




2022

THE BGS 13 MUTANT OF PICHIA PASTORIS AND ITS EFFECT ON STRUCTURAL CHANGES OF THE REPORTER PROTEIN B- LACTOGOLBULIN

Bushra Irshad
University of the Pacific

Follow this and additional works at: https://scholarlycommons.pacific.edu/uop_etds

 Part of the [Biology Commons](#), and the [Microbiology Commons](#)

Recommended Citation

Irshad, Bushra. (2022). *THE BGS 13 MUTANT OF PICHIA PASTORIS AND ITS EFFECT ON STRUCTURAL CHANGES OF THE REPORTER PROTEIN B-LACTOGOLBULIN*. University of the Pacific, Thesis.
https://scholarlycommons.pacific.edu/uop_etds/3844

This Thesis is brought to you for free and open access by the University Libraries at Scholarly Commons. It has been accepted for inclusion in University of the Pacific Theses and Dissertations by an authorized administrator of Scholarly Commons. For more information, please contact mgebney@pacific.edu.

THE BGS 13 MUTANT OF *PICHLA PASTORIS*
AND ITS EFFECT ON STRUCTURAL CHANGES
OF THE REPORTER PROTEIN B-LACTOGLOBULIN

By

Bushra Irshad

A Thesis Submitted

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

College of the Pacific
Biological Sciences

University of the Pacific
Stockton, California

2022

THE BGS 13 MUTANT OF *PICHLA PASTORIS*
AND ITS EFFECT ON STRUCTURAL CHANGES
OF THE REPORTER PROTEIN B-LACTOGLOBULIN

By

Bushra Irshad

APPROVED BY:

Thesis Advisor: Geoff Lin-Cereghino, Ph.D.

Committee Member: Craig Vierra, Ph.D.

Committee Member: Andy Franz, Ph.D.

Department Chair: Eric Thomas, Ph.D.

THE BGS 13 MUTANT OF *PICCHIA PASTORIS*
AND ITS EFFECT ON STRUCTURAL CHANGES
OF THE REPORTER PROTEIN B-LACTOGLOBULIN

Abstract

By Bushra Irshad

University of the Pacific
2022

Pichia pastoris, a methylotrophic yeast, is an ideal host for recombinant protein expression. It has the capability of performing many eukaryotic post-translational modifications and grows to high cell densities. However, *P. pastoris*'s secretion properties are not always efficient, and its secretory pathway mechanisms have not been thoroughly elucidated. A previously identified mutant strain, *bgs13*, was found to efficiently secrete most recombinant proteins tested, raising the possibility that this *bgs13* mutant is a universal super secretor and understanding its secretion process is needed. In this study, we used a reporter protein, β -lactoglobulin (b-LG), to perform structural analysis and comparisons of protein secreted from wild type and mutant *bgs13* strains of *Pichia pastoris*. Primary, secondary, and tertiary structure of b-LG were examined using Edman sequencing, circular dichroism, and tryptophan fluorescence. Our results suggest that the mutant strain improves the folding of b-LG, affecting its tertiary structure. Understanding the enhanced folding ability of *bgs13* will provide insight into the secretory mechanism of *P. pastoris*, which will allow for further improvement and use of *P. pastoris* as an optimal host for protein production.

Table of Contents

List of Tables.....	6
List of Figures.....	7
CHAPTER 1: Introduction.....	9
Recombinant Protein Expression.....	9
Expression Systems.....	10
Advantages of <i>Pichia Pastoris</i> as an Expression System.....	10
Inefficient Secretion with <i>Pichia Pastoris</i>	12
N-linked Glycosylation.....	12
The <i>BGS13</i> Mutant Strain of <i>Pichia Pastoris</i>	15
Research Objectives and Aims.....	16
CHAPTER 2: Materials and Methodology.....	19
Strains, Media, and Reagents.....	19
BioInformatics.....	19
General Methodology.....	20
CHAPTER 3: Results.....	32
Construction of the PBI-1 Expression Vector (MAT α -bLG-myc-his-6).....	32
Transformation of PBI1 into WT and <i>BGS13 P. Pastoris</i>	38
Small Scale Expression of BLG.....	38
Large Scale Expression.....	43
HisPur Cobalt Column Protein Purification.....	44
Dialysis of b-LG.....	45

	5
CD Spectroscopy.....	45
Melting Temperature Using GloMelt.....	47
EndoHf Assay.....	47
N-Terminal Sequencing.....	49
Melting Temperature Using CD.....	50
Melting Temperature Using Tryptophan Fluorescence (TRP Fluorescence).....	52
Temperature Induced Aggregation Assay.....	54
CHAPTER 4: Discussion.....	56
Proposed Models for <i>BGS13</i>	57
Secondary Structure.....	58
Tertiary Structure.....	59
Conclusion.....	60
Future Work.....	61
References.....	63

List of Tables

Table

1. Forward and Reverse Primer.....	20
2. Dilutions for Circular Dichroism.....	28
3. Amino Acid Sequence at N-terminus.....	48
4. b-LG Dimers and Monomers.....	53

List of Figures

Figure

1. Plasmid Map of pBLHIS SX with Restriction Enzymes.....	34
2. bLG Protein Sequences.....	34
3. b-LG PCR Fragment and pBLHIS SX.....	37
4. Samples of Plasmids.....	38
5. Restriction Digestion with Mph1103I and BamH.....	39
6. Plasmid Map of pBI1.....	39
7. Spot Western Results with Anti-myc Primary Antibody.....	42
8. Coomassie Blue Stain of Extracellular Matrixes of Samples from Deep Well Plate Assay.....	43
9. Western Blot Using Mouse Anti-myc Primary and Goat Anti-Mouse Secondary Antibody.....	44
10. Western Blot Using Anti-bLG Antibody.....	45
11. Coomassie of Large-Scale Expression Showing b-LG Band at About ~30 kDA in lanes 4 and 6.....	47
12. Coomassie Stained SDS PAGE of Purified b-LG from WT P.pastoris and b-LG from bgs13.....	48
13. CD Graph Showing Two Different Concentration of Commercial b-LG at 1.0 µg/µl and 0.5µg/µl.....	50
14. CD Spectra.....	50
15. EndoHf Assay of the WT and bgs13 b-LG Proteins with RnaseB and EndoHf as Controls.....	52
16. Predicted Amino Acid Sequence at the N-terminus of b-LG.....	54

17. b-LG Melting Curves Showing the Melting Temperature (T_m) of Both Proteins.....	55
18. b-LG Melting Curves including the Commercial Protein at 207 nm.....	56
19. Trip Fluorescence Curve.....	57
20. Coomassie Blue Stain of SDS-PAGE Gel Examining Temperature Introduce Aggregation of b-LG.....	58

CHAPTER 1: INTRODUCTION

Recombinant Protein Expression

Proteins, otherwise known as the “building blocks of life,” are the workhorses in biological systems. They facilitate most of the processes in a cell. The blueprint for proteins is stored in DNA and defined sequences of amino acids form a protein through a process known as a gene expression. A recombinant or heterologous protein is a manipulated form of a native protein which is generated in various ways in a different host organism. Recombinant protein production begins at the genetic level, where the sequence for the gene of interest is isolated and then cloned into an expression plasmid vector. Many expression systems can be used to express recombinant protein including bacteria (Baneyx, 1999), yeast (Cregg et al., 2000; Malys et al., 2011), baculovirus/insect (Kost et al., 2005), and mammalian cells (Lackner et al., 2008; Jager et al., 2015). Though each expression system is unique and requires different protocols, the general process of recombinant protein expression is completed by similar steps. First, the gene of interest is isolated and then inserted into a vector that will be compatible with the type of expression system that will be used. After this step, the vector is transformed into the expression system. Consequently, the now-recombinant host is cultured under conditions to trigger the production of the protein.

Recombinant protein expression is a driving force of the biotechnology market with proteins being produced for biomedical research, to understand health and disease as well as to serve as biotherapeutics. Recombinant proteins need to be produced fast, efficiently, and in a biologically active form (Vieira Gomes et al., 2018). Recombinant protein production is mainly driven by the need for cost-effectiveness, simplicity, and speed.

Expression Systems

As previously mentioned, a few different types of expression systems exist. *E. coli*, a commonly used bacterial expression system, efficiently produces many recombinant proteins. It requires inexpensive carbon sources for growth and has a short doubling time (Sahdev S. et al., 2008). However, it lacks the machinery required to perform many of the post-translational modifications associated with higher eukaryotes such as phosphorylation and glycosylation (Sahdev et al., 2008). These post-translational modifications are a critical downstream process required to produce many biologically active proteins (Ferrer-Miralles et al., 2009). The proteins that don't require these post-translational modifications can be expressed in prokaryotic expression systems such as *E. coli*. However, most recombinant proteins require post-translational modifications to ensure proper folding and bioactivity and therefore cannot often be produced correctly in prokaryotic organisms (Jenkin et al., 2008). Compared to bacteria, yeast have significant advantages including growth speed, post-translational modifications, and easy genetic manipulation (Karbalaei et al., 2020). Also, yeast have been utilized in successfully producing many therapeutic recombinant proteins on the market such as human insulin, human serum albumin and hepatitis B (Vieira Gomes et al., 2018).

Advantages of *Pichia Pastoris* as an Expression System

Pichia pastoris, also known as *Komagataella phaffii*, is a methylotrophic yeast that is used as a host for recombinant protein expression. *P. pastoris* can produce milligram to gram per liter quantities of proteins intended for research and industrial purposes (Macauley-Patrick et al., 2005). The production process of *P. pastoris* can be scaled up or down depending on the situation (Macauley-Patrick et al., 2005). Many culturing parameters that influence protein productivity and activity such as pH, aeration, and carbon source feed rate can be controlled in a

fermenter (Macauley-Patrick et al., 2005). In addition, *P. pastoris* can express foreign proteins intra- or extracellularly (Cereghino & Cregg, 1999). Its use in recombinant DNA technology is ideal because it secretes few endogenous proteins, potentially making the secreted heterologous proteins the majority of the total proteins in the medium (Cereghino & Cregg, 1999). This makes downstream processing much faster and simpler. One of its major advantages, relative to the well-studied *Saccharomyces cerevisiae*, is that it has poor fermentation and is an obligate aerobic yeast. This is advantageous because ethanol, which is a product of *S. cerevisiae* fermentation, often builds up to toxic levels, which limits the further growth and production of heterologous proteins in the system (Cereghino & Cregg, 1999). Due to *P. pastoris*'s preference for respiratory growth, it can be cultured to high densities in a controlled environment with little risk of this process becoming toxic, as in *S. cerevisiae* (Cereghino & Cregg, 1999).

Another major advantage that *P. pastoris* has over prokaryotic expression systems is that it has the potential to perform many of the post-translational modifications associated with higher eukaryotes which include folding, disulfide bridge formation and O- and N- linked glycosylation (Cereghino & Cregg, 1999). Lastly, *P. pastoris* possesses one of the strongest and most highly inducible promoters, alcohol oxidase-1 (*AOX1*). Alcohol oxidase-1 (*AOX1*) drives the expression of foreign genes of interest by regulating protein expression through the manipulation of culture medium. In the presence of glucose or glycerol medium, transcriptional expression is repressed; however, changing the culture to methanol initiates induction and causes a 1,000-fold increase in promoter activity (Cereghino & Cregg, 1999). Lastly, *P. pastoris* is a model organism to understand and investigate areas of cell biology research including the organization and function of the secretory pathway in eukaryotes (Cereghino & Cregg, 1999).

Its many advantages such as high protein production, culturing conditions and the *AOXI* promoter allow it to be an ideal host over other expression systems.

Inefficient Secretion with *Pichia Pastoris*

Although there are many advantages of *P. pastoris*, some recombinant proteins are secreted inefficiently by the yeast. Because the secretion pathway of *P. pastoris* has never been fully studied, the mechanisms of *trans*-acting factors involved in secretion in *P. pastoris* are poorly understood. A major problem that arises when using *P. pastoris* as an expression system is that recombinant proteins engineered to be secreted are sometimes retained inside the cell. Instead of being exported, recombinant proteins may accumulate in the ER, Golgi, or endosomal compartments as either folded or misfolded products, promoting ER-associated protein degradation (ERAD) where proteins are translocated from the ER to the cytosol and degraded in proteasome or degraded in the vacuole (Naranjo C. et al., 2019). Another scenario is that the proteins may lack the correct post-translational modifications and folding which causes them to have reduced activity or retained in the cell. Therefore, *P. pastoris* has limitations with insufficient secretion being a major obstacle (Larsen S. et al., 2013).

N-linked Glycosylation

Consequently, while *P. pastoris* can perform many post-translational modifications, a problem often arises with glycosylation. Some proteins can be hyperglycosylated when produced in *P. pastoris*. Glycosylation consists of adding sugars to amino acid residues on a protein while it is being processed within the secretory organelles of a eukaryotic cell. *P. pastoris* has the ability of performing both *O*- and *N*- linked glycosylation (Cereghino & Cregg, 1999). *O*-linked glycosylation in eukaryotes involves the addition of an *O*-linked saccharide on the hydroxyl groups of the amino acids serine and threonine (Cereghino & Cregg, 1999). *O*-

oligosaccharides in *P. pastoris* are composed solely of mannose (Man) residues but the process of O-linked glycosylation is poorly understood (Cereghino & Cregg, 1999). In eukaryotes, the process of N-linked glycosylation, which has been more thoroughly investigated, starts in the endoplasmic reticulum, and involves the transfer of a lipid-linked oligosaccharide unit (Glc3Man9GlcNAc2) (where Glc3=glucose and GlcNAc2= N-acetylglucosamine and Man9=mannose) to asparagine at the recognition sequence Asn-X-Ser/Thr (Cereghino & Cregg, 1999). The lipid-linked oligosaccharide is added to the asparagine, and then the oligosaccharide core is trimmed to become Man8GlcNAc2 (Cereghino & Cregg, 1999). After this has happened, the pathway of N-linked glycosylation begins to differ between higher and lower eukaryotes (Cereghino & Cregg, 1999). In mammals, the Golgi apparatus performs trimming and addition reactions whereas, in yeasts, the N-linked core units are elongated in the Golgi due to the addition of mannose on the outer chains (Cereghino & Cregg, 1999). These outer chains end up being different lengths and cause hyperglycosylation (Cereghino & Cregg, 1999).

Hyperglycosylation poses several problems for the use of *P. pastoris* as an expression system because N-glycosylation plays a role in protein secretion, protein folding, stability, antigenicity, and enzyme activity (Roth et al., 2010). Hyperglycosylated proteins that become antigenic when introduced into mammals can potentially initiate an immune response (Cereghino & Cregg, 1999). Another problem with hyperglycosylation results from the difference between yeast and mammalian N-linked glycosylation pattern. The longer outer chain in yeast can potentially interfere with the folding and function of a foreign protein (Cereghino & Cregg, 1999). Misfolded proteins may have modified or no function at all. This makes folding a rate-limiting factor because a protein cannot be biologically active unless it takes on its native shape. Lastly, glycoproteins that are properly modified and folded in the ER enter the Golgi apparatus

where glycosylation continues and the correctly folded/modified proteins are then secreted outside of the cell (Roth et al., 2010). However, in the presence of misfolded glycoproteins, the unfolded protein response (UPR) is activated, and the misfolded product is degraded through ER-associated degradation (ERAD) and are degraded in the cytoplasm by proteosomes (Roth et al., 2010). Thus, N-hyperglycosylation can lead to improperly folded proteins 1) being secreted outside of the cell in an inactive state or 2) simply being retained inside the cell, initiating ER-associated protein degradation, minimizing the use of *P. pastoris* as a heterologous expression system.

While some studies have shown that reducing the number of native N-glycosylation sites in glycoproteins leads to a decreased level of secretion, the opposite results have been reported in other studies (Roth et al., 2010). The tetanus toxin fragment C (TetC) from *Clostridium tetani*, which is composed of 451 amino acids, was used in a study to examine the effect of N-glycosylation on secretion (Roth et al., 2010). Wild-type TetC and eight mutated gene fragments were inserted into *P. pastoris*, and the level of intracellular and secreted proteins were measured and compared. The results showed that an increase in the number of N-glycosylation sites increased the yield of the secreted protein in most of the mutants (Roth et al., 2010). However, in another study, when some proteins were engineered to reduce glycosylation levels in *Saccharomyces cerevisiae*, the extracellular activities of the recombinant proteins were enhanced significantly (Tang H. et al., 2016). Most of the information surrounding the N-glycan dependent ER quality control system has been derived from *Saccharomyces cerevisiae* or mammalian cells but has not been studied in *P. pastoris* (Roth et al., 2010). More research is needed to examine the effects of N-glycosylation on secretion of proteins in *P. pastoris* specifically.

The *BGS13* Mutant Strain of *Pichia Pastoris*

Scientists have tried to overcome these secretion obstacles in *P. pastoris* by illuminating the protein export process in this yeast. It is noted through a combination of single-cell measurements and computational modeling that protein trafficking through the secretory system is often the rate-limiting step for heterologous protein production in yeast (Gellissen et al., 1992). Therefore, gene products affecting secretion need to be identified. However, only a few *P. pastoris* gene products affecting secretion have been identified (Larsen S. et al., 2013). In a previous study, performed in the Lin-Cereghino lab, the disruption of 12 genes increased the secretion efficiency of a β -galactosidase reporter (Larsen S. et al., 2013). The mutants were chosen using the basis that wild type *P. pastoris* cannot grow on medium containing lactose as a sole carbon source because it lacks β -galactosidase (Larsen S. et al., 2013). β -galactosidase is needed to convert lactose to glucose and galactose. *P. pastoris* strains that were transformed with *E. coli* β -galactosidase (*lacZ*) fused to the MAT α secretion leader under the control of the strong constitutive *GAP* promoter were still unable to grow on medium containing lactose because the enzyme, β -galactosidase, was still not being secreted (Larsen S. et al., 2013). The inability of wild type *P. pastoris* to secrete the enzyme was used to select mutants that had defects in the secretion process that would enable them to export β -galactosidase and grow on lactose medium (Larsen S. et al., 2013). The mutant strains were generated using the restriction enzyme-mediated integration (REMI) method which randomly inserts a Zeocin-resistant vector (pREMI-Z) into genomic DNA to disrupt gene expression (Larsen S. et al., 2013). This allowed for the isolation of mutants that secrete significant amount of β -galactosidase to grow on lactose medium (Larsen S. et al., 2013).

One mutant identified as *bgs13* affects the secretion of a wide range of recombinant reporter proteins, suggesting that it may play a more general role in protein export (Larsen S. et al., 2013). Further analysis of the *bgs13* mutant showed that Bgs13 peptide (Bgs13p) is similar to the *Saccharomyces cerevisiae* protein kinase C 1 protein (Pkc1p), which plays a significant role in cell wall integrity (Naranjo C et al., 2019). It was found that the disruption of *BGS13* in the chromosome led to a truncated protein that had reduced protein kinase C activity and a different location in the cell compared to the wt Bgs13 protein (Naranjo C et al., 2019). Also, the *bgs13* mutant's cell wall was found to have structural problems though its porosity was normal (Naranjo C et al., 2019). Because the *bgs13* strain also had increased levels of normally secreted proteins, as well as endoplasmic reticulum-resident chaperone proteins, this suggested that Bgs13p helps regulate the UPR and is involved in protein sorting (Naranjo C. et al., 2019). Overall, the strains expressing this modified mutant bgs13 protein have lower protein kinase C activity, an abnormal cell wall structure, reduced proteolytic activity, and altered protein sorting within the secretory network of organelles (Naranjo C. et al., 2019). However, because protein kinase C peptides are master regulatory proteins involved in many cellular processes in other eukaryotes, it is unclear if these changes result in increased secretion of proteins or are side effects of the mutant bgs13 protein activity that are unrelated to enhanced peptide export (Naranjo C. et al., 2019).

Research Objectives and Aims

Despite the success of the *P. pastoris* system, opportunities exist to expand the range of proteins that can be expressed in the host system (Cereghino & Cregg, 1999). Studies are needed to address problems associated with the secretion of recombinant proteins from *P. pastoris*, especially those related to N-linked glycosylation (Cereghino & Cregg, 1999). The analysis of

the structural properties of secreted proteins produced by the mutant strain, *bgs13*, will allow for further augmentation of this knowledge. The goal of this research is to further understand the pathway of secretion in *P. pastoris* so it can be genetically altered to make it a more ideal host for recombinant protein expression.

The effects of *bgs13* on N-linked glycosylation can provide insight into whether changes in glycosylation patterns are increasing secretion. β -lactoglobulin (BLG) is a reporter protein, with well characterized glycosylation, which has been previously expressed in *P. pastoris*. Previous research, in *S. cerevisiae* suggested that Pkc1p, the homolog of *bgs13p* physically interacts with the glycosylation enzyme, oligosaccharyl transferase (Park & Lennarz, 2005). Compared to the native BLG which comes from bovine sources, recombinant β -lactoglobulin produced in *P. pastoris* appears to have the same conformation and stability (Kim et al., 1997). The recombinant BLG is also secreted at high levels by the yeast, which should allow it to be obtained at sufficient levels to perform different types of analysis. These factors make BLG an ideal reporter protein to be used in our experiments to examine secretion levels as well as N-linked glycosylation differences between the wild type and *bgs13 P. pastoris*. BLG, containing C-terminal c-myc and His6 tags, will first be expressed in wild type and *bgs13* strains to examine secretion levels. The BLG will then be purified from both strains using cobalt affinity chromatography. We predict that secretion will be higher in *bgs13*; however, the mechanism behind this is unknown. We believe that N-glycosylation and protein stability may play a role. Our experiments will examine the secondary and tertiary structure of BLG, along with N-glycosylation, to determine the folding differences in protein produced by the wild type and *bgs13* strain, if there are any.

In addition, if the *bgs13* strain displays a decrease in N-glycosylation, this would generate even more interest in this mutant because hyperglycosylation of recombinant proteins is a serious problem with therapeutic proteins due to the tendency of these sugar groups to trigger an unwanted immune response. The results of this research have the potential to open up new areas of study and provide further insight into *bgs13*'s role in controlling glycosylation and regulating secretion.

CHAPTER 2: MATERIALS AND METHODOLOGY

Strains, Media, and Reagents

The *Escherichia coli* strain One Shot TOP10 (Invitrogen Corp., Carlsbad, CA) was used for recombinant DNA manipulations. TOP10 was cultured at 37°C in Lennox Broth (LB) medium (0.5% yeast extract, 1% glucose, 0.5% NaCl) supplemented with 30µg/mL ampicillin or plated on LB + Amp plates.

Pichia pastoris yJC100 is a derivative of the original wild type *P. pastoris* strain NRRL Y11430 (North Regional Research Laboratories, US Department of Agriculture (Peoria, IL) and was used as the “parent” or wild type strain in this research. Yeasts were cultured in YPD medium (1% yeast extract, 2% peptone, 2% glucose), or were plated on YPD plates supplemented with 25µg/mL Zeocin (Life Technologies, Carlsbad, CA). Yeast and bacterial cultures were incubated in a New Brunswick Scientific C25 Incubating Shaker (Edison, NJ) at 30°C or 37°C, as necessary.

All the primers used for site-directed mutagenesis and PCR were produced by Sigma-Aldrich (St. Louis, Mo). The restriction enzymes and their corresponding buffers were purchased from MBI Fermentas (Hanover, MD). The optical densities at 600nm were measured by Spectronic Genesys 2 (Spectronic Instruments Inc., Rochester, NY). Also, the DNA and the protein concentrations were measured at 260nm and 280nm using the Nano Drop 2000c (Thermo Scientific, Wilmington, DE).

BioInformatics

All DNA sequences, restriction maps, and oligonucleotide for PCR were generated with SnapGene (GSL Biotech, San Diego, CA).

General Methodology

Sequence Analysis

Using SnapGene, the multi or hyper glycosylated β -lactoglobulin cDNA (Batt, 2001) provided by the Batt laboratory was examined to design forward and reverse primers (Table 1). Primers were diluted to 1 $\mu\text{g}/\mu\text{L}$ and stored at -20°C . Also, the plasmid pBLHIS SX was analyzed with SnapGene to identify unique enzyme cut sites to be later used for restriction digestion and ligation reactions. *EcoRI* and *NotI* were chosen as the restriction enzymes.

Table 1

Forward and Reverse Primers

Name of Primer	Sequence (5'-3')	μg
5BLGEcoR1long	AAGAATTCTCAAAAGAGAGGCTGAAGCTTACGTTAC CCAGACCATGAAGGGC	1055.3
5BLGEcoR1short	AAGAATTCTCAAAAGAGAGGCTGAAGCTTACGTTAC CCAGAC	561.7
3BLGNotI	CCGCGGCCGCGATGTGGCACTGCTCCTCCAG	459.9

PCR of Hyperglycosylated BLG S123 Fragment

3 μL (100 ng/ μL) of each forward and reverse primers were added to 5 μL of 10x PCR buffer, 10 μL 5x Taq Master Enhancer, 1 μL 10 mM dNTP solution, 23 μL of sterile MilliQ water, and 0.5 μL of Taq DNA polymerase per reaction all in a 0.2 ml Eppendorf PCR tube, using the kit components provided by the MasterTaq kit from 5 Prime (San Francisco, CA). PCR was carried out in Applied Biosystems 2720 Thermal Cycler (Foster City, CA). pBLG S123, which

contains the mature BLG coding sequence with three additional N-glycosylation sites, was used as a DNA template. For the first cycle, the PCR tube was incubated 95°C for 5 minutes. Then, for cycles 2-35, the tube was incubated at 95°C for 1 minute, 53°C for 1 minute, and 72°C for 1 minute each cycle. Lastly, the tube was incubated at 72°C cycle for 10 minutes and left at 4°C cycle until the tube was taken out and stored at -20°C

DNA Electrophoresis

Pre-made 1.2% agarose Flash gels (Lonza, Rockland, ME) were used for running and detecting DNA. 2µL of 5x DNA loading dye (50% glycerol, 0.25% bromophenol Blue, 0.35% 30 xylene cyanol, 0.01% SDS, 10 mM Tris pH 8.0, 1 mM EDTA pH 8.0) was added to 10µL of each DNA sample. 6µL of each sample was loaded into a well of the Flash gel along with 5µL of 1Kb Plus ladder from Thermo Scientific (Waltham, Massachusetts). The PCR products were electrophoresed at 275 V for 7 minutes and then visualized using the ChemiDoCTM XRS+ System from Bio-Rad (Hercules, CA).

Bacterial Cultures to Prepare for Miniprep pBLHIS SX Plasmid

3mL of LB culture with ampicillin (final concentration=75 µg/mL) were mixed in sterile 50mL conical tubes. Colonies containing plasmid were picked with P200 micropipette tips and placed in the sterile conical tubes containing LB medium with ampicillin. The conical tubes were then placed in the 37°C shaker at 225 RPM and incubated overnight.

Plasmid Miniprep

Plasmid DNA was isolated using the Zyppy Plasmid Miniprep Kit (Zymo Research, Irvine, CA). 1.5mL cultures of colonies were transferred into 1.5mL Eppendorf tubes. The tubes were centrifuged at 16,000xg for 2 minutes, and the supernatants were discarded. If double pelleting, the procedure was repeated by adding another 1.5mL of overnight LB culture and

centrifuged at 16,000xg for 2 minutes. After discarding the supernatant, the pellet was resuspended into 600 μ L of sterile MilliQ water. 100 μ L of 7x Lysis Buffer was added to the resuspended cells and inverted 6 times. 350 μ L of the cold Neutralization Buffer was added and mixed by inverting 3 times. The tubes were then centrifuged at 16,000xg for 4 minutes. The supernatant (approximately 900 μ L) was transferred to the Zymo-Spin column. The column was placed in a collection tube and centrifuged at 16,000xg for 30 seconds. The flow through was discarded and 200 μ L of Endo-Wash Buffer was added into the column. The column was centrifuged for 30 seconds at 16,000xg. 400 μ L of the Zyppy Wash Buffer was added to the column and centrifuged for 1 minute. After centrifugation, the column was transferred into a new, clean 1.5mL Eppendorf tube. 50 μ L of Elution buffer was added to the column and centrifuged for 30 seconds to elute the purified plasmid DNA.

DNA Purification

PCR products and digested plasmid DNA products were cleaned and concentrated using the Zyppy DNA Clean & Concentrator Kit (Zymo Research, Irvine, CA). In a sterile 1.5mL Eppendorf tube, 2 volumes of DNA Binding Buffer were added to each DNA sample. The solution was then transferred to a Zymo-Spin column in a collection tube. The mixture was centrifuged for 30 seconds at 13,000 rpm, and the flow through was discarded. 200 μ L of the Wash Buffer was added to the column and centrifuged for 30 seconds. The wash step was repeated twice. The column was transferred to a new, sterile 1.5mL Eppendorf tube and 30 μ L of the Elution Buffer was added directly to column matrix. The column was centrifuged at 30 seconds to elute the clean and concentrated DNA.

Restriction Enzyme Digestion

Restriction digestion was performed to prepare plasmids for subcloning, and to confirm correct construction of plasmids. The restriction digests include the following: 500ng of the plasmid or linear DNA, 1 μ L of the restriction enzyme, and the appropriate 10X restriction enzyme buffer specific to the restriction enzyme (MBI Fermentas, Hanover, MD) which were placed in a 1.5mL Eppendorf tube. The total volume was brought to 30 μ L with sterile MilliQ water. The components were mixed via vortex and spun down briefly in a centrifuge. The reaction was incubated at 37°C water bath for 1-2 hours. Pre-cast 1.2% agarose Flash gels (Lonza, Rockland, ME) were used to determine the restriction digest products.

DNA Sequencing

All purified DNA plasmids were sent to Quintarabio (Richmond, CA) for sequencing.

Linearization and DNA Purification

pBLHIS SX-based plasmids were linearized within the *HIS4* region of the plasmid with the restriction enzyme *StuI*. Linearization was accomplished by adding 3 μ L of plasmid DNA (approximately 100-200 ng/ μ L), 3 μ L of 10X blue buffer, 1 μ L of *StuI* restriction enzyme, and sterile water to a final volume of 30 μ L. The reaction mixture was vortexed, centrifuged, and incubated at 37 °C for 2 hours. Linearized plasmids were then purified with the DNA Clean & Concentrator kit (Zymo Research, Irvine, CA) and DNA concentration was measured using the Nano Drop 2000c.

***P. pastoris* Transformation by Electroporation**

Electroporation generates an electric pulse to induce *P. pastoris* to take up clean linearized DNA and allows the yeast to integrate the DNA into the chromosome. The electroporation method was performed using the BTX ECM 630 Electroporator (Harvard

Apparatus, Brookline, MA). 45µL of electrocompetent *P. pastoris* cells were aliquoted into a cold 2.0mm electroporation cuvette. 4µL of linearized DNA (at ≥ 200 ng) was added directly into the sample of cells in the cuvette and then placed on ice for 5 minutes. The cuvette was wiped dry with a kimwipe and placed in the electroporation chamber. The cells were electroporated with set parameters of delivered pulse (200Ω, 1500V, 50µF, 2.0mm) and immediately after electroporation, recovery solution (1,000µL 1M sorbitol) was added to the cells. The cells were transferred to a 1.5mL Eppendorf tube and allowed to recover for 1 hour in the 30°C shaker. After 1 hour, the cells were centrifuged at 4,000 rpm for 30 seconds, 800µL of supernatant was discarded, and the cells were suspended in the remaining 200µL supernatant. 100µL of cells suspension was plated onto YND plates incubated for 48-72 hours in the 30°C incubator. Colonies that arose after 2-3 days were purified by single streaking.

Deep Well Plate and Expression of BLG in WT and *bgs13* Strain of *P. pastoris*

96-well plates were used to express BLG in WT and *bgs13* mutant strains of *P. pastoris*. The deep well plate was performed using the following procedure. 500µL of sterile water was added to all the perimeter wells of sterile 96 deep well plates. Selected strains from YND plates were inoculated in triplicates with 250µL of 1X BMD (buffered minimal dextrose: 1.34% yeast nitrogen base, 4×10^{-5} % biotin, 200mM potassium phosphate buffer, pH 6.0 and 1% dextrose). Deep well plates were sealed with Parafilm® and grown at 28°C shaker for 48 hours at 325 rpm. After 48 hours, the cells were checked for glucose presence by placing 5µL of culture onto Diastix test strips (Bayer, Pittsburgh, PA). The Diastix test strips should be negative for glucose presence. Then, 250µL of 2X BMMY (0.5% methanol, 2% peptone, 1% yeast extract, 1.34% YNB, 4×10^{-5} % biotin, 100mM potassium phosphate, pH 6.0) was added to the wells, and cells were allowed to continue growing for 24 hours. After 24 hours, 50µL of 10X BMM (buffered

minimal methanol 10-fold; 1M K₂HPO₄ and 1M KH₂PO₄ with pH 6.0, 10X YNB, 500X biotin, and 100% methanol) was added to the wells, and cells were placed back in the 28°C shaker to grow for another 24 hours. After a total of 48 hours, cells were harvested, and the OD₆₀₀ values at a 1:20 dilution were taken. To harvest the cells, the cells were resuspended 3-4 times in the wells before transferring them to a 1.5mL Eppendorf tube. Cells were then centrifuged briefly, and supernatants were frozen at -80°C for future analysis.

Spot Western Assay

Frozen, harvested cell cultures were removed from -80°C and thawed in a 37 °C bath briefly, centrifuged at 13,000 rpm and then kept on ice. A membrane was prepared using a piece of Whatman paper and a nitrocellulose membrane cut to the same size. 75µL of sample were spotted onto nitrocellulose using a vacuum spot blot unit from Topac (Bristol, United Kingdom). The membrane was processed using the SNAP i.d®. Protein Detection System from Millipore (Billerica, MA) following manufacturer's instructions. Briefly, the spotted nitrocellulose was allowed to dry in a 60C oven for 5 minutes. Afterwards, the membrane was soaked in 1X PBS and placed on the blot holder, followed by a spacer sheet. The closed blot holder was placed onto the SNAP i.d®. apparatus attached to a vacuum. With the vacuum turned on, 30 mL of I-Block (ThermoFisher Scientific, St. Louis, MO) solution (1xPBS, 0.2% I-Block, 0.1% Tween 20) was added to the membrane. The vacuum was turned off before incubating the membrane with 5mL of I-Block solution containing the appropriate amount of the primary antibody, which was 25µL mouse anti-myc or 10 µL of anti-BLG Santa Cruz biotechnology, Santa Cruz, California). After 10 minutes of incubation with the primary antibody, the vacuum was turned on again, and the membrane was washed three times with a total of 100 mL wash buffer (1X PBS, 0.1% Tween 20). The membrane was then incubated in 5 mL of I-Block containing 2uL of

the goat anti-mouse secondary antibody-alkaline phosphatase (Applied Biosystems, Foster City, CA) for 10 minutes with the vacuum off. Afterwards, the membrane was washed three times with wash buffer again, as previously described, before it was incubated twice for 5 minutes in a petri dish containing 20 mL of femtoLUCENT™ PLUS-AP 1X TBST (G-Biosciences, St. Louis, MO). A flat plastic-wrap surface was prepared where 1-2 mL of the femtoLUCENT™ PLUS-AP detection reagent was added dropwise onto the membrane. After 5 minutes of incubation at room temperature, the detection reagent was drained off. The membrane was placed into a plastic envelope or wrap and developed using the Bio-Rad ChemiDoc XRS+ Imaging System (Hercules, CA) with exposure times of up to 3 minutes. Signals were quantified with ChemiDoc™ XRS+ system (Bio-Rad, Hercules, CA).

SDS-PAGE

20µL of each protein sample was mixed with 4µL of 6x protein sample buffer (PSB) (375 mM Tris-HCl, 9% SDS, 50% Glycerol, 0.03% Bromophenol blue). The samples were vortexed, centrifuged briefly, and heated at 100°C for 10 minutes with sure locks; meanwhile, the running apparatus was assembled. A 12% premade 1.0mm x10 well Tris-Glycine gel (Bio-Rad, Hercules, CA) was locked in the running apparatus and placed into the running tank. The tank was filled with 1L of 1X SDS-PAGE gel running buffer (3g/L Tris base, 14g/L glycine, 1g/L SDS). Each well was rinsed out using SDS-PAGE running buffer to remove any bubbles or impurities in the wells. The heated samples were then allowed to cool at room temperature, centrifuged, and loaded into the wells of the gel. The gels ran at 150V for 65 minutes.

Coomassie Blue Protein Staining

After SDS-PAGE, the gels were transferred to a large petri dish and washed 3 times for 10 minutes with MilliQ water. 20mL of GelCode Blue Stain Reagent solution (Thermo

Scientific, Rockford, IL) was added to the petri dish, which was placed on a tippy tilt overnight. Next day, the gel was destained with MilliQ water 3 times for 10 minutes, and visualized using the ChemiDocTM XRS+ system (Bio-Rad, Hercules, CA).

Western Blot Analysis

SDS-PAGE was ran following protocol mentioned above. Once the gel was run, the manufacturer's instructions for the iBlot system (Invitrogen, Carlsbad, California) were followed for transferring proteins from the gel to a nitrocellulose membrane. The same steps used to detect proteins for the spot western were utilized for the western blots.

Large Scale Expression of BLG and Purification Using Cobalt Affinity Column.

Large scale expression was carried out for BLG produced in WT and *bgs13* strains. Cultures were first grown overnight in 5 ml YPD medium at 28°C in a shaking incubator at 300 rpm. On the second day, 2.5 ml of the YPD culture was used to inoculate 25 ml of 1X BMD (buffered minimal dextrose: 1.34% yeast nitrogen base, $4 \times 10^{-5}\%$ biotin, 200mM potassium phosphate buffer, pH 6.0 and 1% dextrose). On the third day, the cells were pelleted for 5 min at $2,000 \times g$ at room temperature. The supernatant was decanted, and the cells were suspended in 50mL of BMMY (buffered minimal medium with 0.5% methanol, 2% peptone, and 1% yeast extract). The cultures were induced for 48 and 72 h at 28°C with shaking (225 rpm), with the addition of methanol to 0.5% (vol/vol) every 24 h. After the OD₆₀₀ of each culture was measured, 1.0 ml was centrifuged at $10,000 \times g$ for 1 min to separate cells from extracellular supernatant, and the supernatants were transferred to new microcentrifuge tubes. The remaining 49mLs were stored in 50mL conical tubes. Pellets and supernatants were immediately frozen and stored at -80°C.

BLG produced by WT and *bgs13 P. pastoris* was purified with the 3mL PIERCE HisPur™ Cobalt Spin Columns (Rockford, IL) according to manufacturer's instructions. The cobalt column was centrifuged at 700xg for 2 minutes to remove storage buffer. The column was then equilibrated with 6mL of 1X equilibration buffer and centrifuged at 700xg for 2 minutes. 20mL of the equilibrated supernatant, 1 part 10X equilibration buffer (500 mM NaH₂PO₄, 3 M NaCl, pH 7.4) and 9 parts of supernatant, was added to the column. The column was sealed with Parafilm® and placed on a tippy-tilt (Thermolyne, Dubuque, Iowa) at 24-degree angle for at least 30 minutes. The column was then centrifuged for 2 minutes at 700xg and the flow-through was discarded. Then, the column resin was washed 3 times with 6mL of 1X equilibration buffer, followed by 700xg centrifugation for 2 minutes and each wash fraction was collected. Finally, each fraction of BLG protein was eluted 3 times, each time with 2mL of elution buffer (50mM sodium phosphate, 300mM sodium chloride, 150mM imidazole; pH 7.4). The fractions were centrifuged at 700xg for 2 minutes and then collected.

Dialysis

Dialysis was performed to decrease the salt and imidazole concentration of the elutant. The Slide-A-Lyzer Dialysis Cassette, 2,000 MWCO (Thermo Scientific, Rockford, IL) was used for the dialysis process. The cassette was first saturated in sterile water to equilibrate the dialysis membrane for 2 minutes. All 6 mL (3 x 2mL elution's) of eluted proteins were injected with a sterile needle into the dialysis cassette. The dialysis cassette was then allowed to rotate in 1L of sterile water for 2 hours in at 4°C. After 2 hours, the sterile water was changed, and the sample was left to dialyze overnight. The next day, the dialyzed BLG protein was collected with a sterile needle. This process was repeated twice but then second time the dialysis was done in 1X

PBS instead of sterile water. The A_{280} of the protein was taken each time to approximate protein concentration.

Circular Dichroism Spectroscopy.

The final dialyzed product in 1X PBS, 1.3 $\mu\text{g}/\mu\text{L}$ WT and 2.4 $\mu\text{g}/\mu\text{L}$ *bgs13*, was diluted in a phosphate buffer (10 mM NaPO_4 , 100 mM NaCl, pH 7.4) and placed in a 200 μL cuvette (Starna Cells, Inc., Atascadero, CA). Table 2 shows the dilutions that were made.

Table 2

Dilutions for Circular Dichroism

Protein/Protein Concentration	Dilution
WT: BLG 1.3 $\mu\text{g}/\mu\text{L}$	20 μL protein / 180 μL phosphate buffer
<i>bgs13</i> : BLG 2.4 $\mu\text{g}/\mu\text{L}$	10 μL protein / 190 μL phosphate buffer
Commercial BLG 1.0 $\mu\text{g}/\mu\text{L}$	30 μL protein / 170 μL phosphate buffer

Note. Analysis was performed in a Jasco J-815 CD spectropolarimeter between 190 nm to 240 nm. The results from circular dichroism were then graphed using Microsoft Excel and analyzed.

GloMelt Protein Stability Assay

The GloMelt Thermal Shift Protein Stability Kit (Biotium, Fremont, CA) was used to examine the temperature-dependent unfolding of BLG according to the manufacturer's instructions. First 1X GloMelt dye was prepared from the stock solution of 200X GloMelt dye. The GloMelt dye was added to either PBS, TE, or water. Next the protein of interest, either BLG or the control protein, was prepared in the buffer of choice at a final concentration of 0.5-2/5 $\mu\text{g}/\mu\text{L}$. Lastly, after the combining of the dye and protein, the reaction was set up in qPCR tubes

or a qPCR plate with an optical seal. The reaction was maintained on ice until it was placed in a thermocycler with ramp rates between 0.01-0.05°C with the melt/dissociation curve being between 25-99°C.

Endoglycosidase Hf Assay

Endo Hf is a recombinant protein fusion of Endoglycosidase H and maltose binding protein that is used to cleave N-linked glycoproteins. Endo H (New England Bio Labs, Ipswich, MA) was used to look at differences in N-glycosylation between WT and *bgs13* produced BLG with the following protocol: 1-20 µg of glycoprotein was combined with 1 µl of 10X Glycoprotein Denaturing Buffer and H₂O (if necessary) to make a 10 µl total reaction volume. The reaction was then heated at 100°C for 10 minutes to denature the glycoprotein. After the denaturing, 2 µl of 10X GlycoBuffer, 5-9 µl H₂O and 1-5 µl Endo H were added to make a total reaction volume of 20 µl. The reaction was then incubated at 37°C for 1 hour, and the results were visualized by Coomassie-stained SDS PAGE.

N-terminal Sequencing

330 pmol of WT and *bgs13* produced BLG was sent to the Molecular Structure Facility at University of California, Davis for an N terminal sequencing analysis. The sample was loaded on to a PVDF membrane and analysed following a detailed protocol available on the Edman Sequencing Analysis website (Protein Gel Electrophoresis and Electroblothing, <https://msf.ucdavis.edu/edman-sequencing-analysis>).

Determination of Melting Temperatures Using CD

Dialyzed and concentrated BLG proteins (as shown in Table 2) were diluted and placed in a 200µL cuvette (Starna Cells, Inc., Atascadero, CA). The cuvette was placed in the circular dichroism instrument (Jasco J-815 CD spectropolarimeter) and analysis of BLG melting

temperatures between 190nm to 260nm was performed. The temperature parameters for the melting scans were set to start recording from 20°C to 85°C. A scan was taken every 5 seconds until the protein was fully denatured. Data was then plotted using Microsoft Excel and analyzed.

Determination of Melting Temperatures Using Tryptophan Fluorescence.

To measure the melting temperature using tryptophan fluorescence, 3 mLs of commercial b-LG, WT b-LG and *bgs13* b-LG were used at concentrations of 3uM in 1XPBS. The protein was prepared in sterile 15 mL conical tubes and placed on ice before beginning the experiment. The water source for the Fluorescence Spectrometer LS 55 (Perkin Elmer., Waltham, MA) was turned on as well as the spectrometer. The Peltier Temperature Programmer PTP1 (Perkin Elmer., Waltham, MA) was also turned on and set at 25°C. Using the FL WinLab software, the following parameters were set excitation wavelength: 280 nm, emission wavelength: 340 nm, integration time: 15 seconds, excitation bandwidth: 5.0 nm, and emission bandwidth: 5.0 nm. The protein sample was then transferred into the cuvette and inserted into the fluorescence spectrometer. After three reads of intensity on FL WinLab, the fluorescence spectrometer was stopped, and the temperature manually increased 1°C. Data was recorded using Microsoft Excel.

Temperature Induced Aggregation Assay

0.03% solutions of both WT and *bgs13* b-LG were created by diluting the dialyzed, concentrated stock solutions in 100 mM sodium phosphate pH 7.4. 10µL samples of each b-LG were heated for 5 minutes at 45°C, 55°C, 65°C and 75°C followed by the immediate addition of 10µL stop solution (4%SDS, 40 mM iodoacetamide pH 7.4). After 10µL of 100 mM sodium phosphate pH 7.4 and 3.5µL of 10x load dye (75% glycerol, 100 mM sodium phosphate pH 6.8, and 0.01% bromophenol blue) were added to each sample (final volume =35µL), 20µL were resolved on a 12% Tris-glycine gel.

CHAPTER 3: RESULTS

The goals of this thesis project were to 1) express a well characterized protein, β -lactoglobulin (b-LG), in both WT and *bgs13* mutant *P. pastoris* 2) and examine the differences in secretion, *N*-glycosylation, primary, secondary, and tertiary structure between the two proteins to give an insight into the enhanced secretion phenotype of the mutant *bgs13*. The first goal was accomplished by transforming both strains of *P. pastoris* with a hyperglycosylated form of b-LG gene and inducing the expression of the protein. The second goal was, in part, accomplished by purifying b-LG produced from both strains with cobalt affinity chromatography and then using the isolated protein to perform a) N terminal sequencing to determine primary structure, b) a circular dichroism analysis to examine secondary structure as well as c) a melting temperature analysis to look at tertiary structure. Also, the differences in proteins produced by each strain were further characterized by an Endoglycosidase Hf analysis to detect any differences in glycosylation. The purified proteins will also be used in future mass spectrometry analysis to evaluate any *N*-glycosylation differences between WT and *bgs13* mutant strains of *P. pastoris*.

Construction of the PBI-1 Expression Vector (MATalpha-bLG-myc-his-6)

The commercially available plasmid pBLHIS SX (5925 bp) (Figure 1) contains the *AOXI* promoter and a c-myc epitope and 6xHis tag attached to its *C*-terminus. The b-LG coding sequence (790 bp) had three extra glycosylation sites inserted from a previous study (Kalidas et al., 2001). Figure 2 shows the differences in the mature protein (with signal sequence removed), the wild type protein (containing signal sequence) and the hyperglycosylated protein (S123).

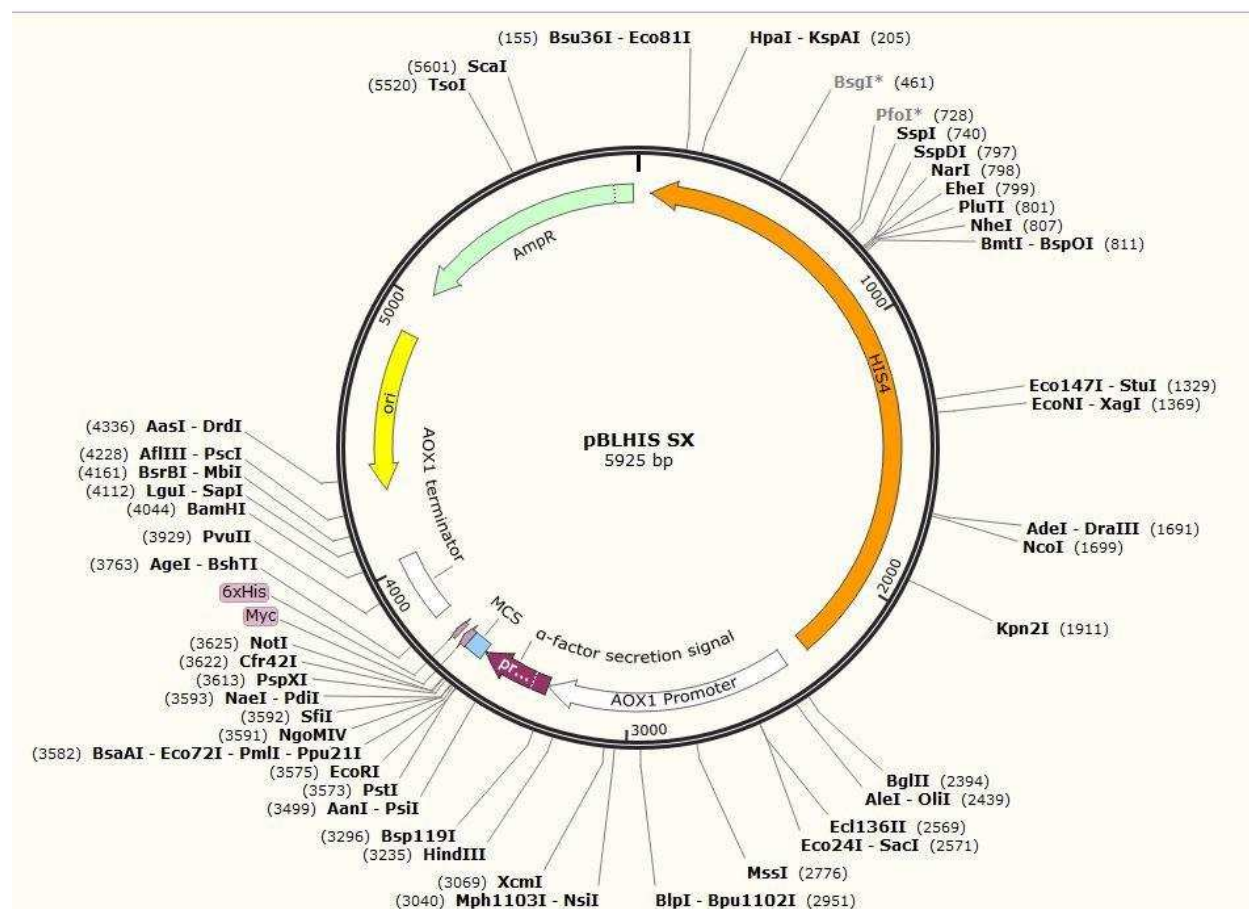
Figure 1*Plasmid Map of pBLHIS SX with Restriction Enzymes*

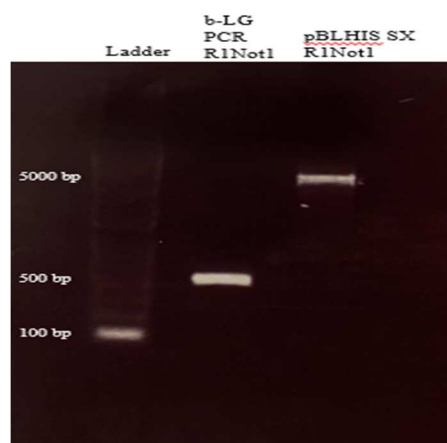
Figure 2*bLG Protein Sequences*

	b-LG Protein Sequence
Genbank 1987 mature protein	IQKVAGTWYSLAMAASDISLLDAQSAPLRVYVEELKPTPEGDLE ILLQKWENDECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDY KKYLLFCMENSAEPEQSLVCQCLVRTPEVDDEALEKFDKALKAL LPMHIRLSFNPTQLEEQCHI
Wild type (Batt, 2001)	LIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYV EELKPTPEGDLEILLQKWENDECAQKKIIAEKTKIPAVFKIDALN ENKVLVLDTDYKKYLLFCMENSAEPEQSLVCQCLVRTPEVDDE ALEKFDKALKALPMHIRLSFNPTQLEEQCHI
S123 Hyperglycosylated (Batt, 2001)	LIVTQTMKGLDIQKVAGTWYSLAMAAS <u>N</u> ISLLDAQSAPLRVYV EELKPTPEGDLEILLQKWEN <u>NS</u> CAQKKIIAEKTKIPAVFK <u>NY</u> TNE NKVLVLDTDYKKYLLFCMENSAEPEQSLVCQCLVRTPEVDDEA LEKFDKALKALPMHIRLSFNPTQLEEQCHI

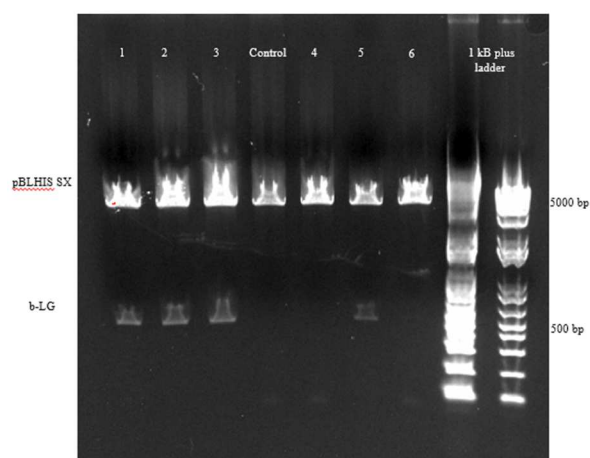
Note. S123 is the name for the hyperglycosylated form of the protein. Mutated residues are shown in red. Underlined N residues indicate asparagine's that are hyperglycosylated.

A colony polymerase chain reaction (PCR) was performed to amplify the b-LG coding sequence (790 bp) which had three extra glycosylation sites using a combination of either a long forward and reverse primer (5bLGEcoR1long + 3bLGNot1) or a short forward and reverse primer (5bLGEcoR1short + 3bLGNot1) (Table 1) from cells containing this gene. The differences between the long forward and short forward primer were that the long forward primer consisted of DNA sequence encoding extra amino acids at its 5' end (Table 1). A flash gel of the PCR results showed that the short forward primer paired with the reverse primer gave a strong band signal of the correct size for b-LG (data not shown). Later results showed that the long forward and short forward primer gave the same results on PCR. The short forward primer PCR product was cloned into the vector.

The b-LG PCR* band was then purified to be consequently ligated into pBLHIS SX. pBLHIS SX was first transformed into TOP10 cells for the purpose of amplifying the plasmid and was minipreped to extract and purify the plasmid DNA. Restriction digestion was then performed on both the b-LG PCR fragment and pBLHIS SX using the restriction enzymes *NotI* and *EcoRI*. The band sized for the b-LG PCR and pBLHIS SX are ~500bp and ~5,000bp respectively, as seen by agarose gel electrophoresis (Figure 3). Following this, the purified digested plasmid and insert were mixed and ligated. The ligation mix pBLHIS SX-bLG was transformed to One Shot Chemically Competent TOP10 *E.coli* cells and plated on LB + Amp. 6 transformants were chosen and cultured in LB + Amp medium. Following this growth, minipreps were performed and the expected plasmid construct, designated pBI1. The pBI1 DNA was digested again with the restriction enzymes *NotI* and *EcoRI* to be later ran on a gel. An empty pBLHIS SX plasmid not containing the b-LG insert was also digested with the same enzymes as a negative control (Figure 4). Agarose gel electrophoresis shows that samples 1, 2, 3, and 5 contain the correct band lengths whereas sample 4 and 6 samples only contained one band for the plasmid, indicating that there is no b-LG insert in these samples (Figure 4). Another restriction digestion was performed to confirm these results using the restriction enzymes *MphI103I* and *BamHI*. The expected band sizes were 5,000 + 1,100 bp for pBLHIS SX and 5,000 + 1,600 bp for pBI1. The digests were electrophoresed and showed that samples 1, 2, 3, and 5 had the correct band sizes for pBI1 (Figure 5). It was confirmed through sequencing that pBI1 plasmid contained the DNA sequence of b-LG in the proper orientation and fused in frame to the N-terminal alpha-factor secretion signal, under the control of the inducible *AOXI* promoter, and in frame with a C-terminal c-myc epitope and polyhistidine sequence (Figure 6).

Figure 3*b-LG PCR Fragment and pBLHIS SX*

Note. b-LG PCR fragment and pBLHIS SX were digested with EcoR1 and Not1 restriction enzymes and resolved on a 1.2 % agarose gel.

Figure 4*Samples of Plasmids*

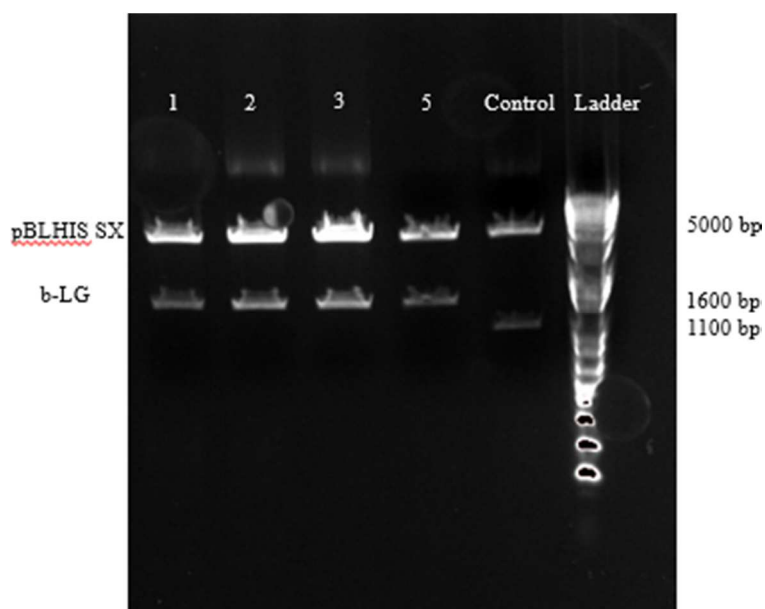
Note. The first 7 lanes contain samples 1-6 as well as the negative control which is an empty plasmid (pBLHIS SX). Lane 8-9 contains the 1 kB plus ladder. Samples 1, 2, 3, and 5 seem to

(Figure 4 Continued)

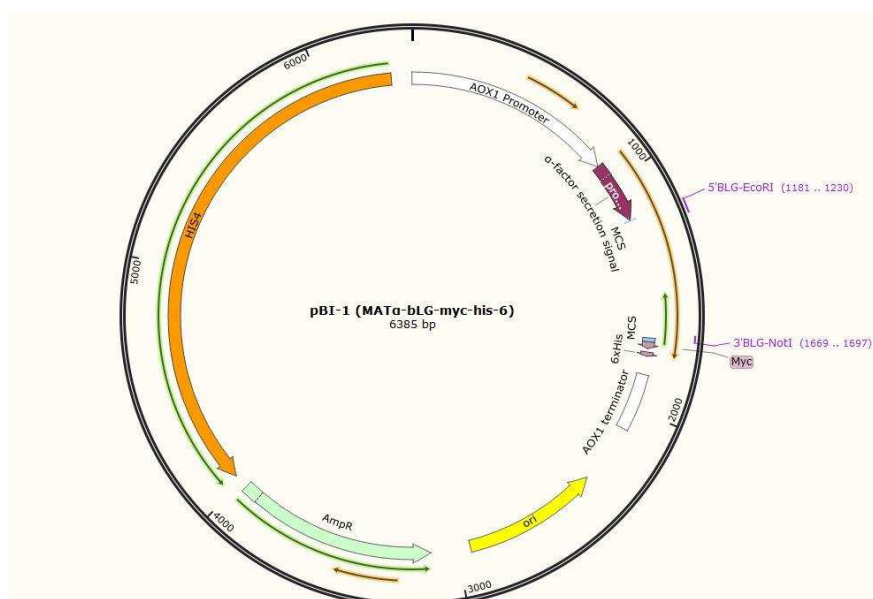
have both fragments of pBLHIS SX and b-LG insert. Sample 4 and 6 do not have the b-LG insert.

Figure 5

Restriction Digestion with Mph1103I and BamHI



Note. Samples 1, 2, 3 and 5 have the correct band sizes for pBLHIS SX and b-LG insert. pBLHIS SX “control” produced the correct band sizes for a plasmid containing no insert.

Figure 6*Plasmid map of pBI1*

Transformation of PBI1 into WT and *BGS13 P. pastoris*

To transform the plasmid into competent *P. pastoris* cells, pBI-1 was linearized with *StuI* restriction enzyme. The linearized pBI-1 DNA was purified and electrophoresed on a gel to confirm that it was linearized and had the size of ~6385 bp (data not shown). The wild type *P. pastoris* strain yGS115 (*his4*) and the *bgs13(his4)* mutant strain were transformed with linearized pBI-1 by electroporation. Once transformed, the transformants were plated on YND plates to select for histidine prototrophy (HIS +) and transformants were purified.

Small Scale Expression of BLG

To express bLG from both WT and *bgs13* mutant strains, a small deep well plate assay was performed. An empty plasmid (pBLHIS SX without the b-LG insert) was transformed into both WT and *bgs13* cells, and these strains were used as negative controls. As a positive control

a *P. pastoris* strain secreting SLPI (secretory protease leukocyte inhibitor), a protein containing a c-myc epitope + His6x tag, was used. Cultures were grown for 2 days in BMD (buffered minimal medium with glucose), then shifted to BMMY (buffered complex medium with methanol) to trigger b-LG expression, and then harvested after 48 hours. Methanol turns on the *AOXI* promoter, which should ultimately result in the production and secretion of b-LG to the extracellular medium. Equal amounts of extracellular medium were taken from each well, spotted onto nitrocellulose for spot western detection. The mouse anti-myc primary antibody followed by the goat anti-mouse secondary antibody-alkaline phosphatase conjugate were used for detection (Figure 7), which showed that several wild type and *bgs13* strains successfully secreted the b-LG reporter.

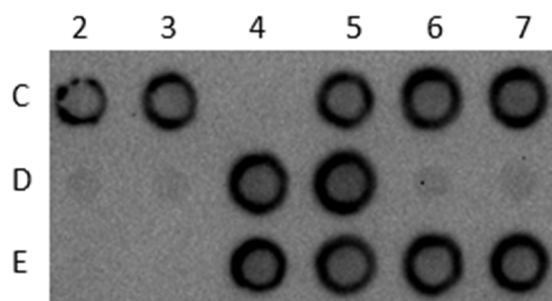
Following the small deep well plate assay another deep well plate assay was performed with more strains. The same protocol was followed as above. Following the deep well plate assay a spot western analysis was performed and confirmed that b-LG was being secreted from both *bgs13* and WT strains. SDS-PAGE was also performed to help further confirm the results of the spot western analysis and visualize the protein bands. The gel was stained with Coomassie blue and all four samples of *bgs13*: pBI1 (lanes 5-8) and a single sample of WT:pBI1 (lane 3) displayed the correct band size for b-LG of about 30 kD (Figure 8). Lane 2 shows the b-LG made by the original S123 strain which appears to be smaller than the b-LG produced by the pBI1 strains (Figure 8). This difference in size will be addressed in later experiments.

A western blot was also performed using the mouse anti-myc primary and anti-mouse-AP secondary antibodies on the same samples shown in figure 8. The western showed that the 30 kD bands detected by the Coomassie stain in figure 8 are also bound by the anti-myc antibody (Figure 9). This was expected because the recombinant b-LG was engineered to contain a C-

terminal c-myc epitope. The bands looked like smears which indicates that there may be multiple bands probably from heterogeneous glycosylation (Figure 9). A second western blot analysis was done to further confirm that the 30 kD band detected by Coomassie was also bound by an anti-b-LG antibody. For this western, an anti-b-LG antibody was used at a 1:1000 dilution after spot westerns indicated that this was the optimal dilution of antibody (data not shown). The western blot showed that the anti-b-LG antibody detected a smear of 30 kD proteins in the extracellular media of both WT and *bgs13* strains (Figure 10). Taken together, the data in figures 8-10 identify the 30 kD bands detected by Coomassie as b-LG proteins containing a myc tag. Given the low sensitivity of Coomassie stain, we conclude that we have created WT and *bgs13* strains which are secreting b-LG at relatively high levels.

Figure 7

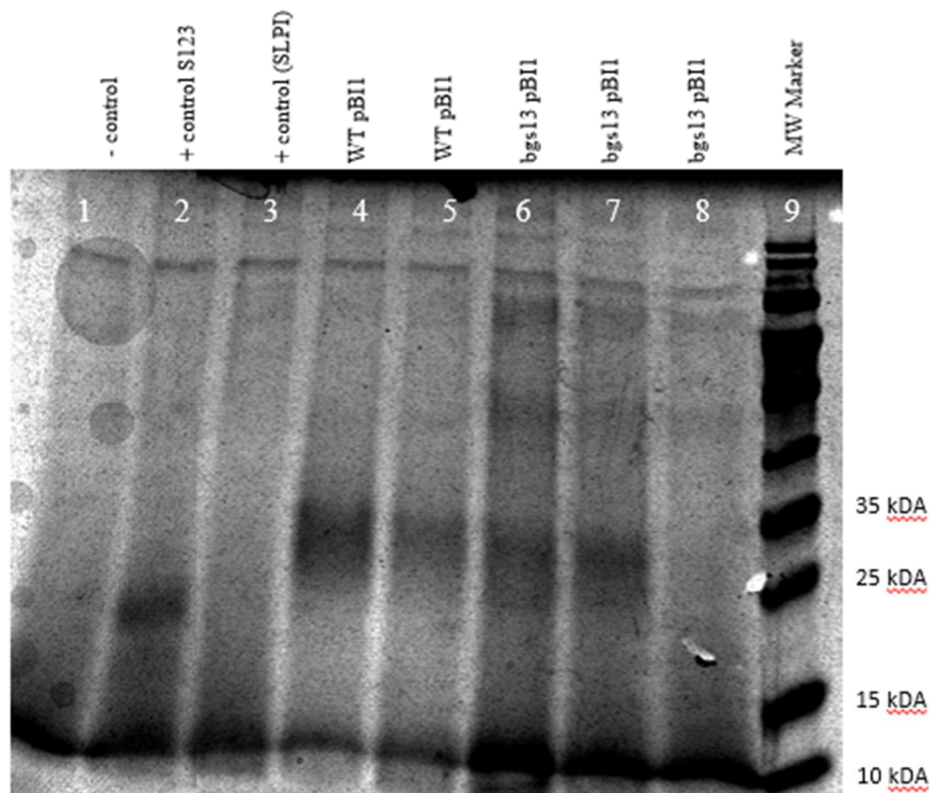
Spot Western Results with Anti-myc Primary Antibody



Note. C2-C3, E6-E7 show secretion of b-LG from WT *P. pastoris*. D4-D5, E4-E5 show secretion of b-LG (containing the myc epitope) from *bgs13*. C2-C3=WT pBI1-1, C4-C5=WT pBI1-2, C6-C7= *bgs13* pBLHIS SX, D2-D3= *bgs13* pBI1-1, D4-D5= *bgs13* pBI1-2, D6-D7= *bgs13* pBI1-3, E2-E3=WT pBI1-3, E4-E5= *bgs13* pBI1-4, E6-E7=WT pBI1-5.

Figure 8

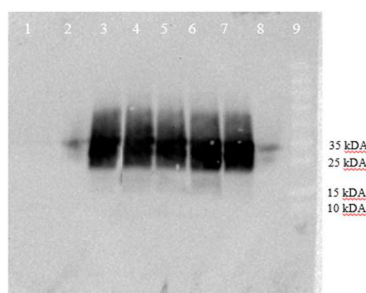
Coomassie Blue Stain of Extracellular Matrixes of Samples from Deep Well Plate Assay



Note. Lanes 4-7 contain the expected band for b-LG of about ~30 kDA. Lane 1= yJC100:pPICZ, Lane 2= b-LG S123, Lane 3= yJC100:pABU1, Lane 4= WT pBI1-1, Lane 5= WT pBI1-2, Lane 6= *bgs13* pBI1-1, Lane 7= *bgs13* pBI1-2, Lane 8= *bgs13* pBI1-3, Lane 9= Molecular weight marker.

Figure 9

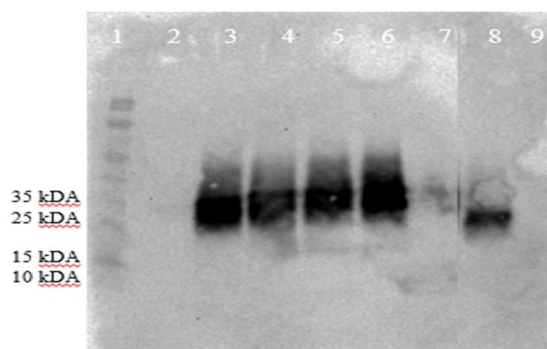
Western Blot Using Mouse Anti-myc Primary and Goat Anti-Mouse Secondary Antibody



Note. Lane 4&5 are b-LG from WT *P. pastoris* and Lane 6&7 are b-LG from *bgs13*. Lane 1=negative control, Lane 2= positive control, Lane 3=yJC100: pABU1, Lane 4= WT pBI1-1, Lane 5=WT pBI1-2, Lane 6= *bgs13* pBI1-1, Lane 7= *bgs13* pBI1-2, Lane 8= positive control, Lane 9= Protein marker.

Figure 10

Western Blot Using Anti-bLG Antibody



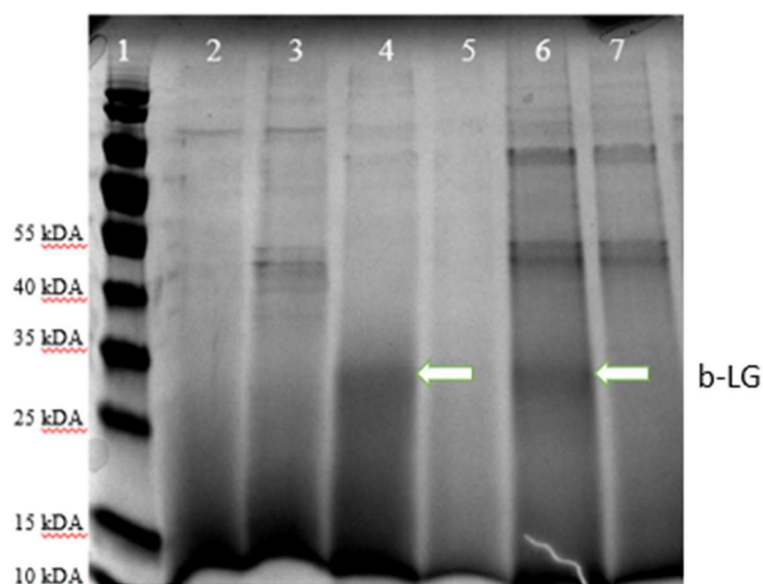
Note. Lane 3&4 *bgs13* b-LG and Lane 5&6 are WT b-LG. Lane 1= protein marker, Lane 2= negative control, Lane 3= *bgs13* pBI1-1, Lane 4= *bgs13* pBI1-2, Lane 5=WT pBI1-1, Lane 6= WT pBI1-2, Lane 7: yJC100: pABU1, Lane 8= b-LG S123, Lane 9= negative control.

Large Scale Expression

After successfully expressing b-LG on a small scale, we moved on to large scale expression. Large scale expression of b-LG was confirmed by a Coomassie blue stain. Lanes 4 and 6, which have higher levels of expression, show secreted WT b-LG and *bgs13* b-LG respectively (Figure 11). The b-LG band was stronger in lane 6, which was expected, because *bgs13* has a super-secretion phenotype (Figure 11). To further understand the mechanisms of *bgs13* and its effects on secretion, the next step was to purify the proteins and perform examine the structures.

Figure 11

Coomassie of Large-Scale Expression Showing b-LG Band at About ~30 kDA in lanes 4 and 6



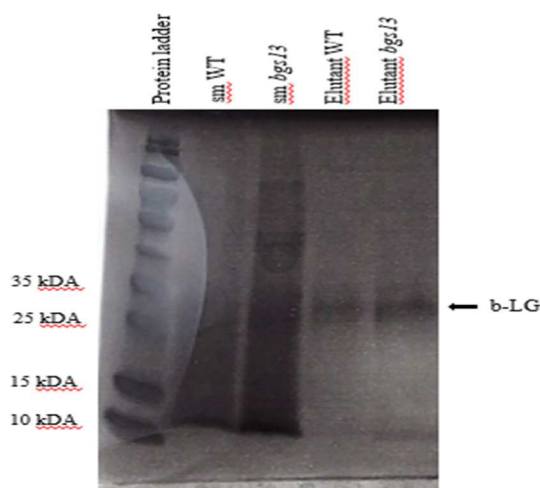
Note. Lane 1= Molecular Weight marker, Lane 2= positive control, Lane 3=negative control, Lane 4=WT pBI1-1, Lane 5=WT pBI1-2, Lane 6=*bgs13* pBI1-1, Lane 7=*bgs13* pBI1-2.

HisPur Cobalt Column Protein Purification

As mentioned previously, β -lactoglobulin (b-LG) contained the C-terminal His6 tag which allowed for purification using a cobalt spin column. While nickel (Ni^{+2}) is commonly used in many immobilized metal affinity chromatography (IMAC) resins to isolate His-tagged proteins, the purity of the protein is sometimes not optimal and further optimization is required (Thermo Scientific, Instructions 1852.3). We chose cobalt due to its ability to achieve higher protein yield as well as purity. Following purification, the samples were analyzed by SDS PAGE and stained with Coomassie to visualize the purified proteins. Figure 12 shows the Coomassie stained SDS PAGE of b-LG from WT *P. pastoris* and b-LG from *bgs13* which has the starting material (sm) before protein purification alongside the purified proteins which were eluted. Large scale expression was performed twice and both times measured OD values were slightly higher for *bgs13*. This indicates that *bgs13* produced higher yields of b-LG compared to the WT which confirms that *bgs13* yields higher levels of protein.

Figure 12

*Coomassie Stained SDS PAGE of Purified b-LG from WT *P. pastoris* and b-LG from *bgs13**



(Figure 12 Continued)

Note. Lane 1= Protein Ladder, Lane 2= starting material WT, Lane 3= starting material *bgs13*, Lane 4= purified b-LG elutant WT, Lane 5= purified b-LG elutant *bgs13*.

Dialysis of b-LG

Following the first large scale protein production and purification, b-LG from WT and *bgs13* were dialyzed in water to remove excess imidazole, which could potentially interfere with circular dichroism (CD) and other future experiments. However, after doing a CD scan, we found that our protein became unfolded in water (see section below). To fix this problem, we re-dialyzed our second large scale protein production batch against 1X PBS, which should induce folding of the protein. The data later showed that it was folded in 1X PBS (data is shown in next section). We concluded that the protein needs ionic strength, which the 1X PBS was providing, to remain folded.

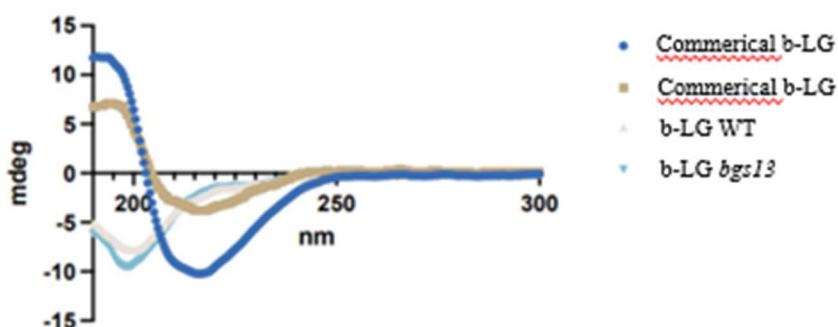
CD Spectroscopy

The different preparations of purified, dialyzed b-LG protein were used for a CD spectroscopy analysis. CD spectra examines the secondary structure of proteins and measures the alpha and beta helix sheets that make up the protein. As a control, commercial bovine b-LG was used at different concentrations. Our first CD scan in Figure 13 shows b-LG produced by WT and *bgs13* strains, which was dialyzed against water, as well as commercial bovine b-LG. Compared to the commercial bovine b-LG dissolved in PBS, which shows more of a beta helix sheet structure (as demonstrated by a minimum at 217 nm) b-LG from WT and *bgs13* show random coils (as demonstrated by a minimum at less than 200 nm). The reason for this difference was that the WT and *bgs13* protein used during this first CD scan was dialyzed against

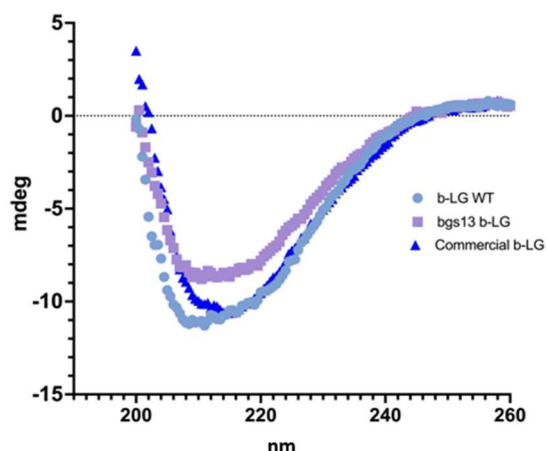
water, causing it to become unfolded. This experiment was performed again using b-LG that was dialyzed against 1X PBS. Figure 14 shows the CD scan of the folded WT and *bgs13* b-LGs, which were dialyzed against PBS, are now similar to commercial b-LG, displaying more beta helix sheets. The secondary structure of *Pichia* produced b-LGs seem similar to each other; there are no obvious differences in structure. We can conclude that the *bgs13* mutation does not cause any identifiable difference in the secondary structure of the protein according to the CD.

Figure 13

CD Graph Showing Two Different Concentrations of Commercial b-LG at 1.0 μ g/uL and 0.5 μ g/uL Respectively



Note. Commercial b-LG is showing beta sheets whereas, b-LG from WT and *bgs13*, which were dissolved in water, show a random coil due to being unfolded.

Figure 14*CD Spectra*

Note. CD spectra showing commercial b-LG along with the b-LG from WT and *bgs13* at the concentrations of 1.0 μ g/uL, 1.3 μ g/uL and 2.4 μ g/uL respectively, which were dissolved in 1X PBS. b-LG from WT and *bgs13 P. pastoris* has the same beta sheet structure as commercial b-LG.

Melting Temperature Using GloMelt

GloMelt is a dye that undergoes fluorescence enhancement when binding to hydrophobic regions of proteins that are denatured. It is useful to detect protein unfolding or measure thermal stability by performing a thermal shift assay (Biotium, 2019). A commercially available kit was used in this experiment with IgG as a control protein. IgG worked well in water, PBS, and TE as buffers; however, we did not see any unfolding of b-LG in any three of the buffers (data not shown). We concluded that this assay did not work for our protein.

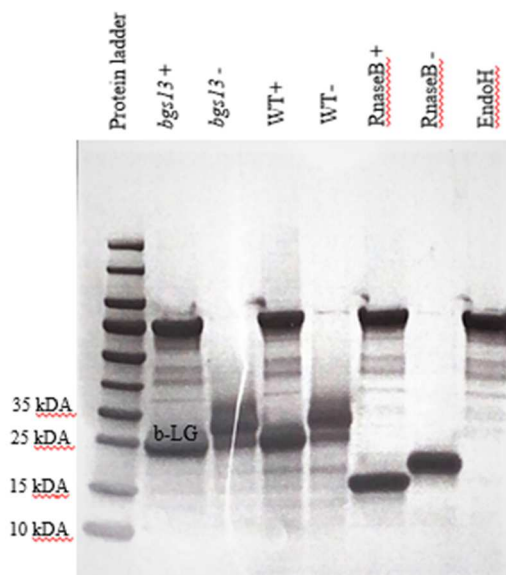
EndoHf Assay

To examine the effect of *N*-glycosylation on the size of b-LG, we performed an EndoHf assay. EndoHf cleaves the chitobiose core of high mannose from *N*-linked glycoproteins (New

England Biolabs). Its activity is identical to Endo H. Rnase B was used as a control in the experiment. Figure 15 shows Rnase B with the addition of Endo H and without Endo H. We found that while the b-LG protein size did decrease with *N*-glycosylation removal, both b-LGs were still larger and exceeded the predicted size for b-LG with an addition of c-myc and His6x, which is about 21 kDA. Bovine b-LG is about ~18 kDA, and the c-myc + His tag account for about ~3 kDA. Previous results showed that the b-LG from both WT and *bgs13* is about ~30 kDA. Thus, while we expected EndoHf cleavage to produce a protein of around ~21 kDA the actual digested b-LGs were about 25 kDA (Figure 15), suggesting the presence of some other posttranslational modification.

Figure 15

EndoHf Assay of the WT and bgs13 b-LG Proteins with RnaseB and EndoHf as Controls



(Figure 15 Continued)

Note. Lane 1= protein ladder, Lane 2= *bgs13*⁺=b-LG +EndoH, Lane 3=*bgs13*⁻=b-LG with no EndoH, Lane 4= WT+=b-LG + EndoH, Lane 5=WT=b-LG with no-EndoH, Lane 6=RnaseB+ EndoH, Lane 7=RnaseB- EndoH, Lane 8= EndoH.

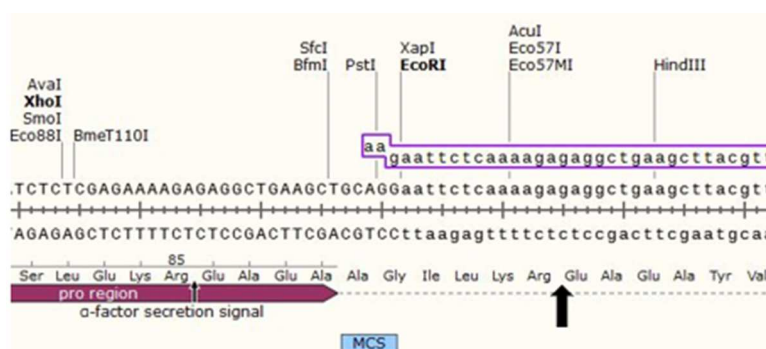
N-Terminal Sequencing

The large size of the protein even after cleavage of the *N*-glycosylation groups raised the possibility that the N-terminal end of the recombinant b-LG might have extra amino acids because the MAT α signal peptide was not fully cleaved during the secretion process at the Lys-Arg sequence (Cregg et al., 1997). This could potentially explain the greater molecular weight of the protein. To examine this possibility, we sent our purified b-LG protein from WT and *bgs13* strains to the Molecular Structure Facility at UC Davis for N-terminal, also known as Edman, sequencing. The predicted amino acid sequence at the N-terminal is shown in Figure 16 as well as pro region of MAT α signal peptide which if not cleaved could explain the large size of the protein. Results, as presented in Table 3, showed that both proteins had the expected amino acid sequence at the N-terminus; thus, the MAT α signal peptide had been cleaved from both WT and *bgs13* proteins as expected at the Lys-Arg sequence, and extra residues from the MAT α signal peptide could not account for the bigger size of the proteins (Table 3).

Table 3*Amino Acid Sequence at N-terminus*

	Predicted sequence at N-terminus	Sequencing results of N-terminus
WT	Glu-Ala-Glu-Ala-Tyr-Val	Glu-Ala-Glu-Ala-Tyr-Val
<i>bgs13</i>	Glu-Ala-Glu-Ala-Tyr-Val	Glu-Ala-Glu-Ala-Tyr-Val

Table 3: The predicted amino acid sequence at the N-terminus after the MAT α signal peptide removal (Batt, 2001) and the N –terminal sequencing results of both WT and *bgs13* b-LG proteins.

Figure 16*Predicted Amino Acid Sequence at the N-terminus of b-LG*

Note. Bold arrow is showing where the MAT α signal peptide should have been cleaved by a Kex2 protease located in the endoplasmic reticulum.

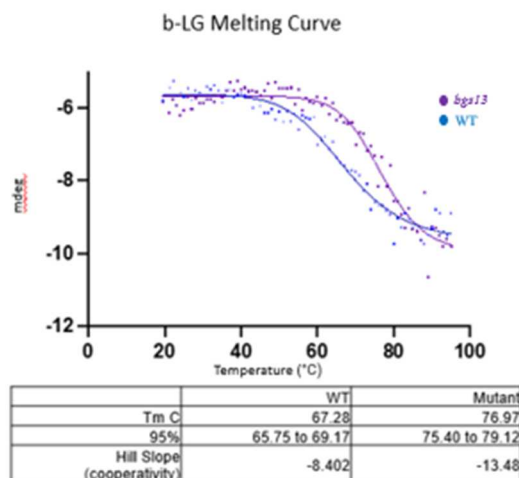
Melting Temperature Using CD

To examine tertiary structure of the b-LG peptides, we analyzed their melting temperature using a modified circular dichroism (CD) procedure. The change in CD as a

function of temperature, at certain wavelengths, allows for the determination of the midpoint of the unfolding transition or melting temperature T_m (Greenfield, 2006). We measured the unfolding of the protein between 190 nm to 260 nm, beginning the melting at 20°C (Figure 17). The results were surprising because not only was the published melting temperature of native bovine b-LG about 63°C, but also because the melting curve was expected to curve upward with increasing temperature as denaturation usually triggers an increase in absorbance by newly exposed amino acids. Our results could be interpreted in two ways: either a) there is a 10-degree difference in the T_m of the WT and *bgs13* b-LG proteins, 67°C vs. 77°C. Our data curved downwards because many beta proteins (those rich in beta sheets) like b-LG exhibit a decrease in absorbance with the loss of beta-pleated sheet. Based on this hypothesis, our results suggest that the unfolding of the b-LG from *bgs13* was more cooperative compared to the b-LG from WT *P. pastoris*; or b) there is not a significant difference in the T_m s of the b-LG proteins but the decrease in CD absorbance stems from the appearance of aggregation, meaning that the denatured proteins can associate with each other to form polymers. Previous studies have demonstrated that native bovine b-LG undergoes simultaneous loss of beta sheet structure and aggregation as it denatures at higher temperatures (Prabakaran et. al, 1997). Based on this hypothesis, our results would indicate that there is a difference in the structure of the two b-LGs that causes the WT protein to aggregate more quickly at a lower temperature than the *bgs13* protein. As a comparison, commercial b-LG shown in Figure 18 did not have a melting curve like the *P. pastoris* produced b-LGs, suggesting a major difference between the bovine and recombinant proteins.

Figure 17

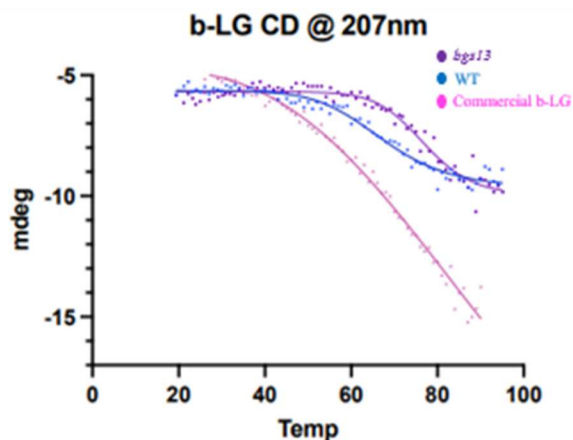
b-LG Melting Curves Showing the Melting Temperature (T_m) of Both Proteins



Note. Both curves start a -6 mdeg which confirms that they were at the same concentration.

Figure 18

b-LG Melting Curves including the Commercial Protein at 207 nm



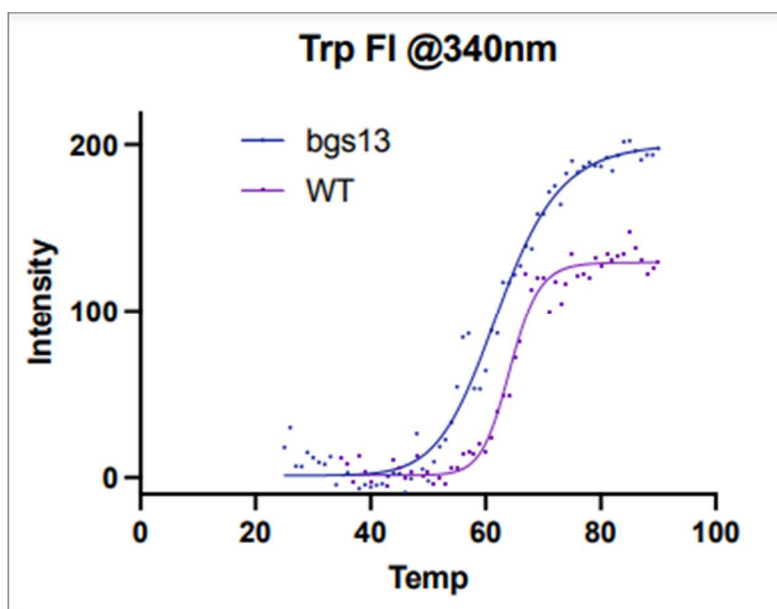
Melting Temperature Using Tryptophan Fluorescence (TRP Fluorescence)

To differentiate between these two hypotheses, tryptophan fluorescence was used to examine melting temperature. This method involves measuring the intensity of the fluorescence that the tryptophan residues of the protein emit as the protein is heated and melts. An advantage

of using tryptophan fluorescence spectroscopy is that trp fluorescence is very sensitive to the conformational changes of a protein such as changes in the secondary and tertiary structure (Garidel et al., 2008). This makes protein fluorescence as useful tool to examine the conformational state of the protein. The trp fluorescence data indicated that the melting temperatures of b-LG from WT and *bgs13 P. pastoris* were relatively similar at 64°C and 62°C, respectively (Figure 19). This would suggest that the difference in the CD melting curves was due to aggregation.

Figure 19

Trp Fluorescence Curve



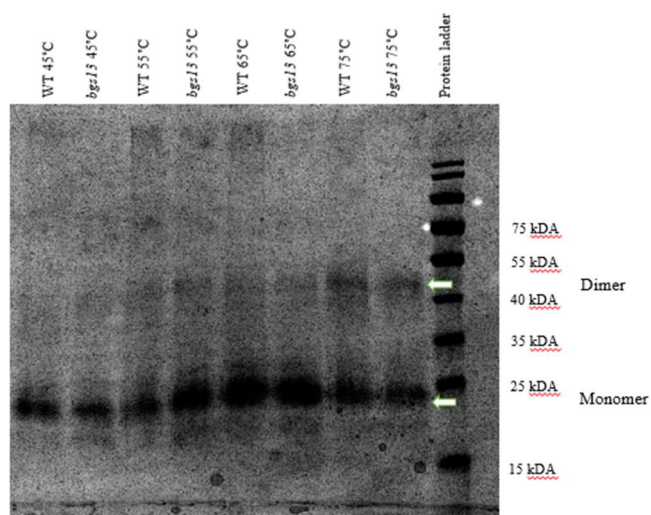
Note. Trp fluorescence curve at 340 nm which is measuring the intensity of the fluorescence as the temperature is increased.

Temperature Induced Aggregation Assay

Bovine-derived b-LG, as it denatures at higher temperatures, has been observed to aggregate from the folded monomer to unfolded dimer, trimer and other polymer forms (Prabakaran et. al, 1997). Interestingly, the b-LG is not completely denatured even at 90°C, suggesting that b-LG maintains some shape, either secondary or tertiary structure, even under high heat (Chattacharjee & Das, 2000). In bovine b-LG, the initiation of the polymerization reaction upon heating involves a critical change in b-LG conformation, which exposes a buried sulfhydryl (SH) in the protein interior and initiates sulfhydryl-disulfide interchange reactions leading to polymerization. To determine if the *P. pastoris* produced forms of b-LG were undergoing aggregation with increased heat, samples of each protein were exposed to increasing temperatures for 5 minutes and then visualized by SDS-PAGE to detect the presence of a dimer form, a sign of aggregation. As shown in figure 20, more of the dimer form appears for both samples as temperature is increased, demonstrating that aggregation is occurring. Moreover, after 75°C treatment, a greater proportion of the WT b-LG is in the dimer form than the monomer form compared to the *bgs13* counterpart (Table 4), suggesting that the *bgs13*-produced b-LG is aggregating less than the WT protein. Taken together, this aggregation and trp fluorescence results indicate that the behavior of the WT and *bgs13* proteins in CD was caused by some difference affecting aggregation, not simply denaturation, of the proteins. The identity of the structural difference(s) that are responsible for this difference in aggregation are presently unknown.

Figure 20

Coomassie Blue Stain of SDS-PAGE Gel Examining Temperature Induced Aggregation of b-LG



Note. WT and *bgs13* samples of b-LG were heated for 5 minutes at the indicated temperatures to induce aggregation followed by the immediate addition of stop solution. Monomer (non-aggregated) and dimer (aggregated) forms are indicated by arrows.

Table 4

b-LG Dimers and Monomers

	WT	<i>bgs13</i>
Monomer	6.9×10^6	7.8×10^6
Dimer	5.3×10^6	4.8×10^6

Table 4: Quantitative measurements of the amount of dimers and monomers present of b-LG after 75°C treatment.

CHAPTER 4: DISCUSSION

P. pastoris is one of the most popular and well-utilized protein expression systems, that has been used for the successful production of biopharmaceuticals, vaccines, and antibodies. (Ahmad et al., 2014; Chahal, et al., 2016; Naranjo, et al., 2019). Its advantages over other more inferior expression systems include its ability to produce large quantities of correctly folded recombinant proteins with the needed post-translational modifications that are required for optimal function. (Chahal et al., 2016). While it has many advantages, limitations exist when using *P. pastoris* as an expression system. One major limitation is inefficient secretion: recombinant proteins that are engineered to be secreted can be either retained or degraded inside the cell (Idiris et al. 2010). This can potentially be due to the recombinant proteins lacking the correct post translational modifications and folding (Naranjo, et al., 2019).

To help illuminate the secretory mechanisms in *P. pastoris*, gene products that are involved in secretion were identified using a novel strategy known as REMI (restriction enzyme-mediated integration) (Larsen et al., 2013). Of the mutant strains identified, *bgs13* showed increased secretion of many recombinant proteins and raised the possibility of being a universal supersecreter, suggesting that it played a general role in protein export and generated interest in this mutant (Larsen et al., 2013).

The 1,035 amino acid Bgs13 peptide's (Bgs13p) predicted amino acid sequence in wt *P. pastoris* is 50% identical and 68% similar to the *Saccharomyces cerevisiae* (*S. cerevisiae*) protein kinase C 1 protein (Pkc1p) in the N-terminal region (Naranjo, et al., 2019). As mentioned previously, the *S. cerevisiae* Pkc1p plays a crucial role in cell wall integrity (Naranjo, et al., 2019). The mutant Bgs13 protein expressed in the *bgs13* strain, due to the insertion of the pREMI plasmid, lacked the first 147 N-terminal amino acids of the wildtype protein. Instead, it

contained 12 amino acids encoded by the plasmid fused in frame with the last 888 amino acids of the Bgs13 peptide. It was found that the mutant bgs13p had reduced activity and was in a different location in the cell compared to the wild type (Naranjo, et al., 2019). Furthermore, while *bgs13*'s cell wall had structural problems, it was not more porous compared to a wild type cell, so this did not explain its super-secretion phenotype (Naranjo, et al., 2019). *bgs13*'s secretion pathway was further analyzed by examining both the wild type and mutant protein's involvement in the Unfolded Protein Response (UPR).

The *bgs13* strain had an increased level of normally secreted proteins as well as endoplasmic reticulum-resident chaperone proteins, which suggested that the Bgs13p is involved in the UPR as a regulator (Naranjo C. et al., 2019). These chaperone proteins serve in accelerating the folding of the secreted proteins as they transit the secretory pathway. However, it was unclear if these changes were the cause of enhanced protein secretion or a side effect of the mutant bgs13p unrelated to the increased protein export because protein kinase C peptides have been shown to regulate many processes in a cell (Naranjo C. et al., 2019).

Because the *bgs13* strain has similarities as well as differences compared to cells with a decreased UPR, it cannot be simply concluded that the Bgs13p activates or deactivates the UPR (Naranjo, et al., 2019). The UPR has two mechanisms that are used to get rid of ER stress that is brought on by the accumulation of proteins in secretory organelles. The first mechanism is to degrade the misfolded proteins and target them for vacuolar degradation while the second mechanism is to enhance the folding of nascent proteins.

Proposed Models for *BGS13*

Therefore, our lab has currently two proposed models to explain how the mutant bgs13p may be affecting the UPR. The first model suggests that the bgs13p is downregulating the UPR,

which is causing misfolded proteins that are intended for vacuolar degradation to be secreted, leading to the export of nonfunctional (aka “garbage”) peptides. The second model suggests that Bgs13p is upregulating the UPR which enhances its folding capacity to allow more proteins to be released in their functional conformations.

To help distinguish between these two models and illuminate the underlying mechanisms involved in the super secretion phenotype of *bgs13*, we used β -lactoglobulin (b-LG), a well-studied reporter protein, to examine structural differences between *bgs13* and wildtype secreted proteins. If the first model is correct, then it is expected that the b-LG proteins that are secreted by the *bgs13* strain will be less correctly folded than those from the wild type strain. However, if the second model is correct, then the b-LG proteins secreted by *bgs13* should be correctly folded. To determine which model was correct, we analyzed secondary and tertiary structures of the secreted b-LG to reveal the folding efficiency.

Secondary Structure

We used circular dichroism to examine the differences in secondary structure of b-LG produced by *bgs13* and wildtype *P. pastoris*. We found that the secondary structures of the *P. pastoris* produced b-LGs seem similar to each other as well as commercial b-LG, which is isolated from cows. The secondary structure of all three samples was dominated by beta-sheets as expected (Hammann and Schmid., 2014). There was no obvious difference in secondary structure; therefore, we concluded that the *bgs13* mutation does not cause a change in folding that control's secondary structure.

Tertiary Structure

Using melting temperature analysis, we examined tertiary structure. First, we utilized both CD and tryptophan fluorescence as a function of temperature to determine the effects of the *bgs13* mutation on b-LG's stability (Greenfield NJ., 2006). Our CD results suggested that there was a ~10 degree difference between the melting temperatures of WT and *bgs13* b-LG. On the contrary, our tryptophan results suggested that the melting temperature of the two proteins was around 64 °C and 62 °C. 63 °C is where critical structural changes in commercial b-LG are observed (Prabakaran, 1997).

Based on the CD data, we first thought that the *bgs13* b-LG had a more gradual melting curve compared to wt b-LG. Because the curve was more cooperative, this suggested that the *bgs13* b-LG was more stable and more correctly folded. b-LG contains five cysteine molecules, and four out of five of these molecules form two disulfide bonds which help determine the tertiary structure of the protein (Hammann and Schmid., 2014). Since the curve was more cooperative, this also suggested that *bgs13* b-LG may have had more of the disulfide bonds present, which allowed it to be folded into its tertiary structure more correctly. Also, there could be other features, such as ionic and hydrogen bonds between secondary structures, that would contribute this enhanced folding, which allows proteins to transition more smoothly from a nascent to a mature protein.

However, tryptophan fluorescence suggested that the CD results may be due to aggregation of the proteins rather than unfolding. As b-LG denatures at higher temperatures, it has been found to undergo polymerization. In the CD data, as temperature was increased, we also saw an increase in secondary structure upon its unfolding. This may have resulted from aggregation of the protein rather than the melting of the protein. While both WT and *bgs13*

forms aggregate at higher temperatures, WT seems to go from a monomer to a dimer form more quickly. Furthermore, based on the CD results, *bgs13* b-LG may be more stable because it takes a longer time to denature and form aggregates. Based on the temperature aggregation assay, at 75°C, WT b-LG existed less in monomer form and more in dimer form whereas *bgs13* b-LG was found more in monomer form than dimer form. In bovine b-LG, aggregate formation was triggered by a structural change in the monomer that exposes a buried sulfhydryl in the protein interior and initiates sulfhydryl-disulfide interchange reactions leading to polymerization. The b-LG produced by *bgs13* may be less prone to aggregation which suggests that a) it may be more stable and therefore, folded better compared to WT b-LG or b) it may contain a posttranslational modification, such as a unique glycosylation feature, that helps deter sulfhydryl interactions, unlike the WT b-LG. Identifying this posttranslational modification may be the key to understanding the enhanced secretion of this strain.

Conclusion

Taken together, our research supports the second proposed model that *bgs13* is upregulating the UPR which is enhancing its folding capacity. This would increase the number of recombinant proteins in a cell with the correct folding and stability that can be secreted. Our hypothesis is consistent with the previous finding that the endoplasmic reticulum-resident chaperone proteins were overexpressed in the *bgs13* strain (Naranjo C., et al 2019). These chaperone proteins such as, protein disulfide isomerase (PDI), aid in the folding of proteins (Naranjo C., et al 2019). PDI acts as a chaperone protein, and as its name suggests, functions in the formation of disulfide in proteins. PDI also aids in promoting other types of protein bond formation, as it has been shown to help the folding of proteins with no disulfide bonds (Imaoka S., 2011). The overexpression of PDI in *bgs13* partially can explain its enhanced secretion and

folding capacity. In addition, our results with b-LG CD suggest that proteins produced by the *bgs13* strain, unlike those secreted by WT, may contain some posttranslational modification which helps stabilize them and prevent aggregation.

Future Work

Overall, this research generates even further interest in the *bgs13* strain because our data demonstrates that it is producing quality proteins that have the correct folding, not “garbage” proteins. However, it is still unclear what are all the players in the secretory pathway that are causing the proteins to be more correctly folded. Further studies need to examine *N*-glycosylation of proteins produced in *bgs13* to examine if changes in this posttranslational modification are enhancing folding and preventing aggregation. Currently, we are examining the *N*-glycosylation of b-LG by digestion and enzymatic release of glycans using PNGase F. The plan is to then use a chemical labeling procedure and mass spectrometry to analyze the differences in the glycans between the two strains of *P. pastoris*, which will give us insight into the effects of *bgs13* on *N*-glycosylation. Another experiment to determine how *N*-glycosylation is affected by *bgs13*'s secretory pathway would be to delete the Golgi mannosyltransferase gene, which would block the extension of the outer chain to avoid hyperglycosylation, in both wildtype and *bgs13* strains to illuminate the effects on secretion.

Future experiments should also examine if the upregulation of the secretory pathway is caused by the activation of the UPR through the classic IRE1/HAC1 pathway. Hac1p is a transcription factor, which is generated from spliced HAC1 mRNA through the intron removal of unspliced HAC1 mRNA (Tang et al., 2016). Hac1p can regulate the UPR by upregulating many genes in secretory pathway (Tang et al., 2016). The splicing level of HAC1 in *bgs13* could be used as indicator of activation of the UPR and the IRE1/HAC1 pathway.

Understanding the secretion pathway and enhanced folding ability of *bgs13* will provide insight into the secretory mechanism of *P. pastoris*. This will allow for further improvement and use of *P. pastoris* as an optimal host for protein production.

References

- Ahmad M., Hirz M., Pichler H., Schwab. *Protein expression in Pichia Pastoris: Recent achievements and perspectives for heterologous protein production*. Applied microbiology and biotechnology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/24743983/>
- Baneyx, F. *Recombinant protein expression in escherichia coli*. Current opinion in biotechnology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/10508629/>
- Bhattacharjee, C., Das K.P. *Thermal unfolding, and refolding of beta-lactoglobulin. an intrinsic and extrinsic fluorescence study*. European journal of biochemistry. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/10866794/>
- Cereghino, J. L., & Cregg, J. M. (2000, January 1). *Heterologous protein expression in the methylotrophic yeast Pichia Pastoris*. OUP Academic. Retrieved October 24, 2022, from <https://academic.oup.com/femsre/article/24/1/45/525926>
- Chahal S;Wei P;Moua P;Park SP;Kwon J;Patel A;Vu AT;Catolico JA;Tsai YF;Shaheen N;Chu TT;Tam V;Khan ZE;Joo HH;Xue L;Lin-Cereghino J;Tsai JW;Lin-Cereghino GP; (n.d.). *Structural characterization of the α -mating factor prepro-peptide for secretion of recombinant proteins in Pichia pastoris*. Gene. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/27984193/>
- Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E., & Villaverde, A. (2009, March 24). *Microbial factories for recombinant pharmaceuticals - microbial cell factories*. BioMed Central. Retrieved October 24, 2022, from <https://microbialcellfactories.biomedcentral.com/articles/10.1186/1475-2859-8-17>

- Gardiel P., Hegyi M., Bassarab S., Weichel M. *A rapid, sensitive, and economical assessment of monoclonal antibody conformational stability by intrinsic tryptophan fluorescence spectroscopy*. Biotechnology journal. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/18702089/>
- Gellissen G;Melber K;Janowicz ZA;Dahlems UM;Weydemann U;Piontek M;Strasser AW;Hollenberg CP; (n.d.). *Heterologous protein production in yeast*. Antonie van Leeuwenhoek. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/1444338/>
- Greenfield, N. *Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions*. Nature protocols. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/17406506/>
- Hammann, F., & Schmid, M. (2014, December 10). *Determination and quantification of molecular interactions in protein films: A Review*. Materials (Basel, Switzerland). Retrieved October 24, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5456426/>
- Idiris, A., Tohda H., Kumagai, H., Takegawa, K. *Engineering of protein secretion in yeast: Strategies and impact on protein production*. Applied microbiology and biotechnology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/20140428/>
- Jäger, V., Büssow, K., & Schirrmann, T. (1970, January 1). *Transient recombinant protein expression in mammalian cells*. SpringerLink. Retrieved October 24, 2022, from https://link.springer.com/chapter/10.1007/978-3-319-10320-4_2

- Jenkins N., Murphy L., Tyther R. *Post-translational modifications of recombinant proteins: Significance for biopharmaceuticals*. Molecular biotechnology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/18327554/>
- Karbalaee M., Rezaee S.A., Farsiani H. *Pichia pastoris: A highly successful expression system for optimal synthesis of heterologous proteins*. Journal of cellular physiology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/32057111/>
- Kim TR;Goto Y;Hirota N;Kuwata K;Denton H;Wu SY;Sawyer L;Batt CA; (n.d.). *High-level expression of bovine beta-lactoglobulin in Pichia pastoris and characterization of its physical properties*. Protein engineering. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/9514124/>
- Kost, T. A., Condreay, J. P., & Jarvis, D. L. (2005, May 5). *Baculovirus as versatile vectors for protein expression in insect and mammalian cells*. Nature News. Retrieved October 24, 2022, from <https://www.nature.com/articles/nbt1095>
- Lackner A;Genta K;Koppensteiner H;Herbacek I;Holzmann K;Spiegl-Kreinecker S;Berger W;Grusch M; (n.d.). *A bicistronic baculovirus vector for transient and stable protein expression in mammalian cells*. Analytical biochemistry. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/18541133/>
- Larsen, S., Weaver, J., de Sa Campos, K., Bulahan, R., Nguyen, J., Grove, H., Huang, A., Low, L., Tran, N., Gomez, S., Yau, J., Ilustrisimo, T., Kawilarang, J., Lau, J., Tranphung, M., Chen, I., Tran, C., Fox, M., Lin-Cereghino, J., & Lin-Cereghino, G. P. (2013, November). *Mutant strains of Pichia pastoris with enhanced secretion of recombinant proteins*. Biotechnology letters. Retrieved October 24, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3814129/>

- Macauley-Patrick S., Fazenda M.L., McNeil B., Harvey L.M. *Heterologous protein production using the Pichia Pastoris Expression System*. Yeast (Chichester, England). Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/15704221/>
- Malys N., Wishart J., Oliver SG, McCarthy. *Protein production in saccharomyces cerevisiae for systems biology studies*. Methods in enzymology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/21943899/>
- Naranjo CA;Jivan AD;Vo MN;de Sa Campos KH;Deyarmin JS;Hekman RM;Uribe C;Hang A;Her K;Fong MM;Choi JJ;Chou C;Rabara TR;Myers G;Moua P;Thor D;Risser DD;Vierra CA;Franz AH;Lin-Cereghino J;Lin-Cereghino GP; (n.d.). *Role of BGS13 in the secretory mechanism of Pichia pastoris*. Applied and environmental microbiology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/31585990/>
- Park H., Lennarz W.J. *Evidence for interaction of yeast protein kinase C with several subunits of oligosaccharyl transferase*. Glycobiology. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/10910977/>
- Prabakaran, S., & Damodaran, S. (1997, January 1). *Thermal unfolding of beta-lactoglobulin: Characterization of initial unfolding events responsible for heat-induced aggregation*. AGRIS. Retrieved October 24, 2022, from <https://agris.fao.org/agris-search/search.do?recordID=US1997079668>
- Roth J;Zuber C;Park S;Jang I;Lee Y;Kysela KG;Le Fourn V;Santimaria R;Guhl B;Cho JW; (n.d.). *Protein N-glycosylation, protein folding, and protein quality control*. Molecules and cells. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/21340671/>
- Sahdev S., Khattar S.K., Saini K.S. *Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies*. Molecular and

cellular biochemistry. Retrieved October 24, 2022, from

<https://pubmed.ncbi.nlm.nih.gov/17874175/>

Shibui, T., Bando, K., & Misawa, S. (2013, May 27). *High-level secretory expression, purification, and characterization of an anti-human her II monoclonal antibody, trastuzumab, in the methylotrophic yeast Pichia Pastoris*. Advances in Bioscience and Biotechnology. Retrieved October 24, 2022, from

<https://www.scirp.org/journal/paperinformation.aspx?paperid=32262>

Tang H;Wang S;Wang J;Song M;Xu M;Zhang M;Shen Y;Hou J;Bao X; (n.d.). *N-hypermannose glycosylation disruption enhances recombinant protein production by regulating secretory pathway and cell wall integrity in saccharomyces cerevisiae*. Scientific reports. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/27156860/>

Vieira Gomes AM;Souza Carmo T;Silva Carvalho L;Mendonça Bahia F;Parachin NS; (n.d.). *Comparison of yeasts as hosts for recombinant protein production*. Microorganisms. Retrieved October 24, 2022, from <https://pubmed.ncbi.nlm.nih.gov/29710826/>