Fall 2020

Determining a Method for Expressing and Purifying Cytochrome P450 4V2: A Protein Involved in Bietti's Crystalline Dystrophy

Cody Lane Turner
Missouri State University, Turner133@live.missouristate.edu

DETERMINING A METHOD FOR EXPRESSING AND PURIFYING CYTOCHROME P450 4V2: A PROTEIN INVOLVED IN BIETTI'S CRYSTALLINE DYSTROPHY

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Chemistry

By

Cody Lane Turner

December 2020

DETERMINING A METHOD FOR EXPRESSING AND PURIFYING CYTOCHROME P450 4V2: A PROTEIN INVOLVED IN BIETTI'S CRYSTALLINE DYSTROPHY

Chemistry

Missouri State University, December 2020

Master of Science

Cody Lane Turner

ABSTRACT

Within the Cytochrome P450 class of enzymes, there are a group known as the "orphan" cytochromes. The "orphan" classification comes from the poorly understood *in vivo* functionality and substrate specificity. Cytochrome P450 4V2 (*CYP4V2*) is one of these "orphans" and belongs to the *CYP4* family. The *CYP4* family is known for the omega oxidation of endogenous fatty acids. This family is most commonly found on chromosome 1 (*CYP4ABXZ*). *CYP4V2* is unique in that its location is bound to chromosome 4 as discovered by Jiao in 2004. Mutations within the *CYP4V2* gene have been associated with the disease known as Bietti's Crystalline Dystrophy (BCD). BCD is an autosomal recessive disorder presenting with crystalline and lipid deposits in the cornea and retina. Gian Bietti first reported the disease in 1937 in three patients, and in 2004 Li and his research group identified the *CYP4V2* gene to be the cause. The unknown nature surrounding wild-type *CYP4V2* is where this research study gets its purpose. To start declassifying its "orphan" status, a method for the expression and purification of protein has been developed and optimized to obtain protein for functional characterization and structural biology studies. Preliminary binding data for wild-type enzyme has yet to be obtained, and a crystal hit has been identified with further optimization required.

KEYWORDS: orphan cytochrome P450, enzyme, Bietti's crystalline dystrophy, molecular biology, expression, purification, protein crystallography

DETERMINING A METHOD FOR EXPRESSING AND PURIFYING CYTOCHROME P450 4V2: A PROTEIN INVOLVED IN BIETTI'S CRYSTALLINE DYSTROPHY

By

Cody Lane Turner

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Chemistry

December 2020

Approved:

Natasha DeVore, Ph. D., Thesis Committee Chair

Gary Meints, Ph. D., Committee Member

Gautam Bhattacharyya, Ph. D., Committee Member

Amy Hulme, Ph. D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to thank the following people for their support during the course of my graduate studies. To all the friends I made on my college journey, you all shaped who I became today and allowed me to discover myself. To the graduate students, faculty, and students I had the pleasure of interacting with, you all deserve a large round of applause. Jessica Bruer deserves special recognition for the near daily support and uplifting comments that have pushed me through our joint journey through graduate school and my final semester. To my family who have always kept up with my life and even pestered me with the after-graduation questions I never knew how to answer, and still sometimes do not.

I would also like the thank Dr. Natasha DeVore for taking me on as her graduate student midway through my program when I was struggling to keep my head above water finding a research lab. She further developed my interest in the area of biochemistry and helped me find my home in the many disciplines offered to chemistry students. To my thesis committee, Dr. Gary Meints, Dr. Gautam Bhattacharyya, and Dr. Amy Hulme, thank you for agreeing to serve on my committee.

I dedicate this thesis to my mother, Lynn Turner, for always being there for me, even nine hundred miles away.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 – INTRODUCTION

Section 1.1 Cytochrome P450

Cytochrome P450s (CYP) are a superfamily of heme proteins catalyzing a range of oxidative reactions. The name P450 comes from the characteristic 450-nm absorption peak in its reduced and carbon monoxide bound state (Raucy and Allen, 2001). The human genome encodes for 57 different isozymes of CYPs each differentiated with a unique alphanumeric code denoting which family and subfamily the protein belongs. A shared sequence identity >40% classifies the cytochrome into the same family, >55% identity classifies it into the same subfamily, the final number in code denotes the polypeptide chain (ex. CYP4V2, family 4, subfamily V, polypeptide 2) (Edson and Rettie, 2013). The interest in cytochrome P450s stems from their roles in many biological processes. Cytochromes oxidize exogenous compounds, drugs, pollutants, and other natural compounds, and endogenous chemicals including steroids, eicosanoids, and other regulatory molecules. These proteins are ubiquitous and the need to understand their diverse function has driven the field of research. Within the superfamily, there is a class called the "orphan" cytochromes. They are given this distinction because their substrates and *in vivo* functions are not understood well. (Raucy and Allen, 2001; Kelly et al., 2011; Edson and Rettie, 2013). This chapter will focus on the structure, redox partner, its catalytic capabilities, and the CYP4 family involved with Bietti's Crystalline Dystrophy (BCD).

Section 1.1.1 Structure. Overall, the structure of CYPs are highly conserved. The closer to the active site, heme, the more conserved the structure becomes. Two helices, I and L, and the β-bulge, are the most conserved due to their interactions with the heme ligand and conservation of heme-thiolate chemistry, Figure 1. Regions not highly conserved are those that are involved

with the substrate specificity of each isoform. Eukaryotic CYPs are usually membrane-bound to the endoplasmic reticulum (ER). A short transmembrane helix (TRH) of about 20 amino acid is what binds the folded protein to the ER. The TRH is not required for catalytic activity, and in recombinant protein studies, is often not included in the insertion gene. There are different conformations that CYPs can exhibit in the cell, Figure 2.

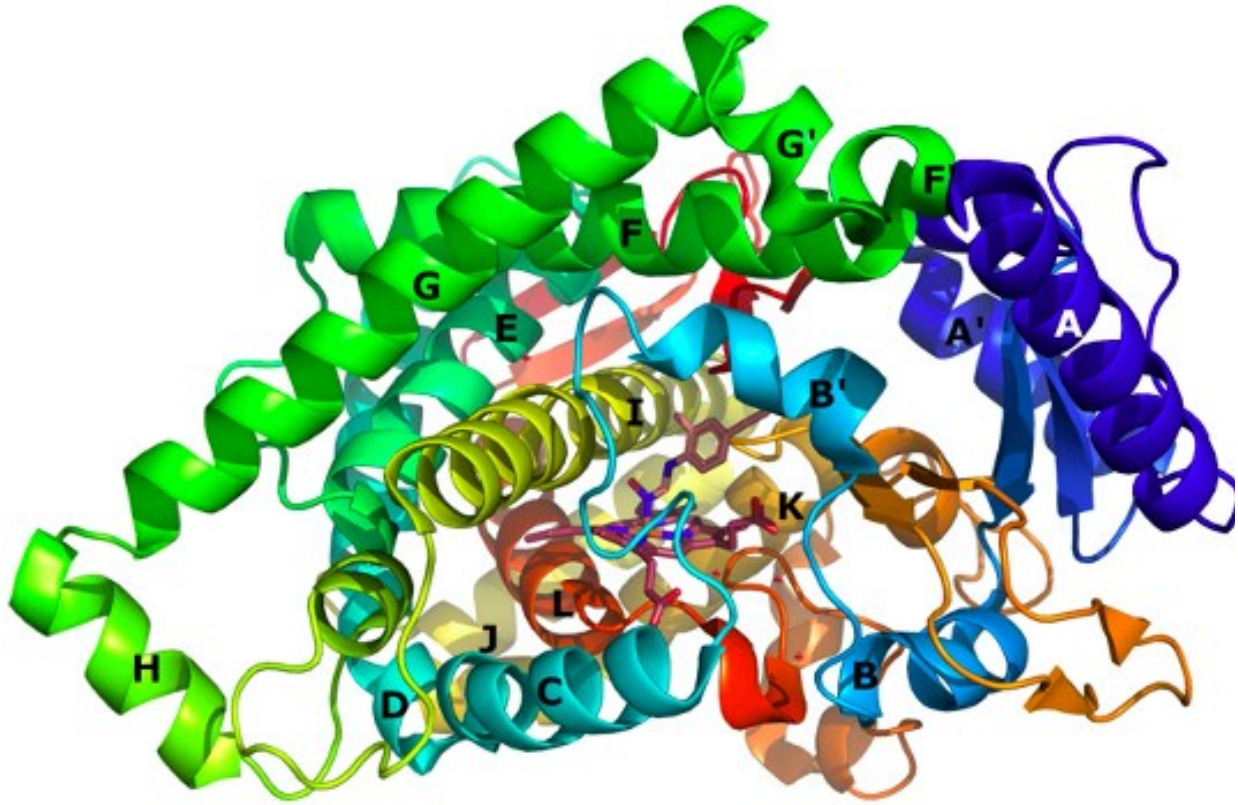

Figure 1. Crystal structure of Cytochrome 4B2 labelling all helices and showing the heme ligand in red.

Shifting between these conformations allows for different substrates to bind to the active site (Raucy and Allen, 2001; Ortiz de Montellano, 2015).

In the closed-state, CYPs exhibit a high amount of thermal mobility in the F and G helices, indicating these structures are involved in the conformational shifts allowing substrates

to enter the active site (Ortiz de Montellano, 2015). Once a substrate binds, the F and G helices move to close the entry. The entryway to different CYPs can accommodate various sized substrates. Once the substrate has entered, it is brought within 4-5 Å of the heme iron. Oxidation of the substrate proceeds with the help of a redox partner that facilitates a series of electron transfer reactions. The redox partner complexes loosely with the CYP complex allowing for quick dissociation and a rapid enzymatic turnover (Ortiz de Montellano, 2015).

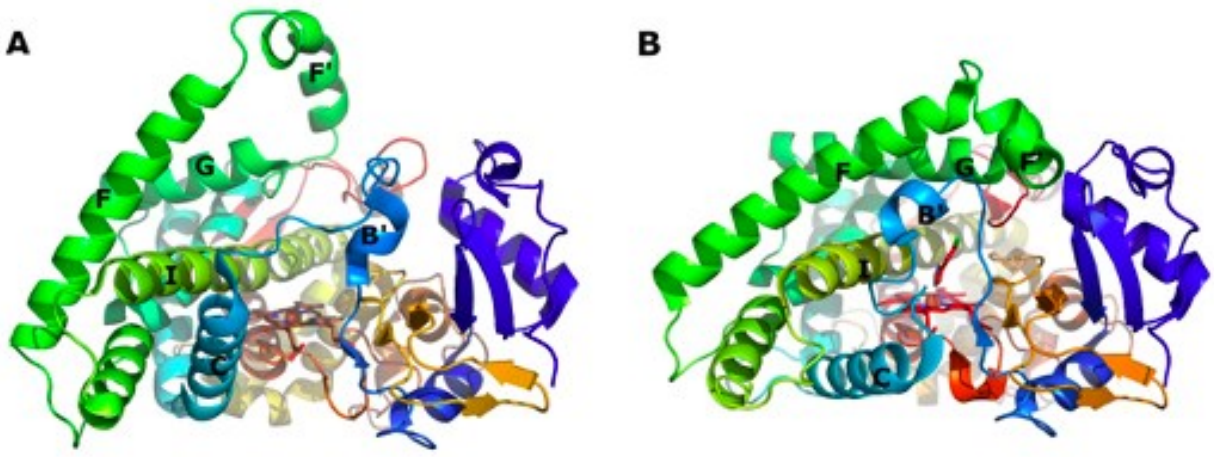

Figure 2. Depiction of Cytochrome P450 CYP2B4 in the open (PDB 1PO5) and closed (PDB 1SUO) conformations.

Section 1.1.2 Redox Partner. Oxidation of a substrate happens through the reduction of nicotinamide adenine dinucleotide phosphate (NADPH) and a series of electron transfer reactions (Ortiz de Montellano, 2015). The NADPH-cytochrome CYP oxidoreductase is one such oxidation partner (POR), Figure 3. Similar to the CYP complex, POR is a membrane-bound complex containing two flavin moieties, flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) (Ortiz de Montellano, 2015). NADPH is a two-electron donor, with CYP heme being a one-electron acceptor. FAD accepts a hydride ion from NADPH and then funnels the electrons one at a time to FMN. FMN then transfers the electrons to the CYP. This series of

transfer reactions gives rise to oxidation of many different compounds. Focus will be put on the oxidation of lipids for purposes here (Ortiz de Montellano, 2015).

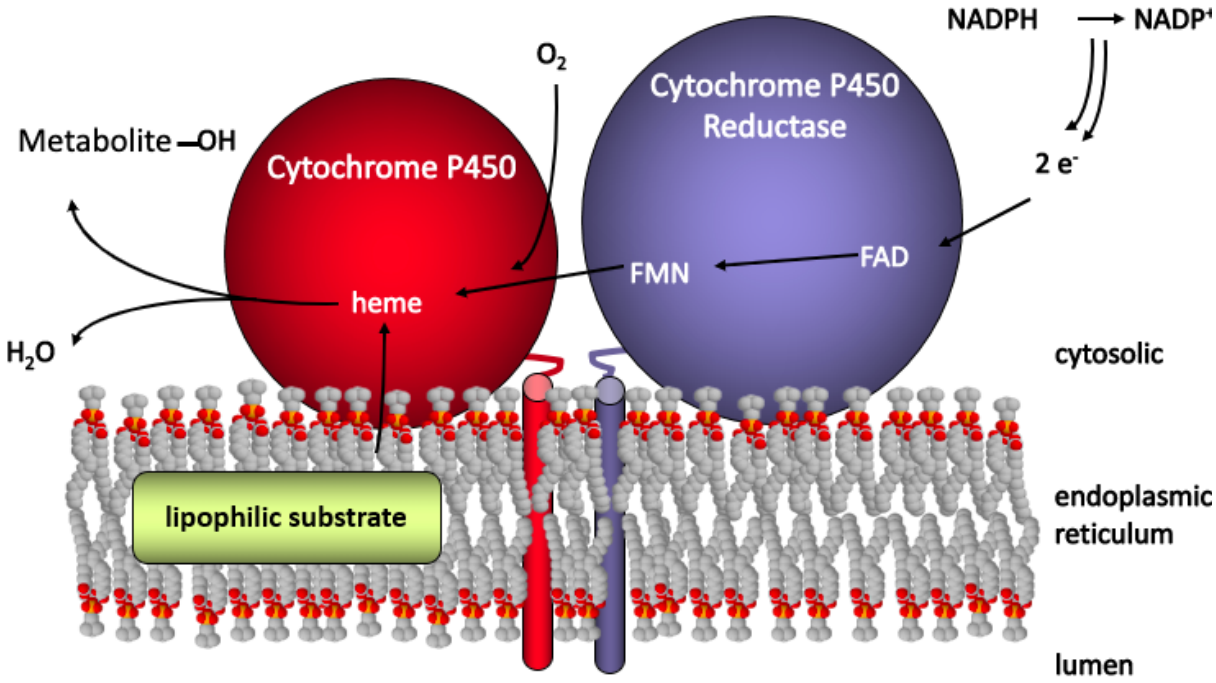

Figure 3. Cartoon depiction of Cytochrome P450 in complex with the redox partner Cytochrome P450 Reductase.

Section 1.1.3 Enzymatic Role. One of the more recent discoveries of CYP oxidative power is the metabolism of arachidonic acid to form eicosanoid products. There are three main sites for cytochrome P450 catalyzed for highly regioselective lipid oxidation: 1. ω/ω-1 hydroxylase, 2. mid-chain hydroxylases, and 3. epoxygenases (Ortiz de Montellano, 2015).

The ω/ω-1 hydroxylase catalyzes the addition of a hydroxy group at the terminal carbon or near the terminal carbon to form 16-20-hydroxyeicosatetranoic acids (HETEs) (Ortiz de Montellano, 2015). The hydroxylation of the terminal carbon is done mainly through the CYP4 enzyme (Ortiz de Montellano, 2015). The formation of terminal HETEs leads a variety of effects in the body such as renal vasodilation, inhibiting adenosine triphosphate (ATP) activity in renal

tubes, and promoting hypertension (Ortiz de Montellano, 2015). Further discussion of the CYP4 is provided in subsequent sections. Mid-chain hydroxylases form 5-, 8-, 9-, 11-, 12, and 15-HETEs. Hydroxylation in the mid-chain carbon is mainly done through the CYP1 and CYP2 families (Ortiz de Montellano, 2015). Mid-chain HETES regulate aqueous humor excretion, increases Na^+ and K^+ excretion in urine, and also promote neutrophil migration (Ortiz de Montellano, 2015). Epoxygenases generate *cis*-/*trans*- epoxyeicosatrienoic acids (EETs). The majority of CYPs that catalyze these transformations are from the CYP2. EETs function as membrane receptors, activate GPCR pathways, and have shown anti-inflammatory properties. (Ortiz de Montellano, 2015).

Section 1.1.4 Cytochrome P450 Family 4. As stated above, the CYP4 family is responsible for the ω-oxidation of saturated and polyunsaturated fatty acids (PUFAs) forming various HETEs. The energetic costs of hydroxylating the ω-carbon is greater than a substituted internal carbon (Edson and Rettie, 2013). Another note of significance is the presence of covalent heme binding. A glutamate residue is esterified to a methyl group on the heme ligand (Henne et al., 2001). This is postulated to give the active site rigidity for substrate hydroxylation.

Genes encoding for CYP4 families are clustered on chromosomes 1, 4, and 19. The orphan members of the CYP4 family are CYP4F22 and CYP4V2 (Kelly et al., 2011; Edson and Rettie, 2013). CYP4V2 is a distant relative to the rest of the CYP4 family. Its sequence identity is ~35%, under the usual cut-off for same family classification (Kelly et al., 2011). It is grouped in with the other CYP4 family members because of its ability to hydroxylate the ω-carbon similar to other CYP4 family members (Edson and Rettie, 2013). It has ubiquitous expression with a large local expression in the eye. This protein has clinical relevance in that it is involved

with a rare ocular disease known as Bietti's Crystalline Dystrophy (Kelly et al., 2011; Edson and Rettie, 2013).

Section 1.2 Bietti's Crystalline Dystrophy

Bietti's Crystalline Dystrophy is a rare ocular disease first reported by Gian Bietti in 1937 (Garcia-Garcia et al., 2010; Kelly et al., 2011; Darki et al., 2019). BCD is a progressive disease the manifests within the 2nd or 3rd decade of life. The disease has an initial clinical presentation of white-yellow crystals in the retina and the cornea, Figure 4 (Garcia-Garcia et al., 2010).

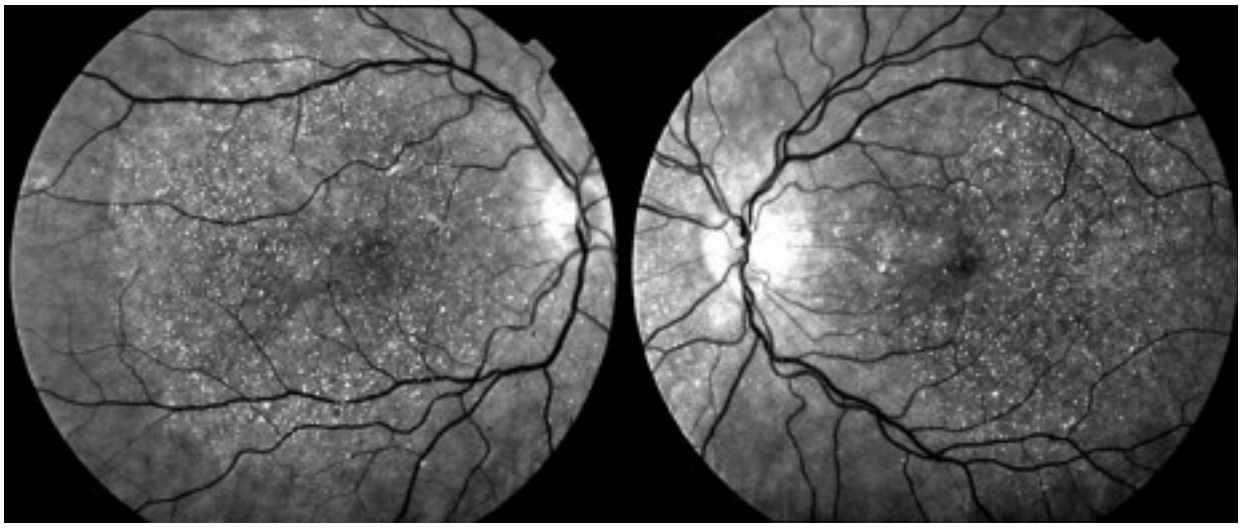

Figure 4. Initial clinical presentation of BCD. Numerous white crystals can be seen throughout the eye.

Near the end stages, central vision declines, night blindness develops, and the visual field is hindered (Garcia-Garcia et al., 2010). The exact composition of these crystals is unknown, but it has been speculated that they may be caused by abnormal lipid metabolism (Nakano et al., 2014). Since its discovery in 1937, the cause for BCD has been discovered to be mutations within the CYP4V2 gene (Garcia-Garcia et al., 2010). Genetic mapping has led to the discovery

of more than 20 mutations believed to cause BCD. (Garcia-Garcia et al., 2010; Kelly et al., 2011; Darki et al., 2019).

CYP4V2 has a large local expression in retinal tissues and metabolizes PUFAs (Garcia-Garcia et al., 2010; Nakano et al., 2014). The outside segment of our photoreceptor cells contains membrane disks rich in phospholipids esterified to docosahexaenoic acid (DHA) and other fatty acids (Garcia-Garcia et al., 2010). This outer segment is shed to regenerate membrane disks (Garcia-Garcia et al., 2010). The lipids from the disks are then take up by the retinal pigment epithelium (RPE) cells. Once inside the RPE cells, the lipids are transferred to the inner segment of photoreceptors where they are used to synthesize new disk membranes (Garcia-Garcia et al., 2010). This exchange of lipid molecules coupled with the role of CYP4V2 in fatty acid metabolism, and the abnormal metabolism of fatty acids gives some insight into how CYP4V2 may have a role in BCD (Garcia-Garcia et al., 2010; Nakano et al., 2014). However, the exact lipids that are metabolized by CYP4V2 in this process are not known.

Bietti's Crystalline Dystrophy can be classified into three main stages. Stage 1 is the death of RPE cells with crystal deposits forming at the center of the retina. This stage appears most commonly during the 2nd or 3rd decade of life (Kelly et al., 2011). The initial stage can be asymptomatic, or it can develop so slowly that an accurate diagnosis is difficult (Garcia-Garcia et al., 2010; Kelly et al., 2011). The presence of symptoms does not need to appear in both eyes at the same time and can develop asymmetrically (Garcia-Garcia et al., 2010; Kelly et al., 2011). Stage 2 has symptoms spreading from the center of the retina into the choriocapillaris with a decrease in the number of crystal deposits where the more advanced stages of atrophy occur (Garcia-Garcia et al., 2010; Kelly et al., 2011). Stage 3 is the most advanced with severe atrophy the RPE and choriocapillaris (Garcia-Garcia et al., 2010; Kelly et al., 2011). There are little to no

crystal deposits at this stage. Visual representation of the three stages is seen in Figure 5 (Garcia-Garcia et al., 2010; Kelly et al., 2011).

Section 1.3 Objectives

This project was designed to begin declassifying the "orphan" status of CYP4V2. An optimized method for expressing and purifying the protein must be determined to give enough protein for substrate binding and metabolism studies. One original goal was to compare the wild-type CYP4V2 enzyme to four mutations that have been identified in connection with BCD. The other main goal is to obtain an optimized crystal to elucidate the 3D structure.

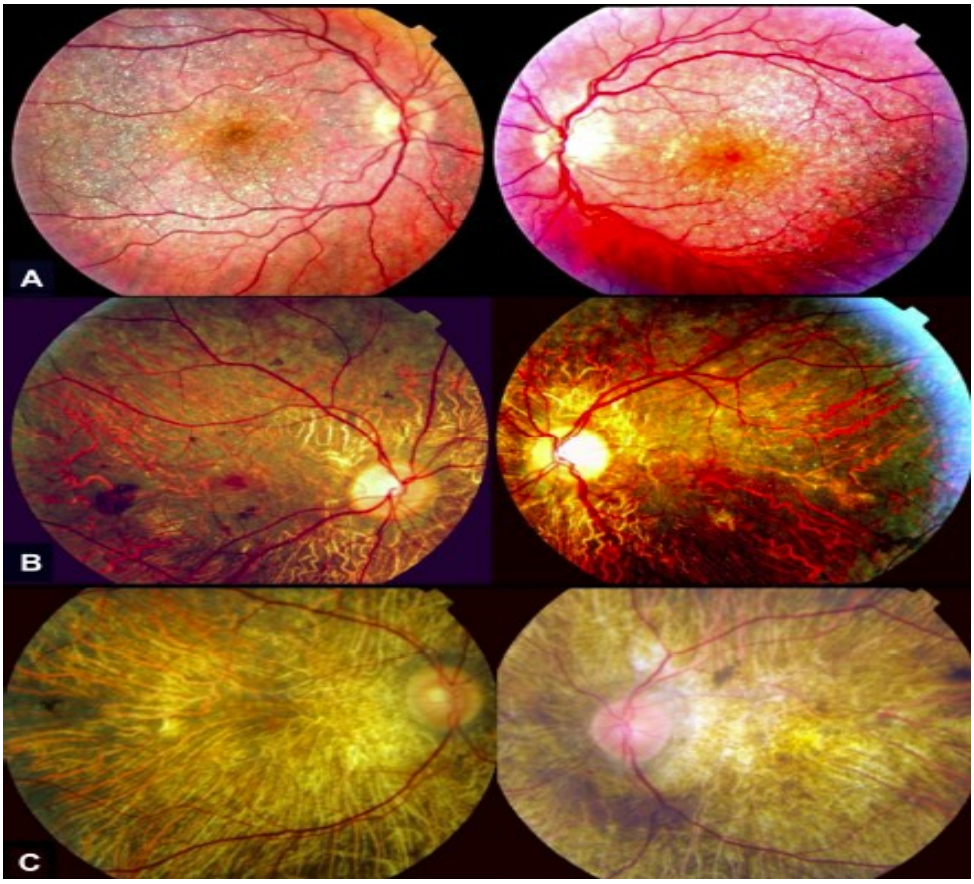

Figure 5. (A) Stage 1 BCD presence of crystals apparent in the eye. (B) Stage 2 crystals begin to disappear and RPE cells atrophy. (C) No crystal deposits and extensive atrophy of RPE and choriocapillaris.

CHAPTER 2 - MOLECULAR BIOLOGY – CLONING AND SITE-DIRECTED MUTAGENESIS

The world of recombinant proteins has revolutionized biochemistry. Previously small amounts of protein were purified from large samples of animal and plant tissue or large volumes of biological fluids. This was an impractical method because of the sheer volume of material needed for small recovery. With the introduction of recombinant technology, the new methods allowed for expression and purification of "large" volumes of protein that could be used for characterization and research while also revolutionizing the industrial and commercial sectors (Rosano and Ceccarelli, 2014).

Recombinant technology is an area of study in which genetic material is enzymatically manipulated through *in vitro* rearrangements (Rodriguez and Tait, 1983). At the basic level, obtaining a recombinant protein is achieved through taking the target DNA, cloning it into a desired expression vector, transforming it into a host cell (bacterial, insect, or mammalian), induction, and then extraction (Rosano and Ceccarelli, 2014; Manz et al., 2015). As simple as the process may seem, researchers may be left with low yields, protein inactivity, or failed growth.

This chapter focuses on the molecular cloning aspects of this project. Theory behind molecular cloning techniques of cloning and site-directed mutagenesis will be introduced. This provides a background to help those understand the processes used to clone the CYP4V2 gene into a suitable vector and the background required to understand the process that is site-directed mutagenesis.

Section 2.1 Polymerase Chain Reaction

Section 2.1.1 Principles of Polymerase Chain Reaction. The *in vitro* technique known as polymerase chain reaction (PCR) allows for the amplification of nucleic acid segments for analysis and cloning. The entire reaction takes place in a 0.2 mL tube which is placed into a thermocycler. The reaction tube contains: 1. DNA template, 2. DNA polymerases, 3. DNA primers, 4. free nucleotides, and 5. buffer. The DNA template is the sample of DNA of interest to be amplified. DNA polymerases are added to act as catalysts for complementary DNA synthesis. DNA primers are oligonucleotide segments that are complementary to the target DNA. Two different primers are used, a forward and reverse. Each primer anneals to one strand of double-stranded DNA (dsDNA). These primers are designed specifically for the target DNA and in such a manner that they do not self-anneal. Free nucleotides in the form of deoxyribonucleotide triphosphates (dNTPs) are added as the fuel for synthesis. These are depleted as the reaction takes place. The final component is a buffer. Buffers are used to provide a suitable environment for the polymerase to be efficient. In the thermocycler, there are three stages to the process of replication (Manz et al., 2015).

Stage 1 – DNA Denaturation. The stage elevates the temperature of the reaction vessel into the range of 95°C. The hydrogen bonds inside the dsDNA are disrupted and the previously annealed DNA separates into two single-stranded DNA (ssDNA) segments. At this stage, the polymerases are inactive (Manz et al., 2015).

Stage 2 – Primer Annealing. At this stage, the temperature drops between 50-60°C. This temperature range is acceptable for the hydrogen bonds between the nitrogenous bases to anneal. With the excess primer segments in solution, the two new ssDNA segments are prevented from

reannealing with each other. One primer is designed to anneal to the 3' end of ssDNA, and the other primer anneals to the 5' end (Manz et al., 2015).

Stage 3 – Primer extension. For this stage, the temperature is raised to around 72℃ which is an optimal temperature for the near perfect reproduction of complementary DNA. The polymerases synthesize the complementary DNA by incorporating matching dNTPs to the template strand. Once the segment is complete, extension stops (Manz et al., 2015).

These three stages make up one complete cycle. A complete reaction usually consists of 20-35 cycles of repeated heating and cooling. The temperature at which each stage takes place and the heating and cooling it takes to reach the temperatures is imperative to the success of reaction. The amount of DNA present in the tube usually doubles with each cycle (Manz et al., 2015).

Section 2.1.2 Rate of Amplification. In a perfect scenario, DNA concentration will double with every cycle, leading to an exponential model of amplification. The theoretical number of DNA molecules can be calculated by multiplying the initial number of DNA segments in the vessel by the exponential function, 2^n, where n is the number of cycles completed. Figure 6 shows the amount a single DNA complex is amplified over the first four cycles of PCR.

This is purely theoretical and the actual number of copies at the end of the reaction is less due to a few factors. As the cycles repeat, the reagents in solution deplete. Another reason is the introduction of errors by the polymerases leading to incorrect base pairing (Manz et al., 2015).

Section 2.2 Restriction Digest

One of the major discoveries in recombinant DNA technology is the presence of restriction enzymes. Restrictions enzymes are endonucleases that cut both strands of DNA at the

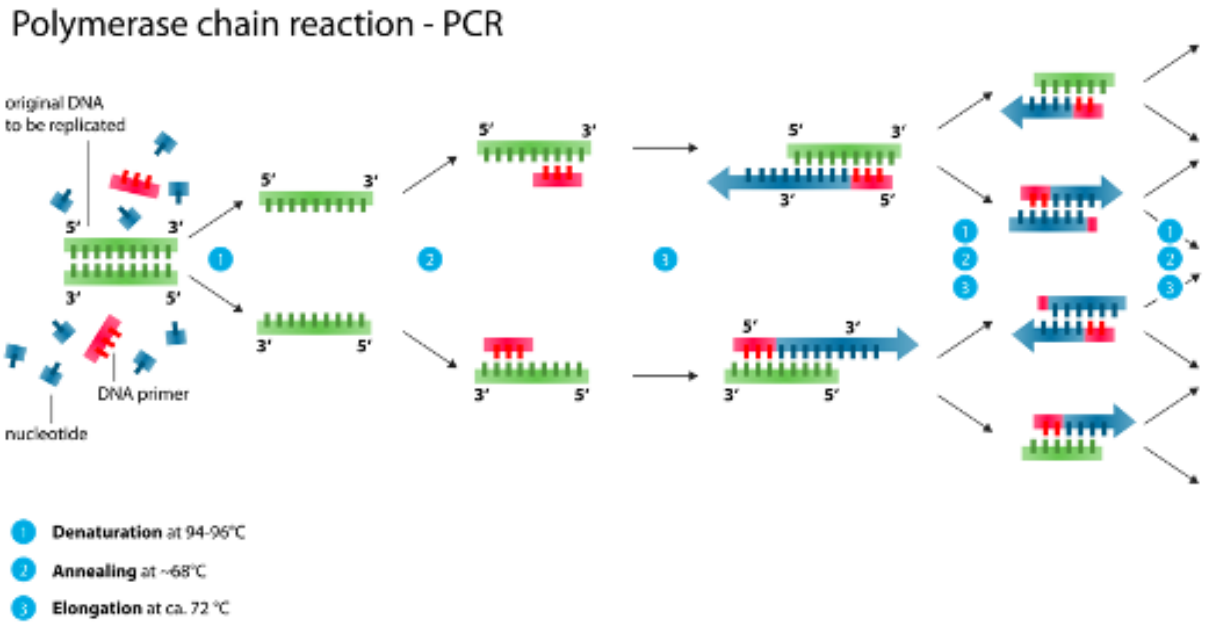

Figure 6. Theoretical amplification of the first four cycles of a PCR reaction. This image was made by Enzoklop, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons, no changes were made.

same time along the sequence at specific sites. These cuts create small fragments of DNA that can then be inserted into a vector and introduced into a host. There are two types of restriction enzymes: Type I and Type II (Birge, 1981; Freifelder, 1987; Griffiths et al., 1993).

Type I enzymes are nonspecific. These enzymes are made up of nonidentical subunits of high molecular weight (Birge, 1981; Freifelder, 1987). This enzyme creates nonspecific cuts along the DNA complex at a randomly located location many base pairs away from the original recognition sequence (Birge, 1981). Type II enzymes are specific restriction endonucleases that makes two single-stranded cuts within the recognition sequence (Birge, 1981). There are two types of cuts that can be made, flush or blunt cuts and cohesive cuts. Cohesive cuts are more desired for cloning. For flush cuts, the enzyme makes a cut directly down the center of symmetry of the recognition sequence. Cohesive cuts are made by placing nicks along the DNA complex in

such a manner that a tail of nucleotides extends past the nick. This done on both strands creating two tails of complementary sequences (Birge, 1981; Freifelder, 1987).

Type II restriction enzymes are important for cloning recombinant methodologies. PCR products are digested with restriction enzymes that match the restriction enzymes used to digest the vector of interest. This creates complimentary ends for the two to fuse together during the ligation process.

Section 2.3 Ligation

With the formation of two complementary components, the target gene now needs to be joined with the vector through a process called ligation. This process can be completed *in vitro* or *in vivo* by the enzyme known as DNA ligase. DNA ligases are found readily in cells and belong to a family of enzymes known as the nucleotidyl transferases (Çağlayan, 2019). These enzymes repair DNA nicks by catalyzing the formation of a phosphodiester bond between a 5' phosphate and 3' hydroxyl group on adjacent nucleotides (Çağlayan, 2019). There are three main steps in the ligation process. Step 1 involves the adenylation of the adenylation domain (AdD) forming a DNA ligase-adenylate complex. Adenosine triphosphate (ATP) is cleaved to form adenosine monophosphate (AMP), which is what binds to the AdD. In step 2, the DNA ligase binds to the nick in the DNA strand at the oligonucleotide binding domain (OBD) and the AMP is transferred to the 5' phosphate nucleotide. This forms an intermediate 5'AMP with DNA. The third, and final, step, DNA ligase catalyzes the reaction between the 5'-AMP and 3'OH releasing AMP in the process (Çağlayan, 2019). This process is irreversible because of the release of highly energetic pyrophosphate coupled with the AdD being adenylated, pushing the ligation process forward. With the process of ligation complete, the gene of interest has been cloned into

a suitable vector for transformation and expression, the final process to obtain a desired recombinant protein (Rodriguez and Tait, 1983; Çağlayan, 2019).

Section 2.4 Mutations and Site-Directed Mutagenesis

Section 2.4.1 Mutations. Mutations in genetics are inherited changes in the observable traits of an organism. At the molecular level, a mutation is an alteration in a base pairing, which can cause a change in the amino acid sequence of a protein. Sometimes mutations can produce the same amino acid by mutating to a different codon encoding for the amino acid. This section will introduce different types of mutations and how they affect polypeptides (Birge, 1981).

The first type of mutation is a base substitution. This is mutation is the replacement of one base with another (Birge, 1981). Specific type base substitutions are transition and transversion (Birge, 1981). Transition mutations are exchanging one pyrimidine for another pyrimidine or a purine for a purine (Birge, 1981). Transversion mutation replace pyrimidines for purines and vice versa. These mutations are the easiest to visualize and think through but can be difficult to detect in sequences. There are multiple codons that encode the same amino acids because of the high redundancy in the genetic code (Birge, 1981). The codon “UUU” encodes for the amino acid phenylalanine, so does the codon “UUC.” Missense mutations, a type of base substitutions, are easier to detect. This mutation causes the original amino acid to be replaced with another. “ACC” encodes for threonine, while “CCC” encodes for proline (Birge, 1981). This mutation would be easy to visualize because replacement with proline has detrimental effects in secondary and tertiary structure, especially in helices (Birge, 1981). The genetic code encodes for three termination codons “UGA,” “UAA,” and “UAG.” The insertion of any of these codons is called a missense mutation (Birge, 1981). When these codons are inserted into a

coding sequence, early termination of polypeptide elongation occurs (Birge, 1981). This creates a truncated protein altering activity (Birge, 1981).

Three other types of mutations are insertions, deletions, and frameshifts. Insertions and deletions are inverse mutations (Birge, 1981). Deletions are the removal of base pairs from the DNA, and insertions are the addition of base pairs in DNA. These can be single mutations or multiple mutations. If these mutations occur in multiples of three, then this adds or eliminates an amino acid from the polypeptide chain. If the mutation causes the insertion or deletion of a single base pair, this induces a frameshift mutation. If the insertion or deletion shifts the reading frame, the normal amino acid sequence is disrupted, forming gibberish proteins (Birge, 1981).

Section 2.4.2 Site-Directed Mutagenesis. Natural mutations can have serious effects on polypeptide structure and function. For the purpose of research interests, the ability to introduce mutations into a sequence of DNA is important for understanding their effects in hosts. The introduction of site-directed mutagenesis (SDM) revolutionized the research world. SDM methods can be either *in vivo* or *in vitro*. Synthetic oligonucleotide primers are designed such that, for a single-base mutation, the primers are complementary to the DNA sequence except for the mutated base. If the primers are designed correctly, they have no problem annealing to the parent DNA. Figure 7 double-primer mutagenesis design with single mutation primer.

Figure 7. Cartoon depiction of double-primer mutagenesis with single-primer mutation. This image was made by Rob Hunt, BY CC-SA 4.0, image not modified.

Single-primer designs create a heteroplasmid with one strand of DNA containing the mutation, while the other strand is a parent strand. Double-primer designs create a new plasmid

with the mutated base on both strands. Current SDM methods utilize the double-primer design wherein two primers are designed that are complementary to the parent DNA and thus to each other. This introduces the complication of primers self-dimerizing in solution which decreases the efficiency of the reaction and could prevent mutagenesis completely. Alternative protocols are developed to circumvent these complications and are commonplace in SDM (El-Gewely et al., n.d.; Carter, 1986; Edelheit et al., 2009).

Section 2.5 Electrophoresis

The use of electrophoresis can be applied to many biomolecules including DNA/RNA, proteins, and carbohydrates (Manz et al., 2015). Matrices for separation are usually done in gels made by cross-linked polymers of either agar, agarose, or polyacrylamide (Stellwagen, 2009; Manz et al., 2015).

Section 2.5.1 Theory of Electrophoresis. In the presence of an applied electric field, charged molecules in a conductive medium will travel through a gel consisting of cross-linked polymers. Conductive mediums are usually represented by the use of aqueous buffers that have high salt content and a pH that charges the molecule either negatively or positively which determines the flow of analytes to either the anode or cathode. Figure 8 shows the main principle behind electrophoresis.

If using a gel, the cross-linked polymers act as molecular sieves that can filter out the analyte based on sample size. Larger molecules are retarded more than smaller molecules, so molecules with similar charge to mass ratio can still be separated out (Manz et al., 2015).

Section 2.5.2 Gel Electrophoresis Instrumentation. There are two main parts to the electrophoresis instrumentation. First is the power supply, and the second is the chamber which

the gel sits with an electrolyte buffer. Power supplies are set to deliver both low and high voltage power. Electrodes are attached to the chamber and fashioned in such a way that they are immersed in running buffer. The chamber can either be vertical or horizontal. The running buffer is chosen to either have a high pH, creating negatively charged molecules, causing migration to happen from cathode to anode. If the pH is low, then the opposite movement results. When introducing sample into the gel, they are placed such that they migrate towards the electrode of analyte charge (Manz et al., 2015).

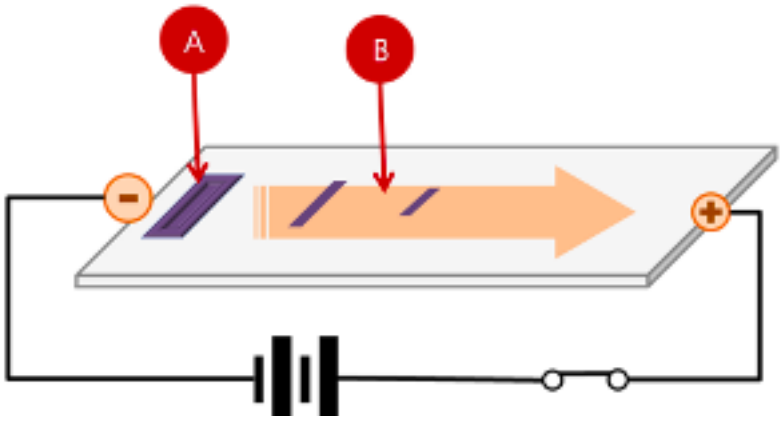

Figure 8. Principle behind electrophoresis. Negatively charged analytes move towards the cathode. Image by Jjw, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons, this image was not modified.

Section 2.5.3 Band Visualization. Sample analyte cannot be visualized under normal conditions and must be treated. The most common method of visualizing bands is staining with either fluorescent dyes, silver dyes, or *Coomassie brilliant blue*. A sample DNA ladder is runs in tandem with the analytes. The DNA ladder is a mixture of defined base pair lengths which give a standard set of molecular weights. The sample analytes can be visually inspected and compared against the ladder to determine approximate number of base pairs in the sample. (Manz et al., 2015).

Section 2.6 Methodology

Section 2.6.1 Polymerase Chain Reaction. To start a polymerase chain reaction (PCR) both the gene and the primers need to be diluted in water. On arrival from Thermo Fisher, the gene contained about five micrograms (μg) of Kan R plasmid DNA. To this, 20 microliters (μL) of water was added to make a 250 ng/μL solution. The forward primer arrived with a mass of 235.73 μg which corresponds to 32.04 nanomoles (nmol). To the primer tube, 1.282 mL of water was added to create a 25 micromolar (μM) solution, tubes were then vortexed to create a homogenous solution. All further dilutions of primers are provided in Table 1.

Table 1. Dilutions of Wild-Type Primers and Mutant Primers

Primer Designation	Molecular Weight (g/mol)	# of moles present (nmol)	mL used for dilution	Final Conc. (μM)
Rvs 4V2[1]	5882.89	30.05	1.22	24.63
Fwd F73L	5215	76.9	1.538	50
Rvs F73L	5162	78	1.56	50
Fwd I111T	5185	82.5	1.65	50
Rvs I111T	5203	52	1.04	50
Fwd T325I	4926	50.8	1.016	50
Rvs T325I	4793	61.7	1.234	50
Fwd H331P	5530	71.4	1.428	50
Rvs H331P	5468	61.8	1.236	50

[1] Due to human error in initial calculations, Rvs 4V2 was incorrectly diluted to 24.63 μM instead of 25 μM.

Once all dilutions were completed, PCR tubes can be set up. Two PCR tubes are run together in case of one failing. In the following order, each tube received 35.5 μL of autoclaved DI water, 1 μL CYP4V2 DNA, 1 μL 4V2 Fwd primer, 1 μL 4V2 Rvs primer, 1 μL dNTP (deoxynucleotide triphosphate), 10 μL 5x GC buffer, and 0.5 μL Phusion Polymerase. The tubes

are then placed in the PCR instrument. Method PCR2 was selected from the screen and the instrument was allowed to run. Method PCR2 initial setup provided in Table 2.

Section 2.6.2 Site-Directed Mutagenesis. To obtain the mutations of CYP4V2, site-directed mutagenesis is performed. Using the mutant primers listed in Table 3. and the Phusion Mastermix PCR kit, the PCR tubes will be set up similarly to the wild-type PCR tubes. Twenty-three microliters of DI autoclaved water, 25 µL of Phusion Mastermix, 0.5 µL forward mutant primer, 0.5 µL reverse mutant primer, and 1 µL of CYP4V2 template DNA is added to a PCR tube and run in duplicate using a modified PCR method. In step 2, the "mutant pcr" method, change the temperature of the annealing step to match the T_m of the mutant primers. T_m of primers given in Table 4 and are calculated using the calculator found on the Thermo Fischer website.

Table 2. PCR Method for Veriti 96 Well Thermal Cycler Wild-Type Reactions.

Stage	Step	Temperature (°C)	Time (s)
1	1	98	30
2	1	98	10
	2	53	30
	3	72	60
3	1	72	120
	2	4	Indefinitely

Section 2.6.3 Restriction Digest. Once PCR product was sequence verified, restriction digest was performed next. To each tube the following was added: 1µL of Dpn1, 1 µL HindIII, 1 µL NdeI, and 5 µL of FD Buffer. When doing a digest of plasmid, Dpn1 is not added. Following the addition of the restriction enzymes, the two tubes were placed in a New Brunswik Galaxy 170S incubator for thirty minutes to an hour.

Section 2.6.4 DNA Gel Electrophoresis. After the PCR product has been obtained, a DNA gel is used to determine if the PCR amplification worked. To make a gel, 1.5 g of agarose solid is weighed out and transferred to a large beaker. 100mL of 1x TAE Buffer was poured in. The entire beaker was taken to the microwave and heated in short 30 second bursts to dissolve the agarose solid. The contents of the beaker start vigorously bubbling near the end, so the beaker was watched through the microwave door, and heating was stopped if needed. Periodically, the beaker was removed from the microwave to check if the solid dissolved. Once it was dissolved, the hot beaker was carefully removed, and the contents were poured into a gel runner. A well comb was placed on one end of the gel and the entire apparatus was allowed to cool to room temperature on the counter. Once cooled down, it was moved into the refrigerator to finish gelling.

Table 3. Primer Sequences for Site-Directed Mutagenesis.

Primer Designation	Primer Sequence[2]
Fwd F73L	GTCGTGAATTATTTCAG
Rvs F73L	CTGAAATAATTCACGAC
Fwd I111T	TTGAAGTGACTCTGACC
Rvs I111T	GGTCAGAGTCACTTCAA
Fwd T325I	GAGGTTGATATCTTTA
Rvs T325I	TAAAGATATCAACCTC
Fwd H331P	GTTTGAAGGTCCTGATAC
Rvs H331P	GTATCAGGACCTTCAAAC

[2] Sequences read 5' → 3' with mutant base pair in red.

After the gelling process completes, the runner was removed, and the comb taken out. The runner with the gel was placed into the Kodak BioMax QS710 electrophoresis chamber, making sure to have the gel wells next to the black anode so they run towards the red cathode. The chamber was filled with 1X TAE buffer to cover the gel. A DNA ladder was made to have a

standard comparison. To 10 µL of DNA ladder 2 µL of EZ-Vision One visualization dye was added. Five microliters of PCR product are mixed with 1 µL of EZ-Vision One. Ten microliters of DNA ladder with dye was added into the first well and 5 µL of each PCR product with dye was added to subsequent wells. The Bio Rad PowerPac™ Basic was turned on and the voltage set to 120. The gel was allowed to run until the dye band reached about halfway through the gel. Once the gel was finished, it was transferred to the Transilluminator for visualization. PCR products in the wells fluoresced when placed in the Transilluminator.

Table 4. Annealing Temperatures of Mutant Primers.

Mutation	Annealing Temperature (°C)
F73L	50.5
I111T	56.0
T325I	46.1
H331P	55.2

Section 2.6.5 DNA Cleaning and Concentrating. After restriction digest, the tubes were removed from the incubator. Using Zymo's DNA Clean & Concentrator™-5 kit DNA is concentrated. The products from the two PCR tubes were combined into a 1.5 mL Eppendorf tube. Binding buffer is added to the combined product in a 5:1 ratio for PCR product (e.g. For 100 µL of PCR product, 500 µL of binding buffer is added). The entire contents of the Eppendorf tube were transferred to a Zymo-Spin™ column. The column was placed in a thermoscientific mySPIN12 centrifuge and balanced with another column. The centrifuge was set to spin at 10,000 rpm for thirty seconds. The eluate was discarded in a small beaker. 200 µL of wash buffer was added to the column and centrifuged again at 10,000 rpm for thirty seconds. The eluate was discarded, and the wash step was repeated once more. After discarding the second eluate 10 µL of DNA elution buffer was added to the column and allowed to incubate at

room temperature for one minute. After incubation, the column was transferred to a 1.5 mL Eppendorf tube and centrifuged again for 30 seconds. The eluate here is the cleaned and concentrated DNA.

Section 2.6.6 Plasmid Preparation. pCW CYP17A1 was provided as a gift through a material transfer agreement with Dr Emily Scott. A single DH5α colony containing pCW 17A1 was selected and used to start a 5 mL LB culture. Using a Thermo Scientific Gene JET Plasmid Miniprep kit, two 1.5 mL Eppendorf tubes, 1000 µL deliveries were made. The tubes were the centrifuged for two minutes. The liquid layer was decanted off and another set of cell delivery were made. This process was repeated once more. After the final centrifuge, the liquid was completely decanted off and 250 µL of resuspension solution was added one tube and the cells were resuspended and transferred to the other tube, and the cells were completely resuspended. Two-hundred and fifty milliliters of Lysis solution was added to the tube and inverted a few times. After inversion, 350µL of Neutralization solution was added and inverted a few times. After inversion, the tube was placed in the centrifuged and allowed to spin for 5 minutes at an rpm of 13,300. The supernatant solution was added to a Thermoscientific Gene JET spin column and centrifuged for one minute. The eluant was poured off and 500 µL of wash solution was added to the column and centrifuged for one minute. This wash step was repeated one time. After washing, the column was transferred to an Eppendorf tube, 50 µL of elution buffer was added, and it was allowed to incubate for two minutes. After incubation, it was centrifuged for two minutes and the eluant was collected.

Section 2.6.7 Gel DNA Recovery Kit. Plasmid collected in Miniprep Kit was digested and run in a DNA gel. To recover the Plasmid, a Zymoclean Gel DNA Recovery Kit was used. Gel slices were removed that contained the completely digested DNA. In a 3:1 ratio of agarose

dissolving buffer to plasmid gel, the gel was dissolved in a tube. This was allowed to incubate at 37℃ until completely dissolved. The dissolved agarose gel was transferred to a zymo-spin column and centrifuged for one minute at 13,300 rpm. Two-hundred microliters of buffer was added, centrifuged, and washed again. Ten microliters of elution buffer are added to the column, the column was transferred to an Eppendorf tube, and centrifuged for one minute.

Section 2.6.8 Sequencing Preparation. A single DH5α colony with pCW4V2 was selected and a culture was grown in 5 mL of LB media overnight. Culture was treated the next day according to Plasmid Preparation above. Using the eluant collected, 10 µL of plasmid was removed and placed in a PCR tube. To every other tube, 5 µL of pCW Fwd primer was added, and to the other tubes, pCW Rvs primer was added. The tubes were capped and wrapped in parafilm to ensure safe delivery. Tubes were then sent GenScript for sequencing analysis.

Section 2.7 Results

The beginning of any expression project starts with molecular cloning. The CYP4V2 gene was designed and ordered from Invitrogen by Thermo Fischer Scientific and arrived in the pMK-RQ vector backbone. To give ideal concentrations for PCR, the gene was diluted to an initial concentration of 250 ng/µL to give a final concentration of ~5 ng/µL in the PCR tube. The first attempt at PCR is highlighted in Figure 9. The first lane of the gel is the DNA ladder, the next two lanes are of PCR product, the remaining five lanes are of digested pCW plasmid. All lanes are visualized by adding a fluorescent dye, visible under a UV light source. The two lanes of 4V2 gene are stuck in the well indicating either a large kb product or digestion with restriction enzymes did not cleave the gene from the insert and the plasmid as a whole is sitting in the well.

The intense fluorescence is a qualitative indication that there is replicated product. The lanes with pCW17A1 being empty could indicate no DNA in solution.

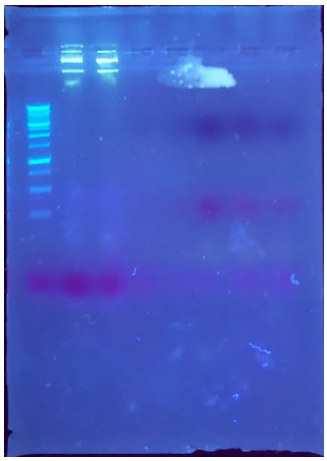

Figure 9. Electrophoresis gel showing the first attempt at PCR of the CYP4V2 gene (lanes 2 and 3) with product of pCW17A1 digestion (lanes 4-8).

For attempt number two, the buffer used in the PCR tube was changed. The first reaction utilized the fast digest (FD) buffer that came with the Phusion™ High Fidelity PCR kit, the second reaction switched from FD buffer to GC buffer. After digestion, another DNA gel was prepared and run. Lane one was DNA ladder, and the next two lanes were PCR product, Figure 10. This attempt had a successful replication and separation of the gene from the vector backbone. A second digestion of the pCW17A1 was attempted. This time, there is clear separation between 17A1 gene insert and the pCW backbone, which can be visualized in Figure 11. The bottom band is the gene insert and the top band is the vector backbone. The top band was dissected, and the DNA was extracted from the gel using the Zymoclean Gel DNA Recovery Kit.

Now that PCR product has been visualized and an empty plasmid backbone has been obtained, the two needed to be ligated together to make the pCW4V2 vector. In the presence of

ligase, the insert and plasmid are ligated together in a 3:1 ratio with 2.5 µL of ligase. After a period of incubation, the presumed ligated DNA was transformed into DH5α cells, and plated onto a lysogeny broth ampicillin (LB-amp) plate.

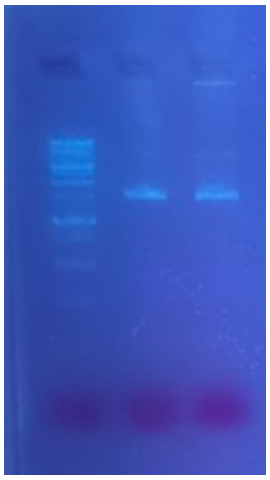

Figure 10. Banding present around the 15 kDa band of the DNA ladder representing the CYP4V2 gene.

The plate was incubated and the next day colonies were observed. Three colonies were picked, grown a small portion of LB media overnight, and the following day the DNA was extracted from the cells. To confirm ligation worked, extracted DNA was sent to GeneScript® for sequence verification. All three colonies picked from the plate were empty, no DNA insert was confirmed.

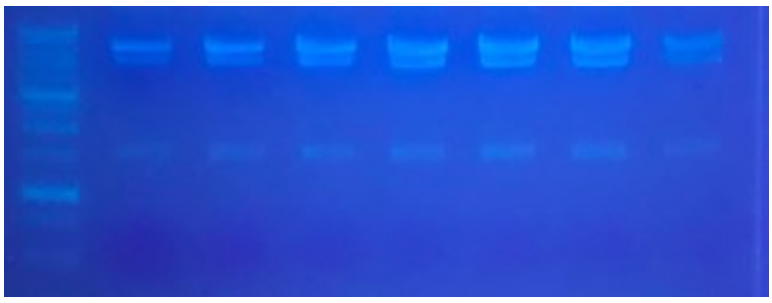

Figure 11. Successful digestion of pCW17A. Top band empty plasmid, bottom band 17A1 insert.

Another PCR was completed with the same conditions as attempt number two, and another ligation attempt was made. Cell transformation had an additional outgrowth step where transformed competent cells are incubated with a small portion of LB media to increase the number of cells. The cells were plated and left to incubate. Two plates were incubated for the two PCR products run, one plate had one large colony and the second plate had no colonies, plate one seen in Figure 12. The one large colony was picked and set for sequencing like the previous attempt, sequencing results showed failed ligation again. Numerous attempts were made with all ligation attempts to grow colonies on plates, with little to no success. If colonies appeared, the sequence analysis showed no insert in the plasmid. After many unsuccessful attempts, it was determined that outside help was needed with the ligation step. One hundred microliters of 10 ng/µL DNA insert, and 5 µg of pCW plasmid was sent to Applied Biological Materials Inc. (ABM, Inc.) who offered subcloning services. They sent back a full plasmid with gene inserted.

Figure 12. LB-amp plate with a single large colony.

With a full expression vector, work on site-directed mutagenesis could start. Four mutants were chosen, choosing no mutations in the transmembrane helix. Figure 13 shows a map of all known mutations assumed to cause BCD (Kelly et al., 2011).

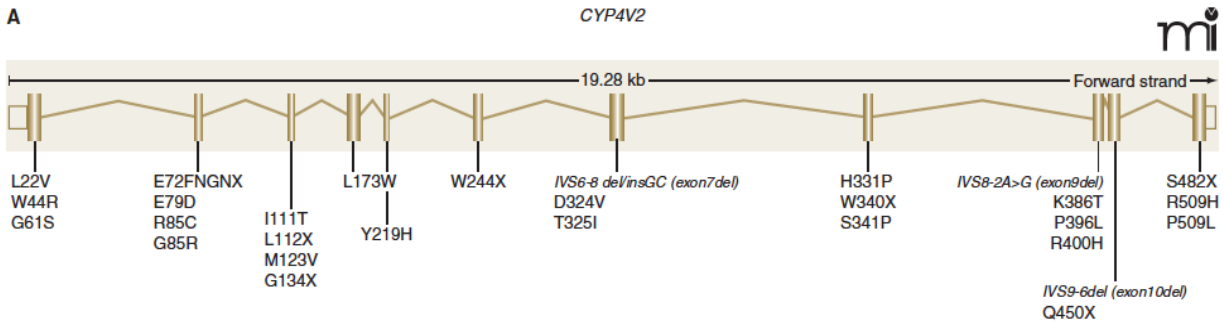

Figure 13. Mutation sites along the CYP4V2 amino-acid chain believed to cause BCD. (Kelly et al., 2011)

Mutations I111T, T325I, and H331P were all chosen from this map. Mutation F73L was chosen from a paper detailing a research group studying 14 different Chinese families with BCD (Yin et al., 2004). The first experiment with mutagenesis mutants F and I were chosen. All necessary PCR reactants were added to the tube and run in the thermocycler. Product was digested with just Dpn1, transformed into DH5α cells, and plated. The next day, there were no cells on the plates. The next attempt, the same two mutants were used, using the Phusion High-Fidelity PCR Master Mix. Using the new master mix, the plated cells had colonies, Figure 14.

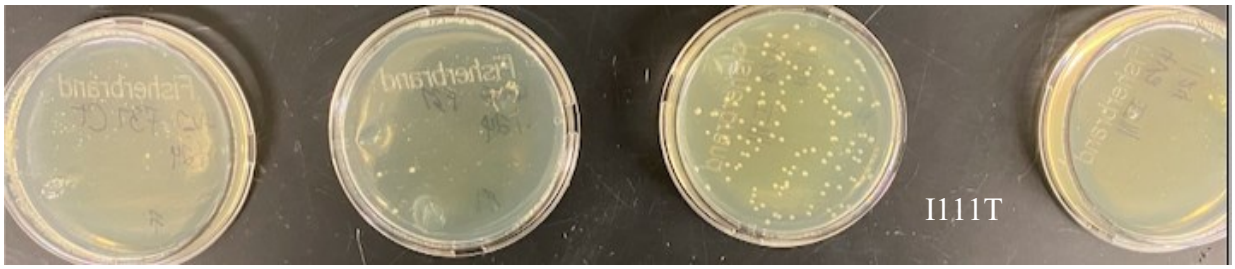

Figure 14. Mutations F73L and I111T LB-amp plates with colonies present.

When grown and DNA sent off for sequencing, it was confirmed that no mutation was present in the plasmid. To remove any parent plasmid in the PCR tubes, they were digested a second time with DpnI, and new cells were transformed. The plates with I111T mutations had

colonies. Sequencing confirmed there was no mutation present. The next two attempts at mutagenesis attempted the T325I and H331P mutations using the same master mix and methods. No colonies were observed on the plates.

For subsequent PCR attempts, only the forward primer was added to the tube. The thermocycler was allowed to run for a few cycles, paused, and the reverse primer was added to the tube, and the cycle was allowed to finish. When transformed and plated, there were a few colonies for each mutation. When sequenced, there were mutations, insertions and deletions, but none were of the desired mutation, so the product was discarded. Upon inspection of the primers, it was discovered that there was an incorrect base pairing. The reverse primers of T325I and H331P were redesigned to correct the pairings. When the new primers were diluted, all four mutations were run using Phusion PCR kit. No colonies were present on the plates.

Instead of running all four mutants at the same time, each mutant was run individually with their own T_m. Before plating, the concentration of DNA in the tube was verified using a nanodrop. All concentrations were greater than 200 ng/μL indicating PCR was working. After each mutant was cycled and plated, only the H331P mutation had any colonies. The other mutations were replated using a longer outgrowth step for the transformation. Except for the T325I mutation, all plates had colonies. When sent off for sequencing, a different company, MCLab, was used, and DNA concentration was too low for sequencing. Using the DNA from the previous plate, new cells were transformed and plated. Colonies were picked, and DNA was extracted. The product was concentrated using the DNA Clean & Concentrator™-5 kit, and concentrations were still too low for sequencing. One more attempt at mutagenesis was made with two tubes set up. One tube had exclusively the forward primer, and the other had the reverse primer. After the first cycle, the two tubes were combined, and heated to denaturing

temperatures. The tubes were cooled in a stepwise manner to promote reannealing. When plated, no colonies were present. Site-directed mutagenesis was taken over by an undergraduate research student as their main project.

Section 2.8 Discussion and Conclusions

Cloning was not an easy step to overcome in this project. Many attempts were made to clone the 4V2 gene into the pCW plasmid with little success. To overcome this, ABM Inc. was contacted to help with the subcloning, and they were able to produce a full vector. There were similar issues with the subcloning step int pCW with another CYP project as well, so it is suspected that the issue had to do with the pCW vector. After successful subcloning, the project was able to move forward. Expression of the wild-type enzyme and site-directed mutagenesis were the next steps.

Mutagenesis has proven difficult. Many different methods of PCR have been attempted and the outcomes are usually no colonies or few colonies. If there are colonies, then the sequencing results show that reactions failed to produce the mutation. The next step in the site-directed mutagenesis is to use a Quickchange site-directed mutagenesis kit that was designed specifically for this purpose.

CHAPTER 3 – PROTEIN EXPRESSION AND PURIFICATION

In protein chemistry, chemists are not working with known substances in a beaker to synthesize novel compounds for characterization. Protein chemists usually use host cells, (bacterial, insect, or mammalian) for protein synthesis. A vector is designed with the gene of interest and this is then transformed into the desired host cell. These cells are then grown in large batches of media. This provides a large sample from which protein can then be harvested. With the use of host cells, many other biomolecules are introduced into the sample. Cells synthesize many other biomolecules and extracting just the molecule of interest is near impossible, and purification methods are needed to isolate them. After isolation, chemists need additional techniques to characterize the isolated material for qualitative and quantitative determinations (Rosano and Ceccarelli, 2014; Manz et al., 2015).

Using classical analytical chemistry techniques such as gas chromatography (GC), liquid chromatography (LC), nuclear magnetic resonance (NMR), mass spectrometry (MS), infrared spectroscopy (IR), UV-visible spectroscopy (UV-Vis), and X-ray crystallography pose challenges for biomolecular chemists. GC usually results in sample analytes destroyed and requires the sample to be volatile for detection (Manz et al., 2015). Destroying research product is non-ideal on the basis that acquiring quantifiable yields can be difficult, so preservation of sample is preferred. IR and UV-Vis are insufficient in that biomolecules have large amounts of functional groups and chromophores that give unambiguous identification near impossible. NMR works by detecting spin active nuclei and for large biomolecules there can be many spin active nuclei (Manz et al., 2015). For these techniques to be utilized, it is important to isolate the particular protein of interest.

For biochemists to use classical analytical methods modifications must be made. UV-Vis is viable for molecules of interest that have very few chromophores that give signal in the visible spectrum and in the area of protein. There have been many advances have been made to now have NMR as a useful method be useful in determining the structure and molecular dynamics of many macromolecules (Cavalli et al., 2007).

Section 3.1 Membrane Protein Expression and Extraction

Section 3.1.1 Protein Expression. Recombinant protein methods are used for human proteins due to lack of abundance in their native environments. To obtain a sizable amount of protein, methods in recombinant expression are used. The first process that must be complete for expression is the choice in host cells. *Escherichia coli* (*E. coli*) is the most common method for the expression of heterologous proteins (Rosano and Ceccarelli, 2014). There are several advantages to using *E. coli* as a host cell. The growth of the cells is quick with a doubling time somewhere in the range of 20-30 minutes (Rosano and Ceccarelli, 2014). Another benefit is the quick and easy methods to transform the cells with the foreign DNA. Plasmid transformations can be as quick as 30 minutes. *E. coli* cells also have a high cell density, so a lot of cells can be grown with the protein of interest in a minimal amount of media (Rosano and Ceccarelli, 2014). All of this combined makes the use of *E. coli* seem to be the cure all for protein expression but there are some problems associated with their use, especially for membrane proteins (Rosano and Ceccarelli, 2014; Delar et al., 2015).

For the expression of eukaryotic membrane proteins there are circumstances that make *E. coli* a less than suitable host (Rosano and Ceccarelli, 2014). If the protein in question has any post-translational modifications, the host cell may not be able to reproduce the modifications

because of a lack of modification capabilities. Translational capabilities are also variable from host organisms. What could be a common codon encoding an amino acid in human cells could be a rare codon in the chosen organism. These rare codons can have a serious impact on the expression levels (Delar et al., 2015). After membrane proteins have been expressed, these proteins must then be inserted into the membrane of the cell. There are differences in the lipid bilayer and transport mechanism that could lead to the protein aggregating instead of being enzymatically active (Delar et al., 2015). Further, expressed membrane proteins can be toxic to their host cells. With most membrane proteins acting as transport systems and membrane channels, the introduction of increased pores and channels can change the influx and efflux of nutrients and other intracellular chemistry of the cells (Delar et al., 2015). This can disrupt metabolic processes leading to detrimental effects. To combat all of these factors, various strains of *E. coli* have been developed to reduce the negative effects and give suitable expression yields (Rosano and Ceccarelli, 2014; Delar et al., 2015).

Section 3.1.2 Protein Extraction. Once expressed in a new host cell, membrane proteins must be extracted. With cell soluble proteins, this is a simple step in which the cells are lysed open and the protein is extracted through various purification methods. Membrane proteins are embedded in the membrane, and if the cells are lysed, then the proteins of interest would stay associated with the cell membrane. These must be extracted from the cell membrane using detergents that provide similar conditions to those found in the bilayers. Using detergents, researchers can mimic the environment of the lipid bilayer, preserving protein structure and functionality. Figure 15 is a cartoon depiction of a protein embedded in the lipid bilayer, and protein enveloped in detergent. Most detergents have a similar structure with two main groups, a hydrophobic tail and a hydrophilic head group. This gives the detergent molecule an amphipathic

nature which allows it to interact with both the solvent and the nonpolar surfaces of the protein. In solution, detergent monomers surround the protein associating with each other and binding to the protein surface at concentrations equal to the detergents critical micelle concentration (CMC) (Delar et al., 2015).

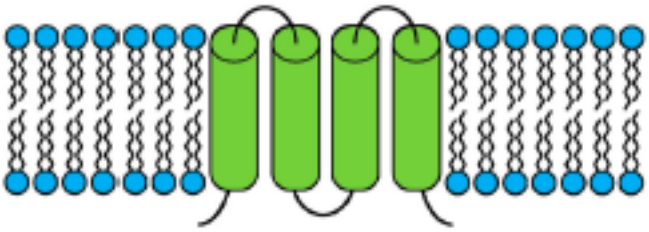

Figure 1b

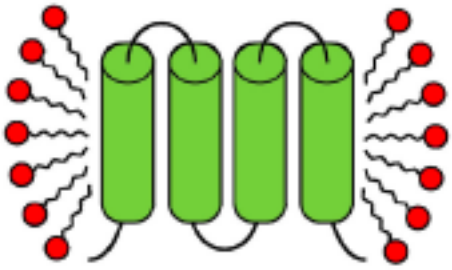

Figure 15. Cartoon depiction of membrane-bound protein, and membrane protein isolated with detergent molecules (Delar et al., 2015).

Section 3.2 Chromatography

Section 3.2.1 Principles of Chromatography. Chromatography in the simplest of forms has an analyte molecule in the mobile phase interacting with the stationary phase. Separation takes place because of the different interactions between the mobile phase and a stationary phase.

Depending on the extent of interactions between the mobile and stationary phase, molecules will elute from the stationary at different times leading to separation of molecules (Manz et al., 2015).

Analyzing samples by chromatography is most commonly used for the separation, isolation, and purification of biomolecules. When using cell host cells for the procurement of usable quantities of protein, other proteins and cell components are dispersed with analyte of interest. These extra components must be separated from the sample and can be done so through many chromatographic methods. The two discussed here will be affinity chromatography and ion-exchange chromatography (Manz et al., 2015).

Section 3.2.2 Affinity Chromatography. Affinity chromatography works on the principle of high-specificity interactions. Recombinant proteins usually have a poly-histidine tag that interacts with a metal ion, cobalt or nickel, incorporated during gene design. There are four general stages that affinity chromatography can be separated into: 1. Sample introduction, 2. Adsorption, 3. Washing, and 4. Desorption. In stage 1, impure sample is loaded onto the column that has been packed with a solid phase of metal ions, antibodies, or protein specific ligands. In stage 2, biomolecules of interest that have an affinity for the solid phase are adsorbed and the impurities that do not have affinity flow through the matrix. Stage 3 has the bound sample washed with different buffers that elute off the weakly bound impurities. Stage 4, desorption, is achieved through various means. pH increase, salt content increase, of introducing a ligand with higher specificity. The introduction of an elution factor competes for binding to the solid phase causing bound analyte to elute off the column. For nickel-affinity chromatography, the elution factor is typically imidazole or histidine.

Section 3.2.3 Ion-Exchange Chromatography. Ion-exchange chromatography separates using a difference in charge between the mobile phase with target analyte and a stationary phase

of porous, functionalized resin beads. There are two main types of ion exchangers: anionic exchangers and cationic exchangers. In anion exchange, the resin beads are coated in diethyl aminoethyl (DEAE) functional groups which are positively charged. Cationic exchangers are coated in carboxy methyl (CM) functional groups that are negatively charge. Choosing a resin depends on the pH of the mobile phase used for separation. Using a mobile phase with a high pH will cause more deprotonation of target analytes causing a negatively charged species. In this case, the resin of choice would be the DEAE beads because the negatively charged molecules would bind in the pores of the positively charged resin beads. The opposite holds true for positively charged molecules and CM beads. Choosing an appropriate mobile phase is important, as it should be noted that the loading capacity is dependent on the charge of the resin. Choosing a mobile phase that has too high or low of a pH can be detrimental to the resin. For the case of DEAE beads, choosing a mobile phase with a pH greater than 8 would cause the amino group to be deprotonated and the beads would no longer bind negatively charged molecules. (Manz et al., 2015).

Proteins are amphoteric in nature with both positive and negatively charged groups. The combination of all these groups charges gives the protein its overall charge. At some point along the pH spectrum, a protein is completely neutral, this is called the protein's isoelectric point (pI). When operating at pH values close to a proteins pI, there is minimal absorption to the beads. Increasing or decreasing the pH by one or two values would vastly increase the charge a protein exhibits, increasing binding greatly. Buffer concentrations should be kept as low as possible to avoid competitive binding (Manz et al., 2015).

As in affinity chromatography, there are four main stages to ion-exchange: 1. Sample introduction, 2. Adsorption, 3. Washing, and 4. Desorption. Desorption of analyte molecules is

done through increasing the salt concentration or changing the pH of the buffer. Increasing the salt concentration, increases the number of ions in solution. The increase in ions creates new competitive binding for the beads. Weaker bound proteins (protein misfolding or incomplete proteins) elute first because of the difference in net charge from correctly folded proteins. The other method of a pH gradient works in a similar fashion. Changing the pH causes the net charge of the protein to change and decrease the interactions with the resin, or it neutralizes the resin also decreasing the interactions eluting the protein (Manz et al., 2015).

Section 3.3 Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The main principle driving the separation in SDS-PAGE is protein size. The introduction of the SDS coats it in an evenly distributed negative charge, eliminating the intrinsic charges of proteins, giving all proteins the same mobility. Another additive in SDS-PAGE is β-mercaptoethanol (2-mercaptoethanol), which is a thiol-reducing agent. When introduced, disulfide bonds are broken, eliminating the tertiary structure of the proteins. Similar to standard gel electrophoresis, a molecular weight ladder runs in tandem that gives a standard set of molecular weights for visual verification. After the chamber finishes running, the entire gel must be stained to visualize any banding that may appear. A staining solution of *Coomassie brilliant blue* is usually poured over the entire gel, heated, and allowed to sit and stain .

Section 3.4 Carbon Monoxide Difference Assay

One method used to determine not only the quantity of a cytochrome P450 protein but also the quality of protein in a purified cytochrome P450 sample or cellular matrix is the use of a carbon monoxide (CO) difference assay. The main principle behind the CO difference assay is

the binding of CO to the heme of hemeproteins causes a wavelength maximum at 450 nm (Soret peak) in the energy spectrum. This assay is mainly used in the characterization of recombinant P450 proteins as this characteristic shift is how they earned their name. It also shows the integrity of the protein over time. As protein degrades in solution, a shift in the Soret peak from 450 nm to 420 nm indicates protein integrity decreasing. One of the problems with this though is that it does not distinguish between hemeproteins in solution. Should there be multiple cytochrome P450s, no difference would be seen (Guengerich *et al.*, 2009).

Section 3.5 Methodology

Section 3.5.1 Media Preparation. To prepare a 100 mL solution of LB media, into a medium-sized beaker, 1 gram of Tryptone, 0.5 g yeast extract, and 1 g of sodium chloride (NaCl) were added. The volume was brought to 100 mL with water. The top of the beaker was sealed with aluminum foil and a piece of autoclave tape. The beaker was placed into the autoclave and a liquid cycle was run. Once the cycle completed, the beaker contents were allowed to cool to room temperature before storing in cold storage for later use.

To prepare TB media, to a 2 L flask, 50.4 g of prepared Terrific Broth powder was supplemented with 4 mL of 100% glycerol. The volume was brought to 1000 mL, sealed with aluminum foil. The flask was run through a liquid autoclave cycle and allowed to cool to room temperature before placing in cold storage for later use.

Section 3.5.2 Buffer Preparation. All purification buffers started with a basic Tris buffer of pH 8.0. This buffer was prepared in a large storage container where 60.57 g of Trizma® and 450 mL of water were added. The pH of the solution was brought to ~8.0 and then diluted to 500 mL. This was used to prepare all other buffers. All purification buffers from the Tris buffer

are pH adjusted to 8.2 after being prepared. In Table 5, the purification buffer components are listed.

Table 5. Reagents Needed for Purification Buffers.

Buffer Name	Tris Buffer (mL)	Glycerol (mL)	Water (mL)	NaCl (g)	Imidazole (g)
Resuspension Buffer	250	50	200	8.77	N/A
Ni-NTA Wash Buffer	50	50	400	8.77	0.34
Ni-NTA Elution Buffer	50	50	400	1.46	6.8
Ion Exchange Wash Buffer	25	50	425	N/A	N/A
Ion Exchange Elution Buffer	50	50	400	14.61	N/A

Section 3.5.3 Cell Transformation. To start the transformation, the plasmid and DNA need to be ligated. To a PCR tube, 2.5 µL of Anza T4 DNA Ligase Mastermix, 6 µL of DNA, and 2 µL of empty plasmid was added. This was allowed to incubate on the benchtop for at least an hour, with an overnight incubation preferred. To 50 µL of DH5α competent cells, 5 µL of ligation mix was added. This new mixture was allowed to incubate on ice for thirty minutes. After the incubation period, the cells were heat shocked at 37°C for thirty to forty seconds in a hot water bath. After heat shock, the cells can be chilled back on ice for two minutes, or 500 µL of LB media with no antibiotic can be added. If placing on ice, incubate for two minutes and then plate onto a prepared LB-Amp plate. Glass beads are added, and the plate is swirled around to deliver a thin layer of cells. The plate was placed inside the incubator overnight. If doing an outgrowth step, after adding the 500 µL of LB media, the cells are allowed to incubate at 37°C in the incubator for at least an hour. After incubation, the cells were centrifuged down. All but 50 µL of solution is removed from the Eppendorf tube and the cells are resuspended in the remaining media. The resuspended cells are plated in the above method.

Section 3.5.4 Expression of CYP4V2. After sequence verification, 1 µL of plasmid from Plasmid Preparation was transformed into the JM109 cell line. The next day, a single colony of transformed cells were picked and grown in 100 µL of LB media with ampicillin. This starter culture of cells is distributed between 2 L of prepared TB media with ampicillin. The cells were grown shaking at 37℃. After an hour and a half, the OD_{600} is measured. Cells need to grow to at minimum 0.4 and at maximum 1.4 AU before induction. This allows to enter peak logarithmic growth. To induce, 0.22 g/L of β-D-1-thiogalactopyranoside (IPTG) and 0.22 g/L aminolevulenic acid (ALA) was added to each liter of cells. The temperature of the incubator was reduced to 32℃ and the cells were set to shake for three days.

Once cells have been expressed for three days, distribute the TB media between four 400 mL centrifuge tubes. Centrifuge the cells at 5000 rpm for 15 minutes. Decant off the upper liquid layer and repeat until all growth media has been centrifuged. Resuspend the pellet with minimal resuspension buffer and either decant into a 50 mL Falcon tube for cold storage or proceed onto protein extraction.

Section 3.5.5 Protein Extraction. Place resuspended cells in an appropriately sized beaker and place into a bed of ice. Take the cells on ice to a probe sonicator and position the tip of the probe near the bottom of the beaker, but not resting on the bottom. Sonicate the cells for thirty seconds, and let them rest for thirty seconds, alternating between the two for at least three minutes. Cells should take on a darker color. Remove the probe, place a stirring rod into the beaker, and place into a shallow glass dish on a magnetic stirrer surrounded by ice. While stirring, add in enough Cymal-5 and Triton X-100 to reach the CMC for both. Example calculation for a dry detergent (Cymal-5) shown in the equation below.

3.2 mM x ____ mL sample x 494.5 g/mol = _____/1000/1000 = _____ g cymal-5

Allow the mixture to stir for an hour before splitting evenly into small centrifuge tube. Place the tubes into the centrifuge and spin at 20,000 rpm for fifteen minutes. Remove and save the upper liquid layer. This contains the now solubilized protein ready for purification.

Section 3.5.6 Protein Purification. Using the BioRad Biologic LP chromatography instrument, connect a purification column with Ni-NTA beads to the delivery line. Flush all lines with water. In order, connect lines A, C, D, and E to the lysate, 50 mL of resuspension buffer, 35 mL of wash buffer, and 35 mL of elution buffer. Using above equation, add the appropriate amount of Cymal-5. Triton X-100 is not added. Using a small pipette, add resuspension buffer to the column to equilibrate the beads. Using the "4V2" method in the instrument, begin the purification, and press run on the fraction collector. Allow the instrument to run through lysate and all the buffer. Once finished, either visually inspect the collected fractions, or using the LP Data View™ MS Software, chose the fractions that eluted off in the latest buffer. Remove the fractions and purge the column and lines with DI water.

Dilute these fractions 10x, if using fractions from the wash buffer, or 5x, if using fractions from the elution buffer. Dilution is done with ion-exchange wash buffer. Once diluted, place line C into the diluted protein, and lines A and B into 50 mL of wash buffer and 25 mL of elution buffer, respectively. All buffers, including dilution buffer, should have detergent added to them. Run the "4V2 DEAE" method to begin the purification. Collect the fractions like above.

Section 3.5.7 Obtaining an Absolute Spectrum. After purification, an absolute spectrum is taken to give a general idea of protein purification and to see if there is P450 enzyme. Using 750 µL of elution buffer with detergent added, autozero and baseline the UV-Vis. After baselining, add anywhere from 5-30 µL of protein sample, depending on visual concentration. Scan the sample at a wavelength range of 750-250 nm.

Section 3.5.8 SDS-PAGE. Once the protein sample has been purified, verification of purity is needed. To start, fractions are collected that are believed to have the purest protein. Remove 20 µL of protein fraction and place into a small Eppendorf tube, keep in mind which fraction is which. To the Eppendorf tubes, add in 20 µL of loading dye with β-mercaptoethanol. Prepare an Eppendorf tube with a protein ladder in the same manner. The gels come with a comb and a piece of tape that both must be removed. Remove the tape located at the bottom of the gel before applying voltage to the chamber. The comb is inserted in the top and must be removed gently as to not disturb the well shapes. Remove the gel runner from the container and place the gel into the runner, ensuring a secure seal is made with the clamps. Place the gel runner into the container and fill with 1x SDS-PAGE running buffer making sure the electrical contacts are covered with buffer. Load the protein ladder into the first well, and the analytes in subsequent wells. Set the voltage to 210 and allow to run. Make sure there are visible bubbles coming from the bottom of the gel. Allow the gel to run until the visible bands reach the bottom of the gel. Once complete, remove the gel runner, unclamp the gel and remove. Crack the four corners of the plastic housing, exposing the gel. Remove the gel with a little bit of water and place into a container for staining.

To the plastic container, add in an arbitrary amount of Coomassie staining solution to cover the gel. Place into a microwave and run for 20-30 seconds to allow the gel to stain. Pour the excess stain back into the bottle. Cover the gel with water, and microwave again for two minutes. Repeat the process of heating and pouring off excess liquid until the bands on the gel can be visualized. Gel may need to be moved to a light box to help visualize the banding. Qualitative protein purity can be determined by the number of bands present in each lane.

Section 3.5.9 CO Difference Assay. To determine if there is active P450 protein in the purified fractions, a CO Difference Assay is used. Place 750 µL of protein fraction into a cuvette with a few crystals of sodium dithionite, place into the UV-Vis instrument and autozero, then baseline with the sample. Remove the cuvette and add in 50 bubbles of CO. Place the cuvette back in the instrument, autozero, and scan over a range of 500-350 nm. Repeat the scan every five minutes until an increase in absorbance is no longer seen.

Section 3.6 Results

With the plasmid from ABM, Inc. JM109 cells were transformed with the DNA and plated. For the first expression, a colony was picked from the plate and a starter culture was grown. The culture was split between 2 L of TB media. OD_{600} was checked over the course of a few hours waiting for an absorption value of 0.4. At $OD_{600} = 0.156$ the cells were induced. This was too early of an induction; cells were not in the appropriate stage of growth. After the three-day growth the cells were pelleted and resuspended in buffer, with a pH of 7.4. The cells had the appearance of a light coffee color. After detergent extraction, the lysate appeared a bright orange color. Once added to the column, there was a dark red band present at the top when protein loading finished. Once the resuspension buffer started to run, the protein started to crash out. After the resuspension, wash, and elution run, white, flocculant material was present indicating crashed protein. Figure 16 shows the pelleted cells, lysate, and banding on the column.

For the second purification, the cells were allowed to grow to an $OD_{600} > 0.8$. The pelleted cells had a much darker appearance. There was a darker brown banding around the outside of the pellet indicating more present protein in the cells. During the resuspension step of

purification, all protein eluted off the column. Unlike previous preparations, the third preparation bound throughout the entire column. Similarly, protein eluted off in the wash buffer.

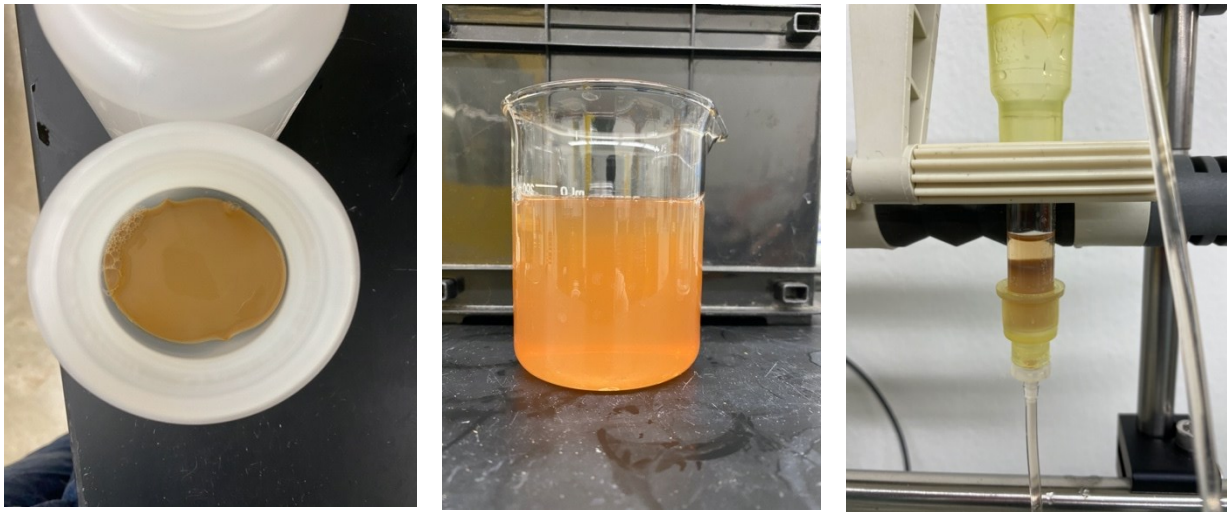

Figure 16. Pictures from the purification process. From left to right, pelleted cells after centrifugation, lysate decanted after detergent extraction, banding present on the affinity column after protein loading.

This wash elution was moved to the ion-exchange column for further purification. For the column matrix, a CM column was used. After a 10x dilution, no binding was present on the column. Using Expasy Protparam, it was discovered that the pI point was 7.26. Using buffers at pH 7.4 not only was not conducive to CM media, but also being at the pI point was causing protein to aggregate. For further purifications, the pH of buffers was adjusted to 8.2, and ion-exchange media was changed to DEAE.

Present in all three preparations, after the final centrifugation, there was a dark band present at the top of the pelleted cells. Protein was still present in the cells and needed to be extracted. Preparation four addressed the extraction issue by adding in Triton X-100. When added alongside Cymal-5, the banding at the top of the cells is less prevalent and the color difference in lysate is drastic, Figure 17.

Figure 17. Differences in lysate color before and after addition of Triton X-100

This was purified using affinity and ion-exchange chromatography. Banding was present on the DEAE beads, and protein eluted off in the wash buffer. Using a portion of saved fractions from the affinity and ion-exchange purifications, an SDS-PAGE was run to give a qualitative purity, Figure 18. Lane 1 is the protein ladder, lanes 2 and 3 are fractions from the affinity column, and lanes 4, 5, and 6 are from the ion-exchange column.

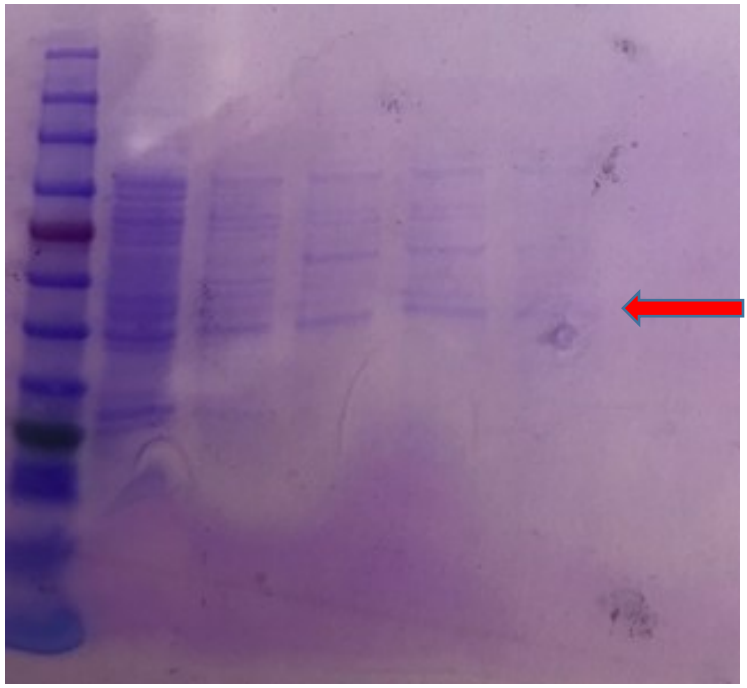

Figure 18. SDS-PAGE gel showing the change in purity from affinity chromatography to ion-exchange chromatography. Red arrow points to band representing CYP4V2.

The purity from affinity to ion-exchange is noticeable, only three main bands are present in the later lanes with multiple bands present in the first two protein lanes. For a more quantitative approach to determining purity, an absolute spectrum was obtained from the ion-exchange fraction. Figure 19 shows the spectrum obtained. The fractions from the ion-exchange column were concentrated down to ~500 µL, and the concentration was determined to be 306 µM or 17.7 mg/mL. This protein was used to set up a crystal tray. Before the concentrated protein could be used further, it lost all activity. Possibly because of concentrating all sample together leading to aggregation. Preparation number five was purified all the way through ion-exchange. A decreased yield of 0.87 mg/mL was obtained from this purification. At this point, the glycerol stock prepared from the first purification was a few months old and with the next starter culture, a fraction was saved to create a new glycerol stock.

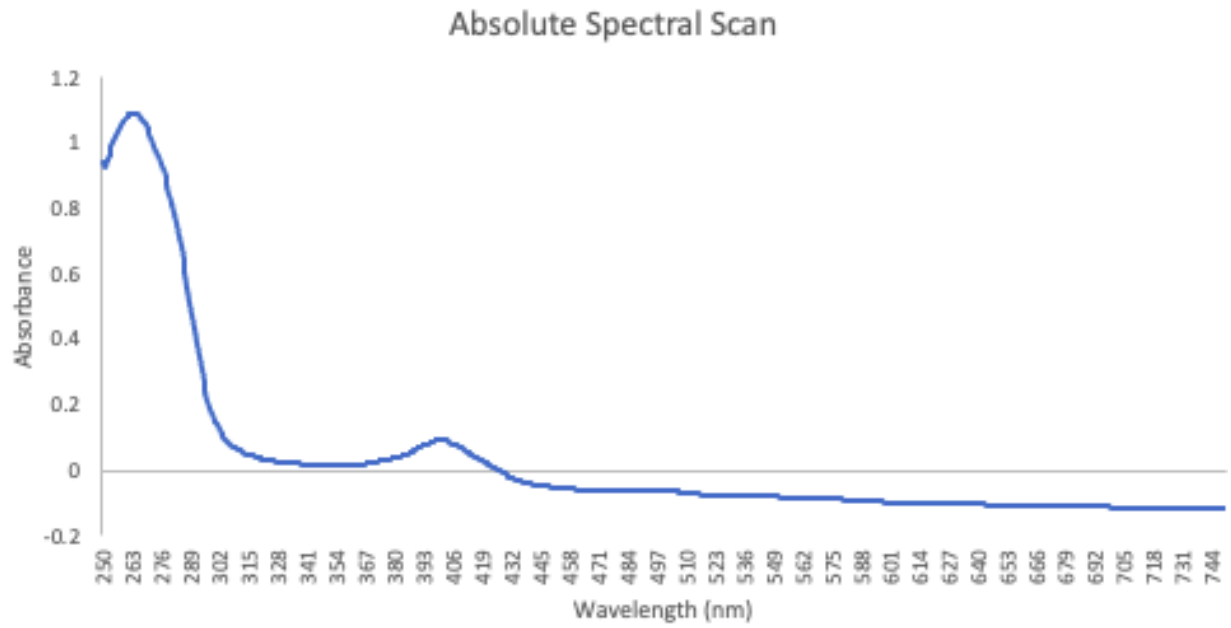

Figure 19. UV-Vis spectrum of protein in solution from ion-exchange chromatography. Soret peak at 400 nm.

With each previous purification, the nickel was being stripped from the beads. The bright blue color was slowly dulling, and binding was become more diffuse through the column. The idea was had to change from a highly basic Tris buffer system to a more acidic phosphate system. Buffer recipes now had 500 mM phosphate and a pH adjusted to 6.5. When purifying a new preparation, pellets were resuspended in the new buffer system. When loading onto the column, no binding was present. The phosphate system was abandoned, and resuspension was switched back to Tris. Since the nickel beads were losing their color, they were recharged using nickel sulfate. Before the decision to switch back to Tris was made, a preparation was already resuspended in phosphate buffer. During the affinity purification, the column was equilibrated with pH 8.2 Tris buffer, and the pH 6.5 resuspended protein was loaded. There was very diffuse binding on the column, when the pH 8.2 resuspension buffer started loading, protein crashed on top of the column. Letting the method run through elution proved useful in that two 2 mL fractions of protein eluted in the elution buffer. These two fractions were purified further, and a protein concentration of 86 μM, or 4.98 mg/mL. A full spectrum graph was obtained, and protein was stored in the fridge for stability studies. The next day, to assess longevity, another full spectrum scan was performed, and protein concentration dropped, Figure 20. This is a sign of stability issues. One more preparation was expressed and purified and again an overnight difference in CO activity showed that there was a decrease in activity.

Section 3.7 Discussion and Conclusions

The first expression with the early induction time was a human error and could have easily been prevented. All subsequent expressions were induced at $0.4 \leq OD_{600} \leq 1.4$. This changed the pellets from just coffee colored to have a dark ring around the edges and dappled

with the darker portions. Purification was the main struggle in this project. Many times, the protein did not bind to the column, or if it did bind it eluted in the resuspension buffer. Except for one case, protein banding was never present. All bindings were very diffuse and weak. To try and increase binding the pH of the buffer was changed. This seemed to allow protein binding to strengthen, and protein stopped crashing on top of the column. The downside being that the nickel column was stripping. Only a few purifications could be done before the column needed to be recharged. To try and offset the stripping, the pH was changed to an acidic buffer made from phosphate. This decreased stripping but again caused protein binding to weaken significantly.

The method that worked best for purification was using the Tris base with a pH of 8.2. The column will have a life of 3-4 purifications before it needs to be recharged due to stripping. After the purification process, the work time is short due to stability issues. Any subsequent testing needs to be done quickly. Changes that I would suggest helping with stability issues would be to increase the glycerol concentration in the ion-exchange elution buffer. This may lead to protein stability.

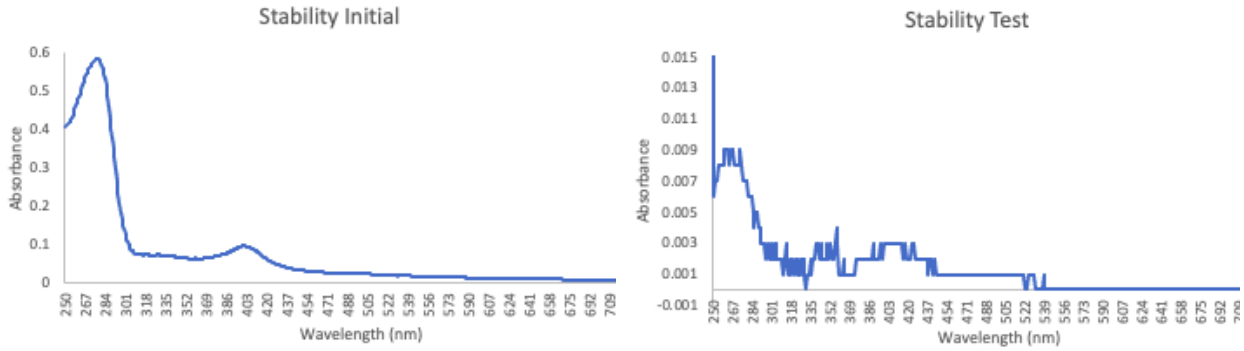

Figure 20. Contrasting absolute spectra showing a decrease in protein stability.

CHAPTER 4 – X-RAY CRYSTALLOGRAPHY

Section 4.1 Membrane Protein Crystallography

In 2015, the Protein Data Bank had close to 100,000 protein structures logged. Of the 100,000, approximately 1% of the structures were those of membrane proteins (Delar et al., 2015). Although the number of solved membrane proteins is small, about 30% of genes encode for membrane proteins (Delar et al., 2015). The main problems associated with membrane protein crystallography are: 1. acquiring sufficient quantities of purified protein and 2. obtaining high-quality crystals for diffraction (Delar et al., 2015). The first problem was discussed in Chapter 3 Section 3.1.1 and will not be discussed here.

In the case of soluble protein samples, proteins not embedded in the membrane, the dominating factor of crystallization are the protein-protein interactions. In the case of membrane proteins, the use of detergent in the extraction introduces protein-detergent and detergent-detergent interactions, while still having protein-protein interactions (Delar et al., 2015). Delar et al. believe that the more effort that is placed in finding a suitable, mild detergent is more important than optimizing other parameters of crystallography. In its most simple form, a crystallization experiment involves a solution of precipitant is mixed with a solution containing a purified protein-detergent complex. The precipitant solution is a mix consisting of a precipitant, salt, and buffer (Delar et al., 2015). The combination of the two solutions causes the drop to enter a state of supersaturation. Once supersaturated, the complex can either aggregate and fall out of solution or crystallize. The most common, and successful method for obtaining high quality crystals is the vapor diffusion method (Delar et al., 2015).

As stated above the precipitant solution is a mix of three different components to aid in the crystallization process: precipitant, salt, and buffer. There are two categories that precipitants are divided into: 1. organic solvents, and 2. long-chain polymers (Delar et al., 2015). Organic precipitants reduce the electrostatic interactions leading to a decrease in ionic compound solubility (Delar et al., 2015). When the concentration of organic precipitants becomes too high, they can denature proteins. The second precipitants are a group of long-chain polymers, with the most common being poly-ethylene glycols (PEGs). Precipitants using PEGs are acting through the volume-exclusion effect (Delar et al., 2015). The second component of precipitant solutions, salts act by modifying the solubility of proteins. While the mechanism of action is unknown, it is speculated that salts cause interactions between the ions, the protein structure, and water. The third component, buffers, changes the pH of the solution which then affects the electrostatic interactions between the protein-detergent complexes (Delar et al., 2015). At a pH equal to the pI of the protein, solubility decreases. This creates a harsh environment for crystallization, while pH values close to the pI has shown to be optimal conditions for crystallization (Delar et al., 2015).

In the vapor diffusion method, small quantities of purified protein samples are mixed in equal proportions with a crystallization solution forming a drop. This drop is now half protein, half crystallization solution. This drop is then either placed in a well, sitting-drop method, or attached above the well, hanging-drop method, and the chamber is then sealed from the outside environment, Figure 21. Volatile components in the drop will evaporate into the chamber and diffuse into the well. This creates a system at equilibrium. At the state of equilibrium, there are different phases the drop can manifest. Crystallization with detergents can crystallize protein, aggregate compounds into micelles, or it can separate into regions being detergent rich or poor. Today, there are readily available crystal screening kits that have preset precipitant solutions that

have proven to crystallize complexes, and act as a starting point for any crystallographer. Once an initial hit has been discovered, the conditions of the well can be modified in small increments to optimize crystal growth, producing large crystals with few imperfections. Multiple rounds of optimization may be performed to find the best conditions. (Delar et al., 2015).

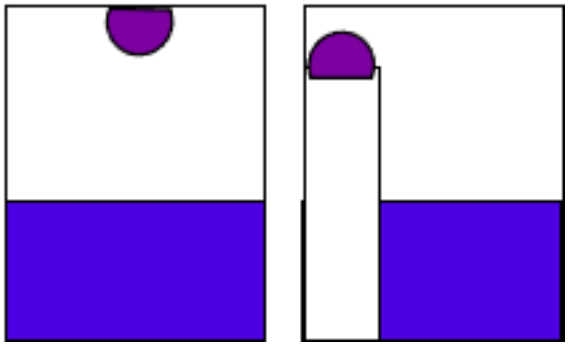

Figure 21. Hanging-drop (left) and Sitting-drop (right) method for crystallization.

Section 4.2 Methodology

Section 4.2.1 Concentrating Protein. Fractions with the purest protein content were concentrated for binding assays and crystallographic setups. The fractions were selected that were determined to be the purest from the SDS-PAGE gel. The fractions were transferred into an Amicon Ultra filter tube with a 50 KDa cuttoff. The tube was centrifuged the tube at 5000 rpm for twenty minutes. Roughly 1 mL of flowthrough was kept and the rest discarded. Repeat the concentration with other fractions purified, saving some fractions to not concentrate. This process was repeated with other fractions purified. Once fractions were concentrated to ~200 µL, the molar concentration of the protein was determined.

Section 4.2.2 Determining Molar Concentration. Using the UV-Vis, the instrument was baselined with 750 µL of saved flowthrough, scanning at a wavelength range of 500-350 nm. 5 µL of concentrated protein was added to the cuvette, pipetting up and down, gently, a few

times to mix the solution. The instrument was scanned to obtain the absorbance at 400 nm. Using Beer's Law, the concentration of the protein was calculated by rearranging the equation and dividing the absorbance by the product of the extinction coefficient and path length. Absorbance was taken from the UV-Vis scan, and the extinction coefficient is given to be 0.1 $\mu M^{-1}cm^{-1}$ if using the absolute spectrum peak for the heme. This provides the concentration after dilution. To determine the protein concentration before dilution, multiply the concentration by the dilution factor. To determine dilution factor, divide the total cuvette volume by the volume of protein sample added.

Section 4.2.3 Crystallographic Setups. Hampton Research Crystal Screen 1 and 2 was used to set up the 96 reagents into a reservoir tray. 75 µL of each reagent was pipetted into a crystal screen tray, making sure to keep the same positioning of reagents. Then, 0.5 µL of purified protein sample was pipetted into the sample well, and then 0.5 µL of each reagent was placed into the sample well. The screening tray was covered tightly and placed into a 19℃ incubator. Each sample well was checked every few days and the appearance of the sample well was described.

A crystal hit can be optimized from the initial conditions using the optimization tool on the Hampton Research website. The optimization tool provides all relevant volumes and concentrations for a 24-well crystal tray. For the optimized tray, the indicated volumes of reagents into the sample well were pipetted. On a cover slide, 0.5 µL of purified protein and 0.5 µL of well reagent were placed. The edge of the sample well was covered with grease or another sealant and inverted over the well to seal airtight. This tray was also incubated, and progress was checked every few days.

Section 4.3 Results and Conclusions

Using the concentrated mentioned above in section 3.6, an initial crystal screen was performed. Of the 96 wells, only one presented with a crystal hit, well H3. The well reagents were 0.2 M magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, and 3.4 M 1,6-hexanediol. These conditions produced a single crystal, Figure 22. An optimized plate has been designed around these conditions. No recent expression has yielded suitable quantities of protein for preparation.

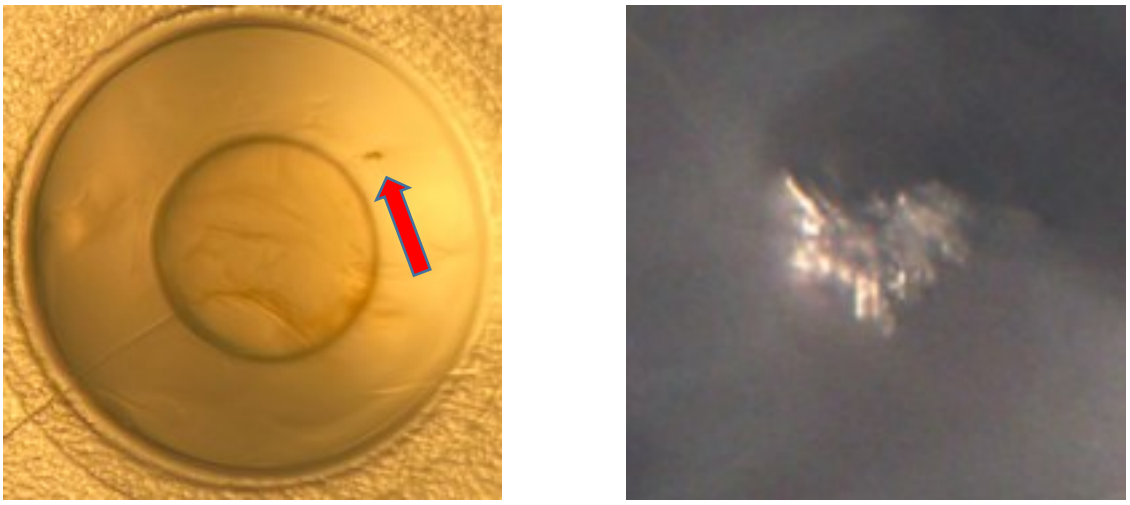

Figure 22. Picture of the crystal well (left) and an up-close picture of the initial crystal hit (right). Red arrow indicating the position of the crystal in the well.

CHAPTER 5 – FUTURE DIRECTIONS

For the future direction of this project, all four mutants need to be mutated through site directed mutagenesis and sequence verified. Once they have been obtained, the expression method for the wild-type enzyme may need to be modified. Modifications may be slight, or a different method may need to be determined.

For the wild-type purification method, fixing the stability issues is the next project. This may be achieved through increasing the glycerol content in the elution buffers or finding a more stable buffer system that can better accommodate the protein's instability. Other detergents beside Cymal-5 and Triton-X-100 may be required. After stabilization, optimization of the crystal screen can happen. The plate has been designed, but protein purity and concentrations need to be obtained before a plate can be made.

REFERENCES

Birge EA (1981) *Bacterial and Bacteriophage Genetics An Introduction* (Starr MP ed), Springer-Verlag New York Inc.

Çağlayan M (2019) Interplay between DNA Polymerases and DNA Ligases: Influence on Substrate Channeling and the Fidelity of DNA Ligation. *J Mol Biol* **341**:2068–2081.

Carter P (1986) Site-Directed Mutagenesis. *Biochem J* **237**:1–7.

Cavalli A, Salvatella X, Dobson CM, and Vendruscolo M (2007) Protein structure determination from NMR chemical shifts. *PNAS* **104**:9615–9620.

Darki F, Fekri S, Farhangmehr S, Ahmadieh H, Dehghan MH, and Elahi E (2019) CYP4V2 mutation screening in an Iranian Bietti crystalline dystrophy pedigree and evidence for clustering of CYP4V2 mutations. *Journal of Current Ophthalmology* **31**:172–179.

Delar JA, Bolla JR, Su C-C, and Yu EW (2015) Crystallization of Membrane Proteins by Vapor Diffusion, in *Methods in Enzymology* pp 363–392, Academic Press.

Edelheit O, Hanukoglu A, and Hanukoglu I (2009) Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnology* **9**.

Edson KZ, and Rettie AE (2013) CYP4 Enzymes as potential drug targets: focus on enzyme multiplicity, inducers and inhibitors, and therapeutic modulation of 20-hydroxyeicosatetraenoic acid (20-HETE) synthase and fatty acid ω-hydroxylase activities. *Current Topics in Medicinal Chemistry* **13**:1429–1440.

El-Gewely MR, Fenton C, Kjeldsen E, and Xu H (n.d.) Mutagenesis: Site-specific. *Encyclopedia of Life Sciences*.

Freifelder D (1987) *Molecular Biology*, Second, Jones and Bartlett Publishers, Inc.

Garcia-Garcia G, Martinez-Rubio M, Moya-Moya M, Perez-Santonja J, and Escribano J (2010) Current perspectives in Bietti crystalline dystrophy. *Clinical Ophthalmology* **13**:1379–1399.

Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC, and Gerlbart WM (1993) *An Introduction to Genetic Analysis*, Fifth, W. H. Freeman and Company.

Guengerich FP, Martin MV, Sohl CD, and Cheng Q (2009) Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. *Nat Protoc* **4**.

Henne KR, Kunze KL, Zheng Y-M, Chistmas P, Soberman RJ, and Rettie AE (2001) Covalent Linkage of Prosthetic Heme to CYP4 Family P450 Enzymes. *Biochemistry* **40**:12925–12931.

Kelly EJ, Nakano M, Priyanka R, Yarov-Yarovoy V, and Rettie AE (2011) Finding Homes for Orphan Cytochromes P450s: CYP4V2 and CYP4F22 in Disease States. *Molecular Interventions* **11**:124–132.

Manz A, Dittrich PS, Pamme N, and Iossifidis D (2015a) *bioanalytical chemistry*, Second, Imperial College Press.

Manz A, Dittrich PS, Pamme N, and Iossifidis D (2015b) *bioanalytical chemistry*, Second, Imperial College Press.

Nakano M, Lockhart CM, Kelly EJ, and Rettie AE (2014) Ocular cytochrome P450s and transporters: roles in disease and endobiotic and xenobiotic disposition. *Drug Metab Rev* **46**:247–260.

Ortiz de Montellano PR (ed) (2015) *Cytochrome P450 Structure, Mechanism, and Biochemistry*, Fourth, Springer International Publishing.

Raucy J, and Allen S (2001) Recent Advances in P450 research. *The Pharmacogenomics Journal* **1**:178–186.

Rodriguez R L, and Tait R C (1983) *Recombinant DNA Techniques: An Introduction*, The Benjamin/Cummings Publishing Company, Menlo Park, CA.

Rosano GL, and Ceccarelli EA (2014) Recombinant protein expression in Escherichia coli: advances and challenges. *frontiers in MICROBIOLOGY* **5**:1–17.

Stellwagen NC (2009) Electrophoresis of DNA in agarose gels, polyacrylamide gels and in free solution. *Electrophoresis* **30**.

Yin H, Jin C, Fang X, Miao Q, Zhao Y, Chen Z, Su Z, Ye P, Wang Y, and Yin J (2004) Molecular Analysis and Phenotypic Study in 14 Chinese Families with Bietti Crystalline Dystrophy. *PloS ONE* **9**:1–10.