

2019

# Plasmid optimization and the localization of the binding site of GPS2-UBC13

---

<https://hdl.handle.net/2144/36298>

*"Downloaded from OpenBU. Boston University's institutional repository."*

BOSTON UNIVERSITY  
SCHOOL OF MEDICINE

Thesis

**PLASMID OPTIMIZATION AND THE LOCALIZATION OF THE BINDING  
SITE OF GPS2-UBC13**

by

**AYESHA ABDULLAH**

B.A., Trinity College, 2009

Ph.D., Pennsylvania State University, 2015

Submitted in partial fulfillment of the  
requirements for the degree of  
Master of Science

2019

© 2019 by  
Ayesha Abdullah  
All rights reserved

Approved by

First Reader

---

Valentina Perissi, Ph.D.  
Associate Professor of Biochemistry

Second Reader

---

Yuan Gao, Ph.D.  
Postdoctoral Fellow of Biochemistry

# **PLASMID OPTIMIZATION AND THE LOCALIZATION OF THE BINDING**

## **SITE OF GPS2-UBC13**

**AYESHA ABDULLAH**

### **ABSTRACT**

The GPS2 protein (G-protein pathway suppressor 2) is a product of the mammalian *gps2* gene. It was originally identified and characterized in the context of G protein mitogen-activated protein kinase (MAPK) signaling pathways. Several studies have linked GPS2 with the inhibition of the ubiquitin conjugating enzyme UBC13. GPS2-mediated inhibition of UBC13 regulates several metabolic and inflammatory pathways. It has been shown that a lack of GPS2 is correlated with an increase in adiposity and inflammation due to the aberrant activity of UBC13 affected pathways. Therefore, understanding the relationship between UBC13 and GPS2 will provide further understanding of the molecular processes involved in adipose tissue levels, inflammation and downstream molecular responses. In this study, we attempt to determine the molecular determinants of GPS2 interaction with UBC13 by optimizing the protein expression protocol required to produce GPS2 protein expression in *Escherichia coli* in quantities viable for biochemical and structural assays. Our results indicate that optimization of the *gps2* sequence is required for efficient GPS2 protein expression in *E. coli* cells. Thanks to this optimization we have been able to successfully express GPS2 full length and several fragments, however, further optimization will be required for assessing GPS2-UBC13 molecular binding via *in vitro* binding assays.

## TABLE OF CONTENTS

TITLE.....	i
COPYRIGHT PAGE.....	ii
READER APPROVAL PAGE.....	iii
ABSTRACT.....	iv
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	vi
LIST OF FIGURES .....	vii
LIST OF ABBREVIATIONS.....	viii
INTRODUCTION .....	1
MATERIALS AND METHODS.....	8
RESULTS .....	18
DISCUSSION .....	33
REFERENCES.....	36
VITA .....	39

## LIST OF TABLES

Table	Title	Page
1	Plasmids Used	20

## LIST OF FIGURES

Figure	Title	Page
1	Mutagenesis cycling parameters	13
2	GPS2 1-99 Purification with Plasmid CVP 358 GPS2 1-133 Purification with Plasmid CVP 361	21
3	Fragments GPS2-1 <sup>st</sup> GG Purification with Plasmid CVP 349 and GPS2-2 <sup>nd</sup> GG Purification with Plasmid CVP 359	21
4	Codon Optimization of Gps2 sequences	24
5	Gps2 and pGEX JDK Vector Digest Test	25
6	Gps2, Gps2-1 <sup>st</sup> GG, Gps2-2 <sup>nd</sup> GG PCR Test	25
7	Optimized GPS2-1 <sup>st</sup> GG and GPS2-2 <sup>nd</sup> GG Purifications	27
8	Optimized Fragments GPS2-1 <sup>st</sup> GG, GPS2-2 <sup>nd</sup> GG, and GPS2 Full-Length Purifications	28
9	Optimized GPS2 Western Blot with Plasmid CVP 451	28
10	Optimized Full-Length GPS2 Purification with Silver Stain (500mL Flow Through; 100uL GST Beads)	29
11	UBC13 (18 kDa) TnT Results	31
12	<i>In-Vitro</i> Binding Assay Results UBC13 (TnT) and Optimized Full-Length GPS2 (Purified)	31

## LIST OF ABBREVIATIONS

AA.....	Amino Acid
AKT.....	Protein Kinase B
BP.....	Base Pairs
C.....	Carbon
<i>E. coli</i> .....	<i>Escherichia coli</i>
EDTA.....	Ethylenediaminetetraacetic acid
DNA.....	Deoxyribonucleic acid
DTT .....	Dithiothreitol
G.....	Glycine
GPS2.....	G-protein pathway suppressor 2
GST.....	Glutathione S-transferase
IKK.....	IKB Kinase
kDa.....	Kilodalton
LB .....	Lysogeny Broth
Lys, K.....	Lysine
Methionine.....	M
MgSO <sub>4</sub> .....	Magnesium Sulfate
N.....	Nitrogen
NaCl.....	Sodium Chloride
NCoR .....	Nuclear Receptor Co-Repressor
NP-40.....	Nonyl Phenoxyethoxyethanol

PBS.....	Phosphate-buffered Saline
PCR.....	Polymerase Chain Reaction
PIC.....	Protein Inhibitor Cocktail
PMSF .....	Phenylmethanesulfonylfluoride
RPM.....	Revolutions Per Minute
SMRT.....	Silencing Mediator of Retinoic Acid and Thyroid Hormone
Receptor TNF- $\alpha$ .....	Tumor Necrosis Factor Alpha
UBE2N.....	Ubiquitin Conjugating Enzyme E2 N
UBC13.....	Ubiquitin Conjugating Enzyme E2 13
UEV.....	Ubiquitin E2 Variant Protein
UV.....	Ultra Violet

## INTRODUCTION

### **Protein Ubiquitination:**

Ubiquitination is the addition of one or more ubiquitin molecules to a target protein. It is a very important biological process that affects a wide array of cellular processes through the covalent conjugation of ubiquitin to cellular proteins.

Ubiquitination occurs by formation of an isopeptide bond between a lysine residue donor and the C-terminal carbonyl group at ubiquitin's glycine 76 position. Three enzymes are required to make the isopeptide bond reaction. These enzymes are E1, a ubiquitin activating enzyme, E2, a ubiquitin-conjugating enzyme, and E3, the ubiquitin protein ligase. The E1 enzyme forms an initial thioester bond with the carboxyl terminus of the ubiquitin molecule. This occurs in an ATP-dependent reaction. After the thioester bond forms, the E2 transfers the ubiquitin to its own active site cysteine. The isopeptide bond then forms as a lysine residue of the substrate protein attacks the E2 bond ubiquitin. Lastly and simultaneously, this attack requires an E3 enzyme which binds the E2 ubiquitin thioester and the protein being ubiquitinated (Hershko and Ciechanover, 1998).

Ubiquitination can occur through the addition of a single ubiquitin molecule (mono-ubiquitination), the addition of several ubiquitin monomers (multiple mono-ubiquitination), or the addition of a polymeric chain of ubiquitin molecules (poly-ubiquitination). Such polyubiquitin chains are connected through a covalent bond between the C-terminus of one ubiquitin and one of the seven lysine side chains of the next ubiquitin (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, or Lys63) or the amino-terminal methionine (M1) of the next ubiquitin (Chau et al., 1989; Eddins M. et al., 2006;

Meinzel T. et al., 2006). Polyubiquitin chains with different lysine linkages are usually referred to by the position of the ubiquitin Lys used for branching, and often have distinct biological roles (VanDemark A. et al. 2001). Best studied are chains with Lys-48 linkages that are known to signal proteins for degradation by the 26S proteasome. Ubiquitination pathways that do not tag proteins for degradation, however, also exist. Chains with Lys-63 linkages, for example, do not signal degradation. This ubiquitination pathway has been linked to post replicative DNA repair after UV damage, IKK inflammation pathways, general translational regulation, as well as some types of ubiquitin-dependent endocytosis (VanDemark A. et al. 2001; Eddins M. et al., 2006).

### **G-Protein Pathway Suppressor 2 (GPS2):**

GPS2 (G-protein pathway suppressor 2) is a 36kDa ubiquitously expressed protein. The protein was initially studied in the context of G protein-mitogen-activated protein kinase (MAPK) signal pathways (Spain et al., 1996). Then it was identified as a transcriptional cofactor that interacts with NCoR/SMRT corepressor complex in gene repression and interacts with nuclear receptors and other transcription factor in gene activation (Zhang et al., 2002; Sanyal et al., 2007; Jakobsson et al., 2009; Cardamone et al., 2018). More recently, non-genomic functions of GPS2 in the cytosol have been reported by our lab. In particular, it was found that GPS2 in adipocytes is a negative regulator of AKT activation, required for restricting its activity and downstream signaling activity at low insulin levels (Cenderquist et. al, 2016). In mice, upon GPS2 deletion, the PI3K/AKT pathway is found constitutively activated, lipid metabolism disrupted, and

adiposity is increased (Cederquist et al., 2016). In cytoplasm, GPS2 inhibits of the enzymatic activity of TRAF/UBC13. The downstream molecular response is the inhibition of inflammatory TNF- $\alpha$  genes (Cardamone et al., 2012). Our lab has also identified GPS2 as an important protein in mitochondrial retrograde signaling as well as a transcriptional activator of nuclear encoded mitochondrial genes for mammals (Cardamone et al., 2018). To this end, GPS2 recruitment to its target genes promotes histone H3K9 demethylation and RNA POL2 activation via UBC13 inhibition. Notably, the same strategy is also important for the activation of the nuclear stress response to mitochondrial depolarization (Cardamone et al., 2018).

### **GPS2 as a regulator of protein ubiquitination:**

Overall, these previous studies indicate that GPS2 is an important protein to study and that determining its mechanisms of action is important to understand how to modulate its functions as a regulator of inflammation, lipid metabolism and mitochondrion homeostasis (Spain et al., 1996; Bi H. et al., 2014). As mentioned, our lab's previous study found GPS2 can regulate insulin signal and pro-inflammatory pathways via inhibition of the UBC13 activity (Cardamone et al., 2012, 2014; Cederquist et al., 2016; Lentucci et al., 2016). Specifically, GPS2 acts as a negative regulator of inflammation by inhibiting the pro-inflammatory TNF- $\alpha$  pathway via inhibition of UBE2N/UBC13-mediated ubiquitination of RIP1 (Cardamone at al., 2012). Similarly, in adipocytes, GPS2 negatively regulates activation of insulin signaling by means of the inhibition of UBE2N/UBC13-mediated ubiquitination of AKT (Cederquist et al., 2016).

In the nucleus as well, GPS2-mediated inhibition of UBC13 activity appear to regulate chromatin remodeling, and thus expression, of target genes via stabilization of a histone demethylase (Cardamone et al., 2018) .

UBE2N, in eukaryotes, or UBC13 in prokaryotes, is the main ubiquitin-conjugating enzyme (E2) that conjugates the synthesis of lysine 63-linked ubiquitin chains (Eddins et al., 2006). GPS2 inhibits Lys-63-linked ubiquitination (Eddins et al., 2006). GPS2 performs its regulatory activities by preventing the synthesis of Lys-63-poly-ubiquitin chains via direct interaction with UBC13 (Lentucci et al., 2016), however the mechanistic details of this inhibition are not known.

#### **UBC13 structural information and mechanism of action:**

A special type of E2, a UEV (ubiquitin E2 variant protein), acts as a subunit E2 in the assembly of K63-linked poly-ubiquitin chains. UEVs function directly with UBC13 in the assembly of K63 chains (VanDemark et al., 2001). In one study, a fundamental eukaryotic UEV protein, the Mms2 protein in yeast (its related and structurally similar human UEV protein is UEV1a), is used to show the binding and structural interaction between UEVs and UBC13 that lead to K63 polyubiquitin chains (VanDemark et al., 2001). The study shows that Mms2 and UBC13 form a T-shaped dimer that allows for the binding of ubiquitin's glycine 76 and the subsequent forming of the isopeptide bond (VanDemark et al., 2001). Furthermore, ubiquitin sequencing shows that there are two glycine residues in succession (at positions 75 and 76) at the binding site during the bonding of ubiquitin to UBC13. Thus, in the process of synthesizing Lys-

63-linked polyubiquitin chains, donor ubiquitin's glycine attaches to UBC13 at its active site to form UEV-E3-Ubiquitin. For our study, it will be important to note that the protein GPS2, which also binds to UBC13, also has two glycine residues, in sequence, at two particular locations on its amino acid primary sequence.

### **Project Outline and Goals:**

Many studies have suggested GPS2 and UBC13 as a potential therapeutic targets for intervention in various different cancers, the alleviation of anti-cancer drug resistance, chronic inflammation – especially that associated with obesity – and some viral infections (Fan R. et al., 2016; Huttlin E. et al., 2017; de Chasse B. et al., 2008). Further characterization of the role of UBC13 and improved mechanistic understanding of its inhibition by GPS2 are essential in order to determine the usefulness of such interventions and to design specific approaches.

The long-term goal of the present project is to study the GPS2-UBC13 molecular interaction in order to identify GPS2's mechanism of inhibition. In order to dissect the GPS2-UBC13 interaction, we planned to use *in-vitro* binding assay. Firstly, a sufficient amount of each protein must be gathered in order to perform binding and functional interaction assays. Additionally, in order to localize the minimal binding surface between these two proteins, we must perform binding assays between various fragments of GPS2 and UBC13. Lastly, we can use site-specific mutagenesis to alter specific residues that may be important for the binding or inhibitory activity and test their relevance. In particular, as mentioned above, it is known that UBC13 binds ubiquitin at the C-terminal

GG site. Because the protein sequence of GPS2 also includes two separate sites with sequential GGs, the specific hypothesis we would like to test is that one of the GGs sites on GPS2 are involved in the binding of GPS2 to UBC13.

Overall, the constructs/mutations we set out to test include: the full length wild type variation of UBC13 and GPS2, three variations of the full length GPS2 with a mutation of both of the GG sites (1<sup>st</sup> GG mutation, 2<sup>nd</sup> GG mutation, as well as double mutation), fragments of GPS2 that include amino acids 1-212 (GPS2 1-212), a fragment that includes amino acids one through GPS2's first GG (GPS2-1<sup>st</sup> GG), a fragment that includes amino acids one through GPS2's second GG (GPS2-2<sup>nd</sup> GG), a fragment that includes the first 99 amino acids of GPS2 (GPS2 1-99), a fragment that includes the first 133 amino acids of GPS2 (GPS2 1-133), and lastly a fragment that includes the last 115 amino acids of GPS2 (GPS2 212-327). GPS2 fragments including amino acids 1-99 and 1-133 were chosen to test N-terminal sufficiency, while GPS2 212-327 was chosen to test C-terminal sufficiency. GPS2 1-133, GPS2 1-99, GPS2 212-327, and GPS2 1-212 were chosen as positive controls from previous mapping results. As we hypothesize that the first or second GG in GPS2's sequence is required in binding, we also planned to generate full-length GPS2 containing mutations at these sites. These GPS2 variations will be central to the long-term goal of determining whether the binding of GPS2 and UBC13 require either of the GG sites.

**Specific Aims:**

In order to investigate the required domain/site on GPS2 binding to UBC13, our goals were:

- (1) Construct plasmids expressing GST-tagged GPS2 full length or deletions;
- (2) Test the expression of targeted proteins in *E. Coli*;
- (3) Purify proteins and test for expression;
- (4) Mutagenesis of either wild type or optimized full length GPS2 at each GG and both GG sites;
- (5) Prepare TnT Quick Coupled Transcription/Translation and begin performing binding assays

## MATERIALS AND METHODS

### **Transformation, Culturing, and Induction of *E. Coli* cells:**

In preparation for protein purification, BL21 competent cells (BioLabs) previously stored in a -80°C refrigerator were used to transform plasmids containing GPS2 or GPS2 fragment plasmids. After thawing on ice, one to two microliters of plasmid was transferred into the competent cells vial. After incubating on ice for thirty minutes, the vial was placed in a 42°C water bath for exactly thirty seconds. The vial was removed and placed immediately on ice for two minutes. After two minutes, one milliliter of LB medium was added to the vial and the vial was then placed in incubation at 37°C and shaken for one hour. Approximately fifty microliters of the media and cells were placed onto a plate with vector-related resistance for easy transformation cells (for difficult transformation cells, i.e. initial plasmid transformation, media was spun down, ninety percent of supernatant removed, and then cells were resuspended in remaining one hundred microliters and spread on plate). Plates were then incubated at 37°C overnight. The following morning, nine to twelve colonies were replated and incubated again for five to eight hours.

For cell culturing and induction, a colony sample was added to five milliliters of LB medium with related resistance and cultured overnight. Approximately 12 hours later, one to five milliliters of the culture was removed and diluted with LB medium in a one to one hundred ratio in a five hundred milliliter flask. Cultures were incubated for one to two hours and shaken at 37°C until concentration reached 0.3-0.5 OD<sub>600</sub>. IPTG (Fisher

Bioreagents) was then added to the cultures, diluted to one to one thousand from a 0.5M stock. Induced cultures were then incubated at 25-30°C for three to four hours.

### **Protein Purification and western blotting:**

To extract whole cell lysates, bacteria cells were harvested and spun down for 15min at 3000 rpm (4C). Cell pellet was resuspended in cold lysis buffer (50mM Tris Hcl, pH7.5, 400mM NaCl, 10% glycerol, 0.1% NP40, 1mM EDTA, 1mM DTT, 0.1mM PMSF and protease inhibitor mix (Roche Diagnostics)) Pellets were kept on ice to preserve proteins during the entire process. In order to lyse the cells fully, pellets suspended in lysis buffer were separated into sonicator tubes and sonicated for six minutes at 60% amplitude for 30 seconds work time and 30 second rest and cool time. Samples were then spun down at max speed for 10 minutes at 4°C. Supernatant was taken as whole cell lysates. Concentration of the protein extracts were measured using the Quick Start™ Bradford 1X Dye Reagent per manufacturer's protocol (Bio-Rad) and the NanoDrop™ 2000. Glutathione Sepharose 4B beads (Thermo Fisher) were washed three times with lysis buffer, then added to the whole cell lysates. The Glutathione Sepharose 4B beads-sample mixture was then place in 4°C to rotate and incubate overnight.

Approximately 12 hours later, the Glutathione Sepharose 4B beads-sample was spun down at approximately 600 rpm in 4°C. The supernatant was removed and labelled as flow through. Beads were washed with wash buffer (PBS, 300mM NaCl) for three times. All samples were boiled with NuPAGE LDS Sample Buffer-4X (Life

Technologies) and DTT at 100°C for 10 minutes. Samples were then loaded into either a stain free 10% gel (if silver stain intended) or a stained 10% gel (Bio-Rad Mini-PROTEAN® TGX™). A Bio-Rad Precision Plus Protein™ Dual Color Standards was used for the ladder and gels were run using Boston BioProducts 1x Tris-Glycine-SDS Running Buffer. Proteins were then transferred onto PVDF membranes (EMD Millipore), followed by standardized western blotting protocol. Chemiluminescent western blot detection was performed using Clarity™ Western ECL Substrate and ChemiDoc™ XRS+ with Image Lab software per manufacturer's protocol (Bio-Rad). GST primary antibody (Santa Cruz Biotechnology) were used in this study.

#### **PCR reaction:**

For plasmid creation, primers were diluted with distilled water to 100uM. Primers were then diluted to 10uM in 1.5mL tubes in preparation for individual PCR reaction experiments. Each PCR reaction formula were made to contain 500ug plasmid template, 1x high fidelity PCR buffer, 200nM dNTP mixture, 2mM MgSO<sub>4</sub>, 1U Taq Polymerase (5u/uL), 40nM forward primer as well as 40nM reverse primer, and nuclease free water to final 50ul (Fisher Bioreagents). Typical PCR protocol was followed at temp 37 °C Celsius. PCR products were cleaned with a Qiagen QIAquick PCR Purification kit. PCR Products were run on DNA gel electrophoresis to test for correct base pair length. Gels were 10% agarose in 1x TBE. A 1:1000 allotment of Denville Scientific Greenglo DNA Dye was added to the gel. Fisher Scientific exACT Gene 1kb Plus DNA was used as a

ladder and 1x TBE was used to run the gel. The correct sized PCR products were extracted using Qiagen QIAquick Gel Extraction Kit.

### **PCR Product Digestion and Ligation Solutions:**

In order to construct plasmid expressing target genes, we digested purified DNA fragment and expressing vector using same restriction endonucleases to generate the adhesive ends. The restriction endonucleases used in our study are BamHI and EcoRI (NEB). For ligation reaction, T4 ligase (BioLabs) were used as mentioned in manufacture introduction.

### **Silver Stain Protocol:**

Silver staining was performed in order to determine protein purification levels down to the nanogram level. Thermo Fisher Scientific Pierce Silver Stain Kit was used for silver staining. In order prepare for silver staining, a fixing solution, ethanol solution, sensitizer working solution, developer working solution, and a stop solution were made. 25ml of each solution was created for each silver stain. The fixing solution contained a 30% pure ethanol and 10% acetic acid. The ethanol solution contained 10% pure ethanol in water. The sensitizer working solution contained 25ul sensitizer in 25ml water. The developer working solution contained 0.5mL of enhancer in 25ml developer. The stop solution contained 5% acetic acid in water.

All steps were performed in a single clean staining plastic tray with constant gentle shaking. Gels were handled with clean gloved hands to manipulate the gel to

reduce contamination. Fixing, ethanol, and stop solutions were prepared in advance.

Other solutions were prepared immediately before use.

First, the gel was washed in pure water for 5 minutes. The water was then replaced and the gel was washed for another 5 minutes. The gel was then fixed in the fixing solution for 15 minutes. The fixing solution was replaced and the gel was fixed for another 15 minutes. The gel was then washed in the ethanol solution for 5 minutes. The solution was replaced and the gel was washed for another 5 minutes. The gel was then washed in pure water for 5 minutes, the water was replaced and the gel was washed for another 5 minutes. The sensitizer working solution was then prepared by mixing one-part Silver Stain Sensitizer with five hundred parts pure water. The gel was then incubated in the sensitizer working solution for exactly 1 minute, then washed with two changes of pure water for 1 minute each. The stain working solution was then prepared by mixing one-part Silver Stain Enhancer with fifty parts Silver Stain. The gel was then incubated in the stain working solution for 30 minutes. The developer working solution was then prepared by mixing one-part Silver Stain Enhancer with fifty parts Silver Stain Developer. The stop solution was also kept ready at hand at this time. The gel was quickly washed with two changes of ultrapure water for 20 seconds each. Immediately following the wash, the developer working solution was added and the gel was incubated until protein bands appear. This took roughly one minute. At that time, the developer was removed and the stop solution was immediately added. The gel was incubated for 10 minutes, the stop solution removed, and more stop solution added. The gel remained in the stop solution for 10 minutes and then placed in pure water for viewing.

## Mutagenesis:

Site-Directed Mutagenesis Kit from Agilent Technologies was used to create mutations. A control and sample solution were both prepared. Using the Agilent kit, a 50ul solution total in a PCR tube, 5ul of 10x buffer, 1.2 ul of the kit's prepared primers, 1ul of the kit's dNTP, 0.5 ul of the kit's DNA, 3ul of the kit's QuikSolution, 38ul of nuclease free water, and 1ul of the kit's PfuTurbo DNA polymerase (2.5u/uL) were mixed into a control PCR solution. For the desired product solution, 5ul of 10x reaction buffer, 10ul of GPS2 (GPS2 or 2<sup>nd</sup> GG mutation) DNA template, 125ng of the forward primer as well as 125ng of the reverse primer. 1ul of dNTP nucleotides, 3ul of QuikSolution and 1ul of Pfu DNA polymerase were added to solution and nuclease free water was added to make a final volume of 50ul. The reaction was cycled using the below parameters with an eighteen-cycle limit. Cycle each reaction using the cycling parameters outlines

Segment	Cycles	Temperature	Time
1	1	95°C	1min
2	18	95°C 60°C 68°C	50 sec 50 sec 1min/kb of plasmid length
3	1	68°C	7 min

**Figure 1: Mutagenesis cycling parameters**

Following temp cycling, the reaction tubes were placed on ice for two minutes to cool reactions to less than 37°C. Dpn I from the Mutagenesis kit was then added to each solution for digestion of parent DNA. After incubation at 37°C for 1 hour, the 1ul of the DNA was then transformed to competent cells (BioLabs).

### **Mini and Midi Plasmid Purifications:**

A Qiagen Hispeed Plasmid Midi Kit was used to purify plasmid products following PCR creation and transformation. For midi purification, 50ml of LB media with 1mM ampiciline was added to a culturing tube. Desired bacteria was added to the tube and it was incubated overnight at 37°C. For mini purification, 5ml of LB media and desired bacteria was cultured. For each midi and mini, culture was placed in 1.5mL tubes. The tubes were then spun at max speed in centrifuge to get a pellet and supernatant was removed. The pellet was then resuspended in P1 buffer and incubated at room temperature for 5 minutes. P2 buffer was added to the solution and the solution was incubated at room temperature for 5 minutes. P3 buffer was then added to the solution and it was again incubated at room temperature for 5 minutes. The solution was then spun down at max speed in a centrifuge held at 4°C for 10 minutes. The supernatant was then removed and added to new tubes. Isopropanol was then added to the supernatant at a 0.7x ratio. The tube was then vortexed for 10 seconds to mix the isopropanol and then incubated at 4°C for 10 minutes to aid precipitation of DNA. Tubes were then spun down by centrifuge for 15 minutes at max speed, at 4°C. The isopropanol was removed and the DNA pellet was washed with 500 ul 70% ethanol. The tubes were then spun again at max

speed for 5 minutes at 4°C. Ethanol was removed and DNA was air dried for 10 minutes and then resuspended in 30ul nuclease free water.

***In-vitro* transcriptional and translational reaction (TnT):**

In-vitro transcriptional and translational reaction was performed by TnT Quick Master Mix (Promega) following the manufacture introduction. Flag-tagged GPS2 fragment 1-212 and myc-tagged Ubc13 were formulated by TnT reaction. TnT Quick Master Mix was rapidly thawed by hand-warming and place on ice after removing from -80°C storage. Plasmid DNA used for TnT reaction and Methionine were thawed at room temp and stored on ice. The TnT reaction was including 1x TnT Quick Master Mix, 1ug plasmid DNA and 20nM methionine. The reaction was incubated at 37°C for 1hour. Generated proteins were analyzed by western blot. Myc antibody (Santa Cruz Biotechnology) and Flag antibody (Santa Cruz Biotechnology) were used for standard western blot analysis.

***In-vitro* binding Assay:**

Binding buffer (20mM Tris, pH8.0, 125mM NaCl, 10% glycerol, 0.1% NP40, 0.5mM DTT and PIC) was used to perform binding assays with products of the TnT reaction and GPS2 protein purification products.

First, Glutathione Sepharose 4B beads were washed with binding buffer twice for 10 minutes each. Beads were then resuspended with 300ul binding buffer. 5ul of the TnT reaction were then added to the prewashed beads. The solution was place in 4°C to rotate

for 1 hour. The beads were then washed with binding buffer three times with ten-minute incubation each time. SDS loading buffer was then added to the samples. Samples were boiled at 100C for 10 minutes and loaded into a stain free gel. Results were analyzed by western blot with the same products used to analyze TnT western results.

## RESULTS

### **GPS2 Plasmids and Protein Expression:**

The expression of proteins is a very important step in determining the structure and the binding of proteins to its partners. Often, protein expression is conducted using *E. coli*. As there is overlap and redundancy in the genetic code for protein expression between all living organisms, expressing proteins endogenous to eukaryotes can be done in bacteria. However, differences in coding and in protein length can affect the efficiency of protein expression. While all organisms have molecular mechanisms for ubiquitination and the enzymes responsible for Lys63 polyubiquitin chains are homologous in eukaryote organisms and bacteria, the specific genetic sequence encoding for the UBC13 protein sequence varies between eukaryotes and prokaryotes. Because this enzyme is homologous in both eukaryotes and prokaryotes, expression of the endogenous eukaryote protein in *E. coli* does not present complications. Moreover, because the protein is small, the smaller bacterial cell does not have a problem synthesizing UBC13. GPS2 is endogenous to mammals but not to bacterial cells, however. Additionally, GPS2 is a larger protein comparatively to UBC13 (37kD and 23kD respectively). Thus, GPS2 expression in *E. coli* and many of its larger fragments, in order to be synthesized in *E. coli*, require optimization.

Sequence optimization is the process of changing the DNA coding for specific amino acids from that preferred by eukaryotes to that preferred by prokaryotes, specifically here *E. coli*. A major goal of this project is the development of an optimized plasmid for the successful purification of GPS2 and various fragments of GPS2. The

development of an optimized plasmid is crucial to purifying GPS2 and various fragments in order to acquire sufficient protein expression by means of protein purification methods (Nieuwkoop T. et al. 2019). Proteins gathered from protein purification are subsequently used to perform binding assays to determine the molecular binding site of GPS2 and UBC13. Thus, the goals of this project are twofold: the creation of plasmids capable of successful (>~50ng) GPS2 (and fragments) purifications and the determination of the binding site of GPS2 and UBC13.

This project begins with the determination of which GPS2 protein fragments (GPS2, GPS2-1<sup>st</sup> GG, GPS2-2<sup>nd</sup> GG, GPS2 1-212, GPS2 1-99, GPS2 1-133, GPS2-212-327) require optimized plasmids for expression in *E. coli*. It is hypothesized that non-endogenous proteins larger than 25 kDa will require optimization and that optimization will provide successful quantities of GPS2 require for binding assays. For the further goal of this project - the determination of the binding site of UBC13 and GPS2 - it is hypothesized that one of the two GG sites on GPS2's sequence may be important in binding to UBC13, thereby competitively inhibiting ubiquitination activity. These are amino acids 136 and 137 and 185 and 186 respectively. Thus, we attempted to purify the entire GPS2 protein, the GPS2 fragment containing amino acids from 1-137 (GPS2-1<sup>st</sup> GG) as well as the fragment containing amino acids from 1-186 (GPS2-2<sup>nd</sup> GG). Also, to confirm previous findings and have positive controls for binding experiments, fragments from the 1st 99 amino acids (to determine N-terminal sufficiency), from the last 212-327 (to determine C-terminal sufficiency), as well as GPS2 1-133 and GPS2 1-212 were

purified. It was then determined which proteins required optimization and which did not. We expect that the full length GPS2 would be purified only in low quantities.

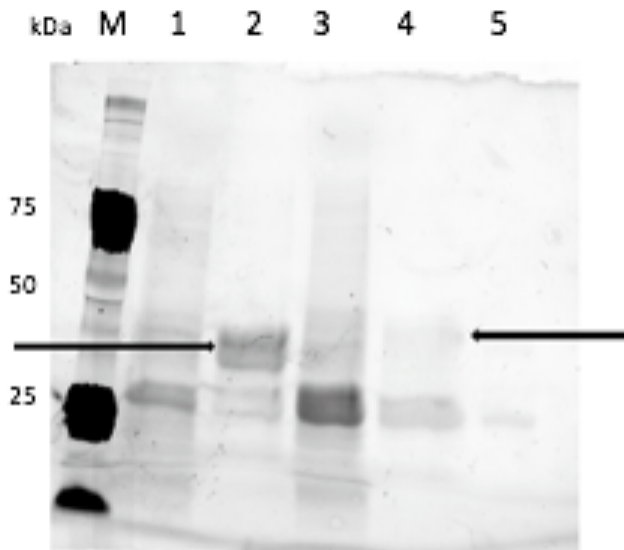
### **Purifications of GST-tagged full length GPS2 and fragments:**

Initial purifications of GPS2 included all of the wild type sequence fragments and wild type full length formats of GPS2. These were plasmids CVP 257-CVP 361 shown in table 1. Each plasmid used for purification was tagged with GST. Protein purification results were deemed successful when bands of sizes corresponding to the known size of the respective protein were observed in gel electrophoresis or Western Blot results. GPS2 fragment AA 2-99 and 2-133 were all purified successfully (Figure 2). Fragments ending at the first GG and second GG were purified with conflicting results in protein size (Figure 3). The conflicting results showed GPS2-1<sup>st</sup> GG as 47 kDa and GPS2-2<sup>nd</sup> GG as 42 kDa.

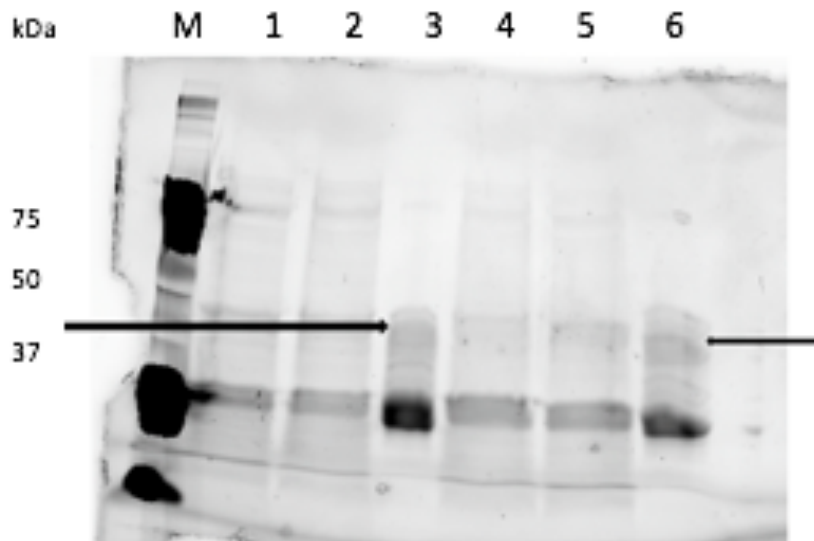
As expected, the full-length wild type GPS2 could not be purified satisfactorily. As the mutant variations of GPS2 are the same size as the wild type, it was concluded without purification that the expression of GPS2 full length mutant variations would not be able to be purified without optimization. Thus, fragments ending at the first GG and second GG as well as the full-length wild type and mutants were considered unable to be purified satisfactorily with their original mammalian sequences.

Plasmid	Full length/Fragment Type	Protein Size	Purification Results
CVP 258	Myc-Ubc13	~41 kDa	-
CVP 263	HA-Flag-GPS2	~45 kDa	-
CVP 266	Flag-GPS2 1-212	~ kDa	-
CVP 292	GST-GPS2 (wild-type)	~64 kDa	Unable to be purified
CVP 346	GST-GPS2 with 1 <sup>st</sup> GG mutation	~64 kDa	Unable to be purified
CVP 347	GST-GPS2 with 2 <sup>nd</sup> GG mutation	~64 kDa	Unable to be purified
CVP 349	GST-GPS2 fragment 1-1 <sup>st</sup> GG (AAs 1-137)	~42 kDa	Purified with poor results
CVP 358	GST- GPS2 fragment (AAs 1-99)	~38 kDa	Purified with original sequence
CVP 359	GST-GPS2 fragment 1-2 <sup>nd</sup> GG (AAs 1-186)	~47 kDa	Purified with poor results
CVP 360	GST-GPS2 fragment (AAs 212-327)	~39.7 kDa	Purified with original sequence
CVP 361	GST-GPS2 fragment (AAs 1-133)	~44 kDa	Purified with original sequence
CVP 451	GST-GPS2 (optimized sequence)	~64 kDa	Purified (low yield)
CVP 452	GST-GPS2 fragment 1 <sup>st</sup> GG (optimized sequence, AAs 1-186)	~42 kDa	Purified w/optimized sequence
CVP 453	GST-GPS2 fragment 2 <sup>nd</sup> GG (optimized sequence, AAs 1-186)	~47 kDa	Purified w/optimized sequence
CVP 454	GST-GPS2 with 1 <sup>st</sup> GG mutation	~64 kDa	To be purified before binding assay
CVP 455	GST-GPS2 with 2 <sup>nd</sup> GG mutation	~64 kDa	To be purified before binding assay
CVP 456	GST-GPS2 with 1 <sup>st</sup> GG and 2 <sup>nd</sup> GG mutation	~64 kDa	To be purified before binding assay

**Table 1: Plasmids Used**



**Figure 2: GPS2 1-99 Purification with Plasmid CVP 358 and GPS2 1-133 Purification with Plasmid CVP 361**  
 Lanes left to right: ladder, whole cell lysis of GPS2 1-99, elution of GPS2 1-133 (38kDa), whole cell lysis of GPS2 1-133, elution of GPS2 1-133 (44kDa), overflow of elution of GPS2 1-133



**Figure 3: Fragments GPS2-1<sup>st</sup> GG Purification with Plasmid CVP 349 and GPS2-2<sup>nd</sup> GG Purification with Plasmid CVP 359**  
 Lanes left to right: ladder, whole cell lysis, flow through, and elution of CVP 349 (42 kDa), whole cell lysis, flow through, and elution of CVP 359 (47 kDa)

### **PCR Construction of Optimized Plasmids:**

Based on the wild type sequences of *gps2* construct, we cannot purify target proteins from *E. coli*. One possible reason is the sequence of *gps2* is from the mammalian gene, which is difficult to express in bacteria (Khow O. et al, 2012). Optimized sequences were made for the first GG and second GG fragment as well as the full length and first GG and second GG mutations. In order to obtain optimized sequences, the Invitrogen Geneart Synthesis database by Thermo Fisher was used to translate eukaryotic protein sequence to an optimized prokaryotic protein sequence. Primers and template DNA for PCR were obtained from MacroGen. Final PCR product sequences were also sent to MacroGen for verification by Sanger sequencing.

In order to construct plasmids for the first GG and second GG fragment as well as the full length and first GG and second GG mutations, optimized sequences were formulated and sent to MacroGen for construction of PCR primers and templates. Typical PCR protocol was performed to create the first GG and second GG fragment as well as all full length *gps2* plasmids. Results yielded the correct construction and sizes of each plasmid.

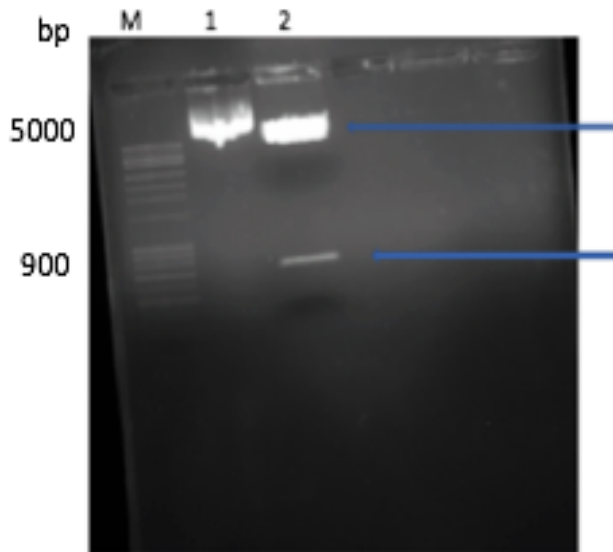
The optimized *gps2* construct was used as template DNA. The forward primer sequence was 5' CGGGATCCCCTGCACTGCTGGAACGT 3' and the reverse primer sequence was 5' CGGAATICACCACCAGGGCTACCC 3'. First, the target sequences were amplified by PCR. The PCR product and a pGEX vector containing a GST tag were then digested with compatible restriction enzymes BamH1 and EcoR1. The digestion reaction was successful as accurate DNA fragment sizes were observed for the vector and

each plasmid (Figures 5 and 6). The digested products were then cleaned by gel electrophoresis followed by the Qiagen PCR clean-up kit. The cleaned digested products were then ligated with T4 ligase buffer and T4 ligase. The solution was placed at 16°C overnight to ligate with the pGEX vector. Approximately 12 hours later, the ligation product was transformed into general, competent *E. coli* cells. Transformants were grown overnight at 37 °C on LB agar plates supplemented with ampicillin. In order to verify the correct PCR products were obtained and correctly oriented in the *E. coli* plasmids, colonies deemed correct were transferred to separate PCR tubes containing a ten-microliter ready-made PCR solution with two microliters of each 5' and 3' sequencing primers added. PCR product was analyzed again – ten microliters each - by agarose gel electrophoresis. For each trial, two to three colonies were sent to Macrogen for screening in order to verify that the insert is in the proper orientation and in the correct reading frame. Figure 6 below shows the optimized *gps2* sequence for primer creation and PCR plasmid creation.

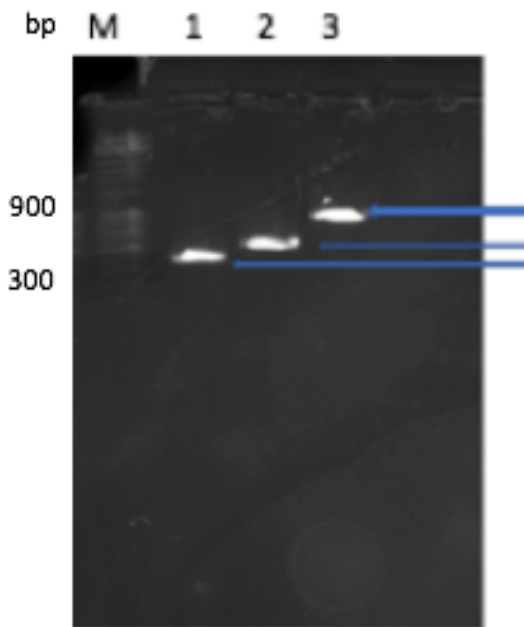
Range 1: 1 to 984 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
637 bits(706)	0.0	734/985(75%)	2/985(0%)	Plus/Plus
Query 1	ATGCCTGCACTGCTGGAACGTCCGAAACTGAGCAATGCAATGGCACGTGCACTGCATCGT	60		
Sbjct 1	ATGCCCGCACTCCTGGAGCGCCCAAGCTTCCAACGCTATGGCCAGGGCTCTGCACCGG	60		
Query 61	CATATCATGATGGAACGTGAACGTAAACGTCAAGAAGAAGAGGAAGTGCACaaaatgatg	120		
Sbjct 61	CACATCATGATGGAGCGGGAGCGCAAACGGCAGGAGGAGGAAGAGGTGGACAAGATGATG	120		
Query 121	gaacagaagatgaagaagaacaagagcgtcgcaaaaagaagaagaatggaagaaCGTATG	180		
Sbjct 121	GAAACAGAAGATGAAAGAAGAGCAGGAGAGAAGAAAGAAAAGGAAATGGAAAGAGAATG	180		
Query 181	AGCCTGGAAGAAACCAAGAGCAGATTCTGAAACTGCAAGAAAACTGAGTGCCCTGCAA	240		
Sbjct 181	TCACTAGAGGAGACCAAGGAACAGATCCTGAAGCTGCAGGAGAAGCTTCTGCTCTACAG	240		
Query 241	GAAGAAAAACATCAGCTGTTTCTGCAGCTGAAAAAAGTGCATGAAGAAGAGAAACGT	300		
Sbjct 241	GAGGAGAAGCACCAGCTTTTCTTGCAGCTCAAGAAAGTTTGCATGAGGAAGAAAAACGG	300		
Query 301	CGTCGTAAGAACAGAGCGATCTGACCACACTGACCAGCGCAGCATATCAGCAGAGCCTG	360		
Sbjct 301	AGGCGAAAGGAACAGAGTGACTTAACCACTCTGACATCAGCTGCATACCAGCAGAGCCTG	360		
Query 361	ACCGTTCATACCGGTACACATCTGCTGAGCATGCAGGGTAGCCCTGGTGGTCATAATCGT	420		
Sbjct 361	ACGGTTCATACAGGAACCCACCTCCTCAGCATGCAGGGGAGCCCTGGAGGACACAATCGC	420		
Query 421	CCGGGTACACTGATGGCAGCAGATCGTGCAAAACAATGTTTGGTCCGCAGGTTCTGACA	480		
Sbjct 421	CCAGGTACTCTGATGGCAGCTGACAGAGCCAAGCAGATGTTTGGACCACAAGTGCTTACG	480		
Query 481	ACCCGTCATTATGTTGGTAGCGCAGCAGCATTTGCAGGTACACCGGAACATGGTCAGTTT	540		
Sbjct 481	ACCCGGCACTACGTGGGATCAGCAGCTGCTTTTGCAGGAACCCAGAACATGGACAATTC	540		
Query 541	CAGGGTTCACCTGGTGGTGCCTATGGCACCGCACAGCCTCCGCCTCATTATGGTCCGACA	600		
Sbjct 541	CAAGGCAGTCCAGGGGGTGCTTATGGGACTGCTCAGCCCCACCTCATTATGGGCCAAC	600		
Query 601	CAGCCTGCATATAGCCCGAGCCAGCAGCTGCGTGCACCGAGCGCATTTCCGGCAGTTT	660		
Sbjct 601	CAACCGCCTATAGTCTAGCCAGCAGCTCAGAGCCCCATCAGCATTCTCTGAGTGCAG	660		
Query 661	TATCTGAGCCAGCCGAGCCTCAGCCGTATGCAGTTTATGGCCATTTTCAGCCGACACAG	720		
Sbjct 661	TACCTATCTCAGCCACAGCCACAACCTATGCAGTGCATGGCCACTTTTCAGCCACTCAG	720		
Query 721	ACCGGTTTTCTGCAACCGGGTAGCACCTGAGCTTACAAAAGCAGATGGAACATGCAAAT	780		
Sbjct 721	ACAGGGTTCCTCCAGCCCGGAGTACCCTCTCTTTGCAGAAACAGATGGAGCATGCCAAC	780		
Query 781	CAGCAGACCAGCTTTAGCGATAGCAGCAGC-CTGCGTCCGATGCATCCGCAGGCGCTGCA	839		
Sbjct 781	CAGCAGACCAGC-TTCTCGGACTCATCTTCTCTGCGGCCATGCACCCCAAGCTCTGCA	839		
Query 840	TCCGGCACCAGGTCTGCTGGCAAGTCCGCAGCTGCCGGTTCAGATTAGGCAGCAGGTAA	899		
Sbjct 840	TCCAGCCCTGGACTCCTGGCTTCCCCCAAGCTTCCCGTACAGATACAAGCAGCAGGGAA	899		
Query 900	AAGCGGTTTTGCAACCACAGTCAGCCTGGTCCGCGTCTGCCGTTTATTCAGCATAGCCA	959		
Sbjct 900	GTCAGGCTTTGCCACCACAGCCAACCTGGCCCCGACTCCCTTTCATCCAACACAGCCA	959		
Query 960	GAATCCGCGTTTTTACCACAAATAA 984			
Sbjct 960	GAACCAAGATTCTATACAAGTAA 984			

**Figure 4: Codon Optimization of Gps2 sequences**  
**Top: Optimized Sequence, Bottom: Original Sequence**



**Figure 5: Optimized Gps2 and pGEX JDK Vector Digest Test**  
 Lanes from left to right: ladder, pGEX vector, pGEX vector (5000 bp) and Gps2 (900bp)



**Figure 6: Gps2, Gps2-1<sup>st</sup>GG, Gps2-2<sup>nd</sup> GG PCR Test**  
 Lanes 1 and 2, ladder, lane 1, Gps2-1<sup>st</sup> GG (411 bp), lane 2 Gps2-2<sup>nd</sup> GG (558 bp), lane 3 Gps2 (900 bp)

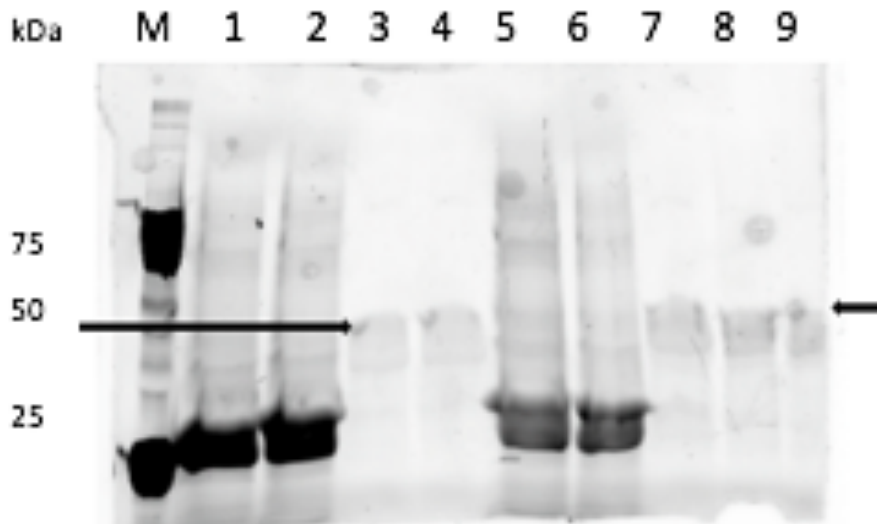
### **Purifications of GPS2 with Optimized Sequences:**

Optimized sequences were made for the first GG and second GG fragment as well as the full length and first GG and second GG mutations. Purification results for optimized sequences were successful with results varying from high (expressed with clear bands on gel electrophoresis) to low yield (bands only confirmed using sensitive techniques such as western blot and silver staining of gel electrophoresis). Plasmids with optimized sequences include CVP 451-CVP 456 shown in table 1.

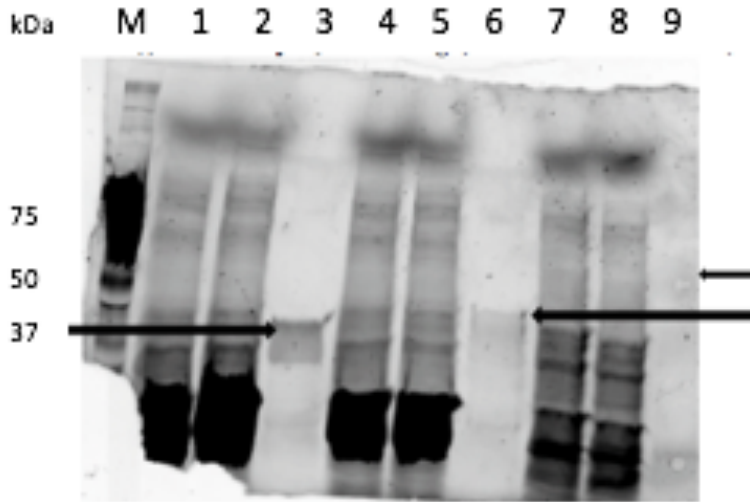
Low yield purification results were further verified for successful protein purification with Western Blot and silver staining, which are more sensitive testing methods to indicate presence of protein purification. These methods are capable of verifying protein expression at levels as low as one nanogram. Three plasmids, CVP 258, CVP 263, and CVP 266 were not used for the purposes of protein purification but for the purposes of TnT Quick Coupled Transcription/Translation in preparation for binding assays. In order to construct plasmids for the first GG and second GG fragment as well as the full length and first GG and second GG mutations, optimized sequences were formulated and sent to Macrogen for construction of PCR primers and templates. Typical PCR protocol was performed to create the first GG and second GG fragment as well as all full length GPS2 plasmids. Results yielded the correct construction and sizes of each plasmid.

After transformation of constructed plasmids into BL21 *E. coli* was performed and cultures plated, purifications on plasmids CVP 45, CVP 452, and CVP 453 were performed. Results showed 1<sup>st</sup> GG and 2<sup>nd</sup> GG fragments purified with expected yield

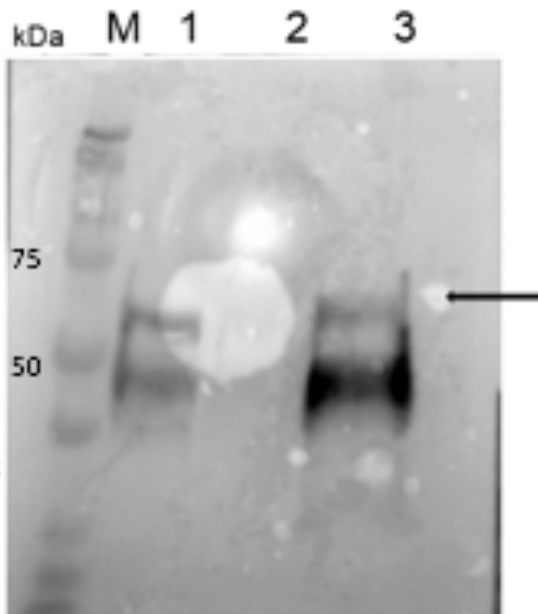
(50+ ng) on gel electrophoresis without sensitive staining techniques. Two trials of the fragments were performed because initial results showed very close bands between 42kDa and 47kDa for GPS2 1<sup>st</sup> GG and GPS2 2<sup>nd</sup> GG respectively (Figures 7 and 8). Full length GPS2 purification showed low yield (~1+ng) and required Western Blotting and silver staining of gels (Figures 9 and 10). Purification results of CVP452 and CVP 453 confirmed the hypothesis that optimized sequences would yield results high enough for viewing on gel electrophoresis. However, purifications of CVP451 showed purification results well below 50ng as expression levels were unable to be seen without silver staining or Western Blotting. High, low and zero yield results are show below (Figures 7 -10).



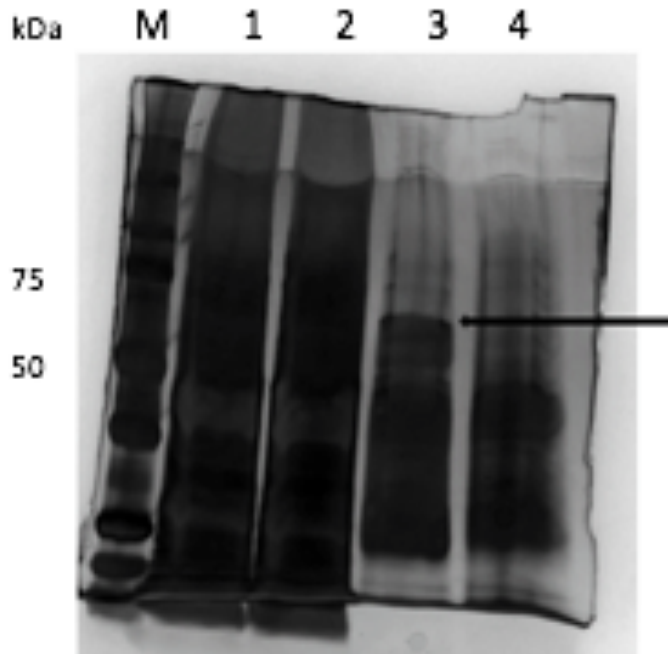
**Figure 7: Optimized GPS2-1<sup>st</sup> GG and GPS2-2<sup>nd</sup> GG Purifications**  
Lanes left to right: ladder, whole cell lysis, flow through, and elution of CVP 452 (42 kDa) (lanes 3 and 4); whole cell lysis, flow through, and elution of CVP 453 (47 kDa) (lanes 7, 8, 9)



**Figure 8: Optimized Fragments GPS2-1<sup>st</sup> GG, GPS2-2<sup>nd</sup> GG, and Optimized GPS2 Full-Length Purifications**  
 Lanes left to right: whole cell lysis, flow through, and elution of CVP 452 (42 kDa), whole cell lysis, flow through, and elution of CVP 453 (47 kDa), and whole cell lysis, flow through, and elution of CVP 451 (64 kDa)



**Figure 9: Optimized GPS2 Western Blot with Plasmid CVP 451**  
 Lanes left to right: ladder, whole cell lysis, flow through, elution (64 kDa)



**Figure 10: Optimized Full-Length GPS2 Purification with Silver Stain (500mL Flow Through; 100uL GST Beads)**  
**Lanes left to right: ladder, whole cell lysis, flow through, elution, elution (64 kDa)**

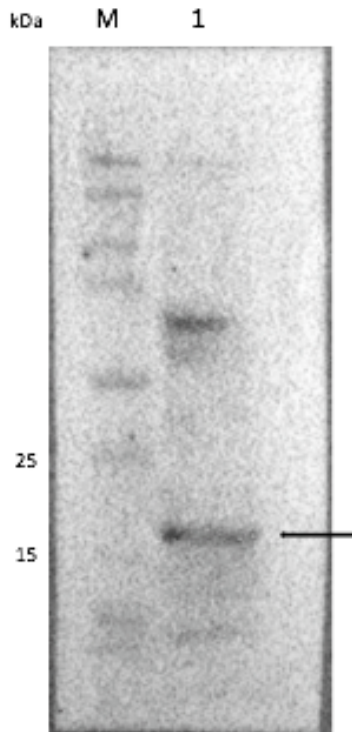
### **Mutagenesis Results:**

The Site-Directed Mutagenesis Kit from Agilent Technologies was used to create mutations for the 1<sup>st</sup> GG, 2<sup>nd</sup> GG, and double mutation of the full length GPS2. Typical PCR construction was then followed. Sequences were also sent to Macrogen for verification. Two complimentary oligonucleotide primers containing the desired mutation were ordered from Macrogen. For the 1<sup>st</sup> GG mutation primers, the forward primer was 5' –GCATGCAGGGTAGCCCTGCAGCCCATAATCGTCCGGGTAC – 3'. The reverse primer for the 1<sup>st</sup> GG mutation was 5' – GTACCCGGACGATTATGGGCTGCAGGGCT ACCCTGCATGC – 3'. The forward primer for the 2<sup>nd</sup> GG mutation was 5' – GGTGCC ATAGGCAGCCGCAGGTGAACCCTGAAACTG – 3' while the reverse primer was 5' – CAGTTTCAGGGTTCACCTGCGGCTGCCTATGGCACC – 3'. These primers were

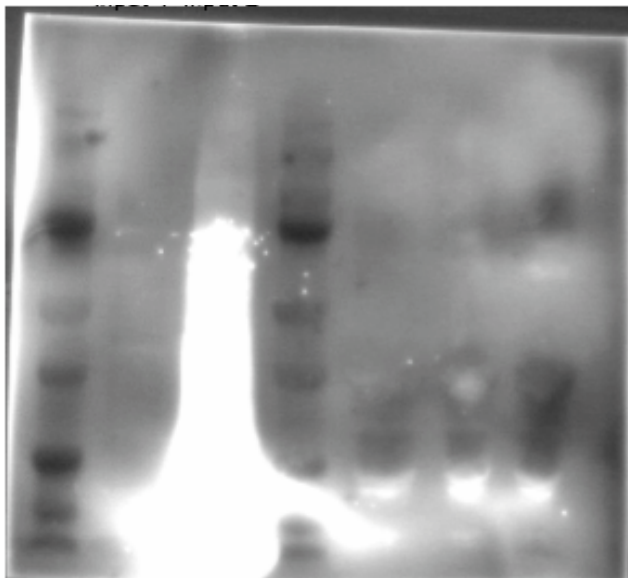
used with the wild type GPS2 DNA template. For the double mutation, the 2<sup>nd</sup> GG mutation DNA plasmid template was used and the 1<sup>st</sup> GG forward and reverse primers were used. For plasmid creation, primers were diluted with distilled water to 100uM. Primers were then diluted to ten micromolar concentration in 1.5mL tubes in preparation for individual PCR reaction experiments. Four PCR vials of PCR reaction formula were made to contain fifty microliters total. Verification from Macrogen showed successful plasmid construction for each mutation type. These plasmids were then transferred to BL21 *E. coli* cells in preparation for purification after initial binding assay results.

#### **TnT and Binding Assay Results:**

As we already have all types of purified gps2, next we need to express another factor for *in-vitro* binding. TnT protocol was followed to construct UBC13 with Myc tagging for a binding assay with GPS2 GST tagging. TnT results were analyzed by Western Blot analysis. Results were deemed successful when the appropriate band size (~18 kDa) appeared on Western Blot (Figure 11). TnT results showed UBC13 had been successfully transcribed and translated by the TnT process. After obtaining UBC13 protein in TnT solution, an *in vitro* binding assay between GPS2 and the UBC13 TnT product was performed as a positive control as it is known that GPS2 and UBC13 bind *in vivo*. Positive results for the binding assay between UBC13 and GPS2 were unable to be attained (Figure 12).



**Figure 11: UBC13 TnT Results**  
**Lane 1: Ladder, Lane 2: UBC13-TnT (18 kDa)**



No corresponding  
band sizes present

**Figure 12: *In-Vitro* Binding Assay Results UBC13 (TnT) and Optimized Full-Length GPS2 (purified)**

**Lanes left to right: ladder, input 1, input 2, ladder, elution 1, elution 2, elution 3.**  
**Elution 1: 5ul TnT+5ul beads, Elution 2: 10ul TnT+5ul beads, Elution 3: 20ul TnT+5ul beads**

## DISCUSSION

GPS2 has been identified as an important protein in both genomic and non-genomic functions by means of the inhibition of K63 ubiquitination. Through inhibiting UBC13, important pathways associated with the regulation of insulin, tumor suppression, and inflammation are affected. Recent studies have confirmed that GPS2 is an inhibitor of the UBC13/TRAF2/TRAF6 activated pro-inflammatory pathways, and UBC13 activated AKT pathways associated with insulin regulation (Bi H. et al., 2014; Fan R. et al., 2016). Accordingly, GPS2 could possibly be an important site of interest due to its downstream effects on metabolic and inflammatory pathways that influence obesity and various cancers. In this study, we set out to dissect the UBC13-GPS2 binding in order to determine the localization of GPS2 minimal interaction domain required for binding and inhibition of UBC13. The goal of my project was to perform protein expression and purifications from *E. Coli* for performing in vitro binding assays.

The first thing I accomplished was to optimize the constructs required for GPS2 expression. While investigating which forms and protein fragments of GPS2 can be purified in quantities high enough to be detected and used for *in-vitro* binding assay, we discovered that full length GPS2 proteins and protein fragments larger than 25 kilodaltons in length were difficult to express and required optimization for successful *E. coli* protein expression (Table 1). We conclude that, due to there not being a corresponding endogenous GPS2 protein in *E. coli*, the bacteria has difficulty expressing this protein and its protein fragments.

I constructed optimized plasmid sequences based on an optimized codon attained from the Invitrogen Geneart Synthesis database by Thermo Fisher. GPS2-1<sup>st</sup> GG and GPS2-2<sup>nd</sup> GG were then successfully purified with their optimized sequences. Unfortunately, the full-length version of GPS2 only ever achieved minimal – approximately 1-10 ng – purification results (Figure 10; Figure 11-14; Table 1). This, we conclude, is due to the fact that GPS2 is not endogenous to *E. coli*. When tasked with expressing non-endogenous gene products, *E. coli* often makes misfolded proteins (Lesley S. et. al, 2002). A common response to the cellular stress of accumulating misfolded proteins is a molecular switch from production of the misfolded protein to induction of protein pathways meant to deal with the translational stress of misfolded proteins (Lesley S. et. al, 2002). This likely affected the levels of protein expression in the present study.

In the process of determining the GPS2-UBC13 binding site with *in-vitro* binding assay, we were only able to get positive UBC13 expression after one TnT Quick Coupled Transcription/Translation trial. Further, we were unable to confirm binding in our positive binding control between the full length optimized GPS2 and UBC13. This, we conclude, is also a result of improper folding of GPS2 into its secondary protein structure. Because secondary and tertiary protein structure can be crucial to protein interaction, the incorrect or lack of folding of GPS2 may seriously inhibit its ability to bind to UBC13. As a result, showing the positive control binding results with full length GPS2 and UBC13 was likely affected by both minimal GPS2 protein expression and incorrect folding. Accordingly, binding assays could not be performed with GPS2 plasmids with

mutations at its 1<sup>st</sup> GG, GPS2 with mutations at its 2<sup>nd</sup> GG, and GPS2 with mutations at both GG sites in order to further test the importance of the GG sites to GPS2-Ubc13 binding. Further work on this study requires investigation on how to optimize large endogenous protein expression and proper protein folding in *E. coli*.

## REFERENCES

- Bi, H., Li, S., Wang, M., Jia, Z., Chang, A. K., Pang, P., and Wu, H. (2014) SUMOylation of GPS2 protein regulates its transcription-suppressing function. *Molecular Biology of the Cell* **25**, 2499–2508.
- Cardamone, M. D., Krones, A., Tanasa, B., Taylor, H., Ricci, L., Ohgi, K. A., Glass, C. K., Rosenfeld, M. G., and Perissi, V. (2012) A protective strategy against hyperinflammatory responses requiring the nontranscriptional actions of GPS2. *Molecular Cell* **46**, 91–104 .
- Cardamone, M. D., Tanasa, B., Cederquist, C. T., Huang, J., Mahdavian, K., Li, W. R., and Perissi, V. (2018) Mitochondrial Retrograde Signaling in Mammals Is Mediated by the Transcriptional Cofactor GPS2 via Direct Mitochondria-to-Nucleus Translocation. *Molecular Cell* **69**, 757-772.
- Cederquist, C. T., Lentucci, C., Martinez-Calejman, C., Hayashi, V., Orofino, J., Guertin, D., Fried, S. K., Lee, M. J., Cardamone M. D., and Perissi, V. (2016) Systemic insulin sensitivity is regulated by GPS2 inhibition of AKT ubiquitination and activation in adipose tissue. *Molecular Metabolism* **6**, 125-137
- Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989) A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**, 1576-1583.
- de Chasse, B., Navrati, V., Tafforeau, L., Hiet, M. S., Aublin-Gex, A., Agaugé, S., Meiffren, G., Pradezynski, F., Faria, B. F., Chantier, T., Le Breton, M., Pellet, J., Davoust, N., Mangeot, P. E., Chaboud, A., Penin, F., Jacob, Y., Vidalain, P.O., Vidal, M., André, P., Rabourdin-Combe, C., and Lotteau, V. (2008) Hepatitis C virus infection protein network. *Molecular Systems Biology* **4**, 230-247.
- Deng L., Wang C., Spencer E., Yang L., Braun A., You J., Slaughter C., Pickart C., and Chen Z. J. (2000) Activation of the IKB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **103**, 351-361.
- Eddins, M.J., Carlie C.M., Comez K. M., Pickart C. M., and Wolberger C. (2006) Mms2-Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature Structural & Molecular Biology* **13**, 915-920.
- Fan, R., Troubal, A., Goñi, S., Drareni, K., Huang, Z., Alzaid, F., Ballaire, R. Ancel, P.,

- Liang, N., Damdimopoulos A., Hainault I., Soprani A., Aron-Wisnewsky J., Foufelle F., Lawrence T., Gautier J.F., Venter N., and Treuter E. (2016) Loss of the co-repressor GPS2 sensitizes macrophage activation upon metabolic stress induced by obesity and type 2 diabetes. *Nature Medicine* **22**, 780-791.
- Finley, D., Sadis, S., Monia, B. P., Boucher, P., Ecker, D. J., Crooke, S. T., and Chau, V. (1994) Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Molecular Cell Biology* **14**, 5501-5509.
- Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. *Annual Review Biochemistry* **67**, 425-479.
- Hodge, C., Spyropoulos, L., and Glover, J. N. (2016) Ubc13: the Lys63 ubiquitin chain building machine. *Oncotarget*. **7**, 664471-664504.
- Hofmann, R.M. and Pickart, C.M. (1999) Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**, 645-653.
- Huang, X. D., Xiao, F. J., Wang, S. X., Yin, R. H., Lu, C. R., Li, Q. F., Liu, N., Zhang, Y., Wang, L. S., and Li, P. Y. (2016) G protein pathway suppressor 2 (GPS2) acts as a tumor suppressor in liposarcoma. *Tumour biology : The Journal of the International Society for Oncodevelopmental Biology and Medicine* **37**, 13333-13343.
- Huttlin, E. L., Ting, L., Bruckner, R. J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., Dong, R., Guarani, V., Vaites, L. P., Ordureau, A., Rad, R., Erickson, B. K., Wüthrich, M., Chick, J., Zhai, B., Kolippakkam, D., Mintseris, J., Obar, R. A., Harris, T., Artavanis-Tsakonas, S., Sowa, M. E., De Camilli, P., Paula, J. A., Harper, J. W., and Gygi, S. P. (2015) The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425-440.
- Jakobsson, T., Venter, N., Toresson, G., Damdimopoulos, A. E., Ehrlund, A., Lou, X., Sanyal, S., Steffensen, K. R., Gustafsson, J. A., and Treuter, E. (2009) GPS2 is required for cholesterol efflux by triggering histone demethylation, LXR recruitment, and coregulator assembly at the ABCG1 locus. *Molecular Cell* **34**, 510-519.
- Khoo, O. and Suntrarachun, S. (2012) Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pacific Journal of Tropical Biomedicine* **2**, 159-166.

- Lentucci, C., Belkina, A., Cederquist, C., Chan, M., Johnson, H., Prasad, S., Lopacinski, A., Nikolajczyk, B., Monti, S., Snyder-Cappione, J., Tanasa, B., Cardamone, M., and Perissi, V. (2016) Inhibition of Ubc13-mediated Ubiquitination by GPS2 Regulates Multiple Stages of B Cell Development. *Journal of Biological Chemistry* **292**, 2754-2772.
- Lesley, S., Graziano, J., Cho, C., Knoth, M., Klock, H. (2002) Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Engineering* **15**, 153-160.
- Meinzel T., Serero A., Giglione C. (2006) Impact of the N-terminal amino acid on targeted protein degradation. *Biological Chemistry* **387**, 839-851.
- Nieuwkoop, T., Claassens, N., Oost, J. (2019) Improved protein production and codon optimization analysis in *Escherichia coli* by bicistronic design. *Microbiology Biotechnology* **12**, 173-179.
- Sanyal, S., Bavner, A., Haroniti, A., Nilsson, L. M., Lundasen, T., Rehnmark, S., Witt, M. R., Einarsson, C., Talianidis, I., Gustafsson, J. A., and Treuter, E. (2007) Involvement of corepressor complex subunit GPS2 in transcriptional pathways governing human bile acid biosynthesis. *Proceedings of the National Academy of Sciences of United States of America* **104**, 15665-15670.
- Spain, B. H., Bowdish, K. S., Pacal, A. R., Staub, S. F., Koo, D., Chang, C. Y., Xie, W., Colicelli, J. (1996) *Molecular Cell* **16**, 6698-6706.
- Spence, J., Sadis, S., Haas, A.L., and Finley, D. (1995) A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Molecular Cell Biology* **15**, 1265-1273
- Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M., and Finley, D. (2000) Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**, 67-76
- VanDemark, A. P., Hofmann, R., Tsui, C., Pickart, C., and Wolberger, C. (2001) Molecular Insights into Polyubiquitin Chain Assembly: Crystal Structure of the Mms2/Ubc13 Heterodimer. *Cell* **105**, 711-720
- Zhang, J., Kalkum, M., Chait, B. T., and Roeder, R. G. (2002) The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Molecular Cell* **9**, 611-623

## VITA

