5-2015

Molecular Mechanisms of Tandem CBD of Clostridium histolyticum

Dawn Weir
University of Arkansas, Fayetteville

Molecular Mechanism of Tandem CBD of *Clostridium histolyticum*

An Honors Thesis submitted in partial fulfillment of the requirements of

Honors Studies in Chemistry

By

Dawn Weir

Spring 2015

J. William Fulbright College of Arts and Sciences

The University of Arkansas

Acknowledgements

I would first like to thank my mentor, Joshua Sakon, for all the support and advice that he has provided over the past three years. Through the good times and the bad he has been there to encourage me, not only in my research, but in all areas of my life. I have a large amount of respect for this mentor, who has shown me how to be a patient, kind, and hardworking scientist.

I would also like to acknowledge Ryan Bauer and Kate Janowska for teaching me proper lab techniques and assisting me whenever I had questions. Also, I would like to thank Keisuke Tanaka (Nippi Inc.) who developed and taught me the crosslinking protocols. I am also thankful for the University of Arkansas Mass Spectrometry facility employees, J. Lay & R. Liyanage, who worked so diligently to provide me with results.

I would also like to thank Osamu Matsushita and Takehiko Mima for allowing me to visit Okayama Japan for the Medical Research Exchange Program. With their help and patience, I was able to learn so much about my research and the culture of Japan. I will never forget their generosity and kindness.

My trip to Okayama, Japan would have never been possible without the funding from the Honors College Undergraduate Research Grant and Honors College Travel Grant. Additionally, I would like to thank ABI, NIH-COBRE, NIH-INBRE, and Nippi Inc. for their funding and resources.

Table of Contents **Page**

I. Abstract

In order to spread infections, bacterial collagenases methodically unravel collagen fibril in tissues. Collagen is the most abundant protein in the body, and can be found in the skin, bone and cartilage [1]. Two collagenases, ColG and ColH, synergistically dismantle collagen fibrils by seeking different weak links in the collagen structure. The collagen-binding domain (CBD) of these collagenases binds to most vulnerable regions in collagen [8]. Without CBDs, collagen fibril cannot be degraded.

Cells express collagen receptors in order to anchor themselves, which is a critical step in cell proliferation. Binding sites for some collagen receptors, such as integrin and interleukin-2, have been determined [4]. The goal of my project is to determine where tandem CBD from ColG binds on rattus type I collagen. Out of the more than 20 different types of collagen in the human body, this type of collagen is the most abundant [4]. Biophysical studies using collagen-like peptides have demonstrated that the CBD targets less tightly wound (undertwisted) regions in the triple-helix [8]. Therefore, it is expected that the CBD also targets proline poor regions in collagen. It is also expected that the CBD does not interfere with binding of important anchoring proteins based on previous studies of ColH.

In order to determine whether these hypotheses are correct, my approach is to use a photo-reactive crosslinker to freeze the CBD-collagen complex [16]. The crosslinker has a sulfhydryl reactive functional group. Adjacent to the collagen-binding cleft in CBD, a cysteine mutation was introduced by means of an overlap extension polymerase chain reaction (PCR). This cysteine mutation allows for the crosslinker to be attached by forming a disulfide bridge. After performing the PCR reactions, the mutant proteins were

successfully overexpressed and purified. Following purification, several crosslinkers were attached to the cysteine residue, and the results were confirmed using mass spectrometry. These attached photoreactive crosslinkers consists of a photo-reactive functional group. After UV light exposure the cross-linker should be activated to attach to the collagen forming a CBD-collagen complex. Future experiments then subjected the complex to digestion by trypsin, and the tryptic fragments were analyzed by LC-MS/MS (ESI mass spectrometry) to identify the location of CBD binding.

Fusion proteins of ColH CBD and signal molecules, such as cytokines, hormones and growth factors, anchor themselves to rapidly remodeling collagen [15]. This anchoring promotes tissue engineering and wound repairs *in vivo* [12, 13]. It also reduces the amount of signal molecules required to achieve the desirable outcome while at the same time reducing side effects. Although the tandem CBD from ColG is safe *in vitro*, fusion proteins of the tandem CBD and signal molecules have not yet been made and tested *in vivo*. My findings may contribute to the continuum of research to lead to efficacious treatment options for wound repair and tissue engineering applications.

II. Introduction

Collagen is the most abundant protein in the body, and is distinct from other proteins in that it comprises a triple helical structure made from three polypeptide chains [1]. In order for the chains to form a triple helical structure, the amino acid Glycine (Gly) has to be at every third position of the amino acid sequence [2]. This requirement causes a repetitive Gly-X-Y sequence to be seen throughout most collagen amino acid sequences, where X and Y are frequently seen to be the imino acids, proline and

hydroxyproline [1]. The sporadic placement of this sequence leads to variability in the structure and stability. [3]

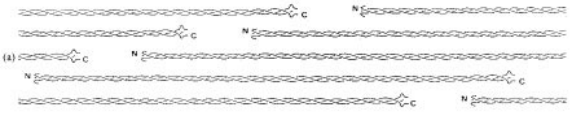


Figure 1. Diagram of triple-helical collagen molecules in a fibril that was determined by using positive and negative stains to produce a pattern [1].

There are more than 20 different types of collagen, but collagen types I, II, III, V and XI are the most abundant in the body [1]. Type I collagen is found throughout the body, except for in structures made out of cartilage [1]. This type of collagen typically interacts with cell surfaces, extracellular matrix molecules, growth factors and differentiation factors [4]. As the most abundant type of collagen in the body these interactions lead to maintaining the integrity of most tissue samples [4]. Type II collagen can be located in the cartilage, vitreous humour, and also during development in the cornea [1]. Type III collagen is a major component of the extracellular matrix of the skin, and is also found in the walls of arteries and many hollow internal organs, including the intestine, liver and spleen. [1, 5, 6] It usually occurs in the same fibril with type I collagen. Liu *et al.* showed that type III collagen plays a critical role in the development of fibrils during the development of organs, because in the absence of type III collagen organs failed to develop properly leading to functional failure [5]. The final two types of collagen are minor components of tissue. In tissues, type V collagen and type XI collagen, mostly form heterotypic fibrils with type I and type II collagen, respectively [1].

Collagen remodeling occurs during development, maintenance and regeneration of tissues [7]. Any structural or metabolic abnormalities in this process can result in various pathologic conditions and genetic diseases. Specifically, excess collagen

remodeling activity can cause tumors and arthritis [7]. Collagen fibrils can act as a major biomechanical scaffold allowing for cell attachment and anchorage of macromolecules [1].

Bacterial collagenases bind to collagen, and it is hypothesized that they also target remodeling collagen. These enzymes methodically unravel collagen fibril in tissues, to spread infections, and cause gas gangrene. Two collagenases from *Clostridium histolyticum*, ColG and ColH, synergistically dismantle collagen fibril by possibly seeking different weak links in collagen [8]. As can be seen by looking at Figure 2, both enzymes have segmental structures but the key difference between these two enzymes is that ColG has two CBDs.

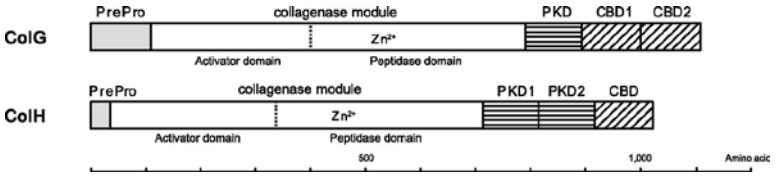


Figure 2. Collagenase ColG and ColH schematic diagram [9] .

The segmental structure for ColG includes the following domains, *i.e*. S1, S2, S3a and S3b [9]. Every segment is expected to possess a different role. The S1 is the catalytic domain that hydrolyzes the collagen. Various functions have been suggested for the S2 domain. It is proposed to be a spacer. It is also proposed to be an assistant for collagen binding and to cause collagen fibril swelling. The S3a and S3b domain anchor the enzyme to collagen [10, 11].

The collagen-binding domain of ColH (S3) has systemic effects, by anchoring fused drugs at the site of injection with prolonged half-life. A fusion protein of collagen binding domain and parathyroid hormone (PTH-s3) is being developed as a treatment option for osteoporosis. Ponnapakkam *et al.* observed that either monthly or weekly administration of PTH-CBD increased bone mineral density (BMD) in mice without

causing hypercalcemia, which is a common side effect in present PTH therapy [12, 13]. Other forms of this drug were used to help prevent chemotherapy-induced hair loss, bone fracture repair, and bone transplantation [14].

ColH domains S2b-S3 bind more tightly to collagen. Dr. Uchida, Dr. Matsushita and others demonstrated that when the collagen-binding segment of ColH (s2b-s3) is used as a drug carrier, it anchors the fused basic fibroblast growth factor (bFGF) at the site of injection with prolonged half-life [15]. The results also showed that bFGF-s2b-s3 promoted bone healing much faster than bFGF-s3 and bFGF alone [15]. Tandem CBD from ColG (S3a-S3b) is known to bind the tightest to collagen, and therefore it may be a better drug delivery vehicle than ColH (s2b-s3).

Collagen-binding cleft of CBD has been identified by X-Ray Crystallography, and it has been demonstrated by Sakon's lab that it targets loosely wound collagen-like peptide [8]. However, binding site on bona fide collagen has not been elucidated. To better control drug and extracellular matrix interaction, collagen-CBD interaction needs to be elucidated. It is expected that tandem CBD will target one of the six regions in type I collagen that are predicted to be rich in underwound triple helices. It is also expected that because remodeling collagen fibril may make undertwisted regions more accessible to CBD, it will bind to these regions, and can be valuable for diagnostic and therapeutic treatments for cancer.

III. Overview of Experimental Approach

Photoreactive crosslinking reagents are used to capture and freeze these momentary collagen-CBD contacts to study the collagen-CBD interactions [16]. The rapid reactivity on crosslinkers allows interactions to be seized in a complex stable

enough for isolation and characterization. To attach such reagent to ColGS3aS3b, cysteine mutations were introduced adjacent to collagen-binding clefts in the following mutants using overlap extension polymerase chain reactions (PCR): ColGS3a(S879C), ColGS3aS3b(Y996A, S997C), ColGS3aS3b(S997C), ColGS3aS3b(S879C), ColGS3b(Y996A, S997C).

Overlap extension using PCR can be used to introduce site-directed mutations into a wild type DNA sequence. This procedure involves two PCR reactions. Looking at Figure 3, the first PCR reaction involves the generation of two mutant DNA strands with overlapping ends. In order to perform this reaction, two complementary primers (Forward and Reverse) must be designed to have the desired nucleotide mutation. A wild type DNA plasmid is used as a base template. After denaturation by heat, separate vials allow the Forward and Reverse primers to anneal individually to the base template along with a complementary primer. In the case of this experiment, a 5' pGEX primer was used with the Reverse primer, and the 3' pGEX was used with the Forward primer. A DNA polymerase is used to elongate the primers until two new mutant DNA strands are formed.

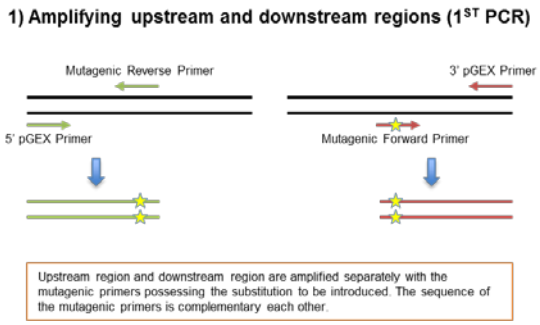


Figure 3. Production of mutant DNA fragments using PCR.

After generating the two mutant DNA strands, a second PCR is performed in order to allow for the recombination and amplification of these strands. Looking at Figure 4, the first step of this reaction is to denature the DNA allowing for the 3' complementary

ends to anneal together. This places the mutation of interest in the overlapping regions. After the two strands anneal together, DNA polymerase elongates the strands. 3'pGEX and 5'pGEX primers anneal to the strands allowing for replication of the two mutant DNA strands. This results in two complementary DNA strands. As can be seen in Figure 5, the resulting complementary DNA strands can be inserted into an expression vector after cleavage using XhoI and EcoRI-HF. By placing, the mutant CBD in the expression vector it forms a plasmid, which can then be inserted into *Escherichia Coli* (*E. Coli)* bacterial cells for DNA production.

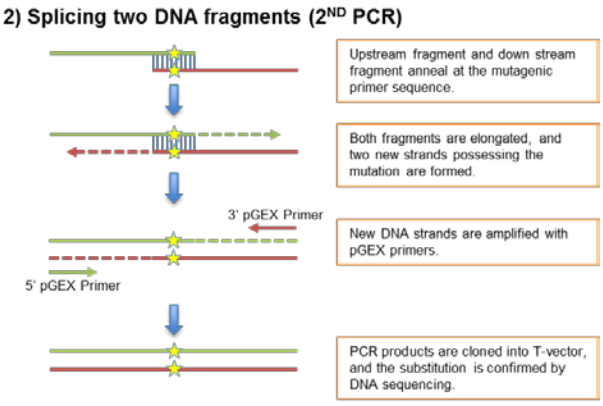


Figure 4. Overlapping fragments allow for elongation of mutant DNA strands.

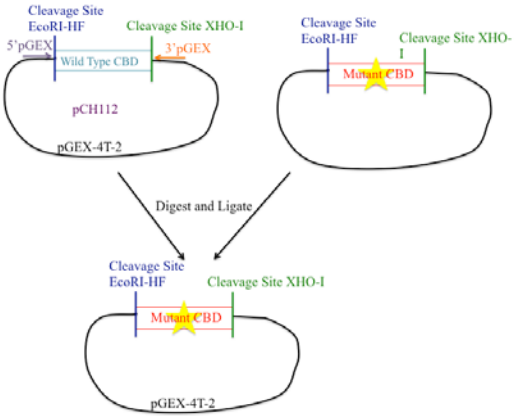


Figure 5. Cleavage by enzymes to place mutant CBD in the expression vector.

In order to confirm that the PCR produced desired results, the DNA strands can be sequenced using The Big Dye Terminator v1.1 Cycle Sequencing Kit and a sequencing machine. This method of analysis uses a DNA polymerase to reproduce the mutant DNA sequences using deoxynucleosidetriphosphates (dNTPs), which are

fluorescently labeled nucleotides. Di-deoxynucleotidetriphosphates (ddNTPs) are used to terminate DNA strand elongation. After running the prepared DNA strands in the sequencing machine, a sequencing chromatogram is produced. As shown in Figure 6, the DNA chromatogram shows different colored peaks that represent the four different nucleotides. This is because all four dNTPs have a different colored fluorescent tag. After analyzing the graph to determining the sequences, the experimental DNA strand sequences and expected DNA sequences can be compared.

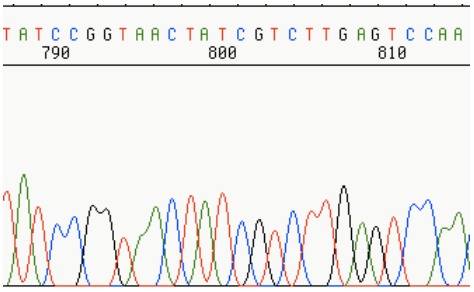


Figure 6. Example of a DNA chromatogram obtained after sequencing.

After confirming the sequence of the mutations, the plasmids were introduced into *E. coli* bacteria in order to produce the proteins. These proteins were purified using GST-tag and glutathione sepharose beads. Glutathione sepharose beads have a high affinity for GST, and allow for the binding of GST tagged proteins to the column. After cleaving the GST tag off the proteins, the GST tag will remain bound to the column, and the eluent will contain purified protein with little to no traces of impurities.

If a large amount of impurities were left in the eluent, the protein was further purified using a strong anionic purification column using Q-sepharose beads. Q-sepharose beads are positively charged and allow for binding of negatively charged molecules. A buffer without salt is added to collect any protein that does not bind to the Q-sepharose beads. Afterwards, a sodium chloride (NaCl) concentration gradient is used to slowly elute the protein from the column.

Each mutant protein was produced for a different purpose. First, in order to determine where the S3a segment interacts with collagen, ColGS3a(S879C) and ColGS3aS3b(S879C) were produced. Also, in order to determine where the S3b segment binds on collagen ColGS3aS3b(S997C) was created. Finally, in order to rule-out any nonspecific binding sites, ColGS3aS3b(Y996A, S997C) and ColGS3b(Y996A, S997C) were produced. These last two mutants should not bind to collagen.

The mutation of focus has been ColGS3aS3b(S997C). A binding assay was performed, and it was determined that ColGS3aS3b(S997C) binds to collagen, as well as the wild type. As mentioned before, this protein consists of two collagen binding domains. These domains exist tandem to each other probably as a result of gene duplication, and as a result binds tightest to collagen fibrils [12]. It is expected that both domains bind to two tropocollagen molecules, and therefore has the potential of being a longer lasting drug delivery carrier.

In order to assess the potential for the tandem CBDs to be a long lasting and localized drug delivery carrier photoreactive crosslinkers were successfully attached to the cysteine mutation. The four crosslinkers used include fluorescein-5-malemide, MTS-BP-Bio Biotin, N-[4-(p-azidosalicylamido) butyl]-3'(2'-pyridyldithio)propionamide (APDP), and azide *N*-((2-pyridyldithio)ethyl)-4-azidosalicylamide (PEAS).

There are three key elements of the APDP and PEAS crosslinkers that can be seen in Figure 7. First there is a cysteine reactive portion that will react with the introduced cysteine mutation in the protein. Second there is a photoreactive element that can be excited by UV light (usually at <360nm) and will bind to collagen. Finally, there is a

disulfide bridge present in both of these crosslinkers that can be reduced and has the potential of being helpful in future experiments.

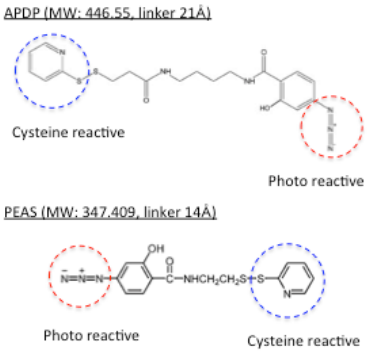


Figure 7. APDP and PEAS crosslinkers; diagram created by Keisuke Tanaka.

The MTS-BP-Bio biotin crosslinker was initially used because it was expected to help purify the CBD complex. There are four major components of this crosslinker that are shown in Figure 8. Like APDP and PEAS, there is a cysteine reactive portion that can bind to the cysteine mutant, and a photoreactive portion that becomes reactive upon UV exposure. Also, present is the disulfide bridge. Unique to this cross-linker is the biotin substituent that can bind to streptavidin beads in a purification column.

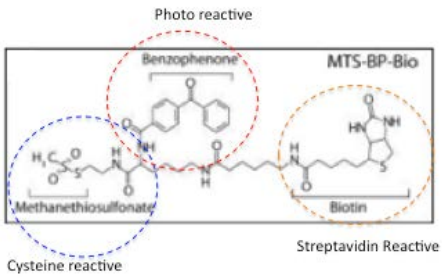


Figure 8. MTS-BP-Bio crosslinker figure.

Fluorescien-5-malemide was initially used to study the distribution of tandem CBD. However, previous *in vivo* work published by Dr. Sakon demonstrated that CBD targeted skin, bone and intestine [13, 14]. It was hypothesized that CBD may preferentially target rapidly remodeling collagen. In order to determine where CBD binds *in vitro*, CBD was labeled with fluorescein-5-maleimide (seen in Figure 9). Fluorescein appears fluorescent under UV light, and therefore can be clearly seen under microscope. The compound allowed

fluorescent labeling of sulfhydryl group on CBD. Fluorescein labeled CBD was purified from unreacted fluorescein-5-maleimide by using size exclusion chromatography column. It was then introduced into pancreatic cancer cells to observe binding.

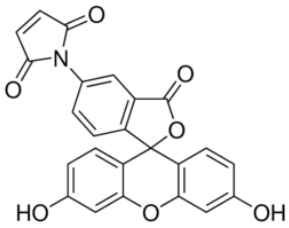


Figure 9. Fluorescein-5-maleimide crosslinker.

In order to attach cysteine reactive crosslinkers to the CBD, the disulfide bridges in the CBD must be reduced using dithiothreitol (DTT), a common redox reagent. After reduction and dialysis the protein was analyzed by 12% SDS-PAGE. If it was determined that the crosslinker was successfully attached, then the crosslinker attached CBD undergoes cross-linking to solubilized tropocollagen from rattus source. As can be seen in Figure 10, the cross-linker will bind to collagen after UV exposure. Once exposed to UV light, the collagen and collagen complex are separated from each other by running SDS gel. Faint bands above β and α bands of collagen correspond to the crosslinked complex. The crosslinking assay can then be analyzed by 2-D electrophoresis using 5% SDS-PAGE to determine the location of linkage.

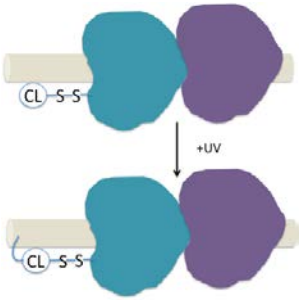


Figure 10. Diagram of cross-linking assay.

After SDS-PAGE confirmation, the crosslinking assay results are produced in large quantity in order to obtain enough sample to analyze by liquid chromatography electrospray ionization (ESI) tandem mass spectrometry (LC-MS/MS). SDS-PAGE gel spots corresponding to the crosslinked complex were reduced/alkylated using DTT/ iodoacetamide before the in gel trypsin digestion. Looking at Figure 11, reduction by DTT releases tandem CBD from the crosslinker. Part of the crosslinker is left behind as the signature to be identified by LC-MS/MS. Alkylation modifies the thiol group by attaching iodoacetamide to form carbamidomethylated crosslinker. Trypsin hydrolyzes the peptide at the C-terminal side of arginine (R) and lysine (K), forming small peptide fragments. These small peptides make identifying the binding site from collagen much easier by LC-MS/MS. The individual tryptic peptides undergo MS/MS fragmentation in order to derive the amino acid sequence information corresponding to individual tryptic peptides from the collagen.

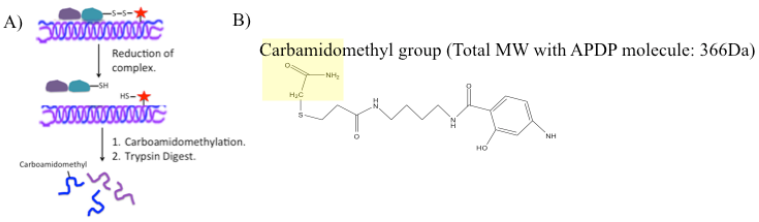


Figure 11. Work flow used in LC-MS/MS analysis to determine the CBD binding site in collagen.

Collagen has several possible post translational modifications such as P and K (oxidation), N and Q (deamidation), K (galactosylation), and K (glucosylgalactosyl). This makes it difficult to map the sequence of collagen because there can be many different post translational modification combinations located in one tryptic peptide. Therefore, it is necessary to not only perform LC but also perform on the fly MS/MS to map the sequence and modifications exits in any given tryptic peptide. LC was performed by

using reverse phase nano column. Reversed phase chromatography involves using a hydrophobic stationary phase, that includes regular phase silica particles modified by –$(CH_2)_{17}$CH3 chain or C18 bonded silica. This non polar stationary phase allows for hydrophobic molecules in the mobile phase to be retained longer, while allowing hydrophilic molecules in the mobile phase to elute first. After being separated by the reverse phase liquid chromatograph, ESI-MS/MS will identify the collagen tryptic peptides with the crosslinker attached by looking at the mass differences. For instance, the APDP crosslinker with the carbamidomethyl group changes the tryptic peptide mass by 366Da. This modification can be searched in the data base to identify the tryptic peptide with APDP.

MASCOT software from matrix science (http://www.matrixscience.com) and skyline software from (http://skyline.gs.washington.edu) are used to filter MS/MS data. MASCOT or skyline software uses mathematical algorithms and some statistical software to identify peptides with at least 95% confidence. For the peptide to be considered a reliable identification, both MASCOT and skyline had to identify the peptide as a good identification. False positive identifications are filtered out by searching MS/MS data for the α_1 and α_2 chains in the same way as the MS/MS data from the crosslinked complex.

IV. Methods

A. Mutagenesis Projects by Overlap Extension PCR

In order to introduce the desired mutations into the amino acid sequence of the wild type DNA, primers were designed that could be amplified using polymerase chain reactions (PCR). In order to design these primers, the DNA sequence for the specific mutagenesis was determined and analyzed in order to determine which codon for the

desired amino acid mutation should be introduced. Ideally, the best way to introduce a mutation is by only changing one of the nucleotides because this increases the likelihood that the wild type DNA will be successfully mutated.

For ColG3aS3b(S997C), the reverse (R) and forward (F) primers designed were as follows, :

R: AACTCATAGTTTCCTGATCCGCAGTATTTATAAACAAGTAGAT

F: ATCTACTTGTTTATAAATACTGCGGATCAGGAAACTATGAGTT

The original DNA sequence that was altered using PCR was the vector pCHG112, which represents the sequence for the wild type ColGS3aS3b.

For ColG3aS3b(Y996A/S997C) and ColGS3b(Y996A/S997C), one primer was designed because the mutations of interest is located in the S3b CBD for both mutations. The primers designed consisted of the following sequences:

R:

TAATAGATGAACAAATATTTCGGAGTCCTAGTCCTTTGATACTCAA

F: ATTATCTACTTGTTTATAAAGCCTCAGGATCAGGAAACTATGA

Additionally, these primers only contain an alanine mutation because the original DNA sequence that was used in the PCR reaction was ColG3aS3b(S997C) and ColGS3b(S997C) respectively. These primers only changed two of three nucleotides in the original sequence from tyrosine (TAC) to alanine (GCC).

For ColG3aS3b(Y994A/S997C), the designed primers were as following:

R:AGCAACAAGTAGATAATATTTTC

F:GAAAATATTATCTACTTGTTGCTAAATACTGCGGATCAGGAAA

The original DNA sequence that was altered was ColGS3aS3b(S997C).

All of these primers were ordered through the company Fasmac Bio (Okada, Atsugi Kanagawa). After receiving the primers, the tubes containing the primers were centrifuged and autoclaved distilled water (MilliQ) was added to solubilize the primers. Working primers were made by taking ten microliters of each primer and diluting it with MilliQ water for a total amount of 100µL working primer.

The reaction mixtures for the first round of PCR were prepared by placing 10µL 5x PrimeSTAR Buffer, 4µL dNTP Mixture, 1µL F Primer (R Primer), 1µL pGEX primer (5' pGEX primer), 1µL DNA Plasmid, and 0.5µL DNA polymerase (PrimeSTAR Enzyme) into two small PCR vials. The reaction vials were then placed in a thermo cycler at 98°C for 10 seconds, 55°C for 5 seconds, and 72°C for 1min/kilobase for a total of 30 cycles.These reaction mixtures formed two mutant DNA strands that were used in a second PCR reaction.

For each mutagenesis project, a different starting DNA plasmid was used. For ColGS3aS3b(Y996A/S997C) and ColGS3aS3b(Y994A/S997C), DNA plasmid E022 was used as the base template. This plasmid was created by Ryan Bauer and consists of the DNA sequence for ColGS3aS3b(S997C). Additionally, for ColGS3b(Y996A/S997C), the DNA plasmid E006 served as the base template. This plasmid was also created by Ryan Bauer and contains DNA plasmid for ColGS3b(S997C). Before using these plasmids, the sequences were confirmed by sequencing the DNA.

Agarose gel electrophoresis was used to analyze and confirm that the amplified DNA plasmids contained the proper amount of base pairs. In order to perform the agarose gel electrophoresis, 1.5% agarose gel was created by measuring out 1.5 g of agarose. 100 mL of Tris/Borate/EDTA (TBE) buffer was added to the Erlenmeyer flask, and the

solution was microwaved for five minutes. Before solidification the mixture was poured into the agarose gel-casting stand. After solidification the gels were removed and placed in a container of TBE solution.

In order to conserve resources, the gel was cut in order to contain the required amount of wells for analysis. The gel was then placed in the horizontal electrophoresis unit. A small amount of amplified primer was mixed with 1:5 volume of primer to blue dye. The two amplified primers were placed into two separate wells, and a DNA molecular weight marker was added to a third well. The electrophoresis unit was allowed to run for approximately 30 minutes. The unit was turned off before the sample could run off the agar gel. The gel was placed into ethidium bromide. From this point on gloves were used to handle the gel because ethidium bromide is a carcinogen. After soaking in the ethidium bromide solution for 15 minutes, the gel was washed off with water and placed into a UV machine that contained a camera. A picture was taken of the gel exposed to UV light, in order to analyze the bands for the correct number of DNA base pairs.

After confirmation that the primers contained the proper amount of base pairs, the DNA plasmids were purified using the QIAqucik PCR Purification Kit. In order to use this kit, two 1.5 mL vials were labeled with reverse and forward so that the solutions could be recognized. The entire PCR mixture from the first PCR was placed into the appropriately labeled 1.5 mL vial. The instructions from the kit were followed until step 7, in which water was used instead of Buffer EB. At the end of the purification, the flow through was collected and a NanoDrop UV-Vis spectrophotometer was used to determine the concentration of DNA in the sample.

In order to join the overlapping ends of the forward and reverse primers, a second PCR reaction was performed. In order to perform this reaction a small PCR vial was prepared containing 10µL 5x PrimeSTAR Buffer, 4µL dNTP Mixture, 1µL 5' pGEX primer, 1µL 3' pGEX primer, forward primer (total concentration of 50ng/ 50µL), reverse Primer (total concentration of 50ng/50µL), 1µL DNA Plasmid, and 0.5 P DNA polymerase (PrimeSTAR Enzyme). After mixing these solutions the total volume was adjusted to 50µL using MilliQ water. The reaction vials were then placed in a thermo cycler at 98°C for 10 seconds, 55°C for 5 seconds, and 72°C for 1min/kilobase for a total of 30 cycles.

The vials were then left in the thermocycler at 4°C overnight. The next day the sample was analyzed by agarose gel electrophoresis using the already mentioned protocol. If the sequence was correct the QIAquick PCR Purification was used to purify the sample. Afterwards, an addition reaction was performed by mixing small PCR vials containing 8µL PCR Product, 1µL 10x Buffer, 0.5µL dATP, and 0.5nd A-Overhang Enzyme. The reaction mixture was then incubated at 65°C for ten minutes. Then 1 µL of the addition reaction mixture was added directly to a Ligation Mixture, which contained 1µL pMD20-T Vector, 3µL MilliQ water, and 5µL Ligation Mighty Mix. This mixture was mixed by flicking the vial with a fingertip, centrifuged and then placed in an incubator at 16°C for thirty minutes.

Afterwards the plasmids were introduced into DH5α *E.Coli* cells. The DH5α cells were used to obtain DNA that could be sequenced to confirm that the mutation was introduced into the plasmid DNA. *E. coli* cells are extremely sensitive to heat and should be kept on ice at all times. In order to prevent contamination, the 10µL solution of

ligated mixture was introduced into the vial of DH5α *E. coli* cells underneath a gas hood. The vial was incubated on ice for thirty minutes, transferred to an incubator at 42°C for ninety seconds and then placed on ice for two minutes. Once again in the hood, 900 μL of SOC media was added to the cells. The mixture was then incubated at 37°C for 1 hour.

In order to create, the SOC media 10g Bacto Tryptone, 2.5g Bacto Yeast Extract, 0.25g Sodium Chloride and 0.093g Potassium Chloride were dissolved in 400 mL of deionized water. The pH was adjusted to 7.0 using Sodium Hydroxide. After adjusting the pH the total volume of the solution was brought up to 500mL. This solution was then divided into five 100mL aliquots. Before using the SOC media, 1.8mL of 20% glucose (a final concentration of 20mM) and 1mL of 1M Magnesium Chloride (a final concentration of 10mM) must be added to the 100mL aliquot.

In order to screen bacteria production, these cells were applied to two LB agarose gel with ampicillin (AMP) plates. In the hood, the plates were labeled with the mutation, date, and either Plate #1 or Plate #2. On Plate #1, 50μL of the cells were placed in the middle of the plate on a spin table. While spinning the plate, an "L" shaped instrument was used to spread the bacteria out evenly over the plate. In order to ensure proper colony formation, this must be done gently without breaking the gel. For Plate #2, the remaining solution was centrifuged for one minute at maximum speed. Then approximately 850μL was removed from the top of the *E. coli* cells. After mixing the cells with the remaining 100μL of solution, the cells were plated onto Plate # 2 following the same procedure as Plate #1. The plates were left overnight in the incubator at 37°C.

In order to prepare the LB agarose gel plates, 100mL of deionized water, 2.5g of LB Medium, and 1.5g of Agar were added to an Erlenmeyer flask. The flask was covered

with aluminum foil and then autoclaved for 15 minutes at 121°C. It was allowed to cool to a temperature that could be handled by hand. Before the mixture solidified 100μL AMP (50 mg/mL) was added to the broth. In the gas hood, the broth was then poured into the plates until the plates were 20% full. The LB agarose was allowed to solidify and then the plates were sealed in a plastic bag to prevent dehydration of the gel.

After allowing the colonies to grow to a sufficient size, two 10mL vials were prepared by labeling #1 and #2. In the hood, 5mL of LB broth was added to both of the vials. Then, 5μL of AMP was added to the broth. After sterilizing an inoculating loop by fire, the loop was used to transfer one colony into vial #1. The loop was sterilized again and a different colony was transferred to vial #2. The vials were placed in a shaker at 37°C to be left overnight. Two different colonies were incubated in order to increase the likelihood of having a correct sequence.

After allowing the culture to grow overnight, the bacterial broth was purified in order to obtain DNA plasmid that could be sequenced. In order to do this 1mL of both cultures were placed in two 1.5 mL vials labeled #1, for Culture #1, and vial labeled #2, for Culture #2. These vials were centrifuged for one minute at high speed. The supernatant was dumped off the top of the bacterial cells, and then another 1 mL portion was added to each vial. The vials were once again centrifuged for one minute at high speed. The SV Mini Preps Isolation kit was used to isolate the plasmid DNA from the protein. The concentration of DNA in the sample was recorded using the NanoDrop UV-Vis spectrophotometer.

After collecting the plasmids a 1.5% agarose gel analysis was performed in order to confirm the number of nucleotides in the DNA sequence. In order to analyze by agar

gel the DNA needed to be cut using the two enzymes EcoRI-HF and XHOI. The following reagents were mixed in a vial containing, ~200ng/20µL Plasmid DNA, 2µL10x buffer (CutSmart B), 1µL Enzyme 1(EcoRI-HF), and 1µL Enzyme 2 (XHOI). One vial was made for each culture. After adding MilliQ water to make the total volume 20µL the solutions were mixed and incubated at 37°C for one hour. The electrophoresis was ran in the same way as mentioned above. After confirmation of the correct number of base pairs present in the DNA sequence, the plasmids were prepared for DNA sequencing. Also, 1200µL of the *E. coli* culture was mixed with 600µL 50% Glycerol. These cultures were then placed in the freezer after labeling with appropriate names.

If the number of nucleotides appeared to be incorrect on the gel, then a procedure known as colony PCR was performed in order to find a colony with the correct number of nucleotides. This procedure was used for mutations ColGS3b(Y996A/S997C) and ColGS3aS3b(Y996A/S997C). In order to do this eight 1.5mL Vials were labeled 1-7 and C. After labeling, 20µL of MilliQ water was added to each vial. A master plate was labeled in order to keep track of the colonies that were placed in each vial. After taking up one colony using a sterilized inoculating loop, the master plate was streaked and then the colony was mixed in the appropriate vial. This was done for vials labeled 1-7. Afterwards, a master buffer was mixed using 18µL10X PCR Buffer, 14.4µL 4X DNTP, 9µL 5' pGEX primer (20µM), 9µL 3'pGEX primer (20µM), 119.7µL deionized Water, and 0.9µL rTAQ Polymerase.

This master buffer was divided into 20µL aliquots by placing it in 8 small PCR vials. Afterwards, 1µL of bacterial water from the vials labeled 1-7 and C were placed

into each separate PCR vial. These PCR vials were placed in the thermocycler at 94°C for 5 min, 94°C for 30 seconds, 50°C for 30 seconds, 74°C for 1 minutes, and 4°C for ∞.

After this protocol was complete, a 1.5% agar gel was run in order to compare colonies. If there was a colony which had the correct number of base pairs the colony from the master plate was transferred to the LB broth and left in the incubator at 37°C overnight to restart the process.

After confirming that the DNA plasmid has the correct number of base pairs, the plasmid DNA was treated using The Big Dye Terminator v1.1 Cycle Sequencing Kit. This kit prepares the plasmid for DNA sequencing. The 3' and 5' primers were sequenced separately. ~200ng/20µL collected double stranded DNA plasmid was mixed with 8µL Terminator Ready Reaction Mix in two separate vials. In one tube labeled 3', 2 µL1.6pmol/µL 3' PGEX primer was added. In a second vial labeled 5', 2µL of 1.6pmol/µL 5'PGEX primer was added. MilliQ water was added to both vials to make the total volume 20µL. The samples were then placed in the thermocycler at 96°C for 1 min. Then, the thermocycler repeated 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 minutes 25 times. It then cooled to 4°C for ∞.

Once this protocol was done a termination sequence was used in order to get rid of extra nucleotides at the end of the sequence. A solution of 20µL PCR Product, 90µL SAM, and 20µL XT was mixed for each primer. The mixtures were placed in a vortex for 30 minutes and centrifuged at 1000G for 2 minutes. The clear solution from each vial were transferred to individual wells in a plate used for sequencing. A buffer was created using 27 mL of water with 3mL of 10x Running Buffer. This buffer was used for the sequencing machine.

After running the sequencing machine, the data was collected and a sequencing chromatogram was analyzed. If the peaks were clear then the sequence could be analyzed to determine if it is the correct sequence. If the experimental sequence was correct, then the plasmids were digested by first preparing a vial with 16μL final PCR product, 2μL Cutsmart Buffer, 1μL Eco-RI-HF, and 1μL XhoI. Then, another vial was prepared with ~500ng/20μL PGEX-4T-2, 2μL Cutsmart Buffer, 1μL Eco-RI-HF, and 1μL XhoI. These vials were incubated for an hour at 37°C, and then an agar gel was run using 20μL of DNA plasmid and 2.2μL10x loading buffer.

Two different wells are used for the electrophoresis for the 3' and 5' DNA plasmid. After finishing the electrophoresis, the gel was placed in ethidium bromide for 15 minutes. The gel was then placed under UV light so that the bands could be clearly seen. The bands were cut out of the gel using forceps and an X-Acto knife, while wearing a facemask to protect the eyes from the UV light, and gloves to prevent exposure to ethidium bromide. This procedure must be done quickly so that the UV light doesn't damage the DNA plasmids which could produce unwanted DNA mutations. The weight of two 1.5mL vials was recorded and then the 5' and 3' plasmid gel pieces were placed into the two separate 1.5mL vials. The weight of the vials with the gel piece was measured. The weight before and after were subtracted in order to determine the weight of the gel pieces. This information was then used to determine the amount of a solution that was used in Qiagan QIAquick Gel Extraction Kit protocol. After extracting the DNA from the gel following the protocol, a vial containing 10μL of insert DNA (pCHC302) and 10μL of expression vector was prepared. 17μL of this solution was placed in a new vial containing 2μL ligation buffer and 1μL T4 DNA Ligase for a total volume of 20μL.

In order to start protein production the plasmids were introduced into BL21-RIL *E. coli* cells following the same procedure that was performed for introducing the cells into DH5 α cells. The cells were then plated on LB agarose plate with Amp and CM.

The above described protocol was followed until certain problems were encountered that needed to be resolved using different experiments. While creating ColGS3b(Y996A,S997C) and ColGS3aS3b(Y996A,S997C) there was problems with getting the correct plasmid after transforming the DNA plasmid in DH5α. After purifying the DNA plasmids, the plasmids were cleaved using the enzymes Eco-RI and XhoI. After running the agar gel it was determined that the plasmids were incorrect because the plasmid was not cleaved into two parts. The first correction method tried was performing the transformation process without the TA-cloning step. In order to do this we prepared a vial containing 16μL final PCR product, 2μL Cutsmart Buffer, 1μL Eco-RI-HF, and 1μL XhoI. Another vial was prepared containing ~500ng/20μL PGEX-4T-2, 2μL Cutsmart Buffer, 1μL Eco-RI-HF, and 1μL XhoI. MilliQ water was added to this vial to make the final volume 20μL. These two vials were incubated at 37°C for 1 hour.

The sample was run in an agar gel and placed in ethidium bromide. The bands were extracted from the gel by using an X-Acto knife. In order to prevent the plasmid from being burned this extraction procedure has to be done quickly. After extracting the band, the QIAquick Gel Extraction Kit was used in order to obtain the plasmid. Then, 10μL of both vials was placed into one vial. 17μL of this sample was placed in another vial with 2μL Ligation Buffer and 1μL T4 DNA Ligase. This vial was then incubated at 16°C for 1 hour. This plasmid was then transformed into DH5α competent cells and plated on LB+ amp plates. The plates were left overnight at 37°C. If colonies were

present then the broth was inoculated in order to collect the plasmids to be sequenced. However, for the two attempts that were made at this method it didn't work.

For ColGS3b(Y996A,S997C) after repeated failures, the experiment was performed in large scale by concentrating the sample using ethanol precipitate. In order to do this 1.5µL 3M sodium acetate (pH 7.0), 37.5µL Absolute Ethanol, and 15µL of two different colonies plasmid were distributed in two different vials. The vials were placed in the freezer at -80°C overnight. The mixture was thawed and then spun down at 15K rpm for 15 minutes. The supernatant was removed and the plasmid was washed with 200µL 70% ethanol. The mixture was spun down at 15K rpm for 5 min. The plasmids were air-dried by vacuum. Then 5µL TE was added to dried plasmids and the plasmid was analyzed by SDS-PAGE. One of the plasmids appeared to have the correct cleavage, therefore the plasmid underwent sequencing to determine if the correct sequence had been introduced.

The sequence was incorrect and therefore two new colonies were chosen from the colony PCR plate. After performing the PCR reactions and running agar gel electrophoresis, the cleaved bands were not clear on the gel but the sequencing solution had already been prepared so the plasmids were sequenced. After analyzing the results, it was clear that the correct sequence had been introduced into the plasmid. The plasmids underwent the rest of the procedure without any major problems and were transformed into BL21-RIL cells to begin protein production.

For ColGS3aS3b(Y996A,S997C) the process was repeated with the TA cloning step and a plasmid was obtained that could be sequenced. After correcting this part, the steps were continued until another major problem was encountered. Before transforming

into DH5α cells, the agar gel electrophoresis did not show the desired results. Recall that several experiments had been performed to get to this point. Therefore each step had to be checked in order to determine which step was having issues. In order to do this, at each step a small sample was taken out of the product to be analyzed by agar gel electrophoresis. Also, all the vectors were analyzed by agar electrophoresis in order to confirm that they were correct. After performing the experiment in large scale by ethanol precipitate, it was determined that extraction step, where the plasmid is extracted from the agar gel pieces, was the step having issues. There were no bands at all. After repeating the extraction step with supervision to analyze the where error was introduce, it was determined that the error was caused by not removing all of the cleaning solution for the Extraction Kit. This error would prevent the DNA plasmid from leaving the gel, and therefore explains the lack of bands in the agar gel. After correcting this step the rest of the procedure was completed without problems and protein production began.

B. Protein Production

In order to start protein production, 10μL of the double stranded plasmids were transformed into BL21-RIL cells using the same procedure that was used for the DH5α cells. These cells were then plated onto LB agarose plates with AMP and chloramphenicol (CM) using the previously stated method. The plates were left overnight at 37°C, BL21 cells grow at quicker rate then the DH5α cells so the transformation process for BL21 cells was started later in the evening.

On the same day that the cells were plated, 2xYT-G medium was prepared by first dissolving 16g/L tryptone, 10g/L yeast extract, and 5g/L sodium chloride in 800mL of deionized water. The pH of this solution was adjusted to 7 using sodium hydroxide. The

volume was adjusted to 900mL and then was autoclaved at 121°C for 20 minutes. After the solution had cooled 100mL of 20% glucose was added to the medium and then agar (1.2-1.5%).

The next day, if colonies were present on the plate, protein production began. One 25 mL autoclaved Erlenmeyer flask with a pink cap was filled with 10mL of 2xYT-G, and then 10μL of CM and AMP were added. One colony was added to the flask and the culture was incubated in a shaker for 12 hours at 37°C. During this time, four 1L Erlenmeyer flasks were autoclaved at 121°C for 20 minutes. These flasks were necessary for the large scale production that occurs the next day. After the 12-hour incubation, one 300mL autoclaved Erlenmeyer flask was filled with 120mL of 2xYT-G+AMP+CM. 6mL of the incubated culture was added to the 300mL flask. The flask was left in the incubated shaker at 37°C for at least twelve hours.

The next day four 1L autoclaved Erlenmeyer flasks were filled with 500mL of 2xYT-G+AMP+CM. 25mL of the culture that was left overnight was added to each flask. These flasks were placed in the incubated shaker at 37°C. After an hour, the optical density was checked. Once the optical density reached 0.7, 5mL of IBTG was added to each flask in order to induce protein production. IBTG is stored in the freezer and therefore must be taken out of the freezer 30 minutes before needed in order to thaw. After adding the IBTG, the flasks were left in the shaking incubator for two hours. During this two hour time period, 20mL of PMSF was made by mixing 20mL of 1-isopropanol with 0.348g PMSF. Also, during this time the large centrifuge was turned on to allow time for cooling. Before using the PMSF solution, all the hairpin needles must be completely dissolved in the solution.

At the end of the two hours, 5mL of PMSF was added to each flask. Then four large centrifuge tubes were obtained. The contents of one Erlenmeyer flask could fit into two of the large centrifuge tubes. After filling four of tubes, they were balanced and then placed into the large centrifuge. The tubes were spun for 10 minutes at 60000rpm. After spinning the tubes, the supernatant from each tube was poured back into the large Erlenmeyer flask. The same centrifuge tubes were filled with the cultures from the other two flasks and the process was repeated. At this point the cells can be frozen in the freezer at -80°C and left overnight. In order to dispose of the bacterial waste, the Erlenmeyer flasks with the supernatant were autoclaved and then the waste dumped down the drain.

The next step involves suspending the cells using PMSF with PBS. This solution was made by first pouring 90mL of deionized water and 10mL of 10xPBS into a 200mL beaker. Afterwards, 1.1mL of 1-isopropanol was mixed with 0.018g of PMSF. These two solutions were mixed together. The cells were suspended using a pipette and the PMSF with PBS solution. 5mL of this solution was added to the first tube. The cells were suspended into the solution, which was then transferred to the next tube. The cells in the next tube were suspended. This was done for all the tubes. Then this process was repeated two more times in order to ensure that all the cells were present in the solution.

After suspending the cells, the French Pressure Cell Press was used to break open the cells. After loading the cells in the French Press cell, the pressure was kept constant at 640psi. In order to ensure that all the cells were broken apart the pressure had to remain constant. Then two smaller tubes were filled with the broken cells, and 1.16mL of 20% Triton was used to wash the protein of any impurities. The tubes were mixed for 30

minutes and during this time the large centrifuge was turned on in order to allow cooling. After the 30 minutes, the tubes were placed in the centrifuge. The tubes were spun down at 15000rpm for 30 minutes. The centrifugation was done to get rid of the cellular debris and to separate the protein. After the thirty minute time period, the supernatant (protein) was poured into two new tubes and they were placed in the centrifuge at 15000rpm for 30 minutes.

During this time, glutathione sepharose beads were washed using PBS. In order to make the PBS, 50mL of 10xPBS was combined with 450mL of deionized water. 10mL of glutathione sepharose beads was divided into four 15mL tubes. Then each tube was filled to the top with the PBS solution. The tubes were placed in a centrifuge at 2000rpm. The supernatant from each tube was dumped into a waste beaker while ensuring that none of the beads were lost in the waste beaker. This cleaning procedure was repeated three times. After the last round of cleaning, the protein from the large centrifuge was poured in equal amounts into the four tubes filled with beads. The beads and protein were mixed for 30 minutes.

At the end of the 30 minutes, the supernatant was poured into the waste beaker. The beads were washed five times following the same cleaning procedure with PBS. Then after the last washing, the beads were placed in a column. This column was prepared by putting all the pieces together. Then in order to ensure that no air bubbles were in the membrane of the column, a syringe filled with 10mL of PBS was connected to the stopcock. Then the PBS was forced into the column using the syringe. 6mL of PBS was left in the column and the beads were pipetted into the column.

After this an elution buffer was made by first mixing 5mL of 50mM TRIS and 1M HCL (pH 8.0) with 95mL of deionized water. Afterwards, 0.307 grams of glutathione was added to this solution. Five 15mL tubes were labeled 1-5. The column was drained of all PBS into a waster beaker. Tube 1 was placed under the column in order to collect the eluent that passed through the column after addition of 5mL of elution buffer. Tube 2 replaced Tube 1 and 10mL of elution buffer was added to the column. This was repeated for Tube 3-5. It was expected that most of the protein would come out in Tube 2. In order to test which tubes had protein in it, 200μL of a Bradford acid was poured into a small tray. 20μL of the flow through was added to the acid. The acid would turn bright blue if the solution contained protein. The tubes containing protein were then placed in the fridge at 4°C overnight.

The next day an SDS-PAGE was ran to confirm the presence of the correct protein. If it was correct then the proteins were concentrated and a Bradford Assay was performed. Materials needed for this assay include a Protein Assay Dye Reagent Concentrate (Bio-Rad, Cat. 500-006, 450mL) and 2mg/mL BSA standard. At the beginning of this procedure, 10mL of 5x diluted working dye reagent was prepared and 2mg/mL BSA standard was diluted to 0.5mg/mL. A BSA standard concentration gradient was formed using five 1.5mL vials. Table 1 was used to make the correct concentration in the 1.5mL vials.

	BSA Standard Final Concentration (microgram/tube)					
	0µg	2µg	4µg	6µg	8µg	10µg
Deionized Water	20µL	16µL	12µL	8µL	4µL	0µL
0.5mg/mL BSA	0µL	4µL	8µL	12µL	16µL	20µL

Table 1. BSA Standard Concentration preparation table.

Three samples containing the protein were created by mixing 8µL of the sample with 72µL deionized water. 20µL of this solution was placed in three 1.5mL vials, and then 1mL of the working reagent was added to each vial. The solution was mixed well and allowed to sit for 10 minutes. A spectrophotometer was used to measure the absorption at a wavelength of 595nm. Water was used as the blank sample.

Afterwards a plot was created to determine the concentration of protein. This information was then used to determine the total amount of thrombin needed to cleave the GST tags of the protein. The thrombin was added to the protein and left overnight. The next day an SDS-PAGE was ran to see if the cleavage was successful.

While the SDS electrophoresis machine was running, a dialysis buffer was made by measuring out 4mL of 500mM calcium chloride and 100mL 1M TRIS. This solution was added to 1896mL of deionized water. The buffer was then placed in the fridge to cool. A dialysis tube was washed using deionized water. The tube was checked for holes by closing one end of the tube with one hand and allowing it to fill with water. This was checked because if holes were present in the dialysis tube, the protein would be lost in the dialysis buffer with no way of recovering the protein. If no holes were present a clamp was placed at one end of the tube. The protein was pipetted into the dialysis tube and a second clamp was placed at the top of the tube. The protein in the dialysis tube was then

placed in 500mL of the dialysis buffer. The protein was left in dialysis for 48 hours with the buffer being changed every 12 hours.

At the end of this dialysis period, the protein was purified by using the same glutathione sepharose bead column. This time the purpose of the column was to separate the GST tag from the protein. This column was cleaned using three 5mL portions of glutathione. The beads were evenly distributed into 4 different 15mL tubes. The tubes were filled with 1M TRIS pH 7.5 with calcium chloride buffer following the same washing procedure that was used to originally clean the beads. The protein was then added to the column in 1mL portions. 1.5mL vials were used to collect the flow through. After each 1mL portion the solution was tested for protein using a 1x Bronstead Dye that turns blue in the presence of protein. Once a portion of the flow through turns the dye blue, a 15mL tube was placed under the column and the rest of the protein was added. After collecting all of the protein an SDS-PAGE was run.

If impurities were still present in the sample, an anionic GSH column was used to purify the protein. The beads for this column were prepared using the same washing procedure, however only 3mL of beads were used. After washing the beads and placing them in the column, the protein was added in 1mL portions. The flow through was collected in 1.5mL portions. The Bradford dye was used to determine which ones had protein. After adding all the protein to the column a salt concentration gradient column was used to push the protein through the beads. The concentration gradient was prepared according to Table 2.

Salt Gradient									
TRIS (pH7.5) $CaCl_2$	900µL	800µL	700µL	600µL	500µL	400µL	300µL	200µL	100µL
TRIS (pH7.5) $CaCl_2$ +NaCl	100µL	200µL	300µL	400µL	500µL	600µL	700µL	800µL	900µL

Table 2. Salt Gradient preparation guide.

The first salt buffer added to the column must contain the lowest salt concentration. For one protein, ColGS3aSb(S997C), the concentration gradient was performed with the highest concentration of salt first. This error was corrected by redoing the dialysis by using 330mL of TRIS with calcium chloride buffer. The dialysis period was 6 hours, while changing the buffer every 3 hours. Once the dialysis period was finished the GSH column was used again. The protein was collected following the same procedure as above while ensuring to perform the concentration gradient the proper way. After collecting all these samples the vials were placed in the fridge and left overnight.

An SDS-PAGE analysis was performed the next day to confirm that the purification columns worked. If the impurities were low enough the concentration of the protein was determined by performing a BCA assay, using the BCA assay reagent kit by ThermoScientific. In order to perform a BCA assay, a water bath was prepared at 60°C and a working reagent was prepared by mixing a 50:1 ratio of Reagent A and Reagent B from the kit. About 5 mL of this solution was needed for one person. The 2mg/mL BSA standard was diluted with deionized water to get 0.25mg/mL BSA. Table 3, was used to prepare the BSA standards.

	BSA Standard (microgram/tube)					
	0µg	1µg	2µg	3µg	4µg	5µg
Deionized Water	20µL	16µL	12µL	8µL	4µL	0µL
0.25mg/mL BSA	0µL	4µL	8µL	12µL	16µL	20µL

Table 3. BSA Standards.

Three samples of protein were created by mixing 8µL of the sample with 72µL deionized water. 20µL of this solution was placed in three 1.5mL vials and then 40µL of the working reagent was added to each sample including the BSA Standards. The vials were mixed and then placed in the water bath for thirty minutes. At the end of thirty minutes, the vials were placed in tap water to cool down for five minutes. The spectrometer was used at a wavelength of 562nm, while water was used for the blank. The readings were plotted and the concentration was determined. After determining the concentration, the protein was placed in vials and placed in the freezer in order to be sent back to the University of Arkansas.

C. Crosslinking Experiments

After returning to the University of Arkansas, ColGS3aS3b(S997C) became the focus of crosslinking experiments. In order to ensure that the protein did bind to collagen, a collagen binding assay was performed by Mary Kate Tucker. During this time, attempts were made to attach crosslinkkers, such as APDP and PEAS, to the introduced cysteine mutation. This was done by following the procedure performed by Keisuke with ColGS3b(S997C). The first step of this protocol involves reducing the disulfide bonds in the protein. A 1.5mL vial was prepared containing ~0.5mg/mL mutant CBD, 10µL 0.5 M DTT (final concentration 10mM), and 12.5µL 0.2 M EDTA (final concentration 0.5mM). PBS was added to bring the total volume to 500µL. The DTT was prepared by mixing

100μL deionized water with 8mg DTT. The solution was incubated on ice for 30 minutes. During the incubation time a PBS/EDTA dialysis buffer was created by adding 40mL 0.5M PB (pH 7.4), 44mL of 20% NaCl, and 5mL 0.2M EDTA. The total volume of the solution was brought up to 1L using deionized water. The incubated solution was then placed into dialysis devices (2 with 250μL of solution). The devices were placed in 500mL PBS/EDTA dialysis buffer and dialyzed for at least 5 hours at 4°C. At the end of this time period, the dialysis buffer was changed and the solution was dialyzed overnight (for at least five hours).

The next day, the crosslinker was introduced to the R+ mutant. Before this occurred 20μL of R+ mutant was placed in a vial and then left in the freezer to be analyzed by SDS-PAGE at the end of the crosslinking experiment. A separate vial containing 500μL Mutant (R+), 425μL PBS, 25μL 0.2M EDTA, and 50μL 10mM crosslinker was mixed. This solution was incubated for three hours at room temperature in the dark. After this step, the protein must have limited light exposure. The crosslinker is photoreactive, and therefore, too much light exposure could affect the experiment. During the three hour incubation time, two dialysis buffers were created.

The first dialysis buffer was a HEPES buffer without calcium. This buffer was created by mixing 40mL of 0.5M HEPES (pH7.5), 44mL of sodium chloride, and enough deionized water to make the final volume 1L. The second buffer was a HEPES buffer with calcium. This buffer was created the same way as the first buffer with the addition of 1mL of 1M calcium chloride. The solution was then centrifuged at 13,500rpm for 10 minutes in order to collect the extra APDP. The supernatant was then placed in 3 dialysis devices, which were then placed in the 1L of HEPES buffer without calcium. The

solution was dialyzed in the dark cold room at 4°C for at least five hours. The devices were then transferred to a beaker containing 1L of HEPES buffer with calcium for at least five hours at 4°C in the dark.

The next step involves concentrating the cross-linked protein in order to analyze if the crosslinker attached to the cysteine mutation. The protein was spun at 6000rpm for around two and a half hours until the final volume was between 50-100µL. The concentration of the protein was analyzed by UV-Vis spectrometer. After determining the concentration, the protein was analyzed by SDS-PAGE. In order to perform this analysis, samples were made from the wild type, wild type with mutation, R+ protein, cross-linked protein before UV radiation, and cross-linked protein after UV radiation. The final concentration of analyzed proteins was 3µg. In order to prepare the cross-linked sample after UV radiation the UV light was turned on and left to warm up for 15 minutes. Then, the protein was placed under the UV light for 30minutes. 4XBuffer and 4XBPB were added to all the samples, which were then heated at 98°C for 4 minutes before applying to a 12% SDS -PAGE.

After getting APDP or PEAS to successfully attach, the cross-linked protein was used in a crosslinking assay with rattus type I collagen or bovine type III collagen. Two vials were prepared containing 12.5µL of 2mg/mL collagen, ~0.1mg/mL cross-linked mutant, and 12.5µL 2x Binding Buffer. Before mixing, 1x Binding Buffer was added to both vials to make the total volume 50µL. The solutions were then incubated at room temperature for 30 minutes. During this time, the UV light was turned on, in order to warm up the bulb. Then, one vial was irradiated for 30 minutes.

During this time, two sample buffers were made. One sample buffer without DTT was prepared by mixing 60µL 2xBuffer, 3µL 4xBPB and 57µL deionized water. The other buffer was prepared using 60µL 2xBuffer, 3µL 4xBPB 12µL DTT and 45µL deionized water. At the end of the radiation time, both of the original vials were split into 25µL aliquots, for a total of four aliquots. The vials were labeled UV+ with DTT, UV+ with no DTT, UV- with DTT, and UV- with no DTT. For each vial, 25µL of appropriate sample buffer was added. The two vials labeled "with DTT" received the sample buffer with DTT, while the other two vials received sample buffer without DTT. After adding the sample buffers, the samples were heated at 98°C for 4 minutes. 30µL of each sample was loaded into the wells of a 7.5% SDS-PAGE by placing a marker (first lane), UV- with no DTT (second lane), UV+ with no DTT (third lane), a marker (fourth lane), UV- with DTT (fifth lane), and UV+ with DTT (sixth lane).

After confirming crosslinking with collagen, a fibril formation assay was performed. This was performed for two reasons. One reason was to ensure that the incubation time period was long enough for fibril formation of collagen. Another reason was to observe the effects that APDP attached ColGS3aS3b(S997C) had on collagen fibril formation. In order to perform this experiment a plate with small wells was used. On this plate, 8 different wells were used. A 2x Binding Buffer was prepared which consists of 40mM HEPES (pH 7.5), 0.3M NaCl and 2mM Calcium Chloride. Wells were first prepared with water and buffer in Dr. Sakon's laboratory, then after walking to Dr. Strieglers laboratory, collagen and APDP attached CBD were added. In the first well 150µL of 2x Binding Buffer was added to 150µL deionized water. The second well was a collagen control, in which 150µL 2x Binding Buffer, 75µL deionized water, and 75µL

rattus type I collagen (2mg/mL) was added to the well. The third well contained 0.1:1 ratio of APDP attached CBD to collagen. Therefore 0.7µL APDP attached CBD (2.68mg/mL), 74.3µL deionized water, 75µL rattus type I collagen (2mg/mL) and 150µL 2x Binding Buffer was added to the well. The fourth well consisted of a 0.2:1 ratio or 1.4µL APDP attached CBD (2.68mg/mL), 73.6µL deionized water, 75µL rattus type I collagen (2mg/mL) and 150µL 2x Binding Buffer. The fifth well (0.3:1 ratio) was prepared with 2.1µL APDP attached CBD (2.68mg/mL), 72.9µL deionized water, 75µL rattus type I collagen (2mg/mL) and 150µL 2x Binding Buffer. The sixth well (0.4:1 ratio) was prepared with 2.8µL APDP attached CBD (2.68mg/mL), 72.2µL deionized water, 75µL rattus type I collagen (2mg/mL) and 150µL 2x Binding Buffer. The seventh well (0.5:1 ratio) was prepared with 3.5µL APDP attached CBD (2.68mg/mL), 71.5µL deionized water, 75µL rattus type I collagen (2mg/mL) and 150µL 2x Binding Buffer. The eighth well (1:1 ratio) was prepared with 7µL APDP attached CBD (2.68mg/mL), 68µL deionized water, 75µL rattus type I collagen (2mg/mL) and 150µL 2x Binding Buffer. The machine was ran at 37°C for 3 hours. After obtaining the data, an excel spreadsheet was created. This experiment was performed two times because the first experiment produced unclean results.

After confirming that a 30 minute incubation period was sufficient time for fibril formation, the crosslinking experiment was analyzed by 2-D diagonal electrophoresis. A vial was prepared containing 6.25µL 2mg/mL collagen, ~0.1mg/mL cross-linked mutant, and 6.25µL 2x Binding Buffer. Before mixing, 1x Binding Buffer was added to make the total volume 25µL. After irradiating the sample for 30 minutes, 25µL of sample buffer without DTT was added and the sample was heated at 98°C for 4 minutes. The sample

was applied to a single well in a 5% SDS-PAGE with marker on both sides. After running the gel, a knife was used to cut out the lane containing the sample. The lane was transferred to 20mL a solution of sample buffer which was prepared by mixing 10mL of 2x Sample Buffer, 0.5mL 4xBPB, 1mL 1M DTT, and 8.5mL deionized water. 2x Sample buffer was prepared by mixing 4g SDS and 15g glycerol with 12.5mL of 1 M Tris pH 6.8. The total volume of this 2x Sample buffer solution was brought up to 50mL using deionized water. The lane was soaked in the sample buffer for 10 minutes. The gel piece was then washed with 20mL of a TES buffer for 7 minutes for a total of three washes. The gel piece was placed in 10mL TES with 10μL fluorescein, and kept in the dark for 20 minutes. From this point on all work was done in the dark, in order to ensure that photoreactive crosslinker was not activated. The gel was then washed in a solution containing 20mL TES and 200μL DTT for 7 minutes for a total of three washes. The gel was then placed horizontally in the well space of a 5% gel. The gel was under electrophoresis until a separate yellow band was visible on the SDS-PAGE. After running the gel, a picture of the SDS-PAGE was taken using UV light to show the fluorescent binding. Afterwards the gel was treated with Coomassie Blue or silver chloride stain.

If the crosslinking did occur then the samples from the crosslinking assay can be mass produced and the bands representing the crosslinking can be cut of the SDS-PAGE to be analyzed cleaved by trypsin digest and analyzed by LC-MS/MS (ESI mass spectrometry).

The next crosslinker that was used was a biotin (MTS-BP-BIO) crosslinker. This crosslinker was attached to ColGS3aS3b(S997C) using the same experimental procedure as for the APDP and PEAS crosslinker. After analysis by SDS-PAGE to see what the

crosslinking experiment produced, the protein was purified using streptavidin magnetic beads. In order to perform this experiment, 50µL of these beads (0.5mg) were placed in a 1.5mL vial. The beads were then placed in a magnetic stand that was included when purchasing the crosslinker. The beads congregated on the wall near the magnet until the solution was clear. 1mg of beads allows ~3500pmol biotynylated binding so in order to ensure binding to the streptavidin beads a sample was prepared. Calculations were performed in order to convert the concentration to picomols. It was determined that the concentration of protein needed to be less than 1750pmol because 0.5mg of beads was used.

20µL of the prepared sample was added to 280µL of washing/binding buffer. The washing/binding buffer was made by using 5µL of 100% Tween and 4.995mL of 25mM TRIS, 0.15M NaCl (pH 7.2). The supernatant from the vial with the beads was then carefully discarded using a pipette, ensuring that no beads were taken out with the supernatant. After taking out the supernatant, 1mL of this buffer was added to the beads. This must be done in a timely manner in order to ensure that the beads don't dry out. The beads were mixed by flicking the vial. The vial was then placed back into the magnetic stand to allow separation from the supernatant. The wash was removed from the magnetic beads and the 300µL sample was added to the beads. The beads and protein sample were incubated for an hour at room temperature. The beads were mixed every five minutes to ensure that the beads were in solution at all times.

Once the incubation period was up a magnetic stand was used to collect the beads. The supernatant was removed and kept in a vial for analysis. The supernatant should contain protein that was not attached to biotin. The beads were then washed three times

with 300μL washing/binding solution while discarding the supernatant after each wash. The biotin attached protein was then eluted off the beads using 100μL of 0.1M Glycine (pH 2). The beads were incubated for 5 minutes at room temperature. The solution was then taken out of the beads and placed in a vial. The beads can be reused, so 300μL washing/binding buffer was added to the beads and they were stored in the refrigerator. The solution containing biotin attached protein was concentrated and the final concentration was determined using UV–Vis spectroscopy.

After one experiment with biotin it became clear that the biotin did not attach to the cysteine mutation of collagen binding domain and several unexpected bands were seen. The sample was sent to mass spectroscopy analysis. It was suspected that somehow the protein was cleaved and in order to rule out a protease the biotin experiment was repeated using a protease inhibitor cocktail.

Another crosslinker used was fluorescein-5-maleimide dissolved in DMSO (fluorescein). The same crosslinking protocol was used except for the concentration of crosslinker was changed to a 25 molar excess of fluorescein. After calculations, it was determined that 2.4μL of fluorescein was needed for tandem CBD. After running the SDS-PAGE, the gel was observed under UV light to observe the distribution of tandem CBD. The crosslinker attached CBD was then analyzed using rattus type I collagen to see if it would crosslink to collagen. The sample was analyzed by SDS-PAGE analysis.

After analysis, the fluorescein attached tandem CBD was purified using size exclusion chromatography. In order to this 2g of G10-sephadex beads was weighed out. The beads were added to 6mL, pH 7.5 HEPES, and 150mM NaCl (Buffer 1). The beads were allowed to swell for 3 hours at room temperature. The beads were then decanted out

of the column. The bead bed was allowed to settle and the column was drained. After draining the column, 5mL of Buffer 1 was allowed to drain through the column. The sample was concentrated (>200uL) and added to the column. Buffer 1 was applied to the column in 1mL aliquots until there was a clear separation between layers. As the bottom layer started to reach the stopcock 250uL samples were collected. The vials were marked in the order that they were collected. The samples were collected until relatively clear, which results in about 13-20 total samples collected. The absorbance was checked at 280nm for the most concentrated protein samples, which were expected to be the first three or four vials collected. The most concentrated samples were pooled together and were concentrated to approximately 100μL. The beads are reusable so they were cleaned using 10mL of 0.2% NaOH. 10mL of Buffer 1 was added to the beads to ensure that the beads remained moist for the next use. After concentrating the newly purified protein a crosslinking assay was ran in order to observe the crosslinking of this protein with rattus type I collagen.

D. Remodeling Collagen Experiments

For the remodeling collagen experiments using pancreatic cancer cells, fluorescein attached ColGS3aS3b(S997C) and fluorescein attached ColGS3aS3b(Y996A, S997C) were produced following the above crosslinking method. After confirming crosslinker attachment, the proteins were taken to Dr. Yuchun Du laboratory at the University of Arkansas. Tandem CBDs were placed in a freezer until cancer cells were grown. In this laboratory, MiaPaCa2 cancer cells were grown in DMEM medium with 10%FBS and 1%Pen Strep. These cells were grown until they showed a loss of contact inhibition at 37°C and 5% carbon dioxide. At this point, 50ng/μL (final concentration) of

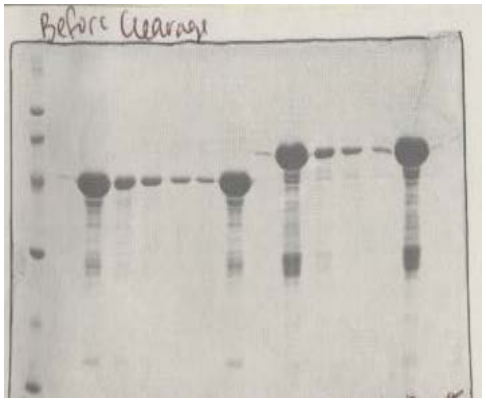
fluorescent CBD was added and incubated overnight at 37°C. The cells were then stained with DAPI stain, which stains the nucleus. Images were taken with an Axio.Vision microscope.

E. LC-MS/MS (ESI mass spectrometry)

Bruker Amazon-SL with captive-spray source coupled to a quadrupole ion trap mass spectrometer was used for MS/MS work. Agilent 1200 series HPLC with a nano pump was used for the chromatographic separation using a C18 revers phase nano column. This column has a dimensions of 0.1 x150 mm and contains particles with a size of 5 micron particle and 300 A^o pore size. A Water/ACN gradient was used to elute the compounds at a flow rate of 1.6 µl/min.

V. Results

A. Mutagenesis Projects and Protein Production



<u>Figure 12.</u> GST tagged ColGS3aS3b after GSH purification. (06/06/2013)

Confirmation of GST tagged ColGS3aS3b, after purification using the glutathione sepharose beads column. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-8, protein for fellow lab partner (i.e. not applicable for this experiment). Lane 9-14, represents the different vials that were used to collect the eluted purified protein. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3aS3b is 27.055kDa [11]. Therefore, the 55kDa band shows that the correct protein was produced.

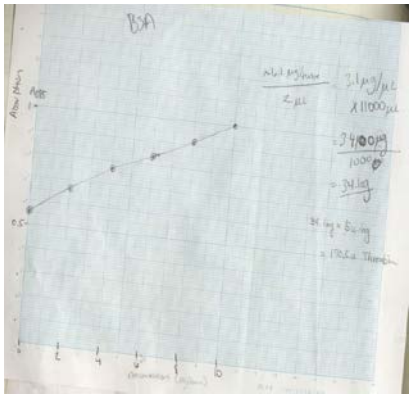


Figure 13. BSA Assay plot for ColGS3aS3b. (06/06/2013)

Plot of Concentration vs Absorption showed that there was approximately 11mL of 3.1µg/µL ColGS3aS3b obtain after the first round of purification using GSH purification column. Using this information, it was determined that 170.5 units of Thrombin needed to be added in order to cleave the GST tag off of Col GS3aS3b.

Figure 14. Cleavage of GST tagged ColGS3aS3b. (06/07/2013)

Confirmation of cleavage after adding thrombin to ColGS3aS3b. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, not applicable to this experiment. Lane 3, cleaved ColGS3aS3b with GST tag still present. Even though the molecular weight of ColGS3aS3b is only 27.055kDa, this protein shows up as a 38KDa band on SDS-PAGE [11]. The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the 38kDa band confirms that ColGS3aS3b was produced.

Figure 15. ColGS3aS3b after GSH purification. (06/10/2013)

Second round of GSH column purification did not get rid of all the impurities in the sample. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-4, shows the four vials that were collected during purification. The presence of the 38kDa band shows that ColGS3aS3b is present in the sample. The high concentration of smaller molecular weight bands indication that an ionic exchange purification column using a salt concentration gradient needed to be performed.

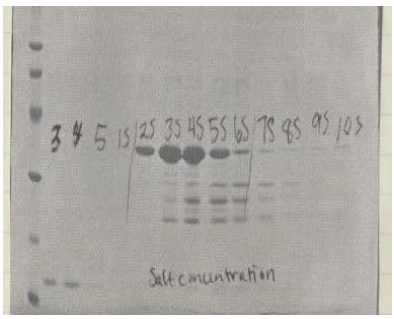
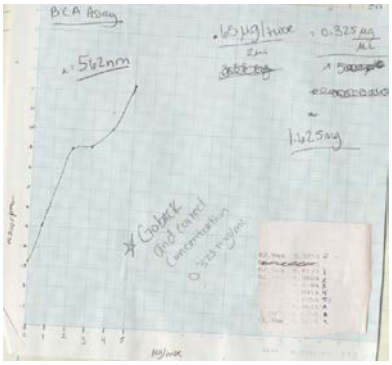


Figure 16. ColGS3aS3b after purification using ionic exchange column. (06/10/2013)

Results show that salt concentration gradient significantly reduced the amount of impurities in the samples. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-13, shows the vials that were collected during purification. The presence of the 38kDa

confirms that our desired protein was still in the collected samples and the protein had a relatively small amount of impurities. Therefore, a BCA assay was ran in order to determine the amount of protein in the collected sample.

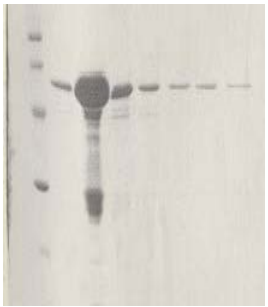


<u>Figure 17.</u> BCA Assay plot for ColGS3aS3b. (06/12/2013)

Plot of Concentration vs Absorption showed that there was approximately 1.625mg/mL of ColGS3aS3b obtained in each vial.

<u>Figure 18.</u> Forward and Reverse plasmids for ColGS3aS3b(Y996A, S997C) after first PCR reaction. (06/19/2013)

First PCR reaction appears to be successful at making the correct base pair plasmid. Lane 1, molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, DNA strand created using 5' primer, Reverse Primer, and E022 DNA plasmid. Lane 3, DNA strand created using 3' primer, Forward primer, and E022 DNA plasmid. The presence of the 754 base pair band in lane 2 confirms that the correct base pair plasmid was created. The 150 base pair band to indicate that the Forward primer was not visible in lane 3. 150 base pairs is very small number of base pairs. It was assumed that the electrophoresis machine had run to long, which allowed for the small plasmid to run off the page. There was mutual consensus to move on to the next PCR reaction.



<u>Figure 19.</u> GST tagged ColGS3aS3b(S997C) after GSH purification. (06/20/2013)

Confirmation that GST tagged ColGS3aS3b(S997C) was successfully produced after confirming the sequence of Ryan Bauer's stock E022. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-8, represents the different vials that were used to collect the eluted purified protein. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3aS3b is 27.055kDa [11]. Therefore, the 55kDa band shows that the correct protein was produced.

Figure 20. Cleavage of GST tagged ColGS3aS3b(S997C). (06/24/2013)

Confirmation of cleavage after adding thrombin to ColGS3aS3b(S997C). Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, cleaved ColGS3aS3b with GST tag still present. Even though the molecular weight of ColGS3aS3b is only 27.055kDa, this protein shows up as a 38kDa band on SDS-PAGE [11]. The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the 38kDa band confirms that ColGS3aS3b(S997C) was produced.

Figure 21. Fused DNA plasmid after second PCR reaction for ColGS3aS3b(Y996A,S997C). (06/19/2013)

Confirmation of fusion between the forward and reverse primers made after first round of PCR. Lane1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, fused DNA plasmid. The presence of a 900 base pair band shows that the PCR reaction was successful in fusing the forward and reverse primers

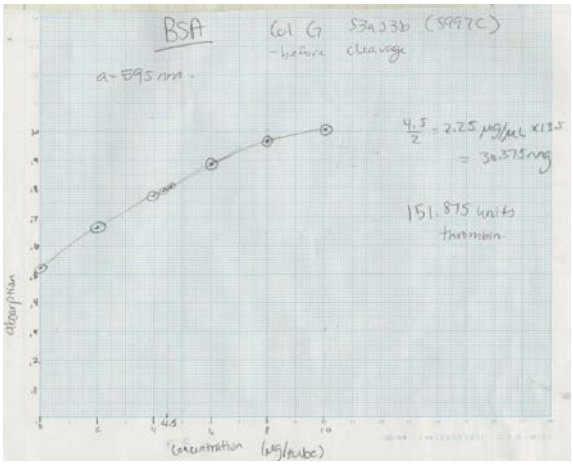
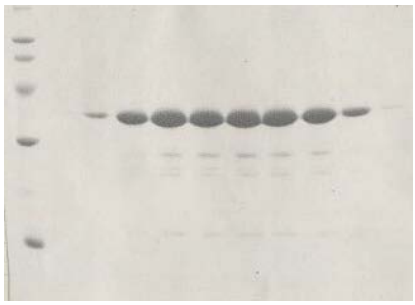


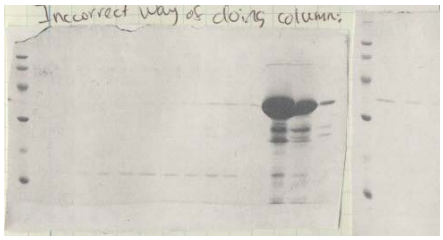
Figure 22. BSA Assay plot for ColGS3aS3b(S997C). (06/21/2013)

Plot of Concentration vs Absorption showed that there was approximately 13.5mL of 2.25µg/µL ColGS3aS3b(S997C) obtain after the first round of purification using a GSH purification column. Using this information, it was determined that 151.875 units of Thrombin needed to be added in order to cleave the GST tag off of Col GS3aS3b(S997C).



<u>Figure 23.</u> ColGS3aS3b(S997C) after GSH purification. (06/28/2013)

Second round of GSH column purification did not get rid of all the impurities in the sample. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-9, shows the vials that were collected during purification. The presence of the 38kDa band shows that ColGS3aS3b(S997C) is present in the sample. In order to get an even purer sample, it was decided that an ionic exchange purification column using a salt concentration gradient need to be performed.



<u>Figure 24.</u> ColGS3aS3b(S997C) after purification using ionic exchange column. (06/28/2013)

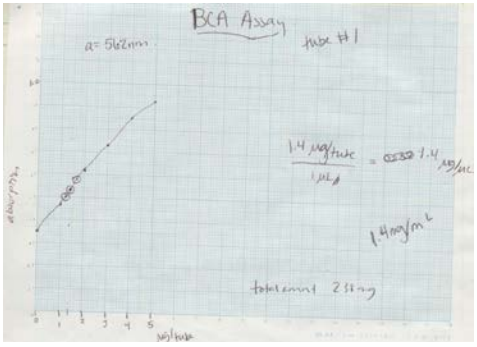
Results indicate incorrect addition of salt concentration gradient. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-14, represents the different vials that were used to collect the eluted purified protein. Lane 15, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 16-17, represents the different vials that were used to collect the eluted purified protein. In order for a salt concentration gradient to work the protein is supposed to be eluted slowly using slowly increasing concentration of salt. In this experiment the highest concentration of salt was used first which is incorrect.

<u>Figure 25.</u> ColGS3aS3b(S997C) after purification using ionic exchange column. (06/29/2013)

Results show that salt concentration gradient significantly reduced the amount of impurities in the sample. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-15, shows the vials that were collected during purification. Lane 16, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 7-23, shows a portion of the samples that were collected during purification. The presence of the 38kDa confirms that our desired protein was still in the collected samples and the protein had a relatively small amount of impurities. Therefore, a BCA assay was ran in order to determine the amount of protein in the collected sample.

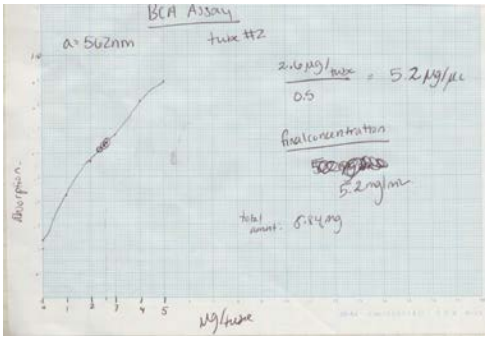
<u>Figure 26.</u> ColGS3aS3b(Y996A,S997C) DNA plasmid from DH5α colonies for sequencing. (06/29/2013)

After cleavage using XhoI and EcoRI-HF, it appears that Colony #1 contains the correct DNA plasmid. Lane 1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, Colony #1 DNA plasmid. Lane 3, Colony #2 DNA plasmid. After cleavage, Colony #1 contains two DNA fragments with a molecular weight of approximately 754 and 150 base pairs, which confirms that the DNA plasmid could contain the correct mutation. After cleavage, Colony #2 only has one band and therefore it does not have the correct DNA plasmid.



<u>Figure 27.</u> BCA Assay Plot for ColGS3aS3b(S997C). (07/01/2013)

Plot of Concentration vs Absorption showed that there was approximately 1.4mg/mL of ColGS3aS3b(S997C) obtained in each vial.



<u>Figure 28.</u> BCA Assay Plot for ColGS3aS3b(S997C). (07/01/2013)

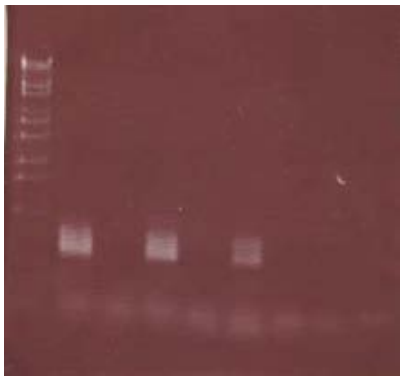
Plot of Concentration vs Absorption showed that there was approximately 5.2mg/mL of ColGS3aS3b(S997C) obtained in each vial.

<u>Figure 29.</u> Forward and Reverse plasmids for ColGS3b(Y996A, S997C) after first PCR reaction. (07/02/2013)

First PCR reaction appears to be successful at -making the correct base pair plasmid. Lane 1, molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, DNA strand created using 5' primer, Reverse Primer, and E022 DNA plasmid. Lane 3, DNA strand created using 3' primer, Forward primer, and E022 DNA plasmid. The presence of the 454 base pair band in lane 2 confirms that the correct base pair plasmid was created. The 150 base pair band that would indicate the Forward primer was barely visible at the bottom of lane 3.

<u>Figure 30.</u> Fused DNA plasmid after second PCR reaction for ColGS3aS3b(Y996A,S997C). (06/19/2013)

Confirmation of fusion between the forward and reverse primers made after first round of PCR. Lane1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, fused DNA plasmid. The presence of a 604 base pair band shows that the PCR reaction was successful in fusing the forward and reverse primers.

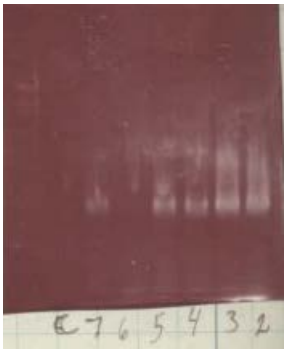


<u>Figure 31.</u> Colony PCR for ColS3b(Y996A,S997C). (07/05/2013)

Colony PCR shows that three DH5α colonies contain correct DNA plasmids. Lane1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, Colony #1. Lane 3, Colony #2. Lane 4, Colony #3. Lane 5, Colony #4. Lane 6, Colony #5. Lane 7, Colony #6. Lane 8, Colony #7. Lane 9, water control. Colony # 1, 3, and 5 show two correct fragments with molecular weight 150 and 454 base pairs. Inoculated tubes with LB broth for all three colonies.

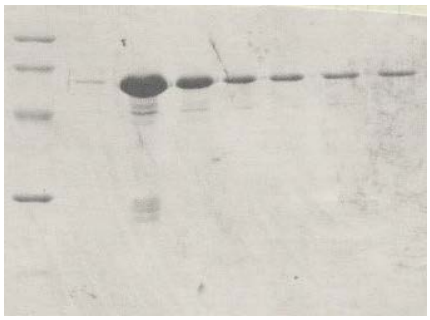
<u>Figure 32.</u> ColGS3b(Y996A,S997C) DNA plasmid from DH5α colonies for sequencing. (07/05/2013)

Incorrect cleavage plasmid #3 and #5 using XhoI and EcoRI-HF, however plasmids were still sequenced because protein had already been prepared. Lane1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, Plasmid #3 after cleavage. Lane 3, Plasmid #5 after cleavage. Even though cleavage is not obvious, after sequencing it became clear that Plasmid #3 and #5 contained the correct DNA sequences.



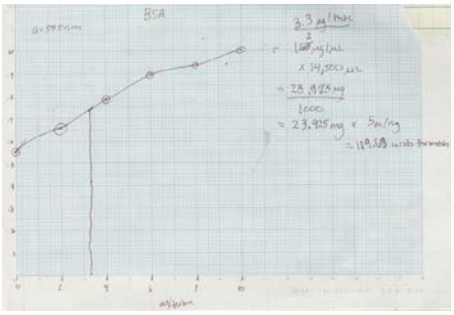
<u>Figure 33.</u> Colony PCR for ColS3aS3b(Y996A,S997C). (07/11/2013)

Colony PCR shows that three DH5α colonies contain the correct plasmid. Lane1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, water control. Lane 3, Colony #7 from master plate. Lane 4, Colony #6 from master plate. Lane 5, Colony #5 from master plate. Lane 6, Colony #4 from master plate. Lane 7, Colony #3 from master plate. Lane 8, Colony #2 from master plate. Colonies 2, 3, and 4 were used to inoculate broth in order to obtain DNA plasmid.



<u>Figure 34.</u> GST tagged ColGS3aS3b(Y996A, S997C) after GSH purification. (07/16/2013)

Confirmation that GST tagged ColGS3aS3b(Y996A, S997C) was successfully produced. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-8, represents the different vials that were used to collect the eluted purified protein. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3aS3b is 27.055 kDa [11]. Therefore, the 55kDa band shows that the correct protein was produced.



<u>Figure 35.</u> BSA Assay Plot for ColGS3aS3b(Y996A/S997C). (07/16/2013)

Plot of Concentration vs Absorption showed that there was approximately 14.5mL of 5µg/µL ColGS3aS3b(Y996A,S997C) obtain after the first round of purification using a GSH purification column. Using this information, it was determined that 179.63 units of Thrombin needed to be added in order to cleave the GST tag off of Col GS3aS3b(Y996A,S997C).

<u>Figure 36.</u> Forward and Reverse plasmids for ColGS3aS3b(Y994A,S997C) after first PCR reaction. (07/16/2013)

First PCR reaction appears to be successful at making the correct base pair plasmid. Lane 1, molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, DNA strand created using 5' primer, Reverse Primer, and E022 DNA plasmid. Lane 3, DNA strand created using 3' primer, Forward primer, and E022 DNA plasmid. The presence of the 150 base pair band that would indicate the Forward primer was barely visible at the bottom of lane 2, confirms that the correct base pair plasmid was created. The 754 base pair band confirms that the correct base pair plasmid was created in lane 3.

<u>Figure 37.</u> Fused DNA plasmid after second PCR reaction for ColGS3aS3b(Y994A,S997C). (07/17/2013)

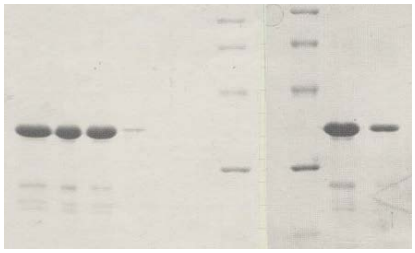
Confirmation of fusion between the forward and reverse primers made after first round of PCR. Lane 1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, fused DNA plasmid. The presence of a 904 base pair band shows that the PCR reaction was successful in fusing the forward and reverse primers.

<u>Figure 38.</u> Cleavage of GST tagged ColGS3aS3b(Y996A, S997C). (07/17/2013)

Confirmation of cleavage after adding thrombin to ColGS3aS3b(Y996A, S997C). Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, cleaved ColGS3aS3b with GST tag still present. Even though the molecular weight of ColGS3aS3b is only 27.055kDa, this protein shows up as a 38kDa band on SDS-PAGE [11]. The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the 38kDa band confirms that ColGS3aS3b(Y996A, S997C) was produced.

<u>Figure 39.</u> Colony PCR for ColS3aS3b(Y996A,S997C). (07/19/2013)

Colony PCR shows that four DH5α colonies contain the correct plasmid. Lane1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, Colony #1 from master plate. Lane 3, Colony #2 from master plate. Lane 4, Colony #3 from master plate. Lane 5, Colony #4 from master plate. Lane 6, Colony #5 from master plate. Lane 7, Colony #6 from master plate. Lane 8, water control. Colonies 2 and 3 were used to inoculate broth in order to obtain DNA plasmid.



<u>Figure 40.</u> ColGS3aS3b(Y996A, S997C) after GSH purification. (07/20/2013)

Second round of GSH column purification did not get rid of all the impurities in the sample. Lane 1-6, eluant captured while using the purification column. Lane 7 and 8, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 9-10, shows the vials that were collected during purification. The presence of the 38kDa band shows that ColGS3aS3b(Y996A,S997C) is present in the sample. In order to get an even purer sample, it was decided that an ionic exchange purification column using a salt concentration gradient need to be performed.

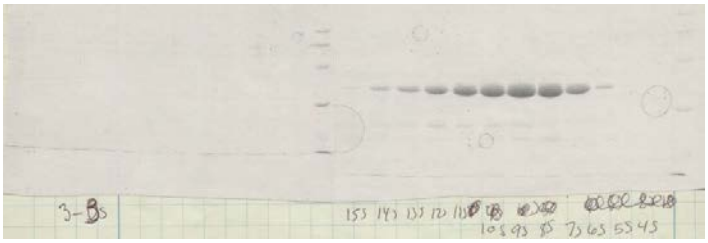


Figure 41. ColGS3aS3b(Y996A, S997C) after purification using ionic exchange column. (07/22/2013)

Results show that salt concentration gradient significantly reduced the amount of impurities in the sample. Lane 1-14, shows the vials that were collected during purification. Lane 15, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 16-29, shows the vials that were collected during purification. Lane 30, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). The presence of the 38kDa confirms that our desired protein was still in the collected samples and the protein had a relatively small amount of impurities. Therefore, a BCA assay was ran in order to determine the amount of protein in the collected sample.

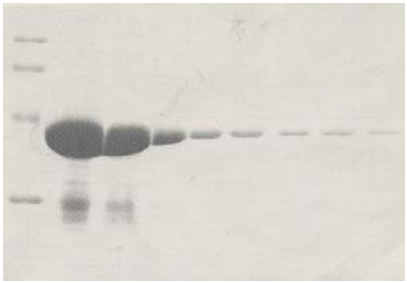


Figure 42. GST tagged ColGS3b(Y996A, S997C) after GSH purification. (07/22/2013)

Confirmation that GST tagged ColGS3b(Y996A, S997C) was successfully produced. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-8, represents the different vials that were used to collect the eluted purified protein. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3b is 13.787kDa [11]. Therefore, the 29kDa band shows that the correct protein was produced

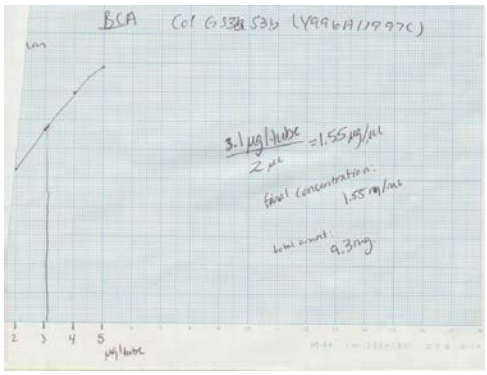
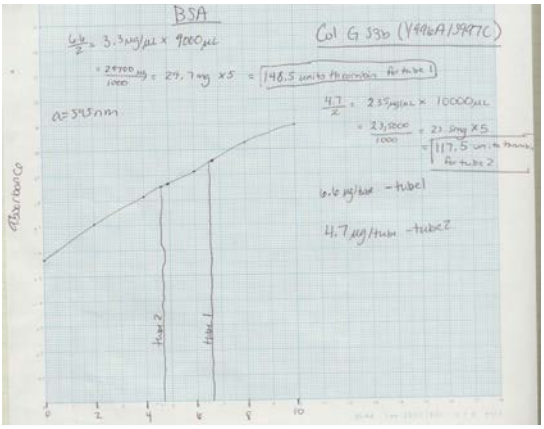


Figure 43. BCA Assay Plot for ColGS3aS3b(Y996A,S997C). (07/23/2013)

Plot of Concentration vs Absorption showed that there was approximately 1.55mg/mL of ColGS3aS3b(Y996A,S997C) obtained in each vial for a total of 9.3mg.

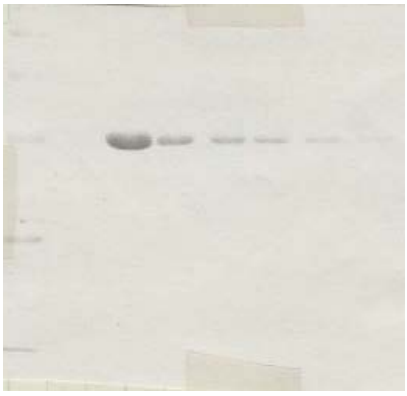
<u>Figure 44.</u> BSA Assay Plot for ColGS3b(Y996A/S997C). (07/23/2013)



Plot of Concentration vs Absorption showed that there was approximately 29.7 of 3.3µg/µL for tube 1 and 23.5mL of 2.353µg/µL of ColGS3b(Y996A,S997C) obtained after the first round of purification using a GSH purification column. Using this information, it was determined that 148.5 units of Thrombin and 117.5 units of thrombin were needed for tube 1 and tube 2 respectively in order to cleave the GST tag off of Col GS3b(Y996A,S997C)

<u>Figure 45.</u> Cleavage of GST tagged ColGS3b(Y996A, S997C). (07/24/2013)

Confirmation of cleavage after adding thrombin to ColGS3b(Y996A, S997C). Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, cleaved ColGS3aS3b with GST tag still present. Even though the molecular weight of ColGS3b is only 13.787kDa, this protein shows up as a 16.5kDa band on SDS-PAGE [11]. The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the band at the bottom of the page confirms that ColGS3b(Y996A, S997C) was produced.



<u>Figure 46.</u> GST tagged ColGS3a(S879C) after GSH purification. (07/25/2013)

Confirmation that GST tagged ColGS3a(S879C) was successfully produced. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-8, represents the samples that were collected in the different vials. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3a is 14. 223kDa [11]. Therefore, the 43kDa band shows that the correct protein was produced.

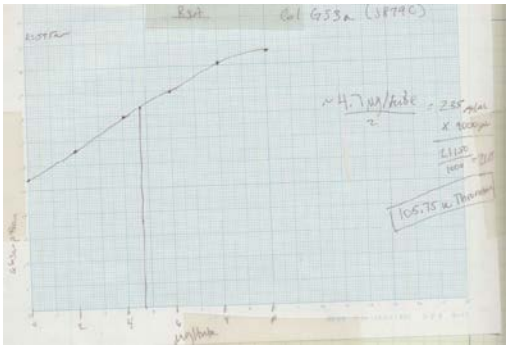


Figure 47. BSA Assay Plot for ColGS3a(S879C). (07/26/2013)

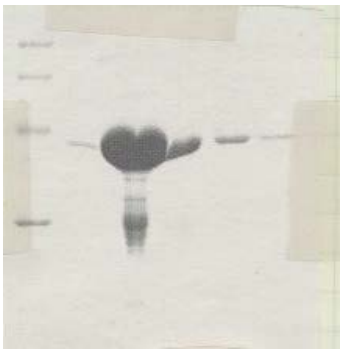
Plot of Concentration vs Absorption showed that there was approximately 21.15mL of 4.7μg/μL of ColGS3a(S879C) obtained after the first round of purification using a GSH purification column. Using this information, it was determined that 105.75 units of Thrombin was needed in order to cleave the GST tag off of ColGS3a(S879C).

Figure 48. ColGS3b(Y996A,S997C) DNA plasmid from DH5α colonies for sequencing. (07/26/2013)

After cleavage using XhoI and EcoRI-HF, it appears that Colony #1 and Colony #2 contains the correct DNA plasmid. Lane 1, DNA molecular weight marker (19329, 7743, 6223, 4254, 3472, 2690, 1882, 1489, 925, 421, 74 base pairs). Lane 2, Colony #1 DNA plasmid. Lane 3, Colony #2 DNA plasmid. After cleavage, Colony #1 and #2 contains two DNA fragments with a molecular weight of approximately 754 and 150 base pairs, which confirms that the DNA plasmid could contain the correct mutation. After sequencing it was determined that they do contain the correct sequences.

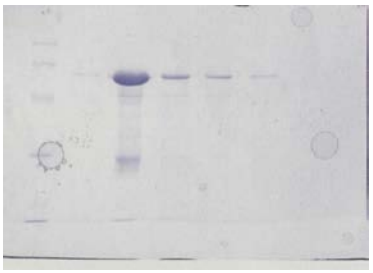
Figure 49. Cleavage of GST tagged ColGS3a(S879C). (07/30/2013)

Confirmation of cleavage after adding thrombin to ColGS3a(S879C). Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, cleaved ColGS3a(S879C) with GST tag still present. Even though the molecular weight of ColGS3a is only 14.223kDa, this protein shows up as a 23kDa band on SDS-PAGE [11]. The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the 23kDa band at the bottom of the page confirms that ColGS3a(S879C) was produced.



<u>Figure 50.</u> GST tagged ColGS3b(Y996A, S997C) after GSH purification. (07/31/2013)

After losing the first batch of ColGS3b(Y996A,S997C) due to a hole in the dialysis tube, reproduced protein confirms that GST tagged ColGS3b(Y996A, S997C) was successfully produced again. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-7, represents the different vials that were used to collect the eluted purified protein. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3b is 13.787kDa [11]. Therefore, the 29kDa band shows that the correct protein was produced.



<u>Figure 51.</u> GST tagged ColGS3aS3b(S879C) after GSH purification. (08/01/2013)

Confirmation that GST tagged ColGS3aS3b(S879C) was successfully produced. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-7, represents the different vials that were used to collect the eluted purified protein. The molecular weight of GST is 26.9kDa and the molecular weight of ColGS3aS3b is 27.055kDa [11]. Therefore, the 55kDa band shows that the correct protein was produced.

<u>Figure 52.</u> ColGS3a(S879C) after GSH purification. (06/28/2013)

Second round of GSH column purification did not get rid of all the impurities in the sample. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-13, shows the vials that were collected during purification. The presence of the 23kDa band shows that ColGS3a(S879C) is present in the sample.

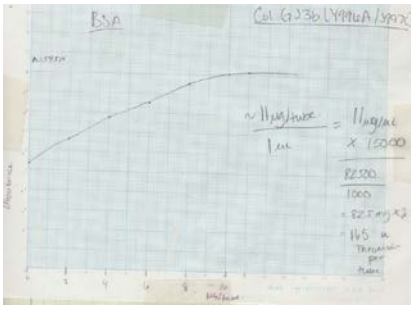


Figure 53. BSA Assay Plot for ColGS3b(Y996A,S997C). (08/01/2013)

Plot of Concentration vs Absorption showed that there was approximately 75mL of 11µg/µL of ColGS3b(Y996A,S997C).obtained after the first round of purification using a GSH purification column. Using this information, it was determined that 165 units of Thrombin was needed in order to cleave the GST tag off of ColGS3b(Y996A,S997C).

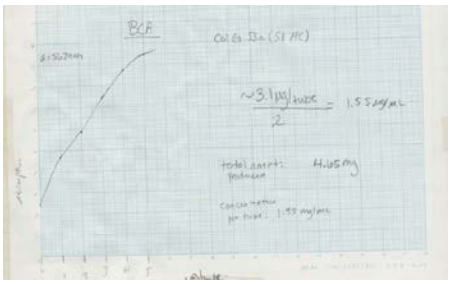
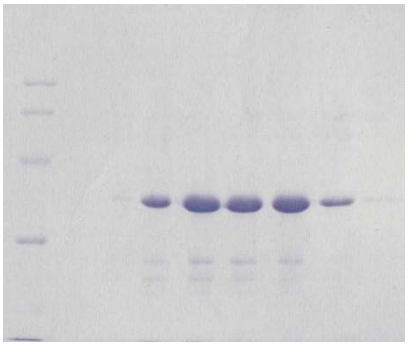


Figure 54. BCA Assay Plot for ColGS3a(S879C). (08/02/2013)

Plot of Concentration vs Absorption showed that there was approximately 1.55mg/mL of ColGS3a(S879C) obtained in each vial for a total of 4.65mg.

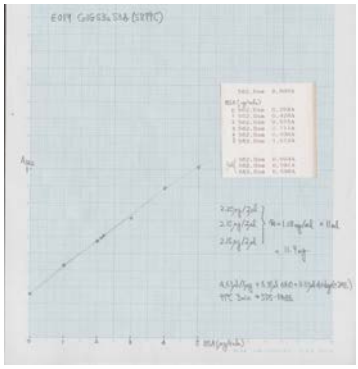
Figure 55. Cleavage of GST tagged ColGS3aS3b(S879C). (08/02/2013)

Confirmation of cleavage after adding thrombin to ColGS3aS3b(S879C). Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, cleaved ColGS3aS3b with GST tag still present. Even though the molecular weight of ColGS3aS3b is only 27.055 kDa, this protein shows up as a 38kDa band on SDS-PAGE [11]. The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the 38kDa band confirms that ColGS3aS3b(S879C) was produced.



<u>Figure 56.</u> ColGS3aS3b(S879C) after GSH purification. (08/02/2013)

Second round of GSH column purification did not get rid of all the impurities in the sample. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa. Lane 2- 8, sample of collected purified protein. The presence of the 38kDa band shows that ColGS3aS3b(S879C) is present in the sample. Vials 2, 3, 7 and 8 were pooled and concentrated down to 0.5mL using an Amicon Ultra 4 (10kDa c.o) device. These were then pooled with 4-6 for a total volume of 11.0mL with a final concentration of 1.08mg/mL.



<u>Figure 57.</u> BCA Assay Plot for ColGS3aS3b(S879C). (08/02/2013)

Plot of Concentration vs Absorption showed that there was approximately 1.08 mg/mL of ColGS3a(S879C) obtained in each vial for a total of 11.9mg.

<u>Figure 58.</u> Cleavage of GST tagged ColGS3b(Y996A, S997C). (08/02/2013)

Confirmation of cleavage after adding thrombin to ColGS3b(Y996A, S997C). Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2, cleaved ColGS3aS3b with GST tag still present. Even though the molecular weight of ColGS3b is only 13.787kDa, this protein shows up as a 16.5kDa band on SDS-PAGE [11] . The molecular weight of GST is 26.9kDa but is represented by a 28kDa band on SDS-PAGE. Therefore the presence of the 28kDa band shows that the GST tag was cleaved off the protein. The presence of the band at the bottom of the page confirms that ColGS3b(Y996A, S997C) was produced.

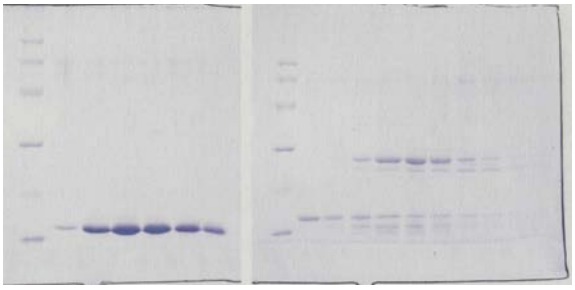


Figure 59. ColGS3b(Y996A, S997C) after purification using ionic exchange column. (08/02/2013) Results show that salt concentration gradient significantly reduced the amount of impurities in the sample. Lane 1, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 2-7, sample of collected protein purified without salt gradient. Lane 8, molecular weight marker (94, 67, 43, 30, 20.1, 14.4 kDa). Lane 9-18, samples of protein collected with salt gradient. The presence of the 16.5kDa confirms that our desired protein was still in the collected samples and the protein had a relatively small amount of impurities. Therefore, a BCA assay was ran in order to determine the amount of protein in the collected sample.

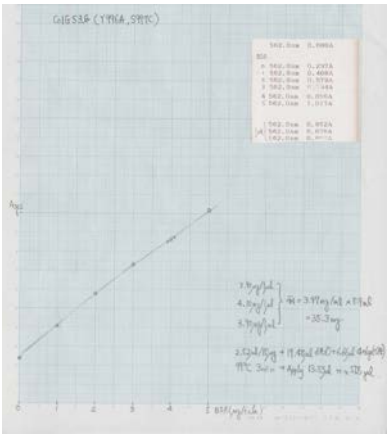


Figure 60. BCA Assay Plot for ColGS3b(Y996A,S997C). (08/02/2013)

Plot of Concentration vs Absorption showed that there was approximately 3.97mg/mL of ColGS3b(Y996A,S997C) obtained in each vial for a total of 35.3mg.

B. Crosslinking and Remodelling Collagen Experiments

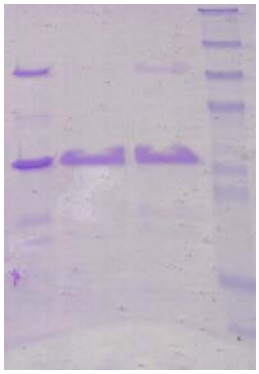


Figure 61. ColGS3aS3b(S997C) +APDP (10/2/13)

Confirmation of attachment of APDP to ColGS3aS3b(S997C). Lane 1, Col GS3aS3b(S997C). Lane 2, ColGS3aS3b(S997C) with APDP crosslinker after no UV radiation. Lane 3, ColGS3aS3b(S997C) reacted with APDP crosslinker after UV radiation. Lane 4, molecular weight marker (175, 80, 58, 46, 30, 25,17 kDa). The apparent molecular weight of ColGS3aS3b(S997C) is 32kDa and 64kDa, because it forms intermolecular disulfide bridge (lane 1). Reaction with APDP in the presence of DTT prevented the intermolecular disulfide bridge formation and allowed for crosslinker

attachment. After crosslinker attachment the disulfide bridge cannot be formed until exposed to UV (lane 2). UV exposure using light source for 30 min triggered intermolecular crosslinking occurred and is shown by the 64kDa band (lane 3).

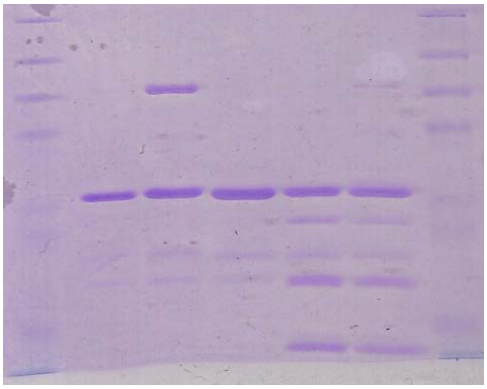


Figure 62. ColGS3aS3b(S997C) + Biotin (10/23/13)

Confirmation of attachment of biotin to ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25,17 kDa). Lane 2, ColGS3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, reduced ColGS3aS3b(S997C). Lane 5, ColGS3aS3b(S997C) with biotin after no UV radiation. Lane 6, ColGS3aS3b(S997C) with biotin after UV radiation. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32kDa and 64kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the intermolecular disulfide bridge formation (lane 4). After crosslinker attachment the disulfide bridge cannot be formed until exposed to UV (lane 5). UV exposure using light source for 30 min triggered intermolecular crosslinking occurred and is shown by the 64kDa band (lane 6).

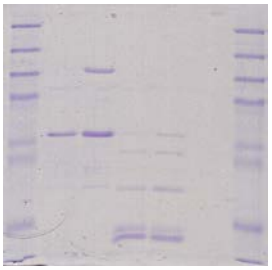
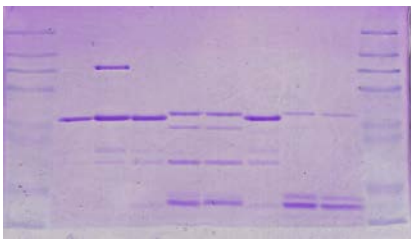


Figure 63. ColGS3aS3b(S997C) + Biotin after Avidin purification (10/26/13)

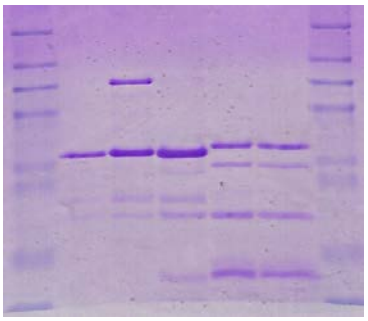
Results showing failed attempt at purification of biotin crosslinked ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, Col G S3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, ColGS3aS3b(S997C) + biotin before streptavidin beads without UV light. Lane 5, supernatant of all protein not attached to Biotin. Lane 6, ColGS3aS3b(S997C) with biotin after elution from streptavidin beads. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32 and 64kDa, because it forms intermolecular disulfide bridge (lane 3). Still apparent in the unpurified sample are unexplained lower molecular weight bands (lane 4). After addition to streptavidin beads it

becomes clear that the lower molecular weight bands were not attached to biotin (lane 5). Elution of biotin from the streptavidin beads shows that the crosslinking was not successful because of the lack of any bands (lane 6).



<u>Figure 64.</u> ColGS3aS3b(S997C) + Biotin (4hour and 5 hour incubation, 11/6/13)

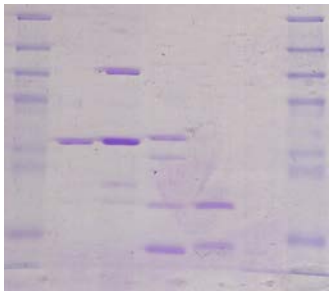
Results from increasing incubation time, in an attempt to increase amount of attached biotin to ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, reduced ColGS3aS3b(S997C). Lane 5, ColGS3aS3b(S997C) with biotin after a four hour incubation time and no UV radiation. Lane 6, ColGS3aS3b(S997C) with biotin after a four hour incubation time and UV radiation. Lane 7, reduced ColGS3aS3b(S997C). Lane 8, ColGS3aS3b(S997C) with biotin after a five hour incubation time and no UV radiation. Lane 9, ColGS3aS3b(S997C) with biotin after a five hour incubation time and UV radiation. Lane 10, molecular weight marker (175, 80, 58, 46, 30, 25,17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32kDa and 64kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the intermolecular disulfide bridge formation (lane 4 and 7). After UV exposure for 30 minutes, it becomes apparent that biotin did not attach for the four and five hour incubation time periods because of the absence of 64kDa bands (lane 6 and 9). The slightly higher masses of 33kDa and the lower molecular weight bands cause concern about the mechanism of crosslinking (lane 5, 6, 8, and 9).



<u>Figure 65.</u> ColGS3aS3b(S997C) + Biotin (double concentration, 11/6/13)

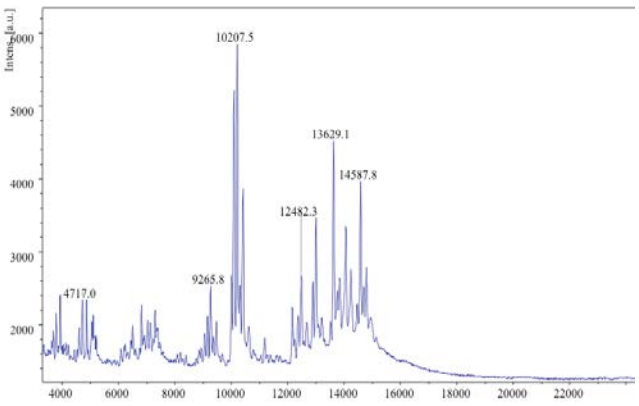
Results from increasing the concentration of mutant CBD to increase amount of contact with biotin to ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, reduced ColGS3aS3b(S997C). Lane 5, ColGS3aS3b(S997C) with biotin after no UV radiation. Lane 6, ColGS3aS3b(S997C) with biotin after UV radiation. Lane 7, molecular weight marker (175, 80, 58, 46, 30,

25, 17 kDa). The apparent molecular weight of wild type ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32 and 64 kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the intermolecular disulfide bridge formation (lane 4). After UV exposure, it becomes apparent that biotin did not attach by observing the lack of 64kDa molecular weight bands (lane 6). The slightly higher 33kDa band and the lower molecular weight bands cause concern about the mechanism of crosslinking (lane 5 and 6). To get a better understanding of what the bands represent, gel pieces from lane 6 were analyzed by mass spectrometry (see figure 7).



<u>Figure 66.</u> ColGS3aS3b(S997C) + Biotin (double concentration) after Avidin Bead purification (11/27/13)

Results showing failed attempt at purification of Biotin attached Col G S3aS3b (S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, Col G S3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, ColGS3aS3b(S997C) + biotin before avidin beads without UV light. Lane 5, supernatant of all protein not attached to biotin. Lane 6, ColGS3aS3b(S997C) with biotin after elution from streptavidin beads. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25,17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32 and 64 kDa, because it forms intermolecular disulfide bridge (lane 3). Still apparent in the unpurified sample are unexplained lower molecular weight bands (lane 4). After addition to streptavidin beads it becomes clear that the lower molecular weight bands were not attached to biotin (lane 5). Elution of biotin from the streptavidin beads, shows that the crosslinking was not successful because of the lack of any bands (lane 6).

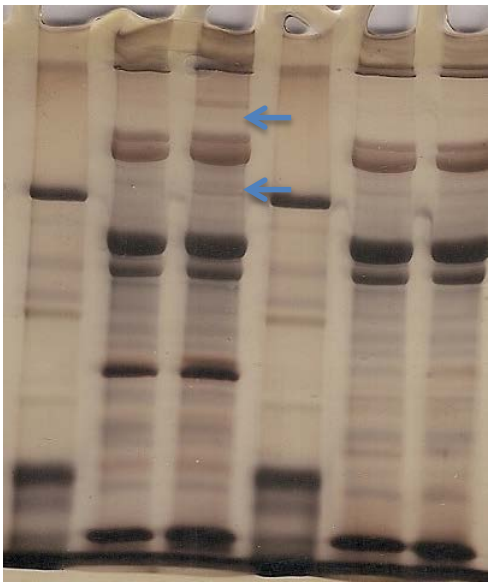


<u>Figure 67.</u> Mass Spectroscopy Analysis of Biotin from 11/6/13 (11/27/13)

The lack of a peak at 32kDa shows that somehow upon binding of the biotin-crosslinker, the collagen binding domains fragmented to pieces. This cleavage explains the unexpected lower molecular weight bands.

<u>Figure 68.</u> ColGS3aS3b(S997C) +Fluorescein (12/3/13)

The distribution of ColGS3aS3b(S997C) is clearly seen by using fluorescein crosslinker. Lane 1, Col GS3aS3b(S997C) with fluorescein under UV light. Lane 2, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of ColGS3aS3b(S997C) with fluorescein is 32kDa. The presence of the lower molecular weight bands shows that this protein had undergone cleavage by a protease, and therefore new protein was ordered and prepared by Dr. Osamu Matsushita. While protein was being produced, experimentation with old protein was done in order to gain experience with procedures.



<u>Figure 69.</u> Col GS3aS3b(S997C) +APDP + Rattus Type I Collagen (Old protein, 1/21/14)

Photocrosslinked ColGS3aS3b(S997C) with APDP plus rattus type I collagen analyzed by 7.5% gel. Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b(S997C) +APDP + rattus type I collagen with no UV exposure. Lane 3, Col G S3aS3b (S997C) +APDP + rattus type I collagen after UV exposure. Lane 4, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 5, Col G S3aS3b (S997C) +APDP + rattus type I collagen with DTT and no UV exposure. Lane 6, Col G S3aS3b (S997C) +APDP + rattus type I collagen with DTT and UV exposure. The β chain of rattus type I collagen is represented by the thicker double bands above the 175kDa marker band. The α chain of rattus type I collagen is represented by the thicker double bands with the lower band having an approximate weight of 80kDa. The band marked by the top arrow represents the crosslinking of tandem with the β chain, and the band marked by the bottom arrow represents the crosslinking of tandem with the α chains (lane 3). The tandem dimer is represented by the dark bands at the 42kDa (lane 2 and 3). Confirmation that the α and β crosslinking did occur is shown after introducing DTT, which causes the breakdown of the disulfide bridge between the

crosslinker and tandem. The crosslinked product at both the α and β chains and the dimer product disappeared after reduction by DTT (lane 5 and 6).

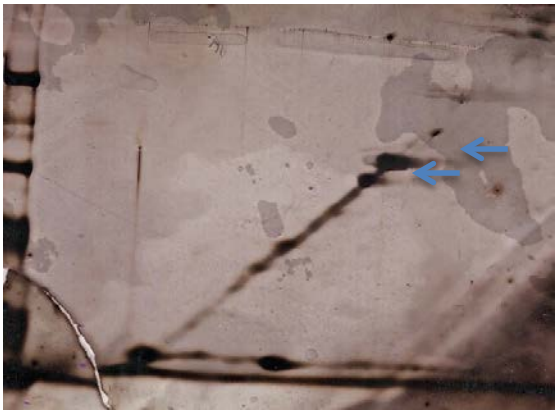


Figure 70. ColGS3aS3b(S997C) +APDP + Rattus Type I Collagen (Old protein, 1/24/14)

2-D diagonal electrophoresis of photocrosslinked ColGS3aS3b(S997C) with APDP plus rattus type I collagen. The arrows indicate where potential crosslinking occurred on the α chains, but crosslinking could not be seen on the β chains.

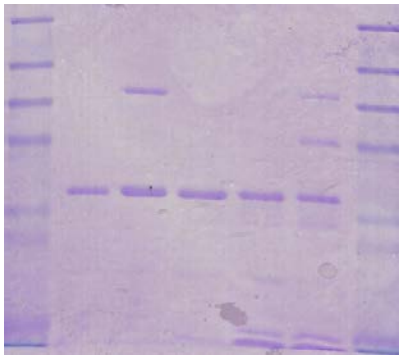
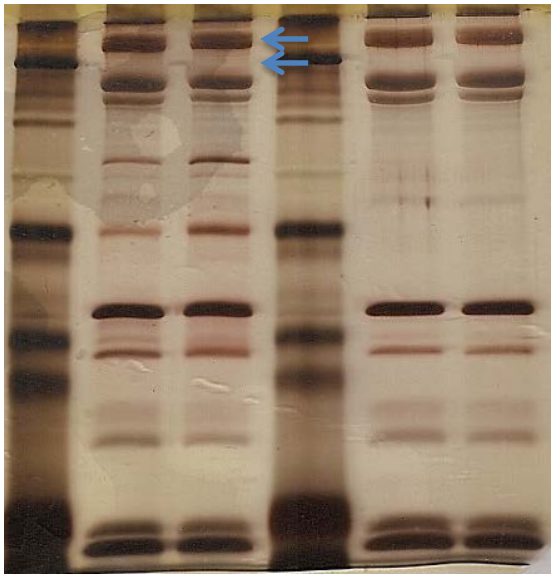


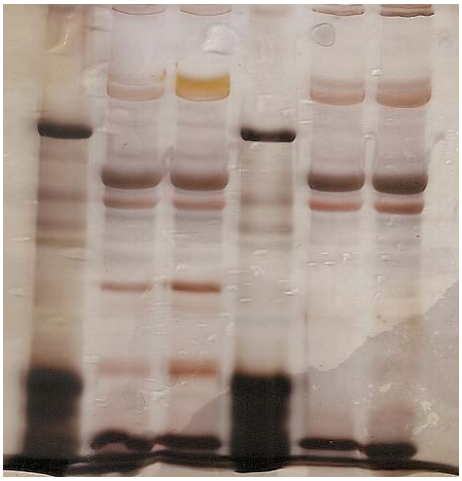
Figure 71. ColGS3aS3b(S997C) +APDP (New Protein, 1/30/14)

Confirmation of attachment of APDP to ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, reduced ColGS3aS3b(S997C). Lane 5, ColGS3aS3b(S997C) with APDP after no UV radiation. Lane 6, ColGS3aS3b(S997C) with APDP after UV radiation. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32KDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32kDa and 64kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the intermolecular disulfide bridge formation (lane 4). After crosslinker attachment the disulfide bridge cannot be formed until exposed to UV (lane 5). UV exposure using light source for 30 min triggered intermolecular cross-linking occurred and is shown by the 64kDa band (lane 6)



<u>Figure 72.</u> ColGS3aS3b(S997C) +APDP + Rattus Type I Collagen (New protein, 2/12/14)

Photocrosslinked ColGS3aS3b(S997C) with APDP plus rattus type I collagen analyzed by 12% gel. Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 KDa). Lane 2, ColGS3aS3b(S997C) +APDP + rattus type I collagen with no UV exposure. Lane 3, ColGS3aS3b(S997C) +APDP + rattus type I collagen with UV exposure. Lane 4, molecular weight marker (175, 80, 58, 46, 30, 25, 17 KDa). Lane 5, ColGS3aS3b(S997C) +APDP + rattus type I collagen with DTT and no UV exposure. Lane 6, ColGS3aS3b(S997C) +APDP + rattus type I collagen with DTT and UV exposure. The β chain of rattus type I collagen is represented by the thicker double bands above the 175KDa marker band. The α chain of rattus type I collagen is represented by the thicker double bands with the lower band having an approximate weight of 80KDa. The band marked by the top arrow represents the crosslinking of tandem with the β chain, and the band marked by the bottom arrow represents the crosslinking of tandem with the α chains (lane 3). The tandem dimer is represented by the dark bands at the 42kDa (lane 2 and 3). Confirmation that α and β crosslinking did occur is shown after introducing DTT, which causes the breakdown of the disulfide bridge between the crosslinker and tandem. The crosslinked product at both the α and β chains and the dimer product disappeared after reduction by DTT (lane 5 and 6).



<u>Figure 73.</u> ColGS3aS3b(S997C) +APDP + Rattus Type I Collagen (New protein, 2/12/14)

Photocrosslinked ColGS3aS3b(S997C) with APDP plus rat type I collagen analyzed by 7.5% gel. Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b(S997C) +APDP + rattus type I Collagen with no UV exposure. Lane 3, ColGS3aS3b(S997C) + APDP + rattus type I collagen with UV exposure. Lane 4, molecular weight

marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 5, Col G S3aS3b (S997C) +APDP + rattus type I collagen with DTT and no UV exposure. Lane 6, Col G S3aS3b (S997C) +APDP + rattus type I collagen with DTT and UV exposure. The β chain of rattus type I collagen is represented by the thicker double bands above the 175 kDa marker band. The α chain of rattus type I collagen is represented by the thicker double bands with the lower band having an approximate weight of 80kDa. Crosslinking did not occur because no additional bands were created after UV exposure for 30 minutes (lane 3).

Figure 74. ColGS3aS3b(S997C) +APDP + Rattus Type I Collagen (New protein, 2/12/14)

2-D diagonal electrophoresis of photocrosslinked ColGS3aS3b(S997C) with APDP plus rat type I collagen. According to this analysis no crosslinking occurred. Gel concentration could have not been high enough to allow for the crosslinker to spread out from the protein in order to observe the crosslinking.

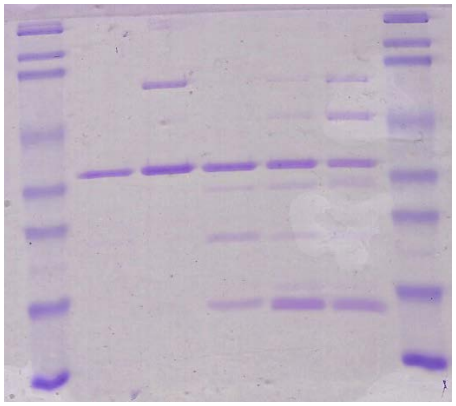
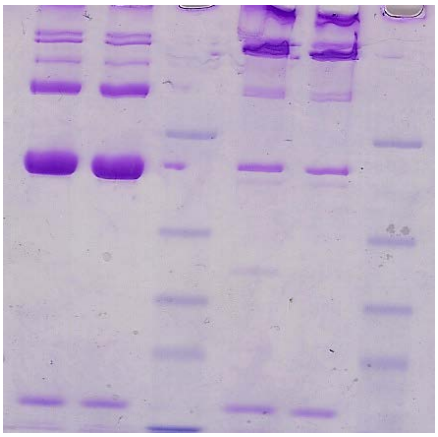


Figure 75. Col GS3aS3b(S997C) + PEAS (New Protein, 1/30/14)

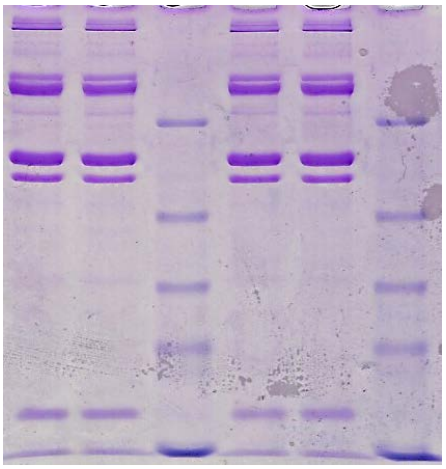
Confirmation of attachment of PEAS to ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, reduced ColGS3aS3b(S997C). Lane 5, ColGS3aS3b(S997C) with PEAS after no UV radiation. Lane 6, ColGS3aS3b(S997C) with PEAS after UV radiation. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25,17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32 and 64 kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the

intermolecular disulfide bridge formation (lane 4). After crosslinker attachment the disulfide bridge cannot be formed until exposed to UV (lane 5). UV exposure using light source for 30 min triggered intermolecular cross-linking occurred and is shown by the 64 kDa band (lane 6).



<u>Figure 76.</u> ColGS3aS3b(S997C) +PEAS + Bovine Type III Collagen (New protein, 4/8/14)

Photocrosslinked ColGS3aS3b(S997C) with PEAS and bovine type III collagen analyzed by 7.5% gel. Lane 1, ColGS3aS3b (S997C) +PEAS + bovine type III collagen with no UV exposure. Lane 2, ColGS3aS3b(S997C) + PEAS + bovine type III collagen with UV exposure. Lane 3, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 4, ColGS3aS3b(S997C) + PEAS + bovine type III collagen with DTT and no UV exposure. Lane 5, ColGS3aS3b(S997C) + PEAS + bovine type III collagen with DTT and UV exposure. Lane 6, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Crosslinking did not occur because no additional bands were created after UV exposure for 30 minutes (lane 2).



<u>Figure 77.</u> ColGS3aS3b(S997C) +PEAS + Rattus Type I Collagen (New protein, 4/8/14)

Photocrosslinked ColGS3aS3b(S997C) with PEAS and rattus type I collagen analyzed by 7.5% gel. Lane 1, ColGS3aS3b (S997C) +PEAS + rattus type I collagen with no UV exposure. Lane 2, ColGS3aS3b(S997C) +PEAS + rattus type I collagen with UV exposure. Lane 3, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 4, ColGS3aS3b(S997C) +PEAS + rattus type I collagen with DTT and no UV exposure. Lane 5, ColGS3aS3b(S997C) +PEAS + rattus type I collagen with DTT and UV exposure. Lane 6, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The β chain of rattus type I collagen is represented by the thicker double bands above the 175kDa marker band. The α chain of rattus type I collagen is represented by the thicker double bands with the lower band having an approximate weight of 80kDa. The arrow indicates crosslinking to α chain (lane 2). The tandem

dimer is represented by the dark bands at the 42kDa (lane 1 and 2). Confirmation that the α crosslinking did occur is shown after introducing DTT, which causes the breakdown of the disulfide bridge between the crosslinker and tandem. The cross-linked product on the α chain and the dimer product disappeared after reduction by DTT (lane 4 and 5).

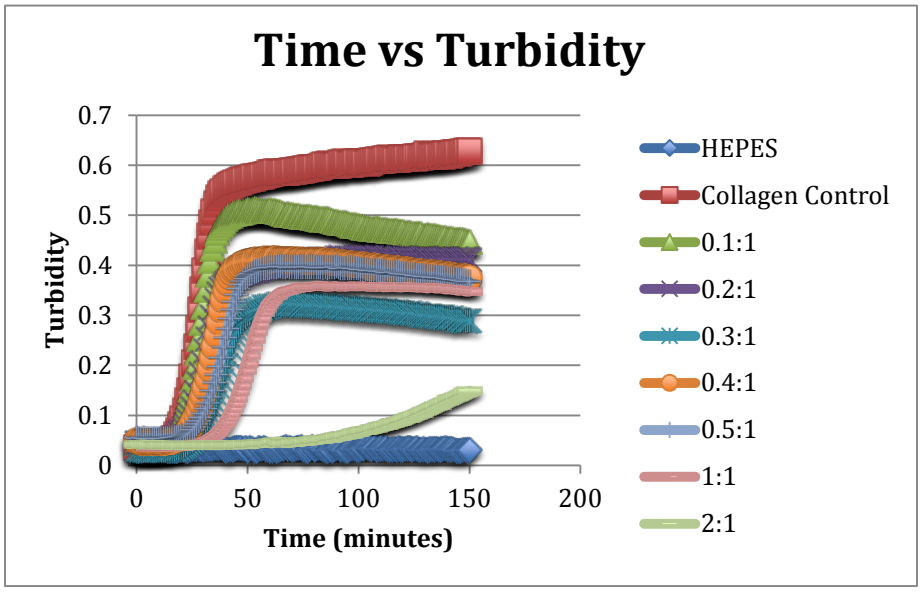


Figure 78. Col GS3aS3b(S997C) +APDP + rattus type I collagen (New Protein, 5/1/14)

Fibril formation assay performed in order to determine if incubation time is long enough for collagen fibril formation. According to the data it can be seen that 30 minutes is sufficiently long enough for the collagen fibril to form in the presence of ColGS3aS3b(S997C) +APDP. Interfering with fibril formation

Figure 79. ColGS3aS3b (S997C) +Fluorescein (9/12/14)

The distribution of ColGS3aS3b(S997C) is clearly seen using the fluorescein crosslinker. Lane 1, ColGS3aS3b(S997C) with fluorescein under UV light. Lane 2, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of ColGS3aS3b(S997C) with fluorescein is 32kDa. This photo does not clearly show the band that was present at 32kDa. However, it can be seen that all the lower molecular weight bands are missing from the protein so this new protein has not been cleaved by a protease.

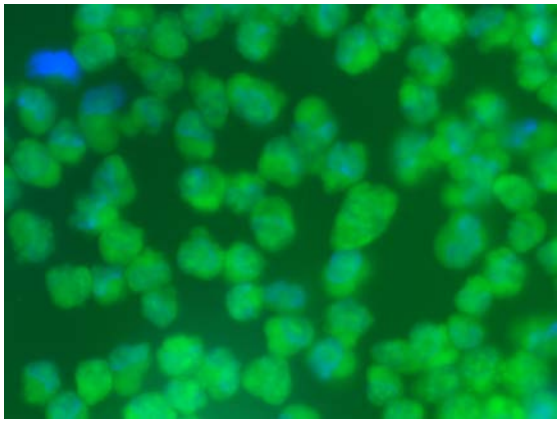


Figure 80. ColGS3aS3b(S997C) + Fluorescein +Pancreatic Cancer Cells (9/13/14)

After addition of fluorescein attached ColGS3aS3b(S997C) it is becomes clear that ColGS3aS3b(S997C) binds to the extracellular matrix of the cancer cells.

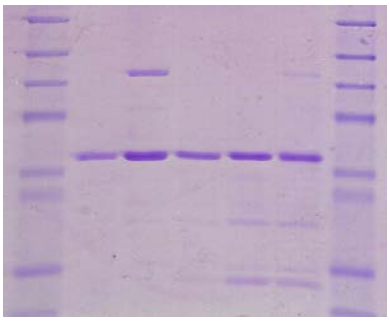


Figure 81. ColGS3aS3b(S997C) +APDP (New Protein, 9/18/14)

Confirmation of attachment of APDP to ColGS3aS3b(S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, ColGS3aS3b. Lane 3, ColGS3aS3b(S997C). Lane 4, reduced ColGS3aS3b(S997C). Lane 5, ColGS3aS3b(S997C) with APDP after no UV radiation. Lane 6, ColGS3aS3b(S997C) with APDP after UV radiation. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of wild type ColGS3aS3b is 32KDa (lane 2). The apparent molecular weight of ColGS3aS3b(S997C) is 32 and 64 kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the intermolecular disulfide bridge formation (lane 4). After crosslinker attachment the disulfide bridge cannot be formed until exposed to UV (lane 5). UV exposure using light source for 30 min triggered intermolecular cross-linking occurred and is shown by the 64kDa band (lane 6).

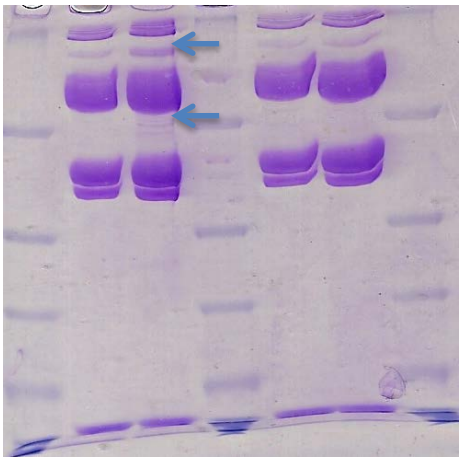
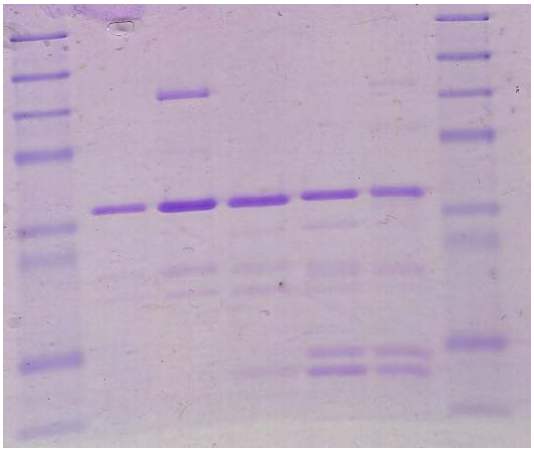


Figure 82. Col G S3aS3b (S997C) +APDP + Rattus Type I Collagen (New protein, 9/18/14)

Photocrosslinked ColGS3aS3b(S997C) with APDP plus rattus type I collagen analyzed by 12% gel. Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25,17 kDa). Lane 2, ColGS3aS3b(S997C) +APDP + rattus type I collagen with no UV exposure. Lane 3, ColGS3aS3b(S997C) +APDP + rattus type I collagen with UV exposure. Lane 4, molecular weight

marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 5, ColGS3aS3b(S997C) +APDP + rattus type I collagen with DTT and no UV exposure. Lane 6, ColGS3aS3b(S997C) +APDP + rattus type I collagen with DTT and UV exposure. The β chain of rattus type I collagen is represented by the thicker double bands above the 175 kDa marker band. The α chain of rattus type I collagen is represented by the thicker double bands with the lower band having an approximate weight of 80kDa. The band marked by the top arrow represents the crosslinking of tandem with the β chain, and the band marked by the bottom arrow represents the crosslinking of tandem with the α chains (lane 3). Confirmation that α and β crosslinking did occur is shown after introducing DTT, which causes the breakdown of the disulfide bridge between the crosslinker and tandem. The crosslinked product at both the α and β chains disappeared after reduction by DTT (lane 5 and 6).



<u>Figure 83.</u> ColGS3aS3b(Y996A,S997C) + APDP (New Protein, 9/18/14)

Confirmation of attachment of APDP to ColGS3aS3b(Y996A,S997C). Lane 1, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). Lane 2, Col G S3aS3b. Lane 3, ColGS3aS3b(Y996A,S997C). Lane 4, reduced ColGS3aS3b(Y996A,S997C). Lane 5, ColGS3aS3b(Y996A, S997C) with APDP after no UV radiation. Lane 6, ColGS3aS3b(Y996A,S997C) with APDP after UV radiation. Lane 7, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of wildtype ColGS3aS3b is 32kDa (lane 2). The apparent molecular weight of ColGS3aS3b(Y996A,S997C) is 32 and 64 kDa, because it forms intermolecular disulfide bridge (lane 3). Introducing DTT to ColGS3aS3b prevented the intermolecular disulfide bridge formation (lane 4). After crosslinker attachment the disulfide bridge cannot be formed until exposed to UV (lane 5). UV exposure using light source for 30 min triggered intermolecular cross-linking occurred and is shown by the 64kDa band (lane 6).

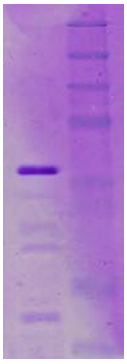


Figure 84. ColGS3aS3b(S997C) +Fluorescien (9/18/14)

Studying the distribution of ColGS3aS3b(S997C) using fluorescein crosslinker. Lane 1, ColGS3aS3b(S997C) with fluorescein under UV light. Lane 2, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of ColGS3aS3b(S997C) with fluorescein is 32kDa. This photo clearly shows the molecular weight of ColGS3aS3b(S997C) at 32kDa.

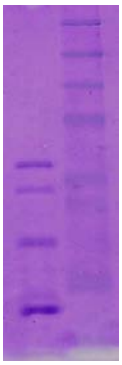


Figure 85. ColGS3aS3b(Y996A,S997C) +Fluorescien (9/18/14)

Studying the distribution of Col G S3aS3b (Y996A/S997C) using fluorescein crosslinker. Lane 1, ColGS3aS3b(Y996A/S997C) with fluorescein under UV light. Lane 2, molecular weight marker (175, 80, 58, 46, 30, 25, 17 kDa). The apparent molecular weight of ColGS3aS3b(Y996A,S997C) with fluorescein is 32kDa.

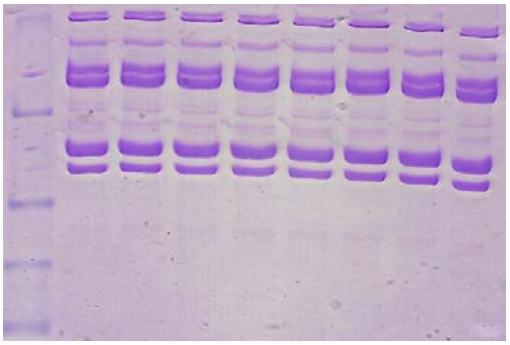


Figure 86. ColGS3aS3b(S997C) +APDP + Rattus Type I Collagen (New protein, 11/7/14)

Mass production of 8 gels with 10 wells each, all containing APDP attached ColGS3aS3b(S997C) with rattus type I collagen after UV radiation. The α_1, α_2, and complex bands were cut out of the gel and sent to mass spectrometry facilities for analysis.

Figure 87. Sequence of rattus type I collagen labeled with possible binding sites from mass spectrometry analysis.

Rattus type I collagen sequence aligned to show collagen fibril formation. Red box indicates the most probable binding site, because crosslinked amino acid for α_1 and α_2 are within three amino acids of eachother. Also, the peptide fragments were confirmed by both Mascot and Skyline software. The orange regions indicate predicted undertwisted regions. As can be seen the red box is located in one of predicted undertwisted regions, which adds support to the hypothesis that collagenases bind to undertwisted regions in collagen. After plotting the possible binding sites on a map of ligand binding sites, it was determined that tandem CBD binds in close proximity to interleukin-2 and integrin binding sites [4].

VI. Discussion

Tandem CBD from ColG binds collagen tighter than all the other clostridial collagen binders. Uchida et al. has recently shown that a tighter binding CBD construct fused with basic fibroblast growth factor enhances periosteal bone formation [13]. The tandem CBD from Col G binds tighter to collagen than the construct used by Uchida's study suggesting that when used it could be more potent in osteogenesis. My research is part of a scheme to understand and to develop novel wound healing in collaboration with various labs and Nippi Inc.

Philominathan *et al.* observed that CBD binds to undertwisted regions of collagen using NMR and SAXS [17]. It was first shown that CBD only binds at the C-terminus of mini-collagen. However, in earlier experiments using gold-labeled CBD, Toyoshima *et al*. showed that CBD could bind to more sites than just the C-terminus of tropocollagen [18]. This promoted further experimentation by Philominathan *et al.* These

experiments showed that CBD binds to collagen at undertwisted regions in the middle as well as at the C-terminus [8]. It was also seen that CBD binds without causing any structural changes and, therefore, does not facilitate the unwinding of collagen fibers. Yasui and Koide developed a method using photoreactive crosslinkers to determine protein binding sites on collagen [16]. It has been speculated that tropocollagne consists of five to six undertwisted regions. It was expected that the same procedure could be used to determine which undertwisted regions the CBD preferred to bind.

In order attach a crosslinker specifically at the lip of the collagen binding pocket of CBD, overlap extension PCR was used to introduce a cysteine mutation in order to allow crosslinker attachment. Using the following method it was confirmed that polymerase chain reactions successfully introduced the desired mutations into each plasmid DNA for all mutagenesis projects, which included ColGS3aS3b (S997C), ColGS3aS3b (Y996A,S997C), and ColGS3aS3b (Y994A,S997C) The Big Dye Terminator v1.1 Cycle Sequencing Kit was used to prepare DNA for a sequencing machine. After determining the sequences from the chromatograms obtained from the sequencing machine, the experimental DNA strand sequences and expected DNA sequences were compared and all mutants were confirmed to have the proper nucleotide orientation.

Afterwards, the above mutants along with Col GS3aS3b, ColGS3a (S879C), ColGS3aS3b (S879C), ColGS3b (Y996A,S997C) were successfully overexpressed and purified, following similar protocol used by Osamu Matsushita, M.D./Ph.D. [10]. Successful production of the proteins was also observed after analysis by SDS-PAGE. A band was present on the gels that showed that the proteins contained the proper molecular

weight. Also bands appeared to on the gel to show that the cysteine mutation was introduced.

Following the procedure developed by Yasui and Koide and modified by Keisuke Tanaka, several different crosslinking experiments were performed [16] . The results from SDS-PAGE showed that four different crosslinkers successfully attached to the introduced cysteine mutations of ColGS3aS3b(S997C). Crosslinker attachment was observed by comparing the masses of mutant CBD without crosslinker to mutant CBD with crosslinker displayed on SDS-PAGE. SDS-PAGE also showed that ColGS3aS3b(S997C) could no longer form dimers without UV exposure. After UV exposure dimer formation became possible again. These two properties are indicative of photocrosslinked cysteine mutations. Even though the crosslinkers were able to successfully attach not all of them were suitable for further studies. A biotin crosslinker was used in an attempt to purify the collagen-CBD complex. For unexplained reasons, the biotin crosslinker caused fragmentation of the protein upon binding to ColGS3aS3b(S997C). In order to rule out protease activity, another experiment was performed with a buffer that prevents proteolytic activity and the same results occurred. Therefore, this fragmentation could have been due to the size of the larger substituents, such as benzophene and biotin. Also, the purification column was unsuccessful.

After repeating several different crosslinking assays with APDP and PEAS attached CBD, it became clear that APDP produced the most consistent results. Even though it appeared that PEAS had become bound to CBD it was unsuccessful in crosslinking assays with rattus type I and bovine type III collagen. Looking at Figure 75, it can be seen in lane 6 that the dimer product of the PEAS attached CBD is lower in

concentration then all the other bands that appear on the gel. This could explain why the crosslinking assays were unsuccessful. The correct PEAS attached CBD product was in such low concentrations that it was unable to make the necessary contacts and bind to collagen. Among all the crosslinkers, fluorescein consistently attached and showed the distribution. It was also successfully used in remodeling collagen studies. Therefore, APDP and fluorescein were used in crosslinking experiments with the control protein, ColGS3aS3b(Y996A,S997C). The crosslinker attachment was confirmed using similar methods to those done for ColGS3aS3b(S997C).

Rattus type I collagen and bovine type III collagen provided by Nippi Inc, were used in different crosslinking assays. Type I collagen is heterotrimer with two α_1 chains and one α_2 chain. The former offers some advantages because it is abundant and a previous study showed that CBD binds to both skin and bone that are rich in type I collagen [4]. In comparison type III collagen is rare but a homotrimer, and therefore could make analysis simpler. The crosslinking assays were analyzed by SDS-PAGE under both non-reduced and reduced conditions. Upon UV exposure, the APDP attached ColGS3aS3b(S997C) crosslinked to the α_1 region of rattus type I collagen, which would appear as a band between the α_1 and β_2 bands of rattus type I collagen. Also, APDP attached ColGS3aS3b(S997C) crosslinked to the β_1 region of collagen, which is observed as a higher molecular weight band above the β_1 band. The reduced conditions showed the disappearance of these bands, and, therefore, confirmed the crosslinking had been successful. Reduction reduces the disulfide bridge between the crosslinker and collagen and therefore the band disappears. The PEAS attached CBD did not successfully bind to bovine type III or rattus type I collagen. Since the PEAS crosslinker was not successful

for both types of collagen it cannot be determined which collagen produced the best results. Further experimentation should be done with APDP attached CBD in order to determine which collagen provides the best results.

After performing the crosslinking assay, the Yasui and Koide method used 2-D electrophoresis to analyze where the proteins became bound to collagen. However, in order to perform a more specific and detailed analysis of where CBD binds to collagen it was expected that mass spectrometry could be used for analysis. In order to do this the the cross-linked CBD-collagen complex was subjected to digestion by trypsin. After analysis of mass spectrometry results, it was determined that this procedure was successful and, therefore, Yasui and Koide's method was improved upon. This new procedure can be used to study other molecular pathways in order to help in the fight for disease prevention.

The results from mass spectrometry show that tandem CBD binds to both the α_1 and α_2 bands of collagen, possibly through the formation of a complex. The data results from comparison of multiple samples. If multiple samples contained the same tryptic fragment identification, it was considered less likely to be a false positive. Multiple samples show that binding occurs in the α_1 chain for the peptide sequence K.GEPGPAGVQGPPGPAGEEGK.R [270, 289]. This peptide sequence is confirmed by both the Mascot and Skyline software. For the α_2 chain, the peptide sequence R.GSPGEPGSAGPAGPPGLR.G [292,309] is also confirmed by many samples. After plotting the collagen sequence, it is seen that the binding amino acids from the α_1 and α_2 chains are within three amino acids from each other. This observation strengthens our assumption that this is the most probable binding site.

After further examination, this peptide fragment is located in a region of collagen that was predicted to contain undertwisted collagen fibrils. This observation indicates that CBD might actually be binding to undertwisted regions, not only in synthesized mini collagen, but also in the full collagen protein. Also, it appears that CBD binds in regions near interleukin 2 and integrin binding sites, which are located on the α_1 and α_2 chain respectively. This observation indicates that using tandem CBD as a drug delivery system could result in negative side effects, however only further experimentation will confirm this assumption.

Upon comparison to work done by Keisuke Tanaka with ColGS3b domain, it is shown that he had located the same binding site while experimenting with bovine collagen and the PEAS crosslinker. This comparison is significant because the cysteine mutation introduce into the tandem CBD is located in the S3b domain. Therefore, it is expected that they should bind to similar locations on collagen. This assumption was proven to be correct because both experiments showed binding to the same peptide fragment. Also, since the results were similar specificity caused by crosslinker and collagen type can be ruled because different crosslinkers and type of collagens were used.

Remodeling collagen is necessary for tissue development, regeneration and maintenance. However, in excess this remodeling can lead to several pathological conditions including cancer [7] . Remodeling collagen consists of undertwisted regions. It was expected that CBD could bind to the undertwisted region and be used for *in vivo* and *in vitro* analysis of cancer cells.

In order to confirm that CBD can bind to remodeling collagen in cancer cells, fluorescein attached ColGS3aS3b(S997C) was used to target remodeling collagen *in*

vitro. After introducing fluorescein attached ColGS3aS3b(S997C) it became clear that tandem CBD recognized collagen secreted by pancreatic cells and became bound to the extracellular matrix of the cells. A control, fluorescein attached ColGS3aS3b(Y996A,S997C) was used and showed that it was in fact the fluorescein attached ColGS3aS3b(S997C) that became bound to the outer part of the cell.

In summary, I was able to generate five mutants to lay foundation for studying the tandem CBD. The tandem CBD binds to collagen the tightest among bacterial collagenase derived collagen binders. X-ray structure of the tandem CBD suggests that it achieves its binding by possibly wedging itself between collagen fibril. Mass spectrometry results confirmed this assumption because the identified binding site is located within the collagen fibril. Also, mass spectrometry showed that it binds to both α_1 and α_2 bands of collagen. In addition, the analysis of the CBD as a diagnostic and therapeutic protein for cancer involving fluorescein attached ColGS3aS3b(S997C) looks promising.

Future experiments will include observing fluorescein attached ColGS3aS3b(S997C) in human pancreatic tissue, which is expected to show binding of CBD to human tissue. These experiments will provide more insight into using the tandem CBD as an imaging technique or for repairing tissues after cancer cell death. Also this could lead to using CBD as a delivery system with a toxin to kill cancer cells.

Along with those experiments, one more crosslinking experiment with APDP and rattus type I collagen should be done in order to solidify the findings of our previous experiments. Also, crosslinking to the β chain should be determined using rattus type I collagen. Additionally, more crosslinking experiments using APDP attached

ColGS3aS3b(S997C) should be performed with bovine and porcine derived type I collagen molecules. These type I collagen share 95% sequence identity and are predicted to possess six underwound regions, thus should result in similar outcome. Experiments should be done to identify the binding preference for tandem CBD to MMP fragment. Furthermore, crosslinking experiments need to be done using the other mutants produced so that it can be determined where S3a binds.

VII. Bibliography

1. Kadler, K.E., et al., *Collagen fibril formation.* The Biochemical journal, 1996. **316 (Pt 1)**: p. 1-11.
2. Bella, J., *A first census of collagen interruptions: collagen's own stutters and stammers.* Journal of structural biology, 2014. **186**(3): p. 438-50.
3. Persikov, A.V., J.A. Ramshaw, and B. Brodsky, *Prediction of collagen stability from amino acid sequence.* The Journal of biological chemistry, 2005. **280**(19): p. 19343-9.
4. Di Lullo, G.A., et al., *Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen.* The Journal of biological chemistry, 2002. **277**(6): p. 4223-31.
5. Liu, X., et al., *Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(5): p. 1852-6.
6. Nowack, H., et al., *Preparation and use in immunohistology of antibodies specific for type I and type III collagen and procollagen.* Journal of immunological methods, 1976. **12**(1-2): p. 117-24.
7. Li, Y., et al., *Targeting collagen strands by photo-triggered triple-helix hybridization.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(37): p. 14767-72.
8. Philominathan, S.T., et al., *Bacterial collagen-binding domain targets undertwisted regions of collagen.* Protein science : a publication of the Protein Society, 2012. **21**(10): p. 1554-65.
9. Ohbayashi, N., et al., *Enhancement of the structural stability of full-length clostridial collagenase by calcium ions.* Applied and environmental microbiology, 2012. **78**(16): p. 5839-44.
10. Matsushita, O., et al., *A study of the collagen-binding domain of a 116-kDa Clostridium histolyticum collagenase.* The Journal of biological chemistry, 1998. **273**(6): p. 3643-8.
11. Matsushita, O., et al., *Substrate recognition by the collagen-binding domain of Clostridium histolyticum class I collagenase.* The Journal of biological chemistry, 2001. **276**(12): p. 8761-70.

12. Ponnapakkam, T., et al., *Monthly administration of a novel PTH-collagen binding domain fusion protein is anabolic in mice.* Calcified tissue international, 2011. **88**(6): p. 511-20.
13. Ponnapakkam, T., et al., *A single injection of the anabolic bone agent, parathyroid hormone-collagen binding domain (PTH-CBD), results in sustained increases in bone mineral density for up to 12 months in normal female mice.* Calcified tissue international, 2012. **91**(3): p. 196-203.
14. Katikaneni, R., et al., *Treatment for chemotherapy-induced alopecia in mice using parathyroid hormone agonists and antagonists linked to a collagen binding domain.* International journal of cancer. Journal international du cancer, 2012. **131**(5): p. E813-21.
15. Uchida, K., et al., *Enhancement of periosteal bone formation by basic fibroblast-derived growth factor containing polycystic kidney disease and collagen-binding domains from Clostridium histolyticum collagenase.* Journal of tissue engineering and regenerative medicine, 2015.
16. Yasui, N. and T. Koide, *Collagen-protein interactions mapped by phototriggered thiol introduction.* Journal of the American Chemical Society, 2003. **125**(51): p. 15728-9.
17. Philominathan, S.T., et al., *Unidirectional binding of clostridial collagenase to triple helical substrates.* The Journal of biological chemistry, 2009. **284**(16): p. 10868-76.
18. Toyoshima, T., et al., *Collagen-binding domain of a Clostridium histolyticum collagenase exhibits a broad substrate spectrum both in vitro and in vivo.* Connective tissue research, 2001. **42**(4): p. 281-90.