

5-2017

Cloning and Expression of C-terminal Fragment of TonB from *Rhizobium leguminosarum* ATCC 14479

Brooke E. Baxter

Follow this and additional works at: <https://dc.etsu.edu/honors>



Part of the [Bacteriology Commons](#)

Recommended Citation

Baxter, Brooke E., "Cloning and Expression of C-terminal Fragment of TonB from *Rhizobium leguminosarum* ATCC 14479" (2017).
Undergraduate Honors Theses. Paper 410. <https://dc.etsu.edu/honors/410>

This Honors Thesis - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Cloning and Expression of C-terminal Fragment of TonB from *Rhizobium leguminosarum* ATCC 14479

By: Brooke Baxter

An Undergraduate Thesis Submitted in Partial Fulfillment
of the Requirements for the
Honors-in-Discipline Microbiology Program
College of Public Health
East Tennessee State University

Brooke Baxter Date

Dr. Ranjan Chakraborty, Thesis Mentor Date

Dr. Sean J. Fox, Reader Date

Mrs. Tory Street, Reader Date

Abstract

The TonB-ExbB-ExbD complex is essential for the siderophore mediated acquisition of iron by Gram negative bacteria. The system provides energy from the proton motive force to the outer membrane in order for the iron siderophore complex to enter into the cell. The main protein involved in energy transduction, TonB, has been extensively studied in the species *Escherichia Coli*. It has been determined that the protein consists of 239 amino acids. In comparison, however, the TonB of *Rhizobium leguminosarum* consists of 457 amino acids with the same conserved regions. What is in question, therefore, is how the additional amino acids effect the structure of the C-terminal region of the protein and how such information can give insight into the way in which the proton motive force functions to provide energy to the outer membrane receptor. The protein regions of *R. leguminosarum* TonB chosen for study were 120 and 248 amino acids from the C-terminal end. Genomic DNA was isolated, primers were designed for each fragment, and polymerase chain reactions were performed. After appropriate restriction enzyme digestion, each DNA fragment was ligated into the plasmid pET-17b and then transformed into *Escherichia Coli* BL21 (DE3). Successful transformation of the 120 amino acid fragment was followed by expression via IPTG induction & extraction of protein. Afterwards, a T7-tag affinity column was attempted to collect the protein for analysis; however, a sufficient amount of protein was not eluted. The procedure will be repeated for obtaining sufficient protein for crystallization or NMR spectrometric analysis.

Background

Iron is an element vital to the survival of most organisms. Iron is required for many biological pathways including photosynthesis, nitrogen fixation, respiration, gene regulation, and DNA biosynthesis (Andrews et al. 2003). In addition, the element has been identified in over one hundred enzymes needed for various metabolic processes (Guerinot 1994). For early life forms, iron was essentially the most abundant and easily acquired element; therefore, its strong presence is justifiable. However, since that time period, readily available iron has changed in chemistry. The previously soluble iron is now found in the environment in an extremely insoluble state. Fe (II) is oxidized in the presence of oxygen to Fe (III). Fe (III) can in turn react with oxygen and water in basic and neutral conditions (Andrews et al 2003). Consequently, iron in the presence of oxygen is not readily available for uptake by microorganisms and can even be potentially toxic. For that reason, microorganisms have developed specific methods to satisfy their need for the element. There are three broad categories for iron uptake including: reduction of the ferric iron to soluble form, lowering of the environmental pH to increase the ferric iron solubility, and production of ferric chelators known as siderophores.

Siderophore production, of the three methods previously listed, is the most commonly seen method in bacteria. Siderophores are small molecules with a high affinity for iron that are secreted by gram negative bacteria. They are produced only when the cells are depleted of iron; and therefore, are a consequence of negative feedback. To date there have been over five hundred different siderophores identified. They can be categorized according to the presence of various functional groups. Their method for iron transport, however, is very similar in all cases. Siderophores are released from the cell and bind tightly to ferric iron to form stable complexes (Methke and Marahiel 2007). They are then accessible for reuptake by the cell. The complexes are too large to enter via porins within the cell, and due to the concentration gradient, require the use of energy for transport. Siderophores, therefore, rely on the use of outer membrane receptors. Outer membrane receptors vary in their specificity towards

different siderophores; however, the general structure has been found to remain consistent. Each has an inner cork or plug portion that is rearranged once a siderophore is bound. This process involves active transport, and thus, requires the use of energy. Gram negative bacteria do not possess an energy source in their outer membrane. The inner membrane is the location of the proton motive force that is used to drive cellular acquisitions. Thus energy must somehow be transferred from the inner membrane to the outer membrane receptor. For siderophore mediated iron transport the energy transducing system, is the TonB-ExbB-ExbD complex (Ferguson and Deisenhofer 2002).

The complex contains three proteins TonB, ExbB, and ExbD, which work in conjunction. TonB is the main protein involved in energy transduction while the other two are considered accessory structures. They are present in a 1:7:2 ratio (Braun and Hantke 2013). ExbB is found in the largest concentration within the complex due to its function. ExbB acts as a scaffold for the complex, holding TonB and ExbD together (Chakraborty 2013). Its C-terminal domain is located in the cytoplasm while the N-terminal domain can be found in the inner membrane. ExbD, on the other hand, is found at a much lower concentration within the system. The protein assists in energy transduction. The N-terminal region of ExbD must be present in order for the proton motive force to be efficient. The C-terminal end directly connects to TonB. TonB is the largest and most studied protein within the complex. TonB itself consists of three identified regions: an N-terminal domain, a proline rich region, and a C-terminal domain. The N-terminal domain is hydrophobic and attaches the protein to the cytoplasmic membrane. The domain is also the main site for association with ExbB and ExbD. The proline rich region consist of repeating Pro-Glu and Pro-Lys amino acids (Chakraborty 2013). The function of the region is to provide support for the complex when binding to the outer membrane receptor. The amino acids form a rigid structure that span the space between the inner and outer membranes. Removal of the proline rich region does not result in loss of energy transduction, therefore, it is not directly involved within the process. The last region of the TonB protein is the C-terminal domain. The C-terminal domain provides the direct connection between TonB and the outer membrane receptor (Andrews et al 2003). The domain connects to the N-terminal region of the TonB box of the outer membrane receptor. This region of TonB is vital for energy transduction, and therefore, overall iron acquisition by the cell.

The TonB complex has been extensively studied in the species *E. coli*. It has been found to be 239 amino acids long (Chakraborty 2013). The C-terminal end makes up the last 155-239 of those amino acids. Importance has been placed on this specific region due to its function. The exact mechanism for energy transduction is currently under debate. What is known is that the outer membrane receptor has a portion called the TonB box that extends once the siderophore binds to the protein. The TonB box consists of seven amino acids of the N-terminal domain of the outer membrane receptor. The extension serves as the available connection to the TonB C-terminal domain. This connection is vital for the energy transduction step. Previous studies have shown removal of the TonB box results in the loss of iron uptake by the cell (Andrews et al 2003). The structure for the C-terminal domain of *E. coli* has been recently obtained along with the structure of the domain combined with FhuA and BtuB. The C-terminal TonB of *E. coli* is composed of two alpha helices and four beta strands. The last 48 residues of the C-terminal TonB provide the direct connection with the outer membrane receptor's TonB box (Change et al 2001). The portions of the outer membrane receptors, FhuA and BtuB, not located within the TonB box, remained unchanged upon binding with TonB. The confirmation of TonB, however, changes considerably after interaction with the N-terminal of the outer membrane receptors. The inner plug structure of TonB does not dissociate completely. Only the necessary number of bonds break to allow

the siderophore to pass into the periplasmic space. While the exact energy transduction mechanism is still not known, it is assumed a larger opening in the TonB box would allow for phages to inject their DNA and colicins to pass into the cell. It is thought that the incomplete dissociation of the TonB box protects the cell from possible invaders during siderophore transport (Chang et al 2001).

While there is still much to learn about the TonB complex in *E. coli*, there is much less information available about TonB from other species of bacteria. One such species that has not been thoroughly studied is *R. leguminosarum*. Rhizobium are nitrogen fixing bacteria found in the soil (Wexler et al 2001). Because of their nitrogen fixation activity and the enzymes involved in the process, they require a large amount of iron. In order to obtain the sufficient concentration of iron, the *R. leguminosarum* has been known to secrete the siderophore vicibactin (Wright et al 2013). The Ton-B protein has been recognized in the species through the common conserved motif "SHLS" at amino acids 103-115. The same sequence has been identified in *E. coli* and is known to connect the three proteins of the TonB complex together. It serves as common identifier for the TonB system. Also similar to *E. coli* is the percentage of proline in the proline rich region of TonB and repetition of the Pro-Glu and Pro-Lys regions. The most interesting finding to date however, is the total number of amino acids composing TonB in comparison to the TonB present in *E. coli*. In *R. leguminosarum* TonB consists of 457 amino acids in comparison to the 239 of *E. coli*. The differences in length are primarily within the N-terminal domain of the protein; however, it is unknown how addition of such amino acids affects the energy transduction via proton motive force. The C-terminal domain has not been comprehensively studied in the species; therefore, such information regarding the outer transmembrane receptors and TonB box is unknown.

Present Work

In order to obtain further information regarding energy transduction of the TonB complex in *R. leguminosarum* the C-terminal domain must be expressed and analyzed. It is proposed that the C-terminal domain expression and purification of the protein will allow for comparison to the TonB of *E. coli* and other thoroughly studied species. The structure determination via crystallization or nuclear magnetic resonance spectroscopy, will probably provide additional information for what is necessary for energy transduction. Comparison of proteins within the C-terminal domain will essentially help to confirm commonalities across all species with TonB dependent energy transport. The determination of the C-terminal structure is the next stage of analysis for the *R. leguminosarum* comprehensive study of TonB in relation to iron acquisition. In order for the structure to be determined and analyzed, the protein must first be expressed and purified in adequate amounts, and is therefore, the goal of this experimentation.

Materials and Methods

Organisms and Media:

Rhizobium leguminosarum ATCC 14479 was obtained from American Type Culture Collection, Manassas, USA. *E. coli* BL21 (DE3) was purchased from New England Biolabs, Ipswich, MA, USA. Yeast Mannitol broth was used for maintenance and growth of *R. leguminosarum* and LB broth was used for growth of the competent *E. coli* BL21 (DE3).

Genomic DNA Extraction

R. leguminosarum ATCC 14479 was grown in Yeast Mannitol Broth in order to acquire sufficient amounts of genomic DNA for experimental application. The YMB media was composed of 1% mannitol, 0.05% K₂HPO₄, 0.02% MgSO₄*7H₂O, 0.01% NaCl, and 0.1% yeast extract, adjusted to a pH of 6.8 using NaOH, and autoclaved prior to use. The 5mL Yeast Mannitol Broth was inoculated with *Rhizobium leguminosarum* ATCC 14479 and placed on a shaker at room temperature. After three days, the genomic DNA was extracted using the CTAB method adapted from Current Protocols in Molecular Biology (Wilson 1987).

Polymerase Chain Reaction

In order to amplify the TonB gene of interest, a polymerase chain reaction was designed and executed. For the polymerase chain reactions, two separate experiments were performed, one for the 120 amino acid fragment and one for the 248 amino acid fragment. In each 5µl of NEB 10x taq buffer, 1µl 10mM dNTP, 1µl of the 10uM forward primer, 1µl of the 10uM reverse primer, 4µl of the isolated *Rhizobium leguminosarum* ATCC 14479 genomic DNA, 0.25µl of NEB taq polymerase, and 37.75µl of ddH₂O were added to a PCR tube. For the 120 amino acid fragment, the forward primer utilized was TonB120. For the 248 amino acid fragment, the forward primer utilized was TonB248. The reverse primer was TonBRev for both fragments. The primer sequences are shown in table 1. The thermocycler program was set to an initial denature temperature of 95°C for two minutes, denature at 95°C for an additional 15 seconds, anneal at 51°C for 20 seconds, extension at 68°C for 45 seconds, and a final extension at 68°C for 5 minutes. The results of the PCR amplification were viewed on a 1% agarose gel. The size of the amplified sequences were analyzed for accuracy of design.

Polymerase Chain Reaction Primers	
Name	Sequence
TonB120	5' GCGCAAGCTTGATGGTGAACGGCCAGGACGGGCCG 3'
TonB248	5' GCGCAAGCTTGATGGAGACCGTTCAGTCGGTGCAG 3'
TonBRev	5' GCGCGAATTCTTATCTAGTGACCTGAACTCCTGG 3'

Table 1: Primers used in PCR Reactions

Purification of PCR Amplification from Agarose Gel

After the polymerase chain reaction was complete and the fragments of desired size were confirmed, the DNA was purified from the agarose for subsequent use. To begin the procedure the bands found at 360 base pairs, corresponding to the 120 amino acid fragment, and the bands located at 744 base pairs, corresponding to the 248 amino acid fragment, were cut out of the gel. UV light was applied and a razor blade was used to carefully remove the minimal amount of agarose that included each DNA band. The bands were then weighed, labeled, and separately placed in microcentrifuge tubes. The DNA was then purified using an MO BIO Laboratories, located in Carlsbad, California, UltraClean15 DNA Purification Kit. The success of the Ultraclean 15 DNA purification kit was assessed for both the 120 amino acid DNA and the 249 amino acid DNA using the NanoDrop before proceeding.

Isolation of pET-17b

Once the DNA coding for the 120 amino acid fragment and the DNA coding for the 248 amino acid fragment of C-terminal TonB from *Rhizobium leguminosarum* ATCC 14479 were successfully amplified and purified, the plasmid needed for cloning was isolated. The plasmid chosen and used in PCR design was pET-17b obtained from Novagen in Madison, Wisconsin. A stock culture of *E. coli* containing pET-17b, therefore, was used to streak a LB plate containing ampicillin. The plate was placed in 37°C to grow overnight. The 5ml LB broth containing 100 mg per ml ampicillin was inoculated with a single colony from the LB plate. The 5ml LB broth was placed on the shaker at 37°C for overnight growth. After sufficient growth and visible turbidity, pET-17b was isolated from the culture using QIA Prep Spin Miniprep Kit. The broth culture was centrifuged in a microcentrifuge tube at 10,000 rpm for three minutes. The supernatant was discarded, and the procedure was repeated utilizing all of the LB culture. For plasmid isolation Qiagen miniprep kit (Qiagen Inc.) was used. The pellet formed was resuspended in 250µl buffer P1. Next, 250µl of buffer P2 was added, and the microcentrifuge tube was inverted 4 to 6 times. Afterwards, 350µl of buffer N3 was added and inverted an additional 4 to 6 times. The sample was then centrifuged at 13,000 rpm for 10 minutes. The supernatant was removed and applied to a spin column. The spin column was centrifuged at 13,000 rpm for 60 seconds. The flow through was discarded and the column was washed with 0.5ml buffer PB. Then, the spin column was centrifuged at 13,000 rpm for 60 more seconds. The flow through was discarded, and the procedure was repeated with a second wash of 0.75ml buffer PE. Afterwards, the spin column was centrifuged at 13,000 rpm for one more minute to ensure complete removal of the wash buffers. Next, to elute the DNA, 50µl of buffer EB was applied to the center of the spin column. The spin column was allowed to sit for one minute before centrifuging for one minute to obtain the purified plasmid isolated in buffer EB. The concentration and size of the isolated pET-17b was confirmed before proceeding.

Restriction Enzyme Digestion

After the DNA from the two fragments of interest was amplified and the plasmid was isolated, restriction enzyme digestion was performed. The two restriction enzymes utilized were chosen based on their location in pET-17b's cloning region. Thus, their restriction sites were also incorporated in the primer design for the fragments as explained in the polymerase chain reaction. In order to clone the 120 amino acid DNA and the 248 amino acid DNA in pET-17b for expression, the two inserts and the plasmid needed complementary "stick ends" for ligation. Therefore, five separate restriction enzyme digestion procedures were performed, two for the two samples of amplified and purified DNA corresponding to the 120 amino acid TonB, two for the two samples of amplified and purified DNA corresponding to the 248 amino acid TonB, and one for the isolated pET-17b sample. To each reaction 1µg of DNA was added. The amount in µl was determined based on the concentration given by the NanoDrop readings. Next, 5µl of CutSmart 10X buffer, 1µl of EcoRI, and 1µl of HindIII were added to each sample. Distilled water was then added to each reaction for a total volume of 50µl. The reactions were placed in 37°C for one hour. Afterwards, the results were assessed with a 1% agarose gel. Bands of appropriate size were cut out and purified using the Ultraclean 15 DNA purification kit as previously explained. The concentration of each was further determined using the NanoDrop.

Ligation

Following successful restriction enzyme digestion of both the inserts and the plasmid, the next step was to ligate the two DNA segments together. Through the ligation reaction the 120 amino acid C-

terminal DNA was cloned into the pET-17b cloning region. The same procedure placed the 248 amino acid C-terminal DNA into the pET-17b cloning region of a separate sample. Therefore, two separate reactions were performed. For each, 50ng of pET-17b was added to a microcentrifuge tube. The amount of plasmid in microliters was determined by the previous NanoDrop readings from the restriction enzyme digest purification. The 120 amino acid restriction enzyme digestion DNA was added to one reaction and the 248 amino acid restriction enzyme digestion DNA was added to the second reaction. Both were added in a concentration 3X that of the pET-17b. The concentration of the inserts was increased in comparison to the plasmid in order to increase the chance the two would combine. In both reactions, 2µl of 10X T4 ligase buffer, 1µl of T4 DNA ligase, and ddH₂O to a final volume of 20µl were added. The T4 DNA ligase was the last of the reagents to be added. Then the reactions were incubated at room temperature for 10 minutes. Next, they were heat inactivated at 65°C for 10 minutes. They were then stored at 4°C until transformation could be performed.

Transformation

To begin transformation, competent cells were thawed on ice for 10 minutes until all crystals were dissolved. Next, 50µl of the cells were placed in two separate transformation tubes on ice. To one reaction, 10µl of the pET-17b ligation with the 120 amino acid insert were added, and to the second reaction, 10µl of the pET-17b ligation with the 248 amino acid insert were added. Both reactions were mixed by flicking 4 to 5 times and placed on ice for 30 minutes. After the allotted time, they were then heat shocked at 42°C for 30 seconds. Immediately following, the reactions were placed on ice again for 5 minutes. Next, 950µl of room temperature SOC was added to each reaction and they were positioned on the shaker at 37°C for 60 minutes. LB plates with ampicillin were also warmed to 37°C during that time. After the one hour, the cells were thoroughly mixed and centrifuged for one minute. All liquid was removed except for 100µl. The cells were resuspended in the 100µl and plated on the warm LB plates with ampicillin. Both plates were placed in 37°C for overnight growth.

To verify transformation, the colonies grown were streaked onto separate LB plates with ampicillin and inoculated into LB broths with ampicillin. The inoculated LB broth was utilized for plasmid isolation. After re-isolating the plasmid from both the 120 amino acid ligation and the 248 amino acid ligation using the QIA Prep Spin Miniprep Kit, restriction enzyme digestion was performed. The protocol for restriction enzyme digestion required 3µl of the isolated plasmid, 5µl of CutSmart 10x buffer, 1µl EcoRI, 1µl HindIII, and 40µl ddH₂O. Two separate reactions were performed, one for the plasmid containing the 120 amino acid insert and one for the plasmid containing the 248 amino acid insert. Both reactions were then placed in 37°C for one hour. Afterwards, the results were assessed using a 1% agarose gel for DNA visualization.

Sequencing

The isolated plasmids from each transformation reaction were sent for sequencing to the core facility of Quillen College of Medicine, ETSU. Both isolated plasmids were diluted to a concentration of 100ng/ml and 10µl were sent to be sequenced. In order to sequence, the corresponding primers from the polymerase chain reaction were also included. For analyzation of the 120 amino acid fragment, 10µl of the primers TonB120 and TonBRev were sent at a concentration of 3µM and for analyzation of the 248 amino acid fragment, 10µl of the primers TonB248 and TonBRev were sent at a concentration of 3µM.

Determination of optimal IPTG concentration for induction

In order to determine the concentration of IPTG required, *E. coli* BL21(DE3) containing pET-17b with the verified 120 amino acid C-terminal was streaked onto an LB plate with ampicillin for single colony growth. The plate was placed in 37°C overnight. The next day, one colony was used to inoculate a 5mL LB broth with ampicillin. The broth was again placed in 37°C overnight. Next, 1.25ml of the broth was utilized to inoculate 4 different 50ml flasks of LB plus ampicillin. A spectrophotometer was used to measure the initial concentration of bacteria in the flasks. The flasks were placed on a shaker at 37°C. After allotted periods of time, 1ml of each culture was removed and the concentration was checked with the spectrophotometer. After approximately seven hours and twenty minutes, all four 50ml cell cultures had reached an optimal density of 0.4 to 0.6. After reaching the desired density, 1ml of culture was removed from each flask. The 1ml was centrifuged at 10,000 rpm for five minutes. The supernatant was discarded, and the pellet was labeled and stored to be utilized as the uninduced control. Next, IPTG was immediately added to the cultures in varying concentrations of 0.1mM, 0.5mM, 1.0mM, and 2mM. To the Lb broth labeled 0.1mM IPTG, 45µl of 100mM stock IPTG were added. Likewise, 250µl of 100mM IPTG were added to the 0.5mM IPTG broth, 500µl to the 1mM IPTG broth, and 1,000µl to the 2mM IPTG broth. All four flasks were then placed on the shaker at 37°C for three hours. After three hours, the cells were collected by centrifugation at 10,000 rpm for five minutes. The supernatant was discarded, and the pellets were stored at -20°C. In order to assess which concentration, 0.1mM, 0.5mM, 1.0mM, or 2.0mM IPTG, was successful at inducing protein expression, SDS-PAGE electrophoresis was conducted.

Expression of the 120 amino acid C-terminal TonB

After the optimal concentration of IPTG was determined as 0.5mM, the procedure was repeated on a larger scale for production of sufficient amounts of protein for purification. The *E. coli* BL21 (DE3) cells from the successful transformation with the pET-17b plasmid containing the 120 amino acid insert, were streaked onto an LB plate with ampicillin. The plate was placed in 37°C overnight. A single colony was then used to inoculate a 2mL LB broth tube with ampicillin. The culture was grown overnight at 37°C. Next, 0.5mL of the grown culture was inoculated into a 50mL LB flask with ampicillin. The culture was again grown at 37°C overnight. The next day, 4 one liter LB broths with ampicillin were each inoculated with 10ml from the overnight 50ml culture. The four flasks were placed on a shaker at 37°C. The concentration of each was monitored over time using a spectrophotometer. After approximately 6 hours, all four flasks had reached optimal density. One milliliter of cell culture was taken from each flask to be utilized as the control before IPTG induction for analysis. The control samples were centrifuged at 10,000 rpm for 5 minutes. The supernatant was discarded, and the pellet was placed in -20°C until further use. The 4 one liter LB flasks were then induced with IPTG to a concentration of 0.5mM. The stock concentration of IPTG utilized was 1 molar; thus, 500µl of IPTG were added to each flask. The flasks were placed on the shaker at 37°C for three hours. After three hours, the cells were centrifuged at 10,000 rpm for 5 minutes. The supernatant was discarded and the pellets were stored at -20°C until confirmation of successful induction. Induction was assessed using SDS-PAGE electrophoresis.

After confirming the 120 amino acid C-terminal TonB was being expressed, the pelleted cells were resuspended in Tris buffer (pH 7.5). The *E. coli* BL21 (DE3) cells were then sonicated on ice for a total five minutes with one minute rest between successive minutes. After sonification was utilized to break open the cells, the cells were centrifuged for 10 minutes at 8,000 rpm. During centrifugation, the

whole cells and cell debris formed a pellet at the bottom of the centrifugation tube. The supernatant, therefore, contained the cellular cytoplasm and membranes. The pellet was discarded and the supernatant was transferred to clean centrifugation tubes. The supernatant was then placed in the ultracentrifuge for one hour and thirty minutes at 30,000 rpm.

Purification of the 120 amino acid C-terminal TonB

After acquiring the cellular cytoplasmic proteins from ultracentrifugation, preparations were made for T7-TAG affinity purification utilizing a Novagen kit. First, 25mL of each 1x T7-Tag bind/wash buffer and T7-Tag elute buffer were prepared. The 1X T7 Tag bind/wash buffer was formulated by diluting 2.5 mL of the 10X T7 Tag bind/wash buffer with 22.5mL of deionized water. The T7-Tag elute buffer was made by combining 2.5mL of 10X T7-Tag elute buffer with 22.5mL of deionized water. Next, the approximately 150mL of sample supernatant from induction with 0.5mM IPTG were passed through a filter to remove any remaining insoluble compounds. The resin within the column was then equilibrated by washing the column with 10ml of 1X T7-Tag bind/wash buffer at room temperature. The 150ml supernatant sample was then loaded into the column in aliquots. The unbound protein or follow through was collected in a 250mL bottle. Next, the column was washed again with 10ml of 1X T7-Tag bind/wash buffer. The flow through was collected. Then, five 1.5mL microcentrifuge tubes were labeled and 150 μ l of 1X T7-Tag Neutralization buffer was dispensed into each. The bound protein was then eluted in five serial 1mL volumes of 1X T7-Tag elute buffer into each microcentrifuge tube. After each fraction was eluted, the sample was capped and inverted several times to mix. The collected samples were then analyzed for purification of the 120 amino acid C-terminal TonB through SDS-PAGE electrophoresis and stored at -20°C.

Results

Genomic DNA Extraction

The genomic DNA of *Rhizobium leguminosarum* ATCC 14479 was successfully extracted using the CTAB method adapted from Current Protocols in Molecular Biology (Wilson 1987). The results are shown below in figure 1.

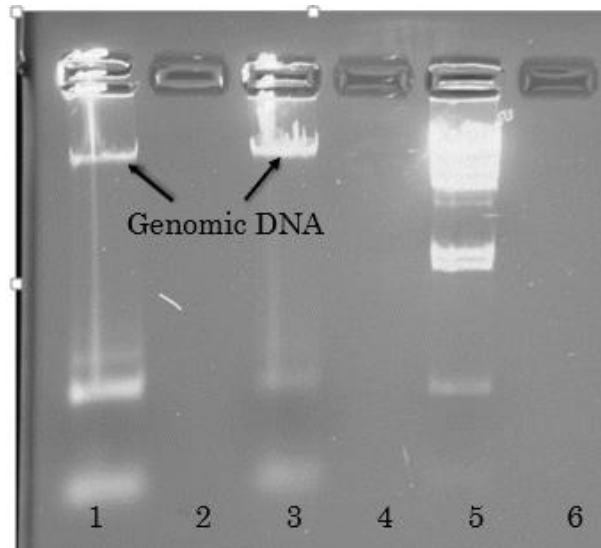


Figure 1: *R. leguminosarum* ATCC 14479 genomic DNA on 1% agarose gel

The 1% agarose gel was prepared according to standard procedure with TAE as the buffer. The DNA was visualized through the use of ethidium bromide and UV light. Lanes 1 and 3 contain 5µl of DNA and 1µl of loading dye. Lane 5 represents the HindIII ladder. Although the bands are not perfectly well-defined, the results clearly indicate the successful extraction of genomic DNA.

Polymerase Chain Reaction

After extracting the genomic DNA from *R. leguminosarum* ATCC 14479, the next step was to utilize the DNA to amplify the sequence corresponding to the protein of interest. The putative TonB in *Rhizobium leguminosarum* ATCC 14479 has the 457 amino acid sequence listed below.

MNPGGHGQTVRGMQMAISAKRSRQVLIGEPDADGSLNDNNMHPGHELSDLRNVQRQPAG
EAVVHYARFAQIPSPDHPEAEPIASVPAPPIDAAVEKQEDERKPMRRRMALTCIGSFFF
HAGLVVALSVFMPTPPNEAIEDAGEAVSVVVYGDSDVDQTSAGDPELERQPEQVASEEVE
PDTVQSEEATELAATTVPPEQAQPVETETMETVQSVQEVTRVSPEAVVAAEPEVLVSESP
AESFVAQPMATAVPEQMPDIAQATVPEEVPTAVQPTAVPPEEVKPVETVEVSPEPEDK
PKPKAKKPKPVEKKQSQRKRAKPAGGKEGSDREDSTRGMVNGQDGPQTDGTSTTTGGTDGM
GSAAVANYPGKIQKRIRRAVRVPDEYKNKSGGMTVRIQLTINGTGRVASVSVARSSGIAE
LDKAVLDGVRRAAPFPPLPSEWGKPSWTFQAQEVQVTR*

*TAA-Terminating code

Two varying lengths of the C-terminal domain were chosen for study based on previous experimentation and future applications. Prior students decided upon a 120 amino acid sequence and a 200 amino acid sequence. In the case that both proteins were overexpressed and purified, the two differing lengths would have a consistent area for comparison. The longer 200 amino acid length would also contain a portion of the proline rich region in addition to the C-terminal domain. Crystallization of both regions would give more credibility to the structural results. The proline rich region would be distinctive within the structure, showing separation of the regions within TonB. Experimentation with the 120 amino acid sequence indicated probable overexpression. The 200 amino acid sequence, however, was never successfully cloned and transformed into competent cells. Therefore, the 120 amino acid sequence was chosen along with a new 248 amino acid sequence for study. Both sequences were selected based on the start codon for translation methionine. The first methionine of each is highlighted in the above sequence.

The vector of choice for cloning was pET-17b. PET-17b was selected for several reasons. The plasmid contains an N-terminal T7-Tag sequence followed by multiple cloning sites. Sequences inserted within the cloning sites can be easily overexpressed by inducing the T7 promoter with Isopropyl β-D-1-thiogalactopyranoside. Isopropyl β-D-1-thiogalactopyranoside mimics lactose within the cells without interfering in metabolic pathways. As a result the cloned protein is attached to a T7 tag easily utilized in affinity chromatography (Kemple 2013). The pET-17b vector also contains ampicillin resistant genes that

can be utilized to ensure the growth of only those cells that have successfully acquired the plasmid during transformation. The pET-17b vector is shown below (Novagen 1998).

pET-17b sequence landmarks	
T7 promoter	333-349
T7 transcription start	332
T7•Tag coding sequence	237-269
Multiple cloning sites (Hind III - Xho I)	141-228
T7 terminator	28-74
pBR322 origin	1480
bla coding sequence	2241-3098

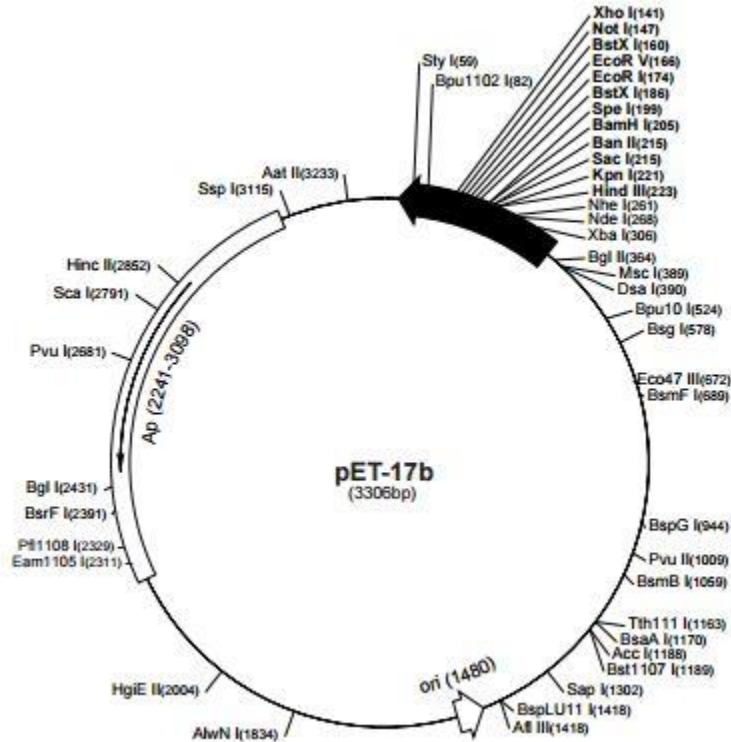


Figure 2: pET-17b Vector Map (Novagen 1998)

Within the cloning region, the two restriction enzymes sites chosen for insertion, and thus utilized in the primers for PCR, were HindIII and EcoRI. HindIII contains the palindromic sequence AAGCTT and was incorporated into the forward primers. EcoRI has the palindromic sequence GAATTC and was included within the reverse primer. The forward primers also contain an additional guanine base after the base sequence for HindIII. The base was incorporated in order to keep the reading frame. The detailed sequence of the cloning region of pET-17b is shown below.

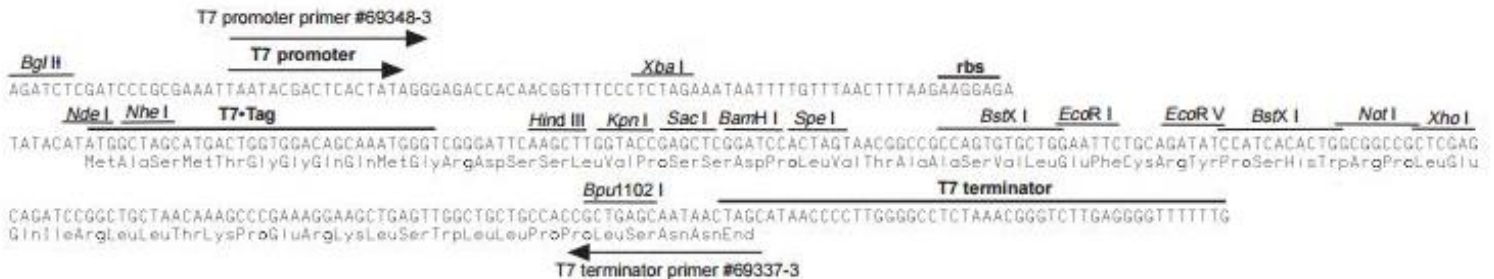


Figure 3: pET-17b Cloning region (Novagen 1998)

The primers for the polymerase chain reaction were designed accordingly. Each contains a GC region on the 5' end. The extra bases ensure the DNA of interest will not be cleaved during future

restriction enzyme digestion. The 3' end also ends in GC base pairing and is identical to the sequence of interest. GC pairing of three hydrogen bonds will confirm tight binding to the genome. The length of each is within the recommended standards. According to software applications, similar sequences are not found at other locations within the genome, and the forward and reverse primers are not complimentary. Hairpin loops should not be formed and additional secondary structures should be minimal. The primers are listed in table 1. Their complimentary regions at the 120 amino acid and 248 amino acid location are highlighted subsequently.

Polymerase Chain Reaction Primers	
Name	Sequence
TonB120	5' GCGCAAGCTTGATGGTGAACGGCCAGGACGGGCCG 3'
TonB248	5' GCGCAAGCTTGATGGAGACCGTTCAGTCGGTGACG 3'
TonBRev	5' GCGCGAATTCTTATCTAGTGACCTGAACTTCCTGG 3'

Table 1: Primers used in PCR Reactions

ATGAACCCCGCGGGCATGGTCAGACTGTCCGAGGCATGCAAATGGCAATTCAGCGAAGTCCAGATCG
 AGACAGGTGCTCATCGGGGAGCCGGACGCTGACGGCAGTCTGAACGACAATAACATGCATCCCGGCCACGAGCT
 TTCCGACTTGCCTAATGTGCAGCGCAGCCGGCTGGCGAGGCGGTGGTCCACTATGCGCGTTTCGCGCAGATAACC
 CTCCTTTCCCGATCATCCGGAAGCCGAGCCGATAGCCTCCGTTCTGCGCCGCCGATCGATGCGGCGGTGGAAAA
 GCAGGAAGACGAGAGGAAGCCCATGCGGCGACGGATGGCGCTGACTTGTATTGGTTCTTTTTCATGCAGG
 TCTGGTGGTTGCTCTTAGCGTGTTTATGCCGACGCCCGAACGAAGCGATTGAAGATGCTGGTGAGGCTGTCAG
 CGTCGTCGTGTATGGCGACTCTGACGTTGACCAGACATCGGCCGGCGATCCAGAGTTAGAGCGCCAGCCCGAGCA
 AGTTGCCTCGGAGGAGGTTGAGCCGGATACAGTTCAATCGGAGGAGGCCACCGAACTAGCTGCCACGACCGTGC
 CGCCGGAGCAGGCCAGCCGGTCGAAACTGAAACGATGGAGACCGTTCAGTCGGTGACGAAAGTTACGCGGGT
 TTCCCTGAGGCCGTTGTTGCCCGGAGCCAGAGGTGCTTGTCTCCGAGAGCCCTGCAGAAATCTTTCGTCGCTCAG
 CCGATGGCAACGGCTGTGCCCGAACAGCCCATGCCTGATATCGCGCAGGCAACAGTACCGGAAGAAGTAGTGCC
 AACGGCTGTACAGCCACCGCAGTTCCGCCGGAAGAAGTAAAGCCTGTCGAGACTGTGAAAGTGTGCGCCGAACC
 TGAAGACAAACCGAAGCCGAAGGCTAAAAAACCGAAACCGTTGAAAAGAAGCAGTCTCAGAAGAGGGCCAAG
 CCTGCTGGTGGAAAGAGGGCTCAGATCGTGAGGACTCGACGAGGGGCATGGTGAACGGCCAGGACGGGCCG
 AAAGTACGGCACGTCAACCACGACAGGCGGAACTGACGGTATGGGAAGTGCTGCCGTAGCCAACTACCCTGGT
 AAGATACAGAAGCGGATCCGTCGTGCTGTTTCGCGTGCCTGATGAATACAAAACAAGAGCGGCGGTATGACCGTA
 AGGATTCAATTGACGATTAACGGAACAGGCAGGGTAGCGTCCGGTCTCTGTTGCTCGCAGTTCGGGTATTGCAGAA
 TTGGATAAAGCTGTTCTGGATGGTGTGCGGCGCGCGGCCGTTTCCGCCGCTGCCCTCGGAGTGGGGTAAGCCC
 AGCTGGACTTTCGCACAGGAAGTTCAGGTCAGTAA

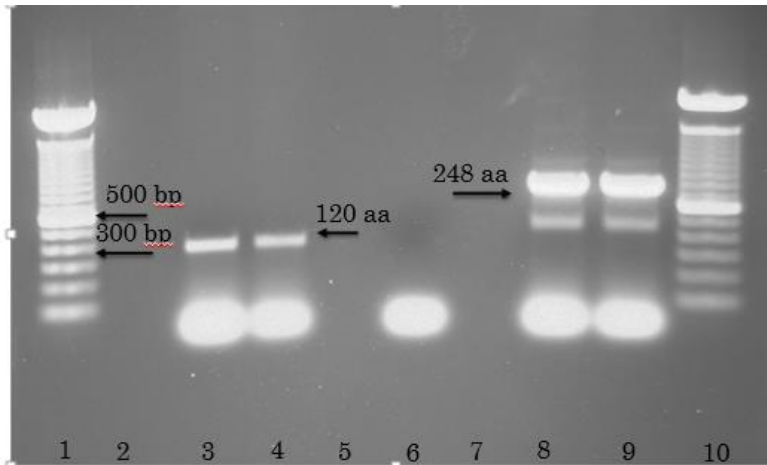


Figure 4: 1% Agarose gel results from 120 amino acid and 248 amino acid PCR reactions

In order to visualize the results from the PCR reaction, a 1% agarose gel was run with the various samples. As with the genomic DNA, TAE was utilized as the buffer, ethidium bromide as the intercalating agent, and UV light as the means of visualization. All samples were loaded in 25µl increments with 5µl of loading dye. Lanes 1 and 10 are NEB 100 base pair ladders used as standards. Lanes 3 and 4 are the 120 amino acid PCR reaction. Lane 6 is the control, and lanes 8 and 9 are the 248 amino acid PCR reaction. The control was performed under the same conditions as the samples and contained the same components with the exception of genomic DNA. Each amino acid is coded for by three nucleotides. Hence, the 120 amino acid fragment of C-terminal TonB should be 360 base pairs, and the 248 amino acid fragment should be located at 744 base pairs. The agarose gel results, when compared to the ladder standard, indicated the proper sizes. The 248 amino acid polymerase chain reaction did yield an additional minor fragment at approximately 400 base pairs. When the 248 amino acid PCR was repeated for subsequent procedures, the thermocycler times and temperatures were modified. However, the fragment was still consistently produced. Regardless, the DNA was removed from the agarose gel at the appropriate location of 744 base pairs and purified to ensure future accuracy.

Purification of PCR amplified DNA

In order to extract the DNA from the agarose gel used in PCR confirmation, the Ultraclean 15 DNA purification kit was utilized (MO BIO Laboratories Carlsbad, California). Two samples of the 120 amino acid PCR and 2 samples of the 248 amino acid PCR were loaded into the agarose gel; therefore, two bands were removed from the gel for purification for each fragment of interest. After purification, the NanoDrop was utilized to assess the concentration of the DNA recovered. The results are shown in table 2.

NanoDrop Results of Purified PCR				
	120 amino acid sample 1	120 amino acid sample 2	248 amino acid sample 1	248 amino acid sample 2
Concentration (ng/µl)	172.2	218.0	189.5	201.1
A 260	3.448	4.360	3.789	4.023
A 280	1.706	2.202	1.804	1.734
260/280	2.02	1.98	2.10	2.32

Table 2: Nanodrop Results of Purified PCR from 1µl of sample

The 260/280 ratio was higher than anticipated for the samples; however, because of the ample concentrations, the purified DNA was kept stored and at 4°C for future restriction enzyme digestion.

Isolation of pET-17b Results

The next step of the cloning process involved isolation of the plasmid chosen for experimentation. pET-17b was isolated from culture using a QIA Prep Spin Miniprep Kit. After following the procedure, the results were assessed for both concentration and accurate size of the DNA. The concentration was measured using the NanoDrop. The results are given in table 3.

NanoDrop Results for pET-17b isolation	
Concentration (ng/μl)	290.4
A 260	5.809
A 280	3.105
260/280	1.87
260/230	2.19

Table 3: NanoDrop Results of Purified PCR from 1μl of sample

In addition to measuring the concentration of DNA isolated, a 1% agarose gel was also ran to ensure the DNA was of the appropriate size corresponding to the plasmid. The gel was made as described previously. The first lane contains a 1kb standard ladder. The last lane contains the 5μl of the isolated pET-17b combined with 1μl of loading dye. It should be noted the lanes in-between contained saved samples from various earlier experiments that were unsuccessful. As shown in figure 1 in materials and methods, pET-17b is 3306 base pairs. The isolated plasmid was approximately that size as observed from figure 5 below.

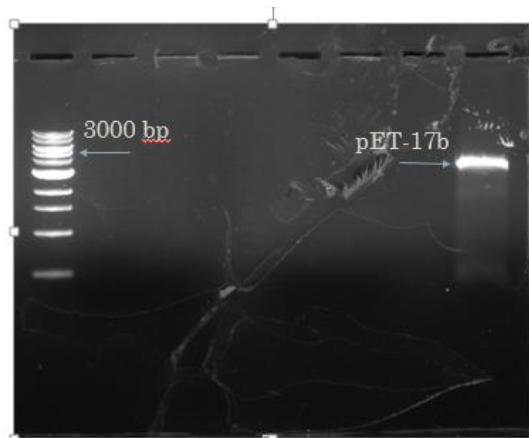


Figure 5: 1% agarose gel of isolated pET-17b

Restriction Enzyme Digest Results

Restriction enzyme digestion reactions were performed for the purified 120 amino acid fragment DNA, the 248 amino acid fragment DNA, and the pET-17b plasmid. A 1% agarose gel was used as a means to assess the results and ensure only the DNA of appropriate size was utilized in the preceding ligation reaction. The results are shown in figure 6.

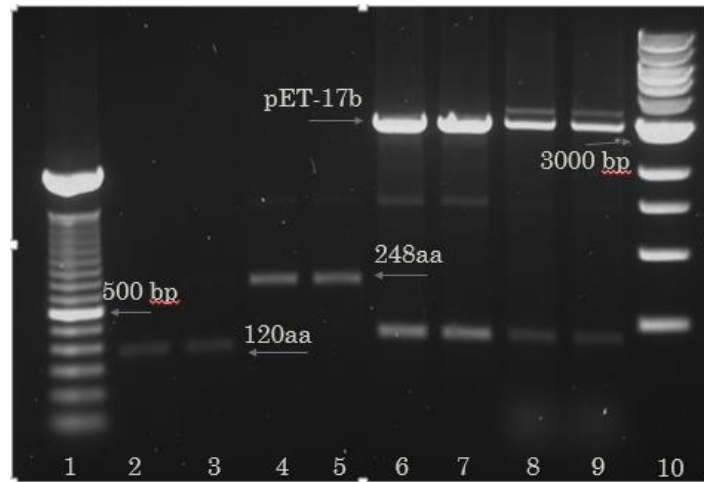


Figure 6: 1% agarose gel of restriction enzyme digestion reactions

In the figure above lane 1 contains the 100bp ladder. Lanes 2 and 3 contain the restriction enzyme digest from the 120 amino acid fragment DNA. Lanes 4 and 5 contain the restriction enzyme digest from the 248 amino acid fragment DNA. Lanes 6, 7, 8, and 9 contain the restriction enzyme digestion of pET-17b, and lane 10 contains the 1kb ladder. The DNA from the 120 amino acid fragment TonB corresponds to the appropriate 360 base pairs. The DNA from the 248 amino acid fragment TonB corresponds to the appropriate 744 base pairs, and the DNA from pET-17b corresponds to the approximate 3,257 base pairs. While the size difference would be extremely difficult to identify, pET-17b after restriction enzyme digestion should be 3,257 base pairs rather than the original 3,306 base pairs. There are a total of 49 base pairs located between the two restriction enzymes sites that should be excised out during the reaction. After confirmation of size, each band was cut out and purified from the agarose by the Ultraclean 15 DNA purification kit. The NanoDrop results from purification are shown in table 4. The concentration was deemed high enough to continue the cloning procedure with ligation.

NanoDrop Results from Restriction Enzyme Digestion				
	120 amino acid fragment	248 amino acid fragment	pET-17b sample 1	pET-17b sample 2
Concentration (ng/μl)	240.9	88.0	143.5	102.6
A 260	4.818	1.760	2.870	2.052
A 280	2.546	0.613	1.280	0.857
260/280	1.89	2.81	2.24	2.39

Table 4: NanoDrop results from purification of agarose gel of restriction enzyme digestion reactions

Transformation Results

The next step in expressing the two TonB fragments of interest from *Rhizobium leguminosarum* ATCC 14479 was to transform the ligated plasmid into competent cells. The competent cells chosen for transformation were *Escherichia coli* BL21 (DE3). *E. coli* BL21 (DE3) cells contain the DE3 lysogen necessary for the T7 RNA polymerase. The T7 RNA polymerase transcribes the genes located downstream of the T7 promoter in the presence of IPTG in pET-17b. The cloning region, and thus the

inserted gene fragments, are located within this region. Consequently, *E. coli* BL21 (DE3) cells are the proper cell choice for the plasmid utilized (Biolabs 2017).

Transformation of both plasmids from ligation was successful. Colonies of the competent cells containing the plasmid with the 120 amino acid insert were grown on one LB ampicillin plate, and colonies of the competent cells containing the plasmid with the 248 amino acid insert were grown on the other LB ampicillin plate. *E. coli* BL21 (DE3) is typically susceptible to the antibiotic ampicillin. Uptake of the plasmid, however, gives the bacteria ampicillin resistance. Therefore, growth on plates with ampicillin verified the cells had the plasmid enclosed. To ensure the plasmid contained the insert of interest, the plasmids were re-isolated from the competent cells. The procedure was followed by restriction enzyme digestion. The results from restriction enzyme digestion of the transformation containing pET-17b with the 120 amino acid fragment are shown in figure 7, and results from the restriction enzyme digestion of the transformation containing the pET-17b with the 248 amino acid fragment are shown in figure 8 and figure 9.

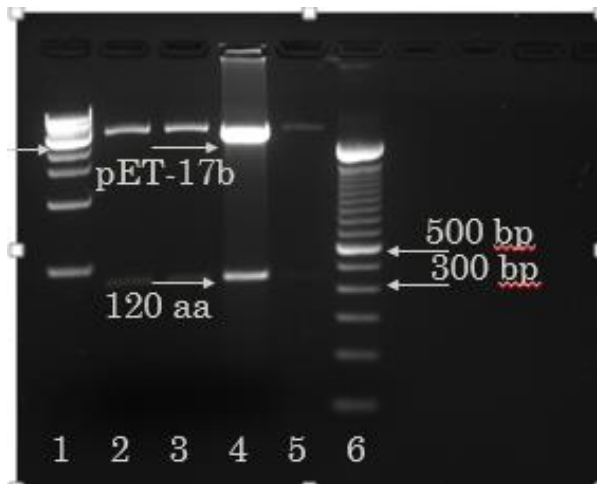


Figure 7: 1% agarose gel of restriction enzyme digestion from the transformation of pET-17b with the 120 amino acid insert.

From the figure above successful transformation of the pET-17b with the 120 amino acid insert from C-terminal TonB was confirmed. Lane 1 contains a 1kb ladder to assess the size of the plasmid. Lanes 2 through 5 contain the restriction enzyme digestion from the plasmid isolation, and lane 6 contains a 100 base pair ladder. The restriction enzymes were utilized to cut the insert out of the cloning region of the plasmid. The two bands observed, therefore, in lanes 2 through 5 represent the plasmid and the 120 amino acid DNA insert separated. pET-17b after restriction enzyme digestion should be approximately 3,257 base pairs, and the 120 amino acid DNA insert should be 360 base pairs in size. Their locations on the agarose gel correspond to the predicted sizes. For further verification, however, the isolated plasmid with the insert was sent for sequencing.

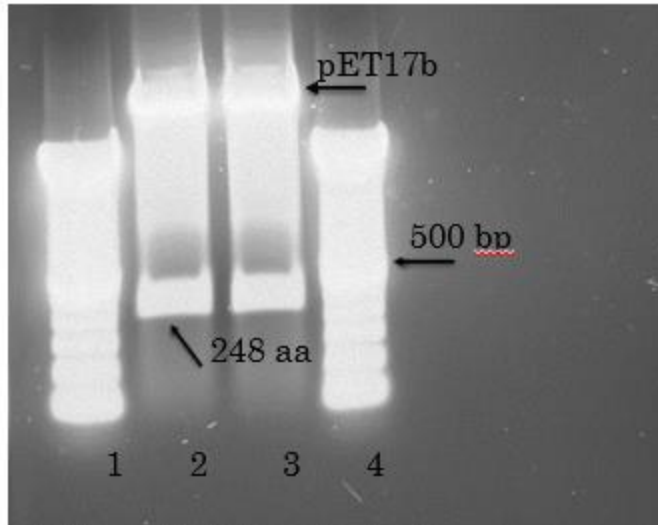


Figure 8 : 1% agarose gel of restriction enzyme digestion from the transformation of pET-17b with the 248 amino acid insert

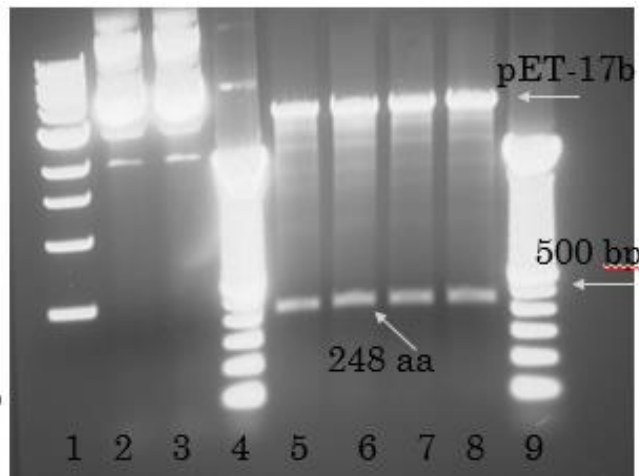


Figure 9: 1% agarose gel for the repeated restriction enzyme digestion from the transformation of pET-17b with the 248 amino acid insert

In figure 6, the first and fourth lanes contain a 100 base pair ladder. The second and third lanes contain the restriction enzyme digestion from the isolated plasmid with the 248 amino acid insert. The restriction enzymes should have removed the 248 amino acid insert from the plasmid. As a result, two DNA fragments should be visualized in the agarose gel. One of the fragments should be approximately 3,257 base pairs corresponding to the linear pET-17b, and the other fragment should be 744 base pairs corresponding to the 248 amino acid C-terminal domain of TonB. In figure 6, lanes 2 and 3 do have two DNA fragments. The larger of which can be assumed to be pET-17b. The smaller fragment, however, was not indicative of the appropriate size. Instead of the predicated 744 base pairs, the fragment appeared to be between 300 and 400 base pairs. Due to the conflicting results, the plasmid was re-isolated from a separate colony and restriction enzyme digestion was repeated. Figure 7 gives the subsequent data produced. In the 1% agarose gel lane 1 contains the 1kb ladder. Lanes 2 and 3 contain the isolated

plasmid from transformation. Lanes 4 and 9 contain a 100 base pair ladder, and lanes 5 through 8 contain the restriction enzyme digestion. The repeated procedure yielded the same results. The insert size was not 744 base pairs as anticipated. In order to gain more information, the isolated plasmid was sent for sequencing to the core facility of Quillen College of Medicine, ETSU.

Sequencing Data Results

The sequencing results for the 120 amino acid C-terminal TonB from *Rhizobium leguminosarum* ATCC 14479 are shown below. The sequence located within the pET-17b plasmid was confirmed as the *Rhizobium leguminosarum* TonB gene when compared to known data. The similarity was 97%. Thus, cloning was deemed successful for the fragment, and subsequent protein expression was pursued.

Rhizobium leguminosarum tonB gene, hmu operon (hmuPV genes) and rpoZ gene
 Sequence ID: [emb|AJ310723.1|](#) Length: 11382 Number of Matches: 1

Range 1: 1034 to 1469 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
739 bits(400)	0.0	425/437(97%)	1/437(0%)	Plus/Minus
Query 210	GGTGAACGGCCAGGACGGGCCGCAAACTGACGGCACGTCAACCACGACAGGCGGAACTGA			269
Sbjct 1469	GGTGAACGGCCAGCACGGGCCGCAAACTGACGGCACGTCAACCACGACAGGCGGAACTGA			1410
Query 270	CGGTATGGGAAGTGCTGCCGTAGCCAACCTACCTGGTAAGATACAGAAGCGGATCCGTCG			329
Sbjct 1409	TGGTATGGGAAGTGCTGCCGTAGCCAACCTACCTGGTAAGATACAGAAGCGGATCCGTCG			1350
Query 330	TGCTGTTGCGCTGCCTGATGAATACAAAAACAAGAGCGGCGGTATGACCGTAAGGATTCA			389
Sbjct 1349	CGCTGTTGCGCTGCCTGATGAATACAAAAACAAGAGCGGCGGTATGACCGTAAGGATTCA			1290
Query 390	ATTGACGATTAACGGAACAGGCAGGGTAGCGTCCGGTCTCTGTTGCTCGCAGTTCGGGTAT			449
Sbjct 1289	ATTGACGATTAACGGAACAGGCAGGGTAGCGTCCGGTCTCTGTTGCTCGCAGTTCGGGTAT			1230
Query 450	TGCAGAAATGGATAAAGCTGTTCTGGATGGTGTGCGGCGCGCGCCGTTTCCGCCGCT			509
Sbjct 1229	TGCAGAGTTGGATAAAGCTGTTCTGGATGGTGTGCGGCGCGCGCCGTTTCCGCCGCT			1170
Query 510	GCCCTCGGAGTGGGGTAAGCCCAGCTGGACTTTCGCCAGGAAGTTCAGGTCACTAGATA			569
Sbjct 1169	ACCCTCGGAGTGGGGTAAGCCTAGCTGGACTTTCACCCAGGAAGTTCAGGTCACTAGATA			1110
Query 570	ATATGATAAAAAAATCAACTTTATACTTTACTCTAAAAATCAAAGTTAATAAGGTCCG			629
Sbjct 1109	ATATGATAAAAGTAATCAACTTTATACTTTACTCTAAAAATCAA-GTTTAATAAGGTCCG			1051
Query 630	CTCGCACACGCAGGAAT	646		
Sbjct 1050	CTCGCACACGCAGGAAT	1034		

The sequencing data for the intended 248 amino acid C-terminal TonB from *Rhizobium leguminosarum* ATCC 14479 is shown below. Sequencing was only successful for the reverse reaction. As indicated by the restriction enzyme digestion results, the fragment was not of the appropriate size. Approximately 400 base pairs were missing. The fragment, however, was consistent with the gene of interest. It was confirmed to be *Rhizobium leguminosarum* TonB gene. The reverse sequence TonB alignment also confirmed the fragment was a portion of the C-terminal domain. Because the forward sequencing was unsuccessful and same end of the fragment was not cloned, the problem may be

associated with the forward primer design. The sequencing results indicated further analysis, repeated procedures, or redesign of experimentation for the larger fragment of the C-terminal TonB was necessary before proceeding with protein expression.

Rhizobium leguminosarum tonB gene, hmu operon (hmuPV genes) and rpoZ gene

Sequence ID: [AJ310723.1](#) Length: 11382 Number of Matches: 1

Range 1: 1149 to 1469 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
544 bits(294)	6e-154	312/321(97%)	0/321(0%)	Plus/Plus
Query 1	GGCTTACCTCACTCCGAGGGCAGCGGCGGAAACGGCGCCGCGCGCCGACACCATCCAGA	60		
Sbjct 1149	GGCTTACCCCACTCCGAGGGTAGCGGCGGAAACGGCGCCGCGCGCCGACACCATCCAGA	1208		
Query 61	ACAGCTTTATCCAATTCTGCAATACCCGAACTGCGAGCAACAGAGACCGACGCTACCCTG	120		
Sbjct 1209	ACAGCTTTATCCAACTCTGCAATACCCGAACTGCGAGCAACAGAAACCGACGCTACCCTG	1268		
Query 121	CCTGTTCCGTTAATCGTCAATTGAATCCTTACGGTCATACCGCCGCTCTTGTTTTTGTAT	180		
Sbjct 1269	CCTGTTCCGTTAATCGTCAATTGAATCCTTACCGTCATACCGTCGCTCTTGTTTTTGTAT	1328		
Query 181	TCATCAGGCACGCGAACAGCAGCAGCGGATCCGCTTCTGTATCTTACCAGGGTAGTTGGCT	240		
Sbjct 1329	TCATCAGGCACGCGAACAGCAGCAGCGGATCCGCTTCTGTATCTTACCAGGGTAGTTGGCT	1388		
Query 241	ACGGCAGCACTTCCCATAACCGTCAGTCCCGCTGTGCGTGGTTGACGTGCCGTGAGTTTGC	300		
Sbjct 1389	ACGGCAGCACTTCCCATAACCATCAGTCCCGCTGTGCGTGGTTGACGTGCCGTGAGTTTGC	1448		
Query 301	GGCCCCGTCCTGGCCGTTCCACC	321		
Sbjct 1449	GGCCCCGTGCTGGCCGTTCCACC	1469		

Reverse sequence (after converting to reverse complement) TonB248_Alignment

EMBOSS_001	1	GGGTTCAAACCTTTATCGAAACACCGGAAACCGAAAACCAAATTCAATGTT	50
EMBOSS_001	1	-----ACGATGGAGACCG-----TTCA----- 	17
EMBOSS_001	51	GTTGGTTCAG---GTT-----GCANACGTT-TTGCGAG	78
EMBOSS_001	18	GTCGGTGCAGGAAGTTACGCGGGTTTCCCTGAGGC---CGTTGTTGCCG . .	64
EMBOSS_001	79	CAG---CAGTNGGTTACGTTTGCTTGGGTATCGGTGATTCAT-----	118
EMBOSS_001	65	CCGAGCCAGAGG-----TGCTT--GTCTCCGAGAGCCCTGCAGAAT	103
EMBOSS_001	119	-TTTGGTAAC-CAG--TAAGGCAAC-----CCCG-CCAGCCTA-GCCG	155
EMBOSS_001	104	CTTTCGTCGCTCAGCCGATGGCAACGGCTGTGCCCGAACAGCCCATGCCT . . .	153
EMBOSS_001	156	GGTCCCAACGACGGGAGCAGATCATGCGCACCCGTGGC--CAG----G	199
EMBOSS_001	154	GAT-----ATC--GCGCA-----GGCAACAGTACC	177
EMBOSS_001	200	GA-----CCCAACG-----CTGCCCAGATCTCGATCCCGC----	230
EMBOSS_001	178	GAAGAAGTAGTGCCAACGGCTGTACAGCCC---ACCGAGTTCCGCCGGA	224
EMBOSS_001	231	-GAAATTAATAC-----GACTCACTATAG-----GGAGACC-----A	261
EMBOSS_001	225	AGAAGTAAAGCCTGTGCGAGACTGTGCGA-AGTGTGCGCCGGA-ACCTGAAGA	272

```

EMBOSS_001      262 CA-----ACGGTTCCCTCTAGAAA---TAATTTGTTTAACTT      297
      ||           |.||           |||.|||   .|.|...|||.||
EMBOSS_001      273 CAAACCGAAGCCGAAGG-----CTAAAAAACCGAAACCGGTTGAA---      312

EMBOSS_001      298 TAAGAAGGAG---ATATACATATGG-CTAGCATGACTGGTGGACAGCAA      342
      |||||.||   |.|   |.|.|| |.|||.|| |||||   |.||
EMBOSS_001      313 -AAGAAGCAGTCTCAGA---AGAGGGCCAAGCCTG-CTGGTGG---GAAA      354

EMBOSS_001      343 ATGGG-----TCG---GGATTCAA-----GCTTGATGGTGAACGGCCA      377
      ..|||   |||   |||.||.||   ||   |||||
EMBOSS_001      355 GAGGGCTCAGATCGTGAGGACTCGACGAGGGGC---ATGGTGAACGGCCA      401

EMBOSS_001      378 GGACGGGCCGCAAACCTGACGGCACGTCAACCACGACAGCGGAACTGACG      427
      |||||
EMBOSS_001      402 GGACGGGCCGCAAACCTGACGGCACGTCAACCACGACAGCGGAACTGACG      451

EMBOSS_001      428 GTATGGGAAGTGCTGCCGTAGCCAACCTACCTGGTAAGATACAGAAGCGG      477
      |||||
EMBOSS_001      452 GTATGGGAAGTGCTGCCGTAGCCAACCTACCTGGTAAGATACAGAAGCGG      501

EMBOSS_001      478 ATCCGTCGTGCTGTTTCGCGTGCCTGATGAATACAAAAACAAGAGCGGCGG      527
      |||||
EMBOSS_001      502 ATCCGTCGTGCTGTTTCGCGTGCCTGATGAATACAAAAACAAGAGCGGCGG      551

EMBOSS_001      528 TATGACCGTAAGGATTCAATTGACGATTAACGGAACAGGCAGGGTAGCGT      577
      |||||
EMBOSS_001      552 TATGACCGTAAGGATTCAATTGACGATTAACGGAACAGGCAGGGTAGCGT      601

EMBOSS_001      578 CGGTCTCTGTTGCTCGCAGTTCGGGTATTGCAGAATTGGATAAAGCTGTT      627
      |||||
EMBOSS_001      602 CGGTCTCTGTTGCTCGCAGTTCGGGTATTGCAGAATTGGATAAAGCTGTT      651

EMBOSS_001      628 CTGGATGGTGTGCGGCGCGGGCGCCGTTTCCGCCGCTGCCCTCGGAGTG      677
      |||||
EMBOSS_001      652 CTGGATGGTGTGCGGCGCGGGCGCCGTTTCCGCCGCTGCCCTCGGAGTG      701

EMBOSS_001      678 AGGTAAGCC-----      686
      .|||
EMBOSS_001      702 GGGTAAGCCCAGCTGGACTTTCGCCCAGGAAGTTCAGGTCACTAGATAA      750

```

Results for Determination of Optimal IPTG Concentration

Attempted expression of the TonB C-terminal domain from *Rhizobium leguminosarum* ATCC 14479 was continued for the 120 amino acid fragment. After successful transformation, the next step was to determine the optimal concentration of isopropyl β -D-1-thiogalactopyranoside required for protein expression. As shown in figure 3, pET-17b has a T7 promoter located upstream of the cloning region. When IPTG is present in a sufficient concentration, the lac repressor falls off the operator site. As a result, the RNA polymerase from *E. coli* BL21 (DE3) cells binds to the T7 promoter. Transcription of the cloning region then occurs, and ultimately leads to translation of the inserted DNA sequence into protein (Bell 2000). Four different concentrations of IPTG were analyzed, 0.1mM, 0.5mM, 1.0mM, and 2.0mM. Before the IPTG was added, however, each culture was grown to an optimal density of 0.4 to 0.6 as measured by spectrophotometry. The results of growth overtime are shown in table 5 below. After seven hours and twenty minutes, all four 50ml cultures had reached OD and the varying concentrations of IPTG were added.

Growth to Optimal Density				
Time (minutes)	Flask 1 Cell Density	Flask 2 Cell Density	Flask 3 Cell Density	Flask 4 Cell Density
0	0.077	0.087	0.068	0.088
2:20	0.076	0.076	0.099	0.079
5:00	0.084	0.088	0.108	0.092
6:40	0.252	0.253	0.275	0.264
7:20	0.623	0.645	0.689	0.84

Table 5: Growth to optimal density for *E. coli* BL21 (DE3) from successful transformation of pET-17b with the 120 amino acid C-terminal TonB.

After three hours of IPTG induction with various concentrations, the cells were harvested and pellet. In order assess the results, the controls taken before induction and small aliquots of pelleted cells after induction, were prepared into samples for SDS page gel electrophoresis. There were a total of eight samples, one from each culture before IPTG was added and one after three hours had elapsed. The eight pelleted samples were resuspended in Tris buffer. Next, 10µl of each sample were combined with 10µl of SDS loading dye. The samples were then placed in a 95°C water bath for five minutes. Afterwards, 10µl of each was added to a prepared 10% SDS gel. The percentage was chosen based on the small protein size. The gel was ran at 120 volts for 70 minutes. The results are shown in figure 10.

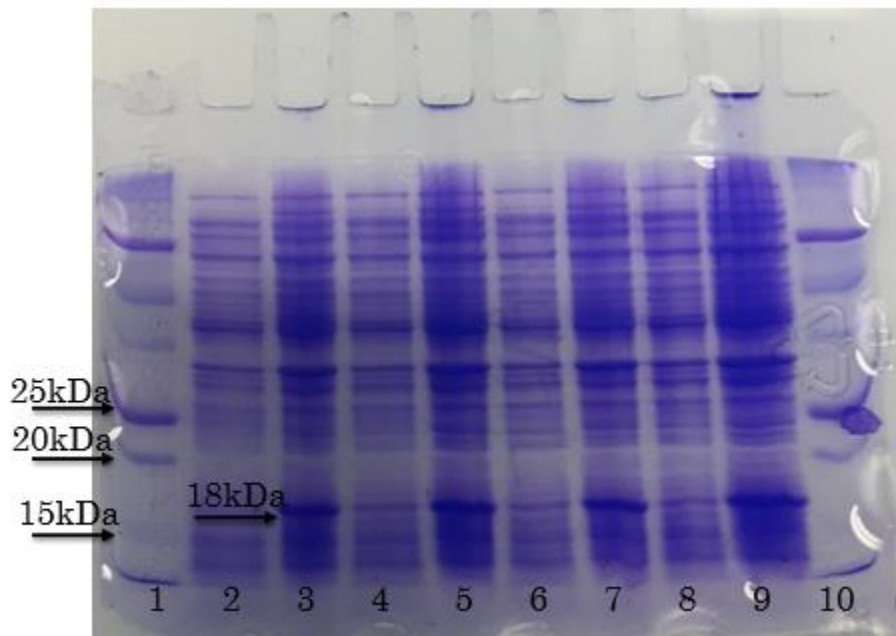


Figure 10: 10% SDS page gel results for determination of IPTG induction concentration

In the figure above lanes 1 and 10 contain the ladder Precision Plus Protein Dual Color Standards. Lane 2 contains the uninduced sample from flask 1. Lane 3 contains flask 1 induced with 0.1mM IPTG. Lane 4 contains the uninduced sample from flask 2. Lane 5 contains flask 2 induced with 0.5mM IPTG. Lane 6 contains the uninduced sample from flask 3. Lane 7 contains flask 3 induced with 1.0mM IPTG. Lane 8 contains the uninduced sample from flask 4, and lane 9 contains flask 4 induced with 2.0mM IPTG. The 120 amino acid C-terminal TonB from *Rhizobium leguminosarum* should be located at approximately 18 kDa, as determined by previous experimentation (Barisic 2015). The TonB C-

terminal protein is approximately 16.6kDa. Because it is expressed in pET-17b, however, the T7 tag should be attached to the N-terminus of the protein. The T7 tag is upstream of the cloning region, and is transcribed with the insert by the T7 polymerase. The T7 tag is approximately 11 proteins corresponding to 1.2kDa. Therefore, with the T7 tag, the protein of interest has a size of 17.8 kDa. From the SDS gel, distinct differences in band width can be observed at that location. For all four concentrations of IPTG, more protein was produced and visualized after induction versus the controls. The optimal concentration of IPTG was determined to be 0.5mM. A concentration of 0.1mM resulted in protein expression; however, it was thought that such a small concentration could cause less protein production in larger scale experiments. Concentrations of 1.0mM and 2.0mM IPTG also clearly resulted in protein expression. The additional IPTG, however, was deemed unnecessary. All further experimentation utilized an IPTG concentration of 0.5mM to induce expression of the 120 amino acid C-terminal TonB.

Expression of the 120 amino acid C-terminal TonB Results

In attempt to produce an adequate amount of protein for purification, 4 liters of the successful transformation cell culture were grown to optimal density. Approximately six hours elapsed between inoculation and a concentration of 0.4 to 0.6. The data collected is shown in table 7.

Growth to Optimal Density				
Time (minutes)	Flask 1 Cell Density	Flask 2 Cell Density	Flask 3 Cell Density	Flask 4 Cell Density
0	0.098	0.095	0.092	0.098
3:30	0.199	0.251	0.209	0.234
5:00	0.311	0.354	0.329	0.343
6:00	0.598	0.632	0.621	0.610

Table 7: Growth to optimal density for *E. coli* BL21 (DE3) from successful transformation for protein expression

After induction with 0.5mM IPTG for three hours, SDS-PAGE electrophoresis was used to confirm protein expression before experimentation was continued. In order to validate IPTG was causing over production of the 120 amino acid TonB, the control samples before induction were compared to samples taken after induction. All pelleted samples were resuspended in Tris buffer. Ten μ l of each was added to 10 μ l of SDS loading dye and placed in a hot water bath at 95°C for 5 minutes. Each well was loaded with 10 μ l of the sample or ladder, and the gel was ran for 75 minutes at 120 volts. The results are shown in Figure 11.

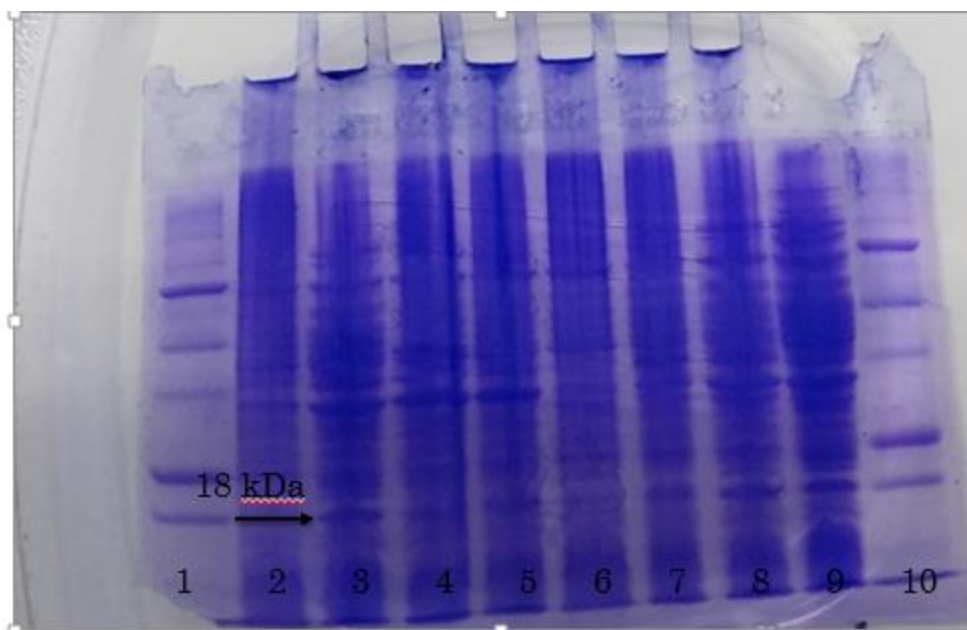


Figure 11: 10% SDS-PAGE gel of 0.5mM IPTG induction

For the SDS-PAGE gel in figure 11, lanes 1 and 10 represent the ladder Precision Plus Protein Dual Color Standards. Lanes 2 and 6 are the controls without IPTG and the remainder of the lanes are the induced cells. From the results, protein expression was deemed successful. There is clear evidence of a higher protein concentration in the induced cells compared to the uninduced. There is a distinct band at approximately 18kDa corresponding to the 120 amino acid C-terminal TonB with T7 Tag.

Following confirmation of induction, the cytoplasmic proteins were collected using sonification and centrifugation. From the ultracentrifuge, the pellet formed contained the cell membrane proteins, while the supernatant was composed of the cytoplasmic proteins. The supernatant was collected in aliquoted proportions and assumed to contain the 120 amino acid C-terminal TonB. While the functional TonB protein is located within the inner membrane, the fragmented TonB being expressed through integration of the plasmid should be present in the cell cytoplasm. The amino terminus is required by the Sec proteins to insert TonB into the inner membrane. The portion of the protein being expressed, however, is from the carboxyl terminus. The amino terminus is not included within the 120 amino acids, and therefore, the protein will not be lodged within the membrane (Mori 2001). Hence, the supernatant was utilized for further analysis and purification.

Protein Purification Results

Successful induction was followed by an attempt to purify the protein for analysis using Novagen T7-Tag affinity purification kit. The basis of the column is reliant upon the notion that the T7-Tag present on the 120 amino acid C-terminal TonB protein should bind to the resin when passed through the column. All other cytoplasmic proteins should leave the column as flow-through. Then, when the elution buffer is added, the protein of interest should be eluted into concentrated samples. A portion of each sample was combined with an equal allotment of SDS loading dye. The samples were placed in a 95°C hot water bath for 5 minutes. Each was then loaded in 5µl increments into a 10% SDS-

PAGE gel. The gel was ran at 120 volts for 75 minutes. The results from the elution are shown in figure 12 below.

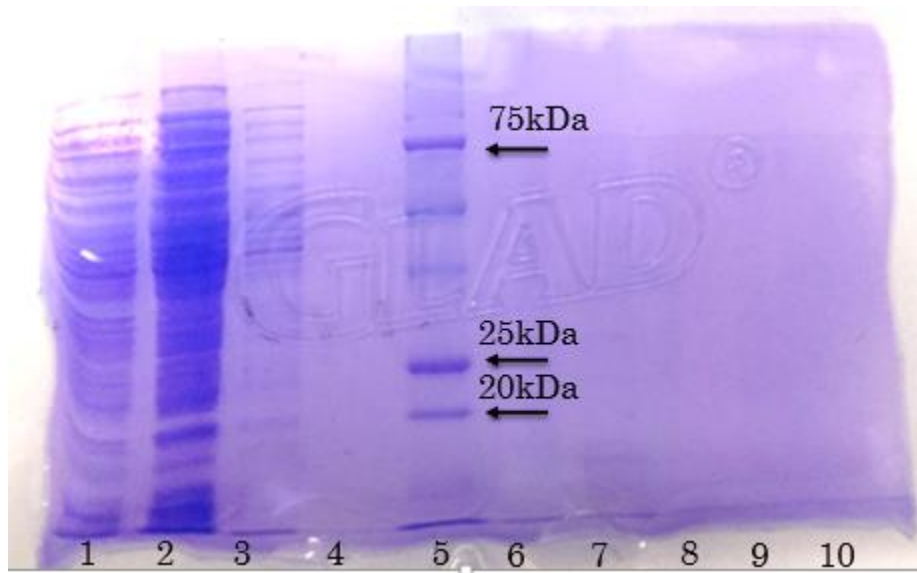


Figure 12: 10% SDS-PAGE gel of T7-Tag affinity purification

In the SDS-PAGE gel in figure 10, lane 1 contains the control taken of the supernatant before the column was utilized. Lane 2 contains the flow through from the sample. Lanes 3 and 4 contain the subsequent collection of the T7-Tag bind/wash buffers. Lane 5 contains the ladder Precision Plus Protein Dual Color Standards, and lanes 6 through 10 contain samples of the five serial elutions. The results ultimately indicated the T7-Tag affinity column was not successful in purification of the 120 amino acid C-terminal TonB expressed. Lane 1 ensures the gel was ran properly and cytoplasmic protein was present before the sample was passed through the column. Lanes 3 and 4 confirm the protein of interest was not washed away in the T7-Tag bind/wash buffer. Lanes 6 through 10, however, should show a distinct band at approximately 18kDa. Yet, no protein is visualized at all, signifying error within the column or within the protocol design. The T7-Tag column had been utilized previously. There is a possibility that the column resin was no longer effective. Also, although the protein size was confirmed after induction, it is conceivable that the T7-Tag was not attached to the 120 amino acid C-terminal TonB. The small difference of 11 additional proteins cannot be accurately assessed by the broad spectrum of the standard ladders and a SDS-PAGE gel. Although unlikely, another possibility remains where the protein assumed to be the 120 amino acid C-terminal TonB is in reality another protein induced by IPTG with a similar size. In order to assess the results, the procedure must be repeated utilizing a new T7-Tag affinity column.

Discussion and Conclusion

A portion of the C-terminal TonB of *Rhizobium leguminosarum* ATCC 14479 was successfully cloned into the plasmid pET-17b and transformed into *E. coli* BL21 (DE3). The 120 amino acid fragment was chosen based on its alleged interaction with the TonB box of the outer membrane receptor and its vital role in energy transduction. From the putative TonB gene sequence, primers were specifically designed to amplify this region of interest. Following a successful polymerase chain reaction, the 120 amino acid C-terminal TonB was prepared for insertion into a plasmid through restriction enzyme

digestion and then ligated into the plasmid pET-17b. Subsequently, the plasmid was cloned into competent *E. coli* BL21 (DE3) cells. Growth after transformation was verified by re-isolating the plasmid and checking the insert within the cloned region for appropriate size. The 120 amino acid C-terminal TonB's presence within the plasmid was confirmed in the successful transformation through sequencing and comparison to known data. The optimal IPTG concentration to induce the *E. coli* BL21 (DE3) cells containing the plasmid with the gene of interest was then determined. An IPTG concentration of 0.5mM for three hours of induction was concluded to be sufficient for protein expression. A large quantity of the cells from the successful transformation were then grown, induced with 0.5mM IPTG, sonicated, and centrifuged to collect the fragment of TonB being produced. Purification of the protein was then attempted through utilization of a T7-tag affinity column. The purification of the protein, however, was unsuccessful as indicated by SDS-PAGE electrophoresis.

In order to collect enough protein for structure analysis, protein expression and purification must be repeated. The *E. coli* BL21 (DE3) cells with the ligated pET-17b containing the 120 amino acid C-terminal TonB can be utilized for subsequent procedures, as the inserted gene was confirmed. For following purification, however, a new column should be used to eliminate the possibility of a nonfunctional resin. It is proposed that a Western Blot technique be applied to the SDS-PAGE gel prior to utilization of the affinity column as well to ensure the presence of the T7-Tag. The T7-Tag should be transcribed and translated as a portion of the N-terminus on the 120 amino acid fragment. However, lack of such an occurrence would render the affinity column ineffective, as the column is reliant upon the presence of the tag to differentiate the protein being expressed from other cytoplasmic proteins. The folding of the C-terminal TonB may also interfere with the exposure of the T7-Tag leading to inadequate binding. Further analysis of repeated procedures will yield more conclusive results.

As for the 248 amino acid C-terminal TonB originally chosen for study, amplification of the gene of interest was successful. The designed polymerase chain reaction produced a fragment from the genomic DNA of *Rhizobium leguminosarum* ATCC 14479 of the appropriate size according agarose gel electrophoresis. Subsequent restriction enzyme digestion was also indicated as effective, and as a result, ligation into pET-17b was performed. Following ligation, the plasmid with the insert was transformed into competent cells. Further analysis, however, indicated the gene was not 248 amino acids long as expected. The fragment inserted into the cloning region during ligation was not the entire 248 amino acid C-terminal TonB, rather it was a shortened sequence of the C-terminal domain composed of only 437 base pairs. Consequently, progression to protein expression was not attempted. In order to continue to study the longer portion of the C-terminal TonB, ligation and transformation must be repeated. Similar difficulties, however, have occurred in attempts to express other portions of the TonB gene containing the proline rich region. Previous research was not successful for a similar experimental design containing a 200 amino acid C-terminal length. It is proposed that the addition of the proline-rich region causes a change in the DNA sequence that possibly leads to the formation of secondary structures inhibiting ligation. The procedure must be repeated for further confirmation. Varying the size of the protein to include more or less of the proline rich region may also prove to be beneficial.

Overall, structural analysis of the C-terminal domain of TonB from *Rhizobium leguminosarum* ATCC 14479 is an attainable goal for the future. The framework and design set forth should provide purified protein product when repeated with modifications. Succeeding nuclear magnetic resonance spectroscopy and crystallization of the protein product could lead to increased understanding of the TonB protein. Because the exact mechanism for the transport of energy from the proton motive force is

still unknown, additional information may help gain insight into how the transduction pathway is utilized by gram negative bacteria. The information could ultimately have various applications as many bacterial species essentially use the siderophore mediated acquisition of iron and TonB as a means of survival and reproduction.

References

- Andrews SC, Robinson AK, Rodriguez-Quinones F. 2003. Bacterial iron homeostasis. *FEMS Microbiology Review* 27(2-3): 215-237.
- Barisic, Valeria. "Characterization of Putative ExbB and ExbD Leads to the Identification of a Potential Tol-Pal System in *Rhizobium leguminosarum* ATCC 14479." (2015). Masters thesis, ETSU
- Bell, C.E. & Lewis, M. A closer view of the conformation of the Lac repressor bound to operator. *Nature structural biology* 7, 209-14 (2000).
- Biolabs, New England. "FAQ: What Is the Difference between BL21 and BL21(DE3) Competent E.coli Cells?" What Is the Difference between BL21 and BL21(DE3) Competent E.coli Cells? | NEB. N.p., n.d. Web. 04 Mar. 2017.
- Chakraborty, Ranjan. 2013. Ferric Siderophore Transport via Outer Membrane Receptors of *Escherichia Coli*: Structural Advancement and A Tribute to Dr. Dick van der Helm- an 'Ironman' of Siderophore Biology. *Iron Uptake in Bacteria with Emphasis on E. coli and Pseudomonas*. 11-19.
- Chang C, Mooser A, Pluchthun A, Wlodawer A. 2001. Crystal structure of the dimeric Cterminal domain of TonB reveals a novel fold. *Journal of Biological Chemistry* 276(29): 27535-27540.
- Ferguson AD, Diefenbacher J. 2002. TonB-dependent receptors-structural perspectives. *Biochimica et Biophysica Acta* 1565: 318-332
- Guerinot ML. 1994. Microbial iron transport. *Annual Review Microbiology* 48: 743-772.
- Kimple, Michelle E., Allison L. Brill, and Renee L. Pasker. "Overview of Affinity Tags for Protein Purification." *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* 73 (2013): Unit-9.9. PMC. Web. 3 Mar. 2017.
- Miethke M, Marahiel MA. 2007. Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Review* 71(3): 413-451.
- Mori, Hiroyuki, and Koreaki Ito. "The Sec protein-translocation pathway." *Trends in microbiology* 9.10 (2001): 494-500.
- "PET-17b Vector." TB047 (1998): n. pag. Novagen. Web.
- Volkmar B, Hantke K. 2013. The Tricky Ways Bacteria Cope with Iron Limitation. *Iron Uptake in Bacteria with Emphasis on E. coli and Pseudomonas*: 32-35.
- Wexler M, Yeoman KH, Stevens JB, de Luca NG, Sawers G, Johnston AW. 2001. The *Rhizobium leguminosarum* tonB gene is required for the uptake of siderophore and haem as sources of iron. *Molecular Microbiology* 41(4): 801-816
- Wilson, Kate. "Preparation of genomic DNA from bacteria." *Current protocols in molecular biology* (1987): 2-4.
- Wilson, Kate. "Preparation of genomic DNA from bacteria." *Current protocols in molecular biology* (1987): 2-4.

Appendices

Media and Buffers

Yeast Mannitol Broth

Mannitol 4.0g
K₂HPO₄ 0.2g
MgSO₄ 0.08g
NaCl 0.04g
Yeast Extract 0.4g
ddH₂O qs to 400mL
mix ingredients, adjust pH to 6.8
with NaOH, and autoclave

100mL of 1x TE

2mL 0.5M EDTA (pH 8)
2.5mL 1M Tris (pH 8)
ddH₂O qs to 100mL, autoclave

CTAB/NaCl

4.1g NaCl
80ml ddH₂O
Heat and stir to 65°C to dissolve
ddH₂O qs to 100mL

5M NaCl

14.6g NaCl
ddH₂O qs to 50mL

1M Tris (pH 8)

6g Tris base
25mL ddH₂O
Add HCl to pH 8 and ddH₂O qs to 50mL

TNE Buffer

0.1M Tris-HCl (pH 8.0)
0.15M NaCl
20mM EDTA

TEST-LR Buffer

0.1M Tris (pH 8.0)
20mM EDTA
0.6M Sucrose
1% Triton x-100
24ug/mL lysozyme
0.8ug/ml RNase A

10% SDS

5g SDS
ddH₂O qs to 50mL

24:1 Chloroform: Isoamyl alcohol

24mL chloroform
1ml isoamyl alcohol

25:24:1 Phenol: Chloroform: isoamyl alcohol

25ml phenol
24mL chloroform
1mL isoamyl alcohol

0.5M EDTA (pH 8)

9.5g ethylenediaminetetraacetic acid tetrasodium salt
25mL ddH₂O
Add NaOH to pH 8 and ddH₂O qs to 50ml

SDS-PAGE Solutions

Stacking Gel Buffer (pH 6.8)

0.5M Tris
ddH₂O

Running Gel Buffer (pH 8.8)

1.5M Tris
ddH₂O

Tris-Glycine-SDS Buffer (10x)

Tris	0.25M
Glycine	1.92M
Sodium Dodecyl Sulfate	1.0% (w/v)
ddH ₂ O	to 1L

Coomassie Blue Staining Solution

Coomassie R-2500	0.1% (w/v)
Methanol	50% (v/v)
Glacial acetic acid	10% (v/v)
ddH ₂ O	40% (v/v)

SDS-PAGE Destaining solution

Methanol	50% (v/v)
Glacial acetic acid	10% (v/v)
ddH ₂ O	40% (v/v)

SDS-PAGE gel

	<u>Stacking gel</u>	<u>Separating gel (10%)</u>
30% bis-acrylamide	1.33mL	6.66mL
Stacking gel buffer	2.5mL	-----
Running gel buffer	-----	5.0mL
ddH ₂ O	6.1mL	8.0mL
10% SDS	0.1mL	0.2mL
10% APS	50μL	100μL
TEMED	5μL	10μL