Production of the Bacteriocin Antibiotic Thuricin CD

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

CHEMISTRY

October 2015

By

Charles E. Holjencin

Thesis Committee:

Joseph Jarrett, Chairperson

Kristin Kumashiro

Ho Leung Ng

Acknowledgements

The past three years have been the most difficult but rewarding years of my young life. The idea of moving half-way around the world both thrilled me and scared me, as I did not know how I would cope to such a culture shock. In the end, I feel as if the combination of everything from intense lab work to trying poi for the first time, from snorkeling with whales to running the mass spec all hours of the night has provided me with a lifetime of experience, ultimately making my decision to move to Hawaii the best decision of my life.

I did not gain all this experience alone. In fact, with out the help of my friends and mentors, I probably would have moved home after my first year. First and foremost, I have to thank my parents for supporting me through this endeavor. I love you. Second, I would like to thank my PI, Dr. Joseph Jarrett, for taking me on as a graduate student and putting up with me for the last three years. Tackling this project together has been fun. Dr. Hemscheidt, you were there for me every second of the way and you have provided me with the most valuable professional and life guidance I could ever ask for. Third, I have to thank Julie Cramer for everything she has done for me. I could not have asked for a better co-worker and friend. I would have never completed this project without you. Fourth, I have to thank the members of the Williams Lab for all the time they spend with me. They helped with any and every sample I had without a moments hesitation. Thank you guys! Lastly, I have to thank the family I developed out there. There is no doubt in my mind that we have the closest chemistry department in the world. I had a lifetime's worth of fun with you all and I cannot wait to get back to see you. Thank you all! This acknowledgement would not be complete without thanking my committee members, Dr. Kristen Kumashiro and Dr. Ho Leung Ng. Thank you for taking time out of your busy schedules to aid in the completion of this project!

Abstract

Bacteriocins are ribosomally produced peptide antibiotics produced by bacteria as highly effective antimicrobial natural products. Many bacteriocins are thought to function by selective insertion of multimers into the cell membrane, generating unregulated ion pores that ultimately trigger cell lysis. Bacteriocins often contain stabilizing post-translational modifications including N- and C-terminal additions, epimerized D-amino acids, chemically modified amino acids, and disulfide and thioether crosslinks.

Thuricin CD is a two-component antimicrobial consisting of peptides, Trnα and Trnβ, produced by *Bacillus thuringiensis*, that specifically targets the bacterium *Clostridium difficile*. Each peptide contains three thioether crosslinks in which cysteine residues are chemically bonded to the α-carbon of distal amino acid residues; the resulting compact structure is highly resistant to proteolysis. The thioether crosslinks are proposed to be introduced by the radical SAM enzymes TrnC and/or TrnD. In this work, we develop a fusion protein strategy for producing the propeptides TrnA and TrnB and further demonstrate that TrnC alone is capable of introducing two of the thioether crosslinks. The successful *in vitro* production of thuricin CD could provide a potential lead for novel antimicrobials targeting *C. difficile* infections.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

A_{280}	Absorbance at 280 nm
ACN	Acetonitrile
ATP	Adenosine triphosphate
BME	2-mercaptoethanol
Bp	Base pair
Btn	Biotinylated
CBD	Chitin Binding Domain
CDAD	*Clostridium difficile*-associated diarrhea
°C	degree Celsius
Cys	Cysteine
ddH_2O	Double distilled water
DEAE	Diethylaminoethanol
Dha	2,3-didehydroalanine
Dhb	(*Z*)-2,3-didehydrobutyrine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	1,4-dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
g	Gram

Gly	Glycine
HPLC	High performance liquid chromatography
hr(s)	Hour(s)
IAD	Iodoacetamide
IPTG	Isopropyl B-D-1-thiogalactopyranoside
kDa	Kilodalton
L	Liter
Lan	Lanthionine
LCMS-TOF	Liquid Chromatography Mass Spectra- Time of Flight
m	Minutes
m/z	Mass to charge ratio
M_{av}	Mass average
MALDI-MS	Matrix assisted laser desorption mass spectroscopy
MBP	Maltose Binding Protein
MeL	Methyllanthionine
mg	Milligram
mL	Milliliter
mM	Millimolar
MWCO	Molecular weight cut-off
NaCl	Sodium Chloride
ng	Nanogram
Ni-NTA	Nickel nitrilotriacetic acid

NMR	Nuclear Magnetic Resonance
PCR	Polymerase chain reaction
Phe	Phenylalanine
PMSF	Phenylmethylsulfonyl fluoride
Rpm	Revolutions per minute
s	Seconds
SAM	*S*-adenosyl-L-methionine
SDS-PAGE	Sodium dodecyl sulfide polyacrylamide gel electrophoresis
Ser	Serine
SKF	Sporulation Killing Factor
TCEP	Tris (2-carboxyethyl) phosphine
TEV	Tobacco etch virus
TFA	Trifluoracetic acid
Thr	Threonine
TLCK	Tosyllysine Chloromethyl Ketone Hydrochloride
TSAP	Thermosensitive Alkaline Phosphatase
µg	Microgram
µL	Microliter
µM	Micromolar
10x	Ten times
5'-dA	5'-deoxyadenosyl

Chapter 1

Introduction, Background, and Research Goals

1.1. Introduction and Background

Thuricin CD is an antimicrobial peptide complex that was recently identified and purified (1) and found to be a two-component system consisting of the peptides TrnA and TrnB, with antimicrobial activity against *Clostridium difficile* in the nanomolar concentration range. Thuricin CD is produced by *Bacillus thuringiensis*, a bacterial strain commonly found in human feces. This bacteriocin (antibacterial proteins or peptides that kill or inhibit the growth of bacteria) is an interesting target because it is not harmful to beneficial flora in the human gut, as opposed to more common broad-spectrum antibiotics used to treat *Clostridium difficile*-associated diarrhea (CDAD) (1). CDAD occurs when *C. difficile* produces spores and germinates in the stomach of patients that have been treated with antibiotics and lack other flora. This lack of flora allows the *C. difficile* to flourish, leading to *C. difficile* infection, which ultimately causes CDAD (2). In recent years, an upsurge of highly virulent, antibiotic resistant *C. difficile* strains has increased annual death rates related to microbial infection dramatically. In fact, it has been shown that CDAD related fatalities increased from 5.7 to 23.7 persons per million in just a five-year span (3). The present lack of efficient treatment necessitates a new antibiotic source.

Bacteriocin peptides are produced by many organisms to control the proliferation of microorganisms in the surrounding environment. Examples of bacteriocins have been

reported in diverse organisms, including bacteria, fungi, and humans. These ribosomally-synthesized peptides are most commonly found to be 30 – 100 amino acid residues in length, displaying compact structures that have been stabilized by the introduction of post-translational modifications. In particular, these post-translational modifications include crosslinks generated by disulfide and/or thioether bond formation. Some well-known bacteriocins shown in Figure 1.01 include Nisin A (4), Subtilosin A (5), and Lacticin 481 (6).

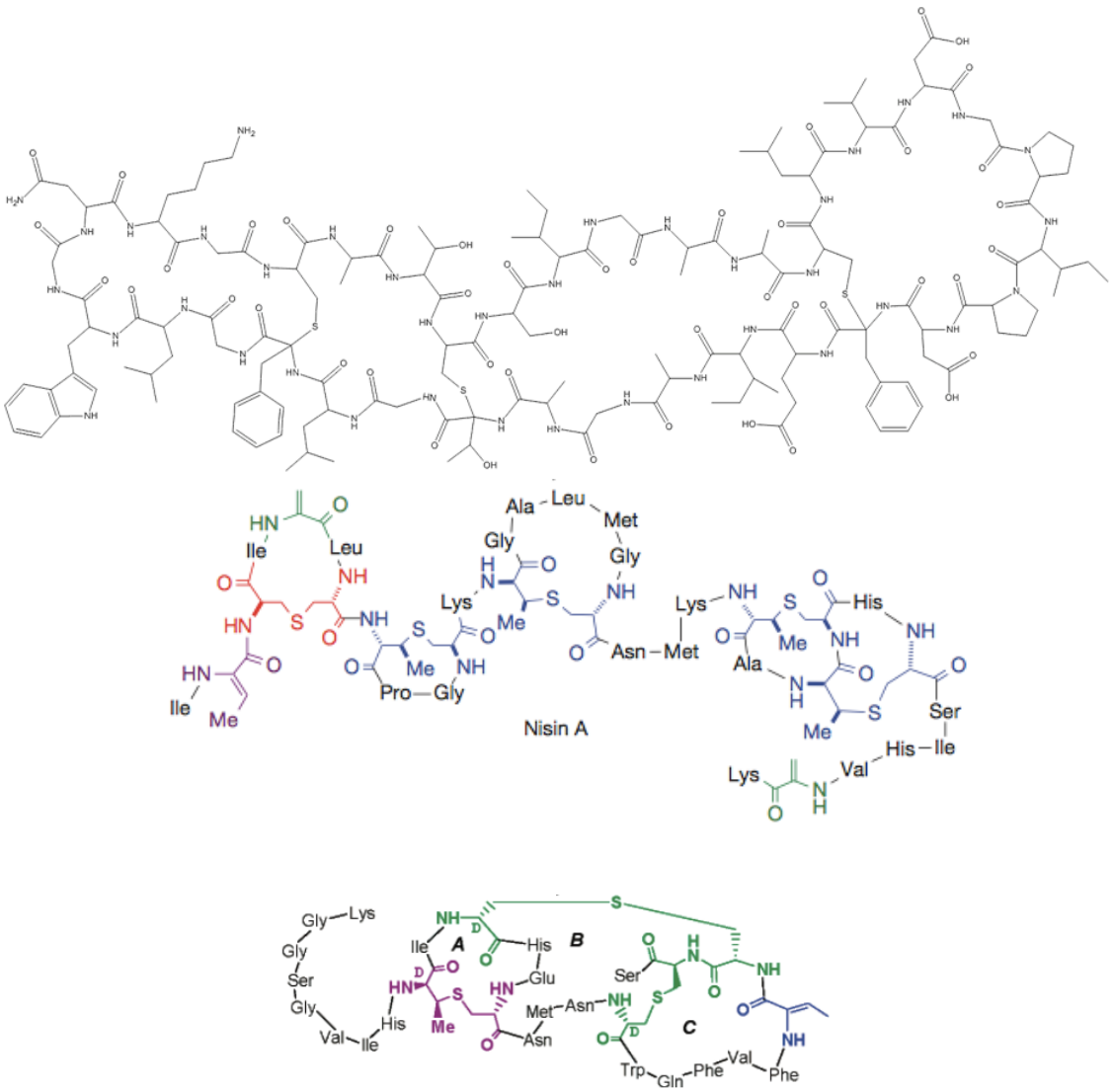


Figure 1.01 Amino acid structures of Nisin (top) (4), Subtilosin A (middle) (5), and Lacticin 481(bottom) (6).

Crosslinks between distal residues in bacteriocin peptides have the effect of stabilizing the peptides against proteolysis, as well as reducing conformational entropy, and thereby facilitating the formation of a single compact, folded structure. Cysteine has proven especially useful in the formation of crosslinks in the form of cysteine disulfide, lanthionine, methyllanthionine, and cysteine-to-α-carbon thioether bonds. Figure 1.02

depicts these four common motifs of sulfur-centered peptide crosslinking. The first structure shows a disulfide linkage created by oxidation of two cysteine residues. The second and third structures are classified as lanthionine/methyllanthionine linkages; formed between cysteine and serine/threonine residues, respectively, (dehydration of serine/threonine to afford the α-β unsaturated amino acid is followed by a Michael addition of the cysteine thiol to the β carbon) (7). The final structure in Figure 1.02 contains an interesting crosslink that presumably occurs between the α-carbon of any non-glycine residue and cysteine, created through radical *S*-adenosyl-L-methionine chemistry.

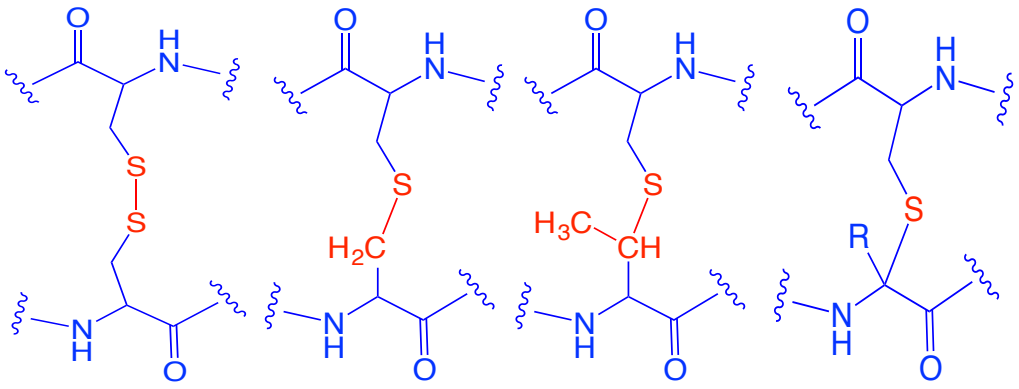


Figure 1.02. (Left to right) disulfide linkage, lanthionine linkage, methyllanthionine linkage, and cysteine to α-carbon linkage.

Initial characterization of Thuricin CD, including NMR studies, indicated that thioether crosslinks exist between cysteines 5, 9, and 13 and the residues at positions 21, 25, and 28 of each peptide (1). The modifications result in thioether crosslinks from cysteines to the α-carbon of these residues. These sulfur to α-carbon crosslinks place thuricin CD in a new category of bacteriocins known as sactibiotics (sac- stands for cysteine sulfur to

alpha carbon). The structures of Trnα and Trnβ, following post-translational modification and removal of the leader sequence, can be observed in Figure 1.03.

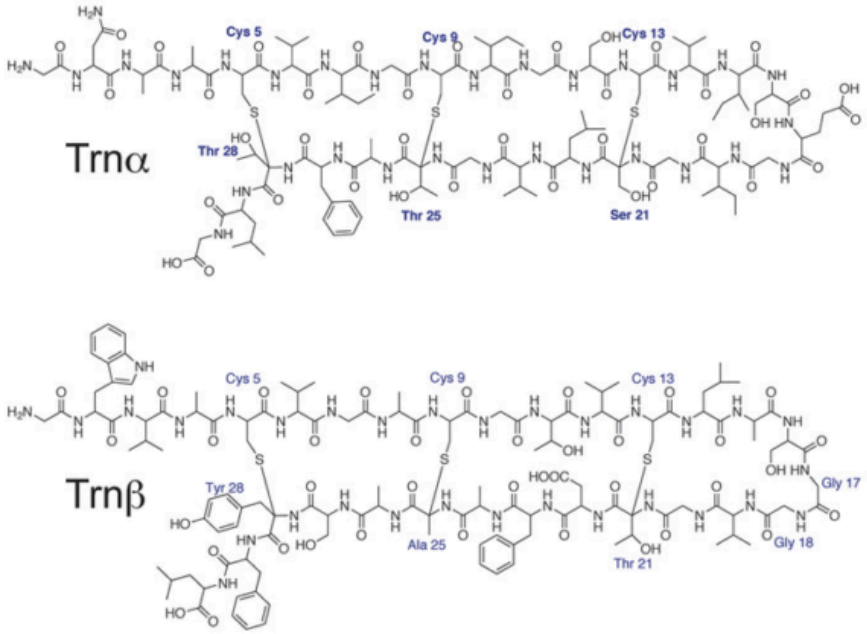


Figure 1.03. Structural representation of thuricin CD components TrnA and TrnB (1).

Sactibiotics are a relatively new discovery among the bacteriocins, as compared to the similar and well-studied lantibiotics. Previously, only two major classes of bacteriocins existed: lantibiotics and unmodified peptides. For the majority of bacteriocins, there exists a propeptide containing a leader peptide that facilitates excretion from the intracellular matrix to the periplasm. The thuricin CD operon includes a gene product for a protease enzyme that cleaves this leader peptide, yielding the mature, modified product (8). Unlike unmodified peptides, with the exception of disulfide formation, lantibiotics undergo a posttranslational modification by which a sulfur bridge is formed between the sulfur of cysteine and either serine or threonine residues. Dehydration of serine and threonine residues yields 2,3-dehydroalanine (Dha) and (*Z*)-2,3-dehydrobutyrine (Dhb),

respectively. Stereospecific, intramolecular attack by the sulfur of cysteine onto Dha or Dhb forms a lanthionine (Lan) or methyllanthionine (MeL) bridge, respectively (7), as can be observed in Figure 1.04 below.

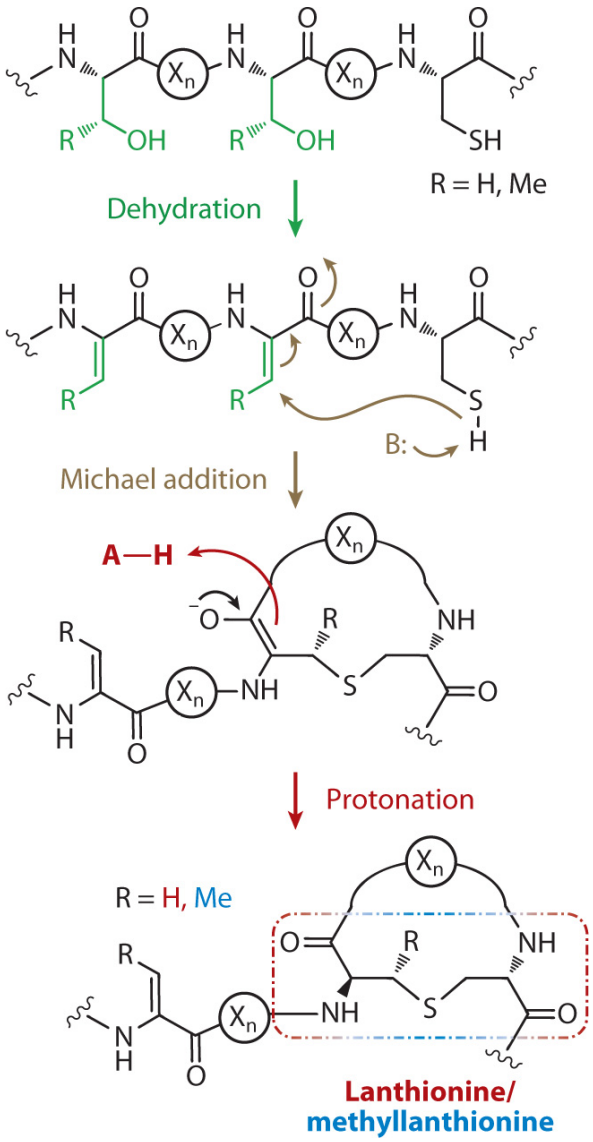


Figure 1.04 Installation of lanthionine (R=H) or methyllanthionine (R= Me) residues into a lantibiotic precursor peptide (7).

Like lantibiotics, sactibiotics undergo posttranslational modifications where the sulfur of cysteine residues become covalently bonded to the α-carbon of other specific

residues, as was first observed in Subtilosin A. After the structure of Subtilosin A was elucidated in 2003, the presence of the thioether crosslinks was revealed, and it was apparent that a new class of bacteriocins had been discovered. Since then, only five other sactibiotics have been isolated with successful structure elucidation. Interestingly, all known sactibiotics are produced by the genus *Bacillus* (9).

Similar to the operons of lantibiotics, the operons of sactibiotics consist of gene products responsible for proteolysis of the leader peptide, and immunity to and transportation of the product. Within the thuricin CD operon are eight possible enzyme coding regions. The protein sequences for which these genes encode show sequence homology to zinc-dependent proteases, self-immunity/carrier proteins and *S*-adenosyl-L-methionine radical enzymes. It is important to note, particularly for the self-immunity/carrier proteins, that these gene products will vary greatly between systems due to the fact that the target antagonist, and therefore, mechanism of action will vary between organisms. What has been shown to remain consistent between systems is the homology to ATP binding cassette transport proteins, with respect to the transport gene products observed. Sactibiotic operons differ from lantibiotics by the presence of gene products that show sequence homology to radical *S*-adenosyl-L-methionine-like enzymes. It was hypothesized that these enzymes are responsible for the cross linking reactions (1).

Common to radical *S*-adenosyl-L-methionine enzymes is a characteristic CXXXCXXC motif that has been found to be responsible for binding three of four irons in a $[4Fe\text{-}4S]^{2+}$ cluster. The fourth iron is responsible for coordinating one molecule of

S-adenosyl-L-methionine. Upon reduction of the $[4Fe\text{-}4S]^{2+}$ cluster, one electron is transferred to *S*-adenosyl-L-methionine, effectively cleaving it into a 5'-deoxyadenosyl radical (5'dA) and methionine (9). A scheme for this reaction can be observed in Figure 1.05.

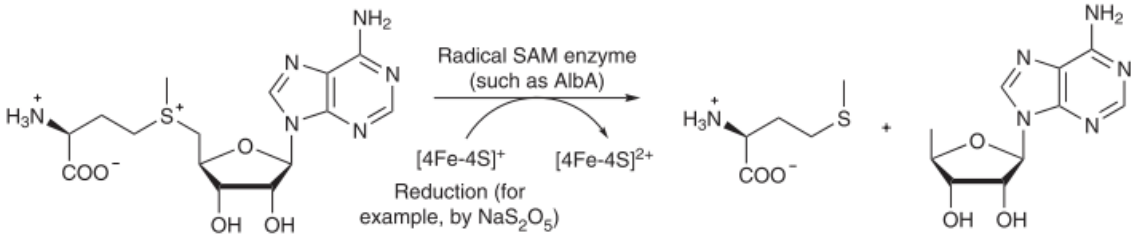


Figure 1.05. Radical cleavage of S-adenosyl-L-methionine into 5'-deoxyadenosyl radical and methionine (9).

An interesting characteristic found in the radical *S*-adenosyl-L-methionine enzymes of the well studied subtilosin A and sporulation killing factor (SKF) systems, both produced by *Bacillus subtilis,* is the presence of a second sequence motif containing 3 – 4 cysteine residues, proposed to bind a second iron-sulfur cluster, a feature found in other radical *S*-adenosyl-L-methionine enzymes such as BioB in the biotin synthase pathway (10). Three possible hypotheses can be formed based on this second FeS binding motif; the second motif is binding a second cluster that transfers an electron to the first cluster, thus activating the enzyme for a second turnover; the second cluster transfers an electron to a secondary source, i.e. sodium dithionite, that is responsible for the reactivation of the first cluster; or the third plays a role in the coordination of substrate.

Setting sactibiotics apart from lantibiotics is the presence in the biosynthetic operon of radical *S*-adenosyl-L-methionine enzymes, but one aspect common to both classes of compounds is the presence of a leader sequence in the unmodified gene product that is not present in the bioactive secondary metabolite. Moreover, it has been shown in lantibiotics that the presence of the leader peptide is necessary for enzymatic modification (4). This raises an important question; is the leader peptide necessary for enzymatic activity in sactibiotics? Studies of the sactibiotics subtilosin A and sporulation killing factor indicate that the cleavage of the leader peptide must occur after post translational modification (9) (11).

As mentioned above, the structure of Subtilosin A was elucidated in 2003 but it was not until 2012 that studies confirmed a radical SAM enzyme was responsible for catalysis of the cross-linking reactions (9). The product contains 35 residues, with thioether linkages between Cys4 and Phe31, Cys7 and Thr28, and Cys13 and Phe22 with retention, inversion, and inversion of stereochemistry, respectively. The leader peptide includes eight residues making a 43-residue precursor peptide (sboA). The enzyme, AlbA, contains two $[4Fe\text{-}4S]^{2+}$ clusters, with the second cluster binding site residues appearing towards the C-terminus of the polypeptide chain, as is common in radical SAM enzymes that modify sactibiotics. Along with AlbA, the operon contains gene products responsible for immunity, namely *albB*, *albC*, and *albD*, and for proteolysis of the leader peptide, *albE* and *albF* (9).

Another notable sactibiotic that has been studied in depth in recent years is sporulation killing factor (SKF), a 26 residue, cyclic peptide containing a thioether crosslink between Cys4 and Met12 and one disulfide bridge. The leader peptide includes 31 residues, making a 55-residue precursor peptide (SkfA). One enzyme in the peptide gene cluster, SkfB, contains two $[4Fe\text{-}4S]^{2+}$ clusters known to catalyze the thioether crosslinking reaction. Along with SkfB the operon encodes gene products responsible for immunity, namely SkfE and SkfF, and for proteolysis of the leader peptide, SkfC (11).

While countless studies can be envisioned, three important questions emerged during the investigation of both sactibiotics: (1) are radical *S*-adenosyl-L-methionine enzymes responsible for catalysis of the thioether crosslinking reactions, (2) what is the significance of the second $[4Fe\text{-}4S]^{2+}$ cluster, and (3) is the crosslinking reaction dependent on the presence of the leader peptide?

To answer any one of these questions, it was necessary to first express and purify both the substrate and enzyme. After observation of the amino acid compositions of SboA and SkfA, it quickly became clear that the peptides contained many hydrophobic residues, as is the case in the majority of sactibiotics. In the case of Subtilosin A, SboA was insoluble at concentrations higher than 200 μM and as a result, DMSO was necessary to solubilize the peptide (9). During the study of SKF, mutational analysis of SkfA with respect to the SkfB binding pocket indicated that large hydrophobic and aromatic residues are most tolerated, smaller hydrophilic residues are moderately tolerated, and large hydrophilic residues are not tolerated at all (11). This suggests that the binding

pocket of SkfB is hydrophobic in nature, which may also be the case for other sactibiotics.

After expression and purification of the radical SAM enzyme, it is necessary to reconstitute the FeS cluster(s). After reconstitution, the iron and sulfur content was determined to be approximately 8 mole Fe/S per mole of protein for both AlbA and SkfB. It can be shown via UV-Vis that the clusters are $[4Fe\text{-}4S]^{2+}$ clusters by the presence of a characteristic shoulder band at 410 nm (Figure 1.06). A $[2Fe\text{-}2S]^{2+}$ cluster would show an absorption band at approximately 460nm. After triple alanine mutants were generated for the first radical SAM $[4Fe\text{-}4S]^{2+}$ cluster, in both AlbA and SkfB, and reconstitution of the enzymes, the Fe-S content was determined to be approximately 4 mole Fe/S per mole of protein. This indicates that the second FeS cluster binding motif also binds a $[4Fe\text{-}4S]^{2+}$ cluster (11).

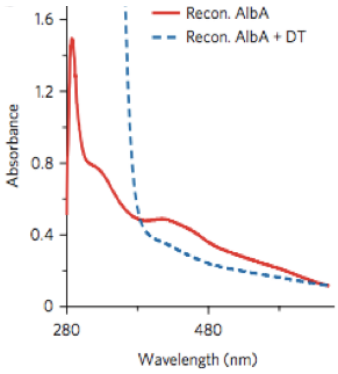


Figure 1.06. Indication of a $[4Fe\text{-}4S]^{2+}$ cluster by a characteristic 410 nm shoulder band.

Once the $[4Fe\text{-}4S]^{2+}$ clusters have been reconstituted, it is necessary to determine their reactivity with SAM. The proposed mechanism is as follows: an external reducing

agent, i.e. sodium dithionite, transfers one electron to the $[4Fe\text{-}4S]^{2+}$ cluster that is coordinated by the first CXXXCXXC motif in the radical *S*-adenosyl-L-methionine enzyme. The *S*-adenosyl-L-methionine molecule, coordinated by the fourth iron of the cluster, is subsequently cleaved, producing methionine and one equivalent of a 5'dA radical, the latter being responsible for abstracting the hydrogen of the α-carbon of the respective residue. This radical center then forms a thioether bond with the sulfur of Cys, coordinated to the second $[4Fe\text{-}4S]^{2+}$ cluster, which acts as a final electron acceptor. It was proposed for both AlbA and SkfB that this electron is then donated back to the original external electron donor, now acting as an acceptor, preparing the system for a second turnover. These findings could provide insight as to the role of the second $[4Fe\text{-}4S]^{2+}$ cluster. This proposed mechanism is shown in Figure 1.07 (9).

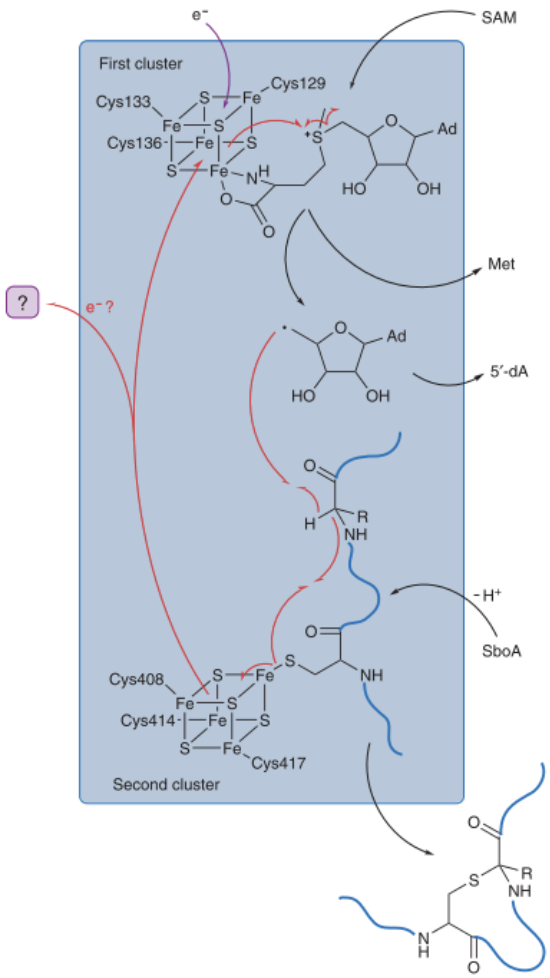


Figure 1.07. Proposed radical mechanism for the formation of sactibiotic like thioether crosslinks (9).

Triple alanine mutants of both FeS cluster binding motifs were used to determine the necessity of each $[4Fe\text{-}4S]^{2+}$ cluster in both AlbA and SkfB. Activity assays conducted on enzymes in which the first CXXXCXXC motif had been mutated resulted in no production of modified peptide. This result suggests that the first $[4Fe\text{-}4S]^{2+}$ cluster plays a role in the catalysis of the thioether crosslinks. Similar results were observed in the triple alanine mutants of the second cluster motif. Activity assays involving the mutant AlbA and SkfB produced no modified peptide precursor. Not only does this result

demonstrate the necessity of both $[4Fe\text{-}4S]^{2+}$ clusters in the catalysis of the thioether crosslinking reactions, but also suggests that the second cluster is also necessary for activity. Moreover, experiments have shown that not only does mutation of these three cysteine residues render the enzyme inactive, but chemical treatment of the native enzyme with iodoacetamide prior to activity assays result in failure to modify the peptide substrate in both AlbA and SkfB (9) (11).

The third important question is whether or not the leader peptide is necessary for activity in the precursor peptide. In both Subtilosin A and SKF systems, activity assays in which the enzymes are fed precursor peptide (post cleavage of the leader sequence) did not result in production of modified peptide (9) (11). These results suggest that the leader peptide in SboA and SkfA plays a role in binding the peptide to AlbA or SkfB, respectively, or more specifically, in active site recognition.

Many of the questions that surround the Subtilosin A and SKF systems are also relevant in investigating the thuricin CD system. As mentioned previously, thuricin CD is a two-component sactibiotic system, consisting of thioether crosslinks between Cys and the α-carbon of respective residues in the 30 residue peptides TrnA and TrnB. The unmodified amino acid sequence of the both peptides can be observed in Figure 1.08 (A). As can be seen in Figure 1.08 (B), the two precursor peptides of TrnA and TrnB contain 47 and 49 residues, with a 17 and 19 residue leader peptide, respectively. The crosslinks occur between cysteine residues and post-cleavage residues 21, 25, and 28 in both components. Like the majority of sactibiotics discovered thus far, the operon shown in Figure 1.08 (C) consists of gene products believed to be responsible for product transport, trnF and trnG; gene products for the bacteriocin, TrnB and TrnA; gene products for

potential radical SAM enzymes, TrnC and TrnD; as well as gene products for the maturation protease, trnE, predicted to cleave at the glycine-glycine sequence (first Gly-Gly in TrnB); and the recently reported gene product for self immunity, TrnI. (1). Transcription start sites were identified by *in silico* analysis and suggest that all of the genes are transcribed from a single mRNA, and therefore initial expression levels of all components and enzymes are approximately equivalent.

pro-Trnα MEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAFTLG*

Trnα GNAACVIGCIGSCVISEGIGSLVGTAFTLG*

pro-Trnβ MEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASYFL*

Trnβ GWVACVGACGTVCLASGGVGTEFAAASYFL*

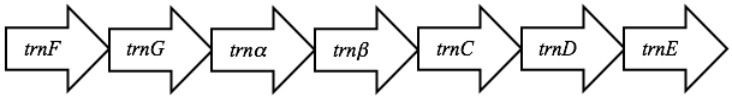


Figure 1.08. A) Amino acid sequences for cleaved TrnA and TrnB. B) Amino acid sequence for propeptides TrnA and TrnB. C). Representation of the thuricin CD operon. *Cysteines and residues involved in crosslinking reactions are marked in red and blue, respectively.

Unlike the preceding two systems discussed previously, there exist two radical *S*-adenosyl-L-methionine enzymes in the thuricin CD operon. It is assumed that because two components in this system are modified, there is a need for two enzymes. However, the relative stereochemistries of the modified α-carbons beg several interesting questions. ^{1}H-NMR studies have confirmed that the biological pathway proceeds with a net retention of stereochemistry at the first two modified residues and inversion of chemistry at the last (1). This would suggest one of the following: (A) one of the enzymes is

responsible for the reactions resulting in retention of stereochemistry and the other for the reaction resulting in inversion, (B) one of the enzymes is an epimerase enzyme, or (C) each enzyme is responsible for one substrate and there exists an epimerase, not present in the operon, that is responsible for the epimerization reaction. It is possible to gain insight into the role of each radical SAM enzyme in the operon by observing activity assays with substrate, but only one of the enzymes, and determining which residues, if any, are modified.

Figure 1.09. Amino acid sequence of TrnC, depicting the radical SAM domain and the SPASM domain.

The thuricin CD system also differs from the previously discussed sactibiotic systems in that the sequences for both TrnC and TrnD indicate the potential for either one, two, or three iron-sulfur cluster binding sites, as can be seen in Figure 1.09. If the enzymes contain only one FeS cluster, the radical SAM cluster, then the thioether crosslinking reactions would presumably need to proceed through an alternate mechanism than that of

SKF and Subtilosin A. A proposed alternate mechanism for this reaction can be observed in Figure 1.10.

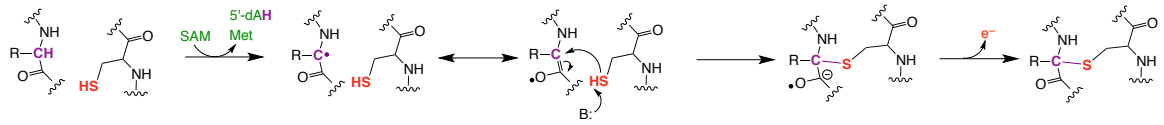


Figure 1.10 Proposed radical mechanism for the formation of sactibiotic link crosslinks in thuricin CD.

1.2 Research Goals

The most important goal in my thesis project is to develop a method for the expression and purification of propeptides TrnA and TrnB. In order to successfully develop this method, *trnA* and *trnB* must first be subcloned into an appropriate vector for recombinant expression of the propeptides. Product isolation is predicted to be a challenging task, as unstructured peptides are notoriously difficult to purify. Nonetheless, successful isolation of TrnA and TrnB will allow for the development of activity assays involving peptide modifying enzymes TrnC and TrnD. With TrnD previously purified by the Jarrett laboratory, it will also be a goal of mine to successfully purify TrnC for use in said activity assays. Development of successful assays may provide valuable insight on the antimicrobial activity of thuricin CD.

Chapter 2

Subcloning, Expression, and Purification of TrnA, TrnB, and Related Fusion Proteins

This chapter describes the various methods and techniques utilized in the attempt to purify propeptides TrnA and TrnB. When early attempts to purify the propeptides failed, it became clear that more definitive methods must be applied. Transitioning into several fusion protein systems and variations of such provided successful purification techniques; however, after unsuccessfully isolating the propeptides (post-cleavage) from said fusion partners, it was believed the overall solubility of the propeptides was contributing most to the failed attempts of isolation. To achieve solubility, constructs were designed to include affinity tags that would contribute to water solubility.

As stated, propeptides TrnA and TrnB were first attempted to be purified alone (with no fusion partner/ affinity tag). When this method failed, a construct was designed to include a biotin recognition site for purification through avidin affinity. This construct was subcloned and expressed, however, the recognition site was not separated from the propeptide by a cleavage site, meaning it could not be later removed. With the potential for negative interactions between this biotin recognition site and the propeptide modifying enzymes, new constructs were designed. These constructs were designed to include TrnA and TrnB as MBP-factorXa fusion proteins. This would allow for more successful purification, followed by removal of the propeptides through factorXa cleavage. These fusion proteins were difficult to purify, as purification was limited to few

techniques, including the utilized ion exchange chromatography. Also, factorXa cleavage was unsuccessful. This led to the development of the same fusion proteins, but with removal of the factorXa protease recognition sequence and the inclusion of a tobacco etch virus protease recognition sequence, as well as an N-terminal hexahistidine affinity tag. While purification of both constructs were successful, cleavage success was low for His_6MBP-TEV-TrnB and was unsuccessful for His_6MBP-TEV-TrnA. It was also believed a heterogeneous His_6MBP-TEV-TrnA product was consistently purified. Mutations at the recognition sequence led to elevated levels of cleavage in both constructs. After observing successful cleavage in both constructs, new constructs were to designed to modify His_6MBP-TEV-TrnA in attempts to purify a single product, but were unsuccessful in doing so. At the point, focus was put exclusively on propeptide TrnB.

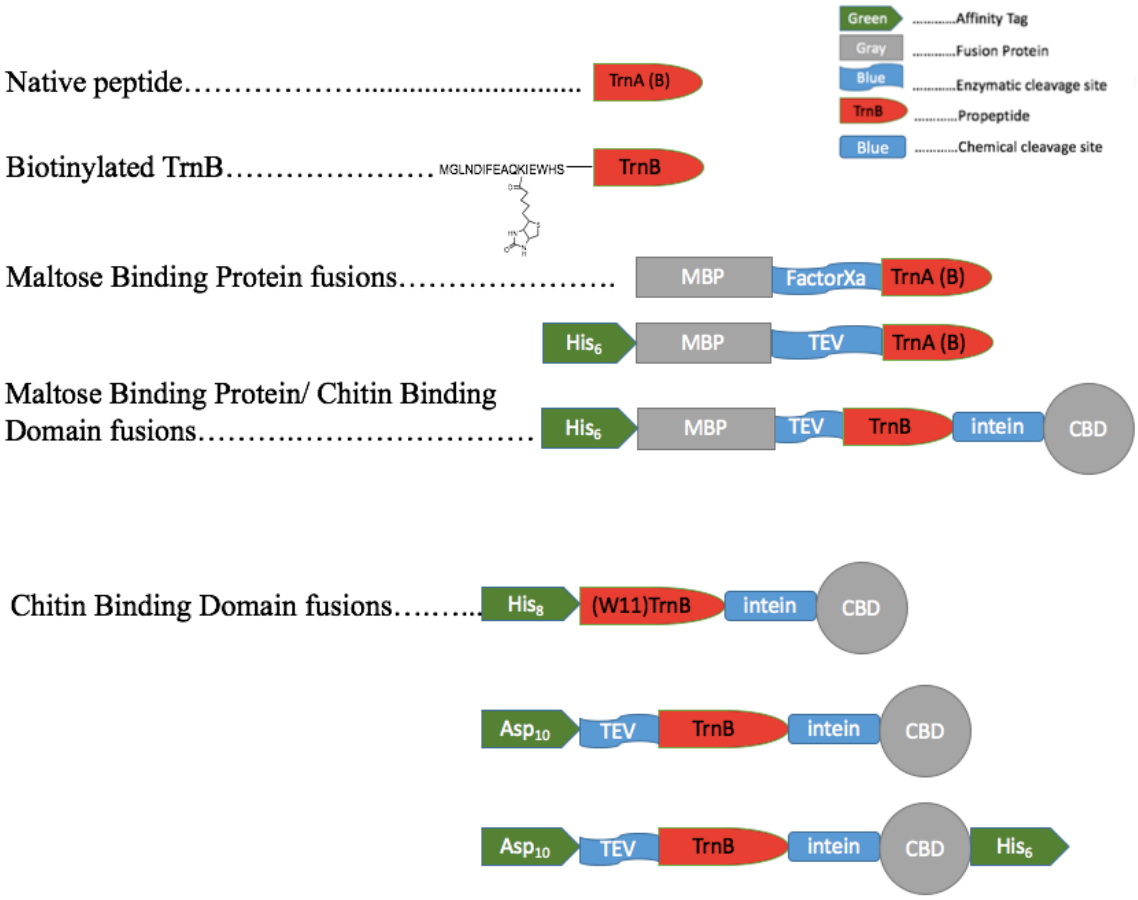
In an effort to allow for multiple methods of purification and cleavage, His_6MBP-TEV-TrnB was subcloned into pTWIN I, which encodes a C-terminus chitin binding domain (CBD), with an intein cleavage site separating the propeptide from the CBD. After successful detection of TrnB through LCMS-TOF was achieved, a construct was designed to remove the His_6MBP-TEV and include an N-terminal octahistidine affinity tag, as well as a tryptophan residue to aid in A_{280} detection. The purpose of this construct was built around the proposal that His_8TrnB could be purified through nickel affinity chromatography. After failed attempts to purify the propeptide, the octahistidine affinity tag was removed and a decaaspartate affinity tag was inserted in its place. This was believed to aid in water solubility, however, the deletion of the hexahistidine affinity tag removed to the ability to purify the fusion through nickel affinity chromatography. As a

result, a final construct was designed to include a hexahistidine affinity tag at the C-terminus of CBD. Table 2.01 displays the affinity tag, fusion partner, cleavage type, and propeptide for each fusion construct described:

Table 2.01 Fusion protein construct design.

Construct	Btn-TrnB	MBP-factorXa-TrnA MBP-factorXa-TrnA	MBP-TEV-TrnA MBP-TEV-TrnA	His_6MBP-TEV-TrnA His_6MBP-TEV-TrnA	His_6MBP-TEV-TrnB-intein-CBD	(W11) His_8TrnB-intein-CBD	Asp_{10}TEV-TrnB-intein-CBD	Asp_{10}TEV-TrnB-intein-CBD-His_6
N-terminal affinity tag	Biotin recognition site			Hexahistidine tag	Hexahistidine tag	Hexahistidine tag	Deca-aspartate tag	Decaaspartate tag
N-terminal Fusion protein		Maltose Binding Protein	Maltose Binding Protein	Maltose Binding Protein	Maltose Binding Protein	Maltose Binding Protein		
Protease recognition		FactorXa	Tobacco Etch Virus	Tobacco Etch Virus	Tobacco Etch Virus		Tobacco Etch Virus	Tobacco Etch Virus
Propeptide	TrnB	TrnA TrnB	TrnA TrnB	TrnA TrnB	TrnB	TrnB	TrnB	TrnB
Cleavage site					Intein	Intein	Intein	Intein
C-terminal fusion protein					Chitin Binding Domain	Chitin Binding Domain	Chitin Binding Domain	Chitin Binding Domain
C-terminal affinity tag								Hexahistidine Tag

To aid in visualization, the following pictorial representations of the different constructs has been included. These pictures are also found directly after the sequence for each construct in the following passages:



2.1. TrnAB

2.1.1. Background

The initial methodology to purify TrnA and TrnB was to co-express the propeptides and isolate the assumed oligomeric product. Following denaturation, the products could be separated through high performance liquid chromatography and further purified.

The TrnA and TrnB genes were ordered from GenScript as codon-optimized synthetic genes in a single pUC57 vector. TrnB preceded TrnA and the genes were separated by a 15 base pair sequence that encodes a ribosome-binding site. Although the gene construct includes two separate stop codons, for the purpose of ease in nomenclature, the construct is referred to as TrnAB.

After polymerase chain reaction amplification, TrnAB was ligated into pET23b(+) and the product was used to transform competent DH5α for DNA amplification. Agarose gel electrophoresis was carried out to confirm the length of the resulting plasmid, which was subsequently used to transform competent BL21(DE3)pLysS for protein expression. After protein expression, the resulting crude product was purified via DEAE ion exchange chromatography and identified with sodium dodecyl sulfate poly acrylamide gel electrophoresis.

A: MEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAFTLG*

B: MEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASYFL*

Figure 2.01. Illustrative representation of propeptides TrnA and TrnB.

2.1.2. Subcloning

The TrnAB genes were received from GenScript in a pUC57 vector. Screening with ampicillin (50 µg/mL), TrnAB_pUC57 was used to transform competent DH5α (Promega) for DNA amplification. The DNA was isolated using the PureYield^{TM} Plasmid Miniprep System from Promega. After DNA amplification and purification, polymerase chain reaction was used to amplify the TrnAB gene following the conditions in Table 2.02 and 2.03 below:

TrnAB Forward Primer: CAATCATATGGAAGTCCTGAACAAACAG

TrnAB Reverse Primer: CAATCTCGAGTTAACCCAGCGTAAAC

Table 2.02 Phusion PCR components.

Component	Volume (µL)	Final concentration
Polymerase chain reaction grade H_2O	32.5	-
5X Phusion HF Buffer	10	1X
dNTPs (10mM each)	1.0	200µM each
Primer A (10 µM)	2.5	0.5 µM
Primer B (10 µM)	2.5	0.5 µM
Template DNA (10 ng/ µL)	1.0	0.2 ng/ µL

Component	Volume (μL)	Final concentration
Phusion DNA Polymerase	0.5	0.02 Units/ μL
Total	50 μL	

Table 2.03 Phusion PCR cycling conditions.

Cycle step	Temperature (°C)	Time	No. Cycles
Initial Denaturation	98	30 s	1
Denaturing	98	7.5 s	30
Annealing	62	20 s	
Extension	72	25 s	
Final Extension	72	5 m	1

*Further discussion will refer to this protocol as the *Phusion polymerase chain reaction protocol*.

The resulting polymerase chain reaction product was purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System from Promega.

Restriction digests were used to prepare pET23b(+) and TrnAB for ligation. The restriction digest conditions are described in Table 2.04:

Table 2.04. Restriction digest components.

DNA	NdeI (1U/μL) (Promega)	XhoI (1U/μL) (Promega)	Buffer D (10X) (Promega)	Alkaline Phosphatase (1MBU/μL) (Promega)	Milli-Q H_2O	Total
pET23B (1μg)	1 μL	1 μL	10 μL	1 μL	84 μL	100 μL

DNA	NdeI (1U/µL) (Promega)	XhoI (1U/µL) (Promega)	Buffer D (10X) (Promega)	Alkaline Phosphatase (1MBU/µL) (Promega)	Milli-Q H_2O	Total
TrnAB (1µg)	1 µL	1 µL	10 µL	-	43 µL	100 µL

After allowing the restriction digest to incubate for 3 hrs at 37°C, 1 µL of Thermosensitive Alkaline Phosphatase (TSAP, 1 unit/µL, Promega) was added to pET23b(+), and the incubation continued for 1 hr. After 4 hrs of incubation, the restriction digest products were purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System (Promega). *Further discussion will refer to this protocol as the *restriction digest protocol*.

The purified products were subsequently ligated using the conditions in Table 2.05:

Table 2.05. Ligation components.

Component	1:1 (Vector: Insert)	1:3 (Vector: Insert)
pET23b (+)	50 ng	50 ng
TrnAB	10 ng	30 ng
Ligase 10x Buffer (Promega)	1 µL	1 µL
T4 Ligase (Promega)	1 µL	1 µL
Milli-Q H_2O	1 µL	2 µL
Total	10 µL	10 µL

The ligations were incubated at 4°C for 16 hrs and later purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System (Promega). *Further discussion will refer to this protocol as the *ligation protocol*.

Following purification, the ligation products were used to transform competent DH5α (Promega) and screened on ampicillin (50 μg/mL) resistant Luria broth agar plates. Successful colonies were then inoculated into Luria broth medium, screening with ampicillin (50 μg/mL) for DNA isolation and purification with the PureYield^{TM} Plasmid Miniprep System (Promega).

As a method to check the correctness of the transformed product, the isolated DNA was digested with XhoI (Promega) and identified through agarose gel electrophoresis (control: pET23b(+), lane 2). Expecting a product approximately 300 base pairs larger than the control, Figure 2.02 shows four bands (lanes 3-6) of roughly the correct size.

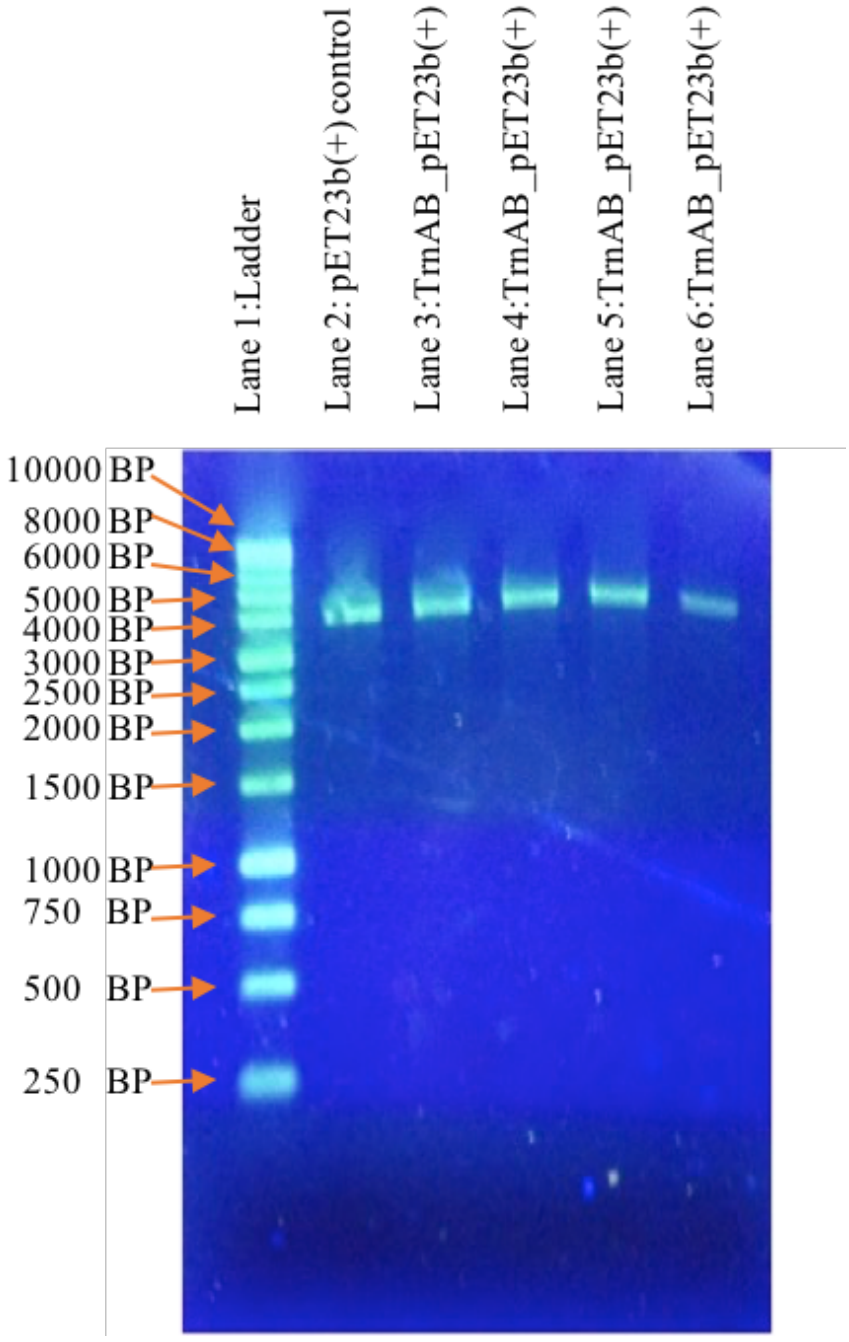


Figure 2.02. Agarose gel depicting TrnAB_pET23b (+) (Lanes3-6). Lane 1, DNA ladder; lane 2, pET23b (+) control.

The TrnAB_pET23b(+) sample from lane 4 (Figure 2.02) was used to transform competent BL21(DE3)pLysS (Invitrogen) for protein expression and purification.

2.1.3. Protein Expression and Purification

A 50 mL Luria broth culture of TrnAB_pET23b in BL21(DE3)pLysS was screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant shaking at 225 rpm. After incubation, 10 mL were added to 4 x 1 L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. Subsequently, the cells were harvested by centrifugation with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

Following defrosting, the pellets were re-suspended in 50 mM Tris-base, pH 8 at 2 mL/gram of material. The re-suspended cells were lysed by sonication with a Fisher Scientific 550 Sonic Dismembrator via 4 x 30 s blasts at 80% power, with 2-minute intervals for cooling. The resulting lysate was centrifuged at 48,400 x *g* for 1 hr to remove all insoluble cellular debris. Further discussion will refer to this protocol as the *sonication protocol*

Following centrifugation, ion-exchange chromatography was performed using FPLC (BioRad). DEAE Sepharose Fast Flow ion exchange resin (GE Healthcare) was first equilibrated with 50 mM Tris-base, pH 8 (buffer A), and after the loading of the lysate followed a 100 mL linear gradient of buffer A to 100% Tris, pH 8, 1.0 M NaCl (buffer B)

to elute bound protein. A 20 mL wash at 100 % buffer B followed. Five mL factions were collected throughout the gradient and were later pooled and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (Millipore) to be identified through sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Figure 2.03).

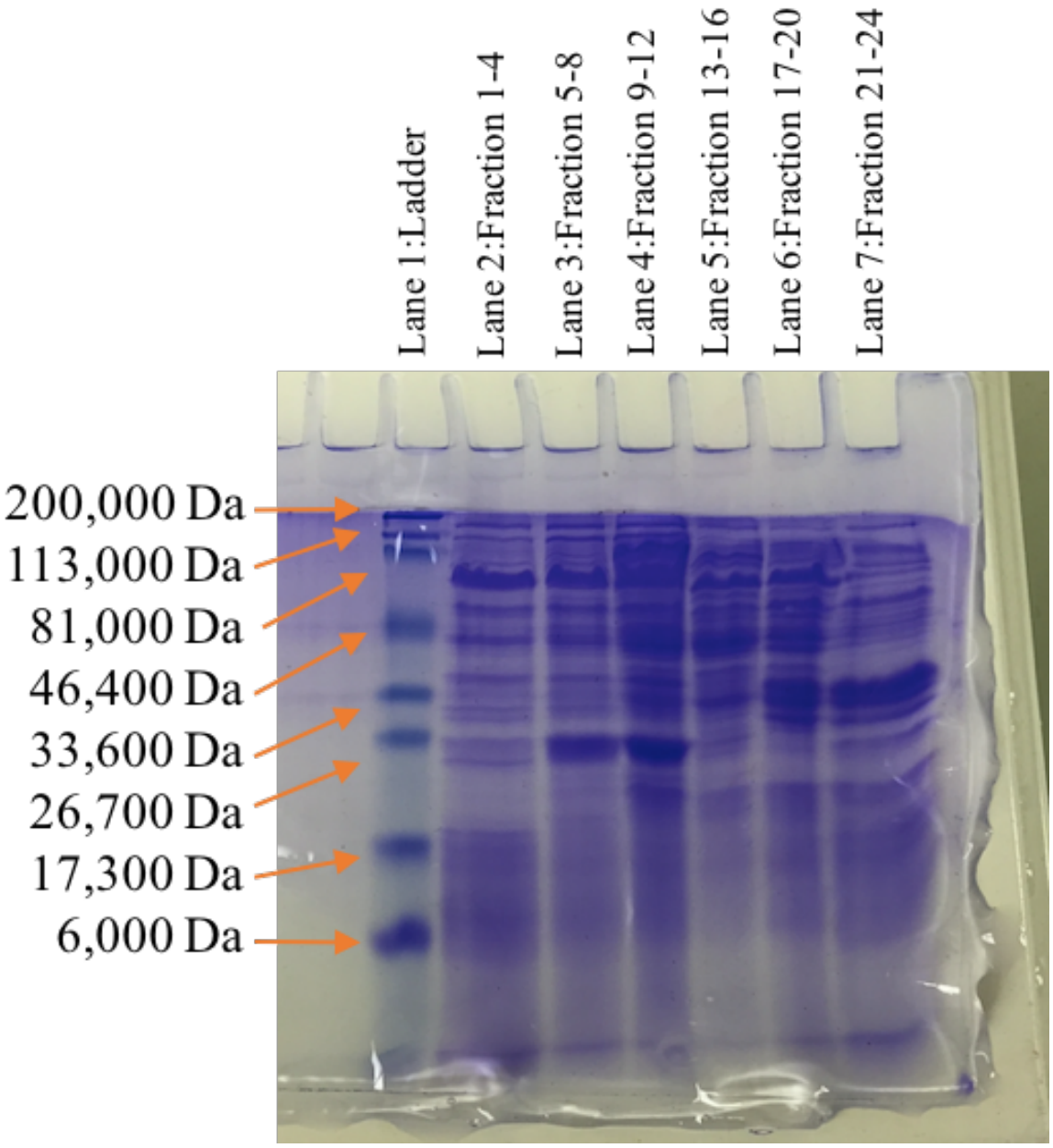


Figure 2.03. SDS PAGE depicting concentrated fractions from DEAE purification of Btn-TrnAB (Lanes2-7). Lane 1, protein standards.

2.1.4. Results

Initial attempts to isolate and purify the propeptides were unsuccessful. After the collection of fractions from ion exchange purification and subsequent attempts to identify through sodium dodecyl sulfate-polyacrylamide gel electrophoresis, it was apparent that a more specific means of purification was necessary to isolate the product. This is also coupled with the lack of a standard to assist in identifying the product using high performance liquid chromatography.

2.2. Biotinylated –TrnB (Btn-TrnB), M_{av} = 6,892.8 g/mol

2.2.1. Background

As a more specific method to purify overexpressed propeptide TrnB, polymerase chain reaction primers were designed to include a leader sequence that would result in the covalent attachment of biotin to the propeptide amino acid sequence; specifically, at the lysine residue. The biotinylated product could then be purified by utilizing avidin affinity chromatography. The addition of the biotinylation site also introduced the restriction enzyme recognition sight for EcoRV.

The Btn_TrnB primers were received from IDT to be used in ligation free polymerase chain reaction (described below). The subsequent product was later used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to verify the base pair length of the amplified DNA, which was next used to transform competent BL21 (DE3) pLysS (Invitrogen) for protein expression. After protein expression, the resulting crude product was purified via DEAE ion exchange chromatography and identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below below is the amino acid sequence for Btn_TrnB. In blue is the newly incorporated leader sequence for binding biotin, which includes the lysine residue highlighted in red, responsible for the binding of biotin; and orange, the TrnB sequence.

B: MGLNDIFEAQKIEWHSMEVLNKQNVNIIPESEEVGGWVACVGACGTVCLAS GGVGTEFAAASYFL*

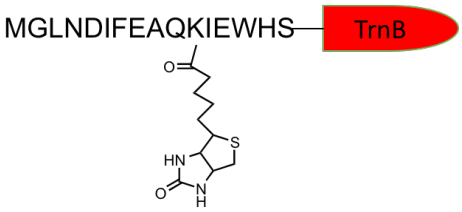


Figure 2.04. Illustrative representation of Btn-TrnB.

2.2.2. Subcloning

Ligation-free polymerase chain reaction was used to introduce the biotinylation sequence using the following primers:

Btn_TrnB Forward Primer:

TATGGGCCTGAATGATATCTTTGAAGCGCAGAAAATTGAATGGCACTC

Btn_TrnB Reverse Primer:

TAGAGTGCCATTCAATTTTCTGCGCTTCAAAGATATCATTCAGGCCCA

Materials and methods used to modify TrnAB_pET23b (+) are found in Table 2.05 and 2.06.

Table 2.06. Ligation-free KOD PCR components.

Component	µL	Final concentration
Polymerase chain reaction grade H_2O	32.0	-
KOD Hot Start 5x Buffer	5.0	1X
MgSO4 (25mM)	3.0	1.5mM

Component	µL	Final concentration
dNTPs (2mM each)	5.0	200 µM each
Primer A (10 µM)	1.5	0.3 µM
Primer B (10 µM)	1.5	0.3 µM
Template DNA (10 ng/ µL)	1.0	0.2 ng/ µL
KOD DNA Polymerase	1.0	0.02 Units/ µL
Total	50 µL	

Table 2.07. Ligation-free KOD PCR cycling conditions.

Cycle step	Temperature (°C)	Time	No. Cycles
Polymerase activation	95	2 m	1
Denaturing	95	20 s	20
Annealing	60	10 s	
Extension	70	100 s	

(After completion of the polymerase chain reaction, samples are treated with DpnI (Promega) for 1 hr at 37°C to remove template DNA; purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System (Promega); and incubated with T4 DNA Ligase (Promega) for three hours at room temperature) *Further discussion will refer to this protocol as the *ligation-free polymerase chain reaction protocol*.

The subsequent product was used to transform competent DH5α (Promega) and screened on ampicillin (50 µg/mL) resistant Luria broth-agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 µg/mL) and incubated overnight for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

To examine the correctness of the isolated DNA samples, Btn-TrnB_pET23b (+) was treated with EcoRV (Promega), following the *restriction digest protocol* (page 40), to generate linear DNA to be analyzed through agarose gel electrophoresis (control: TrnAB_pET23b (+), lane 2). Expecting a product approximately 50 base pairs larger than the control, Figure 2.05 shows six bands (lanes 3-8) of roughly the correct size.

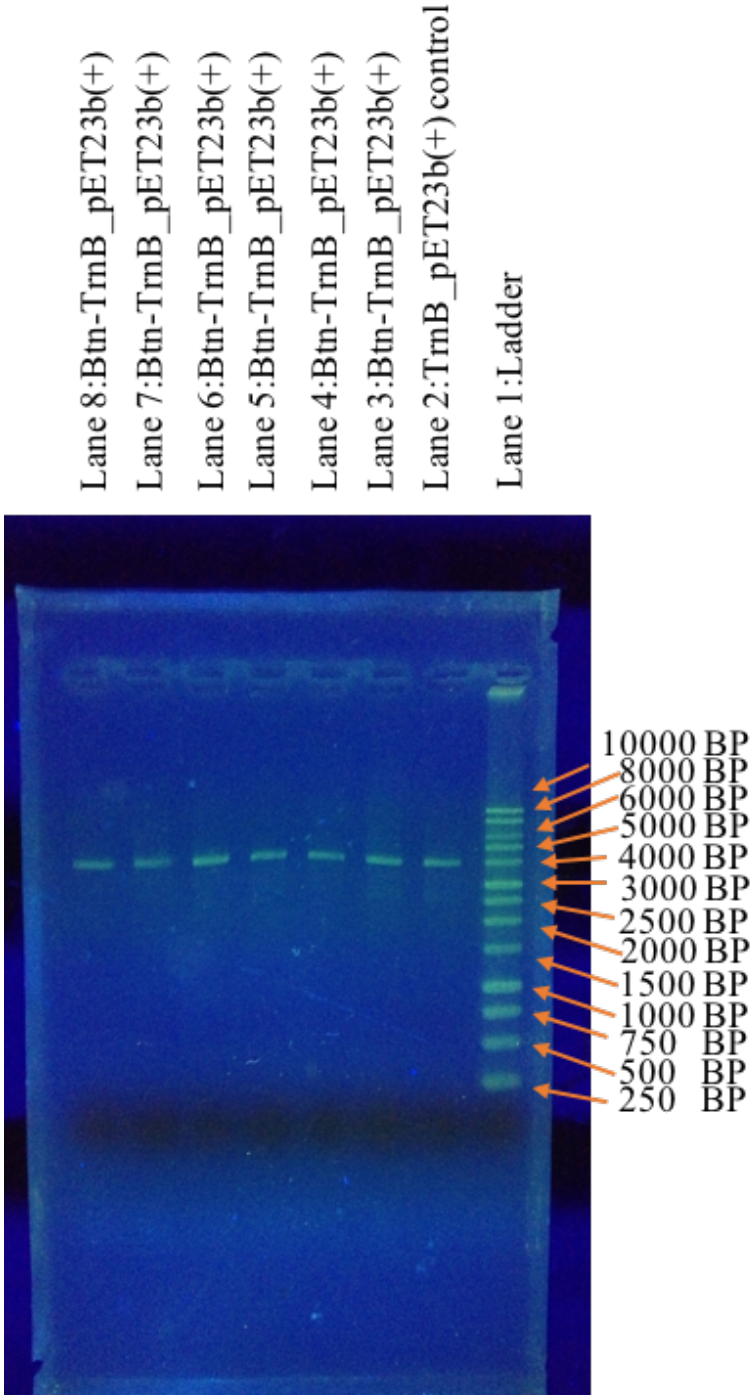


Figure 2.05. Agarose gel depicting Btn-TrnB_pET23b (+) (Lanes3-8). Lane 1, DNA ladder; lane 2, TrnAB_pET23b (+) control.

The Btn-TrnB_pET23b(+) sample from lane 4 (Figure 2.05) was used to transform competent BL21(DE3)pLysS (Invitrogen) for protein expression and purification.

2.2.3. Protein Expression and Purification

One 50 mL Luria broth culture of Btn-TrnB_pET23b(+) in BL21(DE3)pLysS was screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant shaking at 225 rpm. After incubation, 10 mL were added to 2 x 1 L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added to a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

Following defrosting, the pellet was re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Following centrifugation, ion-exchange chromatography was performed using FPLC (Bio-Rad). DEAE Sepharose Fast Flow ion exchange resin (GE Healthcare) was first equilibrated with 50 mM Tris-base, pH 8 (buffer A), and after the loading of the lysate followed a 100 mL linear gradient of buffer A to 100 % Tris-base, pH 8 1.0 M NaCl (buffer B) to elute bound protein. A 20 mL wash at 100 % buffer B followed. Five mL factions were collected through the entirety of the gradient and were later pooled and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit, by Millipore, to be

identified through sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Figure 2.06).

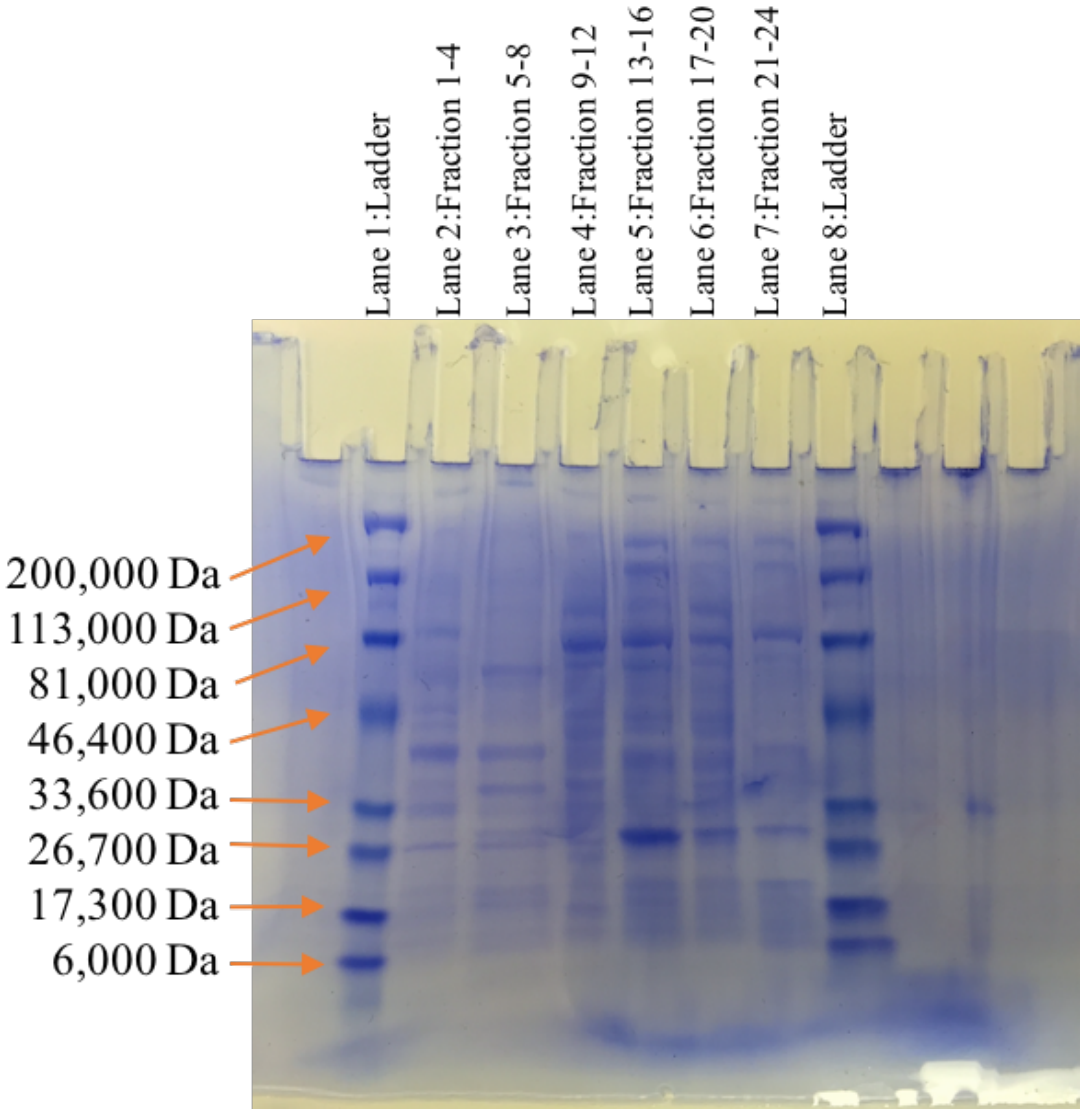


Figure 2.06. SDS PAGE depicting fractions concentrated following DEAE purification of Btn-TrnB (lanes 2-7). Lanes 1 and 8, protein standards.

2.2.4. Results

Observation of an overexpressed Btn_TrnB was not achieved through sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Further attempts to purify the product through avidin affinity chromatography were not made. The goal of this construct is to be utilized in future activity assays and for this reason, no further work was put into this construct; however, after future isolation of the propeptide and optimization of high performance liquid chromatography conditions for observation, the Btn_TrnB may be studied further.

2.3. MBP-factorXa-TrnA, M_{av}= 47,122 g/mol; MBP-factorXa-TrnB, M_{av}= 47,456 g/mol

2.3.1. Background

The maltose binding protein fusion technique, developed by New England Biolabs, is a commonly utilized method to increase protein expression levels and/or help solubilize an otherwise insoluble protein. Ligation of a target gene into the multiple cloning region found downstream of the malE-maltose binding protein coding gene of pMALc2x yields an N-terminal maltose binding protein-factorXa cleavage site-target protein fusion. After expression, the fusion protein can be purified through amylose affinity chromatography, and the subsequent product can be cleaved using the factorXa protease. FactorXa preferentially cleaves after the arginine residue in the Ile- (Glu or Asp)- Gly- Arg recognition sequence. The original purification scheme can be observed in Figure 2.07 (12).

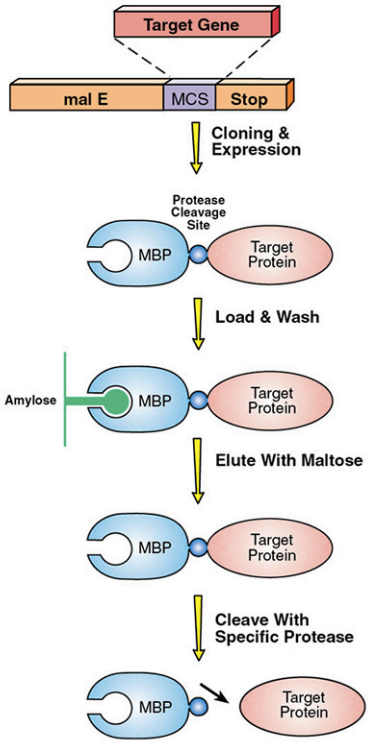


Figure 2.07. Purification scheme of the maltose binding protein fusion system (12).

Utilizing a fusion technique allows for increased levels of expression of the propeptides and provides an efficient means of purification and identification given a mass increase of approximately 42 kDa.

Polymerase chain reaction primers were designed to amplify the TrnA and TrnB genes from TrnAB_pUC57 for ligation into the pMALc2x vector. The ligation products were later transformed into competent DH5α (Promega) for DNA amplification, and the purified DNA was then used to transform competent BL21(DE3)pLysS for protein expression. After protein expression, the resulting crude product was purified via ion exchange chromatography and identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below are the amino acid sequences for both MBP-factorXa-TrnA and MBP-factorXa-TrnB. In black is the newly incorporated maltose binding protein; blue, the newly incorporated factorXa cleavage site; and orange, the TrnA (B) sequence.

A:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAAT
GDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPI
AVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAA
DGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA
FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAAS
PNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENA
QKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNN
NNNLGIEGRMEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAFTLG

B:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAAT
GDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPI
AVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAA
DGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA
FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAAS
PNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENA
QKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNN
NNNLGIEGRMEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAA
ASYFL

Figure 2.08. Illustrative representation of the MBP-factorXa-TrnA(B) fusion protein.

2.3.2. Subcloning

Polymerase chain reaction was used to amplify TrnA and TrnB from TrnAB_pUC57 with the following primers:

TrnA_pMAL Forward Primer: ATGGAAGTCATGAACAATGCACTG

TrnA_pMAL Reverse Primer:

GGCGAAGCTTGAGGCCCTTTCGTCTCGCGCGTTTCG

TrnB_pMAL Forward Primer: ATGGAAGTCCTAAACAAACAGAACG

TrnB_pMAL Reverse Primer:

CGGTGTGAAGCTTCGCACAGATGCGTAAGGAGAAAATACC

The polymerase chain reactions were carried out following the methods described in Figure 2.07 and 2.08:

Table 2.08 KOD PCR components.

Component	µL	Final concentration
Polymerase chain reaction grade H_2O	32.0	-
KOD Hot Start 5x Buffer	5.0	1X

Component	µL	Final concentration
MgSO4 (25mM)	3.0	1.5mM
dNTPs (2mM each)	5.0	200 µM each
Primer A (10 µM)	1.5	0.3 µM
Primer B (10 µM)	1.5	0.3 µM
Template DNA (10 ng/ µL)	1.0	0.2 ng/ µL
KOD DNA Polymerase	1.0	0.02 Units/ µL
Total	50 µL	

Table 2.09. KOD PCR cycling conditions.

Cycle step	Temperature (°C)	Time	No. Cycles
Polymerase activation	95	2 m	1
Denaturing	95	20 s	20
Annealing	60	10 s	
Extension	70	1.5 s	

*Further discussion will refer to this protocol as the *polymerase chain reaction protocol*.

The polymerase chain reaction products, along with pMALc2x, were digested with restriction enzymes XmnI (Promega) and HindIII (Promega) following the *restriction digest protocol* (page 40). Thermosensitive Alkaline Phosphatase (Promega) was added to the pMALc2x digestion after 3 hours of incubation. After digestion, the products were purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System (Promega).

Following the restriction digest purification, the products were ligated following the *ligation protocol* (page 40). Subsequent products were used to transform competent

DH5α (Promega) and screened on ampicillin (50 μg/mL) resistant Luria broth-agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 μg/mL) and incubated overnight for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

To examine the correctness of the isolated DNA samples, TrnA_pMALc2x and TrnB_pMALc2x were treated with XmnI (Promega), following the *restriction digest protocol* (page 40), to generate linear DNA to be analyzed via agarose gel electrophoresis (control: pMALc2x). Expecting a product approximately 1300 base pairs larger than the control, Figure 2.09 shows two bands (lanes 3-4) of the correct size.

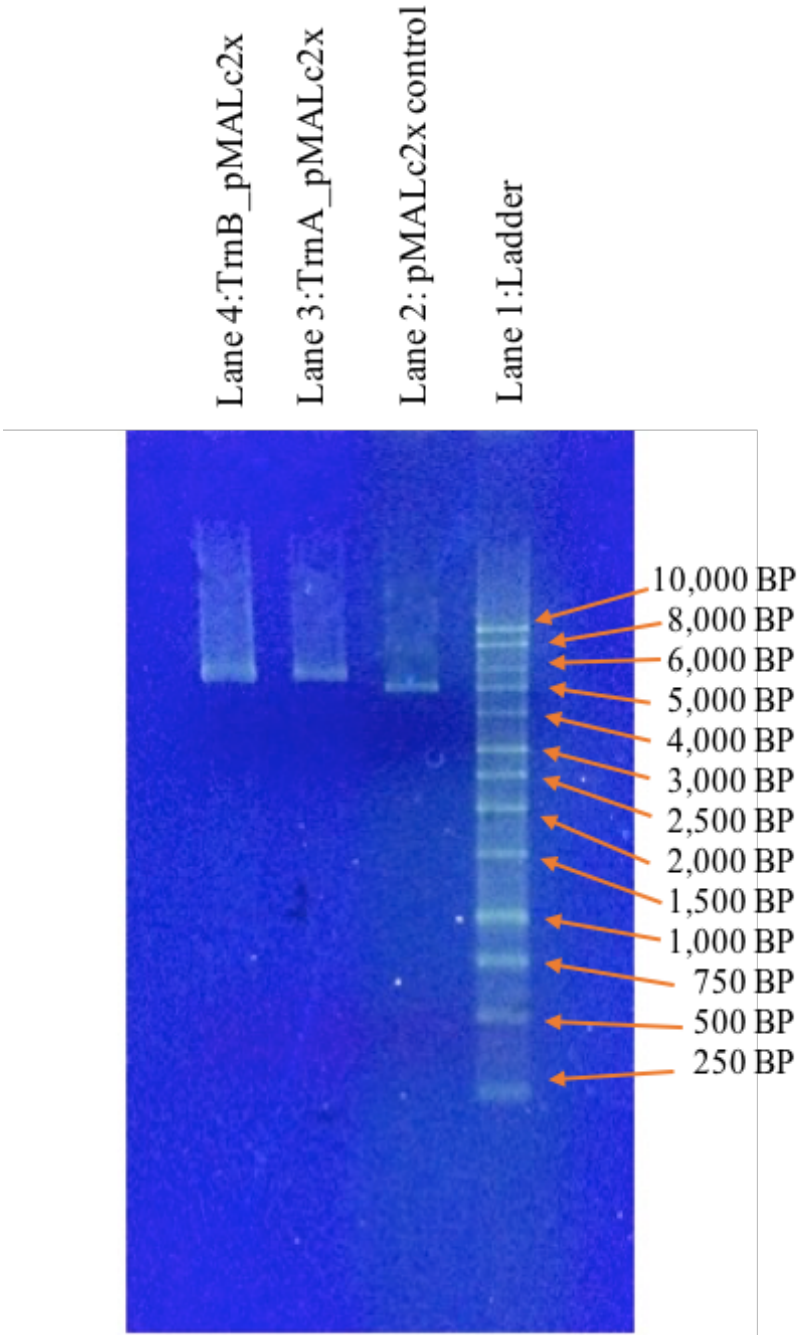


Figure 2.09. Agarose gel depicting TrnA_pMALc2x (lane 3) and TrnB_pMALc2x (lane 4). Lane 1, DNA ladder; lane 2, pMALc2x control.

The TrnA_pMALc2x and TrnB_pMALc2x samples from lanes 3 and 4 (Figure 2.09), respectively, was used to transform competent BL21(DE3)pLysS for protein expression and purification.

2.3.3. Protein Expression and Purification

Two 50 mL Luria broth cultures of TrnA_pMALc2x and TrnB_pMALc2x in BL21(DE3) pLysS were screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant shaking at 225 rpm. After incubation, 10 mL each were added separately to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

Following defrosting, the pellets were re-suspended in 50mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Following centrifugation, ion-exchange chromatography was performed using FPLC (BioRad). DEAE Sepharose Fast Flow ion exchange resin (GE Healthcare) was first equilibrated with 50 mM Tris-base, pH 8, and after the loading of the lysate followed a 300 mL linear gradient of 50 mM Tris-base, pH 8 to 50 mM Tris-base, pH 8 0.6 M NaCl to elute bound protein. Ten mL factions were collected through the entirety of the gradient and were later pooled and concentrated using an Amicon Ultra-15 Centrifugal

Filter Unit, by Millipore, to be identified through sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Figure 2.10 depicts full length MBP-factorXa-TrnA (lane 4; M_{av}: ~47,100 Daltons) and MBP-factorXa-TrnB (lane 2; M_{av}: ~47,450 Daltons), and factorXa cleavage products MBP-factorXa-TrnA (lane 5) and MBP-factorXa-TrnB (lane 1).

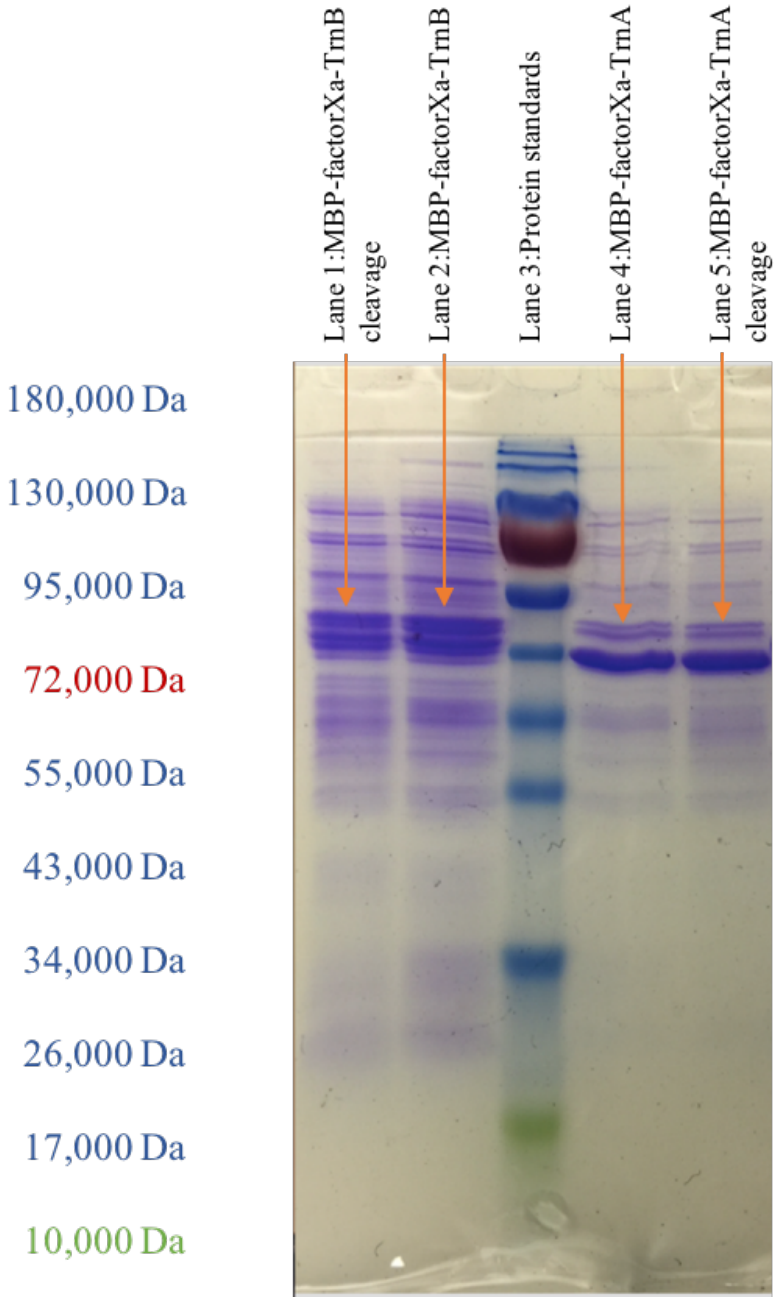


Figure 2.10. SDS PAGE depicting full length MBP-factorXa-TrnA (lane 4) and MBP-factorXa-TrnB (lane 2), and the factorXa cleavage product of MBP-factorXa-TrnA (lane 5) and MBP-factorXa-TrnB (lane 1). Lane 3, protein standards (colors observed in the ladder to the left correspond to the colors observed on the SDS PAGE).

2.3.4. Cleavage

The factor Xa cleavage conditions used are as follows:

Factor Xa cleavage conditions: 1mg factorXa protease/ 50 mg target protein in 50 mM Tris-base, pH 8. Incubate for 16 hrs at 30-37°C.

Lanes 1 and 5 in Figure 2.10 are the factor Xa cleavage products of MBP-factorXa-TrnB and MBP-factorXa-TrnA, respectively. A successful cleavage would yield MBP-factorXa (M_{av}= 42,325.8) and TrnA (M_{av}= 4,814.5); and MBP-factorXa (M_{av}= 42,325.8)and TrnB (M_{av}= 5,148.8).

2.3.5. Results

After expression and purification of MBP-factorXa-TrnA and MBP-factorXa-TrnB, observation of the sodium dodecyl sulfate-polyacrylamide gel electrophoresis suggested the product was isolated in low-moderate yield, as is implied by the presence of bands at approximately 47,500 Daltons. As a result of using the low resolution method of anion exchange chromatography, many contaminant proteins were present amongst the MBP-factorXa-TrnA and MBP-factorXa-TrnB, further indicating a more efficient means of purification was needed.

Initial attempts to cleave the propeptides from maltose binding protein with factorXa were unsuccessful. It was unclear whether this result was from hindrance of the factorXa recognition site due to secondary structure formation by the TrnA and TrnB propeptides, or a result of general low activity of the protease, as has been previously reported.

After analysis of the results observed in the MBP-factorXa-TrnA and MBP-factorXa-TrnB constructs, it became necessary to introduce not only a more efficient method of purification, but also a more specific recognition site for cleavage of the propeptides.

2.4. MBP-TEV-TrnA, M_{av}= 47,405.6 g/mol; MBP-TEV-TrnB, M_{av}= 47,739.9 g/mol

2.4.1. Background

The previous maltose binding protein-fusion constructs contained a factorXa cleavage site for removal of the maltose binding protein from propeptides TrnA and TrnB. After failed attempts to observe the cleavage of the fusion protein, it was determined a more efficient and specific protease should be utilized. Tobacco etch virus protease is a commonly used protease that shows high sequence specificity. The most common recognition site is ENLYFQG, with cleavage occurring between Q/G.

Polymerase chain reaction primers were designed to remove the factorXa cleavage site and replace it with the Tobacco Etch Virus protease recognition site. Ligation-free polymerase chain reaction was utilized to make this modification, and the products were used to transform competent DH5α for DNA amplification. Agarose gel electrophoresis was carried out to confirm the base pair length of the resulting product.

Pictured below are the amino acid sequences for both MBP-TEV-TrnA and MBP-TEV-TrnB. In black is the maltose binding protein; blue, the newly incorporated tobacco etch virus protease recognition site; and orange, the TrnA (B) sequence.

A:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAAT
GDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPI
AVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAA

DGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA
FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAAS
PNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENA
QKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNN
NNNG**ENLYFQG**MEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAF
TLG*

B:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAAT
GDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPI
AVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAA
DGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA
FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAAS
PNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENA
QKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNN
NNNG**ENLYFQG**MEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEF
AAASYFL*

Figure 2.11. Illustrative representation of MBP-TEV-TrnA(B).

2.4.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **MBP-factorXa-TrnA** and **MBP-factorXa-TrnB** to **MBP-TEV-TrnA** and **MBP-TEV-TrnB** from TrnA_pMALc2x and TrnB_pMALc2x, respectively, with the following primers:

MBP-TEV-TrnA Forward Primer:

GAGAATCTTTATTTTCAGGGAATGGAAGTCATGAACAATGCACTGATTACC

MBP-TEV-TrnB Forward Primer:

GAGAATCTTTATTTTCAGGGAATGGAAGTCCTGAACAAACAGAACG

MBP-TEV-Trn (A, B) Reverse Primer:

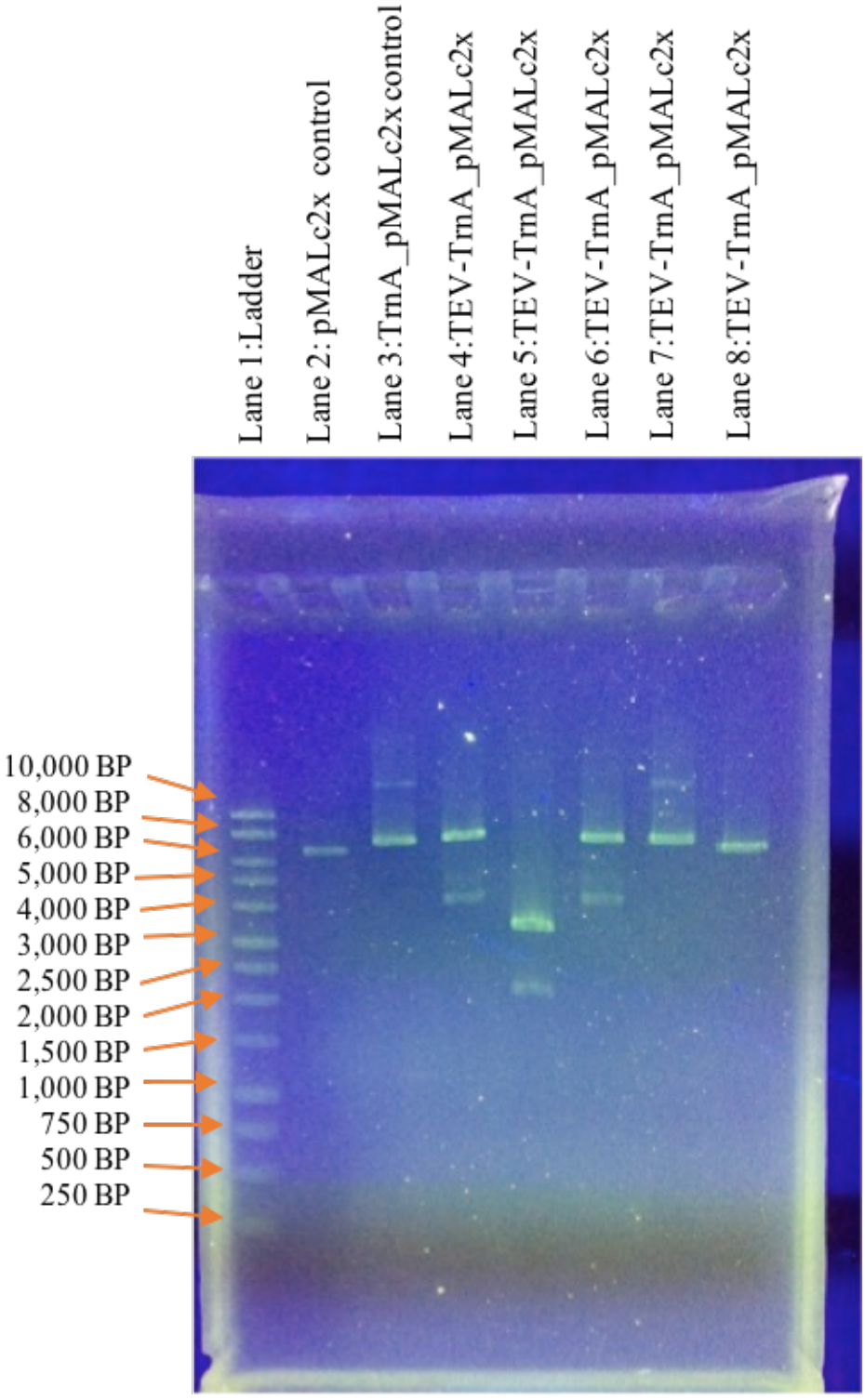
TCCCTGAAAATAAAGATTCTCCCCGAGGTTGTTGTTATTGTTATTGTTGTTGTT

The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol* (page 46) with the following modifications:

- Annealing temperature: 55°C.
- Extension: 3 m

Products were used to transform competent DH5α (Promega), screening with ampicillin (50 μg/mL) resistant Luria agar plates. Successful colonies were inoculated into Luria broth media for overnight incubation, screening with ampicillin (50 μg/mL), for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests with HindIII (Promega) were carried out to determine if the isolated product contained the correct number of base pairs, following the *restriction digest protocol* (page 40). DNA was assumed correct if the band length was approximately equal to the respective TrnA_pMALc2x or TrnB_pMALc2x controls. As can been observed in the agarose gel electrophoresis images in Figure 2.12, both TEV-TrnA_pMALc2x (lanes 6 and 7) and TEV-TrnB_pMALc2x (lane 5) appear to be the correct base pair length (controls: lane 2, left and right, pMALc2x; lane 3, left, TrnA_pMALc2x; lane 3, right, TrnB_pMALc2x)


Lane 1:Ladder
Lane 2: pMALc2x control
Lane 3:TrnA_pMALc2x control
Lane 4:TEV-TrnA_pMALc2x
Lane 5:TEV-TrnA_pMALc2x
Lane 6:TEV-TrnA_pMALc2x
Lane 7:TEV-TrnA_pMALc2x
Lane 8:TEV-TrnA_pMALc2x
10,000 BP
8,000 BP
6,000 BP
5,000 BP
4,000 BP
3,000 BP
2,500 BP
2,000 BP
1,500 BP
1,000 BP
750 BP
500 BP
250 BP

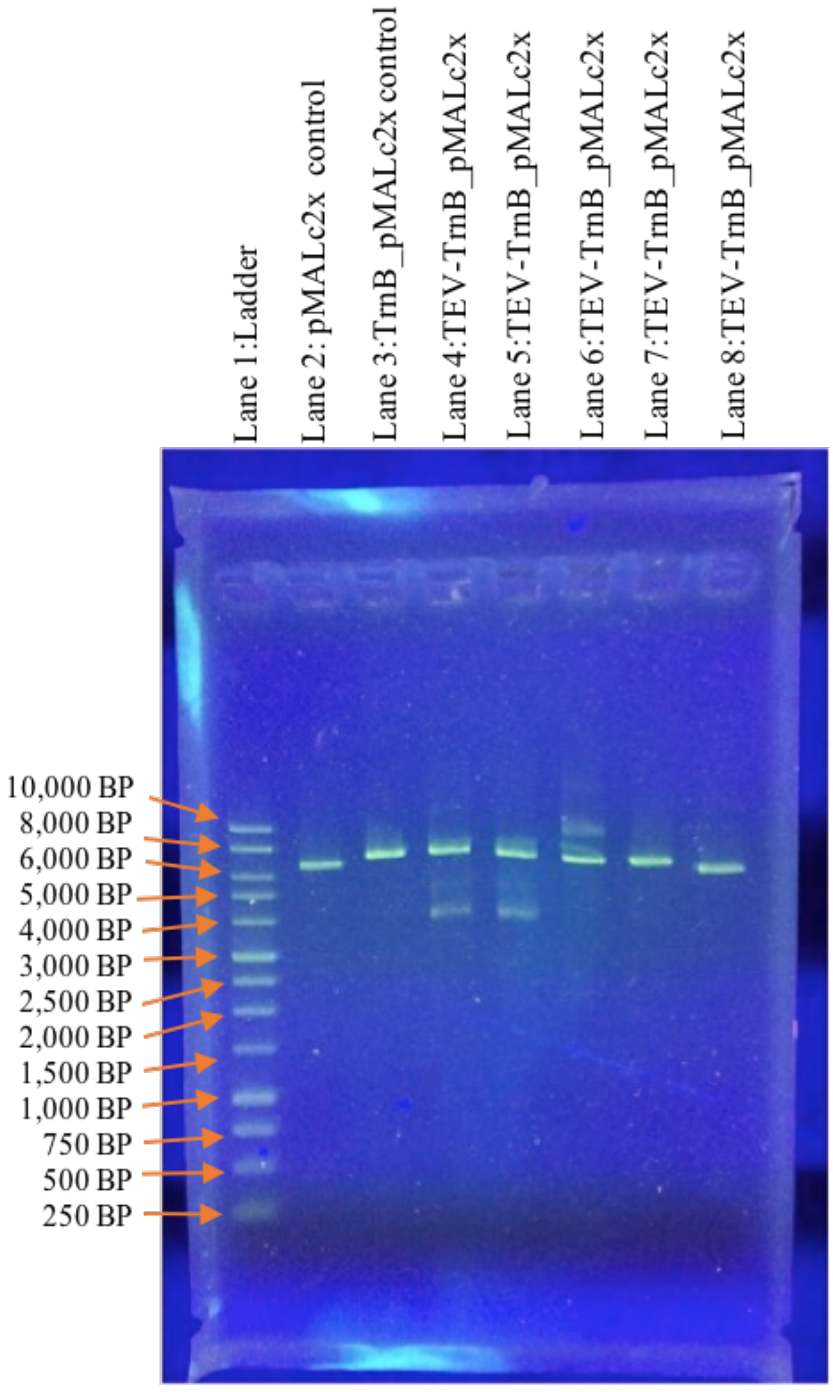


Figure 2.12. (top) Agarose gel depicting TEV-TrnA_pMALc2x (lanes 4-8). Lane 1, DNA ladder; lane 2, pMALc2x control; lane 3, TrnA_pMALc2x control. (bottom) Agarose gel depicting TEV-TrnB_pMALc2x (lanes 4-8). Lane 1, DNA ladder; lane 2, pMALc2x control; lane 3, Trn B_pMALc2x control.

2.4.3. Results

After observation of the agarose gel electrophoresis images (Figure 2.12), it was evident that the isolated DNA in lanes 6-7 in TEV-TrnA_pMALc2x and lane 5 in TEV-TrnB_pMALc2x (above) contained approximately the correct base pair count. With no method to confirm or deny the success of the ligation-free polymerase chain reaction removal of the factorXa recognition site and addition of the tobacco etch virus recognition site, other than by sequencing, it was assumed the reaction was successful.

2.5. His_6MBP-TEV-TrnA, M_{av}= 48,187.3 g/mol; His_6MBP-TEV-TrnB, M_{av}= 48,521.6 g/mol

2.5.1. Background

Constructs described previously made use of DEAE ion exchange as a means of purification. As this is a useful method to initially purify a protein expression, it is not a practical approach for specific isolation of the target protein. Making use of an affinity tag is a method used to purposely target a protein, which greatly enhances the purity. A common affinity tag is a poly-histidine tag, which utilizes 6-10 consecutive histidine residues at either the N-terminus or C-terminus to bind to Ni^{2+}, Co^{2+}, or Zn^{2+} containing resins, a method sometimes referred to as Nickel (Cobalt, Zinc) affinity chromatography. After the resin is washed, the protein is eluted with elevated concentrations of imidazole (150-300 mM). Desalting is commonly necessary as high concentrations of imidazole interfere with ice crystal formation during freezing, which may lead to protein degradation throughout storage. If the histidine becomes charged, the affinity tag will not bind to the resin, therefore, buffers with a pH $\geq$ 8, are necessary to keep the histidine in an uncharged state.

Ligation-free polymerase chain reaction primers were designed to incorporate a hexahistidine affinity tag at the N-terminus of maltose binding protein in the TEV-TrnA_pMALc2x and TEV-TrnB_pMALc2x constructs. The products were later used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to verify the base pair length of the amplified DNA, which was next used to transform competent BL21 (DE3) pLysS (Invitrogen) for protein expression.

After protein expression, the resulting crude product was purified via nickel affinity chromatography and identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below are the amino acid sequences for both His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB. In red is the newly incorporated hexahistidine affinity tag; black, the maltose binding protein; blue, the tobacco etch virus protease recognition site; and orange, the TrnA (B) sequence.

A:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEK
FPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYN
GKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYF
TWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADT
DYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRI
AATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSS
NNNNNNNNNNGENLYFQGMEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGI
GSLVGTAFTLG*

B:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEK
FPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYN
GKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYF
TWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADT

DYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL

SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRI

AATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSS

NNNNNNNNNNGENLYFQGMEVLNKQNVNIIPESEEVGGWVACVGACGTVCLA

SGGVGTEFAAASYFL*

Figure 2.13. Illustrative representation of His_6MBP-TEV-TrnA(B).

2.5.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **MBP-TEV-TrnA** and **MBP-TEV-TrnB** to **His_6MBP-TEV-TrnA** and **His_6MBP-TEV-TrnB** from TEV-TrnA_pMALc2x and TEV-TrnB_pMALc2x, respectively, with the following primers:

His_6-TEV-Trn (A, B) _pMALc2x Forward Primer:

AGCCATCACCATCACCATCACCACATCGAAGAAGGTAAACTGGTAATCTGGA

TTAACG

His_6-TEV-Trn (A, B) _pMALc2x Reverse Primer:

ATGTGGTGATGGTGATGGTGATGGCTCATATGCTATGGTCCTTGTTGGTGAAG

TGC

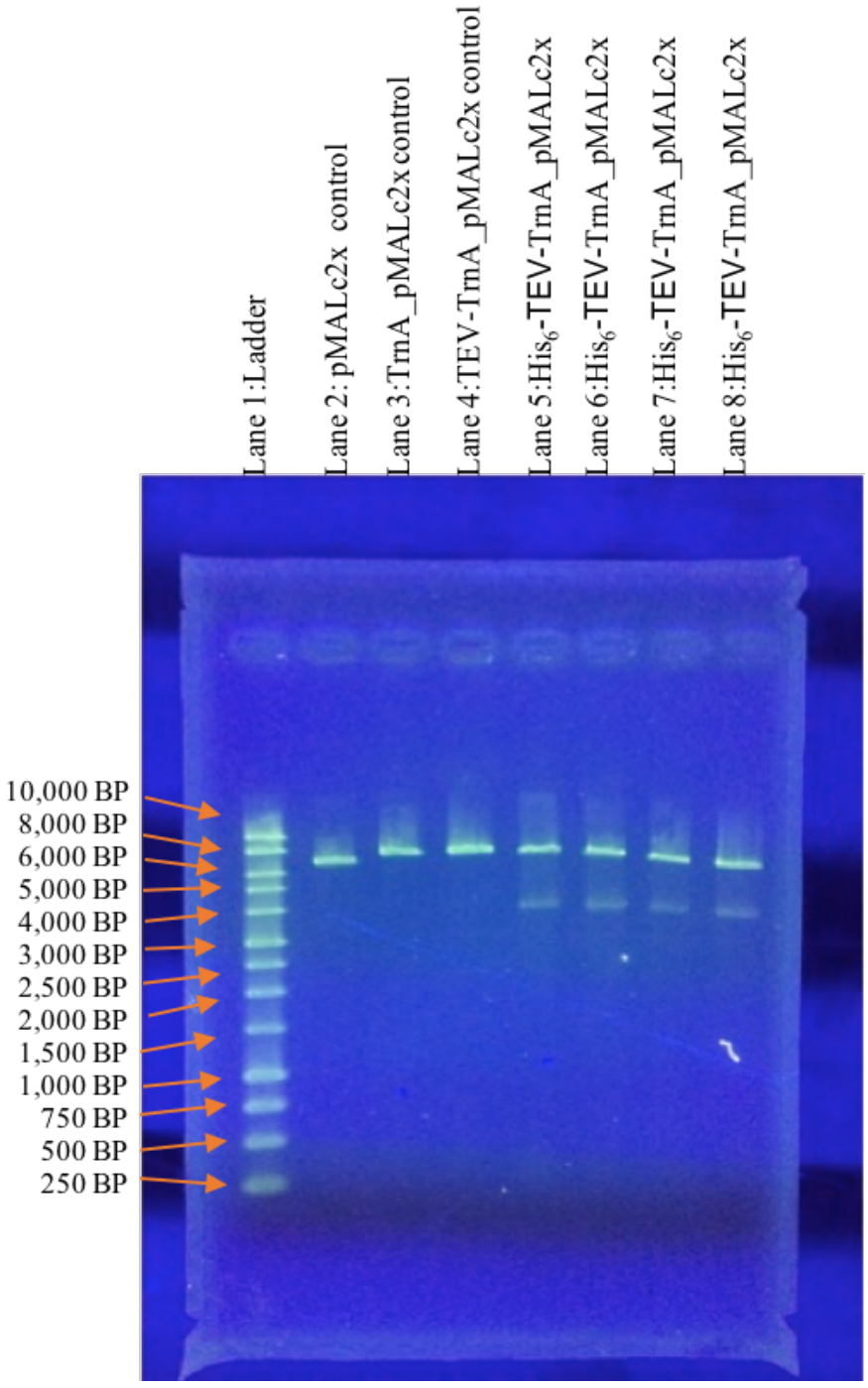
The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol* (page 46) with the following modifications:

- Annealing temperature: 55°C.
- Extension: 3 m
- Separate reactions were carried out for the forward and reverse primers for each construct, e.g., His_6-TEV-TrnA_pMALc2x was carried out as a normal ligation-free polymerase chain reaction; however, only the forward primer was added (additional water replaced the volume for the reverse primer). The same was repeated for the TEV-TrnA_pMALc2x reverse primer, as well as both forward and reverse primers for His_6-TEV-TrnB_pMALc2x. The reactions were allowed to cycle five times, at which point they were stopped and each reverse primer reaction was added to the respective forward primer reaction. The polymerase chain reaction was then allowed to continue for the remaining 15 cycles. This technique was utilized to prevent the forward and reverse primers from binding before the polymerase chain reaction was started. Further discussion will refer to this as the *ligation-free polymerase chain reaction protocol**.

Products were used to transform competent DH5α (Promega), screening with ampicillin (50 μg/mL) resistant Luria both agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 μg/mL) and incubated overnight at 37°C with constant spinning at 225 rpm for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests were carried out following the previous methods described in the TEV-TrnA_pMALc2x and TEV-TrnB_pMALc2x constructs; however, TEV-TrnA_pMALc2x and TEV-TrnB_pMALc2x were also used as controls (lanes 4) (pMALc2x, lanes 2; TrnA_pMALc2x and TrnB_pMALc2x control, lanes 3). As

observed in Figure 2.14, the His_6-TEV-TrnA_pMALc2x and His_6-TEV-TrnB_pMALc2x in lanes 5 appeared to be the correct base pair length, given they are approximately equal to the respective TEV-TrnA_pMALc2x and TEV-TrnB_pMALc2x controls.


Lane 1:Ladder
Lane 2: pMALc2x control
Lane 3:TmA_pMALc2x control
Lane 4:TEV-TmA_pMALc2x control
Lane 5:His6-TEV-TmA_pMALc2x
Lane 6:His6-TEV-TmA_pMALc2x
Lane 7:His6-TEV-TmA_pMALc2x
Lane 8:His6-TEV-TmA_pMALc2x
10,000 BP
8,000 BP
6,000 BP
5,000 BP
4,000 BP
3,000 BP
2,500 BP
2,000 BP
1,500 BP
1,000 BP
750 BP
500 BP
250 BP

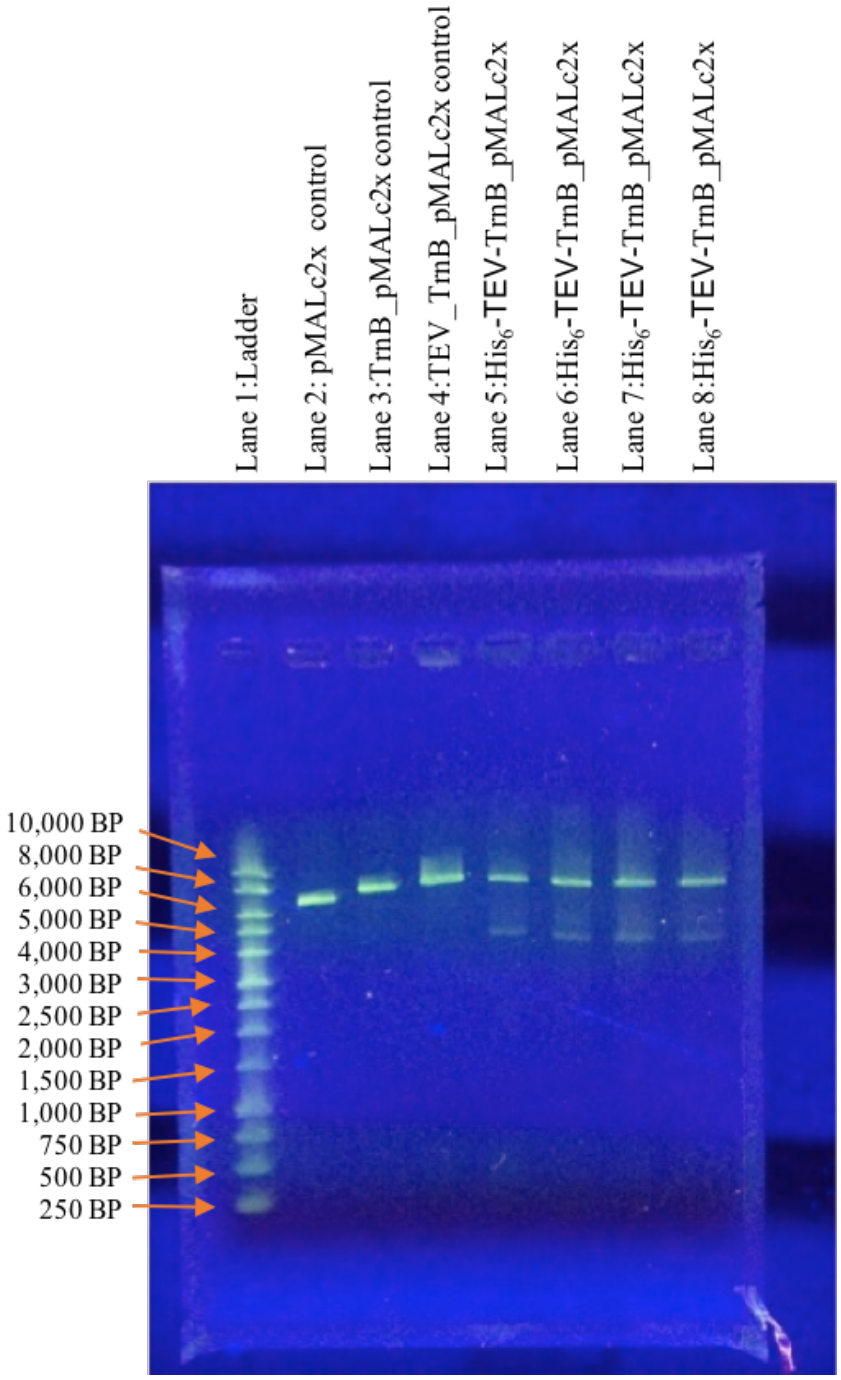


Figure 2.14. (top) Agarose gel depicting His_6-TEV-TrnA_pMALc2x (lanes 5-8). Lane 1, DNA ladder; lane 2, pMALc2x control; lane 3, TrnA_pMALc2x control; lane 4, TEV-TrnA_pMALc2x control. (right) Agarose gel depicting His_6-TEV-TrnB_pMALc2x (lanes 5-8). Lane 1, DNA ladder; lane 2, pMALc2x control; lane 3, TrnB_pMALc2x control; lane 4, TEV-TrnB_pMALc2x control.

After analyzing the base pair length of the samples, His_6-TEV-TrnA_pMALc2x and His_6-TEV-TrnB_pMALc2x from lanes 5 (Figure 2.14) were used to transform competent BL21 (DE3) pLysS (Invitrogen) and screened on ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) resistant Luria broth agar plates.

2.5.3. Protein Expression and Purification

Two 50 mL Luria broth cultures of His_6-TEV-TrnA_pMALc2x and His_6-TEV-TrnB_pMALc2x in BL21 (DE3) pLysS were screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL each were added separately to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellets were re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

The centrifuged His_6-MBP-TEV-TrnA and His_6-MBP-TEV-TrnB crude lysates were subsequently purified by nickel affinity chromatography using FPLC (Bio-Rad) under the following protocol (further discussion will refer to this protocol as the *nickel affinity chromatography protocol*):

-Ni-NTA-agarose (2 x 15 cm) (Bio-Rad) is first pre-equilibrated with buffer A (50 mM Tris-base, 0.5 M NaCl, pH 8.0). Protein lysate is then loaded and washed with 90:10 buffer A: buffer B (50 mM Tris-base, 200 mM imidazole, 0.5 M NaCl, pH 8.0) until flat baseline is achieved. Elution is achieved with 100% buffer B. Eluted protein is collected and concentrated to approximately 10 mL with a Millipore Amicon Ultra-15 Centrifugal Filter Unit. Imidazole is removed by passing through BioGel P2 desalting column and concentrated to desired concentration with Millipore Amicon Ultra-15 Centrifugal Filter Unit.

After purification was completed following the *nickel affinity chromatography protocol* (page 67), the concentrated His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB samples were identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Figure 2.15 depicts His_6MBP-TEV-TrnA (left; lane2; M_{av}: ~48,200 Daltons), His_6MBP-TEV-TrnB (right; lane 2; M_{av}: ~48,500 Daltons)

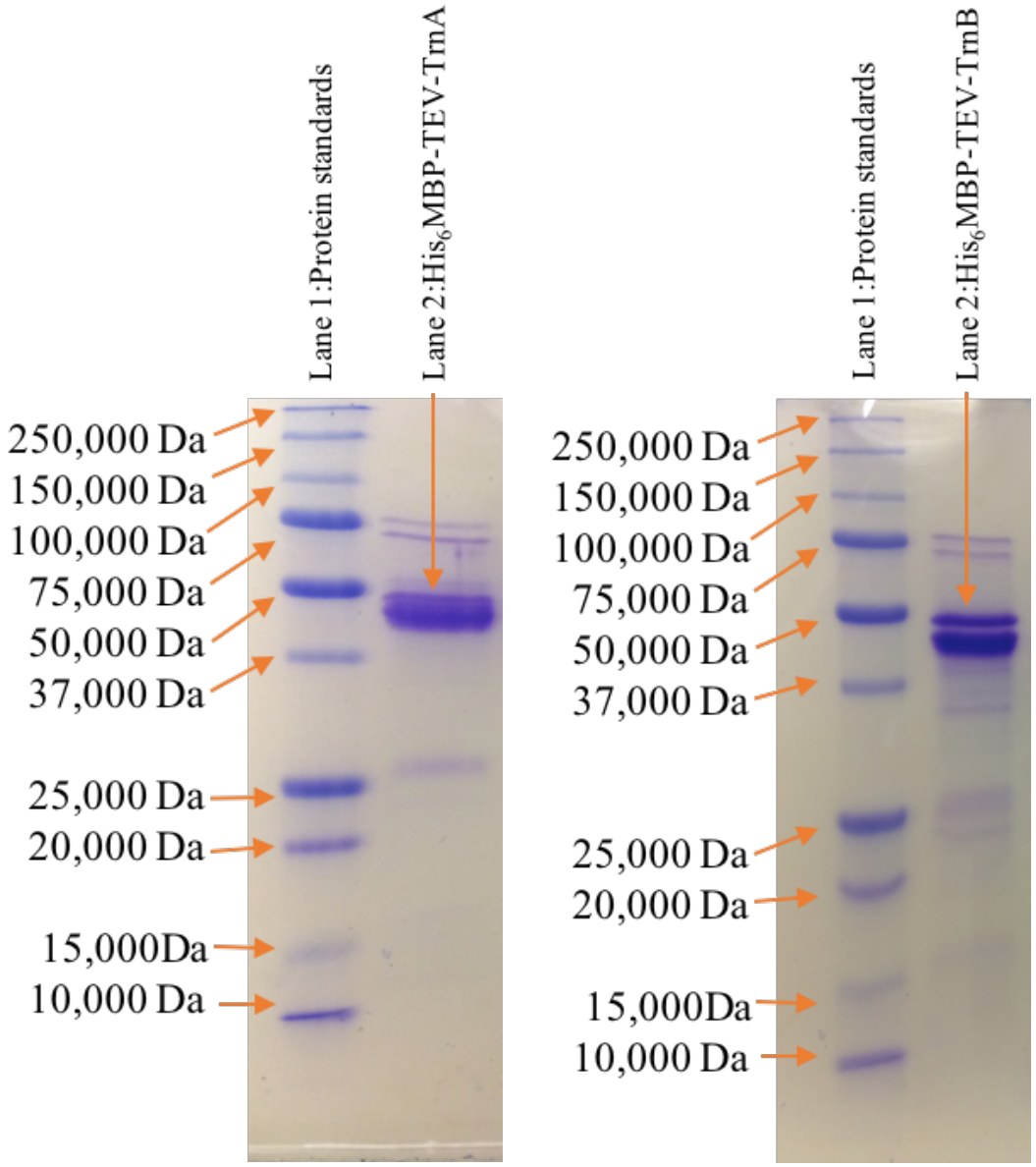


Figure 2.15. (Left) SDS PAGE depicting His_6MBP-TEV-TrnA. Lane 1, protein standards. (Right) SDS PAGE depicting His_6MBP-TEV-TrnB. Lane 1, protein standards.

2.5.4. Cleavage

The tobacco etch virus cleavage conditions used are as follows (further discussion will refer to these conditions as *Tobacco Etch Virus cleavage conditions*):

> *Tobacco Etch Virus cleavage conditions*: 0.005 mg TEV protease/ 0.1 mg target protein in 50 mM Tris-base, pH 8. Incubate for 3 hrs at 30-37°C or 16 hrs 4°C.

Lanes 1 and 5 in Figure 2.16 are the tobacco etch virus cleavage products of His_6MBP-TEV-TrnB and His_6MBP-TEV-TrnA, respectively. A successful cleavage would yield His_6MBP-TEV (M_{av}= 43,489.9 Daltons) and TrnA (M_{av}= 4,658.4 Daltons); and His_6MBP-TEV (M_{av}= 43,489.9 Daltons) and TrnB (M_{av}= 4,992.6 Daltons).

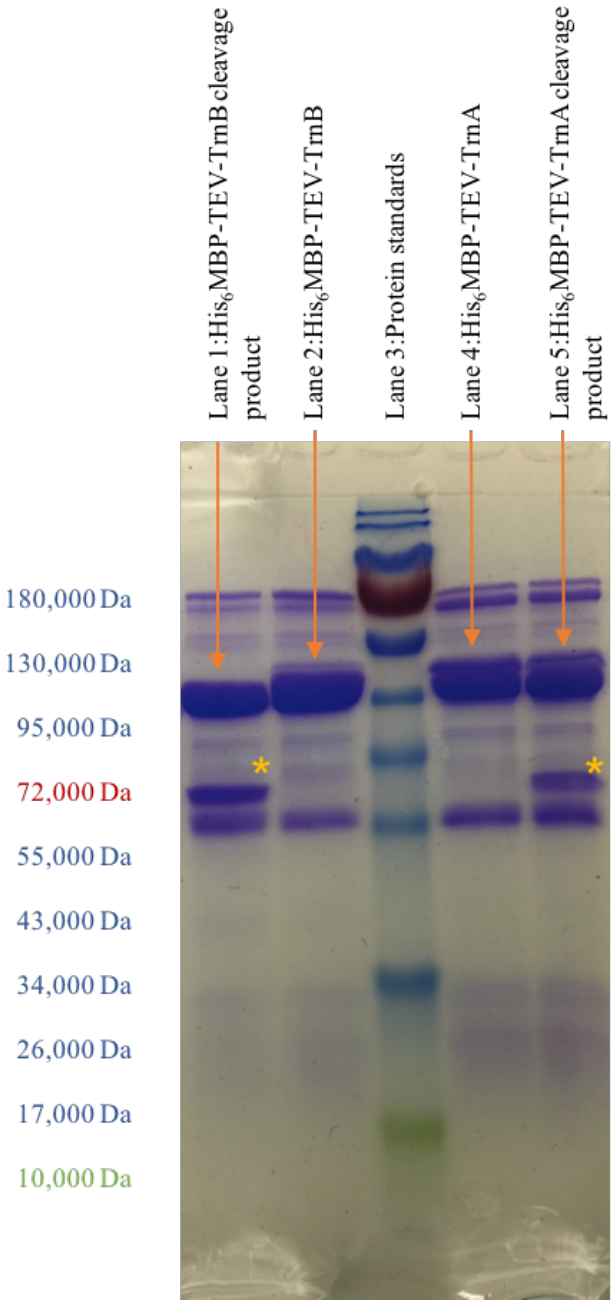


Figure 2.16. SDS PAGE depicting full length His_6MBP-TEV-TrnA (lane 4) and His_6MBP-TEV-TrnB (lane 2), and the TEV cleavage product of His_6MBP-TEV-TrnA (lane 5) and His_6MBP-TEV-TrnB, resulting in His_6MBP-TEV (lane 1). Lane 3, protein standards (colors observed in the ladder to the left correspond to the colors observed on the SDS PAGE).

* Protein bands correspond to TEV protease.

2.5.5. Results

After expression and purification of His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB, observation of the sodium dodecyl sulfate-polyacrylamide gel electrophoresis suggested the product was isolated in moderate yield, as is implied by the presence of bands at approximately 43,800 Daltons. Furthermore, it was evident the introduction of the hexahistidine sequence was successful as suggested by the elution peak observed during nickel affinity chromatography purification.

Likewise, the introduction of the tobacco etch virus recognition site was assumed successful as proteolysis reactions suggested cleavage was occurring, though not at a desirable percentage. After incubating for periods that should result in 80-90% cleavage, only 40-50% cleavage was observed. This indicated that the recognition site was somehow hindered such that the tobacco etch virus protease was unable to consistently access it.

Due to TrnA and TrnB occurring at the C-terminus of the maltose binding protein, the methionine-1 of both propeptides was no longer necessary, as it is a product of the AUG start codon in the original constructs. Mutation of the methionine to a smaller residue, such as glycine, was believed to offer less hindrance to tobacco etch virus protease recognition and therefore result in more successful cleavage. It has also been previously reported by Kapust et al. (2002) (13) that the presence of a serine residue, rather than a glycine, in the ENLYFQG' recognition sequence results in higher levels of cleavage.

2.6. G395S/M396G His_6MBP-TEV-TrnA, M_{av}= 48,143.2 g/mol; G395S/M396G His_6MBP-TEV-TrnB, M_{av}= 48477.5 g/mol

2.6.1. Background

In attempts to optimize the tobacco etch virus protease recognition sequence, ENLYFQG'(M), site directed mutagenesis primers were designed to mutate the glycine residue to serine, as studies have shown that the protease has more specificity to the serine containing site, rather than the glycine. Also, the same primers were designed to mutate the methionine residue to glycine, as it was believed the methionine may be creating negative steric interactions during cleavage. These mutations also introduce a TCCGGA recognition site for the restriction enzyme BspE1.

Ligation-free polymerase chain reaction primers were designed to incorporate G395S and M396G mutations in the TEV-TrnA_pMALc2x and TEV-TrnB_pMALc2x constructs. The products were later used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to verify the base pair length of the amplified DNA, which was next used to transform competent BL21 (DE3) pLysS (Invitrogen) for protein expression. After protein expression, the resulting crude product was purified via nickel affinity chromatography and identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below are the amino acid sequences for both G395S/M396G His_6MBP-TEV-TrnA and G395S/M396G His_6MBP-TEV-TrnB. In red is the hexahistidine affinity tag; black, the maltose binding protein; blue, the tobacco etch virus protease recognition site;

purple, the newly incorporated and G395S and M396G mutations; and orange, the TrnA (B) sequence.

A:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSN
NNNNNNNNNGENLYFQSGEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGS
LVGTAFTLG*

B:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSN
NNNNNNNNNGENLYFQSGEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASG
GVGTEFAAASYFL*

Figure 2.17. Illustrative representation of G395S/M396G His_6MBP-TEV-TrnA(B)

2.6.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **His_6MBP-TEV-TrnA** and **His_6MBP-TEV-TrnB** to **G395S/M396G His_6MBP-TEV-TrnA** and **G395S/M396G His_6MBP-TEV-TrnB** using His6-TEV-TrnA_pMALc2x and His_6-TEV-TrnB_pMALc2x, respectively, with the following primers:

G395S/M396G His_6-TEV-TrnA_pMALc2x Forward Primer:

TTC AGT CCG GAG AAG TCA TGA ACA ATG C

G395S/M396G His_6-TEV-TrnB_pMALc2x Forward Primer:

TTC AGT CCG GAG AAG TCC TGA ACA AAC

G395S/M396G His_6-TEV-Trn (A, B) _pMALc2x Reverse Primer:

GGA CTT CTC CGG ACT GAA AAT AAA GAT TCT CC

The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol** (page 65) with the following modifications:

- Annealing temperature: 57.5°C.

Resulting products were used to transform competent DH5α (Promega), screening with ampicillin (50 µg/mL) resistant Luria broth-agar plates. Successful colonies were

inoculated into Luria broth medium, screening with ampicillin (50 μg/mL), and incubated overnight at 37°C with constant spinning at 225 rpm for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Isolated DNA was subsequently used to transform competent BL21(DE3)pLysS (Invitrogen) and screened on ampicillin (50 μg/mL) + chloramphenicol (25 μg/mL) resistant Luria broth agar plates.

2.6.3. Protein Expression and Purification

Two 50 mL Luria broth cultures of G395S/M396G His_6-TEV-TrnA_pMALc2x and G395S/M396G His_6-TEV-TrnB_pMALc2x in BL21 (DE3) pLysS were screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL each were added separately to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellets were re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

The centrifuged G395S/M396G His_6-MBP-TEV-TrnA and G395S/M396G His_6-MBP-TEV-TrnB crude lysates were subsequently purified by nickel affinity chromatography *nickel affinity chromatography protocol* (page 67)

After purification was completed, the concentrated G395S/M396G His_6MBP-TEV-TrnA and G395S/M396G His_6MBP-TEV-TrnB samples were identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Lanes 2 and 5 in Figure 2.18 depict G395S/M396G His_6MBP-TEV-TrnA (M_{av}: ~48,200 Daltons) and G395S/M396G His_6MBP-TEV-TrnB (M_{av}: ~48,500 Daltons), respectively.

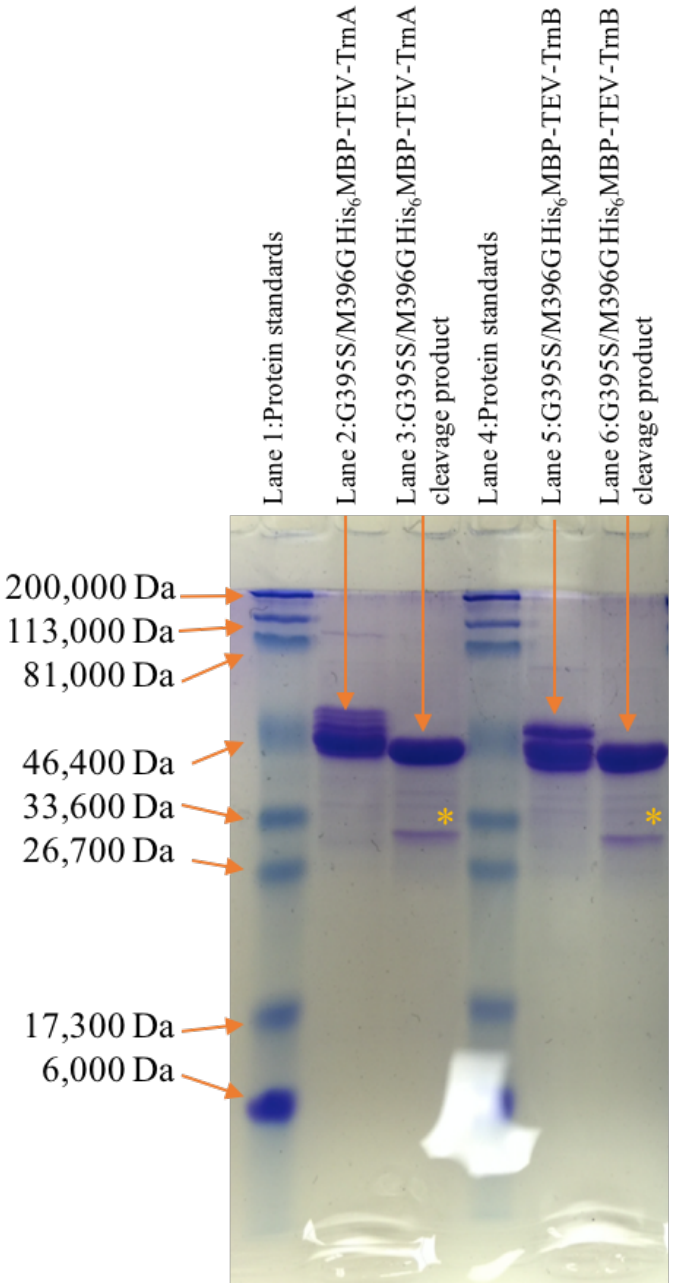


Figure 2.18. SDS PAGE depicting full length G395S/M396G His_6MBP-TEV-TrnA (lane 2) and G395S/M396G His_6MBP-TEV-TrnB (lane 5), and the TEV cleavage product of G395S/M396G His_6MBP-TEV-TrnA, resulting in His_6MBP-TEV (lane 3) and G395S/M396G His_6MBP-TEV-TrnB, resulting in His_6MBP-TEV (lane 6). Lanes 1 and 4, protein standards.

* Protein bands correspond to TEV protease.

2.6.4. Cleavage

Lanes 3 and 6 in Figure 2.18 are the tobacco etch virus cleavage products of His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB, respectively. A successful cleavage would yield His_6MBP-TEV (M_{av}= 43,489.9 Daltons) and TrnA (M_{av}= 4,671.3 Daltons); and His_6MBP-TEV (M_{av}= 43,489.9 Daltons) and TrnB (M_{av}= 5005.6 Daltons).

2.6.5. Results

With sequencing results indicating the DNA sequences were correct, efforts were next placed in the expression of the constructs. As can be observed in the sodium dodecyl sulfate polyacrylamide gel electrophoresis image His_6MBP-TEV-Trn_A (page 75), a heterogeneous product is co-purified. In the maltose binding protein sequence, there exists a decaasparagine sequence prior to the Tobacco Etch Virus recognition site. It was believed that the sequence was in some way interfering with the translation of the fusion protein, and therefore resulting in the observed heterogeneous product. Attempts to inhibit this potential interference are described in the following construct.

2.7. $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA, Mav= 48,578.6 g/mol; $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB, Mav= 48,912.9 g/mol

2.7.1. Background

As can be observed below in large bold print, both amino acid sequences contain a $poly_{10}$-asparagine sequence just prior to the Tobacco Etch Virus recognition site. This is also present in all previous pMALc2x constructs described above. It was believed that the sequence was interfering with the translation of His_6MBP-TEV-TrnA. To determine if this was occurring, polymerase chain reaction primers were designed to introduce a pentaserine sequence directly after the decaasparagine sequence in attempts to inhibit any structure formation.

Ligation-free polymerase chain reaction primers were designed to incorporate the pentaserine sequence into the G400S/M401G His_6MBP-TEV-TrnA and G400S/M401G His_6MBP-TEV-TrnB constructs. The products were later used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to verify the base pair length of the amplified DNA, which was next used to transform competent BL21 (DE3) pLysS (Invitrogen) for protein expression. After protein expression, the resulting crude product was purified via nickel affinity chromatography and identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis.

As mentioned, pictured below are the amino acid sequences for both $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA and $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB. In red is the hexahistidine affinity tag; black, the maltose binding protein; yellow, the newly

introduced pentaserine sequence; blue, the tobacco etch virus protease recognition site; purple, G400S and M401G mutations; and orange, the TrnA (B) sequence.

A:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSS**N**
NNNNNNNNNSSSSSGENLYFQSGEVMNNALITKVDEEIGGNAACVIGCIGSCVIS
EGIGSLVGTAFTLG*

B:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSS**N**
NNNNNNNNNSSSSSGENLYFQSGEVLNKQNVNIIPESEEVGGWVACVGACGTVC
LASGGVGTEFAAASYFL*

Figure 2.19. Illustrative representation of both $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA(B).

2.7.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **G395S/M396G His_6MBP-TEV-TrnA** and **G395S/M396G His_6MBP-TEV-TrnB** to **$Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA** and **$Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB** using G395S/M396G His_6-TEV-TrnA_pMALc2x and G395S/M396G His_6-TEV-TrnB_pMALc2x, respectively, with the following primers:

$Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-Trn (A, B) _pMALc2x Forward Primer:

AGC TCG AGT AGC AGT GGT GAG AAT CTT TAT TTT CAG TCC GGA GAA G

$Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-Trn (A, B) _pMALc2x Reverse Primer:

ACT GCT ACT CGA GCT GTT GTT GTT ATT GTT ATT GTT GTT GTT GTT CGA GC

The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol** (page 65) with the following modifications:

- Annealing temperature: 60.5°C.

Resulting products were used to transform competent DH5α (Promega), screening with ampicillin (50 µg/mL) resistant Luria agar plates. Successful colonies were inoculated

into Luria broth medium, screening with ampicillin (50 μg/mL), and incubated overnight at 37°C with constant spinning at 225 rpm for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests with HindIII (Promega) were carried out to generate linear DNA to determine if the correct product was generated following the *restriction digest protocol* (page 40). As can been observed in Figure 2.20, expecting a product approximately equal in base pairs to the control, the image below demonstrates two bands (lanes 3 and 6, $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6-TEV-TrnA_pMALc2x and $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6-TEV-TrnB_pMALc2x, respectively) of roughly the correct size (Control: TrnB_pMALc2x).

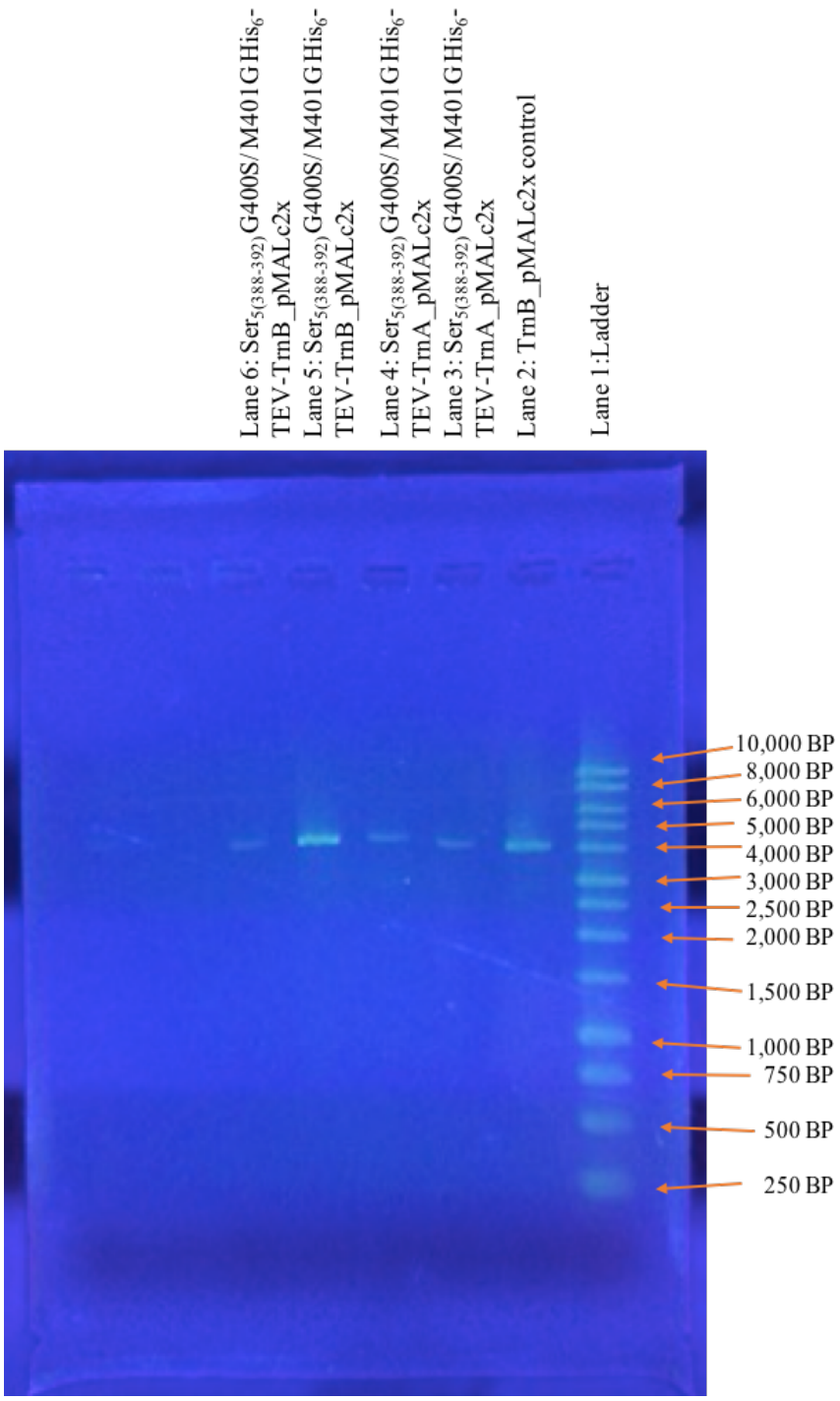


Figure 2.20. Agarose gel depicting $Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-TrnA_pMALc2x (lanes 3and 4) and $Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-TrnB_pMALc2x (lanes 5 and 6). Lane 1, DNA ladder; lane 2, pMALc2x control.

The DNA observed in lanes 3 and 6 ($Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-TrnA_pMALc2x and $Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-TrnB_pMALc2x, respectively) (Figure 2.20) was subsequently used to transform competent BL21 (DE3) pLysS (Invitrogen) and screened on ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) resistant Luria broth agar plates.

2.7.3. Protein Expressions and Purification

Two 50 mL Luria broth cultures of $Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-TrnA_pMALc2x and $Ser_{5(388-392)}$ G400S/ M401G His_6-TEV-TrnB_pMALc2x in BL21 (DE3) pLysS were screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL each were added separately to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellets were re-suspended in 50mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by following the *Nickel affinity chromatography protocol* (page 67).

2.7.4. Proteolysis inhibition

From the glycine gel sodium dodecyl sulfate polyacrylamide gel electrophoresis results for $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6MBP-TEV-TrnA, it was suggested the heterogeneous fusion protein product might possibly be undergoing C-terminal proteolysis during purification. To inhibit this proteolysis, several methods were employed, the first of which including the use of chemical inhibitors. Two separate protein expressions of $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6MBP-TEV-TrnA were carried out following the methods described above; however, before purification, 1 mg of TLCK was added to one protein lysate and 10 mg of PMSF were added to the other. After purification was completed following the *Nickel affinity chromatography protocol* (page_), the concentrated $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6MBP-TEV-TrnA samples were identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis. As observed in Figure 2.21, no significant inhibition of proteolysis was observed.

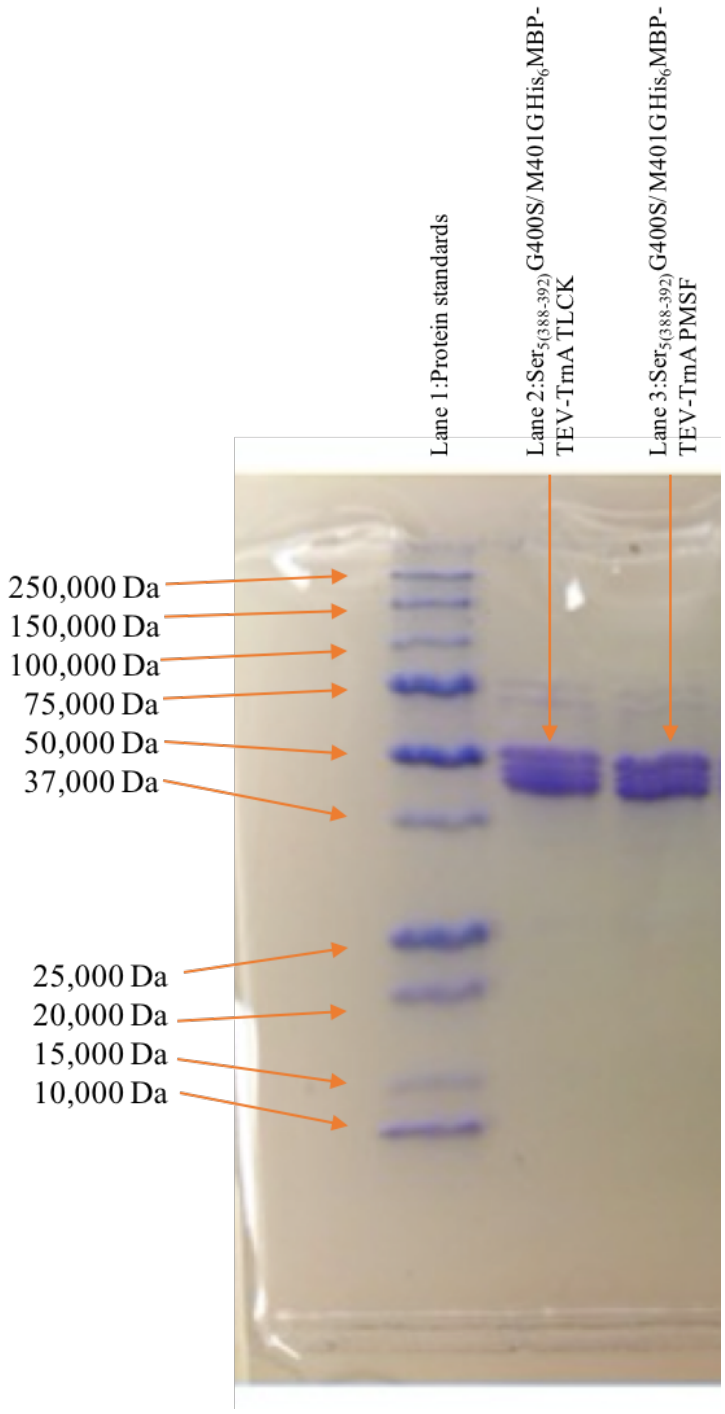


Figure 2.21. SDS PAGE depicting $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA sonication and purification with TLCK (lane 2) and PMSF (lane 3). Lane 1, protein standards.

In a second attempt to inhibit C-terminal proteolysis, five separate protein expressions of $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6MBP-TEV-TrnA were carried out following the methods described above; however, modifications were made as follows:

- Media contained 0.8% glucose; 2) EDTA (5 mM) was added during cell pellet re-suspension; 3) combination of methods 1 and 2; 4) protein expression was incubated for 16 hrs at 20 °C; 5) combination of methods 1 and 2, protein expression was incubated for 16 hrs at 20°C.

After purification was completed following the *Nickel affinity chromatography protocol* (page 67), the concentrated $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6MBP-TEV-TrnA samples were identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis. As observed in Figure 2.22, no complete inhibition of proteolysis was observed.

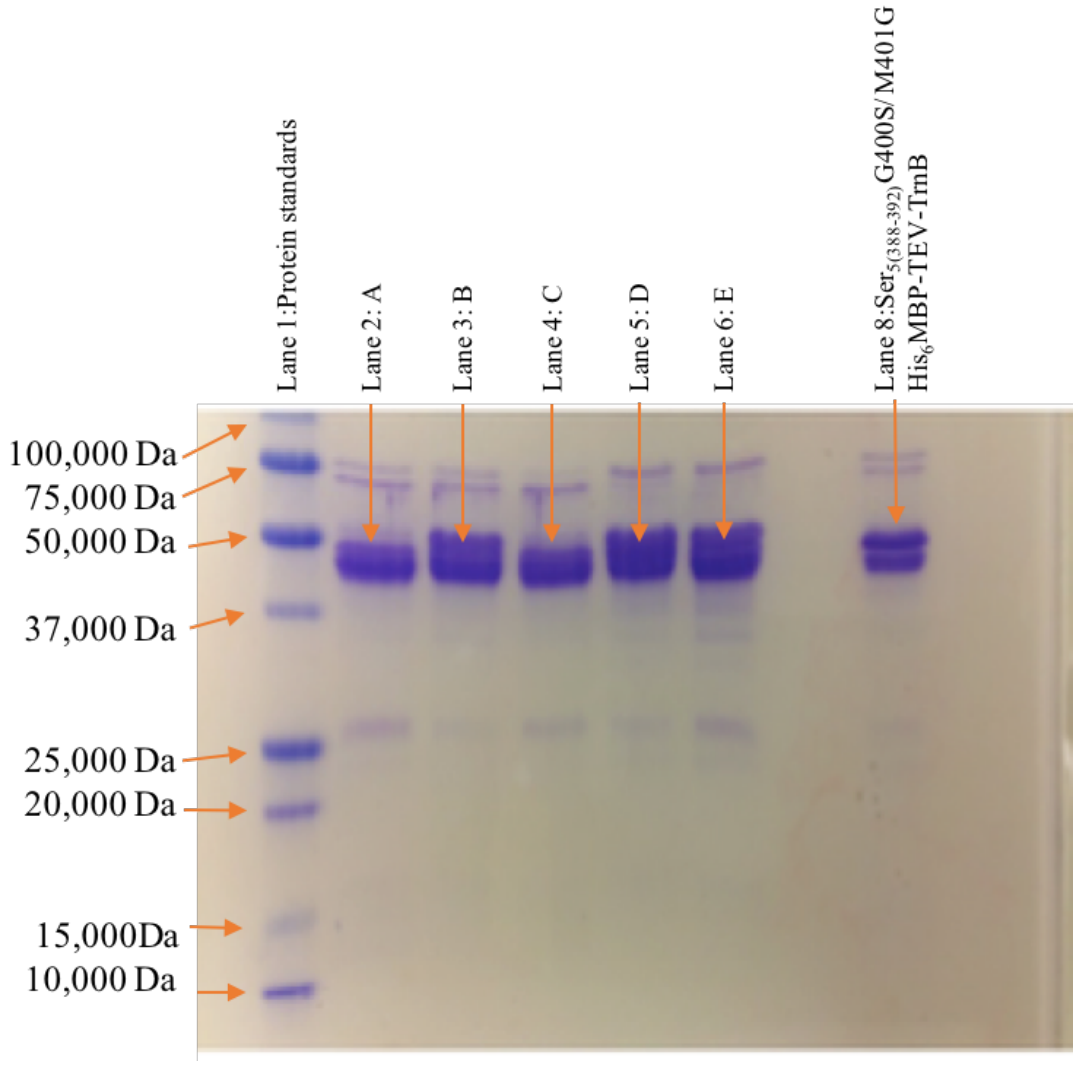


Figure 2.22. SDS PAGE depicting modified expressions and purifications of $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA. A) Media contained 0.8% glucose (lane 2); B) EDTA (5 mM) was added during cell pellet re-suspension (lane 3); C) combination of methods 1 and 2 (lane 4); D) protein expression was incubated for 16 hrs at 20 °C (lane 5); E) combination of methods 1 and 2, protein expression was incubated for 16 hrs at 20°C (lane 6). Lane 1, protein standards; lane 8, $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB.

2.7.5. Results

Attempts to inhibit the production of a heterogeneous $Ser_{5(388\text{-}392)}$ G400S/ M401G His_6MBP-TEV-TrnA product by both the addition of a $poly_5$-serine sequence and the use of methods to prevent proteolysis were moderately successful only in expressions containing 0.8% glucose. As can be observed in the sodium dodecyl sulfate polyacrylamide gel electrophoresis images, there still existed a mixture of products ranging from approximately 48,000- 52,000 Daltons. In a final effort to inhibit the production of these ambiguous products, complete removal of the $poly_{10}$-aspargine sequence was attempted.

2.8. (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA, Mav= 47,002.2 g/mol and (-Asn_{10}) G385S/M386G His_6MBP-TEV-TrnB, M_{av}= 47,336.4 g/mol

2.8.1. Background

After analysis of the results from the $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA construct, it was determined that complete removal of the $poly_{10}$-asparagine sequence may yield better results. To achieve this product, polymerase chain reaction primers were designed remove the sequence from the maltose binding protein through ligation-free polymerase chain reaction.

Ligation-free polymerase chain reaction primers were designed to remove the decaasparagine sequence from the G400S/M401G His_6MBP-TEV-TrnA and G400S/M401G His_6MBP-TEV-TrnB constructs. The products were later used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to verify the base pair length of the amplified DNA, which was next used to transform competent BL21 (DE3) pLysS (Invitrogen) for protein expression. After protein expression, the resulting crude product was purified via nickel affinity chromatography and identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Observed below are the amino acid sequences for both (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA and (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnB. In red is the hexahistidine affinity tag; black, the maltose binding protein; blue, the tobacco etch virus protease recognition site; purple, G385S and M386G mutations; and orange, the TrnA

(B) sequence. It should be observed that the poly$_{10}$-asparagine sequence has been removed.

A:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSG
ENLYFQSGEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAFTLG*

B:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG
KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSG
ENLYFQSGEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASY
FL*

Figure 2.23. Illustrative representation of (-Asn_{10}) G385S/M386G His_6 MBP-TEV-Trn A(B).

2.8.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **G395S/M396G His_6MBP-TEV-TrnA** and **G395S/M396G His_6MBP-TEV-TrnB** to **(-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA** and **(-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnB** using G395S/M396G His_6TEV-TrnA _pMALc2x and G395S/M396G His_6TEV-TrnB _pMALc2x, respectively, with the following primers:

(-Asn_{10}) G385S/M386G His_6TEV-Trn (A, B) _pMALc2x Forward Primer: TTC GGA ATT CGG TGA GAA TCT TTA TTT TCA GTC C

(-Asn_{10}) G385S/M386G His_6TEV-Trn (A, B) _pMALc2x Reverse Primer: CAC CGA ATT CCG AAT TAG TCT GCG CG

The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol**(page 65) with the following modifications:

- Annealing temperature: 56.0°C.
- Extension time: 3 m 30 s

Resulting products were used to transform competent DH5α (Promega), screening with ampicillin (50 µg/mL) resistant Luria broth agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 µg/mL) and incubated

overnight at 37°C, with constant spinning at 225 rpm, for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests with HindIII (Promega) were carried out following the *restriction digest protocol* (page 40). As can been observed in Figure 2.24, expecting a product approximately 6,500 base pairs, the image below demonstrates lanes 2-3 [(-Asn_{10}) G385S/M386G His_6TEV-TrnA _pMALc2x] and lane 5 [(-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x] are of roughly the correct size.

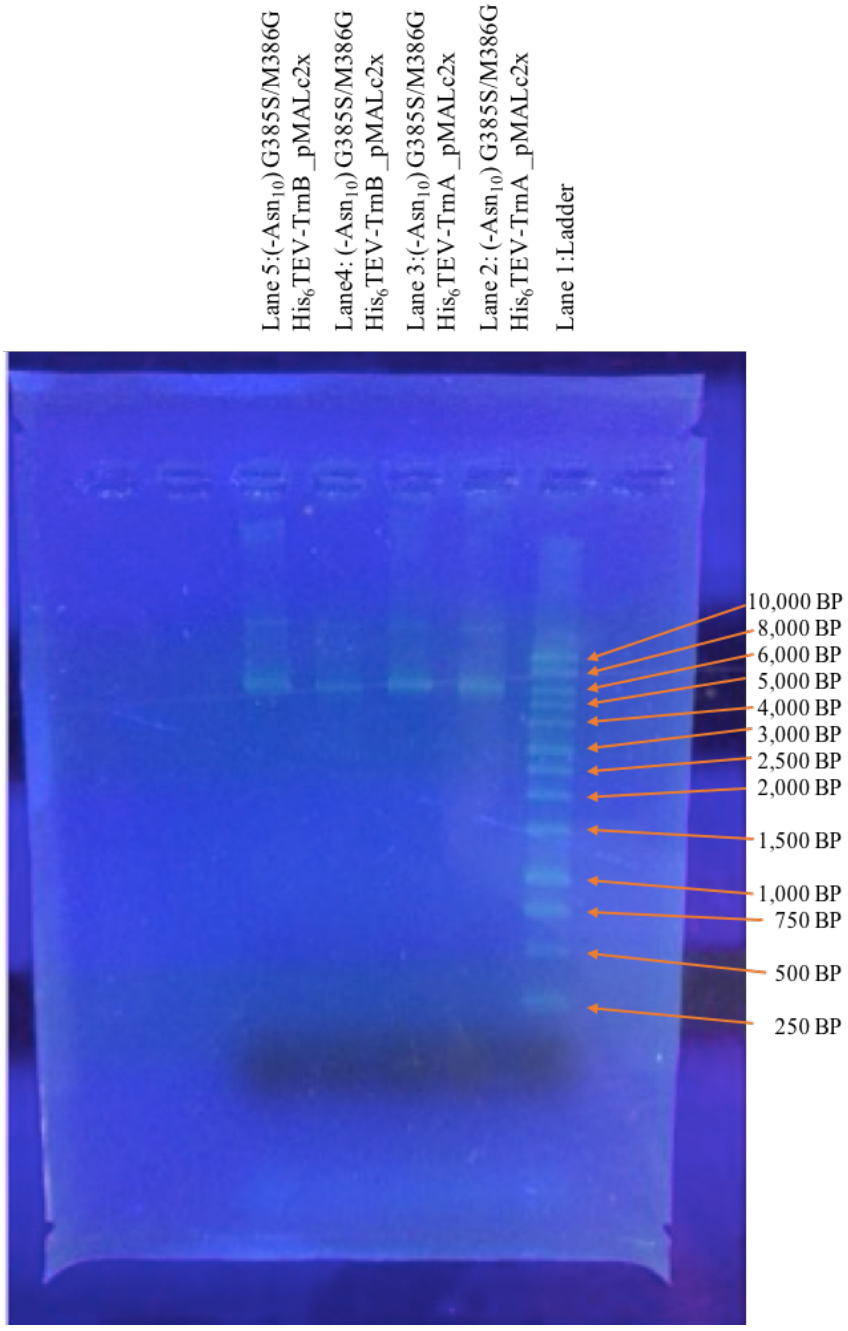


Figure 2.24. Agarose gel depicting (-Asn_{10}) G385S/M386G His_6TEV-TrnA _pMALc2x (lanes 2 and 3) and (-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x (lanes 4 and 5). Lane 1, DNA ladder.

(-Asn_{10}) G385S/M386G His_6TEV-TrnA _pMALc2x and (-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x from lanes 3 and 5, respectively, (Figure 2.24) were used to transform competent BL21 (DE3) pLysS (Invitrogen) and screened on ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) resistant Luria broth agar plates.

2.8.3. Protein Expressions and Purification

Two 50 mL cultures of (-Asn_{10}) G385S/M386G His_6TEV-TrnA _pMALc2x and (-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x in BL21 (DE3) pLysS were screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL each were added separately to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellets were re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by following the *Nickel affinity chromatography protocol* (page 67) and the resulting product was identified via glycine gel sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Figure 2.25 depicts purified (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA (lane 2; M_{av}: ~ 47,000 Daltons) and (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnB (lane 3; M_{av}: ~47,300 Daltons).

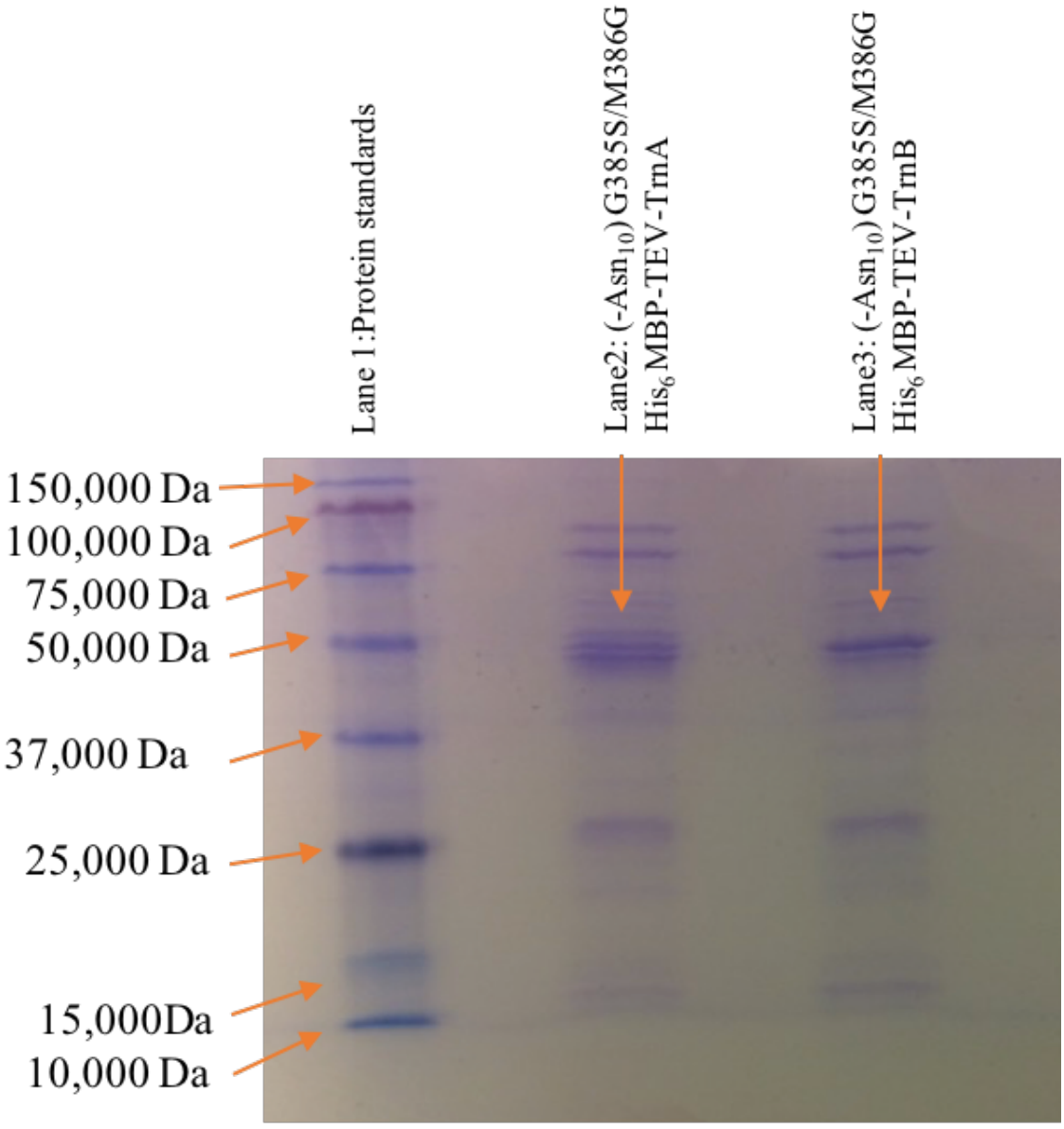


Figure 2.25. SDS PAGE depicting (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA (lane 2) and (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnB (lane 3). Lane 1, protein standards.

2.8.4. Results

After purification of (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA and (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnB and identification via sodium dodecyl sulfate polyacrylamide gel electrophoresis, it was app1arent the removal of the $poly_{10}$-aspargine sequence did not successfully inhibit the expression of a heterogeneous product in (-Asn_{10}) G385S/M386G His_6TEV-TrnA _pMALc2x.

Following multiple failed efforts to isolate a single fusion protein-TrnA product, it was determined that transitioning to a multiple fusion protein technique would better facilitate this process. **All further constructs utilizing the maltose binding protein fusion make use of the (-Asn_{10}) G385S/M386G mutant. It was also decided that focus would be put on TrnB only until a method for purification could be developed.**

2.9. His_6MBP-TEV-TrnB-intein-CBD, M_{av}= 75,279.0 g/mol

2.9.1. Background

To completely eliminate the possibility of C-terminal proteolysis, polymerase chain reaction primers were designed to amplify the (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnB gene from the (-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x vector for ligation into the pTWIN I vector. This vector includes the gene for a chitin binding domain protein that is bound to the C-terminus of TrnB, separated by an intein cleavage site. Intein cleavage sites allow for chemical cleavage by a thiol source following the mechanism described in Figure 2.26 (14):

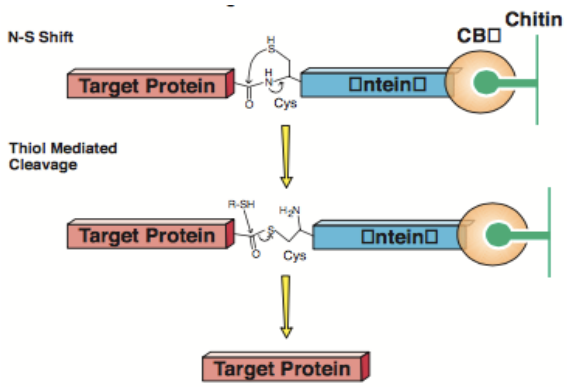


Figure 2.26. Scheme depicting the N-S shift and thiol mediated cleavage at the intein cleavage site of CBD (14).

As can be observed from the mechanism, the sulfur and nitrogen found in the N-terminal cysteine of the intein must first undergo a nitrogen-sulfur shift. This newly formed

thioether bond is susceptible to nucleophilic attack by an outside thiol source, most commonly 1,4-dithiothreitol.

The added advantage to utilizing this expression system is the presence of both an N-terminal and C-terminal fusion protein, allowing for multiple methods of purification and cleavage, as well as eliminating the possibility for C-terminal cleavage of the propeptide during protein expression and purification.

Polymerase chain reaction primers were designed to amply the His_6MBP-TEV-TrnB gene from G400S/M401G His_6MBP-TEV-TrnB_pMALc2x. The product was later ligated into pTWIN I and subsequently used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to verify the base pair length of the amplified DNA, which was next used to transform competent BL21(DE3)pLysS, BL21* (Invitrogen), HMS174 (Invitrogen), and Rosetta (DE3) II (Invitrogen) for protein expression. After protein expression, the resulting crude product was purified via nickel affinity chromatography and identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below is the amino acid sequence for His_6 MBP-TEV-TrnB-intein-CBD. In red is the hexahistidine affinity tag; black, the maltose binding protein; blue, the tobacco etch virus protease recognition site; purple, G385S and M386G mutations; orange, the TrnB; green, the *Mxe* GyrA intein; and brown, the chitin binding domain.

B:

MSHHHHHHIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNG

KLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT
WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTD
YSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS
AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIA
ATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSEFG
ENLYFQSGEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASY
FLCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFHS
GEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQR
SAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGRF
YYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLTTNPGVS
AWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ*

Figure 2.27. Illustrative representation of His_6MBP-TEV-TrnB-intein-CBD.

2.9.2. Subcloning

Polymerase chain reaction was used to amplify the **His_6 (-Asn_6) MBP- (G/S) TEV- (M/G) TrnB** gene from the **His_6- (-Asn_{10}) (G/S) TEV- (M/G) TrnB_pMALc2x** vector with the following primers:

(-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x Forward Primer: CCA TAG CAT ATG AGC CAT CAC CAT CAC C

(-Asn_{10}) G385S/M386G His_6TEV-TrnB _pMALc2x Reverse Primer: GGT GGT GCT CTT CCG CAC AGG AAA TAA GAT GCC GCC GCA AAT TC

The polymerase chain reaction followed the *polymerase chain reaction protocol* (page 53) with the following modification:

- Extension time: 26 s

The resulting polymerase chain reaction product was purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System (Promega).

A restriction digest was used to prepare pTWIN I and (-Asn_{10}) G385S/M386G His_6TEV-TrnB for ligation following the *restriction digest protocol* (page 40) with the following modifications:

- Restriction enzymes: NdeI (Promega) and SapI (Promega); buffer 1X CutSmart® Buffer (Promega)
- pTWIN I was incubated with TSAP Thermosensitive Alkaline Phosphatase (Promega)

After the completion of the restriction digest, the sample was ligated following the *ligation protocol* (page 40).

The resulting product was subsequently used to transform competent DH5α (Promega), screening with ampicillin (50 μg/mL) resistant Luria broth agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 μg/mL) and incubated overnight at 37°C with constant spinning at 225 rpm for DNA isolation with the PureYield™ Plasmid Miniprep System (Promega).

A restriction digest with SapI (Promega) was carried out following the *restriction digest protocol* (page 40). As can been observed in Figure 2.28, the intended product of approximately 7900 base pairs for His_6MBP-TEV-TrnB_pTWIN I is suggested in lane 5.

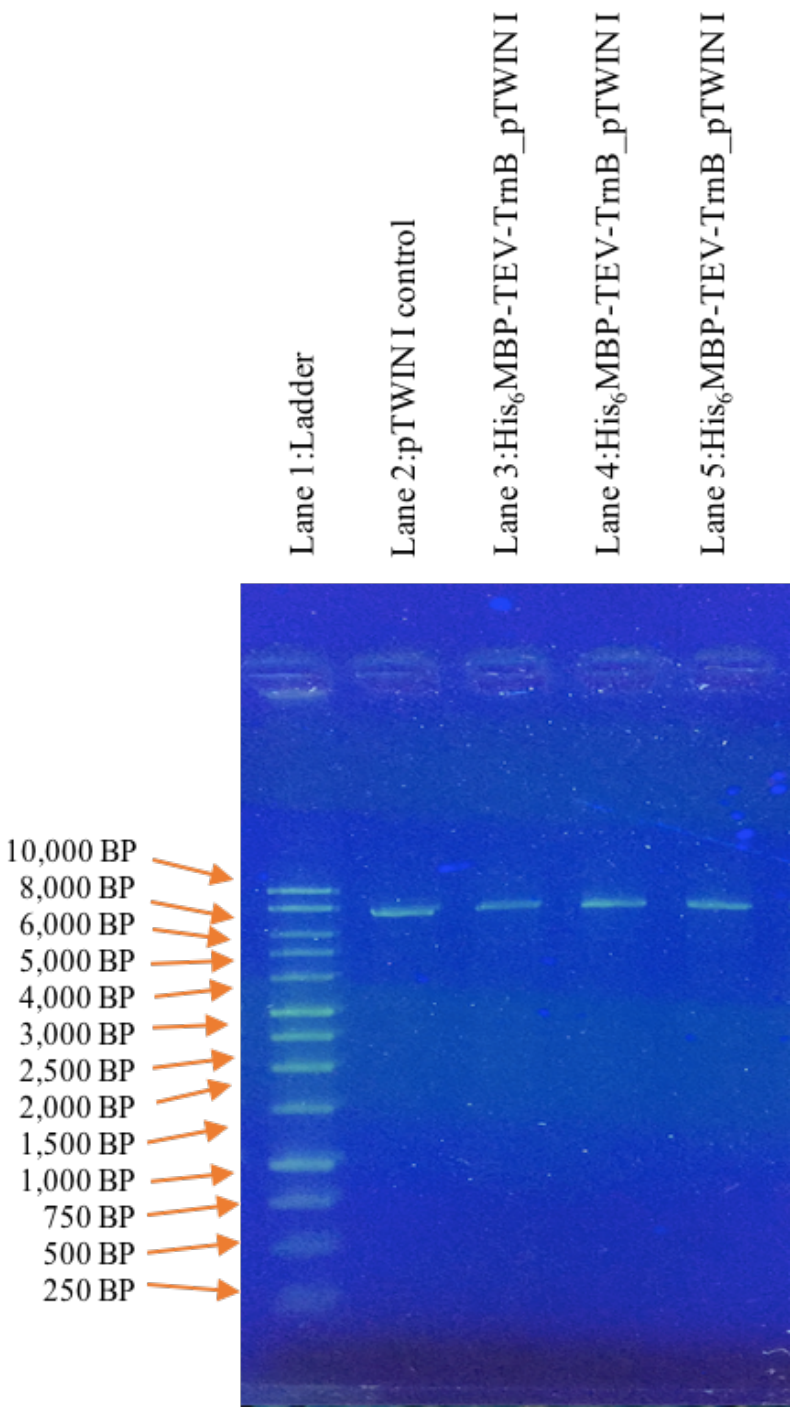


Figure 2.28. Agarose gel depicting His_6MBP-TEV-TrnB_pTWIN I (lanes 3-5). Lane 1, DNA ladder; lane 2, pTWIN I control.

His_6MBP-TEV-TrnB_pTWIN I from lane 5 (Figure 2.28) was used to transform protein expression stains BL21 (DE3) pLysS (chloramphenicol) (Invitrogen), BL21*(no resistance) (Invitrogen), HMS174 (no resistance) (Invitrogen), and Rosetta (DE3) II (chloramphenicol) (Invitrogen) and screened on respective antibiotic resistant Luria broth agar plates.

2.9.3. Protein Expression and Purification

One 50 mL culture of His_6MBP-TEV TrnB_pTWIN I in Rosetta (DE3) II was screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL were added to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellet was re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by following the *Nickel affinity chromatography protocol* (page 67)

Following purification, His_6 MBP-TEV-TrnB-intein-CBD was identified via sodium dodecyl sulfate polyacrylamide gel electrophoresis to analyze the purity and success of the protein expression.

Figure 2.29 depicts His_6MBP-TEV-TrnB-intein-CBD (lane 2; M_{av}: ~75,280 Daltons).

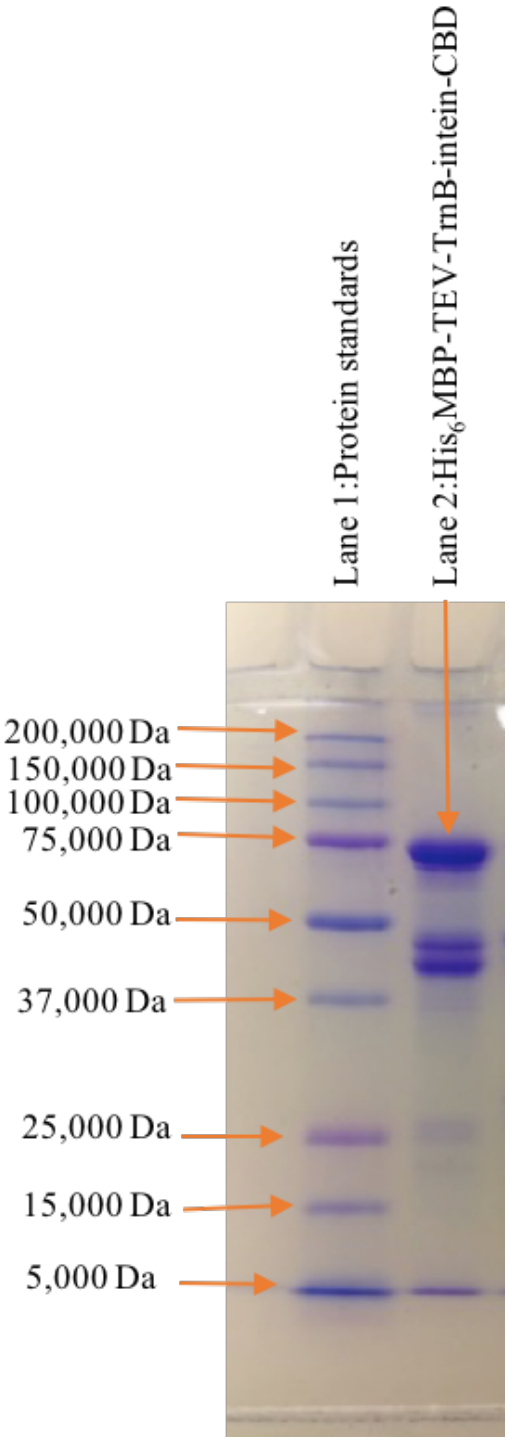


Figure 2.29. SDS PAGE depicting His_6MBP-TEV-TrnB-intein-CBD (lane 2). Lane 1, protein standards.

2.9.4. Cleavage

His_6MBP-TEV-TrnB-intein-CBD was first subjected to cleavage following *Tobacco Etch Virus cleavage conditions* (page 68), as can be observed in lane five on the sodium dodecyl sulfate polyacrylamide gel electrophoresis below (this lane contains product that has been cleaved with dithiothreitol as well). The fusion proteins were also cleaved with dithiothreitol, as can be observed in lane three.

Note: Lower activity is observed in tobacco etch virus protease cleavage when incubated with elevated concentrations 12of dithiothreitol. As a result, all tobacco etch virus protease cleavages were performed prior to any cleavages using dithiothreitol.

Lane 3 in Figure 2.30 is the DTT cleavage product of His_6MBP-TEV-TrnB-intein-CBD. A successful cleavage would yield His_6MBP-TEV-TrnB (M_{av}= 47,354.1 Daltons) and CBD (M_{av}= 27,858.4 Daltons). Lane 5 in Figure 2.30 is the TEV and DTT cleavage product of His_6MBP-TEV-TrnB-intein-CBD. A successful cleavage would yield His_6MBP-TEV (M_{av}= 42,348.8 Daltons), CBD (M_{av}= 27,858.4 Daltons) and TrnB (M_{av}= 5005.6 Daltons).

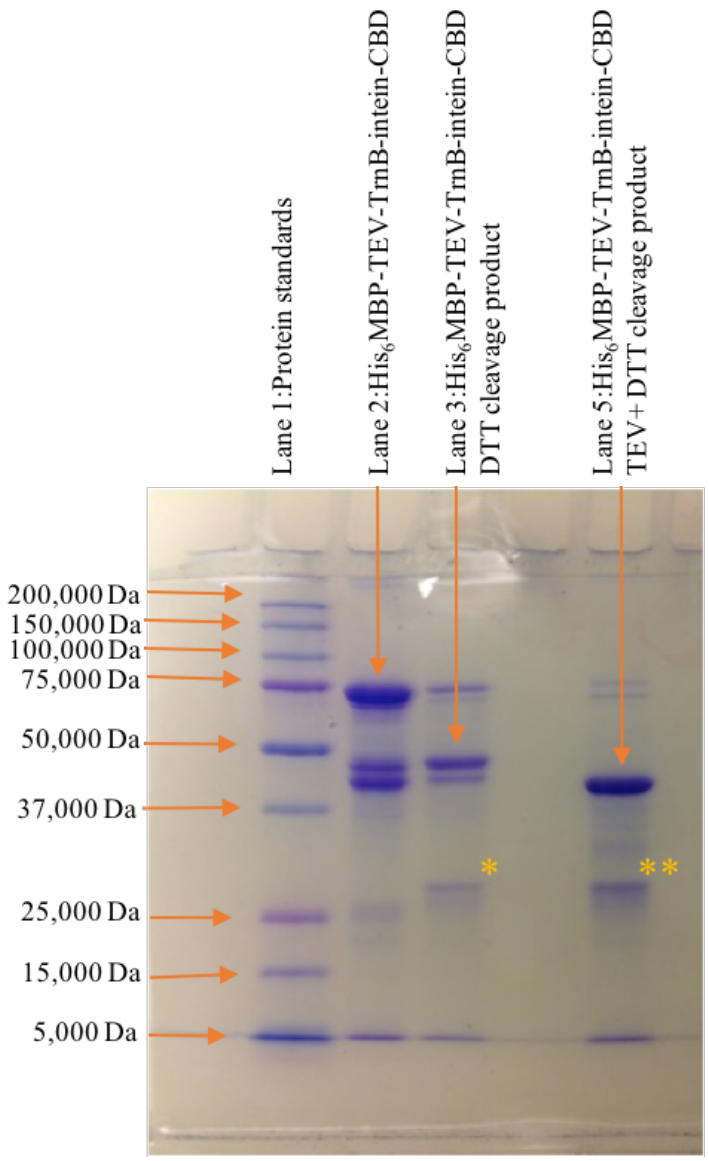


Figure 2.30. SDS PAGE depicting full length His_6MBP-TEV-TrnB-intein-CBD (lane 2); DTT cleavage product of His_6MBP-TEV-TrnB-intein-CBD, resulting in His_6MBP-TEV-TrnB (lane 3); TEV cleavage product of His_6MBP-TEV-TrnB, resulting in His_6MBP-TEV (lane 5). Lane 1, protein standards.

* Protein band corresponds to CBD after DTT cleavage of His_6MBP-TEV-TrnB-intein-CBD

** Protein band corresponds to CBD and TEV protease after cleavage of His_6MBP-TEV-TrnB-intein-CBD with TEV, followed by DTT

2.9.5. Isopropanol (70%) Extraction

We were unable to isolate TrnB from these cleavage reactions using standard HPLC conditions. With respect to conditions developed by Rea et al. 2010 for the isolation of the natural product (thuricin CD) from the native organism, attempts to isolate TrnB by extraction in 70% isopropanol were employed following the method described below:

- His_6MBP-TEV-TrnB was first prepared. His_6MBP-TEV-TrnB-intein-CBD was expressed and purified following the *Nickel affinity chromatography protocol (page 67)*; however, before desalting the elutant from the NTA-Ni^{+2} resin, dithiothreitol was added to the sample until a concentration of 50 mM and was incubated overnight at 4°C. The sample was subsequently desalted to remove imidazole, NaCl, and 1,4-dithiothreitol, and re-purified following the *Nickel affinity chromatography protocol* to purify the His_6MBP-TEV-TrnB from intein-CBD. Samples were subsequently concentrated to approximately 5 mg/mL. Next, 10 mg of His_6MBP-TEV-TrnB were subjected to cleavage following the *Tobacco Etch Virus cleavage conditions* (page 68), allowing for 8 hrs of incubation. Following incubation, samples were flash frozen at -78°C and lyophilized overnight. After lyophilization, 2 mL of 70% isopropanol were added to the samples, which were subsequently incubated at 4°C for 16 hrs with gentle stirring. Following incubation, the samples were centrifuged at 3,500 rpm for 20 minutes to pellet any undissolved material, and the isopropanol was removed from the resulting supernatant through rotary evaporation. Afterwards, the samples were analyzed by LCMS [injection volume, 20µL; Atlantis C18 column; flow rate 0.4 mL/ m; solvent A being *dd*H_2O (0.1 % formic acid) and solvent B being

acetonitrile (0.1 % formic acid)] by applying a linear gradient from 0-100 % B over 40 minutes. Observed in Figure 2.31, the mass of the propeptide TrnB [M-H_2O =4984; [M-H_2O + $2H^+]^2$ = 2493; [M-H_2O +$3H^+]^3$ =1662; [M- H_2O +$4H^+]^4$=1247; [M- H_2O +5H+]^5=998] was identified.

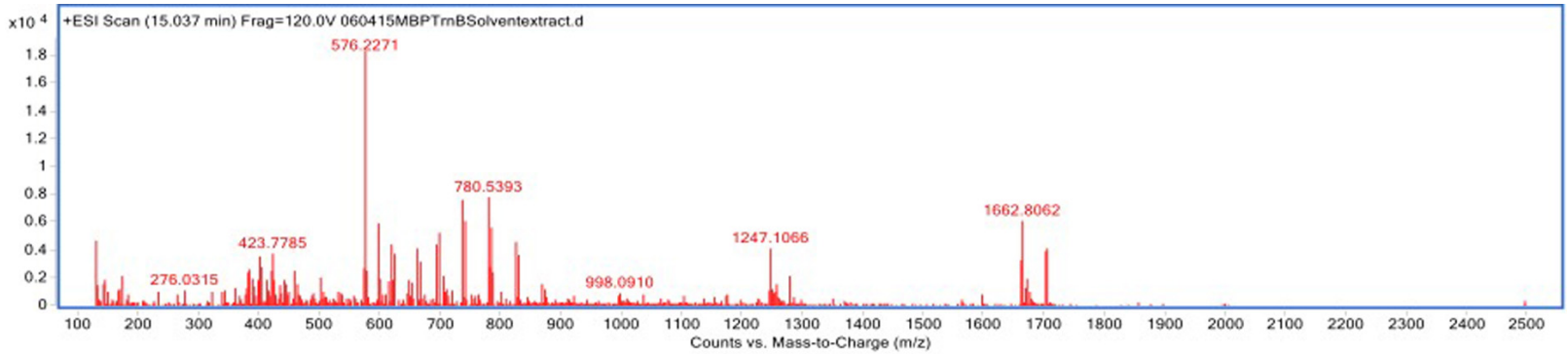


Figure 2.31. LCMS-TOF spectrum of the isopropanol extracted TrnB cleavage reaction; [M-H_2O = 4984; [M-H_2O + $2H^+$]^2 = 2493; [M-H_2O +$3H^+$]^3 = 1662; [M- H_2O +$4H^+$]^4= 1247; [M- H_2O +5H+]^5 = 998].

2.9.6. Results

Expression and purification of His_6 MBP-TEV-TrnB-intein-CBD was successful and in high yield. Average 4 x 1 L Luria broth medium protein preparations resulted in approximately 80 mg total of purified fusion protein. Cleavage reactions were likewise successful. Tobacco etch virus protease cleavages were approximately 90-99% effective after 30 m at 30°C. Similarly, dithiothreitol cleavages were approximately 80-90% effective after 16 hrs of incubation at 4°C. Appropriate fusion protein-propeptide masses were isolated and observed after cleavage of the second fusion partner, as well as isolation and observation of the appropriate fusion protein masses after subsequent cleavage of the propeptide. However, isolation of the propeptides through size exclusion chromatography and reverse-phase chromatography as well as detection through MALDI-MS and LCMS-TOF was unsuccessful under biological conditions. Detection of the propeptide was achieved through LCMS-TOF after extraction with 70% isopropanol, though isolation of a significant quantity of the product was not accomplished.

In an effort to purify propeptide TrnB from the fusion partner, constructs were designed to remove the maltose binding protein fusion and incorporate a octahistidine sequence to the N-terminus of the propeptide for purification through nickel affinity chromatography.

2.10. (W11) His_8TrnB-intein-CBD, M_{av}= 34,347.6 g/mol

2.10.1. Background

Ligation free polymerase chain reaction primers were designed to incorporate the poly-histidine tag and an additional tryptophan residue on the N-terminal side of the TrnB while removing the maltose binding protein and TEV cleavage site using the DNA from the His_6 MBP-TEV-TrnB_pTWIN I as a template. The products were later used to transform DH5α for DNA amplification. Agarose gel electrophoresis was used to confirm the base pair length of the resulting product, which was subsequently used to transform Rosetta 2 (DE3) for protein expression. After protein expression, the resulting crude product was purified via NTA-Ni affinity ion exchange chromatography and identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below is the amino acid sequence for (W11) His_8TrnB-intein-CBD. In red is the newly introduced octahistidine affinity tag; blue, the newly introduced tryptophan residue; orange, the TrnB propeptide sequence; green, the *Mxe* GyrA intein; and brown, the chitin binding domain.

B:

MSHHHHHHHHWSGEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGT
EFAAASYFLCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVL
ADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPG
DYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIAD
ELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGL

TTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ

*

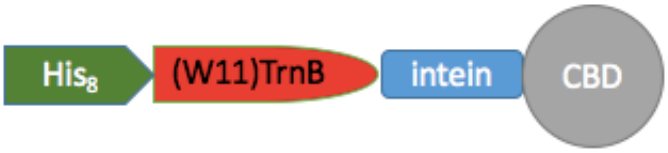


Figure 2.32. Illustrative representation of (W11) His_8TrnB-intein-CBD.

2.10.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **His_6 MBP-TEV-TrnB-intein-CBD** to **(W11) His_8TrnB-intein-CBD** using His_6 MBP-TEV-TrnB _ pTWIN I, with the following primers:

(W11) His_8TrnB _pTWIN I Forward Primer:

TCCGGACCAATGGTGGTGATGGTGATGGTGATGGCTCAT

(W11) His_8TrnB _pTWIN I Reverse Primer:

CACCATTGGTCCGGAGAAGTCCTGAACAAACAGAACG

The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol* (page 65)* with the following modifications:

- Annealing temperature: 55.0°C.
- Extension time: 3 m 10 s

Resulting products were used to transform competent DH5α (Promega), screening with ampicillin (50 µg/mL) resistant Luria broth agar plates. Successful colonies were

inoculated into Luria broth medium, screening with ampicillin (50 μg/mL) and incubated overnight at 37°C with constant shaking at 225 rpm for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests with SapI (Promega) were carried out following the *restriction digest protocol* (page 40). As can been observed in Figure 2.33, the correct His_8 (W)(M/G) TrnB _pTWIN I product of approximately 6,800 base pairs was assumed to be observed in lanes 3, 4, and 6-8. The control, His_6 MBP-TEV-Trn B_pTWIN I, is approximately 1,100 base pairs larger than the expected (W11) His_8TrnB_pTWIN I product, as can be observed in lane 2.

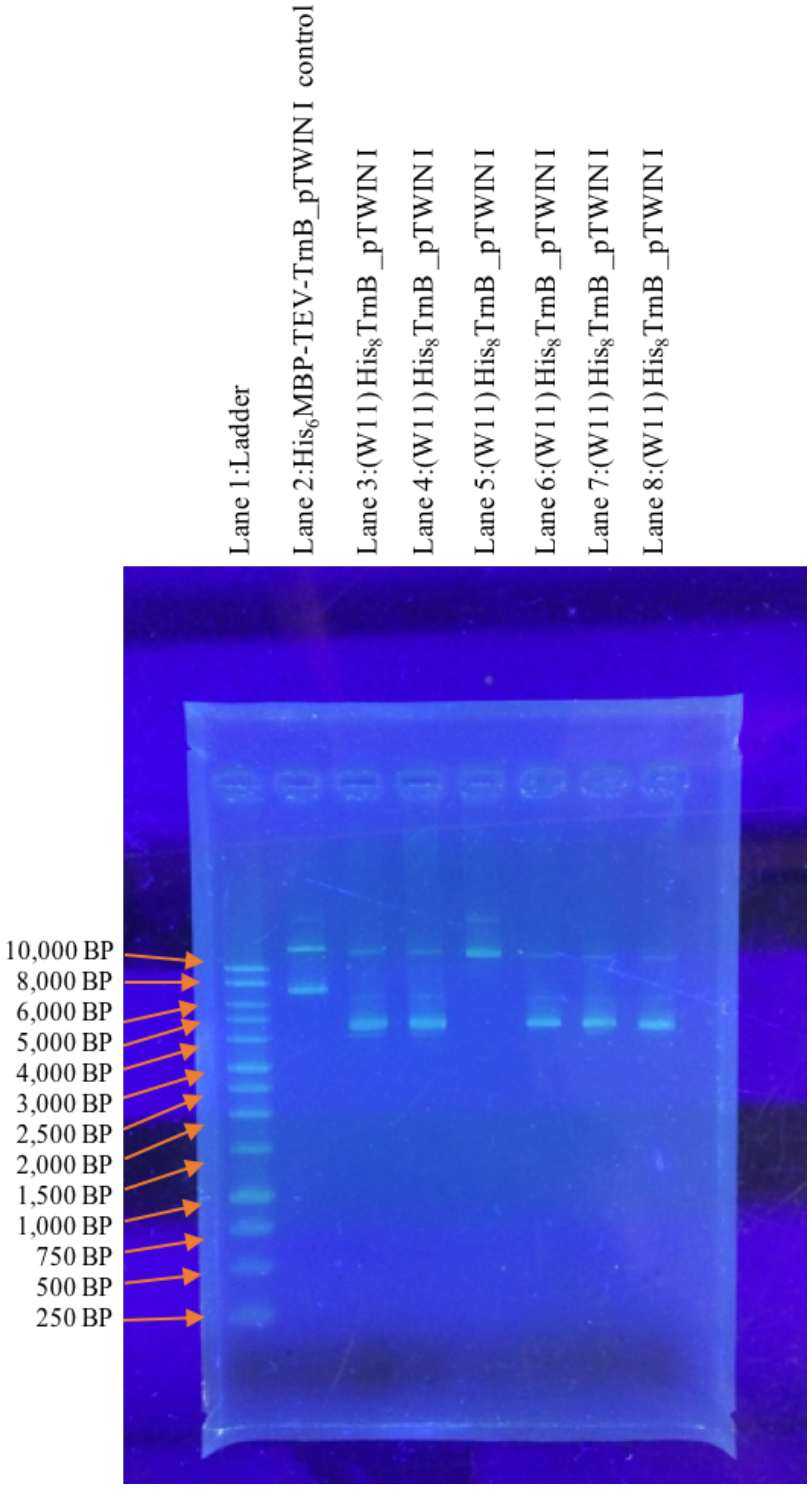


Figure 2.33. Agarose gel depicting (W11) His_8TrnB_pTWIN I (lanes 3-8). Lane 1, DNA ladder; lane 2, His_6 MBP-TEV-Trn B _pTWIN I control.

(W11) His_8TrnB_pTWIN I, observed in lane 4 (Figure 2.33), was subsequently used to transform competent BL21*(no resistance), HMS174 (no resistance), and Rosetta (DE3) II (chloramphenicol) and screened on respective antibiotic resistant Luria broth agar plates.

2.10.3. Protein Expression and Purification

One 50 mL Luria 1broth culture of (W11) His_8TrnB_pTWIN I in Rosetta (DE3) II was screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL were added to 2 x 1L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant shaking at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellet was re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by following the *Nickel affinity chromatography protocol (page 67)*.

Following purification, (W11) His_8TrnB -intein-CBD was identified with sodium dodecyl sulfate polyacrylamide gel electrophoresis to analyze the purity and success of the protein expression.

Figure 2.34 depicts (W11) His_8TrnB-intein-CBD (lane 2; M_{av}: ~34300 Daltons).

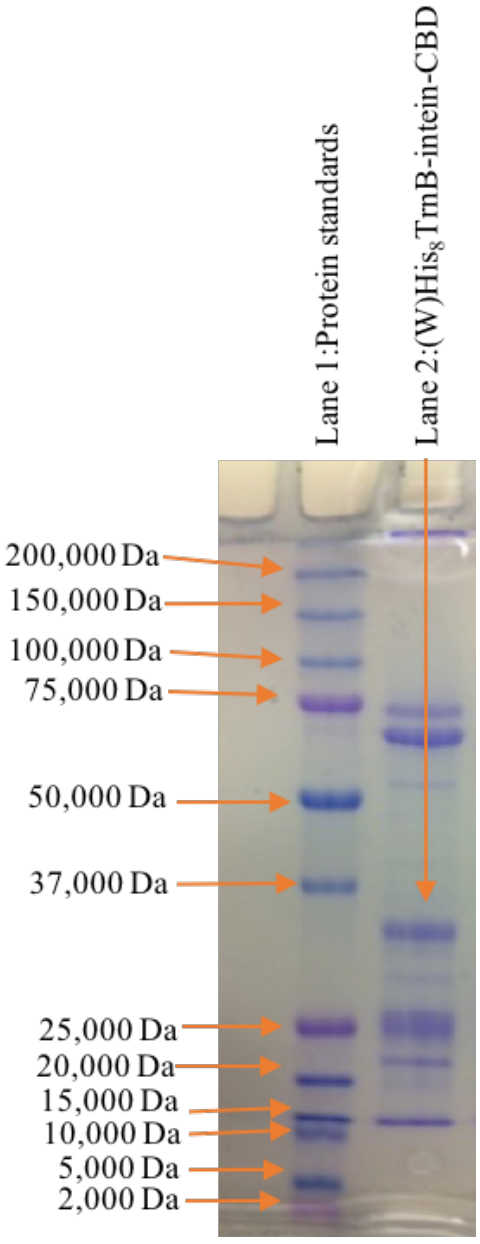


Figure 2.34. SDS PAGE depicting (W11) His_8TrnB-intein-CBD (lane 2). Lane 1, protein standards.

2.10.4. Cleavage

(W11) His_8TrnB -intein-CBD was subjected to 50 mM 1,4-dithiothreitol cleavage. Lane 3 in Figure 2.35 is the DTT cleavage product of (W11) His_8TrnB-intein-CBD. A successful cleavage would yield CBD (M_{av}= 27,858.4 Daltons) and (W11) His_8TrnB (M_{av}= 6,507.2 Daltons).

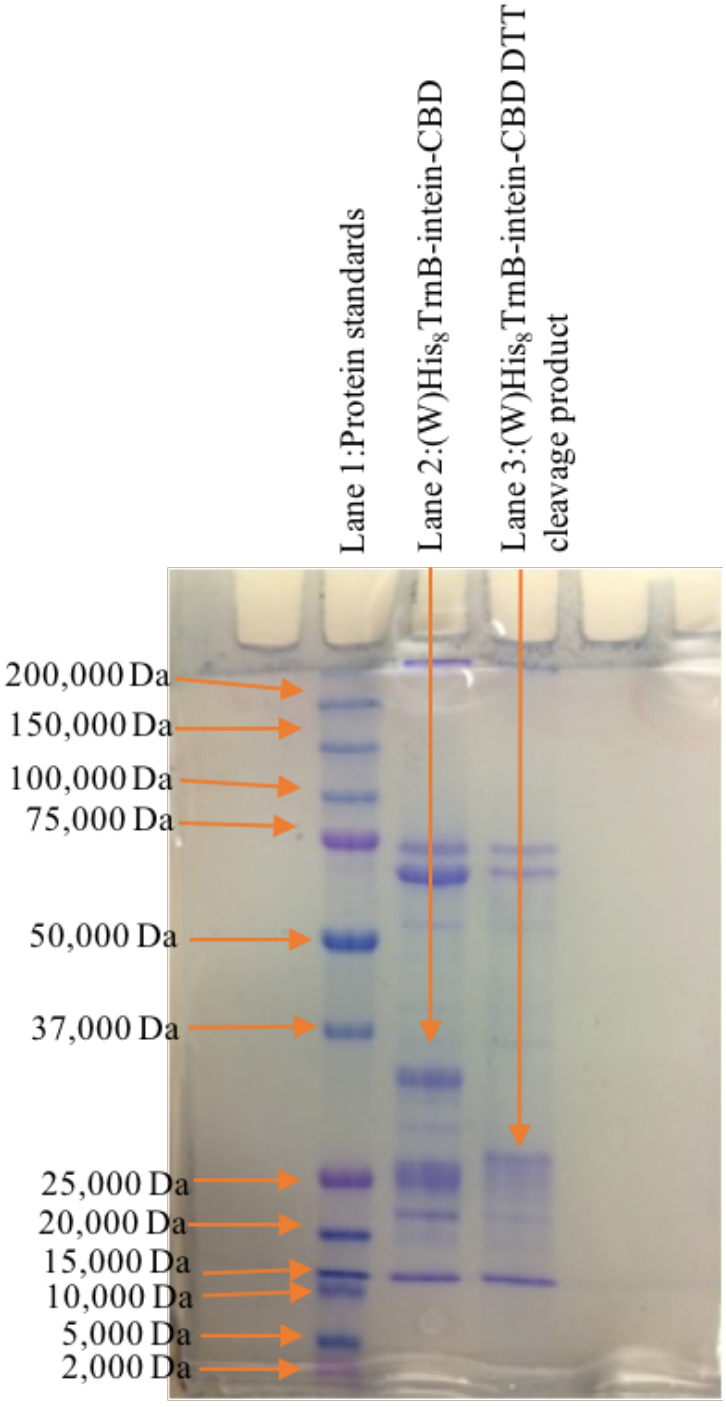


Figure 2.35. SDS PAGE depicting full length (W11) His_8TrnB-intein-CBD (lane 2); and the DTT cleavage product of (W11) His_8TrnB-intein-CBD, resulting in CBD (lane 3). Lane 1, protein standards.

2.10 5. Results

Expressions and purifications of (W11) His_8TrnB -intein-CBD were successful and in high yield. Average 4 x 1 L Luria broth medium protein preparations resulted in approximately 75 mg total of purified fusion protein. Cleavage reactions were likewise successful. Dithiothreitol cleavages were 80-90% effective after 16 hrs of incubation at 4°C. Isolation and observation of the appropriate fusion protein mass after cleavage of the propeptide was observed; however, isolation of the propeptides through nickel affinity chromatography, size exclusion chromatography, and reverse-phase chromatography as well as detection through MALDI-MS and LCMS-TOF was unsuccessful.

2.11. Asp_{10}TEV-TrnB-intein-CBD, M_{av}= 35,010.0 g/mol

2.11.1. Background

After many unsuccessful attempts to isolate the propeptide, it became apparent that the problem may be in the solubility of the product in aqueous solutions buffered at neutral pH. Previous constructs have incorporated poly-histidine affinity tags and multiple enzymatic cleavage sites, all of which contributing little to the overall hydrophobicity, and therefore solubility, of the propeptides. The inclusion of a decaaspartate tag was proposed to aid in solubilizing the propeptide in biological conditions.

Ligation free polymerase chain reaction primers were designed to incorporate the poly-aspartate tag and TEV cleavage site while removing the poly-histidine affinity tag and maltose binding protein using the DNA from the His_6 MBP-TEV-TrnB_pTWIN I vector as a template. The inclusion of the TEV protease recognition site was to allow for removal of the decaaspartate tag.

The product was later used to transform competent DH5α (Promega) for DNA amplification. Agarose gel electrophoresis was used to confirm the base pair length of the resulting product, which was subsequently used to transform Rosetta 2 (DE3) (Invitrogen) for protein expression. After protein expression, the resulting crude product was purified via DEAE ion exchange chromatography and identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Pictured below is the amino acid sequence for Asp_{10}TEV-TrnB-intein-CBD. In blue is the newly introduced decaaspartate affinity tag; green, the newly introduced tobacco etch virus protease recognition site; orange, the TrnB propeptide sequence; green, the *Mxe* GyrA intein; and brown, the chitin binding domain.

B:

MS**DDDDDDDDDDENLYFQS**GEVLNKQNVNIIPESEEVGGWVACVGACGTVCL
ASGGVGTEFAAASYFLCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLD
RHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWK
LIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDP
DAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGL
TGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVP
ALWQLQ*

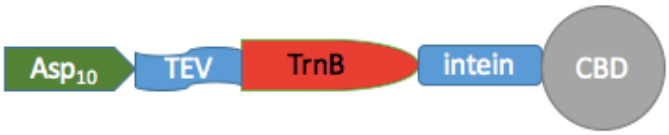


Figure 2.36. Illustrative representation of Asp_{10}TEV-TrnB-intein-CBD.

2.11.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **His_6MBP-TEV-TrnB-intein-CBD** to **Asp10-TEV-TrnB-intein-CBD** using His_6-TEV-TrnB_ pTWIN I with the following primers:

Asp10-TEV-TrnB _pTWIN I Forward Primer:

GAC GACGATGATGACGATGATGAGAATCTTTATTTTCAGTCCGG

Asp10-TEV-TrnB _pTWIN I Reverse Primer:

TCGTCGTCATCATCATCACTCATATGTATATCTCCTTCTTAAAGTTAAAC

The polymerase chain reaction followed the *ligation-free polymerase chain reaction protocol* (page 65)* with the following modifications:

Annealing temperature: 50°C

The resulting product was used to transform competent DH5α (Promega), screening with ampicillin (50 μg/mL) resistant Luria broth agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 μg/mL) and incubated overnight at 37°C with constant spinning at 225 rpm for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests with SapI (Promega) were carried out following the *restriction digest protocol* (page 40). As can been observed in Figure 2.37, the correct Asp10-TEV-Trn B _pTWIN I product of approximately 6,800 base pairs was assumed to be observed in lanes 3-5. The control, His_6MBP-TEV-Trn B _pTWIN I, is approximately 1,100 base pairs larger than the expected Asp_{10}-TEV-TrnB _pTWIN I product, as can be observed in lanes 2 and 6.

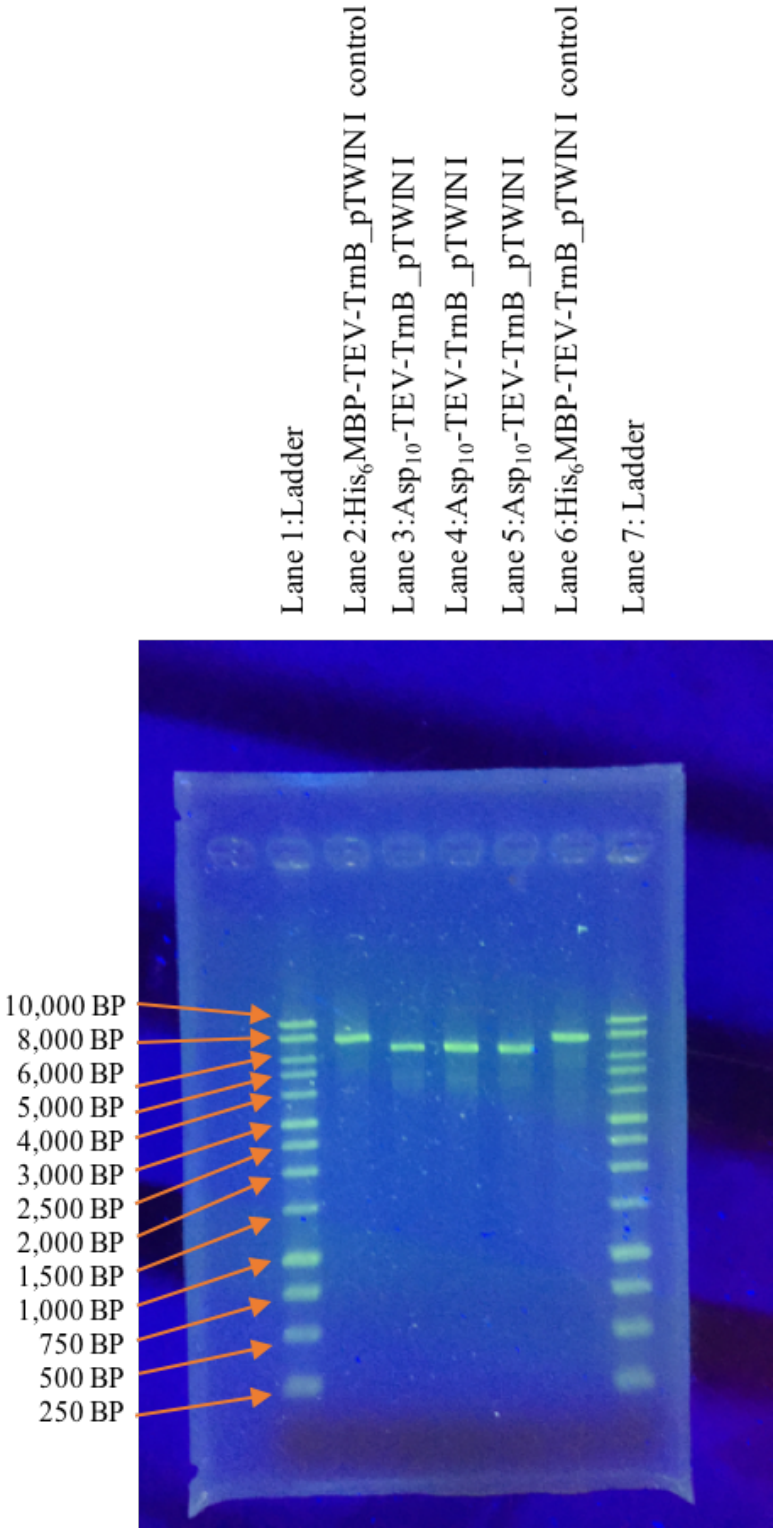


Figure 2.37. Agarose gel depicting Asp_{10}-TEV-TrnB _pTWIN I (lane 3-5). Lanes 1 and 7, DNA ladder; lanes 2 and 6 His_6MBP-TEV-Trn B _pTWIN I.

Asp_{10}–TEV-Trn B_pTWIN I, observed in lane 4 (Figure 2.37), was subsequently used to transform competent Rosetta (DE3) II and screened on chloramphenicol (25 μg/mL) resistant Luria broth agar plates.

2.11.3. Protein Expression and Purification

One 50 mL Luria broth culture of Asp_{10}-TEV-TrnB _pTWIN I in Rosetta (DE3) II was screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL each were added separately to 2 x 1 L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final 1concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellets were re-suspended in 50 mM Tris-base, pH 8 at 2 mL/ 1-gram of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by ion-exchange chromatography using FPLC (Bio-Rad). DEAE Sepharose Fast Flow ion exchange resin (GE Healthcare) was first equilibrated with 50 mM Tris-base, pH 8, and after the loading of the lysate followed a 500 mL linear gradient of 50 mM Tris-base, pH 8 to 50 mM Tris-base, pH 8 1.0 M NaCl to elute bound protein. Ten mL factions were collected through the entirety of the gradient and were later pooled and concentrated using an 10,000 Molecular Weight Amicon Ultra-15

Centrifugal Filter Unit (Millipore) to be identified on glycine gel sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Figure 2.38, lane 1).

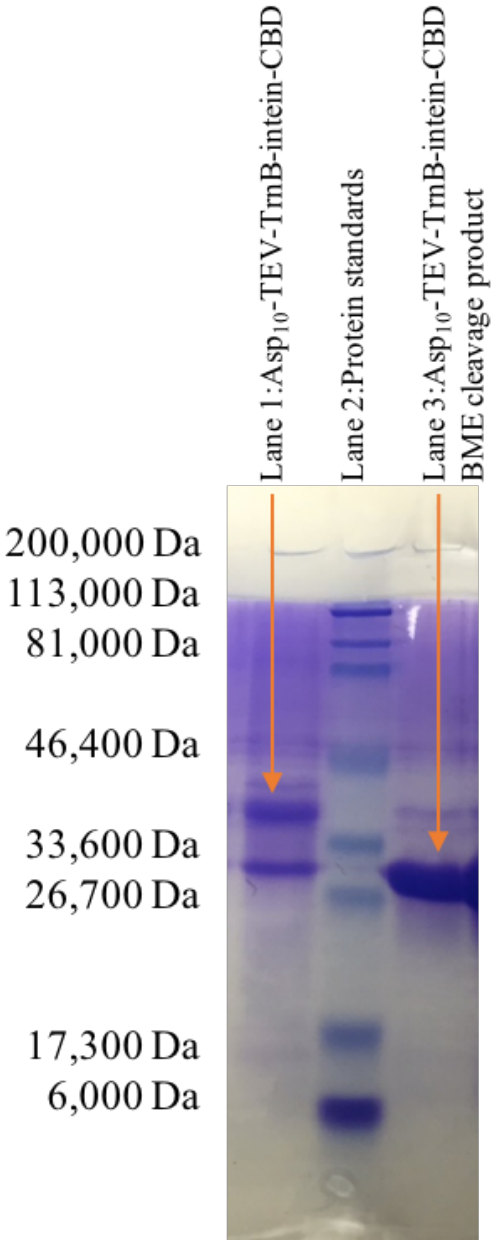


Figure 2.38. SDS PAGE depicting Asp_{10}-TEV-TrnB-intein-CBD (lane 1); and the BME cleavage product of Asp_{10}-TEV-TrnB-intein-CBD, resulting in CBD (lane 3). Lane 2, protein standards.

2.11.4. Cleavage

Asp_{10}-TEV-TrnB-intein-CBD was incubated with 100 mM 2-mercaptoethanol for 16 hrs at 4°C. Lane 3 in Figure 2.38 is the BME cleavage product of Asp_{10}-TEV-TrnB-intein-CBD. A successful cleavage would yield CBD (M_{av}= 27,858.4 Daltons) and Asp_{10}-TEV-TrnB (M_{av}= 7,169.9 Daltons).

2.11.5. Results

Expressions and purifications of Asp_{10}-TEV-TrnB -intein-CBD were successful and in high yield; however, purification through ion exchange chromatography results in mixture of product and contaminant proteins. While this was not an issue during observation of the full-length fusion protein, it became a problem during isolation of the cleaved propeptide. Isolation of the propeptides through ion exchange chromatography, size exclusion chromatography, and reverse-phase chromatography were not successful. In attempts to isolate a purer product, a modified construct was designed to include a hexahistidine affinity tag to the C-terminus of chitin binding domain.

2.12. Asp_{10}TEV-TrnB-intein-CBD-His_6, M_{av}= 35,991.0 g/mol

2.12.1. Background

After initial attempts to purify Asp_{10}TEV-TrnB-intein-CBD through ion exchange chromatography yielded a heterogeneous product, it became apparent a more specific means of purification was necessary. The addition of a hexahistidine tag to the C-terminus of chitin binding domain was believed to provide this specificity while not interfering with the desired N-terminal sequence. The addition of –Gly-Thr- prior to the hexahistidine tag corresponds to the restriction site KpnI, purposely added to allow for a means of determining if the polymerase chain reaction was successful.

Pictured below is the amino acid sequence for Asp_{10}TEV-TrnB-intein-CBD-His_6. In blue, the decaaspartate affinity tag; green, the tobacco etch virus protease recognition site; orange, the TrnB propeptide sequence; green, the *Mxe* GyrA intein; brown, the chitin binding domain; purple, the newly introduced KpnI restriction site; and red, the newly introduced hexahistidine affinity tag.

B:

MS**DDDDDDDDDDENLYFQS**GEVLNKQNVNIIPESEEVGGWVACVGACGTVCL
ASGGVGTEFAAASYFLCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLD
RHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWK
LIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDP
DAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGL
TGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVP
ALWQLQ*GTHHHHHH*

Figure 2.39. Illustrative representation of Asp_{10}TEV-TrnB-intein-CBD-His_6.

2.12.2. Subcloning

Ligation-free polymerase chain reaction was used to modify **Asp_{10}–TEV-TrnB-intein-CBD** to **Asp_{10}–TEV-TrnB-intein-CBD-His_6** using His_6-TEV-TrnB_ pTWIN I with the following primers:

Asp_{10} TEV-Trn B-intein-CBD-His_6 _pTWIN I Reverse Primer:

GGACGGAACACCGTCGAAGTTCCATGGGTGGTGGTGGTGG

The polymerase chain reaction followed the *ligation free polymerase chain reaction protocol (page 46)* with the following modifications:

Annealing temperature: 50°C

The resulting product was used to transform competent DH5α (Promega), screening with ampicillin (50 µg/mL) resistant Luria broth agar plates. Successful colonies were inoculated into Luria broth medium, screening with ampicillin (50 µg/mL) and incubated overnight at 37°C with constant shaking at 225 rpm for DNA isolation with the PureYield™ Plasmid Miniprep System (Promega).

Asp_{10}–TEV-Trn B-intein-CBD-His_6 _pTWIN I was subsequently used to transform competent Rosetta (DE3) II and screened on ampicillin (50 µg/mL) and chloramphenicol (25 µg/mL) resistant Luria broth agar plates.

2.12.3. Protein Expression and Purification

One 50 mL Luria broth culture of Asp_{10} –TEV-Trn B-intein-CBD-His_6_pTWIN I in Rosetta (DE3) II was screened with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant spinning at 225 rpm. After incubation, 10 mL were added to 2 x 1 L of Luria broth medium and screened with ampicillin (50 μg/mL). The cultures were allowed to incubate at 37°C with constant spinning at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellets were re-suspended in 50 mM Tris-base, pH 8 at 2 mL/g of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by following the *Nickel affinity chromatography protocol (page 67)*.

Following purification, Asp_{10}-TEV-Trn B-intein-CBD-His_6 was identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis to analyze the purity and success of the protein expression.

Figure 2.40 depicts Asp_{10}-TEV-Trn B-intein-CBD-His_6 (lane 4) as well as crude lysate (lane 2).

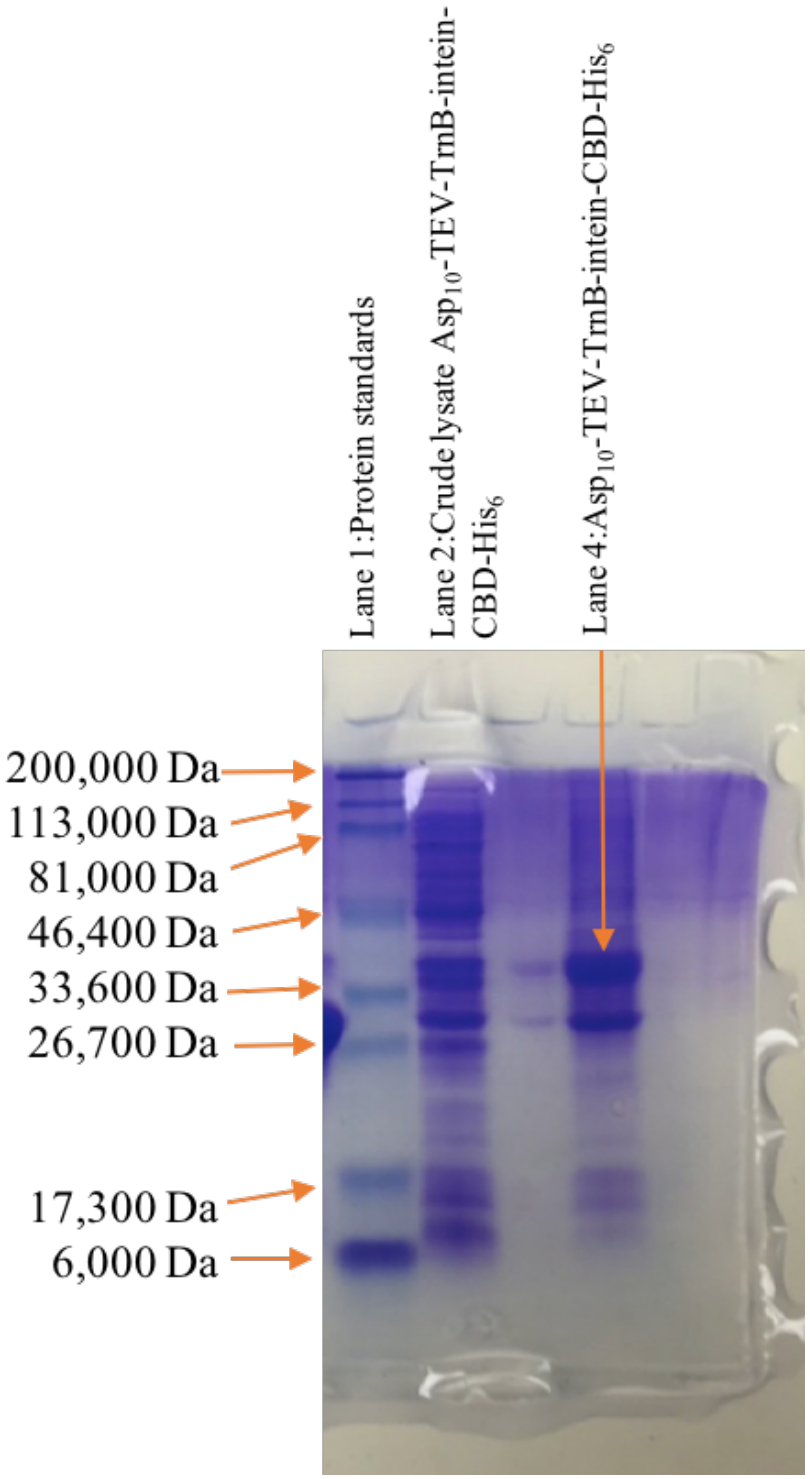


Figure 2.40. SDS PAGE depicting Asp_{10}-TEV-TrnB-intein-CBD-His_6 (lane 4). Lane 1, protein standards; lane 2, crude lysate of Asp_{10}-TEV-TrnB-intein-CBD-His_6.

2.12.4. Cleavage

Asp_{10}-TEV-TrnB-intein-CBD- His_6 was incubated with 100 mM 2-mercaptoethanol for 16 hrs at 4°C. Lane 3 in Figure 2.41 is the BME cleavage product of Asp_{10}-TEV-TrnB-intein-CBD-His_6. A successful cleavage would yield CBD (M_{av}= 28,839.4Daltons) and Asp_{10}-TEV-TrnB (M_{av}= 7,169.9 Daltons).

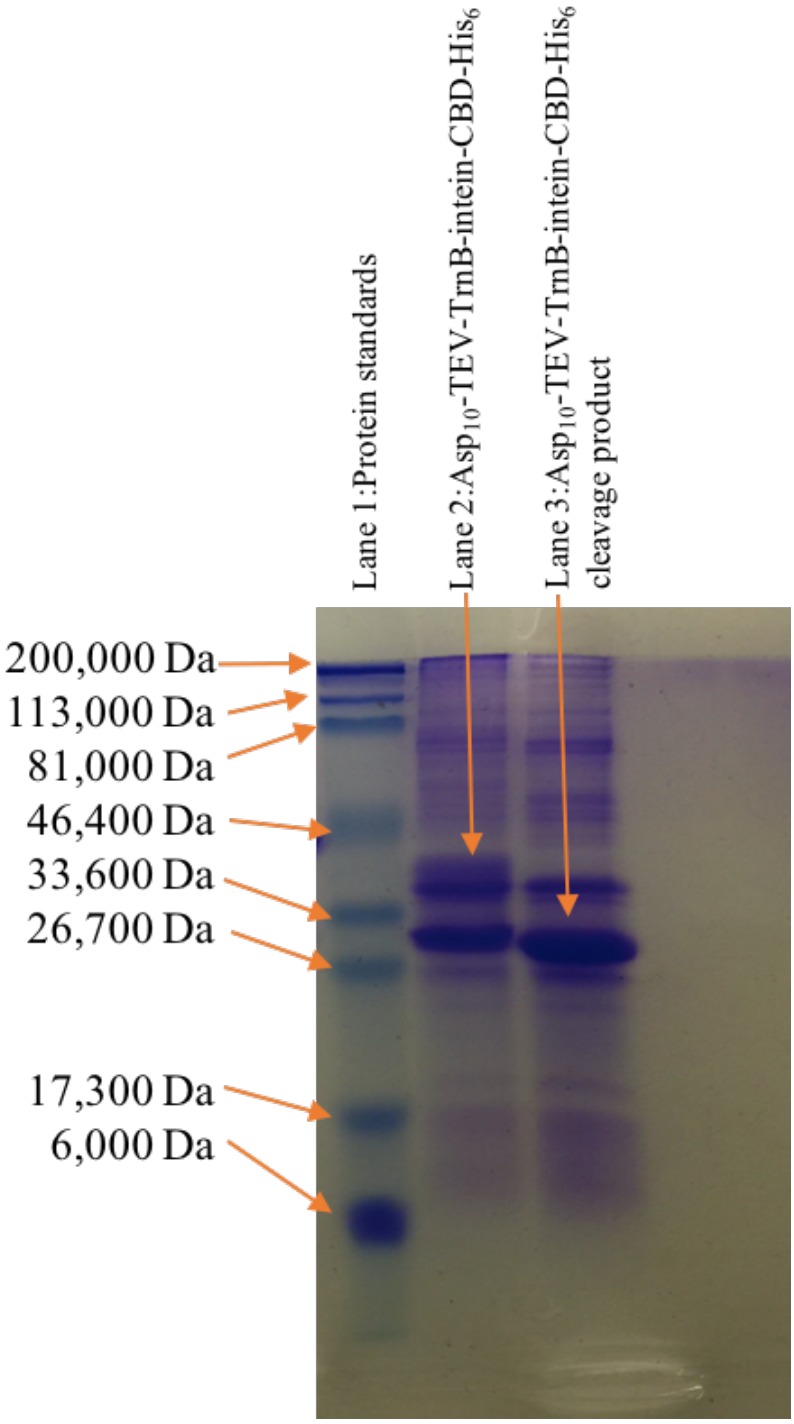


Figure 2.41. SDS PAGE depicting full length Asp_{10}-TEV-TrnB-intein-CBD-His_6 (lane 2) and the BME cleavage product of Asp_{10}-TEV-TrnB-intein-CBD-His_6 (lane3) . Lane 1, protein standards.

2.12.5. Results

Expression and purification of Asp_{10}-TEV-TrnB -intein-CBD-His_6 was successful and in high yield. The addition the hexahistidine was successful in providing an efficient and specific method for purification.

Though expression and purification, as well as cleavage of Asp_{10}-TEV-TrnB -intein-CBD-His_6 was successful, purification of TrnB was not achieved. Isolation of the propeptide through ion exchange chromatography, size exclusion chromatography, and reverse-phase chromatography was not successful.

2.13. Conclusion

In conclusion to the constructs described in chapter 2, a variety of attempts at isolation of propeptides TrnA and TrnB were unsuccessful. Purification of the native propeptides was first attempted. Lacking a specific method for purification, a second construct was designed to include a biotin recognition site for purification through avidin affinity. This construct was subcloned and expressed, but purification through avidin affinity was not attempted. With respect to the lack of a cleavage site separating the biotinylation recognition sequence and the propeptide, constructs were designed to include TrnA and TrnB as MBP-factorXa fusion proteins. This would allow for elevated expression levels as well as the ability to remove the propeptides from the MBP fusion through factorXa cleavage. Purification was limited to ion exchange chromatography, which led to an impure product. This, coupled to unsuccessful factorXa cleavage, led to the development of fusion proteins with removal of the factorXa protease recognition sequence and the insertion of a tobacco etch virus protease recognition sequence. An N-terminal hexahistidine affinity tag was also included to aid in purification. While purification of both His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB constructs was successful, successful cleavage was not observed in His_6MBP-TEV-TrnA. It was also believed a mixed His_6MBP-TEV-TrnA product was present. Mutations at the TEV recognition sequence allowed for elevated levels of cleavage in both constructs. After successful cleavage in both constructs was observed, separate constructs to include a pentaserine sequence and to remove a decaasparagine sequence to/from maltose binding protein were to designed in attempts to purify a single His_6MBP-TEV-TrnA product. These constructs

were unsuccessful in inhibition of a mixed product. At that point, focus was put exclusively on propeptide TrnB.

To allow for multiple methods of purification and cleavage, His_6MBP-TEV-TrnB was subcloned into pTWIN I, which includes a C-terminal chitin binding domain (CBD), with an intein cleavage site separating the propeptide from the CBD. Purification of the construct was achieved through nickel affinity chromatography and both cleavages were observed to be successful. Methods were then designed to extract TrnB in 70% isopropanol from His_6MBP-TEV-TrnB-intein-CBD following cleavage of both MBP and CBD. The extract was analyzed through LCMS-ESI-TOF, where detection of the TrnB was successful. A construct was then designed to remove the His_6MBP-TEV and include an N-terminal octahistidine affinity tag for purification through nickel affinity chromatography, as well as the addition of tryptophan residue to aid in A_{280} detection. After failed attempts to purify the propeptide, the octahistidine affinity tag was removed and an N-terminal decaaspartate affinity tag was inserted to aid in water solubility. The deletion of the hexahistidine affinity tag removed to the ability to purify the fusion through nickel affinity chromatography and as a result, a final construct was designed to include a hexahistidine affinity tag at the C-terminus of CBD. Purification levels increased as a result of the hexahistidine affinity tag, however isolation of the propeptide was unsuccessful.

Chapter 3

Purification of TrnC and Evidence for *In Vitro* Activity Using His_6MBP-TEV-TrnB as Substrate

This chapter describes the subcloning, expression, and purification of the propeptide modifying enzyme TrnC, and early development of methods to modify propeptide TrnB as a His_6MBP-TEV-TrnB fusion, with subsequent detection using mass spectrometry after proteolytic digestion of the modified product.

The subcloning of TrnC followed conventional ligation techniques involving restriction digestions of both the cloning vector and the polymerase-chain-reaction-amplified TrnC gene, followed by ligation of the two products. After propagation of the ligation product and transformation into an expression strain, TrnC was expressed and then purified through nickel affinity chromatography. The correct molecular weight was confirmed through sodium dodecyl sulfate polyacrylamide gel electrophoresis.

After TrnC was purified, a method for detecting the activity of the modifying assays through enzymatic digestion of modified His_6MBP-TEV-TrnB was investigated. The approach was to determine which protease would produce unique fragments both before and after treatment with the propeptide modifying enzymes, such that it would be possible to determine if a reaction had occurred (by a predicted mass gain/loss), and if so, approximately which residues were involved in the reaction. It is not possible to determine exactly which residue has been modified using this approach, but knowing the predicted mass of fragments containing the residues expected to undergo modification

and then observation of a predicted mass gain/loss would suggest the expected residue is being modified. After investigations using trypsin, glu-C, chymotrypsin, and proteinase K, it was determined that chymotrypsin would be the most suitable protease to produce the necessary fragments to determine both if a reaction is occurring and at approximately which residue.

Next, the reactivities of TrnC and the previously purified TrnD were investigated by monitoring the production of 5'-deoxyadenosine and methionine after the enzyme-catalyzed reductive cleavage of S-adenosyl-L-methionine. After determining a very low level of background production of 5'-deoxyadenosine, the data demonstrated that both enzymes were successfully reducing SAM in high yield. This cleavage did not require a peptide or protein substrate, and the source of the hydrogen atom on the 5'-deoxyadenosine product is unknown.

After demonstrating that the modifying enzymes were reactive with SAM, propeptide modifying assay methods were explored. To successfully detect the modified product, it was necessary to purify the His_6MBP-TEV-TrnB from the modifying enzymes before enzymatic digestion with chymotrypsin to avoid ambiguous results during detection. Due to the fact that a hexahistidine affinity tag was present on both the fusion protein and the modifying enzymes, it was not possible to purify using nickel affinity chromatography. Amylose affinity chromatography, which retains the maltose-binding protein, allowed for purification of the MBP fusion protein.

Detection of what is believed to be a crosslinked product was achieved after optimization of activity assays and purification. These crosslinks were detected after incubation with

TrnC and were observed between Cys9 and Ala29, and Cys13 and Thr21. No results suggested TrnD was catalyzing a crosslinking reaction or any other reaction that altered the chromatographic elution time or mass of the peptides derived from digestion with chymotrypsin.

3.1. Purification of His_6TrnC

3.1.1. Background

TrnC and/or TrnD are believed to catalyze the crosslinking reaction(s) observed in TrnA and TrnB via a radical SAM enzyme-like mechanism. As can be observed in the sequence below, the highlighted cysteines (blue) appear in the motif CXXXCXXC (X, red), which is the conserved iron-sulfur cluster binding motif observed in radical SAM enzymes.

TrnC:

MGSSHHHHHHSSGLVPRGSHMSKKRLSMMPFKTDKFSYALDGNTGRVIVADKP
TLYIISHFHKFEKEELLKKTKGKFAELHQDYLTTYNYVSSLINMGMFYLSEKEDS
DSPIDSKELAINSNQSQLILILTEKCNLRCEYCIYNDKYPKEMGYSDEEMDFETAK
KAVDMYYELHMERVKRGHKRFPVITMYGGEPLLKFDLIKKVMEYAKGLMPDTL
FYTTTNGTLLSEKMMDYFINNRIIITFSIDGFKENHDRNRVFVNGMPTFERAFKNI
KRLQEKKKEQNIEQIISFNCCFDQYTDVYKVAKFFEEHYDLFNPFFVLYNQINPFD
TLYFDWCDEQVKTGKWNFDKNNFKNAMQKIEHELYEAETCDDHFQQVAGPLV
MKDFVLSIRNKDGQQQITRNSCIPTSKIAVSPDGTLTLCEKMCKKYPIGTVEKGL

DWKAVDGVTEKLVRHFNSDSCKYCPIRTMCEACFMFLDENGRIKPSFCKSKKM

AVKKNLESYFAKKEKGFDVMKVYNHTSDLDSVKEMVK*

A synthetic TrnC gene, codon optimized for expression in *E. coli*, was received from GenScript in a pUC57 vector. Polymerase chain reaction primers were used to amplify the gene before ligation into pET28b (+). The optional hexahistidine affinity tag (green residues above) present in the multiple cloning site of pET28b (+) was included for ease of purification.

3.1.2. Subcloning

Polymerase chain reaction was used to amplify the His_6TrnC genes from TrnC_pUC57 using the following primers:

TrnC_F

CGGCAGCCATATGAGCAAGAAGCGCCTGTCG

TrnC_R

GTGGTGCTCGAGTCATCATTTCACCATTTCC

The polymerase chain reaction followed the *polymerase chain reaction protocol* (page 53) with the following modification:

- Annealing temperature: 60.0°C
- Extension time: 30 s

The resulting polymerase chain reaction product was purified using the Wizard® SV Gel and Polymerase Chain Reaction Clean-Up System (Promega).

Restriction digests were used to prepare pET28b (+) and TrnC for ligation, following the *restriction digest protocol* (page 40) with the following modifications:

- Restriction enzymes: NdeI and XhoI; buffer D (Promega)
- pET28b (+) was incubated with Thermosensitive Alkaline Phosphatase (Promega)

After the completion of the restriction digests, the samples were ligated following the *ligation protocol* (page 40).

Resulting products were subsequently used to transform competent DH5α (Promega), screening with kanamycin (25 μg/mL) resistant Luria broth agar plates. Successful colonies were inoculated into Luria broth medium, screening with kanamycin (25 μg/mL), and incubated overnight at 37°C with constant shaking at 225 rpm for DNA isolation with the PureYield^{TM} Plasmid Miniprep System (Promega).

Restriction digests with NdeI (Promega) were carried out to determine if the correct product was generated, following the *restriction digest protocol* (page 40). As can been observed in Figure 3.01, the correct product of approximately 1500 base pairs greater than the control, pET28b (+) can be observed.

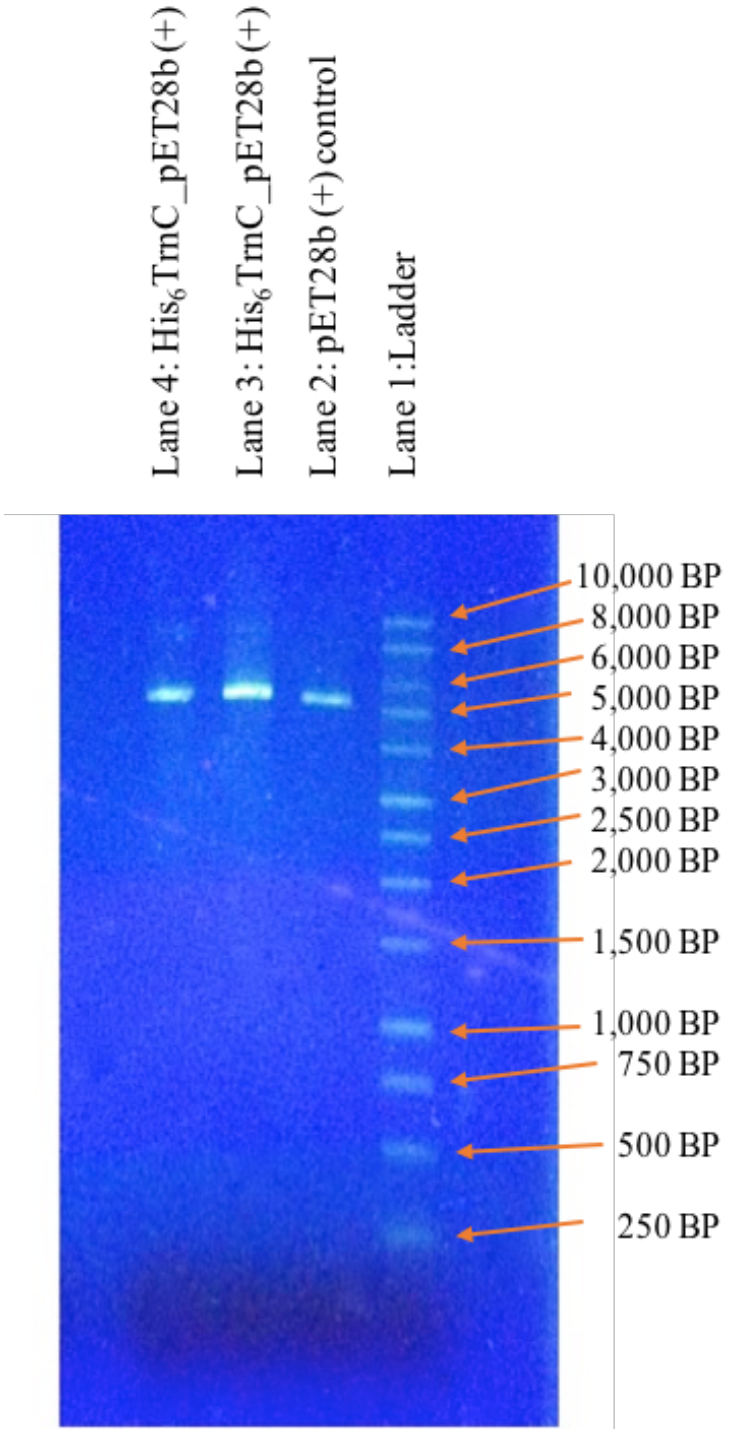


Figure 3.01. Agarose gel depicting His_6TrnC_pET28b (+) (lanes 3 and 4). Lane 1, DNA ladder; lane 2, pET28b (+) control.

3.1.3. Protein Expression and Purification

One 50 mL Luria broth culture of His_6TrnC_pET28b(+) BL21(DE3)pLysS was screened with kanamycin (25 μg/mL) and chloramphenicol (25 μg/mL) and allowed to incubate for 16 hrs at 37°C with constant shaking at 225 rpm. After incubation, 10 mL were added to 4 x 1 L of Luria broth medium and screened with kanamycin (25 μg/mL). The cultures were allowed to incubate at 37°C with constant shaking at 225 rpm until an optical density of 0.5 was observed. At this point, IPTG was added until a final concentration of 0.2 mM and the cultures were incubated for an additional 3 hrs. After 3 hrs of protein expression, the cells were harvested with the Beckman Coulter Avanti J-E Centrifuge and frozen at -80°C.

After defrosting, the pellet was re-suspended in 50 mM Tris-base, pH 8 at 2 mL/g of material and lysed following the *sonication protocol* (page 42).

Purification was achieved by following the *Nickel affinity chromatography protocol* (page 67) and the product was identified through sodium dodecyl sulfate polyacrylamide gel electrophoresis to analyze the purity and success of the protein expression. The product can be seen in Figure 3.02

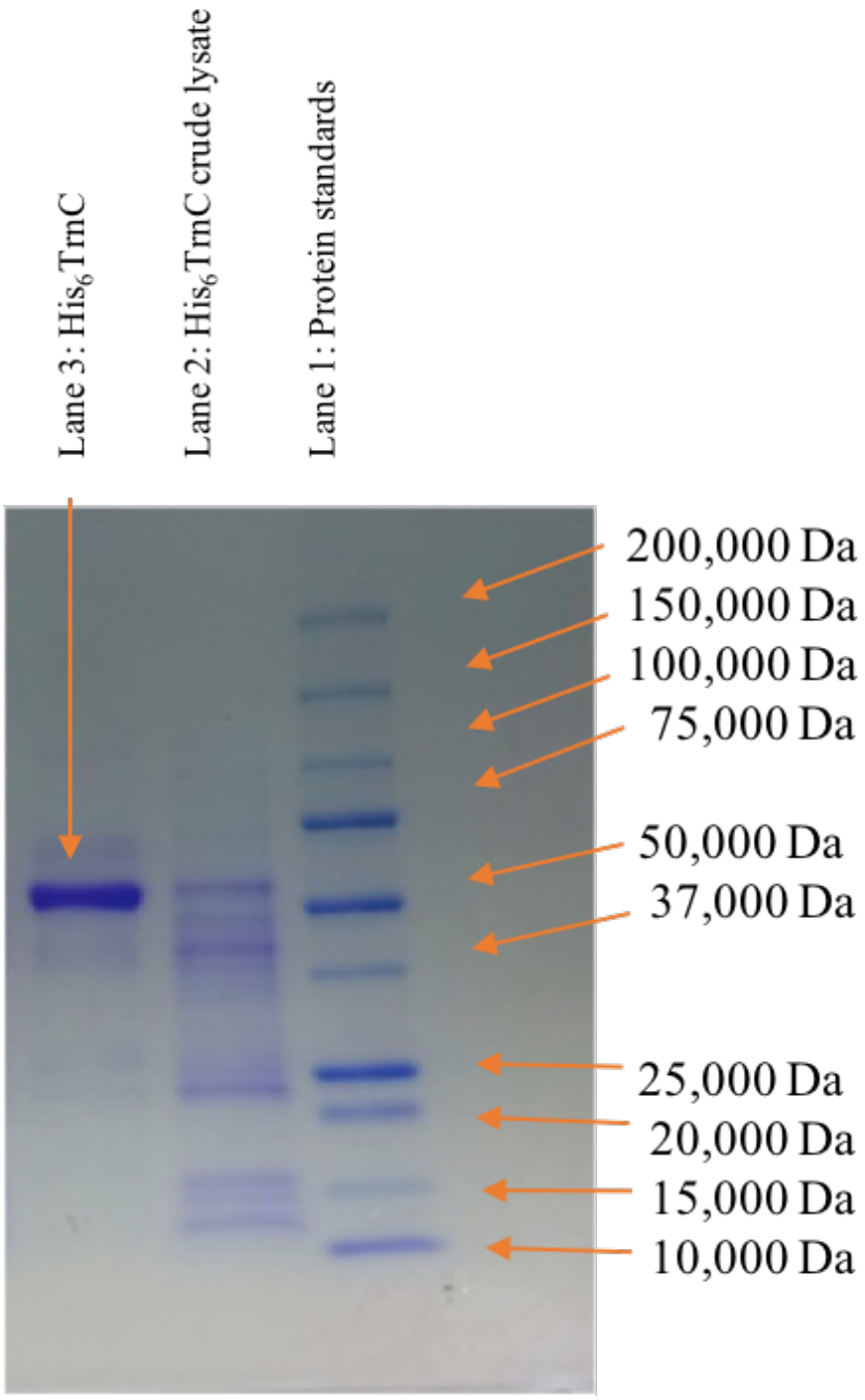


Figure 3.02. SDS PAGE depicting His_6TrnC (lane 3). Lane 1, protein standards; lane 2, crude lysate of His_6TrnC.

3.1.5. Results

Expression and purification of TrnC was successful and in high yield. The hexahistidine affinity tag allowed for specific binding through nickel affinity chromatography, which yielded a highly pure product. This product will be used in propeptide modifying assays, as well as the previously purified TrnD.

3.2. Enzymatic digestion studies of His6MBP-TEV-TrnB

In efforts to confirm the presence of full length His_6MBP-TEV-TrnB, a series of digestion enzyme studies were developed. Listed below are the predicted fragments from digestion with trypsin, Glu-C, chymotrypsin, proteinase K, and pairwise combinations of each. In order to carry out these digestions, the His_6MBP-TEV-TrnB was first prepared. His_6MBP-TEV-TrnB-intein-CBD was expressed and purified following the *Nickel affinity chromatography protocol (page 67)*; however, before desalting the eluent from the Ni-NTA column, dithiothreitol was added to the sample until a concentration of 50 mM and then was incubated overnight at 4°C. The sample was subsequently desalted using a Biogel P2 column to remove imidazole, NaCl, and 1,4-dithiothreitol, and then re-purified following the *Nickel affinity chromatography protocol* (page 67) to purify the His_6MBP-TEV-TrnB from intein-CBD. Samples were subsequently concentrated to approximately 5 mg/mL.

For successful digestion with trypsin, the following method was used:

> First, 20 μL of 8 M urea was added to 20 μL His_6MBP-TEV-TrnB (5 mg/mL). Next, 20 μL digest buffer (50 mM Tris-base, 1 mM $CaCl_2$) was added and then incubated at room temperature for 5 min with slow agitation. After incubation, 4 μL 100 mM TCEP was added and the solution incubated for 1 hr at room temperature. Subsequently, 5 μL of 200 mM iodoacetamide was added for cysteine alkylation and incubated in the dark for 1 hr at room temperature. To avoid any unwanted protease alkylation, 1 μL 100 mM DTT was added to quench the remaining iodoacetamide, and the solution incubated for 10 min at room

temperature. Next, 110 µL digest buffer were added followed by the addition of 2 µg trypsin. The samples were then incubated overnight at 37℃ with slow agitation. Further discussion will refer to the method as the *enzymatic digestion protocol.*

For successful enzymatic digestion with chymotrypsin and proteinase K, the *enzymatic digestion protocol* was carried out with the following modifications:

Chymotrypsin: A 1:60 (w/w) ratio of chymotrypsin: propeptide was incubated at 30 °C for up to 24 hrs with gentle agitation.

Proteinase K: Sufficient enzyme was added for a final concentration of 50–100 µg/mL and incubated at 37–56°C for at least 1 hr.

```
        10         20         30         40         50         60
SHHHHHHHIE EGKLVIWING DKGYNGLAEV GKKFEKDTGI KVTVEHPDKL EEKFPQVAAT

        70         80         90        100        110        120
GDGPDIIFWA HDRFGGYAQS GLLAEITPDK AFQDKLYPFT WDAVRYNGKL IAYPIAVEAL

       130        140        150        160        170        180
SLIYNKDLLP NPPKTWEEIP ALDKELKAKG KSALMFNLQE PYFTWPLIAA DGGYAFKYEN

       190        200        210        220        230        240
GKYDIKDVGV DNAGAKAGLT FLVDLIKNKH MNADTDYSIA EAAFNKGETA MTINGPWAWS

       250        260        270        280        290        300
NIDTSKVNYG VTVLPTFKGQ PSKPFVGVLS AGINAASPNK ELAKEFLENY LLTDEGLEAV

       310        320        330        340        350        360
NKDKPLGAVA LKSYEEELAK DPRIAATMEN AQKGEIMPNI PQMSAFWYAV RTAVINAASG

       370        380        390        400        410        420
RQTVDEALKD AQTNSEFGEN LYFQSGEVLN KQNVNIIPES EEVGGWVACV GACGTVCLAS

       430
GGVGTEFAAA SYFL
```

Figure 3.03. Amino acid sequence of His_6MBP-TEV-TrnB.

3.2.1. Trypsin Digest

Table 3.01. His_6MBP-TEV-TrnB trypsin digest products.

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
SHHHHHHHIEEGK	13-Jan	1620.7406	1620.7431	17383
LVIWINGDK	14-22	1056.5967	1056.5941	995603
GYNGLAEVGK	23-32	1006.5083	1006.5043	2121363
FEK	34-36	422.2165	422.2174	171838
DTGIK	37-41	532.2857	532.2873	674897
VTVEHPDK	42-49	923.4712	923.4732	49003
LEEK	50-53	517.2748	-	-
FPQVAATGDGPDIIFWAHDR	54-73	2212.0701	2212.0683	30779
FGGYAQSGLLAEITPDK	74-90	1765.8886	1765.8909	649718
AFQDK	91-95	607.2966	607.2966	145451
LYPFTWDAVR	96-105	1266.6397	1266.6379	707311
YNGK	106-109	480.2332	-	-
LIAYPIAVEALSLIYNK	110-126	1890.0866	1890.0802	1125441
DLLPNPPK	127-134	892.5018	892.5026	1153583
TWEEIPALDK	135-144	1200.6026	1200.5986	1025526
ELK	145-147	388.2322	388.2326	97875
AK	148-149	217.1426	-	-
GK	150-151	203.127	203.1525	41776
SALMFNLQEPYFTWPLIAADGGYAFK	152-177	2949.4411	2949.4265	92495
YENGK	178-182	609.2758	-	-
YDIK	183-186	537.2798	537.2823	310203

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
DVGVDNAGAK	187-196	944.4563	944.4584	620345
AGLTFLVDLIK	197-207	1188.7118	1188.7132	2156154
NK	208-209	260.1484	260.1497	47684
HMNADTDYSIAEAAFNK	210-226	1896.8312	1896.8299	498842
GETAMTINGPWAWSNIDTSK	227-246	2178.0051	2178.0048	209846
VNYGVTVLPTFK	247-258	1336.7391	1336.7385	1305363
GQPSKPFVGVLSAGINAASPNK	259-280	2138.1483	2138.14	328066
ELAK	281-284	459.2693	459.2704	443306
EFLENYLLTDEGLEAVNK	285-302	2096.0313	2096.0235	787376
DKPLGAVALK	303-312	1010.6124	1010.6085	1294074
SYEEELAK	313-320	967.4498	697.4514	584871
DPR	321-323	386.1914	386.2645	13998
IAATMENAQK	324-333	1075.5332	-	-
GEIMPNIPQMSAFWYAVR	334-351	2109.0176	2109.0233	136786
TAVINAASGR	352-361	958.5196	958.5163	24395
QTVDEALK	362-369	902.4709	902.4685	621221
DAQTNSEFGENLYFQSGEVLNK	370-391	2489.1346	2489.1362	221559
QNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASYFL	392-434	4274.9962	-	-

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
Modified with Iodoacetamide QNVNIIPESEEVGGWVACVG ACGTVCLASGGVGTEFAAAS YFL	392-434	4446.0606	-	-

3.2.2. Trypsin/ Chymotrypsin Digest

Table 3.02. His_6MBP-TEV-TrnB trypsin/chymotrypsin digest products.

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
SHHHHHHHIEEGK	1-13	1620.7406	-	-
VIW	15-17	416.2423	415.2395	179801
INGDK	18-22	545.2809	544.2935	17553
GY	23-24	238.0953	283.0952	44984
NGL	25-27	302.159	302.22	23477
AEVGK	28-32	502.2751	501.2754	47164
EK	35-36	275.1481	275.1488	11509
DTGIK	37-41	532.2857	533.284	768546
VTVEHPDK	42-49	923.4712	922.462	30198
EEK	51-53	404.1907	404.2257	19515
FPQVAATGDGPDIIF	54-68	1546.7667	1546.7598	119607
AHDR	70-73	497.2346	497.2335	28552
GGY	70-77	295.1168	296.1184	515754
AQSGL	78-82	474.2438	474.2405	169001
AEITPDK	84-90	772.3967	772.3961	168119
AF	91-92	236.1161	236.116	138521
QDK	93-95	389.191	390.1907	16055
YPF	97-99	425.1951	424.1777	28352

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
TW	100-101	305.1375	305.1355	356649
DAVR	102-105	459.2441	459.2686	219916
NGK	107-109	317.1699	317.1711	13242
IAYPIAVEAL	111-120	1058.6012	1058.6025	475494
SL	121-122	218.1266	218.1258	95023
IY	123-124	294.1579	294.1569	119467
NK	125-126	260.1484	260.1478	31833
DL	127-128	246.1216	-	-
LPNPPK	139-134	664.3908	-	-
TW	135-144	305.1375	305.1355	356649
EEIPAL	137-142	670.3537	-	-
DK	143-144	261.1325	-	-
EL	145-146	260.1372	260.1478	31833
AK	148-149	217.1426	-	-
GK	150-151	203.127	203.1507	32460
AL	153-154	202.1317	202.1308	93622
NL	157-158	245.1375	245.1509	49423
QEPYF	159-163	682.2962	-	-
TWPL	164-167	515.2744	515.2735	318977
IAADGGY	168-174	665.302	665.3007	392713
AF	175-176	236.1161	236.116	138521
ENGK	179-182	446.2125	-	-
DIK	184-186	374.2165	374.2214	40687
DVGVDNAGAK	187-196	944.4563	944.4565	296985

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
AGL	197-199	259.1532	259.1523	141926
TF	200-201	266.1266	266.1265	246514
VDL	203-205	345.19	-	-
IK	206-207	259.1896	259.1523	141926
NK	208-209	260.1484	260.1478	31833
HM	210-211	286.11	286.4066	80864
NADTDY	212-217	697.2555	697.3083	32846
SIAEAAF	218-224	707.349	707.3446	904691
NK	225-226	260.1484	260.1478	31833
GETAM	227-231	507.1999	507.2661	269106
TINGPW	232-237	686.3388	686.3338	556872
AW	238-239	275.127	275.1252	281955
SNIDTSK	240-246	763.3712	763.3701	177062
VNY	247-249	394.1852	394.184	139571
GVTVLPTF	250-257	832.4694	832.463	363140
GQPSKPF	259-265	759.3915	759.3902	129833
VGVL	266-269	386.2529	386.2522	1071563
SAGINAASPNK	270-280	1028.525	1029.5098	16970
EL	281-282	260.1372	260.1478	31833
AK	283-284	217.1426	-	-
EF	285-286	294.1216	294.1226	28213
ENY	288-290	424.1594	424.1776	28124
TDEGL	293-297	533.2333	533.284	768546
EAVNK	298-302	559.2966	559.2924	44005
DKPL	303-306	471.2693	471.2689	195629
GAVAL	307-311	429.2587	429.2565	513161
SY	313-314	268.1059	-	-
EEEL	315-318	518.2224	-	-
AK	319-320	217.1426	-	-
DPR	321-323	386.1914	-	-
IAATM	324-328	505.257	-	-
ENAQK	329-333	588.2867	589.2899	40242
GEIMPNIPQM	334-343	1128.5307	1128.5294	100016

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
SAF	344-346	323.1481	323.1477	156725
AVR	349-351	344.2172	344.2169	59939
TAVINAASGR	352-361	958.5196	959.4744	30858
QTVDEAL	362-368	774.3759	-	-
DAQTNSEF	370-377	910.3668	910.3599	44099
GENL	378-381	431.2016	431.2359	42932
QSGEVL	384-389	631.3177	631.3123	61613
NK	390-391	260.1484	260.1478	31833
QNVNIIPESEEVGGW	392-406	1669.7947	1669.7854	111498
VACVGACGTVCL	407-418	1094.4922	-	-
Modified with Iodoacetamide VACVGACGTVCL	407-418	1265.5566	-	-
ASGGVGTEF	419-427	823.3712	823.37	331074
AAASY	428-432	481.2172	481.2156	178905

3.2.3. Trypsin/ Glu-C Digest

Table 3.03. His_6MBP-TEV-TrnB trypsin/glu-c digest products.

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
SHHHHHHHIEE	1-11	1435.6242	-	-
GK	12-13	203.127	-	-
LVIWINGDK	14-22	1056.5967	1056.5978	120658
GYNGLAE	23-29	722.3235	722.3261	62121
VGK	30-32	302.1954	-	-
FE	33-34	294.1216	-	-
DTGIK	37-41	532.2857	532.2901	83669

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
VTVE	42-45	446.2376	-	-
HPDK	46-49	495.2441	-	-
LEE	50-52	389.1798	389.1322	154691
FPQVAATGDGPDIIFWAHDR	54-73	2212.0701	2212.0869	211092
FGGYAQSGLLAE	74-85	1211.5822	1211.5849	295290
ITPDK	86-90	572.317	572.3191	44500
AFQDK	91-95	607.2966	607.3016	63865
LYPFTWDAVR	96-105	1266.6397	1266.6467	176391
YNGK	106-109	480.2332	-	-
LIAYPIAVE	110-118	987.564	987.5707	490728
ALSLIYNK	119-126	920.5331	920.5401	653528
DLLPNPPK	127-134	892.5018	-	-
TWEE	135-138	563.2227	563.2772	34575
IPALDK	139-144	655.3904	-	-
LK	146-147	259.1896	259.2529	13008
AK	148-149	217.1426	-	-
GK	150-151	203.127	-	-
SALMFNLQEPYFTWPLIAADGGYAFK	152-177	2949.4411	-	-
YE	178-179	310.1165	-	-
NGK	180-182	317.1699	-	-
YDIK	183-186	537.2798	537.2816	94611
DVGVDNAGAK	187-196	944.4563	944.4637	37378
AGLTFLVDLIK	197-207	1188.7118	1188.721	826956
NK	208-209	260.1484	260.1498	10052
HMNADTDYSIAE	210-221	1365.5507	1365.5518	43200
AAFNK	222-226	549.2911	549.2956	180610

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
GE	227-228	204.0746	-	-
TAMTINGPWAWSNIDTSK	229-246	1991.9411	-	-
VNYGVTVLPTFK	247-258	1336.7391	1336.7443	568735
GQPSKPFVGVLSAGINAASP NK	259-280	2138.1483	2138.1511	154382
LAK	282-284	330.2267	330.2287	20384
FLE	286-288	407.2056	407.2065	8504
NYLLTDE	289-295	866.4021	866.4048	19035
GLE	296-298	317.1587	-	-
AVNK	299-302	430.254	-	-
DKPLGAVALK	303-312	1010.6124	1010.6148	213255
SYEEE	313-317	655.2337	-	-
LAK	318-320	330.2267	330.2291	24397
DPR	321-323	386.1914	-	-
IAATME	324-329	634.2996	-	-
NAQK	330-333	459.2441	459.2469	43904
GE	334-335	204.0746	-	-
IMPNIPQMSAFWYAVR	336-351	1922.9535	-	-
TAVINAASGR	352-361	958.5196	958.521	220253
QTVDE	362-366	590.2548	-	-
ALK	367-369	330.2267	330.2287	20384
DAQTNSE	370-376	763.2984	-	-
FGE	377-379	351.143	-	-
NLYFQSGE	380-387	956.4239	-	-
VLNK	388-391	472.3009	472.3041	14272
QNVNIIPE	392-399	925.4869	925.4928	131091
SEE	400-402	363.1278	-	-
VGGWVACVGACGTVCLASGG	403-422	2151.9751	-	-

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
Modified with Iodoacetamide VGGWVACVGACGTVCLASGG	403-422	2323.0395	-	-
FAAASYFL	427434	888.4381	888.4388	115718

3.2.4. Chymotrypsin Digest

Table 3.04. His_6MBP-TEV-TrnB chymotrypsin digest products.

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
SHHHHHHHIEEGKL	1-14	1733.8247	-	-
INGDKGY	18-24	765.3657	765.3565	58939
AEVGKKF	28-34	777.4385	777.4291	43816
EKDTGIKVTVEHPDKL	35-50	1807.9679	1807.9606	33723
EEKFPQVAATGDGPDIIF	51-68	1932.9468	1932.9346	46439
AHDRF	70-74	644.303	644.2963	149517
AEITPDKAF	84-92	990.5022	990.4936	629744
QDKL	93-96	502.2751	502.2379	27478
DAVRY	102-106	622.3075	622.3035	382306
IAYPIAVEAL	111-120	1058.6012	1058.5986	1448054
LPNPPKTW	129-136	951.5178	-	-
EEIPAL	137-142	670.3537	670.351*	67015*
DKEL	143-146	503.2591	502.2379	27478
KAKGKSAL	147-154	801.5072	-	-
QEPYF	159-163	682.2962	-	-
TWPL	164-167	515.2744	515.2676	1413214
IAADGGY	168-174	665.302	665.2949	564554
ENGKY	179-183	609.2758	610.2711	58425

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
DIKDVGVDNAGAKAGL	184-199	1541.8049	1541.8021**	46669**
IKNKHM	206-211	769.4269	-	-
NADTDY	212-217	697.2555	697.2508	51341
SIAEAAF	218-224	707.349	707.3482**	341945**
NKGETAM	225-231	749.3378	749.3295	74700
TINGPW	232-237	686.3388	686.3359**	270459**
SNIDTSKVNY	240-249	1139.5459	1139.5383	349406
GVTVLPTF	250-257	832.4694	832.4696**	289703**
KGQPSKPF	258-265	887.4865	887.4796	104850
SAGINAASPNKEL	270-282	1270.6517	-	-
TDEGL	293-297	533.2333	533.2856	139682
EAVNKDKPL	298-306	1012.5553	-	-
EEEL	315-318	518.2224	-	-
AKDPRIAATM	319-328	1072.5699	-	-
ENAQKGEIMPNIPQM	329-343	1698.8069	1698.7998**	49838**
AVRTAVINAASGRQTV DEAL	349-368	2041.0916	-	-
KDAQTNSEF	369-377	1038.4618	-	-
QSGEVL	384-389	631.3177	631.3124	219851
NKQNVNIIPESEEVGGW	390-406	1911.9326	1911.9182	275951
VACVGACGTVCL	407-418	1094.4922	1094.5783	37325
Modified with Iodoacetamide VACVGACGTVCL	407-418	1265.5566	-	-
ASGGVGTEF	419-427	823.3712	823.366	544322

3.2.5. Proteinase K Digest

Table 3.05. His_6MBP-TEV-TrnB proteinase K digest products.

Sequence	Position	Predicted mass	Observed mass	Height
SHHHHHHHI	1-9	1177.539	-	-
EEGKL	10-14	574.2962	574.2408	27577
NGDKGY	19-24	652.2816	-	-
NGL	25-27	302.159	302.158	61903
EV	29-30	246.1216	247.1242	23982
GKKF	31-34	478.2903	478.2802	30049
EKDTGI	35-40	661.3282	662.3396	1600083
KV	41-42	245.1739	245.1818	337553
TV	43-44	218.1266	218.1269	381038
EHPDKL	45-50	737.3708	-	-
EEKF	51-54	551.2591	551.3098	32551
PQV	55-57	342.1903	-	-
TGDGPDI	60-66	673.2919	-	-
HDRF	71-74	573.2659	573.2664	146821
GGY	75-77	295.1168	-	-
QSGL	79-82	403.2067	-	-
EI	85-86	260.1372	260.1384	140070
TPDKA	87-91	530.27	-	-
QDKL	93-96	502.2751	502.2303	144989
PF	98-99	262.1317	262.1347	182116
TW	100-101	305.1375	305.1378	312348
DA	102-103	204.0746	-	-
RY	105-106	337.175	337.1714	54611
NGKL	107-110	430.254	430.2573	72169
PI	114-115	228.1474	-	-
EA	118-119	218.0903	-	-
SL	121-12	218.1266	218.1269	381038
NKDL	125-128	488.2594	488.2725	21614
PNPPKTW	130-136	838.4337	-	-
EEI	37-139	389.1798	-	-
PA	140-141	186.1004	-	-

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
DKEL	143-146	503.2591	503.2594	210315
KA	147-148	217.1426	-	-
KGKSA	149-153	489.2911	489.2587	95795
MF	155-156	296.1195	-	-
NL	157-158	245.1375	245.1379	322742
QEPY	159-162	535.2278	535.23	113542
TW	164-165	305.1375	305.1378	312348
PL	166-167	228.1474	-	-
DGGY	171-174	410.1437	410.2369	22260
KY	177-178	309.1688	309.1684	65789
ENGKY	179-183	609.2758	609.315	25945
DI	184-185	246.1216	-	-
KDV	186-188	360.2009	360.2001	66425
GV	189-190	174.1004	174.1002	55342
DNA	191-193	318.1175	-	-
GA	194-195	146.0691	-	-
KA	196-197	217.1426	-	-
GL	198-199	188.1161	188.1166	440621
TF	200-201	266.1266	266.125	733398
DL	204-205	246.1216	-	-
KNKHMNA	207-213	841.4228	-	-
DTDY	214-217	512.1754	512.2941	156589
SI	218-219	218.1266	218.1269	381038
EA	221-222	218.0903	-	-
NKGETA	225-230	618.2973	618.3362	54453
MTI	231-233	363.1828	-	-
NGPW	234-237	472.207	-	-
SNI	240-242	332.1696	-	-
DTSKV	243-247	548.2806	-	-
NY	248-249	295.1168	-	-
GV	250-251	174.1004	174.1002	55342
TV	252-253	218.1266	218.1269	381038
PTF	255-257	363.1794	-	-
KGQPSKPF	258-265	887.4865	-	-
GV	267-268	174.1004	174.1002	55342

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
SA	270-271	176.0797	-	-
GI	272-273	188.1161	188.116	47901
NA	274-275	203.0906	203.1262	28572
SPNKEL	277-282	686.3599	686.3605	73034
KEF	284-286	422.2165	422.2742	43320
ENY	288-290	424.1594	-	-
TDEGL	293-297	533.2333	533.2318	41777
EA	298-299	218.0903	-	-
NKDKPL	301-306	713.4072	713.4067	30125
GA	307-308	146.0691	-	-
KSY	312-314	396.2009	-	-
EEEL	315-318	518.2224	518.2208	58114
KDPRI	320-324	627.3704	-	-
TMENA	327-331	564.2214	-	-
QKGEI	332-336	573.3122	573.2664	146821
MPNI	337-340	473.2308	473.2322	473.2322
PQMSA	341-345	532.2315	-	-
RTA	351-353	346.1965	346.1862	30194
NA	356-357	203.0906	-	-
SGRQTV	359-364	646.3398	646.3199	82269
DEA	365-367	333.1172	333.188	241637
KDA	369-371	332.1696	-	-
QTNSEF	372-377	724.3028	-	-
GENL	378-381	431.2016	431.2096	97118
QSGEV	384-388	518.2336	518.2208	58114
NKQNV	390-394	601.3184	-	-
NI	395-396	245.1375	245.1374	46228
PESEEV	398-403	688.2915	688.4139	361057
GGW	404-406	318.1328	-	-
CV	409-410	220.0882	220.6284	37761
Modified with Iodoacetamide CV	409-410	277.1096	-	-
GA	411-412	146.0691	-	-
CGTV	413-416	378.1573	378.2121	31600

Sequence	Position	Predicted mass (amu)	Observed mass (amu)	Height
Modified with Iodoacetamide CGTV	413-416	435.1788	-	-
CL	417-418	234.1038	-	-
Modified with Iodoacetamide CL	417-418	291.1253	-	-
SGGV	420-423	318.1539	-	-
GTEF	424-427	452.1907	452.1243	28065
SY	431-432	268.1059	268.1063	31323

3.3. Activity Assays: His_6MBP-TEV-TrnB and His_6TrnC/ His_6TrnD

3.3.1. Investigation of His_6TrnC/ His_6TrnD reactivity with S-adenosyl-L-methionine

A characteristic reactivity of radical SAM enzymes is the reduction of S-adenosyl-L-methionine to 5'-deoxyadenosine, both in the presence and absence of other substrates. To probe whether His_6TrnC and His_6TrnD could react with S-adenosyl-L-methionine when not in the presence of substrate, a series of assays were completed, and the production of 5'-deoxyadenosine was monitored. Initially, SAM was incubated with various components of the assay to determine the basal production of 5'-deoxyadenosine in the absence of either enzyme (Please see materials and methods below for assay conditions and components). At the completion of these assays, it was determined that approximately 0.03 equivalents of SAM could cleaved to 5'-deoxyadenosine through background reaction with buffer components. Subsequent assays involving all assay components, including enzyme but lacking TrnA or TrnB, indicated His_6TrnC successfully produced an average of 0.90 equivalents, while His_6TrnD successfully produced an averaged 0.62 equivalents of 5'-deoxyadenosine. The hydrogen donor for this abortive production of 5'-deoxyadenosine was not identified, but is most likely dithiothreitol present at high concentration in the buffer.

3.3.2. Materials and methods

Activity assays were carried out under strictly anaerobic conditions in a MBraun Lab Master 130 and/or under argon. Prior to the beginning of the assay, His_6TrnC and His_6TrnD iron-sulfur clusters were reconstituted. All required reagents were freshly

prepared in buffer containing 25 mM Tris-base, 150 mM NaCl, and 10 mM 1,4-dithiothreitol, pH 8. The method was carried out as follows:

Reconstitution assay (under argon): initially, 100 equivalents of sodium dithionite were added to 150 μM protein and incubated on ice for 1 hr. Next, 10 equivalents of ammonium iron citrate were gradually added during gentle mixing, and the protein was incubated for 5 min on ice. After which, 10 equivalents of lithium sulfide were added and the protein was incubated for 1 hr on ice.

Following reconstitution, the activity assays were carried out as follows:

> Activity assay: sufficient amounts were added to achieve final concentrations of 10 μM His_6TrnC (or His_6TrnD), 1 mM SAM, 1 mM sodium dithionite, and 100 μM His_6MBP-TEV-TrnB. The reaction was then incubated for 4 hr at room temperature.

Activity assays containing His_6TrnC only, His_6TrnD only, and both His_6TrnC and His_6TrnD were completed to gain a potentially understanding of which enzymes, if either, are responsible for catalyzing the thioether crosslinking reactions.

3.3.3. Purification

After the completion of the activity assay, it was necessary to purify the His_6MBP-TEV-TrnB from His_6TrnC (or His_6TrnD). The reaction was loaded onto an Amylose flash column (1 ml) and washed with 25 mM Tris-base, pH 8 (5 column volumes). The His_6MBP-TEV-TrnB was subsequently eluted in 25 mM Tris-base, pH 8, 10 mM maltose.

Samples were concentrated using Millipore Biomax-10k NMWL membrane centrifugal filters.

3.3.4. Cleavages and Expected Products

As a result of undefined LCMS conditions for detection of the unmodified/modified His_6MBP-TEV-TrnB, it was necessary to monitor activity by first incubating with digestion enzymes chymotrypsin or proteinase K following the *enzymatic digestion protocol* (page 164)

Tables 3.06, 3.07, 3.08, and 3.09 show the predicted masses for fragments generated during the enzymatic digest of His_6TrnC and His_6TrnD by chymotrypsin and proteinase K, as a result of unsuccessful amylose resin purification.

3.3.4.1. Chymotrypsin Digestion of His_6TrnC

Table 3.06. His_6TrnC chymotrypsin predicted digest products.

Sequence	Mass (amu)
GSSHHHHHHSSGL	1415.6191
VPRGSHM	782.3857
SKKRL	630.4177
SM	236.0831
MPF	393.1722
KTDKF	637.3435
SY	268.1059
AL	202.1317
DGNTGRVIVADKPTL	1554.8365
IISHF	615.338
HKF	430.2328
EKEEL	646.3173
KKTKGKF	835.5279
AEL	331.1743

Sequence	Mass (amu)
HQDY	561.2183
TTY	383.1692
NY	295.1168
VSSL	404.2271
INM	376.178
GM	206.0725
SEKEDSDSPIDSKEL	1677.7581
AINSNQSQL	973.4829
IL	244.1787
IL	244.1787
TEKCNL	706.332
RCEY	569.2268
CIY	397.1671
NDKYPKEM	1023.4695
GY	238.0953
SDEEM	609.1952
DF	280.1059
ETAKKAVDM	991.5008
EL	260.1372
HM	286.11
ERVKRGHKRFPVITM	1853.057
GGEPL	471.2329
KF	293.1739
DL	246.1216
IKKVM	617.3934
EY	310.1165
AKGL	387.2481
MPDTL	575.2625
TTTNGTL	706.3497
SEKM	493.2206
DY	296.1008
INNRIIITF	1102.6499
SIDGF	537.2435
KENHDRNRVF	1313.6589
VNGMPTF	764.3527
ERAF	521.2598
KNIKRL	770.5126

Sequence	Mass (amu)
QEKKKEQNIEQIISF	1860.9945
NCCF	485.1403
DQY	424.1594
TDVY	496.2169
KVAKF	591.3744
EEHY	576.218
DL	246.1216
NPF	376.1747
VL	230.163
NQINPF	731.3602
DTL	347.1692
DW	319.1168
CDEQVKTGKW	1192.5546
NF	279.1219
DKNNF	636.2867
KNAM	462.226
QKIEHEL	895.4763
EAETCDDHF	1065.3709
QQVAGPL	711.3915
VM	248.1195
KDF	408.2009
VL	230.163
SIRNKDGQQQITRNSCIPTSKIAVSPDGTL	3226.6728
TL	232.1423
CEKM	509.1978
CKKYPIGTVEKGL	1434.7904
DW	319.1168
KAVDGVTEKL	1058.5971
VRHF	557.3074
NSDSCKY	815.312
CPIRTM	719.3458
CEACF	571.177
DENGRIKPSF	1161.5778
CKSKKM	723.3771
AVKKNL	671.433
ESY	397.1485
AKKEKGF	806.465

Sequence	Mass (amu)
DVM	363.1464
KVY	408.2373
NHTSDL	685.3031
DSVKEM	707.316
VK	245.1739

3.3.4.2. Proteinase K Digestion of His_6TrnC

Table 3.07. His_6TrnC proteinase K predicted digest products.

Sequence	Mass (amu)
GSSHHHHHHSSGL	1415.6191
PRGSHMSKKRL	1295.7244
SMMPF	611.2447
KTDKF	637.3435
SY	268.1059
DGNTGRV	717.3405
DKPTL	572.317
SHF	389.1699
HKF	430.2328
EKEEL	646.3173
KKTKGKF	835.5279
EL	260.1372
HQDY	561.2183
TTY	383.1692
NY	295.1168
SSL	305.1587
NMGMF	598.2243
SEKEDSDSPI	1105.4775
DSKEL	590.2911
NSNQSQL	789.3617
TEKCNL	706.332
RCEY	569.2268
CI	234.1038
NDKY	538.2387

Sequence	Mass (amu)
PKEMGY	723.3261
SDEEMDF	871.2905
ETA	319.1379
KKA	345.2376
DMY	427.1413
EL	260.1372
HMERV	670.3221
KRGHKRF	927.5515
PV	214.1317
TMY	413.1621
GGEPL	471.2329
KF	293.1739
DL	246.1216
KKV	373.2689
MEY	441.157
KGL	316.211
MPDTL	575.2625
TTTNGTL	706.3497
SEKMMDY	902.3514
NNRI	515.2816
TF	266.1266
SI	218.1266
DGF	337.1274
KENHDRNRV	1166.5904
NGMPTF	665.2843
ERA	374.1914
KNI	373.2325
KRL	415.2907
QEKKKEQNI	1143.6248
EQI	388.1958
SF	252.111
NCCF	485.1403
DQY	424.1594
TDV	333.1536
KV	245.1739
KF	293.1739
EEHY	576.218

Sequence	Mass (amu)
DL	246.1216
NPF	376.1747
NQI	373.1961
NPF	376.1747
DTL	347.1692
DW	319.1168
CDEQV	592.2163
KTGKW	618.3489
NF	279.1219
DKNNF	636.2867
KNA	331.1856
MQKI	518.2886
EHEL	526.2387
EA	218.0903
ETCDDHF	865.2912
QQV	373.1961
GPL	285.1688
MKDF	539.2414
SI	218.1266
RNKDGQQQI	1085.5578
TRNSCI	692.3276
PTSKI	544.322
SPDGTL	588.2755
TL	232.1423
CEKMCKKY	1031.4602
PI	228.1474
GTV	275.1481
EKGL	445.2536
DW	319.1168
KA	217.1426
DGV	289.1274
TEKL	489.2798
RHF	458.239
NSDSCKY	815.312
CPI	331.1566
RTMCEA	709.2887
CF	268.0882

Sequence	Mass (amu)
MF	296.1195
DENGRI	702.3296
KPSF	477.2587
CKSKKMA	794.4143
KKNL	501.3275
ESY	397.1485
KKEKGF	735.4279
DV	232.1059
MKV	376.2144
NHTSDL	685.3031
DSV	319.1379
KEMV	505.257

3.3.4.3. Chymotrypsin Digestion of His_6TrnD

Table 3.08. His_6TrnD chymotrypsin predicted digest products.

Sequence	Mass (amu)
GSSHHHHHHSSGL	1415.6191
VPRGSHM	782.3857
KKY	437.2638
F	165.079
RL	287.1957
YPYCHL	794.3421
EIGETNSCL	964.4171
Y	181.0739
DISSGKM	736.3425
IRVNRENAEL	1212.6575
L	131.0946
RQCQENVPIESINM	1659.7709
DL	246.1216
GIL	301.2001
DEL	375.1641
IKM	390.2301
NL	245.1375
GTY	339.143
Y	181.0739

Sequence	Mass (amu)
ANPQF	575.2703
IEPF	504.2584
F	165.079
ETNDTKNRIF	1236.6098
GKNNIL	657.381
RQM	433.2107
F	165.079
IL	244.1787
TSTDCNM	770.2575
NCKHCNTDSTVF	1367.5598
RKTGCKIWPKSINL	1642.9341
NAL	316.1747
TQSHW	657.2871
RKIL	528.3747
EAF	365.1587
Y	181.0739
NL	245.1375
HGEEL	583.2602
TF	266.1266
IGGEPF	618.3013
L	131.0946
EF	294.1216
DF	280.1059
IKNIVEIAQEVGISKF	1787.0192
SIF	365.1951
TNGSIINDTIL	1159.6084
NF	279.1219
L	131.0946
M	149.051
ENKIKVY	892.5018
IQIF	519.3057
EVDENKF	879.3974
KAF	364.211
TNSDIPSIQIIDNIKKL	1911.0677
NNHHL	633.2983
DL	246.1216
QL	259.1532
RIL	400.2798
ITRDNDNNL	1073.5101
KKIVNTL	814.5276

Sequence	Mass (amu)
QKETNVKDIKIEF	1590.8617
L	131.0946
YPKPDNSY	982.4396
Y	181.0739
SKKY	524.2958
IPL	341.2314
M	149.051
Y	181.0739
DKKREF	821.4395
SHVNVQKM	941.4753
QF	293.1375
L	131.0946
HQY	446.1914
NPSF	463.2067
F	165.079
GQITIRRDGKVVPHPM	1802.9937
L	131.0946
TRVIGDL	772.4443
QQDDL	617.2657
F	165.079
TIINTEEY	981.4655
QEY	438.1751
STL	319.1743
NKEKISKCSTCAY	1473.6956
KY	309.1688
NCM	366.1032
DDRVIENF	1006.4719
ATGDL	475.2278
Y	181.0739
GM	206.0725
EY	310.1165
CNF	382.1311

3.3.4.4. Proteinase K Digestion of His_6TrnD

Table 3.09. His_6TrnD proteinase K digest products.

Sequence	Masses (amu)
GSSHHHHHHSSGL	1415.6191
V	117.079
PRGSHMKKY	1102.5706
F	165.079
RL	287.1957
Y	181.0739
PY	278.1266
CHL	371.1627
EI	260.1372
GETNSCL	722.2905
Y	181.0739
DI	246.1216
SSGKMI	621.3156
RV	273.1801
NRENA	602.2772
EL	260.1372
L	131.0946
RQCQENV	875.3919
PI	228.1474
ESI	347.1692
NMDL	491.205
GI	188.1161
L	131.0946
DEL	375.1641
I	131.0946
KMNL	504.273
GTY	339.143
Y	181.0739
A	89.0477
NPQF	504.2332
I	131.0946
EPF	391.1743
F	165.079
ETNDTKNRI	1089.5414
F	165.079

Sequence	Mass (amu)
GKNNI	544.2969
L	131.0946
RQMF	580.2791
I	131.0946
TSTDCNMNCKHCNT DSTV	1972.7383
F	165.079
RKTGCKI	804.464
W	204.0899
PKSI	443.2744
NL	245.1375
NA	203.0906
L	131.0946
TQSHW	657.2871
RKI	415.2907
L	131.0946
EA	218.0903
F	165.079
Y	181.0739
NL	245.1375
HGEEL	583.2602
TF	266.1266
I	131.0946
GGEPF	505.2172
L	131.0946
EF	294.1216
DF	280.1059
I	131.0946
KNI	373.2325
V	117.079
EI	260.1372
A	89.0477
QEV	374.1801
GI	188.1161
SKF	380.206
SI	218.1266
F	165.079
TNGSI	490.2387
I	131.0946
NDTI	461.2122

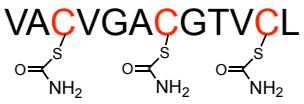
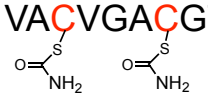
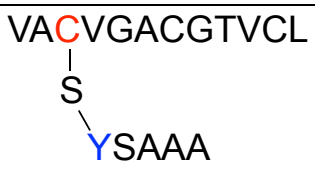
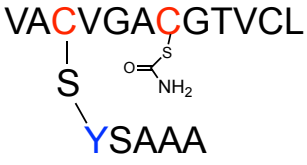
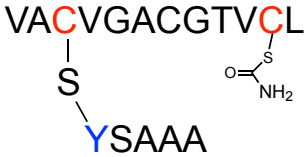
Sequence	Mass (amu)
L	131.0946
NF	279.1219
L	131.0946
MENKI	633.3156
KV	245.1739
Y	181.0739
I	131.0946
QI	259.1532
F	165.079
EV	246.1216
DENKF	651.2864
KA	217.1426
F	165.079
TNSDI	548.2442
PSI	315.1794
QI	259.1532
I	131.0946
DNI	360.1645
KKL	387.2845
NNHHL	633.2983
DL	246.1216
QL	259.1532
RI	287.1957
L	131.0946
TRDNDNNL	960.4261
KKI	387.2845
V	117.079
NTL	346.1852
QKETNV	717.3657
KDI	374.2165
KI	259.1896
EF	294.1216
L	131.0946
Y	181.0739
PKPDNSY	819.3763
Y	181.0739
SKKY	524.2958
I	131.0946
PL	228.1474
MY	312.1144

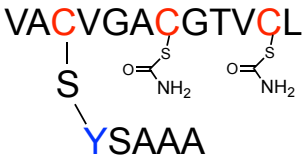
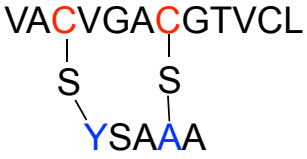
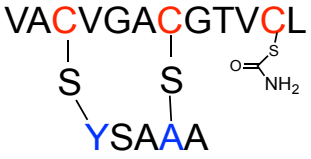
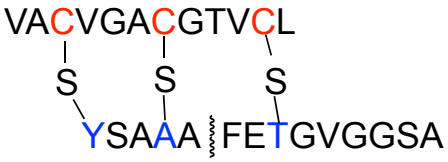
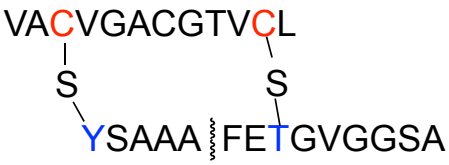
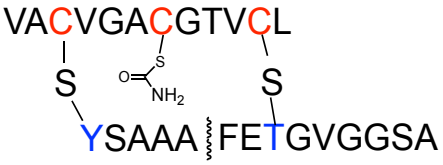
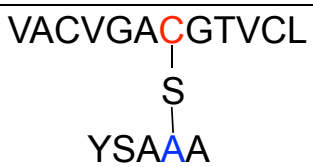
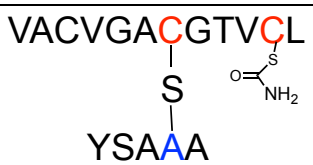
Sequence	Mass (amu)
DKKREF	821.4395
SHV	341.1699
NV	231.1219
QKMQF	680.3316
L	131.0946
HQY	446.1914
NPSF	463.2067
F	165.079
GQI	316.1747
TI	232.1423
RRDGKV	729.4245
V	117.079
PHPML	593.2995
L	131.0946
TRV	374.2278
I	131.0946
GDL	303.143
QQDDL	617.2657
F	165.079
TI	232.1423
I	131.0946
NTEEY	654.2497
QEY	438.1751
STL	319.1743
NKEKI	630.3701
SKCSTCA	698.2727
Y	181.0739
KY	309.1688
NCMDDRV	851.3266
I	131.0946
ENF	408.1645
A	89.0477
TGDL	404.1907
Y	181.0739
GMEY	498.1784
CNF	382.1311

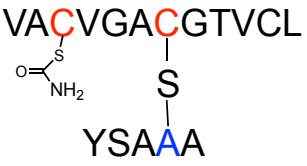
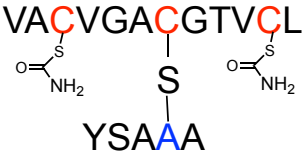
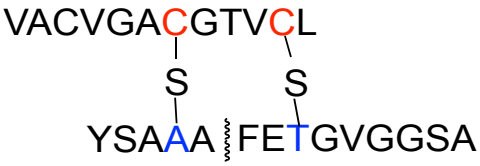
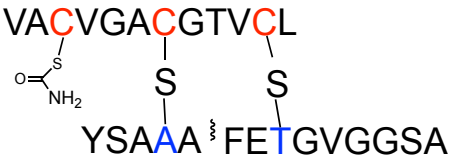
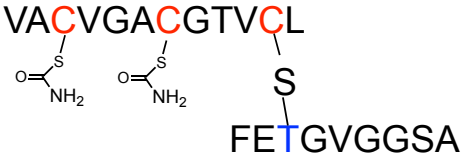
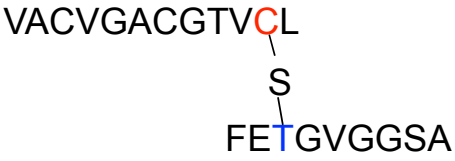
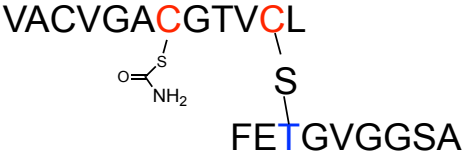
3.3.5. Expected assay products

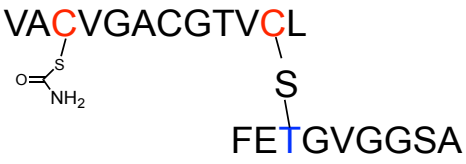
Table 3.10 displays the cysteine-containing fragment resulting from enzymatic digestion with chymotrypsin, and the predicted cross-linked or uncross-linked products of the activity assay incubated with iodoacetamide.

Table 3.10. Activity assay expected products post chymotrypsin digestion.

Conditions	Predicted peptide fragment	Mass (amu)
No assay + chymo 23-34	VACVGACGTVCL	1094.6083
No assay + IOA		1265.6083
No assay + IOA		1208.6083
Assay: C25-Y48 only		1573.8679
Assay: C25-Y48 only (IOA at one Cysteine)		1630.8679
Assay: C25-Y48 only (IOA at one Cysteine)		1630.8679

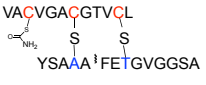
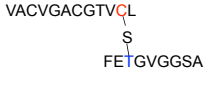
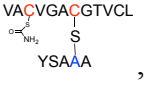
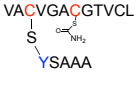
Conditions	Predicted peptide fragment	Mass (amu)
Assay: C25-Y48 only (IOA at C29 and C33)		1687.8679
Assay: C25-Y48 + C29-A45 only		1571.8679
Assay: C25-Y48 + C29-A45 only (IOA)		1628.8679
Assay C25-Y48 + C29-A45 + C33-Y48 only		2393.3235
Assay: C25-Y48+C33-T41 only		2395.3235
Assay: C25-Y48+C33-T41 only (IOA)		2452.3235
Assay: C29-A45 only		1573.8679
Assay: C29-A45 only (IOA one cysteine)		1630.8679

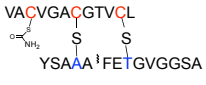
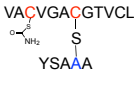
Conditions	Predicted peptide fragment	Mass (amu)
Assay: C29-A45 only (IOA one cysteine)		1630.8679
Assay: C29-A45 only (IOA C25+C33)		1687.8679
Assay: C29-A45+C33-T41 only		2395.3235
Assay: C29-A45+C33-T41 only (IOA)		2452.3235
Assay: C33-T41 only (IOA C25 and C29)		2030.0639
Assay C33-T41 only		1916.0639
Assay: C33-T41 only (IOA one cysteine)		1973.0639

Conditions	Predicted peptide fragment	Mass (amu)
Assay: C33-T41 only (IOA one cysteine)		1973.0639

3.3.6. Analysis

Treatment of MBP-TrnB with chymotrypsin liberates characteristic peptides, ASGGVGTEF, containing Thr21, and AAASY, containing Ala25 and Tyr28, which are predicted to become crosslinked to Cys13, Cys9 and Cys5, respectively. Chymotrypsin also liberates the characteristic peptide, VACVGACGTVCL, containing Cys13, Cys9, and Cys5. Treatment with iodoacetamide generates a product with predicted [M-H_2O+H^+] = 824.4 for the liberated peptide, ASGGVGTEF. Treatment of MBP-TrnB with TrnC suggests the formation of crosslinks, as observed through LCMS-ESI-TOF.

Treatment of the products with iodoacetamide, followed by treatment with chymotrypsin generates the following crosslinked products:  ,with predicted [M-H_2O+3H^+] = 814.4;  , with predicted [M] = 1916.0639; and  , with predicted [M] = 1630.8679. It should be stated that a predicted [M] = 1630.8679 also corresponds to the fragment  , but given the detection of the unique mass

specific to fragment  , it is assumed the mass corresponds to  . These results suggests that TrnC is responsible for catalysis of the crosslinking reactions observed between Cys9 and Ala29, and Cys13 and Thr21, as this product is observed in propeptide modifying assays containing TrnC and TrnC + TrnD, but not in assays containing TrnD.

Chapter 4

Discussion, Conclusion, and Future Directions

4.1. Discussion

The process of purifying propeptides TrnA and TrnB began first with direct cloning of the gene(s) into an appropriate vector. After early failed attempts to isolate the products through expression of the genes in a pET23b(+) vector, it became clear that transitioning to a fusion protein technique would yield the desired products. Peptide products that lack secondary structure are susceptible to proteolysis. In the native organism, the propeptides undergo post-translational modifications to inhibit this proteolysis and to give the final product the structure necessary for activity; however, during recombinant expression, these post-translational modifications to do not occur. Therefore, while the cloning of TrnA and TrnB into the pMALc2x vector to create an N-terminal maltose binding protein fusion was successful, it did not take into account the susceptibility of the propeptides to C-terminal proteolysis during expression and purification. However, in addition to the successful cloning, modifications to the pMALc2x fusion construct including the removal of the factorXa cleavage site and addition of the Tobacco Etch Virus cleavage site, addition of a hexahistidine affinity tag, optimization of the Tobacco Etch Virus cleavage site, addition of a pentaserine spacer, and removal of the decaasparagine spacer were successful using ligation free polymerase chain reaction techniques. To prevent the possibility of C-terminal proteolysis, maltose binding protein-TrnA (B) were successfully

cloned into the pTWIN I vector, using conventional polymerase chain reaction techniques, to create a maltose binding protein- TrnA (B)- intein-chitin binding domain fusion. From this design, constructs including the removal of the maltose binding protein and addition of an N-terminal hexahistidine affinity tag, removal of the N-terminal hexahistidine tag and addition of an N-terminal decaaspartate affinity tag, and addition of a C-terminal hexahistidine tag were successfully cloned using modern ligation free polymerase chain reaction techniques. In addition to cloning of the propeptides, cloning of the radical SAM enzyme, TrnC was also successful.

Purifications of the overexpressed His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB; G395S/M396G His_6MBP-TEV-TrnA and G395S/M396G His_6MBP-TEV-TrnB; $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA and $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB; (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA and (-Asn_{10}) G385S/M386G His_6MBP-TEV-TrnB; His_6MBP-TEV-TrnB-intein-CBD; (W11) His_8TrnB-intein-CBD; Asp_{10}TEV-TrnB-intein-CBD-His_6; and His_6TrnC were completed with the use of nickel affinity chromatography, as well as MBP-factorXa-TrnA and MBP-factorXa-TrnB; and Asp_{10}TEV-TrnB-intein-CBD with the use of DEAE ion exchange chromatography, based on Coomassie staining of the sodium dodecyl sulfate-polyacrylamide gel electrophoresis, with respect to the masses observed in Table 4.01:

Table 4.01. Fusion protein constructs and the expected average mass.

Construct	Approximate expected M_{av} (Daltons)
His_6MBP-TEV-TrnA His_6MBP-TEV-TrnB	48,000 48,500
G395S/M396G His_6MBP-TEV-TrnA; G395S/M396G His_6MBP-TEV-TrnB	48,000 48,500
$Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA; $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB	48,500 49,000
(-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA; (-Asn_{10}) G385S/M386G His_6MBP-TEV-TrnB	47,000 47,500
His_6MBP-TEV-TrnB-intein-CBD	75,000
(W11) His_8TrnB-intein-CBD	34,500
Asp_{10}TEV-TrnB-intein-CBD-His_6	36,000
MBP-factorXa-TrnA; MBP-factorXa-TrnB	47,000 47,500
Asp_{10}TEV-TrnB-intein-CBD	35,000
His_6TrnC	60,000

Following purification, successful cleavages were also observed in His_6MBP-TEV-TrnA and His_6MBP-TEV-TrnB; G395S/M396G His_6MBP-TEV-TrnA and G395S/M396G His_6MBP-TEV-TrnB; $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnA and $Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV-TrnB; (-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA and (-Asn_{10}) G385S/M386G His_6MBP-TEV-TrnB; His_6MBP-TEV-TrnB-intein-CBD;

(W11)His_8TrnB-intein-CBD; Asp_{10}TEV-TrnB-intein-CBD-His_6 and Asp_{10}TEV-TrnB-intein-CBD with respect to the cleaving agent and masses observed in Table 4.02:

Table 4.02. Fusion protein constructs, cleavage products, and the expected average mass.

Construct	Approximate expected M_{av} (Daltons)	Cleavage agent	Products	Approximate expected product M_{av}(Daltons)
His_6MBP-TEV-TrnA	48,000	TEV	His_6MBP-TEV	43,489.9
			TrnA	4,658.4
His_6MBP-TEV-TrnB	48,500		His_6MBP-TEV	43,489.9
			TrnB	4,992.6
G395S/M396G His_6MBP-TEV-TrnA;	48,000	TEV	G395S/M396G His_6MBP-TEV	43,489.9
			TrnA	4,671.3
G395S/M396G His_6MBP-TEV-TrnB	48,500		G395S/M396G His_6MBP-TEV	43,489.9
			TrnB	5005.6
$Ser_{5(388-392)}$ G400S/M401G	48,500	TEV	$Ser_{5(388-392)}$ G400S/ M401G His_6MBP-TEV	43,925.3
His_6MBP-TEV-TrnA;			TrnA	4,671.3
$Ser_{5(388-392)}$ G400S/M401G	49,000		$Ser_{5(388-392)}$ G400S/ M401GHis_6MBPTEV-	43,925.3
His_6MBP-TEV-TrnB			TrnB	5,005.3
(-Asn_{10}) G385S/M386G His_6 MBP-TEV-TrnA;	47,000	TEV	(-Asn_{10}) G385S/M386G His_6 MBP-TEV	42,348.8
			TrnA	4,671.3

Construct	Approximate expected M_{av} (Daltons)	Cleavage agent	Products	Approximate expected product M_{av}(Daltons)
(-Asn_{10}) G385S/M386G His_6MBP-TEV-TrnB	47,500		(-Asn_{10}) G385S/M386G His_6MBP-TEV TrnB	42,384.8 5,005.3
His_6MBP-TEV-TrnB-intein-CBD	75,000	TEV; DTT	His_6MBP-TEV TrnB intein-CBD	42,348.8 5,005.3 27,858.4
(W11) His_8TrnB-intein-CBD	34,500	DTT	W11) His_8TrnB intein-CBD	6,507.2 27,858.4
Asp_{10}TEV-TrnB-intein-CBD-His_6	36,000	BME	Asp_{10}TEV-TrnB intein-CBD-His_6	7169.9 28,839.4
Asp_{10}TEV-TrnB-intein-CBD	35,000	BME	Asp_{10}TEV-TrnB intein-CBD	7169.9 27,858.4

While observation of the fusion partner through coomassie staining of the sodium dodecyl sulfate-polyacrylamide gel electrophoresis following cleavages was successful in all constructs, observation of the target propeptides was unsuccessful. Presumably, observation of the propeptide through coomassie staining is unlikely due to the lack of aromatic residues, resulting in a low binding affinity for Coomassie dye; however, it is a useful method to determine the success of cleavage reactions based on the mass loss of the fusion partner when compared to the full length construct.

Isolation of the propeptide TrnB was achieved only after extraction with isopropanol and was only detectable using LCMS. After cleavage of chitin binding domain, re-

purification, and cleavage of maltose binding protein, the propeptide was extracted in 70% isopropanol and analyzed through LCMS-TOF. Mass peaks corresponding to $[M+H^+]^{+1}$ = 5002]; [M-OH=4984]$^{1+}$; $[M+2H^+]^{+2}$ = 2493; $[M+3H^+]^{+3}$ =1662; $[M+4H^+]^{+4}$ = 1247; $[M+5H^+]^{+5}$ = 998] were identified. Given that the propeptide could only be isolated and observed after extraction utilizing non-biological conditions and at a microgram scale, direct assay of the physiologic propeptide seemed unfeasible, and methods for performing activity assays with the substrate bound to maltose binding protein were developed.

To analyze the success of the activity assays, a method for observation of the cysteine-containing segment and the predicted modified residues of TrnB was first developed. A series of enzymatic digests were performed with various protease enzymes to determine if A) the propeptide could in fact be cleanly digested, and B) which enzyme(s) were best for producing an observable product. After analysis of cleavage products produced by trypsin, Glu-C, chymotrypsin, and proteinase K, it was determined that chymotrypsin alone not only produced a small fragment containing all three cysteines of interest in the propeptide, but also two additional small fragments containing the modified residues of of interest. The remaining three digestion enzymes yielded results that further confirmed the presence of the propeptide in the fusion system, as well as providing confirmation of the amino acid sequence of the fusion partner MPB, however, these digests provided no evidence as to why the isolation of the cleaved product was unsuccessful.

Resulting from the necessity to enzymatically digest the fusion system, it was essential to purify the product away from the radical SAM enzymes following the assay, in order to reduce the number of additional peptides present in the digestion mix. Modified

His_6MBP-TEV-TrnB was successfully purified away from His_6TrnC and His_6TrnD using amylose affinity chromatography. Subsequent enzymatic digestion and analysis through LCMS-TOF provided results that suggest that the enzyme TrnC is responsible for catalyzing the crosslinking reactions between Cys 9 and Ala 25, and Cys 13 and Thr 21, two of the three crosslinks observed in thuricin CD originally purified from *B. thuringiensis*.

4.2. Conclusion

Thuricin CD is two-component bacteriocin consisting of two peptides, TrnA and TrnB, produced by *Bacillus thuringiensis* and specifically targeting the bacterium *Clostridium difficile*. Both TrnA and TrnB contain three thioether crosslinks in which cysteine residues are chemically bonded to the α-carbon of distal amino acid residues. This places thuricin CD in the recently discovered category of ribosomally produced peptide antibiotics known as sactibiotics. Sactibiotic thioether crosslinks have been show to be introduced by the radical SAM enzymes AlbA and SKF in two other natural products, subtilosin and sporulation killing factor, respectively. Analogous gene products in the thuricin CD system are the proposed radical SAM enzymes TrnC and/or TrnD.

The development of a method to purify propeptides TrnA and TrnB was unsuccessful. Transitioning from single expression of the propeptides to an N-terminal fusion method was believed to increase expression levels and allow for a definitive approach to purification. While expression levels successfully increased, isolation of the target product was not achieved. From there, the development of double fusion partner construct was made in an effort to allow for multiple methods of purification, cleavage, and ultimately isolation of the target propeptides. While this approach permitted multiple angles to purify the product, isolation was unsuccessful. To account for the overall hydrophobicity of TrnA and TrnB, the propeptides were modified to include an N-terminal hydrophilic affinity tag. Though purification of this construct was successful, post-cleavage isolation of the more “water soluble” propeptides was not achieved.

Despite failing to develop a method for purification of TrnA and TrnB, detection of TrnB was successfully achieved. With the use of the double fusion His_6MBP-TEV-TrnB-intein-CBD, TrnB was extracted in isopropanol following cleavage of both His_6MBP and CBD, and was subjected to LCMS-TOF where it was successfully detected. This detection suggests the development of a more efficient extraction method may allow for future isolation of the propeptides.

In addition, the expression and purification of TrnC was successful. With TrnD previously purified, development of assay conditions was attempted. Due to the inability to purify TrnA and TrnB, the development of these assays made use of the His_6MBP-TEV-TrnB fusion. Initial results indicate that both TrnC and TrnD are successfully catalyzing the cleavage of S-adenosyl-L-methionine into 5'-deoxyadenosine and methionine. Furthermore, preliminary results suggest TrnC is responsible for catalyzing the crosslinking reactions between Cys9 and Ala29, and Cys13 and Thr21.

Future work on this project undoubtedly involves the purification of propeptides TrnA and TrnB. While development of a more efficient extraction method may prove successful, an interesting approach involving fluorescent probes may be applied. Nucleophilic attack at the intein by a fluorescent probe would successfully cleave the propeptide while leaving the probe bound. Making use of a highly water soluble probe would not only provide a fluorescent tag to aid in detection, but may also help solubilize the propeptide.

Furthermore, more extensive characterization of propeptide modifying radical SAM enzymes TrnC and TrnD is also necessary. Studies involving the mutation of cysteines

present in the catalytic radical SAM and SPASM domains may provide further insight on the mechanism by which TrnA and TrnB are post-translationally modified.

References

(1) Rea, M. C.; Sit, C. S.; Clayton, E.; O'Connor, P. M.; Whittal, R. M.; Zheng, J.; Vederas, J. C.; Ross, R. P.; Hill, C. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (20), 9352–9357.

(2) Ling, Z.; Liu, X.; Jia, X.; Cheng, Y.; Luo, Y.; Yuan, L.; Wang, Y.; Zhao, C.; Guo, S.; Li, L.; Xu, X.; Xiang, C. *Sci. Rep.* **2014**, *4*, 7485.

(3) Redelings, M. D.; Sorvillo, F.; Mascola, L. **2007**, *13* (9), 7–9.

(4) Willey, J. M.; van der Donk, W. a. *Annu. Rev. Microbiol.* **2007**, *61*, 477–501.

(5) Yang, Y.-L.; Xu, Y.; Straight, P.; Dorrestein, P. C. *Nat. Chem. Biol.* **2009**, *5* (12), 885–887.

(6) Chatterjee, C.; Miller, L. M.; Leung, Y. L.; Xie, L.; Yi, M.; Kelleher, N. L.; Van Der Donk, W. a. *J. Am. Chem. Soc.* **2005**, *127*, 15332–15333.

(7) Knerr, P. J.; van der Donk, W. A. *Annu. Rev. Biochem.* **2012**, *81* (1), 479–505.

(8) Kawulka, K. E.; Sprules, T.; Diaper, C. M.; Whittal, R. M.; McKay, R. T.; Mercier, P.; Zuber, P.; Vederas, J. C. *Biochemistry* **2004**, *43* (12), 3385–3395.

(9) Flühe, L.; Knappe, T. a; Gattner, M. J.; Schäfer, A.; Burghaus, O.; Linne, U.; Marahiel, M. a. *Nat. Chem. Biol.* **2012**, *8* (8), 737–737.

(10) De Gruyte. **2015**.

(11) Flühe, L.; Burghaus, O.; Wieckowski, B. M.; Giessen, T. W.; Linne, U.; Marahiel, M. a. *J. Am. Chem. Soc.* **2013**, *135* (3), 959–962.

(12) Protein Expression and Analysis; pMAL Protein Fusion and Purification System. Instructional Manual. New England Biolabs.

(13) Kapust, R. B.; Tözsér, J.; Copeland, T. D.; Waugh, D. S. *Biochem. Biophys. Res. Commun.* **2002**, *294* (5), 949–955.

(14) Protein Expression and Analysis; IMPACT Kit. Instructional Manual. New England Biolabs.