the digestions taken at 0 min, 30 min, 60 min, 90 min, and 120 min. Each sample was heated to 90°C for 10 minutes prior to being frozen at -20°C. After all samples were collected, they were analyzed via a 12% SDS-PAGE.

To optimize the amount of ADP in the digestion, reactions were set up as follows: 0.5M HEPES pH 7.9 with a constant concentration of 16.8 μM activase and 7.2 x 10^{-11} M chymotrypsin. The concentration of ADP in the digestions was varied from 8.3 μM to 38.3 μM in 5 μM increments. This was done by adding specific volumes from a stock of 0.5 mM ADP.

2.3.2 Limited Proteolysis Fragment Isolation

A digestion was performed under optimized conditions (ratio of activase to α-chymotrypsin 233:1, 8.3μM ADP for 60 minutes). The digestion progression was quenched by the addition of phenylmethanesulfonylfluoride (PMSF). The digestion was prepared for analysis using a 0.8 μm syringe filter to remove any

possible large aggregation that may be present in the sample. The filtered sample was analyzed by reverse phase high performance liquid chromatography (RP-HPLC) utilizing a reverse phase C18 column (Vydac 218TP45) attached to a Waters 600S controller paired with a Waters 996 detector. The mobile phase started as 100% A [water with 0.1% (v/v) trifluoroacetic acid (TFA)] and was transitioned to 70% B (acetonitrile) at a rate of 1% per minute. Absorbance was monitored at 220 nm and 280 nm.

SDS gels to be used for electrophoretic protein transfer (blotting) were allowed to polymerize for 24 hours at room temperature to assure complete polymerization. Running buffer, used for SDS gels, to be used for blotting consisted of 1X Tris-Glycine-SDS (TGS) and 5.13 mg of sodium thioglycolate (reducing agent). Blotting was performed in a Bio-Rad tank transfer system. Transfer buffer was made using 3.03 g of Tris base and 14.4 g glycine in 200 ml of methanol with the addition of distilled, deionized water until a volume of 1 L was reached. Polyvinylidene fluoride (PVDF) membrane was immersed in 100% methanol for 60 seconds and then soaked in transfer buffer for 5 minutes to hydrate and facilitate protein transfer. Proteins were allowed to transfer for 75 minutes at 350 mA constant current. The PVDF membrane was then stained with coomassie brilliant blue R-250 in a 40% methanol solution. De-staining solution consisted of 50% methanol and 50% water.

2.3.3 Limited Proteolysis Fragment Identification

Several PVDF blot samples were taken to the ASU proteomics lab for N-terminal sequencing with several analogous samples being saved for mass spectrometry.

Mass spectrometry was performed using matrix-assisted laser desorption/ionization time of flight (MALDI TOF) on a Voyager-DE STR TOF. The samples were retained on the PVDF membrane. An aliquot of cytochrome-c was placed on the PVDF membrane next to the digest sample in order to have a baseline to correct for the thickness of the PVDF membrane. The matrix used was α-Cyano-4-hydroxycinnamic acid mixed in toluene and methanol. The measurements were obtained in linear positive mode.

2.3.4 Recombinant Protein ($NtA\beta_{(182-383)}$) Based on Limited Proteolysis

2.3.4.1 $NtA\beta_{(182-383)}$ Recombinant Protein Cloning

Polymerase chain reaction (PCR) primers were ordered from Integrated DNA Technologies to amplify part of a DNA sequence. The amplified sequence coded for a 202 amino acid long protein creating an N-terminal truncated variant of *Nicotiana tabacum* (tobacco) Rubisco activase. The forward primer utilized for this experiment was 5'-CAC CAT GGG CGG AAC TAC ACA ATA CAC-3'. The reverse primer used was 5'-CTA ACT AGC AAA GAA GGA TCC ATT GTT AAT G-3'. The PCR was performed using the previously mentioned primers along with a pET23a(+) plasmid containing a gene that coded for full-length tobacco activase protein. The PCR product was confirmed to be present using agarose gel electrophoresis. TOPO cloning of the PCR product into the

pET151/dTOPO vector was performed as according to the directions of Invitrogen's Champion pET Directional TOPO Expression Kit. TOP10 cells were then transformed with the pET151/dTOPO vector using heat shock and spread onto a carbenicillin supplemented (100 mg/mL carbenicillin) agar plate and allowed to grow at 37°C for 12 hours. Multiple colonies were subsequently selected and allowed to grow in 25 ml of LB media. The resulting plasmids were then extracted from the media using a Qiagen spin miniprep kit. Samples of each plasmid were taken to the ASU DNA facility to be sequenced and the concentrations of plasmid were determined using a NanoDrop 1000 spectrophotometer.

2.3.4.2 NtAβ$_{(182\text{-}383)}$ Recombinant Protein Expression

The initial step involves transformation of BL21*(DE3) cells. For the transformation, 50 µl of BL21*(DE3) cells were put into an Eppendorf tube on ice. Then, 0.5 µl of plasmid, identified by the ASU sequencing facility to contain the gene of interest, was added to the cells, mixed by stirring with a pipette tip, and let sit on ice for 30 minutes. The cells were heat shocked for 45 seconds at 42°C then placed on ice for 2 minutes. Next, 0.5 ml of SOC was added to the cells in the Eppendorf tube and the cells were placed into an incubator/shaker for 1 hour at 37°C with a rotation rate of 250 rpm. After 1 hour, 100 µl of cell growth was plated onto a carbenicillin supplemented (100 mg/mL carbenicillin) LB agar plate and incubated at 37°C overnight. The following day, using sterile technique, 4 of the colonies were picked and put into 4 separate culture tubes, containing 25 ml of LB supplemented with carbenicillin (100 mg/mL carbenicillin) and allowed

to grow overnight in an incubator/shaker at 37°C and 250 rpm. Following this, four 1 L LB cultures were created by inoculating each 1 L of LB supplemented with carbenicillin (100 mg/mL carbenicillin) using the 4 previously created 25 ml "overnight" cultures. The 1 L cultures were allowed to grow in the incubator/shaker at 37°C and 250 rpm until an optical density (OD) of 0.6 at 600 nm was reached. The cultures were then cooled to 25°C and protein production was induced by adding isopropyl β-D-1-thiogalactopyranoside to a 1 mM final concentration. Cultures were placed back into the incubator/shaker with the temperature set to 25°C and rotations set at 200 rpm for 16 hours.

2.3.4.3 NtAβ$_{(182\text{-}383)}$ Recombinant Protein Purification

Cells were harvested by centrifugation and frozen at -80°C. To purify the protein, the cell pellet was suspended in lysis buffer consisting of 50 mM HEPES pH 8.0, 0.3 M KCL,10% glycerol, 1 mM DTT, 20 mM imidazole at pH 8.0 and 0.1% Triton X-100 and sonicated. The mixture was then centrifuged and the supernatant was harvested. The supernatant was then filtered through a 0.8 μm syringe filter. The filtered supernatant was loaded onto a 5 ml nickel column (Qiagen Ni-NTA superflow) and eluted fractions were collected. The first wash was with 20 mM imidazole. The protein was then eluted with a 200 mM imidazole wash. A final wash was performed with 500 mM imidazole. A quick Bradford assay was performed to try to confirm the presence of protein in the fractions where the protein was expected to elute. Purification for this protein stopped here because there was no protein present in the eluent. SDS-PAGE was performed which showed very little protein expression.

This experiment was repeated with expression temperatures of 30°C and 37°C which was not found to help the expression levels.

2.3.4.4 NtAβ$_{(216-383)}$ Recombinant Protein Cloning

Polymerase chain reaction (PCR) primers were ordered from Integrated DNA Technologies to amplify part of a DNA sequence. The amplified sequence coded for a 168 amino acid long protein creating an N-terminal truncated variant of *Nicotiana tabacum* (tobacco) Rubisco activase. The forward primer utilized for this experiment was 5'-CAC CAA CAA ACA AGA GAA TGC CAG GG-3'. The reverse primer used was 5'-CTA ACT AGC AAA GAA GGA TCC ATT GTT AAT G-3'. The PCR was performed using the previously mentioned primers along with a pET23a(+) plasmid containing a gene that coded for full-length tobacco activase protein. The PCR product was confirmed to be present using agarose gel electrophoresis. TOPO cloning of the PCR product into the pET151/dTOPO vector was performed according to the directions of Invitrogen's Champion pET Directional TOPO Expression Kit. TOP10 cells were then transformed with the pET151/dTOPO vector using heat shock and spread onto a carbenicillin supplemented (100 mg/mL carbenicillin) agar plate and allowed to grow at 37°C for 12 hours. Multiple colonies were subsequently selected and allowed to grow in 25 ml of LB media. The resulting plasmids were then extracted from the media using a Qiagen spin miniprep kit. Samples of each plasmid were taken to the ASU DNA facility to be sequenced and the concentrations of plasmid were determined using a NanoDrop 1000 spectrophotometer.

2.3.4.5 NtA$\beta_{(216-383)}$ Recombinant Protein Expression

The initial step involves transformation of BL21*(DE3) cells. For the transformation, 50 µl of BL21*(DE3) cells were put into an Eppendorf tube on ice. Then, 0.5 µl of plasmid, identified by the ASU sequencing facility to contain the gene of interest, were added to the cells, mixed by stirring with a pipette tip, and let sit on ice for 30 minutes. The cells were heat shocked for 45 seconds at 42°C then placed on ice for 2 minutes. Next, 0.5 ml of SOC was added to the cells in the Eppendorf tube and the cells were placed into an incubator/shaker for 1 hour at 37°C with a rotation rate of 250 rpm. After 1 hour, 100 µl of cell growth was plated onto a carbenicillin supplemented (100 mg/mL carbenicillin) LB agar plate and incubated at 37°C overnight. The following day, using sterile technique, 4 of the colonies were picked and put into 4 separate culture tubes, containing 25 ml of LB supplemented with carbenicillin (100 mg/mL carbenicillin) and allowed to grow overnight in an incubator/shaker at 37°C and 250 rpm. Following this, four 1 L LB cultures were created by inoculating each 1 L of LB supplemented with carbenicillin (100 mg/mL carbenicillin) using the 4 previously created 25 ml "overnight" cultures. The 1 L cultures were allowed to grow in the incubator/shaker at 37°C and 250 rpm until an optical density (OD) of 0.6 at 600 nm was reached. The cultures were then cooled to 25°C and protein production was induced by adding isopropyl β-D-1-thiogalactopyranoside to a 1 mM final concentration. Cultures were placed back into the incubator/shaker with the temperature set to 25°C and rotations set at 250 rpm for 16 hours.

2.3.4.6 NtAβ$_{(216-383)}$ Recombinant Protein Purification

Cells were harvested by centrifugation and frozen at -80°C. To purify the protein, the cell pellet was suspended in lysis buffer consisting of 50 mM HEPES pH 8.0, 0.3 M KCL,10% glycerol, 1 mM DTT, 20 mM imidazole at pH 8.0 and 0.1% Triton X-100 and sonicated. The mixture was then centrifuged and the supernatant was harvested. The supernatant was then filtered through a 0.8 μm syringe filter. The filtered supernatant was loaded onto a 5 ml nickel column (Qiagen Ni-NTA superflow) and eluted fractions were collected. The first wash was with 20 mM imidazole. The protein was then eluted with a 200 mM imidazole wash. A final wash was performed with 500 mM imidazole. A quick Bradford assay was performed to try to confirm the presence of protein in the fractions where the protein was expected to elute. Purification for this protein stopped here because there was minimal protein present in the eluent. SDS-PAGE was performed which showed that the majority of the protein was in the pellet.

CHAPTER 3

RESULTS, DISCUSSION, CONCLUSION, AND FUTURE WORK

3.1 Results and Discussion

3.1.1 NtAβ$_{(248-383)}$ Fragment Based on Homology Model

NtAβ$_{(248-383)}$, a 136 amino acid long soluble protein fragment which was designed based on a homology model was successfully cloned into pET151/dTOPO. After expression 18.6 mg of protein from an 11.4 g cell pellet was obtained. The N-terminus of the recombinant protein that was created starts with the three aromatic amino acids, predicted to be near the hinge region, and ends at the full-length C-terminus. This fragment was selected to be cloned first instead of the limited proteolysis identified fragments because this fragment was predicted to be more stably folded based on the homology model than the fragments that were identified using limited proteolysis. The fragments identified by limited proteolysis are suspected to contain parts of both the αβα subdomain and the all α subdomain making them unlikely candidates to be stably folded.

Characterization of the recombinant protein involved determining its dispersity, melting temperature, and secondary structure. Its dispersity was investigated using size exclusion chromatography (SEC). Figure 1 below shows a chromatogram from full-length activase with broad, non-gaussian shaped bands. Figure 2 below shows the SEC data of the NtAβ$_{(248-383)}$ fragment. The Gaussian peak at about 26.5 minutes is the NtAβ$_{(248-383)}$ fragment. The curve from the NtAβ$_{(248-383)}$ is representative of a monodisperse protein having a single narrow peak and no trailing peak, whereas the curve from the full-length activase shows a

wide trailing peak. This wide trailing peak is actually several, smaller peaks following the first. These trailing peaks represent multiple assembly states present in the solution, hence the polydispersity. The retention time of the $NtA\beta_{(248-383)}$ is typical for a monomeric protein of that size protein (15.3 kDa).

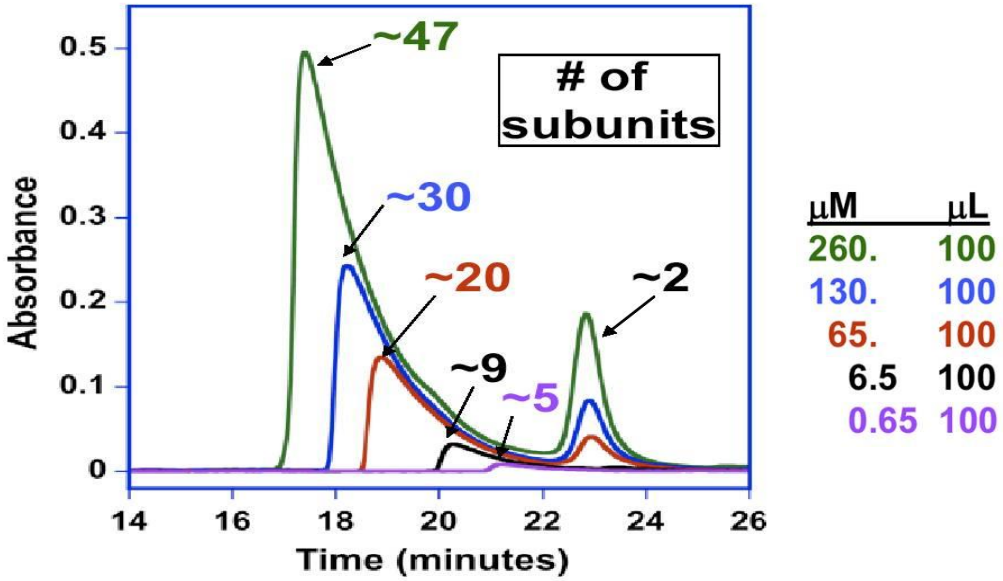


Figure 1: SEC data of full-length activase, showing that retention time changes with concentration. Different multimers of full-length activase as indicated. Chromatogram courtesy of Alison Dunkle.

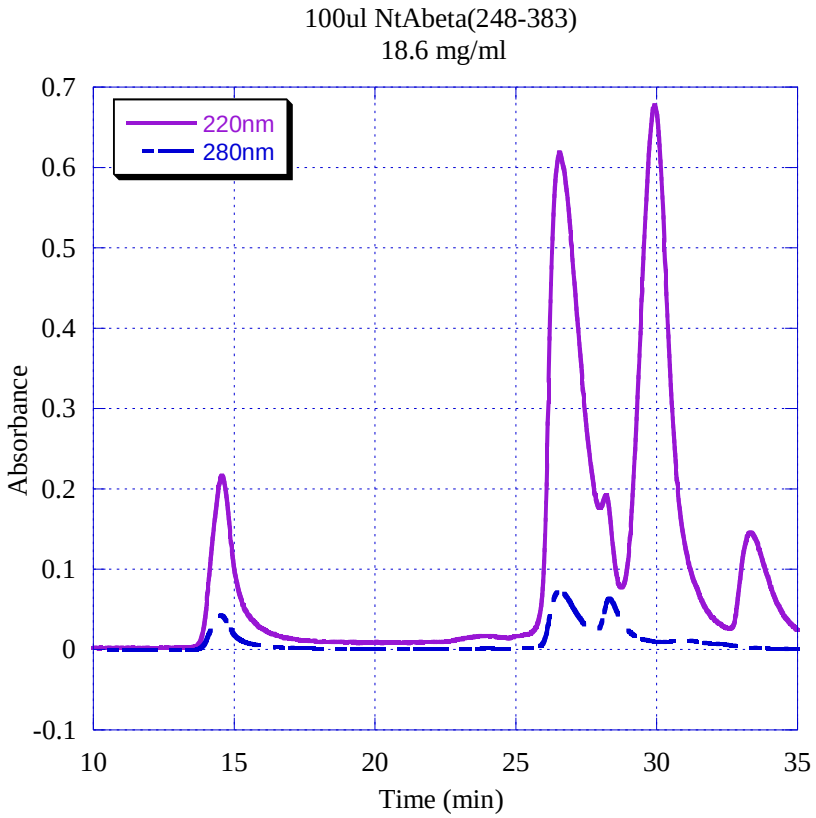


Figure 2: SEC data from NtAβ$_{(248-383)}$ showing 220 nm and 280 nm absorbance. The 280 nm gaussian peak at 26.5 minutes is representative of a NtAβ$_{(248-383)}$ fragment in a single assembly state with a 15 kDa retention time. Chromatogram courtesy of Alison Dunkle.

A thermoflour assay was performed on the NtAβ$_{(248-383)}$ fragment. Thermofluor assays use a dye that has a low fluorescence quantum yield in hydrophilic environments and high fluorescence quantum yield in hydrophobic environments such as near the hydrophobic residues of a denatured protein.[28] The thermofluor assay, as can be seen in Figure 3 below, shows the increase in fluorescence of the SYPRO Orange dye upon heat denaturation. The melting

temperature based on this assay was estimated to be around 30°C, which is relatively low compared to some of the more stable full-length activase species such as cotton or creosote activase, which melt at 39°C and 42°C respectively. In general the melting temperature is representative of the stability of the protein, the lower the melting temperature the less stable the protein. The low thermal stability is probably due to a large amount of disordered amino acids. An example of a stably folded protein is hen egg white lysozyme, which has a melting temperature of about 72°C.[29] It is possible that more thermally stable proteins could be created by cloning activase constructs of various lengths.

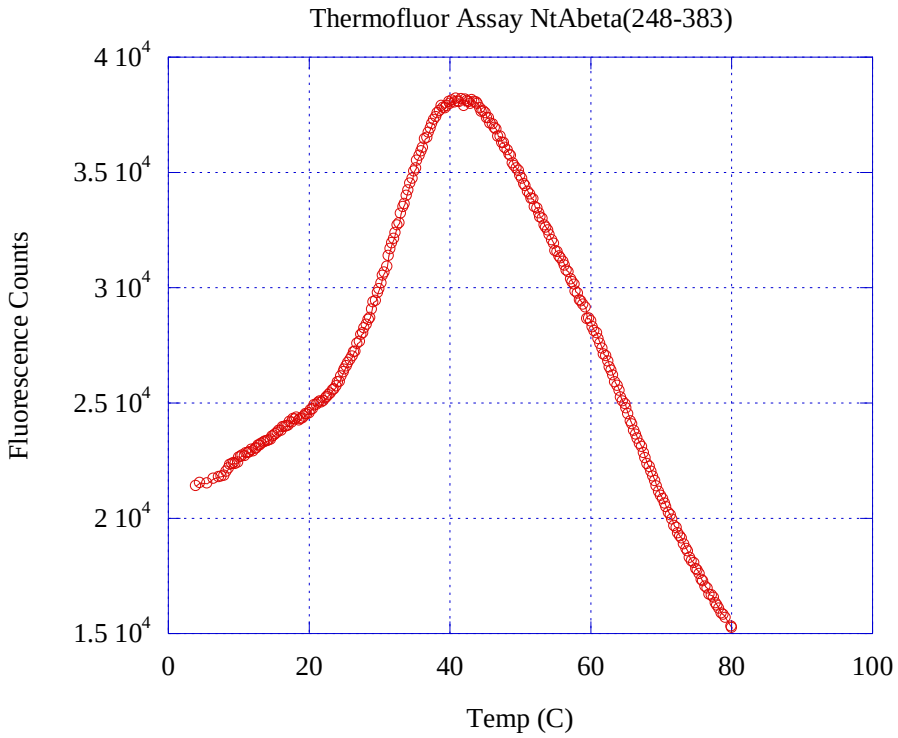


Figure 3: Thermofluor assay of NtA$\beta_{(248\text{-}383)}$ data shows changes in fluorescence as temperature of the fragment is increased.

A circular dichroism (CD) experiment was performed on the NtA$\beta_{(248-383)}$ fragment. Results can be seen below in Figure 4. The 15°C CD data was put through several deconvolution programs: CDSSTR, SELCON3, and ContnII. The computer-calculated percent composition of α-helix, β-sheet, turn, and unordered regions were averaged. The plurality was α-helix with a composition of 34% α-helix, the next highest was unordered regions at 30%, followed by turns at 19%, and lastly β-sheets at 15%. This data suggests that there is a large amount of unordered amino acids. This could mean that removing some of the N-terminal or C-terminal amino acids could result in the loss of some of the unordered amino acids and improvement of thermal stability.

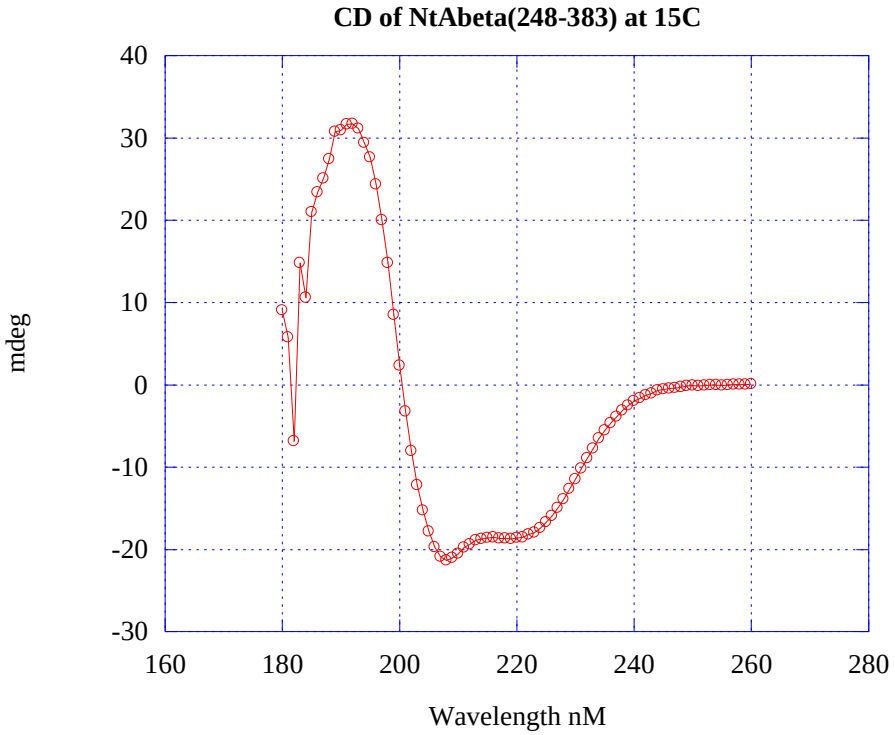


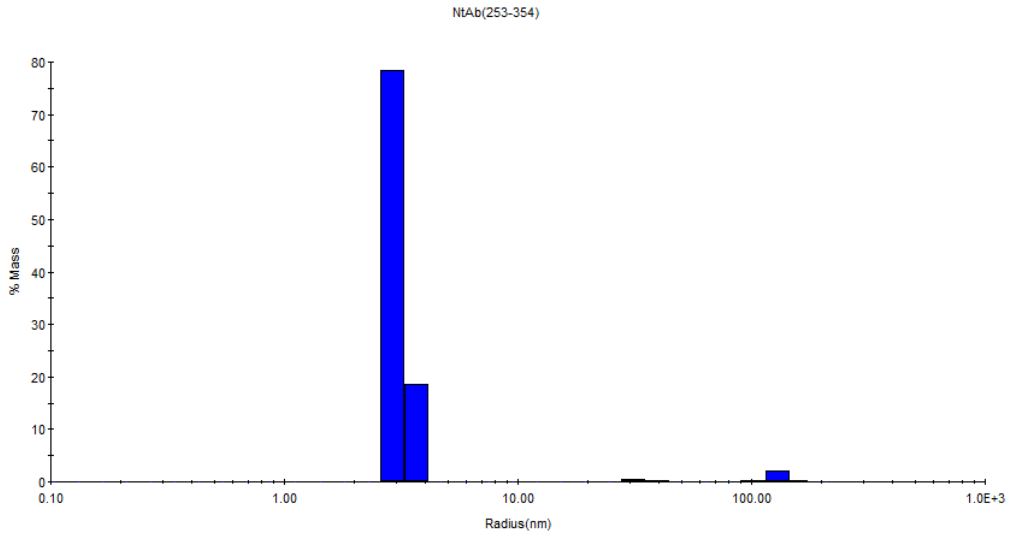
Figure 4: 15°C CD data of NtA$\beta_{(248-383)}$ shows results similar to an all α-helix secondary structure.

Two crystal screens were set up with different concentrations for the NtA$\beta_{(248-383)}$ fragment. In the first Index and Cryo crystal screens that were setup, nearly all of the protein drops precipitated instantly. That was an indication that the concentration of the protein was likely too high. For the next trial, the screens were setup using 6.2 mg ml^{-1} instead of 18.6 mg ml^{-1}. The number of drops that precipitated instantly decreased to about half. The screens were checked after 3 days, then 6 days, then 2 weeks with only signs of phase separation—no crystals. It is possible that the amount of disordered protein is too large, thus preventing

crystal growth. In order for protein crystals to form, the proteins must adopt a compact fold.

3.1.2 NtAβ$_{(253-354)}$ Based on Creosote Fragment

Dynamic light scattering (DLS) is a tool that can be used to estimate the sizes of various particles in a solution. DLS measures fluctuations of the average intensity of light being scattered by particles in solution, moving randomly through a focused laser beam. Globular proteins are assumed to be spherical in shape and that allows the use of the Stokes-Einstein relation to calculate the apparent hydrodynamic radius.[30] The apparent hydrodynamic radius was calculated by the DLS software and the results can be seen below in Figure 5. The data shows that the DLS experiment on NtAβ$_{(253-354)}$ identified 97% of the particles in solution as particles with a size of about 46 kDa, likely a tetramer; however, particles with a mass of about 307,000 kDa scattered the majority of the light, roughly 52% of it. This is interesting because the sample was filtered with a 0.1 μm filter, which means that the protein fragment must aggregate spontaneously. This experiment shows similar results to DLS experiments with Henderson's creosote variant fragment in solution.



Mass Distribution	Radius (nm)	%Pd	Mw-R (kDa)	%Intensity	%Mass	Pd (nm)
Peak 1	3.0	10.0	46	7.4	97.0	0.3
Peak 2	33.8	11.5	12744	40.2	0.5	3.9
Peak 3	131.8	9.5	306905	52.4	2.4	12.5

Figure 5: Dynamic Light Scattering results of NtA$\beta_{(253\text{-}354)}$ show a large amount of small multimers of activase and a few large aggregates.

While no crystals for NtA$\beta_{(253\text{-}354)}$ have been observed to date, this protein is still under investigation. It was designed based on a successfully crystallized fragment of Rubisco activase from the creosote plant. *Henderson et al.* were able to clone, express, purify, and crystallize the substrate recognition domain of higher plant Rubisco activase.[21] This fragment has the expected four helix bundle all α-domain; however, its helix number 3 is interestingly unique. Helix 3 extends outward further than was expected. Helix 3 of Henderson's model is important because it contains amino the acids of the species specificity recognition domain. There have been two higher plant Rubisco activase structures published with different conformations for this specific helix. While Henderson's structure has a roughly 9-turn helix, Stotz's paper has a structure of the same helix with only

about 6 turns connected by a linker to a small helix of about 1.5 turns. Stotz's small helix is the helix that contains the amino acids responsible for species specificity. It is possible that either structure is the physiological relevant structure. The only way to determine the purpose of the long helix and the physiologically relevant structure is by solving more activase structures and comparing them. Hopefully with a little more truncation or different crystallization parameters, the NtA$\beta_{(253-354)}$ fragment will crystallize so that it can be compared to both the creosote structure and Stotz's structure. The most recent paper by Stotz did not comment on the differences between his model's helix 8 and helix 9 and Henderson's model's equivalent helix 3.

3.1.3 Limited Proteolysis

A chart of the optimized conditions to obtain the maximum amount of limited proteolysis products can be seen below in Table 1. The molar ratio of proteolytic enzyme to activase was the most essential condition to optimize. The optimized ratio of enzyme to activase leads to the highest concentration of products with the least amount of completely digested protein. Having too much proteolytic activity leads to only very small protein fragments; not having enough proteolytic activity leaves the majority of the proteins completely intact.

	Moles (1ml rxn)	Molar Concentration (1ml rxn)	Mass (1ml rxn)	Molar Ratios (1ml rxn)
Truncated Tobacco Activase	16.8 nM	16.8 mM	625 µg	Activase:Chymotrypsin 233:1
ADP	8.3 nM	8.3 mM	3.5 µg	Activase:Trypsin 224:1
Chymotrypsin /Trypsin	72 fM /75 fM	72 nM /75 nM	1.8 µg /1.7 µg	Activase:ADP 2.02:1

Table 1: Table of optimized digestion reaction, including: moles, concentrations, mass, and molar ratio of protein to proteolytic enzyme.

The concentration of ADP was not as important as it was expected to be. Varying the amount of ADP did not have a noticeable effect on the digestion products based on SDS-PAGE results. By reducing the amount of ADP present in the digestion, fewer activase proteins should be bound to ADP. Activase not bound to ADP should be more flexible, allowing better access for the proteolytic enzyme to bind to the parts of activase that may be inaccessible while bound to ADP. This does not appear to be the case as can be seen in Figure 6.

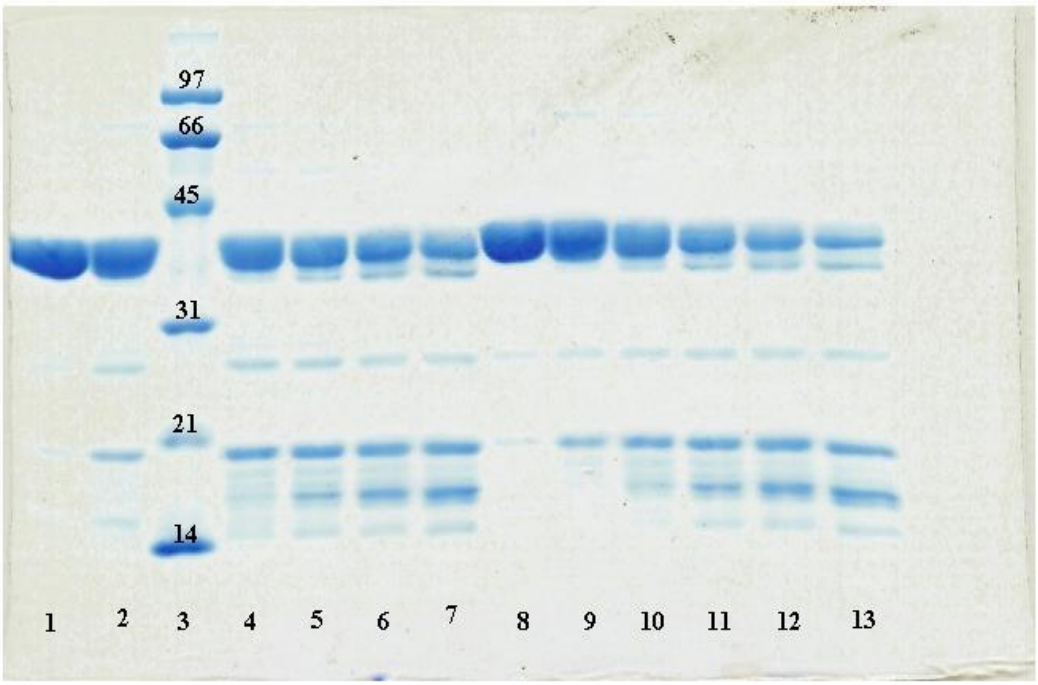


Figure 6: SDS-PAGE result of two different ADP concentration digestions via α-chymotrypsin over 120 min. Activase and α-chymotrypsin concentrations are held constant at 16.8 μM and 72 nM respectively. Lanes 2,4,5,6, and 7 are 0 min, 30 min, 60 min, 90 min, and 120 min with 8.3 μM ADP. Lanes 9, 10, 11, 12, and 13 are 0 min, 30 min, 60 min, 90 min and 120 min with 28.3 μM ADP. Lanes 1 and 8 are controls (only activase after 120 min). Lane 3 is the molecular weight marker.

The samples of activase provided by Dr. Michael Salvucci contained 1 mM ADP. In order to reduce the ADP level to less than a 1:1 ratio, a sample of activase was dialyzed against buffer with no ADP and subsequently digested. By looking at the gel above in Figure 6, it is apparent that no substantial amount of new protein fragments were created by reducing the amount of ADP from the digestion. It may be possible that activase has a very high affinity for ADP and that the dialysis did not remove sufficient amounts of ADP. Along the same lines, it is possible that if activase has a high affinity for ADP, then as the activase proteins that are bound to ADP are cleaved, they could lose their association with

ADP. This would create more available ADP for intact activase proteins to associate with. One way to fix this is to create a recombinant protein and purify it without ADP.

After digestion conditions were optimized, digestions were performed so that the fragments could be isolated and characterized. Isolation of fragments via RP-HPLC was unsuccessful. It seems as though most of the sample became aggregated and was removed by either the syringe filter or the column pre-filter. Several different volumes of digestion were injected. None of the injection volumes provided better results than another. Next, protein blotting was used to isolate the fragments. Numerous samples were cut out of the PVDF membrane with the protein blotted onto it and were given to the proteomics lab for N-terminal sequencing. The lab was able to identify the N-terminal sequence on two of the samples. One fragment from an α-chymotrypsin (NKQENA) digest and one fragment from a trypsin digest (MGGTTQYTVN). The other samples were too dilute to receive accurate results. In an attempt to try and increase the amount of protein on the PVDF membrane, several digestions were performed with increased protein concentration and increased blotting time, although the concentrations of the fragments on the PVDF membrane could not be increased enough to be N-terminally sequenced. The minimum amount of protein required for sequencing is about 12 picomoles so the digestion products must not be very prominent. The mass, determined by MALDI-TOF, of the trypsin digest fragment was 19.78 kDa, and the mass of the α-chymotrypsin digest fragment was 16.71 kDa. The fragment of protein identified from the trypsin digest was a 175 amino

acid fragment believed to consist of some of the αβα subdomain and most of the all α subdomain based on the homology model. The fragment of protein identified from the α-chymotrypsin digest was a 146 amino acid fragment, which again consisted of some of the αβα subdomain and most of the all α subdomain.

While being able to identify two protein fragments using limited proteolysis was a good start, it does not look as though any more fragments will be able to be identified in the same manner. It is likely that the digestion products that we are interested in are intermediate products which are only around for a short period of time before they are digested into smaller protein fragments.

3.1.4 NtAβ$_{(182-383)}$ and NtAβ$_{(216-383)}$

The two fragments NtAβ$_{(182-383)}$ and NtAβ$_{(216-383)}$ were successfully cloned into pET151/dTOPO. Expression of the NtAβ$_{(182-383)}$ fragment yielded no visible bands via SDS-PAGE correlating to a fragment of this size. Expression temperatures of 25°C, 30°C, and 37°C were tested, and no bands for this fragment were visible via SDS-PAGE with any of these temperatures. Expression of the NtAβ$_{(216-383)}$ fragment yielded a large amount of protein correlating to a fragment of this size. However, this large amount of protein was in a non-soluble fraction. While trying to purify and crystallize non-soluble protein fragments is an option, that course of action could be time consuming and there is nothing to suggest that this fragment will crystallize.

3.2 Conclusion and Future Work

While limited proteolysis worked for the group that isolated the *E. coli* Lon protease all-α subdomain, it was not successful in this study to isolate

fragments of activase and crystallize them. Creating a recombinant protein based on the homology model worked well; however, often times the truncated proteins that are created are not very soluble. The next step towards the crystallization of activase is to design and express more truncated recombinant activase proteins. The recombinant proteins can be designed based on the newest published activase structures and protein stability prediction programs such as Expasy's protein parameter tool or FoldIndex. The recombinant proteins' thermal stability and secondary structure compositions will be used to help guide the protein design process.

Once a stable, soluble recombinant protein is expressed, crystallization screens will need to be performed and crystallization conditions will need to be optimized. Once the crystallization conditions are optimized and protein crystals of suitable quality are produced, X-ray diffraction data will be collected. The phase problem may be solved with heavy atom soaks from isomorphous crystals or from selenomethionine substituted proteins. Then, using the electron density map, the structure of the recombinant protein can be computed.

REFERENCES

1. Barta, C.; Dunkle, A. M.; Wachter, R. M.; Salvucci, M. E. Structural changes associated with the acute thermal instability of Rubisco activase Archives of Biochemistry and Biophysics 2010, 499, 17-25.

2. Feller, U.; Anders, I.; Mae, T. Rubiscolytisc: fate of Rubisco after its enzymatic function in a cell is terminated *Journal of Experimental Botany* **2008**, 59, 1615-1624.

3. Lorimer, G.H.; Miziorko, H.G. Carbamate Formation on the €-Amino Group of a Lysyl Residue as the Basis for the Activation of Ribulosebisphosphate Carboxylase by CO_2 and Mg^{2+}. *Biochemistry* **1980,** 19, 5321-5328.

4. Parry, M. A.; Keys, A. J.; Madgwick, P. J.; Carmo-Silva, A. E.; Andralojc, P. J. Rubisco regulation: a role for inhibitors *Journal of Experimental Botany* **2008**, 59, 1569-1580.

5. Griffiths, H. Designs on Rubisco *Nature Plant Biology* **2006**, 441, 940-941.

6. Eichelmann, H.;Talts, E.; Oja, V.; Padu, E.; Laisk, A. Rubisco in planta k_{cat} is regulated in balance with photosynthetic electron transport. *Journal of Experimental Botany* **2009**, 60, 4077-4088.

7. Houtz, R. L.; Portis, Jr. A. R. The life of ribulose 1,5-bisphosphate carboxylase/oxygenase — posttranslational facts and mysteries *Archives of Biochemistry and Biophysics* **2003**, 414, 150-158.

8. Salvucci, M. E.; Portis Jr. A.R.; Ogren W. L. Light and CO_2 response of ribulose-1,5-bisphosphate carboxylase oxygenase activation in *Arabidopsis* leaves. *Plant Physiol.* **1986**, 80, 655-659.

9. Parry, M.A.; Reynolds, M.; Salvucci, M. E.; Raines, C.; Andralojc, P. J.; Zhu, X. G.; Price, G. D.; Condon, A. G.; Furbank, R. T. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency *Journal of Experimental Botany* **2011**, 62, 453-467.

10. Andersson, I. Catalysis and regulation in Rubisco *Journal of Experimental Botany* **2008**, 59, 1555-1568.

11. Portis, Jr. A. R. Rubisco activase – Rubisco's catalytic chaperone *Photosynthesis Research* **2003**, 75, 11-27.

12. Portis, Jr. A. R. Rubisco activase *Biochimica et Biophysica Acta* **1990**, 1015, 15-28.

13. Neuwald, A. F.; Aravind, L.; Spouge, J. L.; Koonin, E. V. AAA+: A Class of Chaperone-Like ATPases Associated with the Assembly, Operation and Disassembly of Protein Complexes. *Genome Research* **1999**, 9, 27-43.

14. Snider, J.; Houry, W. A. AAA+ proteins: diversity in function, similarity in structure *Biochemical Society Transactions* **2008**, 36, 72-77.

15. Iyer, L. M.; Leipe, D. D.; Koonin, E. V.; Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. *Journal of Structural Biology* **2004**, 146, 11-31.

16. Li, C.; Salvucci, M. E.; Portis, Jr. A. R. Two residues of Rubisco activase involved in recognition of the Rubisco substrate *The Journal of Biological Chemistry* **2005**, 280, 24864-24869.

17. Melnikov, E. E.; Andrianova, A. G.; Morozkin, A. D.; Stepnov, A. A.; Makhovskaya, O. V.; Botos, I.; Gustchina, A.; Wlodawer, A.; Rotanova, T. V. Limited proteolysis of *E. coli* ATP-dependent protease Lon – a unified view of the subunit architecture and characterization of isolated enzyme fragments *Acta Biochimica Polonica* **2008**, 55, 281-296.

18. Yamaguchi, H.; Miyazaki, M.; Honda, T.; Nagata, M. P. B.; Arima, K.; Maeda, H. Rapid and efficient proteolysis for proteomic analysis by protease-immobilized microreactor *Electrophoresis* **2009**, 30, 3257-3264.

19. Botos, I.; Melnikov, E. E.; Cherry, S.; Khalatova, A. G.; Rasulova, F. S.; Tropea, J. E.; Maurizi, M. R.; Rotanova, T. V.; Gustchina, A.; Wlodawer, A. Crystal structure of the AAA^+ α domain of *E. coli* Lon protease at 1.9 Å resolution. *Journal of Structural Biology* **2004**, 146, 113-122.

20. Ammelburg, M.; Frickey, T.; Lupas, A. N. Classification of AAA+ proteins *Journal of Structural Biology* **2006**, 156, 2-11.

21. Henderson, J. N.; Kuriata, A. M.; Fromme, R.; Salvucci, M. E.; Wachter, R. M. Atomic Resolution X-ray Structure of the Substrate Recognition Domain of Higher Plant Ribulose-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase *The Journal of Biological Chemistry* **2011**, 286, 35683-35688.

22. Portis, Jr. A. R.; Li, C.; Wang, D.; Salvucci, M. E. Regulation of Rubisco activase and its interactions with Rubisco *Journal of Experimental Botany* **2008**, 59, 1597-1604.

23. Stotz, M.; Mueller-Cajar, O.; Ciniawsky, S.; Wendler, P.; Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. Structure of green-type Rubisco activase from tobacco *Nature Structural & Molecular Biology* **2011**, 18, 1366-1371.

24. Parry, M. A. J.; Hawkesford, M. J.; Food security: increasing yield and improving resource use efficiency *Proceedings of the Nutritional Society* **2010**, 69, 592-600.

25. Long, S. P.; Zhu, X. G.; Naidu, S. L.; Ort, D. R. Can improvement in photosynthesis increase crop yields *Plant, Cell and Environment* **2006**, 29, 315-330.

26. Sage, R. F.; Way, D. A.; Kubien, D. S. Rubisco, Rubisco activase, and global climate change *Journal of Experimental Botany* **2008**, 59, 1581-1595.

27. Rhodes, G. *Crystallography made crystal clear* 2nded.; Academic Press: San Diego, 2000.

28. Cummings, M. D.; Farnum, M. A.; Nelen, M. I. Universal screening methods and applications of Thermofluor *Journal of Biomolecular screening* **2006**, 11, 845-863.

29. Maeda, Y.; Yamada, H.; Ueda, T.; Imoto, T. Effect of additives on the renaturation of reduced lysozyme in the presence of 4 M urea. *Protein engineering* **1996**, 9, 461-456.

30. Wilson, W. W. Light scattering as a diagnostic for protein crystal growth-A practical approach *Journal of Structural Biology* **2003**, 142, 56-65.