
Theses and Dissertations

Student Publications

2017

Design, Cloning, and Expression of Rat Preptin and Alanine Analog

Jahbo M. Love

Follow this and additional works at: https://csuepress.columbusstate.edu/theses_dissertations

 Part of the [Chemistry Commons](#)

Recommended Citation

Love, Jahbo M., "Design, Cloning, and Expression of Rat Preptin and Alanine Analog" (2017). *Theses and Dissertations*. 291.

https://csuepress.columbusstate.edu/theses_dissertations/291

This Thesis is brought to you for free and open access by the Student Publications at CSU ePress. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of CSU ePress.

DESIGN, CLONING, AND EXPRESSION OF RAT PREPTIN
AND ALANINE ANALOGS

Jahbo M. Love

COLUMBUS STATE UNIVERSITY

DESIGN, CLONING, AND EXPRESSION OF RAT PREPTIN AND ALANINE ANALOGS

Copyright © 2017 Jahbo M. Love
All Rights Reserved.

A THESIS SUBMITTED TO
THE COLLEGE OF LETTERS AND SCIENCES
IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF CHEMISTRY

BY

JAHBO M. LOVE

COLUMBUS, GEORGIA

2017

DESIGN, CLONING, AND EXPRESSION OF RAT PREPTIN AND ALANINE ANALOGS

Copyright © 2017 Jahbo M. Love

All Rights Reserved.

Dr. Jonathan Meyer

Committee Members

Dr. Neil Sauer
Dr. David Halley

DESIGN, CLONING, AND EXPRESSION OF RAT PREPTIN AND ALANINE ANALOGS

By

Jahbo M. Love

Committee Chair:

Dr. Jonathan Meyers

Committee Members:

Dr. Anil Banerjee
Dr. Danial Holley

TO MY MOTHER

ABSTRACT

This work is the culmination of 26 years of schooling in which I was never thought to be able to succeed. Preptin, a newly discovered polypeptide found on the proinsulin-like growth factor II. The alanine scanning mutagenesis was accomplished by the use of, cloning, affinity chromatography, gel electrophoresis, gene sequencing, and mass spectra. However, due to inconclusive results, more research is required before the epitopes can be elucidated.

INDEX WORDS: Preptin, Diabetes, Cancer, Alanine Scanning Mutagenesis

ACKNOWLEDGMENTS

This work is the culmination of 26 years of schooling in which I was never thought to be able to succeed. Perhaps in the last few years, the above statement is not entirely true, due to my learning disability I was not expected to be accepted at a University, nor was I expected to make it to the Masters level. I have done so through the determination of myself and my single-parent Mother. She has supported and given me the opportunity to pursue my dreams of being a scientist. This work and all works after stem from the foundations you first laid down Mother.

Spectrometry, for her assistance in analyzing mass spectra data.

ACKNOWLEDGEMENTS

I would like to take a moment to thank everyone who helped me throughout my bachelors and Masters degree at Columbus State University. Special thanks to my mentor Dr. Meyers for all of his assistants in furthering my understanding of the biochemical world. Special thanks to Dr. Banerjee my advisor for my undergrad. Special thanks to Dr. Holly for the understanding given to my often overly curious nature. Finally special thanks to Melissa Boersma of Auburn University COSAM - Chemistry and Biochemistry Director, Mass Spectrometry, for her assistance in analyzing mass spectra data.

A. Overview	8
B. Plasmid Selection	8
C. Gene Construction	9
D. Selection of Expression & Amplification Systems	11
E. Plasmid Amplification	11
F. Plasmid Digestion	12
G. Isolation of the Linearized Plasmid	12
H. In Fusion Cloning	13
I. Transformation of <i>E. coli</i> (DH5 α)	13
J. Isolation of Mutated Plasmids	14
K. Determination of Plasmid Purity and Concentration	14
L. DNA Sequencing	14
M. Transformation of <i>E. coli</i> BL21	14
N. Protein Expression	15
O. Nickel Affinity Chromatography	15
P. Purification: Sephadex G 25 Size Exclusion Chromatography	16
Q. Expression of TEV Protease	17
R. Removal of GFP Fusion Tag	17

TABLE OF CONTENTS

2. Resuspension of Precipitate	18
3. SDS-PAGE	18
RESULTS DISCUSSION	20
ACKNOWLEDGMENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
I. INTRODUCTION AND BACKGROUND INFORMATION	1
A. Diabetes and Cancer	2
B. Preptin	3
C. Alanine Scanning Mutagenesis	5
II. METHODS AND MATERIALS	7
A. Overview	8
B. Plasmid Selection	8
C. Gene Construction	9
D. Selection of Expression & Amplification Systems	11
E. Plasmid Amplification	11
F. Plasmid Digestion	12
G. Isolation of the Linearized Plasmid	12
H. In Fusion Cloning	13
I. Transformation of <i>E. coli</i> (DH5 α)	13
J. Isolation of Mutated Plasmids	14
K. Determination of Plasmid Purity and Concentration	14
L. DNA Sequencing	14
M. Transformation of <i>E. coli</i> BL21	14
N. Protein Expression	15
O. Nickel Affinity Chromatography	15
P. Purification: Sephadex G 25 Size Exclusion Chromatography	16
Q. Expression of TEV Protease	17
R. Removal of GFP Fusion Tag	17

S.	Resuspension of Precipitate from Cleavage Reaction	18
T.	SDS-PAGE	18
III.	RESULTS DISCUSSION.....	20
A.	Rationale	21
B.	Expression of Native Preptin without GFP.....	21
C.	Introduction of GFP	22
D.	Cloning of Fusion Protein.....	22
E.	Sequencing of Plasmids	22
F.	Transformation of E. coli (B121)	25
G.	Expression of Fusion.....	25
H.	Nickel Affinity Chromatography	25
I.	Tobacco Etch Virus Protease Cleavage Reaction.....	27
J.	Analysis of the Cleavage Reaction Precipitate.	27
K.	Mass Spectrometry Analysis of TEV Reaction.	31
1.	Precipitate from TEV Cleavage of GFP-preptin-Ala ²¹ in Basic Solution.....	31
2.	Supernatant of GFP-preptin-Ala ²¹ TEV protease reaction supernatant.	32
3.	Supernatant of GFP-preptin-Ala ²³ TEV protease reaction supernatant.	33
IV.	CONCLUSIONS AND FUTURE WORK	45
A.	Conclusions.....	46
B.	Future Work.....	46
V.	BIBLIOGRAPHY	47

LIST OF TABLES

Figure 1	Preptin Sequence	2
Table 1	Nickel Affinity Chromatography Purification Buffers.....	21
Figure 2	Synthetic Plasmid	7
Table 2	Masses of Preptin Target & Fusion Proteins.....	30
Figure 3	DNA Insert/Plasmid Complementarity	9
Figure 4	Sequencing results for the mutated GFP-Preptin Ala21 gene	22
Figure 5	Sequencing results for the mutated GFP-Preptin Ala22 gene	22
Figure 6	Sequencing results for the mutated GFP-Preptin Ala23 gene	23
Figure 7	Sequencing results for the mutated GFP-Preptin Ala24 gene	23
Figure 8	Affinity Chromatography Isolation of GFP-Preptin Ala21 Fusion.....	24
Figure 9	Affinity Chromatography Isolation of GFP-Preptin Ala23 Fusion and Post Cleavage Analysis	25
Figure 10	Analysis of TEV Reaction Precipitant by PAGE on 12% gel.....	26
Figure 11	Analysis of TEV Reaction Precipitant by PAGE on a 4-20% gradient gel.....	26
Figure 12	Analysis of second TEV reaction precipitant by PAGE on a 4-20% gradient gel	27
Figure 13	Analysis of precipitant from the TEV reaction in acid an base solutions.....	28
Figure 14	Chromatogram of PB(isocratic elution)	29
Figure 15	Spectrum of polypeptides eluted from the chromatography column at 0.265 minutes from the PB sample.....	29

LIST OF FIGURES

Figure 16	Spectrum of polypeptides eluted from the chromatography column at 3.494 minutes from PB sample.....	30
Figure 1	Preptin Sequence	3
Figure 2	Synthetic Plasmid	7
Figure 3	DNA Insert/Plasmid Complementarity	9
Figure 4	Sequencing results for the mutated GFP-Preptin Ala21 gene	22
Figure 5	Sequencing results for the mutated GFP-Preptin Ala22 gene	22
Figure 6	Sequencing results for the mutated GFP-Preptin Ala23 gene	23
Figure 7	Sequencing results for the mutated GFP-Preptin Ala24 gene.....	23
Figure 8	Affinity Chromatography Isolation of GFP-Preptin Ala21 Fusion.....	24
Figure 9	Affinity Chromatography Isolation of GFP-Preptin Ala23 Fusion and Post Cleavage Analysis	25
Figure 10	Analysis of TEV Reaction Precipitant by PAGE on 12% gel.....	26
Figure 11	Analysis of TEV Reaction Precipitant by PAGE on a 4-20% gradient gel.....	26
Figure 12	Analysis of second TEV reaction precipitant by PAGE on a 4-20% gradient gel.....	27
Figure 13	Analysis of precipitant from the TEV reaction in acid and base solutions	28
Figure 14	Chromatogram of PB(isocratic elution)	29
Figure 15	Spectrum of polypeptides eluted from the chromatography column at 0.265 minutes from the PB sample.....	29

Figure 16 Spectrum of polypeptides eluted from the chromatography column at 5.484 minutes from PB sample.....	30
Figure 17 Spectrum of polypeptides eluted from the chromatography column at 12.785 minutes from the reaction of GFP-preptin Ala ²¹ with TEV protease	36
Figure 18 Chromatogram of PB without PEG (gradient elution	37
Figure 19 Mass Spectrum at 5 mins for PB w/o PEG.....	38
Figure20 Mass Spectrum at 6.953 mins for PB w/o PEG.....	39
Figure 21 Chromatogram of GFP-preptin-Ala ²¹ TEV cleavage reaction supernatant.....	40
Figure 22 Mass Spectrum at 6.065 mins for TEV cleavage reaction supernatant.....	41
Figure 23 Chromatogram of GFP-preptin-Ala ²³ fusion sample before TEV cleavage.....	42
Figure 24 Chromatogram of GFP-preptin-Ala ²³ TEV cleavage reaction supernatant.....	43

Design, Cloning, and Expression of Rat Preptin and Alanine Analogs

A thesis submitted to the College of Letters and Sciences in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

DEPARTMENT OF CHEMISTRY

by

Jahbo M. Love

2017

Dr. Jonathan M. Meyers, Chair

Date

Dr. Anil C. Banerjee, Member

Date

Dr. D. Wade Holley, Member

Date

A. Diabetes and Cancer **I. Introduction and Background Information**

Diabetes and cancer are two leading causes of death in the United States, and recent epidemiological data suggest that people with diabetes are at an elevated risk for most cancers. When viewed separately, these diseases have a devastating effect on the quality of life of and place a huge economic burden on the American population. Diabetes was ranked as the seventh leading cause of American deaths in 2016. It was listed as the underlying cause of death for 69,071 Americans and was mentioned as a contributing factor for an additional 234,051 deaths.¹ In addition, diabetes carries with it the risk for a host of complications including hypertension, dyslipidemia, heart attack, stroke, blindness, kidney disease, and amputations.¹ While diabetes can be a costly and potentially deadly disease, many of the acute symptoms and effects can be alleviated by current therapies. In contrast, cancer remains a devastating disease and was responsible for more than 595,000 deaths in 2016. This, coupled with the cost associated with treating the more than 1.6 million new cases in 2016, highlights the extreme burden these two diseases place on our population.

Diabetes Mellitus is a group of diseases characterized by elevated blood glucose levels resulting from a range of defects in the insulin signaling pathway. The two most common forms of diabetes are Type I and Type II. Type I diabetes is responsible for about 5% of total cases diagnosed and is characterized by defects in insulin production so that the body cannot meet glycemic demand.^{2,3} In contrast, type II diabetes occurs when the body becomes "insulin resistant" resulting in an abnormally high blood sugar concentration.⁴ As the disease progresses, the pancreas produces excessive insulin in an attempt to combat the rising blood sugar levels. However, the pancreas is not able to keep up and may eventually cease producing insulin entirely.⁴

A. Diabetes and Cancer

Diabetes and cancer are two leading causes of death in the United States, and recent epidemiological data suggest that people with diabetes are at an elevated risk for most cancers. When viewed separately, these diseases have a devastating effect on the quality of life of and place a huge economic burden on the American population. Diabetes was ranked as the seventh leading cause of American deaths in 2010. It was listed as the underlying cause of death for 69,071 Americans and was mentioned as a contributing factor for an additional 234,051 deaths.¹ In addition, diabetes carries with it the risk for a host of complications including hypertension, dyslipidemia, heart attack, stroke, blindness, kidney disease, and amputations.¹ While diabetes can be a costly and potentially deadly disease, many of the acute symptoms and effects can be alleviated by current therapies. In contrast, cancer remains a devastating disease and was responsible for more than 595,000 deaths in 2016. This, coupled with the cost associated with treating the more than 1.6 million new cases in 2016, highlights the extreme burden these two diseases place on our population.

Diabetes Mellitus is a group of diseases characterized by elevated blood glucose levels resulting from a range of defects in the insulin signaling pathway. The two most common forms of diabetes are Type I and Type II. Type I diabetes is responsible for about 5% of total cases diagnosed and is characterized by defects in insulin production so that the body cannot meet glycemic demand.^{2,3} In contrast, type II diabetes occurs when the body becomes “insulin resistant” resulting in an abnormally high blood sugar concentration.⁴ As the disease progresses, the pancreas produces excessive insulin in an attempt to combat the rising blood sugar levels. However, the pancreas is not able to keep up and may eventually cease producing insulin entirely.⁴

Much like diabetes mellitus, cancer is a broad term for many specific diseases which share characteristics. The most prominent of these characteristics is the uncontrolled growth and spread of cells resulting from both internal and external factors. These factors can act either together or sequentially, and the instigating event may occur years before the detection of symptoms.^{3,5}

Although the link between diabetes and cancer has not been fully elucidated, there is a large body of epidemiological data which would suggest a relationship between the two disorders.⁶⁻⁹ Diabetes has been associated with a higher risk for liver, pancreas, endometrium, colon/rectum, breast, and bladder cancers.⁷ The links between cancer and diabetes may also be a result of shared risk factors such as aging, diet, obesity, and physical inactivity.⁷ In addition to these shared risk factors, an increased risk of some cancers has been associated with diabetes therapies.⁷

For decades research has shown a strong association between obesity and type II diabetes. However, it was only in recent years that a link between obesity and cancer was more firmly established.^{7,8} Higher metabolic activity generally associated with excess caloric states such as those experienced during obesity can cause tumorigenesis particularly through the production of radical oxidative species.⁸ Further epidemiological evidence has shown links between obesity and liver, gallbladder, and pancreatic cancers.⁹

B. Preptin

With the prevalence of these diseases in today's life, it is important for us to understand the metabolic pathways involved in these processes. To this end, a better understanding of the structure-activity relationship was sought for the newly discovered hormone, Preptin, which has been shown to participate in both mitogenic and metabolic processes. Preptin is a 34-amino acid

peptide (3948 Da) corresponding to residues 69-102 of the proinsulin-like growth factor IIE-peptide (pro-IGF-IIE). Insulin-like growth factor II (IGF-II; also known as somatomedinA) is cleaved into three chains: IGF-II, IGF-II Ala-25 Del, and Preptin.²

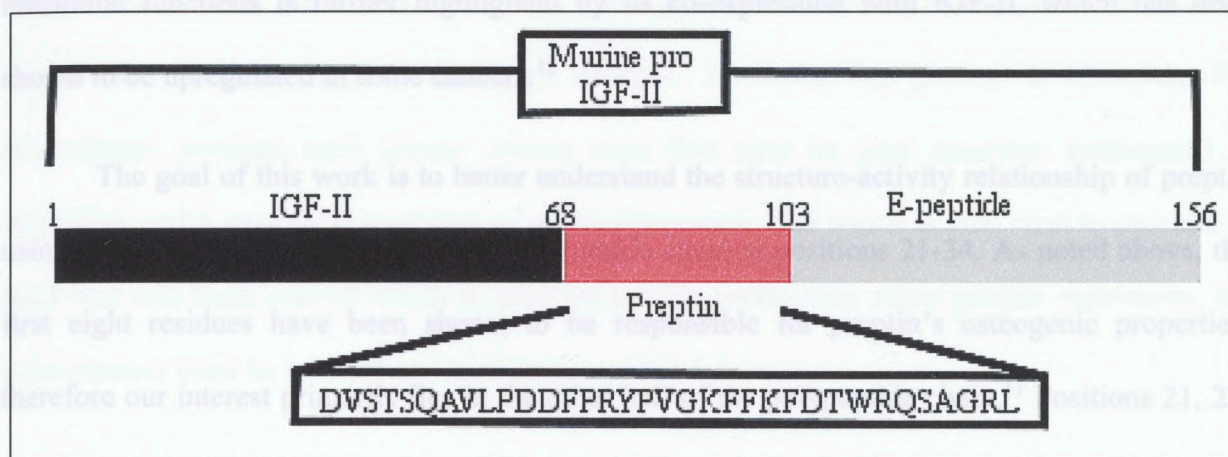


Figure 1: Preptin Sequence. The proinsulin-like growth factor which is cleaved into IGF-II, Preptin, and an E-peptide.

Preptin is primarily secreted by kidneys, liver, pancreas, salivary gland, and mammary tissue.¹⁰ Its main biological and psychological effects are an increase in cell differentiation and activity of osteoblasts, regeneration of bone density, limiting of glucose production in the liver, modulation of insulin sensitivity, and regulation of energy homeostasis.¹⁰ Preptin undergoes glucose-mediated co-secretion with and is a physiological amplifier of insulin.² It has been shown that synthetic preptin causes an increase in insulin secretion from glucose-stimulated β TC6-F7 cells in a saturable manner.²

In addition to its metabolic activity preptin also possesses osteogenic properties. Preptin's osteogenic properties were found to require only the first eight residues which promote increased bone growth by increasing the survival rate and proliferation of osteoblasts in both *in vitro* and *in*

in vivo studies.¹¹ An uncontrolled increase in osteoblast proliferation and survival over time could lead to a cancerous condition. In addition to its osteogenic properties, preptin has been implicated in the pathology of gestational diabetes. Its association with both osteogenic and metabolic functions is further highlighted by its co-expression with IGF-II, which has been shown to be upregulated in some cancers.¹²

The goal of this work is to better understand the structure-activity relationship of preptin using alanine scanning mutagenesis of the peptide chain at positions 21-24. As noted above, the first eight residues have been shown to be responsible for preptin's osteogenic properties, therefore our interest primarily lies in the other half of the polypeptide chain.¹¹ Positions 21, 22, and 24 are phenylalanine residues which possess large hydrophobic centers that could be used for interaction with a hydrophobic pocket on the target receptor. Position 23 is a lysine which has a long basic side chain that could have significant interaction with the receptor.

C. Alanine Scanning Mutagenesis

Protein libraries are powerful tools in determining the relationship between structure and function for proteins.¹⁴ Protein libraries are generated through mutagenesis of wild-type proteins at specific residues.¹¹ These mutations make it possible to probe the contributions of each individual amino acid to the function of the polypeptide.¹⁴ A special type of protein library, an Alanine scanning mutagenesis library, has been used to gain insight into important biological systems such as the binding of human growth hormone (hGH) to hGH-binding protein, CD4-binding to HIV-gp120, elucidation of the enzymatic activities of kinases, and determination of lysosome stability, just to name a few. Alanine scanning mutagenesis is a method in which alanine is systematically substituted for native residues in a step-wise manner.¹⁴ Alanine effectively allows for the removal of all side chain atoms past the beta carbon. Additionally,

alanine can be introduced without fear of inducing increased conformational flexibility in the protein backbone as it does not have a preference for unusual dihedral backbone angles. Replacing the native residue with alanine facilitates an inference into the role of the native side chain's effect on the activity of the protein. Through alanine scanning mutagenesis it is possible to provide a detailed map of functional epitopes. However, this process is laborious. For recombinant proteins, each alanine mutant must first have its gene sequence synthesized or amplified, and it must be cloned into an expression vector. The vector is then used to generate at least two cell lines, one of which is used for protein production. After protein expression, the gene product must be purified and sometimes refolded.

II. Methods and Materials

A. Overview

Recombinant DNA technology was employed to express preptin and preptin analogs in *E. coli*. Gene fragments encoding for preptin and its analogs were designed using a free translation program at ExPasy.org. Synthetic recombinant DNA was cloned into a commercially available vector and fusion proteins were isolated using affinity and size exclusion chromatography.

B. Plasmid Selection

A synthetic plasmid (pD444-SR) was purchased from ATUM and offered ampicillin resistance, a mutated pUCori replication site, lacI regulation gene, a T3 derived promoter, and a strong ribosomal binding site.¹⁵ The ampicillin resistance gene encodes for a beta-lactamase which is secreted into the periplasmic space. This beta-lactamase catalyzes the hydrolysis of the beta-lactam ring in ampicillin allowing for the selection of *E. coli* transformed with this plasmid as they exhibit ampicillin resistance of up to 100 µg/ml. The mutated pUCori replication site is derived from *E. coli* plasmid pBR322 and allows for greater than 500 copies of plasmid per cell during amplification. The lacI regulation gene codes for a repressor that binds tightly to a short DNA sequence just downstream of the beginning promoter in LacZ. This site is known as a lac operon and reduces transcription by interfering with the binding of RNA polymerase. The lac operon provides tighter repression. It also allows controlled expression by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the growth media. IPTG binds to the above mentioned repressor in the lac operon site, causing conformational changes in the repressor which results in its release. The strong ribosomal binding site (RBS) allows for the RNA translation into protein after transcription from DNA by RNA polymerase. These pre-incorporated tools made pD444-SR an excellent choice for our plasmid.

A. Overview

Recombinant DNA technology was employed to express preptin and preptin analogs in *E. coli*. Gene fragments encoding for preptin and its analogs were designed using a free translation program at ExPasy.org. Synthetic recombinant DNA was cloned into a commercially available vector and fusion proteins were isolated using affinity and size exclusion chromatography.

B. Plasmid Selection

A synthetic plasmid (pD444-SR) was purchased from ATUM and offered ampicillin resistance, a mutated pUCori replication site, lacI regulation gene, a T5 derived promoter, and a strong ribosomal binding site.¹⁵ The ampicillin resistance gene encodes for a beta-lactamase which is secreted into the periplasmic space. This beta-lactamase catalyzes the hydrolysis of the beta-lactam ring in ampicillin allowing for the selection of *E. coli* transformed with this plasmid as they exhibit ampicillin resistance of up to 100 µg/ml. The mutated pUCori replication site is derived from *E. coli* plasmid pBR322 and allows for greater than 500 copies of plasmid per cell during amplification. The lacI regulation gene codes for a repressor that binds tightly to a short DNA sequence just downstream of the beginning promoter in LacZ. This site is known as a lac operator and reduces transcription by interfering with the binding of RNA polymerase. The lac operon provides tighter repression. It also allows controlled expression by the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to the growth media. IPTG binds to the above-mentioned repressor in the lac operon site, causing conformational changes in the repressor which results in its release. The strong ribosomal binding site (RBS) allows for the RNA translation into protein after transcription from DNA by RNA polymerase. These pre-incorporated tools made pD444 – SR an excellent choice for our plasmid.

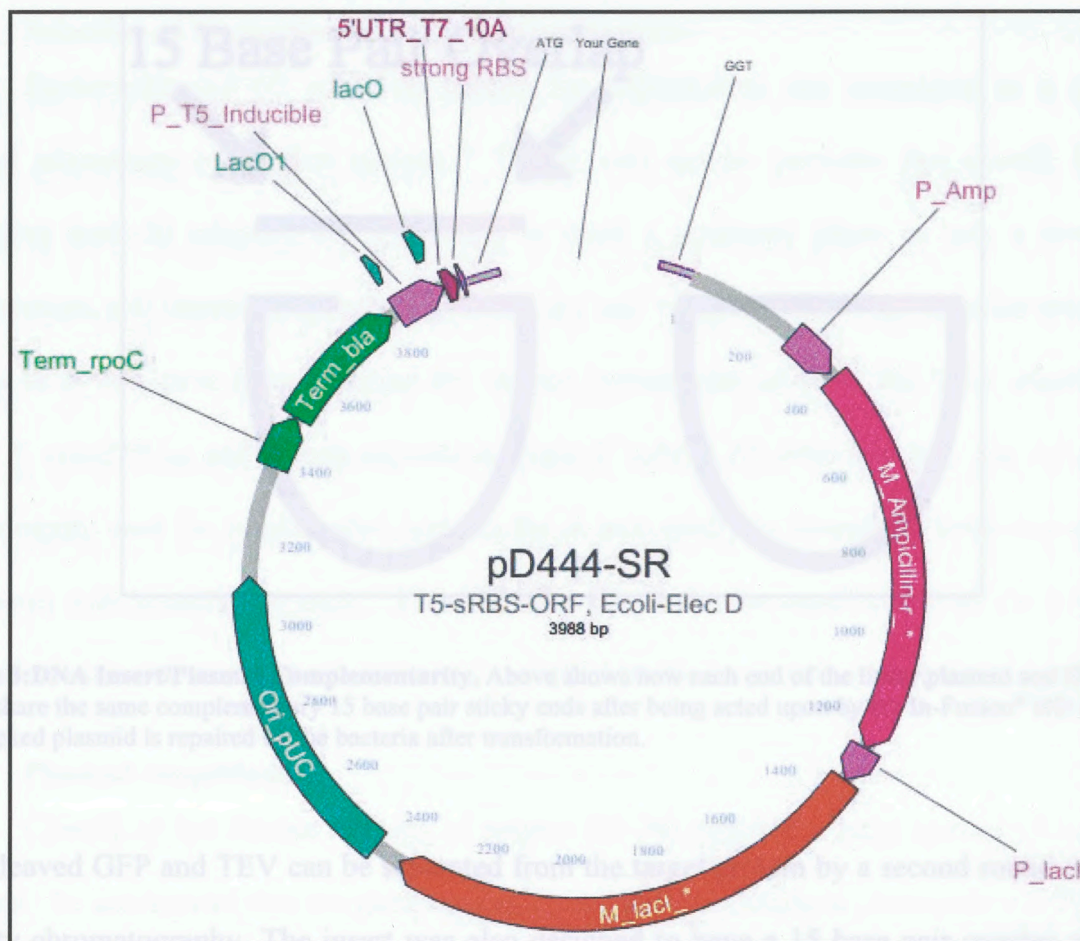


Figure 2: Synthetic Plasmid. The pD444-SR vector which offers Ampicillin resistance, a mutated pUCori replication site, lacI regulation gene, a T5 derived promoter, and a strong ribosomal binding site are indicated.

C. Gene Construction

A series of DNA inserts were designed to be incorporated into the synthetic plasmid. These inserts encode for an N-terminus six histidine nickel affinity tag, green fluorescent protein (GFP), Tobacco Etch Virus protease (TEV) cleavage site, and one of the four preptin alanine mutants. The six residue histidine tag provides the ability to isolate the fusion from the cell lysate using nickel affinity chromatography. GFP allows for real-time tracking of the fusion during chromatography. A TEV cleavage site between GFP and preptin allows for the removal of the fusion partner from the target protein.¹⁶

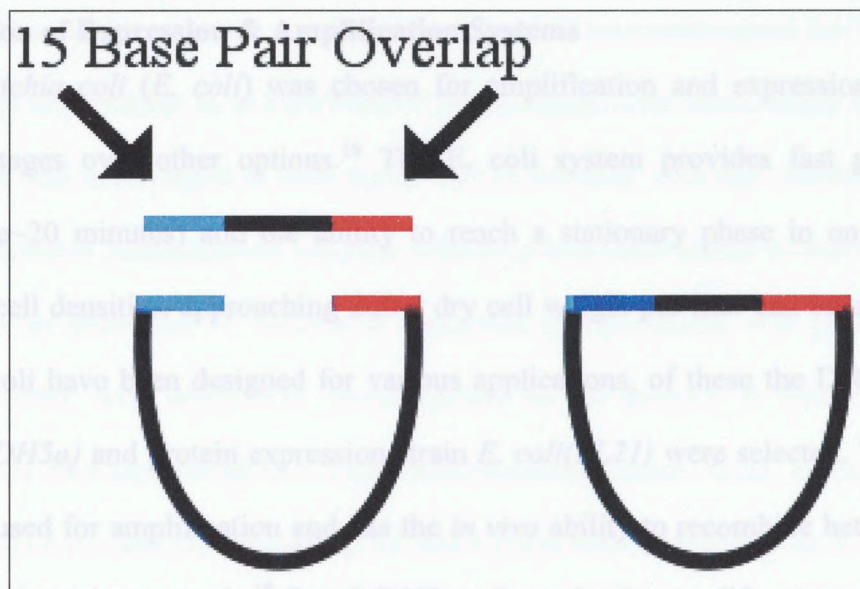


Figure 3: DNA Insert/Plasmid Complementarity. Above shows how each end of the linear plasmid and DNA insert share the same complementary 15 base pair sticky ends after being acted upon by the In-Fusion[®] HD enzyme. The nicked plasmid is repaired by the bacteria after transformation.

E. Plasmid Amplification

The cleaved GFP and TEV can be separated from the target protein by a second round of nickel affinity chromatography. The insert was also designed to have a 15 base pair overlap with the insertion site of the digested pD444-SR plasmid. These unique 15 base pair complementary sequences are required for use of the In-Fusion[®] HD Cloning Kit (Clontech). This cloning kit uses a proprietary enzyme which degrades a single strand of double-stranded linear DNA by 15 residues allowing for interaction between built-in complementary ends of the properly oriented open plasmid and inserts.¹⁷ The 15 base pair interaction is enough to hold the plasmid and DNA insert together during transformation into competent cells. Once transformed the cell repairs the sugar-phosphate backbone closing the circular plasmid with the insert incorporated.

D. Selection of Expression & Amplification Systems

Escherichia coli (*E. coli*) was chosen for amplification and expression as it provides several advantages over other options.¹⁸ The *E. coli* system provides fast growth kinetics, (doubling time~20 minutes) and the ability to reach a stationary phase in only a few hours. Furthermore, cell densities, approaching 200 g dry cell weight per liter can be achieved.¹⁸ Many strains of *E. coli* have been designed for various applications, of these the DNA amplification strain *E. coli*(*DH5 α*) and protein expression strain *E. coli*(*BL21*) were selected. The *DH5 α* strain is commonly used for amplification and has the *in vivo* ability to recombine heterologous DNA fragments with homologous ends.¹⁹ *E. coli*(*DH5 α*) allows for the amplification of plasmid DNA. Amplified plasmids were introduced into *E. coli*(*BL21*) cells for protein expression.

E. Plasmid Amplification

Cloning of the alanine mutants of preptin (21-24) required a large amount of the linear plasmid. To accomplish this the parent plasmid was used to transform chemically competent *E. coli*(*DH5 α*) cells. Transformation and growth conditions are described in following sections. Plasmid extraction was accomplished by use of the Promega Pureyield Miniprep kit. An alternative protocol was used, which allowed for isolation of plasmid DNA from larger culture volumes. This was necessary as the normal protocol did not yield sufficient DNA to support automated sequencing. An aliquot (1.5 mL) of an overnight culture was placed into a microcentrifuge tube and centrifuged for 30 seconds to collect the cell pellet. The supernatant was then discarded, and an additional aliquot (1.5 mL) of culture was added to the microcentrifuge tube. The sample was centrifuged under the same conditions, with subsequent discarding of the supernatant. The resulting cell pellet was then thoroughly resuspended in 600 μ L of TE buffer, before adding 100 μ L of cell lysis buffer (blue) and mixed by inverting the tube six times. 350 μ L of cold (4–8°C) neutralization solution was then added and mixed thoroughly

by inversion. The resulting sample was then centrifuged at maximum speed for 3 minutes before transferring the supernatant to a PureYield™ Minicolumn. The minicolumn was placed in a collection tube and centrifuged at max speed for 15 seconds with the resulting flow-through discarded. Endotoxin neutralization solution (200 μ L) was then added to the minicolumn and centrifuged for an additional 15 seconds at maximum speed. Finally, 400 μ L of column wash solution was added with subsequent centrifugation for 30 seconds at maximum speed. The minicolumn was then transferred to a clean microcentrifuge tube; 30 μ L of elution buffer was added and allowed to stand on the column for one minute before centrifuging at maximum speed for 15 seconds eluting the plasmid.

F. Plasmid Digestion

The amplified plasmid sample was then digested with the enzymes XMal and BsrGI-HF (New England Biolabs, NEB). The enzymes were introduced simultaneously to the sample containing the plasmid in CutSmart buffer (NEB), followed by incubation for two hours at 37°C. The enzymes were then deactivated via incubation at 80°C for 30 minutes.

G. Isolation of the Linearized Plasmid

Agarose gel electrophoresis was used to confirm complete digestion at the correct restriction sites. It is a commonly used method for DNA separation and is achieved by the aid of an electric field. Negatively charged DNA molecules migrate away from the cathode (negative pole) and towards the anode (positive pole). DNA fragment migration is determined solely by molecular weight, with smaller DNA fragments moving faster down the gel than larger ones. This allows for both estimates to be made of the size of DNA fragments with respect to DNA ladders, and to isolate specific bands by removing them from the gel. These estimates can be compared against calculated fragment size to assess endonuclease digestions. Moreover, this

method can be used to show the presence of DNA and protein contaminants in isolated recombinant plasmid samples, or lack thereof. After digestion, the linearized plasmid was separated on a 0.8% agarose gel (Invitrogen) with 0.5 $\mu\text{g}/\text{mL}$ EtBr. The gel was run in TAE buffer with 0.5 $\mu\text{g}/\text{mL}$ EtBr at 130 volts. DNA was collected from the gel using the PureLink™ Gel Extraction Kit.

H. In Fusion Cloning

The In-Fusion® HD cloning kit (Clontech) was employed to introduce the synthetic gene construct into the digested plasmid. 25 μL Linear plasmid was mixed with 1 μL of the synthetic insert in the presence of 8 μL In-Fusion® HD enzyme and 6 μL double distilled water. This sample was then allowed to incubate for 15 minutes at 50°C. The In-Fusion® reaction mixture is then used immediately for transformation.

I. Transformation of *E. coli* (DH5 α)

E. coli(DH5 α) cells were transformed directly with the In-Fusion® reaction mixture. Competent *E. coli*(DH5 α) cells were first thawed on ice before mixing gently. Competent cells (50 μL) were transferred into a falcon tube, followed by addition of 2.5 μL in-fusion reaction mixture produced above. This solution was placed on ice for 30 minutes before being heat shocked at 42°C for 45 seconds. The sample was then placed on ice for one to two minutes before the addition of 447.5 μL of warm at 37°C SOC medium. The sample was then incubated for an hour with shaking at 37°C. The transformation reaction was then diluted by 1/5 and 100 μL of the solution was spread on an ampicillin LB-agar plate spiked with IPTG and incubated overnight. The following day, a colony that fluoresced under handheld long wave UV light was selected and used to inoculate a 4 mL culture in the presence of ampicillin. Once the growth reached exponential phase, the culture was used to make a frozen glycerol stock.

J. Isolation of Mutated Plasmids

The mutated plasmids were isolated from *E. coli*(DH5 α) cells as described above.

K. Determination of Plasmid Purity and Concentration

Concentration and purity of the isolated plasmid samples were assessed using A260/A280 ratio on a NanoDrop™ 2000 spectrophotometer (ThermoFisher Scientific). The A260/A280 Ratio refers to a method where spectrophotometric absorbance of a sample is taken at both 260 nm and 280 nm and used to relate the concentrations of nucleic acids and proteins in a sample.²⁰The extinction coefficients for a 1 μ g/ml nucleic acid sample should be 20 and 10 for 260 nm and 280 nm, respectively.²⁰ This is in contrast to a 1 μ g/ml sample of protein, the trend is reversed and the extinction coefficients are 0.57 and 1.0 for 260 nm and 280 nm, respectively.²⁰ These values suggest that a pure sample of nucleic acid would have an A260/A280 Ratio of 2.0, whereas a pure protein sample would have a ratio of 0.57, and a mixture of the samples would have a value in between.

L. DNA Sequencing

Each of the mutated plasmids was sequenced by the Genomics & Sequencing Lab in the Department of Entomology and Plant Pathology, College of Agriculture at Auburn University.

M. Transformation of *E. coli* BL21

E. coli(BL21) cells were transformed with the purified plasmid from the *E. coli*(DH5 α) miniprep using the same method as described previously. The transformation reaction was spread on an ampicillin LB-agar plate spiked with IPTG and incubated overnight. The following day, a colony that fluoresced under handheld long wave UV light was selected and used to inoculate a 4 mL culture in the presence of ampicillin. Once the growth reached exponential phase, the culture was used to make a frozen glycerol stock.

N. Protein Expression

The newly established *E. coli* (BL21) cell lines were used to express the GFP-preptin fusions. A sample of the frozen stock was used to inoculate an overnight starter culture (incubated at 37°C, with shaking). The following morning, the starter culture was used to inoculate a 1 L culture which was incubated at 37°C with shaking until an optical density at 600 nm (OD₆₀₀) of between 0.6 and 0.8 was achieved. The culture was cooled to room temperature and 1 mL IPTG was added. The culture was incubated at 18°C overnight with shaking. The following day the cells were collected via centrifugation using the GSA rotor for 15 minutes at 5000 RPMs. The resulting pellet was resuspended in NPI-0 (recipe in Table 1) before being lysed. A sonicator was used, with a 2 second on 1-second rest cycle for 9 minutes while the sample was in an ice bath to lyse the cells. The resulting solution was centrifuged in a Sorval RC-5B instrument with an SS-34 rotor for 20 minutes at 15,000 RPMs and the supernatant collected.

O. Nickel Affinity Chromatography

Nickel affinity chromatography was used to separate the target protein from the unwanted native proteins, and to remove the digested fusion tag from the target protein in later steps. As mentioned previously the fusion protein was designed with a six-histidine affinity tag. This tag chelates with immobilized nickel on a gel support. Proteins immobilized by the nickel affinity chromatography can be later removed by the addition of imidazole or with solutions of lowering pH.²¹ In this way the target protein can be purified from all non-histidine tag proteins in a single step. This also allows for the later separation of the histidine-tagged fusion protein partner and digestion enzyme from preptin after digestion. Preptin was isolated from the supernatant using gravity flow over Ni-NTA Agarose resin (Qiagen). The column was washed with increasing

concentrations of imidazole (see table 1). The GFP fusion was eluted with NPI-500 until the absorbance at 280 nm was less than 0.05%.

Table 1: Nickel Affinity Chromatography Purification Buffers. Below is a list of the buffers used throughout this experiment during nickel affinity purification stage.

Buffer	[NaCl]	[NaH ₂ PO ₄]	[Tween 20]	[Imidazole]	pH
NPI-0	300 mM	50 mM	.05% (v/v)	0	pH 8
NPI-10	300 mM	50 mM	.05% (v/v)	10 mM	pH 8
NPI-20	300 mM	50 mM	.05% (v/v)	20 mM	pH 8
NPI-40	300 mM	50 mM	.05% (v/v)	40 mM	pH 8
NPI-500	300 mM	50 mM	.05% (v/v)	500 mM	pH 8
NPI-0 No PEG	300 mM	50 mM	0	0	pH 8
NPI-10 No PEG	300 mM	50 mM	0	10 mM	pH 8
NPI-20 No PEG	300 mM	50 mM	0	20 mM	pH 8
NPI-40 No PEG	300 mM	50 mM	0	40 mM	pH 8
NPI-500 No PEG	300 mM	50 mM	0	500 mM	pH 8

P. Purification: Sephadex G 25 Size Exclusion Chromatography

Although nickel affinity chromatography allows for the easy separation of the target protein from the rest of the native proteins found in *E. coli*. It introduces a large amount of imidazole and other salts to the sample. Thus a desalting step was required. Sephadex G 25 size exclusion chromatography provides a convenient desalting method. This method relies on porous beads of dextrose that construct two major spaces, the gaps between the beads (void volume) and the pores within them (pore volume).²² The void volume is significantly smaller than that of the pore volume and so if the size difference between the molecules that are to be separated is large, such that the smaller molecules can enter the pore volume. Then the larger molecules which cannot enter the pore volume will be eluted faster because their exclusion provides them a shorter volume in which to traverse. This allows for the quick and efficient separation of proteins from salts and small molecules such as Imidazole. Further, because of the incorporated GFP, tracking the progress of the fusion protein over the course of chromatography becomes as simple

as exposing fractions to UV light in order to see if they fluoresce. Samples which fluoresce can be pooled and later lyophilized. The preptin sample was then desalted via size exclusion chromatography on a Sephadex G-25 coarse resin (Sigma Aldrich) to facilitate a buffer exchange. Fractions were collected every thirty seconds and their absorbance at 280 nm measured; samples containing high absorbance were pooled and flash frozen in a dry ice/methanol mixture. The frozen sample was then lyophilized and stored at 4°C.

Q. Expression of TEV Protease

A starter culture of Tobacco etch virus protease (TEV) BL21-pnk793-000-009 (Addgene) was prepared by inoculating 100 mL of LB with 10 μ L ampicillin and 3 μ L chloramphenicol. The starter culture was incubated overnight at 32°C before being used to start a 1 L growth of LB inoculated with 100 μ L ampicillin and 30 μ L chloramphenicol. This growth was incubated until its OD 600 was between 0.6 and 0.8 after which 1 mL IPTG was added and the sample was allowed to incubate overnight at 32°C. The enzyme was purified and stored as described by Expression and Purification of Soluble His6 -Tagged TEV Protease.¹⁶

R. Removal of GFP Fusion Tag

The fusion protein was designed with a TEV restriction site between GFP and preptin. To carry out digestion 0.005 g TEV, 100 μ L of 10x urea/EDTA (4.0 M urea/.5 mM EDTA), and 100 μ L of 10x TECP (10 mM TECP) were added to 800 μ L of the purified fusion sample. The reaction was allowed to incubate at 4°C overnight. These samples were subjected to Polyacrylamide gel electrophoresis (PAGE), and electrospray time-of-flight mass spectrometry to test the validity of the digestion process.

S. Resuspension of Precipitate from Cleavage Reaction

Precipitation was observed during the cleavage reaction. Precipitant was separated from the supernatant and subsequently resuspended in a buffer of 10x urea/EDTA and split into two samples subjected to strongly basic and acidic conditions, respectively. A serial dilution of both samples was performed with one half, one fourth, and one-eighth concentrations respectively. All samples and the supernatant were subjected to standard workup procedures in preparation for PAGE.

T. SDS-PAGE

Similarly to the extraction of plasmid, a method of characterizing the purity and relative mass of the fusion protein samples was required. To achieve this sodium dodecyl sulfate (SDS) PAGE was employed. As with all electrophoresis SDS-PAGE takes advantage of the movement of negatively charged particles away from the cathode and towards the anode.²³ However, proteins have a much wider variety of structural configurations due to hydrophobic and hydrophilic interactions forming secondary tertiary and even quaternary structures. This hurdle can be overcome by use of the detergent sodium dodecyl sulfate (SDS) which when boiled in presence of the protein denatures them and imparts a uniform charge/mass ratio. This uniform charge mass ratio imparted by SDS is a result of its hydrophobic carbon chain permeating and interfering with the binding of hydrophobic groups causing the denaturation of the protein into a random coil coated with negatively charged detergent molecules along the length. This allows for the separation based on the mass of proteins found in a complex mixture. Utilizing all of these tools it will be possible to test if the *E. coli* amplification and expression system is capable of performing alanine scanning mutagenesis on preptin in preparation for further research into the activity of preptin. Although alanine scanning mutagenesis is a well-established method, and the use of *E. coli* expression systems is likewise well-established, these methods have never been

applied to nor used to express preptin or its alanine mutants. Therefore a determination on whether or not they are viable methods in the biosynthesis of preptin and preptin's alanine mutants is required.

III. Results Discussion

A. Rationale

In the light of a growing body of evidence linking diabetes and cancer, there is an increased demand to unravel the links between these two devastating diseases. With this in mind, the goal of this project was to elucidate the role of a prominent hydrophobic patch in the final third of the preproinsulin hormone by expressing and testing the alanine mutants. It is expected that the information gleaned from these novel analogs will be used to design two preproinsulin derived peptides that could elicit separate mitogenic and metabolic processes. Such peptides would provide powerful tools for exploring the individual pathways associated with bone growth and insulin signaling amplification. Additionally, these molecules could serve as starting points for potential treatments of osteoporosis and diabetes.

The recent discovery of preproinsulin limits the number of available assays for assessing activity. Our collaborators at the University of Auckland have established assays using both rat and mouse cell lines. To give us the most flexibility we chose to begin our work with the rat sequence since it displays activity in both cell lines (human and mouse sequences do not).

B. Expression of Native Preproinsulin without GFP

Initial attempts to produce preproinsulin were stymied during the purification phase due to a lack of access to a preparative scale HPLC or FPLC system and on-site mass spectrometer. Instead, gravity flow nickel affinity and size exclusion chromatography were used in conjunction with a standalone fraction collector and bench top UV-Vis spectrophotometer. This process proved to be time-consuming and correct identification of the preproinsulin containing samples delayed pending off-site mass spectral analysis.

A. Rationale

In the light of a growing body of evidence linking diabetes and cancer, there is an increased demand to unravel the links between these two devastating diseases. With this in mind, the goal of this project was to elucidate the role of a prominent hydrophobic patch in the final third of the preptin hormone by expressing and testing the alanine mutants. It is expected that the information gleaned from these novel analogs will be used to design two preptin derived peptides that could elicit separate mitogenic and metabolic processes. Such peptides would provide powerful tools for exploring the individual pathways associated with bone growth and insulin signaling amplification. Additionally, these molecules could serve as starting points for potential treatments of osteoporosis and diabetes.

The recent discovery of preptin limits the number of available assays for assessing activity. Our collaborators at the University of Auckland have established assays using both rat and mouse cell lines. To give us the most flexibility we chose to begin our work with the rat sequence since it displays activity in both cells lines (human and mouse sequences do not).

B. Expression of Native Preptin without GFP

Initial attempts to produce preptin were stymied during the purification phase due to a lack of access to a preparative scale HPCL or FPLC system and on-site mass spectrometer. Instead, gravity flow nickel affinity and size exclusion chromatography were used in conjunction with a standalone fraction collector and bench top UV-Vis spectrophotometer. This process proved to be time-consuming and correct identification of the preptin containing samples delayed pending off-site mass spectral analysis.

C. Introduction of GFP

In order to accelerate the purification steps and allow for initial assessment of purity to be carried out in-house, a designer gene fragment was designed that encoded for a removable green fluorescent protein tag on the N-terminus of the preptin mutants. The addition of the GFP tag to allowed purifications to be carried out without UV detectors. An inexpensive UV pen light (Bio-Rad) was used to track the fusion protein as it moved through the columns. This provided a means to identify the pertinent fractions for further use.

D. Cloning of Fusion Protein

The In-Fusion[®] HD cloning kit (Clontech) was designed to provide rapid cloning requiring fewer steps than traditional cloning methods. However, cloning of the synthetic gene fragments into the PD444-SR plasmid failed more than 50% of the time and required significant optimization before correctly transformed cells were identified. The most common result of the cloning reactions was the recovery of the original parent plasmid. Significant effort was applied to ensuring that the plasmid was completely digested and purified prior to the introduction of the synthetic fragment. In all, more than 10 cloning reactions were carried out at a cost of over \$500 before all four mutated plasmids were confirmed.

E. Sequencing of Plasmids

Successful cloning reactions were confirmed by sequencing of the amplified plasmids. Samples were submitted to the Genomics & Sequencing Lab in the Department of Entomology and Plant Pathology, College of Agriculture at Auburn University. Results were analyzed using the free sequencing software Chromas (Technelysium). The sequence encompassing the mutation site for each mutated gene can be found in Figures 4-7 which confirm the presence of alanine mutations at positions 21-24 respectively.

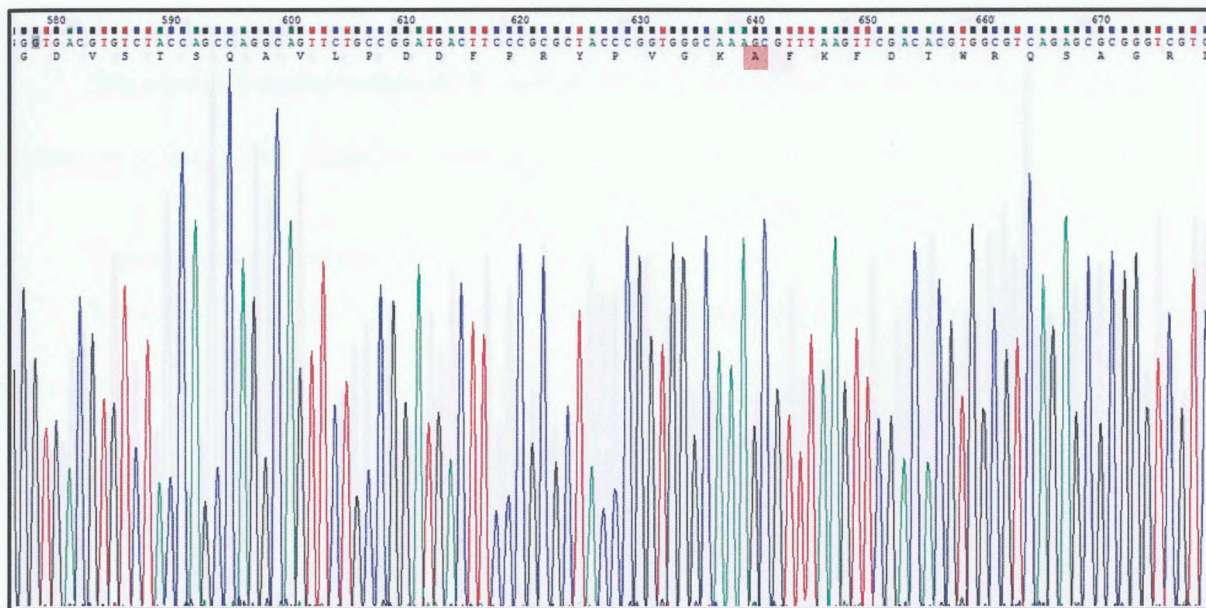


Figure 4: Sequencing results for the mutated GFP-Preptin Ala21 gene. Sequencing confirms mutation of the codon encoding for phenylalanine at peptide position 21 with the codon for alanine (red box). The reading frame for preptin is 577-679.

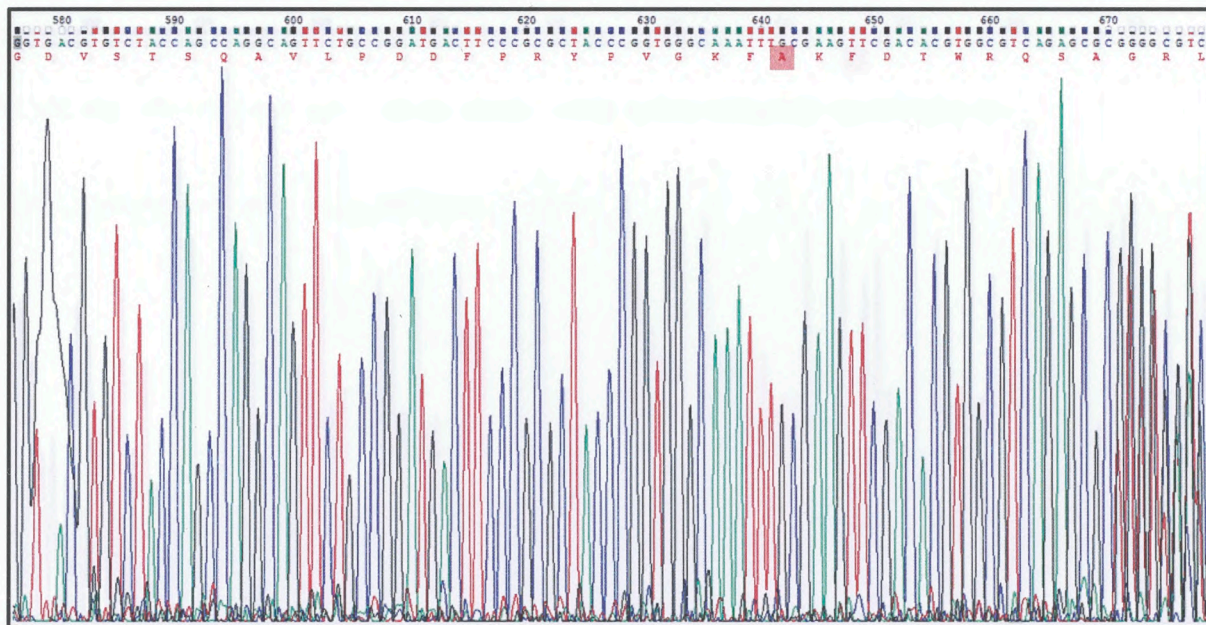


Figure 5: Sequencing results for the mutated GFP-Preptin Ala22 gene. Sequencing confirms mutation of the codon encoding for phenylalanine at peptide position 22 with the codon for alanine (red box). The reading frame for preptin is 576-678.

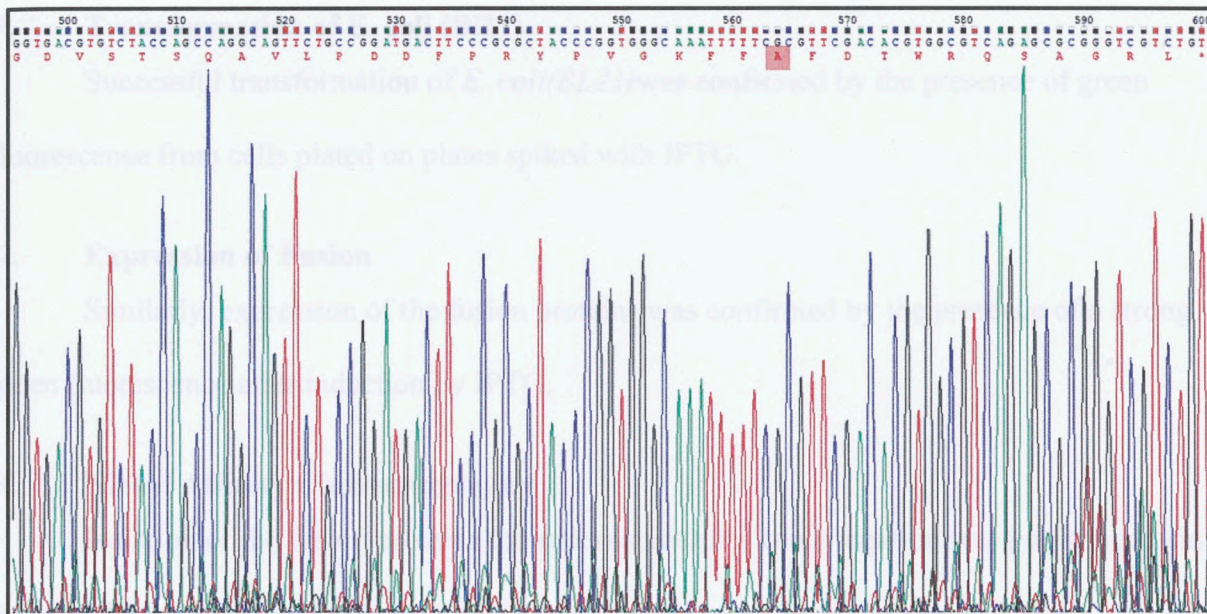


Figure 6: Sequencing results for the mutated GFP-Preptin Ala23 gene. Sequencing confirms mutation of the codon encoding for lysine at peptide position 23 with the codon for alanine (red box). The reading frame for preptin is 498-597.

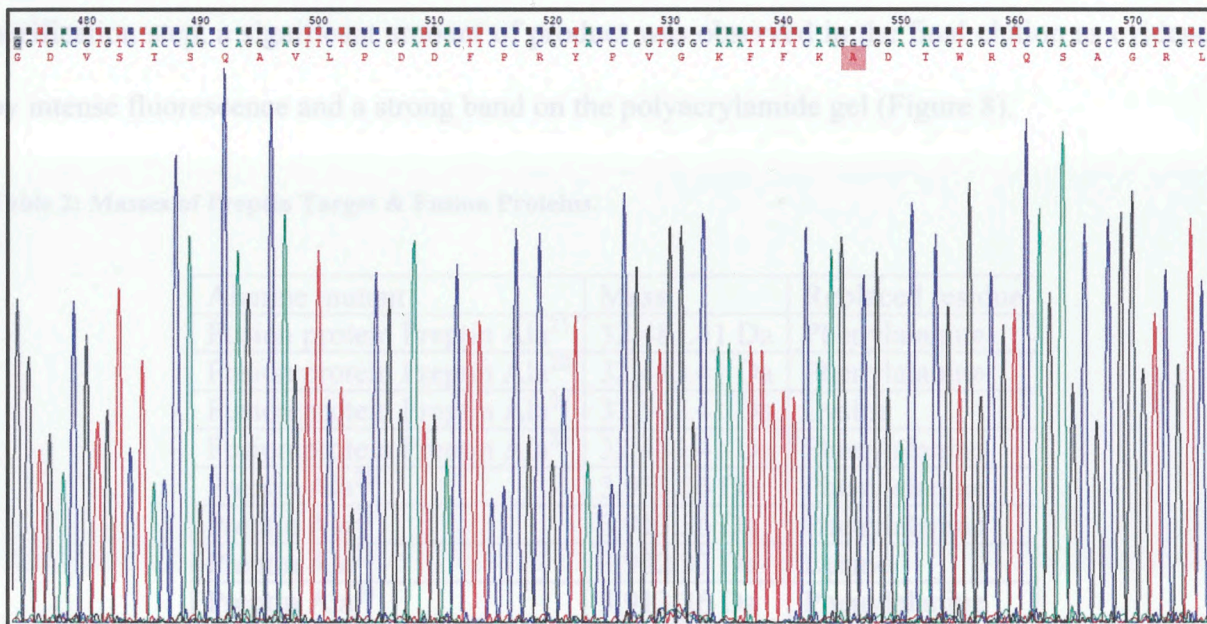


Figure 7: Sequencing results for the mutated GFP-Preptin Ala24 gene. Sequencing confirms mutation of the codon encoding for phenylalanine at peptide position 24 with the codon for alanine (red box). The reading frame for preptin is 474-576.

F. Transformation of *E. coli* (BL21)

Successful transformation of *E. coli*(BL21) was confirmed by the presence of green fluorescence from cells plated on plates spiked with IPTG.

G. Expression of Fusion

Similarly, expression of the fusion proteins was confirmed by the presence of a strong green fluorescence after induction by IPTG.

H. Nickel Affinity Chromatography

Isolation of the fusion proteins the supernatant after centrifugation was performed using a nickel affinity chromatography. Purification progress was monitored with PAGE. It is worth noting that even at the lowest concentrations of imidazole; fluorescence was observed in the eluent from the affinity column. It is postulated that this is a result of partial misfolding of the poly-histidine containing fusion tag. Despite the loss of some fusion during the affinity purification step, a significant amount of product was observed in the final elution as evidenced by intense fluorescence and a strong band on the polyacrylamide gel (Figure 8).

Table 2: Masses of Preptin Target & Fusion Proteins.

Alanine mutant	Mass	Replaced residue
Fusion protein Preptin Ala ²¹	32,483.41 Da	Phenylalanine
Fusion protein Preptin Ala ²²	32,483.41 Da	Phenylalanine
Fusion protein Preptin Ala ²³	32,502.41 Da	Lysine
Fusion protein Preptin Ala ²⁴	32,483.41 Da	Phenylalanine
Preptin Ala ²¹	3,913.36 Da	Phenylalanine
Preptin Ala ²²	3,913.36 Da	Phenylalanine
Preptin Ala ²³	3,932.37 Da	Lysine
Preptin Ala ²⁴	3,913.36 Da	Phenylalanine

Figure 3: Affinity Chromatography Isolation of GFP-Preptin Ala23 Fusion and Peptide Cleavage Analysis. A sample from the various column washes was analyzed by PAGE. An increase in the migration of the major band is evident after treatment with TBE (lane 2) with respect to the untreated sample (lane 4) indicating successful cleavage. However, a band corresponding to the preptin peptide (~4 kDa) was not observed. The remaining samples are column washes #4 through #8 (lanes 5-8 respectively). With the flow through from column loading (lane 9) and the supernatant after cell lysis and centrifugation (lane 10).

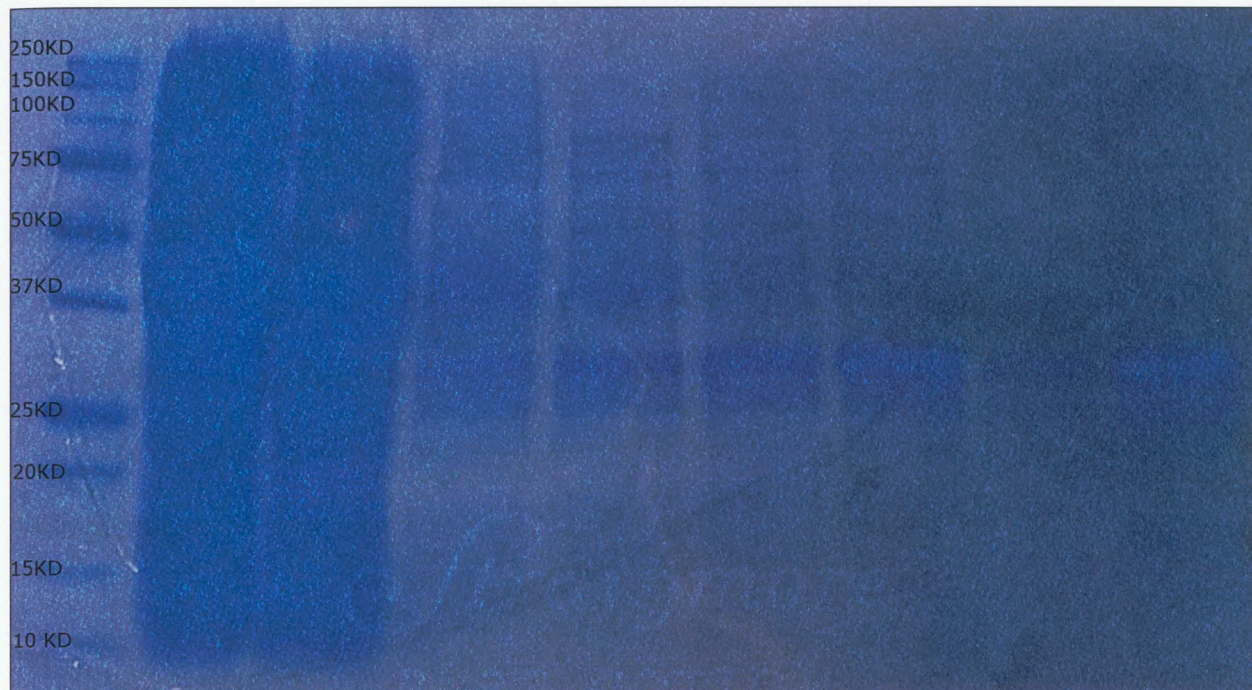


Figure 8: Affinity Chromatography Isolation of GFP-Preptin Ala21 Fusion. A sample from the various column washes was analyzed by PAGE. From left to right there is a decrease in extraneous proteins as we move to our purified product. (Lane 1) contains a ladder, (lane 2) lysis solution, (lane 3) lysis supernatant after centrifugation, (lane 4) eluent from wash one, (lane 5) eluent from wash two, (lane 6) eluent from wash three, (lane 7) eluent from wash four, and (lane 8) the purified fusion protein.

The samples were subsequently analyzed on a polyacrylamide 4-20% gradient gel made by Bio-Rad to resolve the bands found near the bottom



Figure 9: Affinity Chromatography Isolation of GFP-Preptin Ala23 Fusion and Post Cleavage Analysis. A sample from the various column washes was analyzed by PAGE. AN increase in the migration of the major band is evident after treatment with TEV (lane 2) with respect to the untreated sample (lane 4) indicating successful cleavage. However, a band corresponding to the preptin peptide (~4 kDa) was not observed. The remaining samples are column washes #4 through #1 (lanes 5-8 respectively). With the flow through from column loading (lane 9) and the supernatant after cell lysis and centrifugation (lane 10).

I. Tobacco Etch Virus Protease Cleavage Reaction

The purified fusion protein was subjected to treatment with TEV protease to cleave the fusion tag from the target peptide. Subsequent analysis by PAGE confirms a shift in the retention time of the GFP protein (Figure 9). However, the target peptide was not observed on the gel. It is worth noting that a significant amount of non-fluorescing precipitate was produced during the cleavage reaction. Subsequently, attempts to solubilize the precipitant in solutions covering a range of pH values with and without urea meet with minimal success. Samples which were reduced the amount of precipitate were analyzed by PAGE.

J. Analysis of the Cleavage Reaction Precipitate.

The precipitant observed after the reaction with TEV was analyzed to see if it contained our target peptide. When the solutions were run on a 12% polyacrylamide gel, a smear in the proximity of the 10kDa region was observed (Figure 10). However, that mass does not match the predicted mass of the target peptides. The samples were subsequently analyzed on a polyacrylamide 4-20% gradient gel made by Bio-Rad to resolve the bands found near the bottom of the 12% gel (Figure 11).

Figure 11: Analysis of TEV Reaction Precipitant by PAGE on a 4-20% gradient gel. A gradient gel was used in an attempt to visualize the protein peptide target (~4 kDa) and the fusion tag on the same gel. The unreacted GFP-fusion is seen in (lane 2), but no band is present in the supernatant of the TEV reaction mixture (lane 3) suggesting precipitation of GFP. Resuspension of the precipitant in an acidic solution restores a band at the mass expected for cleaved GFP and a dark smear centered around 15 kDa (lane 4). No bands corresponding to fusion or cleaved GFP are observed in the sample of precipitant resuspended in basic solution. However, a smear in the 10kDa range is observed (lane 5). Unfortunately, no evidence of the protein product was observed.

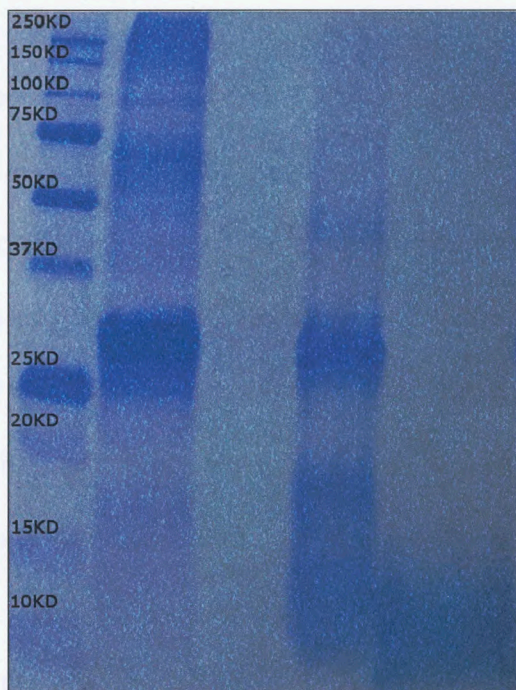


Figure 10: Analysis of TEV Reaction Precipitant by PAGE on 12% gel. The untreated GFP-Fusion is seen in (lane 2), but no band is present in the supernatant of the TEV reaction mixture (lane 3) suggesting precipitation of GFP. Resuspension of the precipitant in an acidic solution restores a band at the mass expected for cleaved GFP (lane 4). No bands corresponding to fusion or cleaved GFP are observed in the sample of precipitant resuspended in basic solution. However, a smear in the 10kDa range is observed (lane 5)

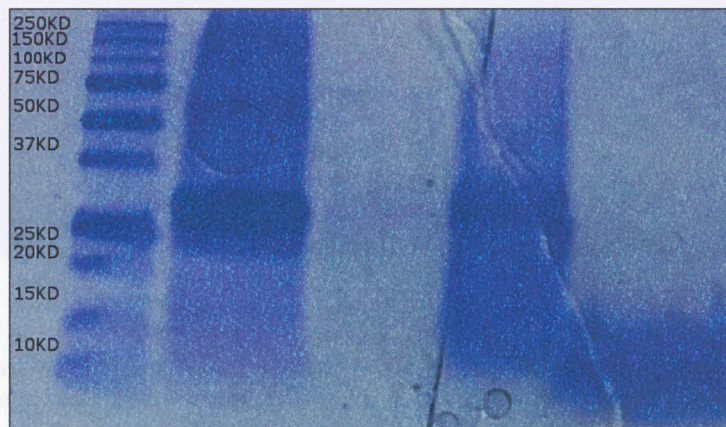


Figure 11: Analysis of TEV Reaction Precipitant by PAGE on a 4-20% gradient gel. A gradient gel was used in an attempt to visualize the preptin peptide target (~4 kDa) and the fusion tag on the same gel. The untreated GFP-Fusion is seen in (lane 2), but no band is present in the supernatant of the TEV reaction mixture (lane 3) suggesting precipitation of GFP. Resuspension of the precipitant in an acidic solution restores a band at the mass expected for cleaved GFP and a dark smear centered around 15 kDa (lane 4). No bands corresponding to fusion or cleaved GFP are observed in the sample of precipitant resuspended in basic solution. However, a smear in the 10kDa range is observed (lane 5). Unfortunately, a no evidence of the preptin product was observed.

From these two experiments, it was found that the vast majority of proteins had been precipitated out of solution after the TEV cleavage reaction (lane 3 in figure 10 and 11). When exposed to a low pH solution the majority of the proteins were resuspended; this includes what was presumed to be the fusion partner GFP. However, when exposed to a high pH buffer only proteins with masses below 15kDa were resuspended. This result was confounding, under both conditions, the precipitant had seemingly been dissolved, and yet the electrophoresis showed distinctly different proteins within the acid and base samples respectively. To test these incongruities an additional cleavage reaction on GFP-preptin Ala²¹fusion was performed and analyzed by PAGE on a 4-20% gradient gel (Figure 12).



Figure 12: Analysis of second TEV reaction precipitant by PAGE on a 4-20% gradient gel. Unlike the first two reactions, neither the fusion nor cleaved GFP was detected in any of the samples after reaction with TEV protease. Light bands in the regions of the smears in figures 10 and 11 were detected, but do not correspond to the expected mass for preptin.

Although a tight band below 10 kDa was observed, there was still a distinct absence of any larger proteins found in the samples tested. To test the validity of these results the digestion reaction on GFP-preptin Ala²¹ fusion was repeated with six times the volume. Because the precipitant was thought to contain the vast majority of the proteins formerly in solution, this increase in volume was intended to increase the concentration of proteins once redissolved. In addition to these modifications, the precipitate from the TEV reactions was dissolved in acidic and basic solutions, and the precipitant that did not dissolve in the acid solution was dissolved in a basic solution (Figure 13).

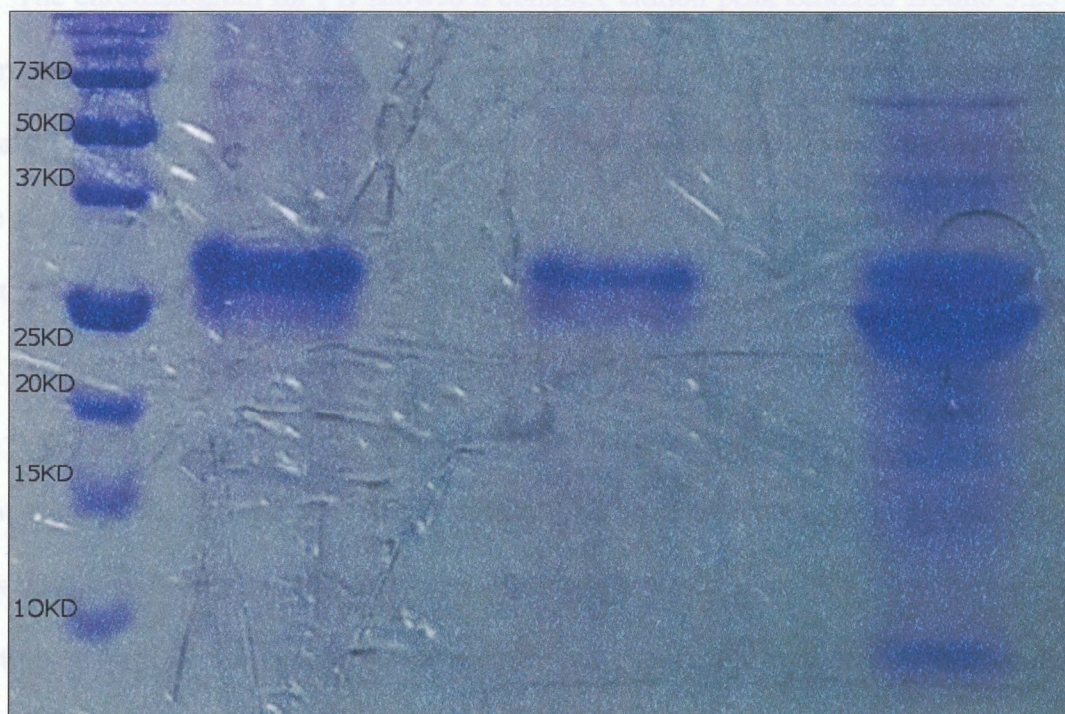


Figure 13: Analysis of precipitant from the TEV reaction in acid and base solutions. The GFP-preptin Ala²¹ fusion is in (lane 1). (Lane 3) is supernatant from the TEV protease reaction. Unlike previous attempts, GFP and/or fusion is observed in the supernatant. A faint band at approximately the mass of preptin was observed in the precipitant sample dissolved in basic solution.

K. Mass Spectrometry Analysis of TEV Reaction.

Samples from the three TEV cleavage reactions were tested using liquid chromatography time-of-flight electrospray mass spectrometry (LC-MS). Although both the raw and de-convoluted mass spectra are presented, the discussion will be restricted to the de-convoluted mass spectra when available.

1. Precipitate from TEV Cleavage of GFP-preptin-Ala²¹ in Basic Solution.

The absence of a preptin band in the TEV cleavage reaction supernatant coupled with the appearance of a low molecular weight smear when the precipitant was dissolved in a basic solution (PB) suggested that the sample should be analyzed by LC-MS.

The chromatogram for PB showed a complex elution peak centered around 5.25 minutes when running on an isocratic gradient of acetonitrile (Figure 14). The raw data was compared to the mass of preptin-Ala²¹, and the predicted degradation masses. This list included mass to charge ratios up to ten, and yielded 54 peaks of interest, with masses less than 2 Da from the predicted masses: 13 peaks at 0.265 minutes (Figure 15), 8 peaks at 5.484 minutes (Figure 16), and 21 peaks at 12.785 minutes (Figure 17). The elution peak at 0.265 minutes produced two m/z values corresponding to the fifth and sixth charge states of preptin which suggested that preptin did not stick to the column and eluted during loading. The remaining peaks had several m/z signals that could be attributed to enzymatic cleavage products of preptin. However, upon further inspection, it was determined that the sample had significant contamination of polyethylene glycol (PEG) that was carried over from the lysis buffer which contained Tween[®] 20.

The second round of expression and purification of GFP-preptin-Ala²¹ was carried out without the addition of Tween[®] 20. The precipitant from the TEV cleavage reaction (PB w/o PEG) was submitted for analysis by LC-MS and an acetonitrile gradient was used to elude the

polypeptides from the reverse phase column. This resulted in a shift of the majority of the signal to the column wash portion of the run (Figure 18). The chromatography peak at 5.751 minutes produced two m/z signals (14,308 Da and 28,616 Da) that could represent the +1 and +2 peaks for a mutated GFP (Figure 19). The cleaved GFP with the polyhistidine tag has an expected mass of 28,588.06 Da, or 28 Da less than observed. This discrepancy suggests a mutation in the GFP region of the gene.

More interestingly, two m/z signals (3,971 Da and 7939 Da) were observed at 6.593 minutes (Figure 20). The smaller signal is approximately 57 Da heavier than the expected mass of preptin-Ala²¹ (3,913.36 Da). This is suggestive of a mutation in which glycine is added somewhere on the preptin-Ala²¹ mutant. The larger peak could be the dimer of this glycine heavy mutant. However, there is no previous evidence that preptin dimerizes.

2. Supernatant of GFP-preptin-Ala²¹ TEV protease reaction supernatant.

The failure to observe the preptin analog in the PB samples prompted the analysis of the TEV cleavage reaction supernatant to determine if the target peptide was present but not observed by PAGE due to the small size. The supernatant from a fresh TEV cleavage reaction (without PEG) was analyzed by LC-MS using an acetonitrile gradient (Figure 21). Despite the presence of several peaks at low acetonitrile concentrations, no m/z values indicative of the preptin analog product were observed. A 32,461 Da peak was observed at 6.065 mins. (Figure 22), which is 22.4 Da less than the expected for the cleaved fusion tag.

Two masses that could be the cleaved fusion tag were observed in the PB and supernatant samples. However, they differ from the expected mass and do not agree with each other. The significance of these observations is unclear at this time. Additionally, the preptin analog was not identified in any of the samples.

3. Supernatant of GFP-preptin-Ala²³ TEV protease reaction supernatant.

Analysis of the GFP-preptin-Ala²¹ TEV protease reactions yielded an m/z signal that matched the target peptide mass plus 57 Da. It was unclear if the mass represented an unintentional preptin mutation or if it was an unidentified bacterial peptide. To differentiate between the two possibilities, the GFP-preptin-Ala²³ mutant was expressed, purified, and subjected to TEV protease. The K23A analog has a mass that is approximately 10 Da heavier than the phenylalanine to alanine analogs. If the previously observed mass of 3,913.36 Da was observed it would rule out a glycine addition and indicate the presence of an unidentified bacterial peptide.

The fusion protein sample before and after TEV cleavage was analyzed by LC-MS. The untreated sample showed two broad overlapping peaks centered around 9 mins (Figure 12). The chromatogram for the TEV reaction mixture (Figure 24) was markedly different than the one for the GFP-preptin-Ala²¹ cleavage reaction, but strikingly similar to the sample before reaction with TEV. While there were numerous difference in the m/z profiles for the untreated and treated samples, neither the 3,913.36 Da nor the preptin-Ala²³ signals were observed.

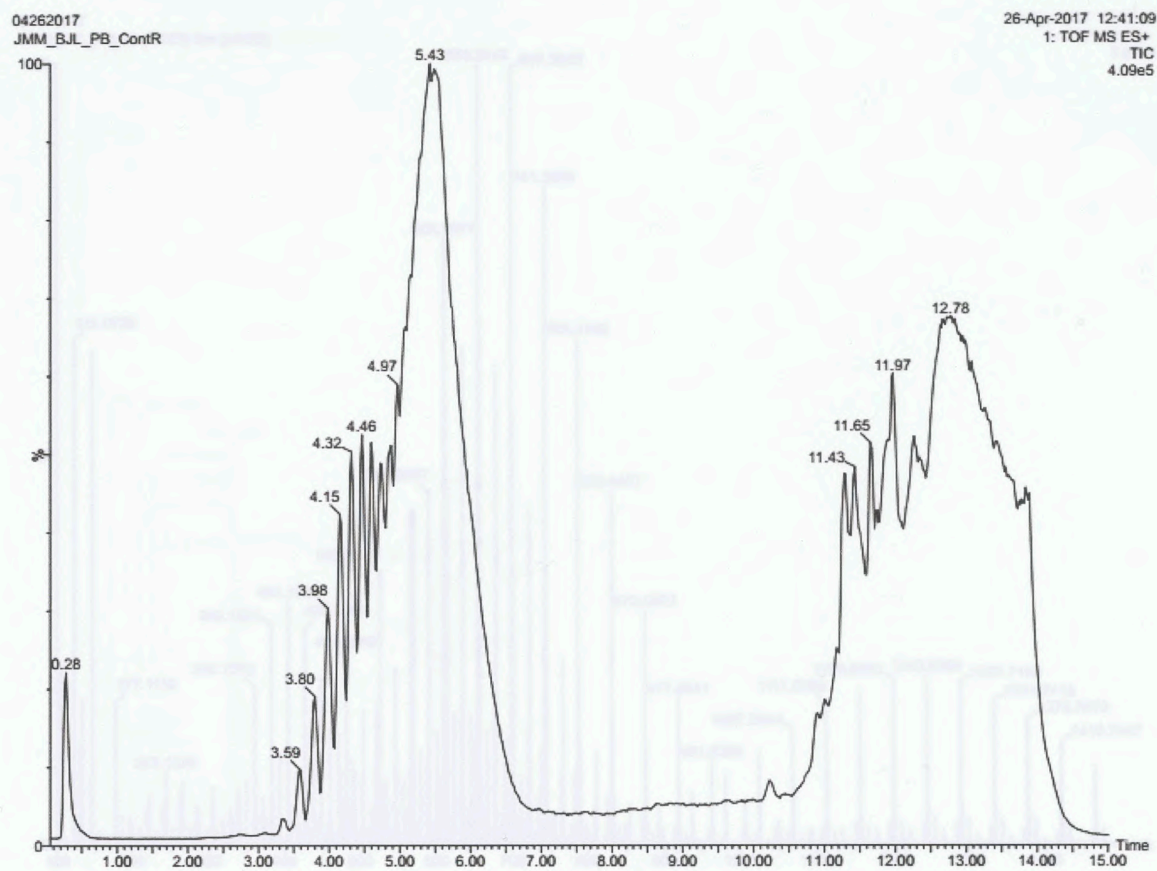


Figure 14: Chromatogram of PB(isocratic elution). A convoluted elution peak centered around 5.25 minutes is evidence of a complex protein mixture. The large signal seen starting in minute 11 are attributed to washing the column.

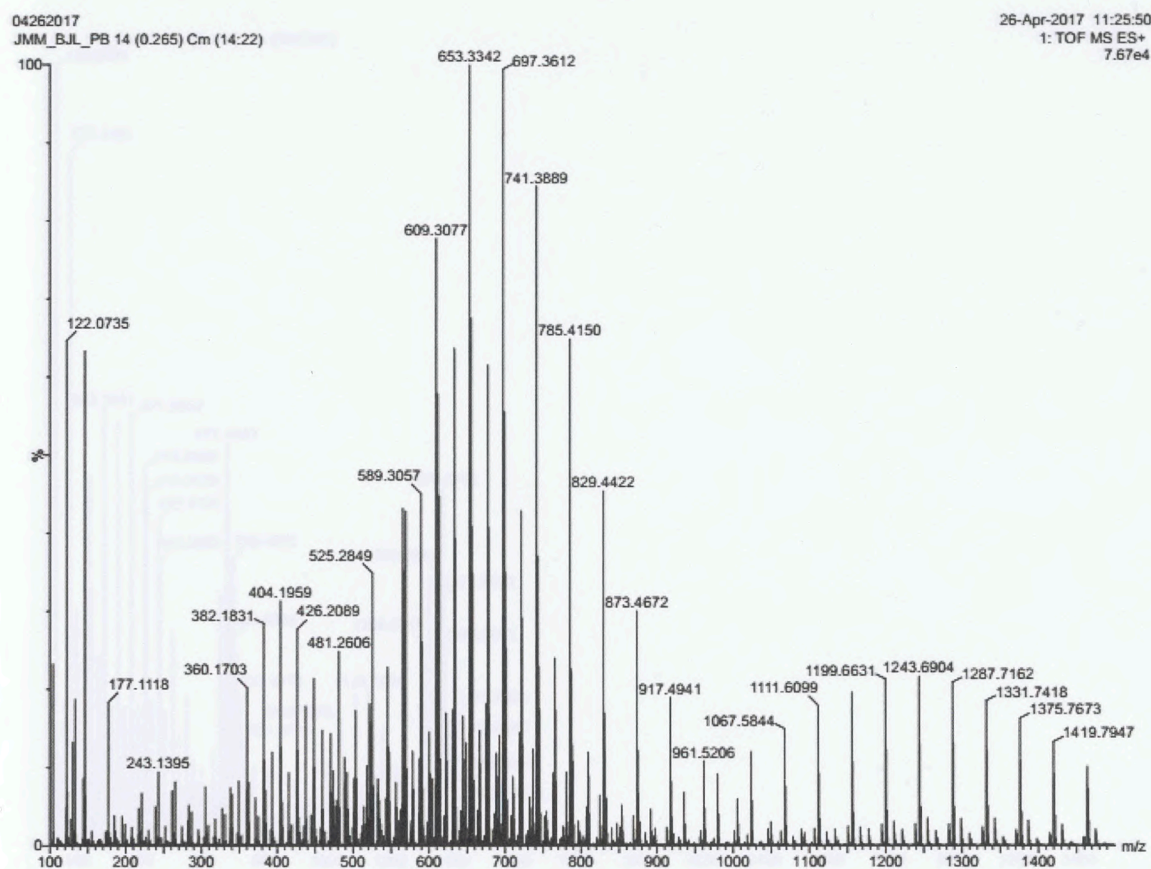


Figure 15: Spectrum of polypeptides eluted from the chromatography column at 0.265 minutes from the PB sample.

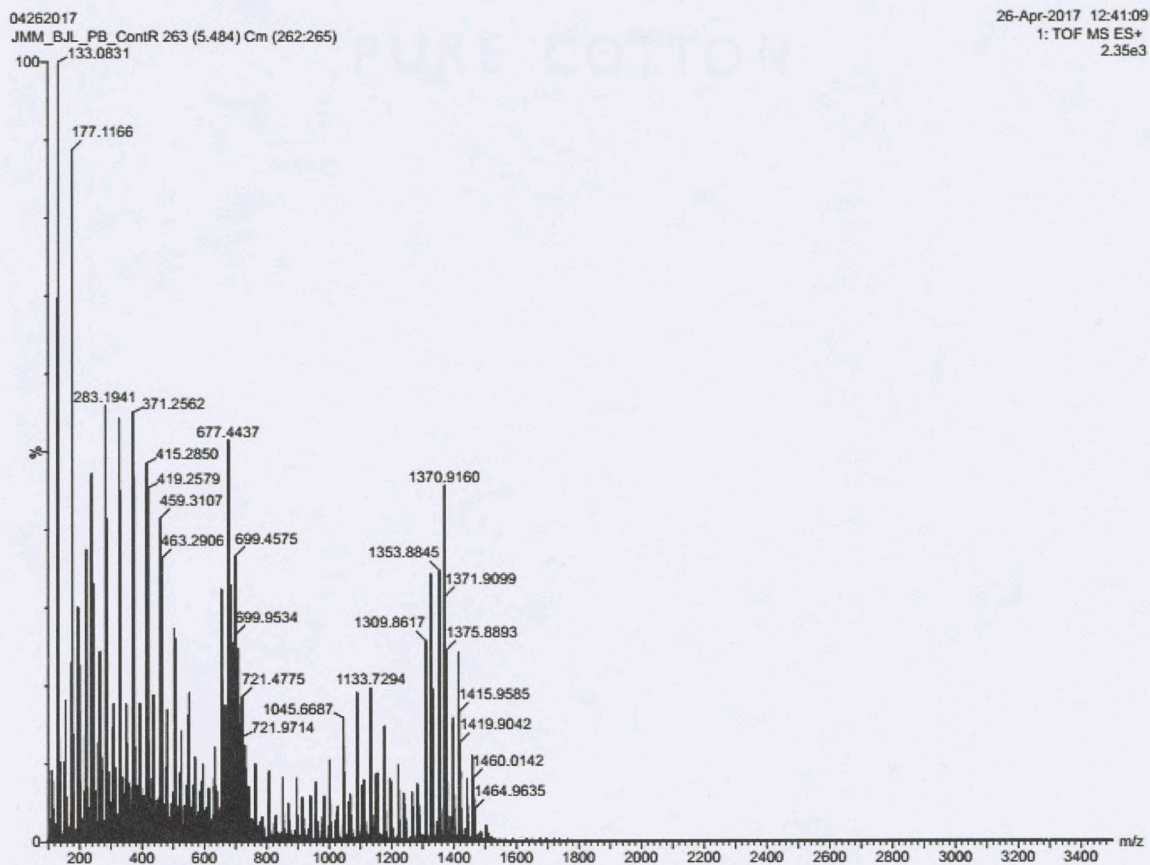


Figure 16: Spectrum of polypeptides eluted from the chromatography column at 5.484 minutes from PB sample.

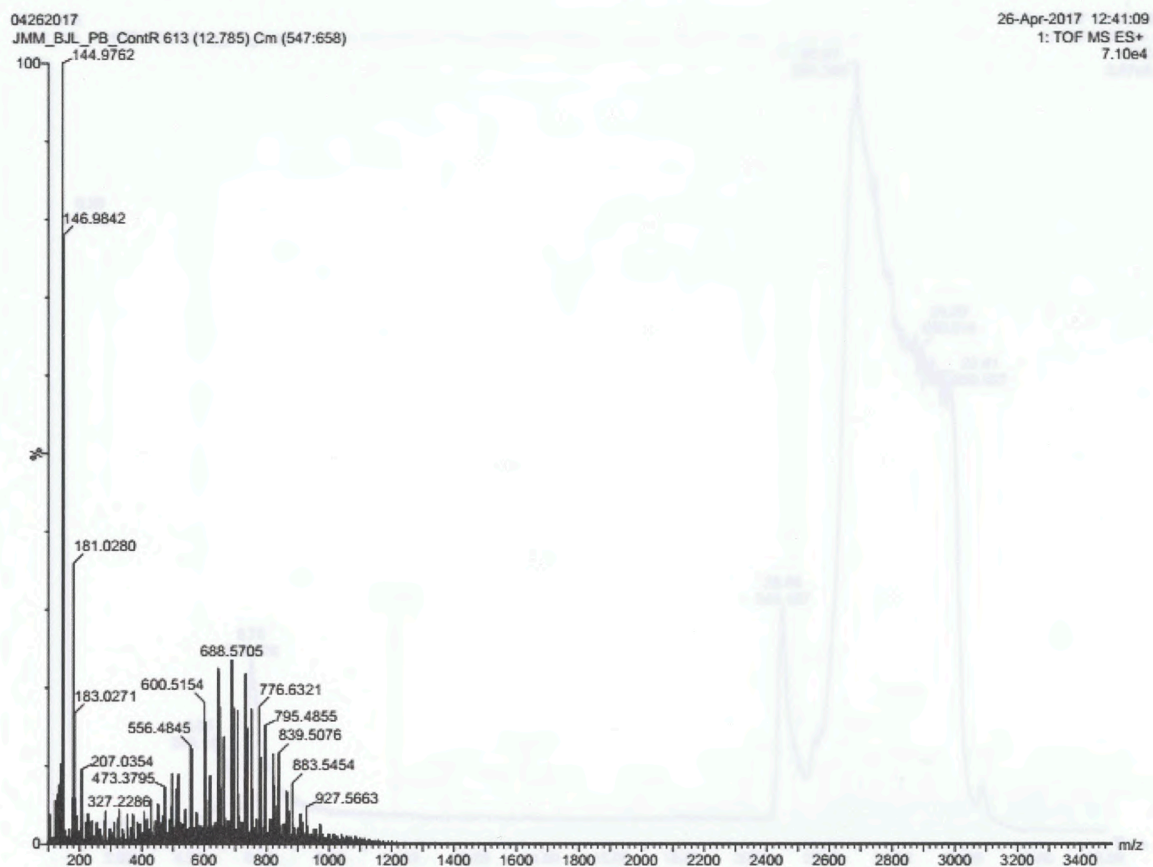


Figure 17: Spectrum of polypeptides eluted from the chromatography column at 12.785 minutes from the reaction of GFP-preptin Ala²¹ with TEV protease.

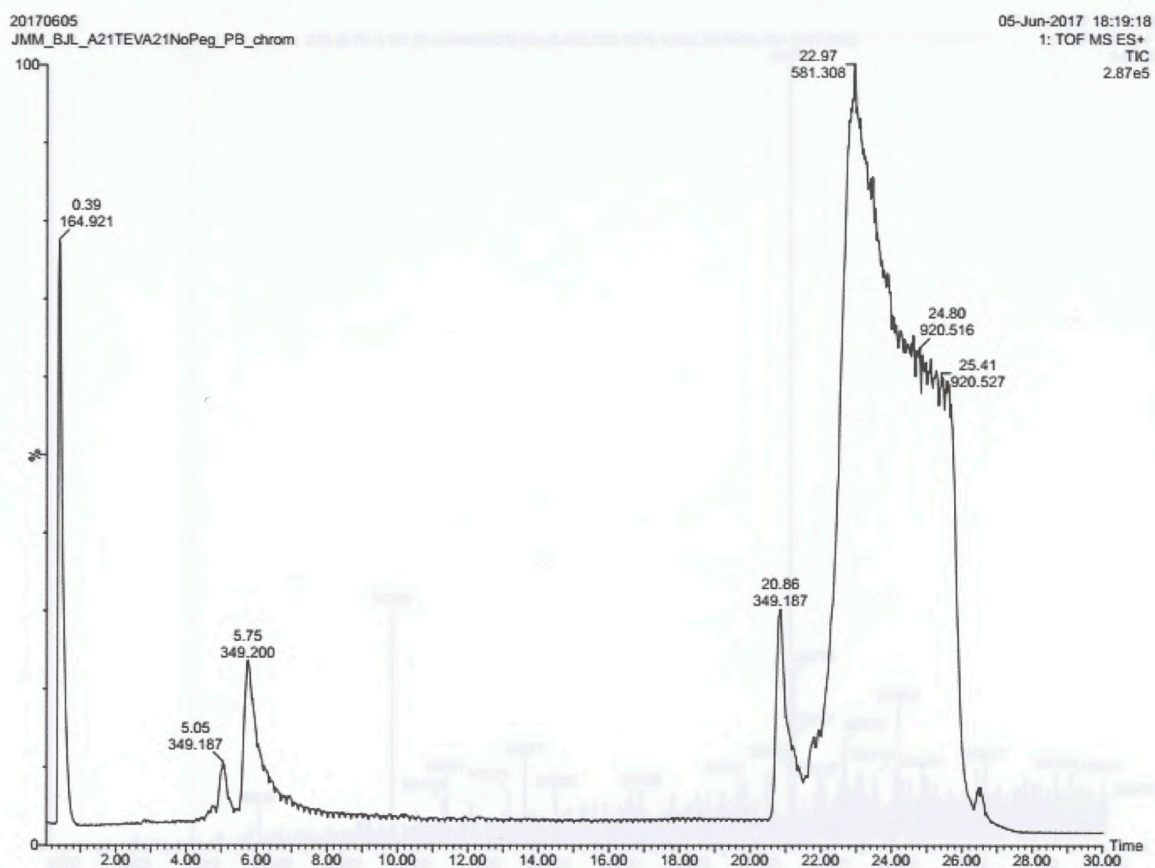


Figure 18: Chromatogram of PB without PEG (gradient elution). The majority of polypeptide signal was shifted to the column wash region of the gradient run.

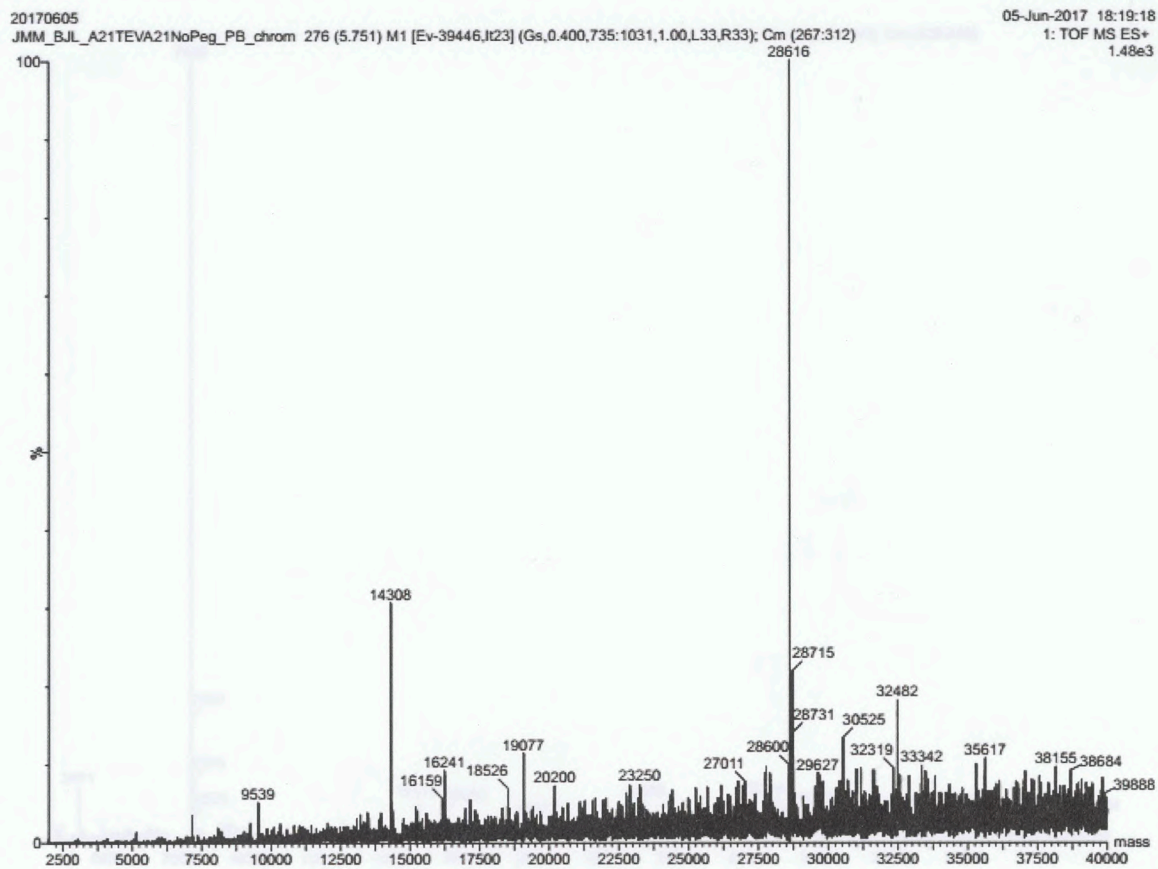


Figure 19: Mass Spectrum at 5 mins for PB w/o PEG. The peak at 5.048 minutes contained two m/z signals that could correspond to a mutated GFP fusion partner.

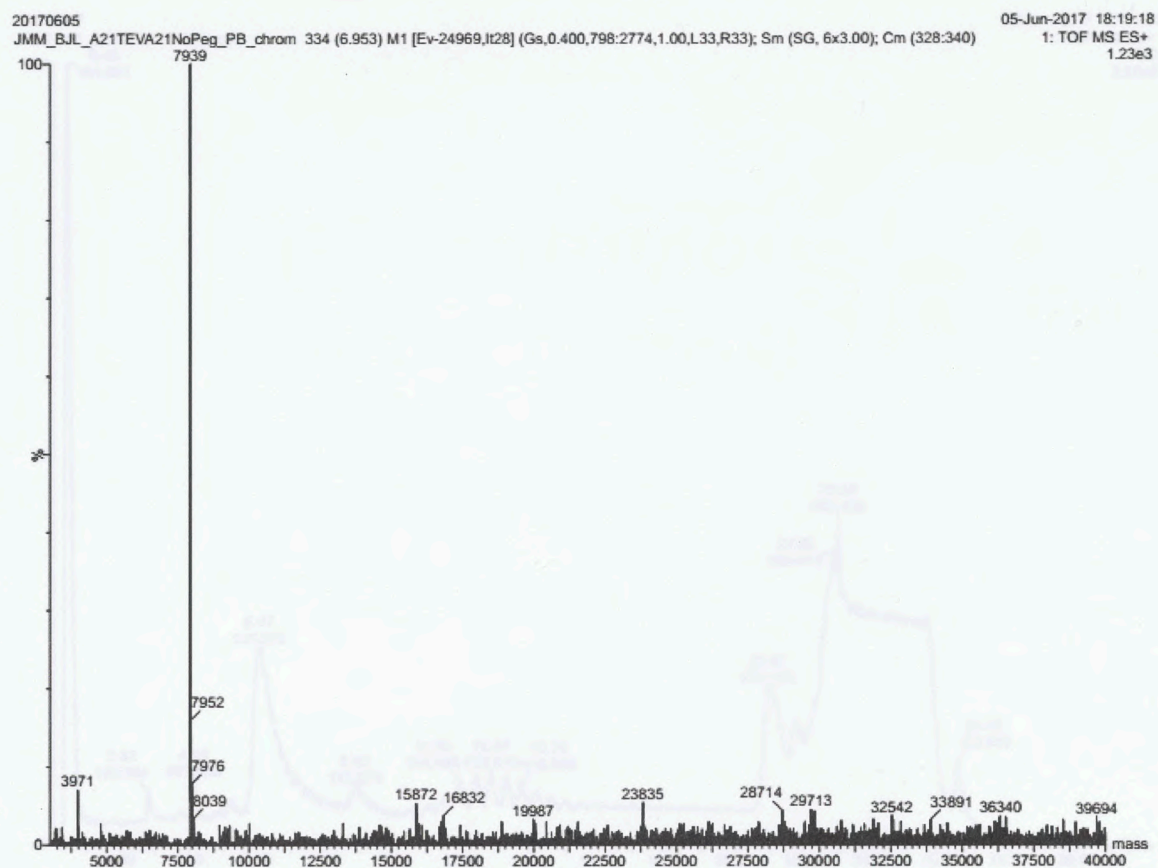


Figure 20: Mass Spectrum at 6.953 mins for PB w/o PEG. The peak contained two m/z signals that could correspond to a glycine heavy preptin-Ala²¹ mutant and its dimer.

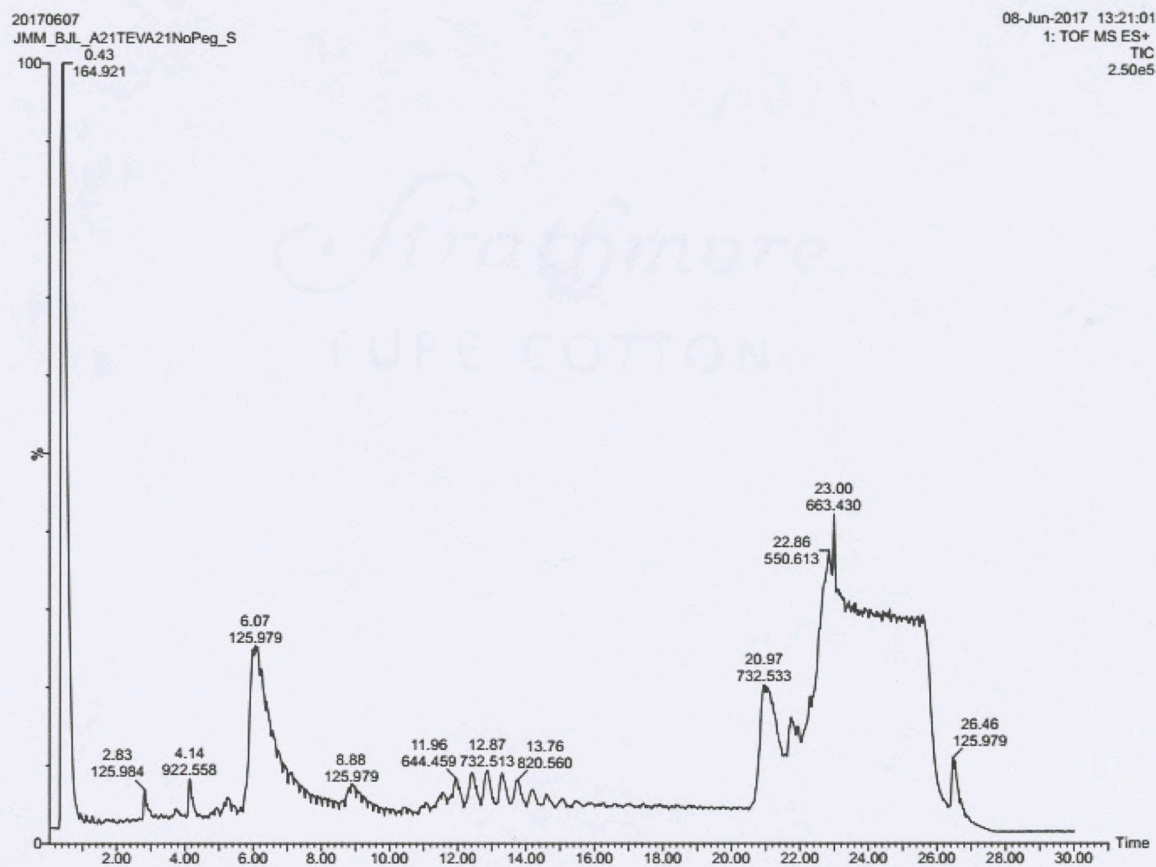


Figure 21: Chromatogram of GFP-preptin-Ala²¹ TEV cleavage reaction supernatant. The chromatograph showed an increase in the number of small and/or hydrophilic peptide signals at low acetonitrile concentrations. This is in contrast to the PB samples.

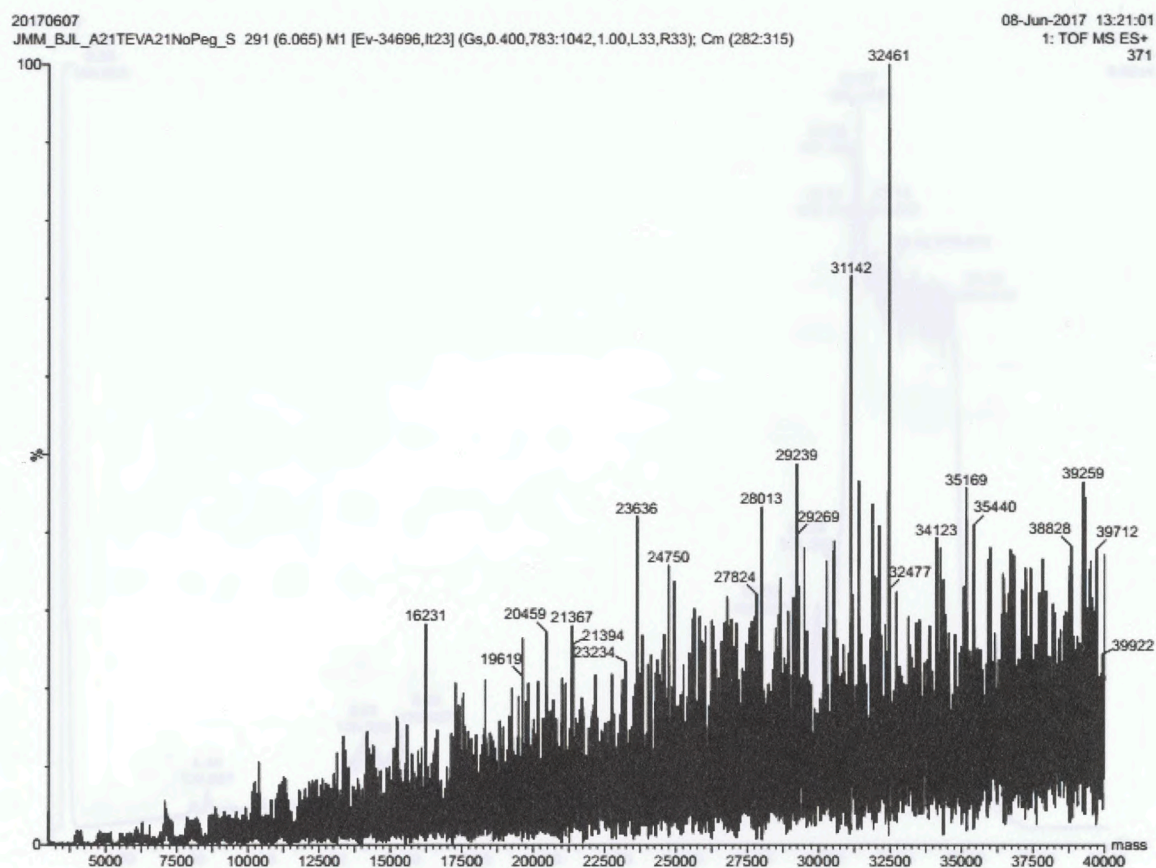


Figure 22: Mass Spectrum at 6.065 mins for TEV cleavage reaction supernatant. The peak at 32,461 Da lighter by 22.4 Da than the expected mass for the cleaved fusion tag.

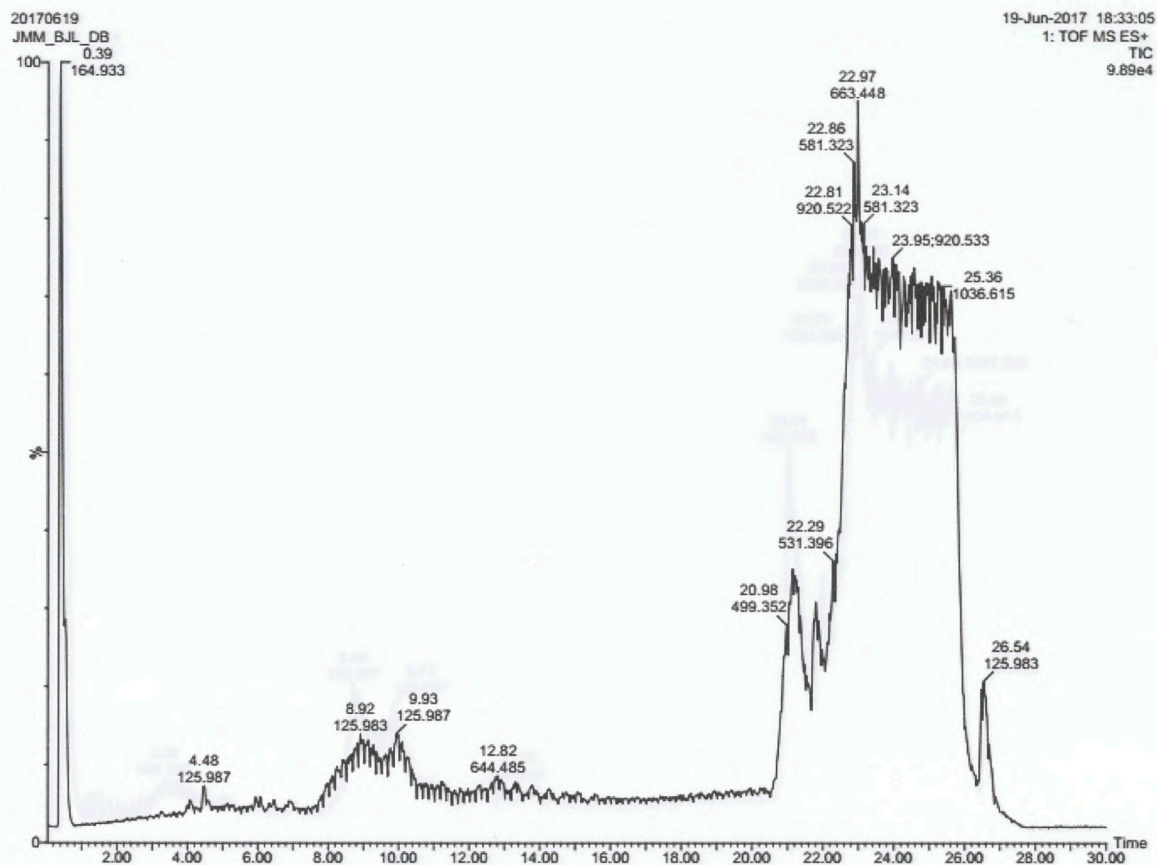


Figure 23: Chromatogram of GFP-preptin-Ala²³ fusion sample before TEV cleavage. The chromatogram was markedly different than the one observed for the Ala²³ mutation but very similar to the untreated sample.

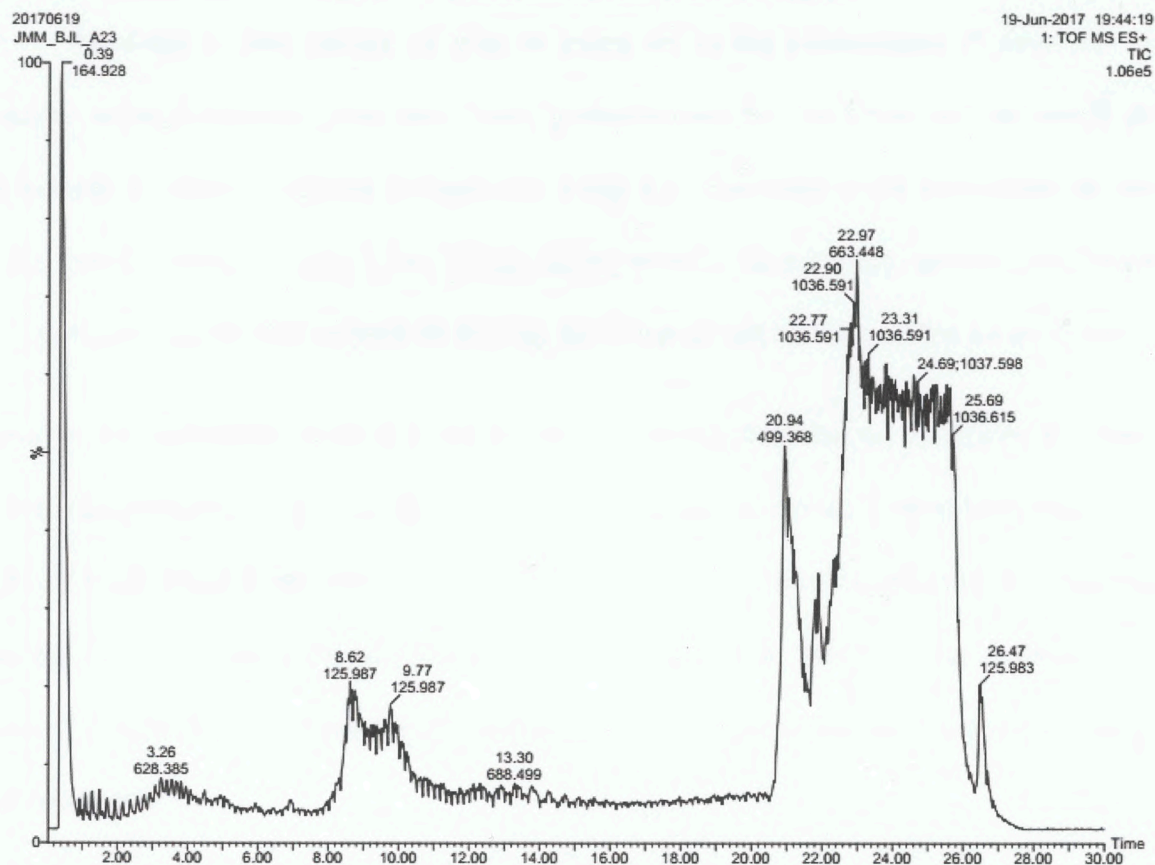


Figure 24: Chromatogram of GFP-preptin-Ala²³ TEV cleavage reaction supernatant. The chromatograph was markedly different than the one observed for the Ala²¹ mutation but very similar to the untreated sample.

A. Conclusions **IV. Conclusions and Future Work**

Although a clear picture of what is going on in the biosynthesis of these for alanine mutants remains elusive, given time these problems can be overcome and the initial goal of performing an alanine acetylating photogenisis using an E. coli expression system will be realized. In the interim, however, only a few things can be stated with certainty, several possibilities can be suggested, and the next steps in furthering the research can be mapped for future researchers.

Based on the sequencing there is a fair degree of certainty that no mutations occurred prior to the transformation of E. coli BL 21 cells. If mutations occurred, it must have been after this point, if at all. There is no major evidence to suggest a mutation of the plasmid. It is fairly certain that the fusion partner GFP was produced in an active form based on fluorescence observed under UV light. Based on the mass observed for TEV there is fair confidence that this protein was produced.

B. Future Work

Based on the data there are several suggestions that can be made as to what may have occurred in the samples. The precipitant seems to contain a large portion of the proteins found after the reaction has taken place. Although the early analysis of the first mass spectra data was more or less thrown out, the idea of protease digesting the fusion protein is still a valid one. The observations in later mass spectra suggest the presence of an extra glycine which may have occurred during or after translation, as well as other unknown modifications to the GFP fusion partner, assuming the masses mentioned above belong to the fusion partners and not native proteins still retained in the sample. The final mass spectra also suggest the possibility of protease digestion prior to the reaction using TEV.

A. Conclusions

Although a clear picture of what is going on in the biosynthesis of these for alanine mutants remains elusive, given time these problems can be overcome and the initial goal of performing an alanine scanning mutagenesis using an *E. coli* expression system will be realized. In the interim, however, only a few things can be stated with certainty, several possibilities can be suggested, and the next steps in furthering the research can be mapped for future researchers.

Based on the sequencing there is a fair degree of certainty data that no mutations occurred prior to the transformation of *E. coli* BL 21 cells. If mutations occurred, it must have been after this point, if at all. There is no major evidence to suggest a mutation of the plasmid. It is fairly certain that the fusion partner GFP was produced in an active form based on fluorescence observed under UV light. Based on the mass observed for TEV there is fair confidence that this protein was produced.

B. Future Work

Based on the data there are several suggestions that can be made as to what may have occurred in the samples. The precipitant seems to contain a large portion of the proteins found after the reaction has taken place. Although the early analysis of the first mass spectra data was more or less thrown out, the idea of protease digesting the fusion protein is still a valid one. The observations in later mass spectra suggest the presence of an extra glycine which may have occurred during or after translation, as well as other unknown modifications to the GFP fusion partner; assuming the masses mentioned above belong to the fusion partners and not native proteins still retained in the sample. The final mass spectra also suggest the possibility of protease digestion prior to the reaction using TEV.

V. Bibliography Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* 2014, 5 (APR), 1–17.

1. American Diabetes Association. Overweight: American Diabetes Association® <http://www.diabetes.org/are-you-at-risk/lower-your-risk/overweight.html?loc=atrisk-slabnav?referrer=http://www.diabetes.org/diabetes-basics/genetics-of-diabetes.html?loc=db-slabnav?referrer=http://www.diabetes.org/diabetes-basics/?loc=db-slabnav> (accessed Apr 10, 2016).
2. Buchanan, C. M.; Phillips, A. R.; Cooper, G. J. Preptin derived from proinsulin-like growth factor II (proIGF-II) is secreted from pancreatic islet beta-cells and enhances insulin secretion. *Biochem. J.* **2001**, 360 (Pt 2), 431–439.
3. VanBuecken, D. E.; Greenbaum, C. J. Residual C-peptide in type 1 diabetes: What do we really know? *Pediatr. Diabetes* **2014**, 15 (2), 84–90.
4. Zealand, N.; Zealanders, N. Type 2 diabetes <http://www.diabetes.org/diabetes-basics/type-2/?loc=superfooter> (accessed Apr 10, 2016).
5. O'Byrne, K. J.; Dalglish, A. G. Chronic immune activation and inflammation as the cause of malignancy. *Br. J. Cancer* **2001**, 85 (4), 473–483.
6. Huxley, R.; Ansary-Moghaddam, A.; Berrington de González, A.; Barzi, F.; Woodward, M. Type-II diabetes, and pancreatic cancer: a meta-analysis of 36 studies. *Br. J. Cancer* **2005**, 92 (11), 2076–2083.
7. Giovannucci, E.; Harlan, D. M.; Archer, M. C.; Bergenstal, R. M. Diabetes and Cancer : A Consensus Report. *a Cancer J. Clin.* **2010**, 60 (4), 207–221.
8. Dang, C. V. Links between metabolism and cancer. *Genes Dev.* **2012**, 26 (9), 877–890.
9. Calle, E. E.; Kaaks, R. Overweight, obesity, and cancer: epidemiological evidence and proposed mechanisms. *Nat. Rev. Cancer* **2004**, 4 (8), 579–591.
10. Aydin, S. Three new players in energy regulation: Preptin, adropin and irisin. *Peptides*. Elsevier Inc. 2014, pp 94–110.
11. Kowalczyk, R.; Yang, S. H.; Brimble, M. A.; Callon, K. E.; Watson, M.; Park, Y. E.; Cornish, J. Synthesis of truncated analogues of preptin-(1-16), and investigation of their ability to stimulate osteoblast proliferation. *Bioorganic Med. Chem.* **2014**, 22 (14), 3565–3572.
12. Singh, S. K.; Brito, C.; Tan, Q. W.; León, M. De; León, D. De. Receptor Isoforms A and B : A Link between Breast Cancer and. **2011**, 29 (6), 278–289.
13. Baykus, Y.; Gurates, B.; Aydin, S.; Celik, H.; Kavak, B.; Aksoy, A.; Sahin, I.; Deniz, R.; Gungor, S.; Guzel, S. P.; et al. Changes in serum obestatin, preptin and ghrelins in patients with Gestational Diabetes Mellitus. *Clin. Biochem.* **2012**, 45 (3), 198–202.
14. Morrison, K. L.; Weiss, G. A. Combinatorial alanine-scanning. *Curr Opin Chem Biol* **2001**, 5 (3), 302–307.
15. Map, P. IP-Free E . coli Expression Vectors with the IPTG-inducible T7 Promoter T7 Induction Protocol T7 Vector Controls Electra Cloning System. 7, 1–4.
16. Kapust, R. B.; Tözsér, J.; Fox, J. D.; Anderson, D. E.; Cherry, S.; Copeland, T. D.; Waugh, D. S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **2001**, 14 (12), 993–1000.
17. Clontech. In-Fusion® HD Cloning Kit User Manual. *In-Fusion Cloning* **2012**, 1 (11614), 1–15.

18. Rosano, G. L.; Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **2014**, *5* (APR), 1–17.
19. Kostylev, M.; Otwell, A. E.; Richardson, R. E.; Suzuki, Y. Cloning should be simple: *Escherichia coli* DH5 α -mediated assembly of multiple DNA fragments with short end homologies. *PLoS One* **2015**, *10* (9), 1–15.
20. Held, P. G. Nucleic Acid Purity Assessment using A 260 / A 280 Ratios. *BioTek-Application Note* **2001**, No. 3, 1–5.
21. Gaberc-Porekar, V.; Menart, V. Perspectives of immobilized-metal affinity chromatography. *J. Biochem. Biophys. Methods* **2001**, *49* (1–3), 335–360.
22. Sciences, G. H. L. Sephadex g25. *Desalt. buffer Exch. with Sephadex G-252000*, 1–7.
23. Yilmaz, M.; Ozic, C.; Gok, I. Principles of Nucleic Acid Separation by Agarose Gel Electrophoresis. *Gel Electrophor. - Princ. basics* **2012**, 36–40.

MASTER OF SCIENCE

DEPARTMENT OF CHEMISTRY

by

Jahbo M. Love

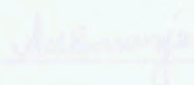
2017



Dr. Jonathan M. Meyers, Chair



Date



Dr. Anil C. Benerjee, Member



Date



Dr. D. Wade Hulley, Member



Date

Design, Cloning, and Expression of Rat Preptin and Alanine Analogs

A thesis submitted to the College of Letters and Sciences in partial fulfillment of the
requirements

for the degree of

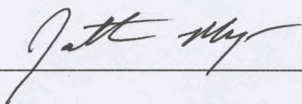
MASTER OF SCIENCE

DEPARTMENT OF CHEMISTRY

by

Jahbo M. Love

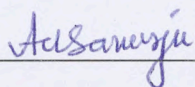
2017



Dr. Jonathan M. Meyers, Chair

11/7/2017

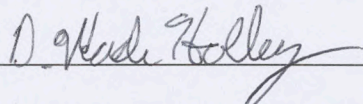
Date



Dr. Anil C. Banerjee, Member

11/7/2017

Date



Dr. D. Wade Holley, Member

11/7/2017

Date

Archives

QH

465

.A1

L68

2017

CSU

Dissertations &

Theses

Love, Jahbo M.

Design,
Cloning, and
Expression of
Rat Preptin and
Alanine Analogs

COLUMBUS STATE UNIVERSITY LIBRARY



3 7337 00322 6126

