10-20-2017

Kinetic Characterization Of Listeria Monocytogenes 2-C-Methyl-D-Erythritol 4-Phosphate Cytidylyltransferase (cms) Enzymes Using High Performance Liquid Chromatography (hplc)

Mark Oblazny
Illinois State University, moblazn@ilstu.edu

KINETIC CHARACTERIZATION OF *LISTERIA MONOCYTOGENES* 2-C-METHYL-D-ERYTHRITOL 4-PHOSPHATE CYTIDYLYLTRANSFERASE (CMS) ENZYMES USING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)

Mark Oblazny

94 Pages

Infectious diseases, with increasing prevalence of antibiotic resistant bacteria, coupled with the declining rate in discovery of antimicrobial agents, impose one of the most significant threats to human health. Here we identify 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (CMS) as a valid target for antibiotic development which is an enzyme in pathogenic organisms that leads to the biosynthesis of isoprene precursor molecules. Isoprene molecules are one of nature's most common building blocks that are vital to many biological metabolic processes and are synthesized via the mevalonic acid dependent (MVA), or methylerythritol phosphate (MEP) pathway. Vertebrates utilize the MVA pathway, while many pathogenic bacteria use the MEP pathway for isoprenoid biosynthesis. With separate pathways expressed between vertebrates and bacteria, the MEP pathway is an attractive target for the development of effective novel antimicrobial therapeutics. This report kinetically characterizes two putative isoforms of CMS from *Listeria monocytogenes*. This study validates a previously developed high performance liquid chromatography (HPLC) method for the kinetic characterization of cytidylyltransferase enzymes. Structural homology and sequence analysis of both putative isoforms of CMS provide new evidence suggesting the lmo0235 gene encodes a

CMS and the lmo1086 gene from *Listeria monocytogenes* encodes a putative ribitol 5-phosphate cytidylyltransferase.

KEYWORDS: CMS, Cytidylyltransferase, HPLC, Isoprenoid, *Listeria*, Antibiotics.

KINETIC CHARACTERIZATION OF *LISTERIA MONOCYTOGENES* 2-C-METHYL-D-ERYTHRITOL 4-PHOSPHATE CYTIDYLYLTRANSFERASE (CMS) ENZYMES USING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)

MARK OBLAZNY

A Thesis Submitted in Partial
Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Chemistry

ILLINOIS STATE UNIVERSITY

2017

KINETIC CHARACTERIZATION OF *LISTERIA MONOCYTOGENES* 2-C-METHYL-D-ERYTHRITOL 4-PHOSPHATE CYTIDYLYLTRANSFERASE (CMS) ENZYMES USING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)

MARK OBLAZNY

COMMITTEE MEMBERS:

Jon Friesen, Chair

Marjorie Jones

Jeremy Driskell

ACKNOWLEDGMENTS

I would like to thank Dr. Jon Friesen for granting me the opportunity to work in his research lab. He has been an exceptional mentor to me throughout my tenure at Illinois State University, having provided expert guidance, encouragement, and unwavering support. He has played an integral role in shaping my most valuable characteristics. In addition, I am thankful for Dr. Marjorie Jones who challenged me to become a better scientist.

I am deeply grateful to my peers in the chemistry department at Illinois State for creating a welcoming and engaging environment, which became pivotal in the development of my thesis. Lifelong friendships have been created within the chemistry department, and I am thankful for their support.

Above all, I want to express sincere appreciation to my family, especially my parents, who I am forever indebted. Because of their support and sacrifice, I am have been able to accomplish my educational goals. They have played an integral role in the development of my identity and work ethic.

M. O.

CONTENTS

TABLES

FIGURES

CHAPTER I: INTRODUCTION

2-C-methyl-D-erythritol-4-phosphate Cytidylyltransferase

2-C-methyl-D-erythritol-4-phosphate cytidylyltransferase (CMS EC # 2.7.7.60) is a member of the cytidylyltransferase family of enzymes, and is essential to many prokaryotic organisms. CMS is an enzyme that catalyzes the third reaction in the methylerythritol phosphate (MEP) pathway, a biosynthetic pathway for production of isoprenes. Other names the enzyme is referred to in literature are IspD (isoprenoid synthesis domain containing), MEP cytidylyltransferase, CDP-ME synthetase, or YgbP (open reading frame/gene). The MEP pathway forms the products isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are the phosphate charged precursor molecules of isoprenes, shown in Figure 1. The pathway is also commonly known as the 1-deoxy-D-xylulose 5-phosphate (DOXP/DXP) pathway. The alternative pathway for isoprenoid synthesis is the mevalonic acid (MVA) dependent pathway, which is present in vertebrates (Figure 2). The MVA pathway does not contain the enzyme CMS, thus making it a great target for antimicrobial therapeutic development.

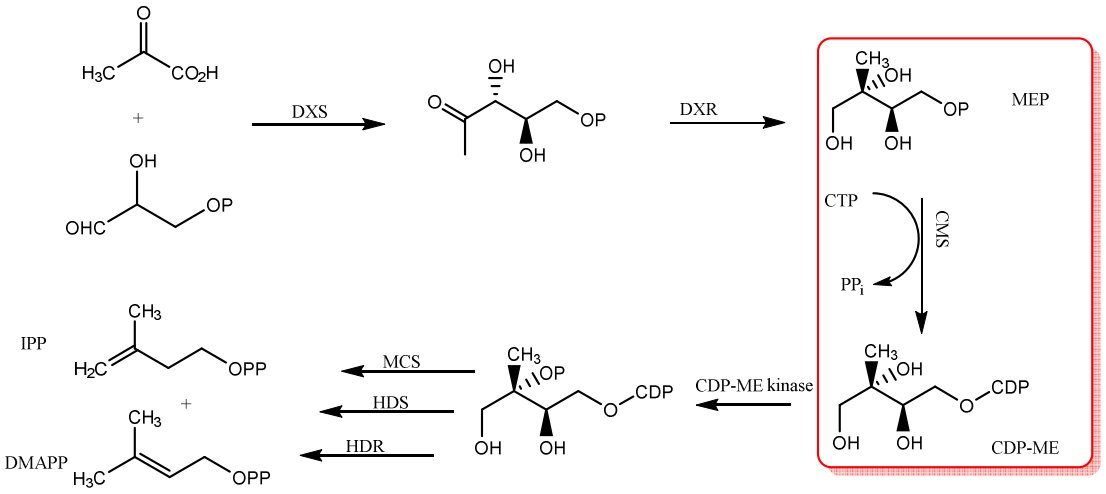

Figure 1. Methylerythritol Phosphate Pathway (MEP). Pathway for the biosynthesis of isoprenoid precursor molecules isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Each "P" symbolizes a phosphate group. "CDP" represents Cytidine diphosphate. Enzyme names: DXS is DOXP synthase, DXR is DOXP reductase, CMS is 2-C-methyl-D-erythritol-4-phosphate cytidylyltransferase, CDP-ME kinase is 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, MCS is 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, HDS is 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, and HDR is 4-Hydroxy-3-methylbut-2-enyl diphosphate reductase. Pathway was made using ChemDraw.

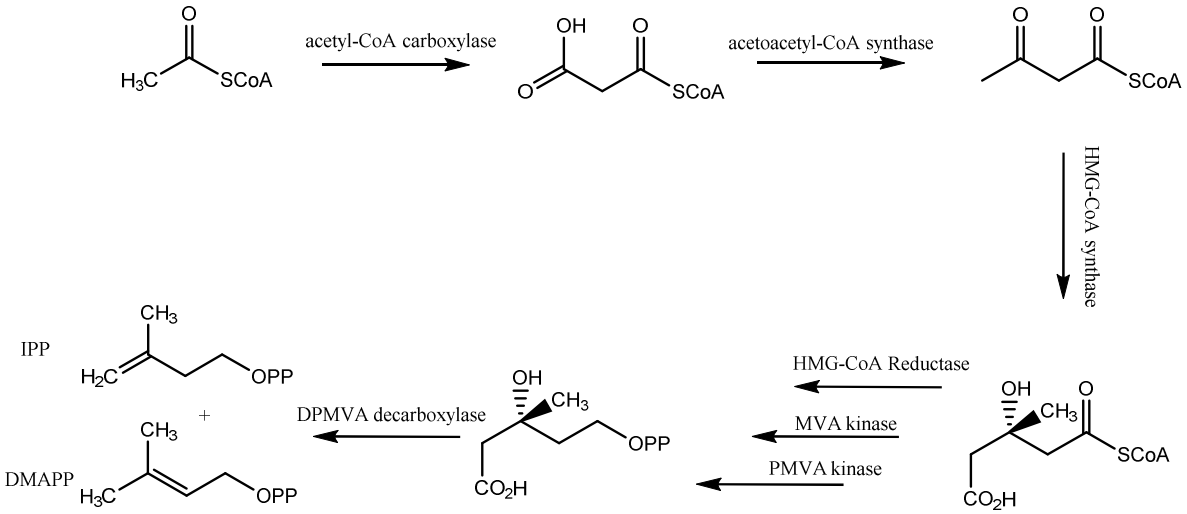

Figure 2. Mevalonic Acid (MVA) Dependent Pathway For Isoprenoid Synthesis In Vertebrates. Each "P" symbolizes a phosphate group. HMG stands for 3-hydroxy-3-methyl-glutaryl. CoA is coenzyme A. MVA is mevalonate-3-phosphate-5-kinase. PMVA is phosphomevalonate kinase. DPMVA ismevalonate-5-pyrophosphate decarboxylase. Pathway constructed using ChemDraw.

Isoprenoids

Isoprenes, the single unit of terpenes, are one of the most common structural motifs in biological systems and one of nature's most common building blocks. Simply, these isoprenoid molecules are necessary for an organism's survival and have over 35,000 primary and secondary metabolites (Chang & Keasling, 2006). Isoprene units can form carbon skeletons ranging from a few linkages to thousands. Isoprenoids are used in many other biosynthetic pathways and build vital metabolites for functionality. Some important metabolites include sterols, vitamins, heme, and ubiquinone (Coenzyme Q) (Ad Heuston, Begley, Gahan, Hill, & Ie, 2012). Sterols are the third class of lipids that are required for a plethora of cellular processes. Sterols are components of plasma membranes found abundantly in eukaryotic cells but less in prokaryotic cells. They play a role in sustaining a liquid ordered state of the membrane that preserves cell membrane fluidity and functionality by creating lipid "rafts" (Simons & Ehehalt, 2002). Ubiquinone is found in mitochondrial membranes and plasma membranes of prokaryotes which serves as an electron carrier in the electron transport chain for the eventual synthesis of ATP.

Isoprene is the most abundant hydrocarbon measurable in the breath of humans, which emphasizes its significance in biological systems (Gelmont, Stein, & Mead, 1981). The isoprene emission accounts for around one-third of all hydrocarbons released into the atmosphere by all organisms (King et al., 2010). Molecules that are biosynthetically synthesized from isoprene subunits play crucial roles in an organism, such as cell signaling, cell maintenance, enzyme cofactors, and cell wall synthesis. An organism without some of these functions will undergo cell death since it will not be able to maintain homeostasis. With the absence of isoprene derivatives, major metabolic pathways will be halted, thus making CMS a great potential target for antibiotic therapeutics.

Isoprene subunits are used in industry today to produce synthetic forms of rubber found in many daily products. Isoprene-derived compounds are found in common household items such as cosmetics, food, baby bottles, tires, and essential oils. These compounds are responsible for some of the pleasant aromas in some cosmetic products. Studies are starting to research terpenoids for their anti-cancer and anti-inflammatory properties (Sharma, Thulasingam, & Nagarajan, 2017). Terpenes are being evaluated in cannabis as they can become useful in the pharmaceutical industry.

Several pathogenic organisms such as *Mycobacterium tuberculosis,* which causes tuberculosis (TB), *Plasmodium falciparum*, the parasite that causes malaria, and *Listeria monocytogenes,* which leads to listeriosis, use the MEP pathway for isoprene biosynthesis. Listeriosis, commonly known as food poisoning, affects about 1,600 individuals per year and is most likely to sicken pregnant women and newborns (Center for Disease Control and Prevention [CDC], 2017). Tuberculosis is one of the top 10 causes of death worldwide. In 2015, 10.4 million people fell ill with TB and 1.8 million died from the disease (World Health Organization [WHO], 2017). In 2015, there were 214 million cases of malaria worldwide resulting in an estimated 438,000 deaths (WHO, 2017). These are just a few organisms of interest although many more can be included as potential targets. Since the biosynthesis of isoprenoids occur in separate pathways it opens up possibilities for next generation antibacterial therapeutics that can target CMS without having adverse effects on vertebrates (humans) which utilize the mevalonate-dependent pathway. Growing antibiotic resistance is becoming an issue of concern which makes the development of novel antibiotics a pressing issue for human society. An enzyme target like CMS may not lead to just curing diseases but improving the quality of life of millions of individuals around the world..

Listeria Monocytogenes

Listeria monocytogenes is the pathogenic gram-positive bacterium that causes listeriosis. Gram-positive bacteria simply means the bacterial cell wall contains peptiodoglycan. *Listeria* is very tough to evade as it has the capability to grow in absence or presence of oxygen (facultative anoerobe), and to replicate and grow at 0 °C (Ramaswamy et al., 2007). This is of great concern as typical refrigeration temperatures around 0 °C to prevent bacterial growth, although this temperature will not affect the ability of *Listeria monocytogenes* to grow. The bacteria are motile with a flagella at temperatures less than 30 °C but not usually under host biological conditions of about 37 °C (Gründling, Burrack, Bouwer, & Higgins, 2004). With the bacteria's ability to grow in low temperatures, it is recognized as a hazard to the food industry. *Listeria monocytogenes* has been associated with such foods as raw milk, pasteurized milk, cheeses, raw vegetables, raw and cooked poultry, and raw meats. It is known to cause food-borne illness, enter and escape cells frequently, and evade the body's immune response systems. It seems that *Listeria* originally evolved to invade membranes of the intestines, as an intracellular infection, and developed a chemical mechanism to do so. This involves the bacterial protein internalin that acts as a ligand that attaches to a protein, cadherin on the intestinal cell membrane and allows the bacteria to invade the cells through a zipper mechanism (Robbins et al., 1999). A zipper mechanism is a form of phagocytosis when a cell engulfs surrounding material and allows it to flow freely in a small vacoule inside a cell. Once inside the cell, *Listeria* is able to escape the vacuole and enter the cytosol through the use of listeriolysin O (LLO) (Portnoy, Auerbuch, & Glomski, 2002). Motility in the intracellular space is due to actin assembly-inducing protein (ActA) which allows the bacteria to recruit the host cell's actin polymerization machinery to polymerize the cytoskeleton to give the bacterial cell motility (Portnoy et al., 2002). The same ActA mechanism

also allows the bacteria to travel from cell to cell. Although not one of the most common diseases, listeriosis is of concern for pregnant patients, neonates, elderly, and immunocompromised individuals.

While CMS is a promising target for novel drug therapeutics, it is far from being well understood and is very understudied. Much like bacterial cell wall biosynthesis, which has been so successfully targeted by antibacterial agents, the MEP pathway is essential for microbial growth, and none of the enzymes of this pathway are present in mammalian cells. Some of the recent discoveries for antibiotic development generally proceed through five phases, including target identification, target validation, lead molecule identification, lead molecule optimization, and preclinical and clinical trials. This thesis project will focus on target identification, which will be choosing the CMS enzyme in the MEP pathway, and target validation studies by obtaining kinetic information. In order to design effective CMS therapeutics, foundational kinetic research must be performed.

Structure Overview of *E. coli* CMS

The *E. coli* CMS isoform is a good model for *Listeria monocytogenes* CMS 1086 and CMS 0235. Although using the *E. coli* as a model for bi-functional CMS enzymes that contain 2 functions due to gene fusion events is less comparable, although this is not the case for CMS 1086 and CMS 0235. Using CMS from *E.coli* as a model, the enzyme is a homodimeric protein with an important structural interlocking arm (Richard et al., 2004). This interlocking arm is essential to protein stability and function. The homodimer of the *E. coli* variant contains two active sites with one on each identical subunit and the interlocking arm lodges itself into the active site of the neighboring subunit to perform catalysis. The interlocking arm domain is a small segment of residues 137-159, while the larger domain is the globular domain with residues 1-136 and 160-236. The protein contains a modified Rossmann motif that is commonly found in proteins that bind nucleotides. A Rossmann motif is an alternating pattern of parallel β-sheets connected by α-helices in a βαβ folding pattern (Hanukoglu, 2015). The *E. coli* CMS appears to contain this motif with the exception of one β-sheet oriented in an antiparallel direction (Figure 3). The interlocking arm region contains a series of β-sheets from one subunit interacting in an antiparallel orientation with the β-sheets of the neighboring subunit.

Figure 3. Structure of *E. coli* CMS Monomeric Unit Displaying the Rossmann Motif (PDB 1INJ). Rossmann fold β-sheets are numbered, 1-7. Structure image made using RasMol, α-helices are yellow, β-sheets are green, and the interlocking arm region is red.

Ribitol 5-phosphate Cytidylyltransferase (RCT)

Ribitol 5-phosphate cytidylyltransferase *(TarI* or RCT EC # 2.7.7.40) catalyzes a similar reaction to CMS and is part of the pathway necessary for biosynthesis of wall techoic acids (WTAs) in gram positive bacteria. The reaction catalyzed is the conversion of CTP and ribitol 5-phosphate to CDP-ribitol, a vital intermediate in the biosynthesis of WTA (Baur, Marles-Wright, Buckenmaier, Lewis, & Vollmer, 2009). The WTA synthesis pathway is separate from the MEP pathway, but upon closer analysis RCT and CMS enzymes display structural similarities. Inhibition of RCT is a potential target for antibiotic development through inhibition of the WTA pathway. The structure of RCT is similar to CMS, which is a homodimer protein containing the interlocking arm and Rossmann like motif, however only three structures have been solved for a ribitol 5-phosphate cytidylyltransferase.

Cell Wall Techoic Acid

Teichoic acids are the major cell wall components of the pathogenic bacterial *Staphylococcus aureus, Staphylococcus saprophyticus,* and *Listeria monocytogenes*. Teichoic acids were discovered in 1958 by Armstrong and co-authors (Armstrong, Baddiley, Buchanan, Carss, & Greenberg, 1958). Techoic acids are made up of anionic polymers of glycerol phosphate or ribitol phosphate, with phosphodiester linkages (Neuhaus & Baddiley, 2003). Wall teichoic acid (WTA), which is covalently linked via a phosphodiester bond to C-6 of N-acetylmuramic acid, plays many roles in the peptidoglycan layer. WTA decorates the outer cell wall and represents about 60% of the total cell wall mass (Neuhaus & Baddiley, 2003). WTA contains anionic character from phosphate groups that can sequester cations from the surrounding environment (Swoboda, Campbell, Meredith, & Walker, 2010). This provides the cell with a reservoir of cations that can be easily accessible for the cells in times of stress. It has

been suggested that cell wall teichoic acids can create localized changes in pH, indirectly modulating the function of some enzymes (Biswas et al., 2012). The polyribitol cell WTA is able to contain modifications on the free hydroxyl groups of ribitol that would allow for additional cell regulation mechanisms. The WTA is the cell's protective layer that stabilizes the cell membrane and provides docking sites for proteins and nutrients as well (Swoboda et al., 2010). *S. aureus* strains that do not express WTAs are unable to adhere to nasal, kidney, and spleen epithelial cells, indicating the reduced virulence of the pathogen (Weidenmaier et al., 2005). WTA is not vital for cell proliferation but it has been seen to play roles in antibiotic resistance and increased virulence. WTAs are structurally diverse between organisms and contain various modifications. A study was performed on the analysis of *Listeria monocytogenes* WTAs confirming the presence of a polyribitol-phosphate backbone (Eugster & Loessner, 2011). Disrupting the biosynthesis of WTAs is a promising target for unique antimicrobial therapeutics. The pathway of WTA synthesis is not well characterized and many of the mechanisms by which WTA is able to regulate cell physiology are still ambiguous. Many genes that encode enzymes within the techoic acid and peptidoglycan synthesis pathways have yet to be identified.

β-lactam Antibiotics

β-lactams are one of the most successful and widely used antibiotics in history (Geddes, Klugman, & Rolinson, 2007). This class of antibiotics is, however, becoming less effective with emerging resistance through multiple mechanisms. A pathway of resistance consists of a β-lactamase that inactivates β-lactam antibiotics by hydrolyzing the β-lactam ring (Drawz & Bonomo, 2010). Despite the emergence of β-lactamase inhibitors that have restored efficacy of the β-lactam class of antibiotics, other mechanisms of antibiotic resistance have emerged. Methicillin resistant *Staphylococcus aureus* (MRSA) has achieved β-lactam resistance through

additional mechanisms, including penicillin-binding protein (PBP2A) which is not susceptible to β-lactam antibiotics. PBP2A is a transpeptidase that functions in secondary crosslinking of the peptidoglycan layer (PG). A secondary transpeptidase PBP4 is also found at the division septum, which generates highly crosslinked PG for cell division. The division septum is the middle portion of the cell after elongation occurs, and this is the region where cell replication is occurring. Localization of the PBP4 at the division septum is reliant on the existence of cell WTA (Atilano et al., 2010). This implies that WTAs play roles in the emergence of antibiotic resistance in MRSA, and inhibition of WTA synthesis can become the forefront in fighting antibiotic resistant bacteria.

Research Objectives

The primary goal of this research was to characterize CMS in the MEP pathway from *Listeria monocytogenes* responsible for isoprenoid synthesis. Characterization of the enzymes included, conducting kinetic experiments to extrapolate kinetic constants K_m and $V_{max.}$ This research lays a foundation for the development of potential inhibitors for CMS enzymes. A recently developed HPLC method for the characterization of cytidylyltransferase enzymes was utilized for the characterizations of CMS 0235 and 1086 (Brault Thesis, 2016). The importance of isoprenoids in biology and industry are apparent with a broad range of applications as discussed earlier. A secondary goal was to identify and explain some of the variation in catalytic activity of CMS 0235 and 1086 by using multiple sequence alignments and structural homology modeling. The objectives were to detect highly conserved amino acids among CMS enzymes and indicate areas that would compromise CMS activity of the enzymes. By incorporating protein modeling we can identify regions of secondary structure that are vital for functionality and compare and contrast secondary and tertiary structure elements.

CHAPTER II: METHODS

Cell Transformation of lmo0235

Plasmid DNA containing the gene for CMS 0235 (pet45b-lmo0235) from Shilpa Maroju (Maroju Thesis, 2008) was inserted into BL21 (DE3)CodonPlus RIL *E. coli* cells (Agilent), which are in the class of B strain *E. coli*. These cells encode a T7 RNA polymerase from bacteriophage that uses double stranded DNA to catalyze the synthesis of RNA in the 5'-3' direction. The T7 polymerase recognizes the lacUV5 promoter sequence TAATACGACTCACTATAG and does not require any additional activators, thus making it ideal for high-level protein expression systems. Another advantage of using this strain of *E. coli* is these cells contain a ColE1-compatible, pACYC-based plasmid containing extra copies of the *argU*, *ileY*, and *leuW* tRNA genes. Due to the increased expression of a heterologous protein using the T7 polymerase, some rare tRNAs may be depleted and stall protein expression. These cells are engineered to be deficient in Lon and OmpT protease proteins which are known to degrade recombinant proteins (Agilent Technologies, 2017).

BL21 (DE3)CodonPlus RIL *E. coli* cells (Agilent) were placed on ice and allowed to thaw slowly and 50 µL of cells were then placed in a 14 mL sterile polypropylene tube, containing 1 µL of pet45b-lmo0235. The mixture of DNA and cells were incubated on ice for 30 minutes then placed in a 42 ºC water bath for 30 seconds to heat shock the cells. This allows for the membranes to become more fluid and porous allowing the plasmid DNA to enter the host cell. Following the heat shock phase, the tube is placed back on ice for 2 minutes as a recovery phase. Following the recovery phase an additional 1 mL of LB medium is pipetted into the tube and incubated in a shaker at 37 ºC for about 1 hour while shaking at about 250 rpm. The cells were plated on agar plates containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. The

plates contained the antibiotics to serve as a selection process for the cells. Two plates were used, one contained 20 µL and the other 200 µL aliquots of cells, since transformation efficiency is unknown. The cells are spread using sterile glass beads and plates are incubated overnight at 37 ^{o}C. Colonies that grow on the plates are presumably the cells that have uptaken the pet45b-lmo0235 plasmid containing the antibiotic resistant genes.

Cell Transformation of lmo1086

Plasmid DNA containing the lmo1086 gene for CMS 1086 (pet45b-lmo1086) from Shilpa Maroju (Maroju Thesis, 2008) was inserted into ArcticExpress (DE3) RIL competent cells (Agilent) for improved protein expression. When using high-level expression systems for heterologous protein, it can result in the synthesis of misfolded and insoluble protein. When lmo1086 is expressed in BL21 (DE3) CodonPlus RIL *E. coli* cells, the protein produced is insoluble and is not suitable for kinetics assays (Maroju Thesis, 2008). Some labor intensive *in vitro* methods call for the unfolding and refolding of these insoluble aggregate proteins. To bypass this, ArcticExpress (DE3) RIL cells can be utilized to yield properly folded soluble recombinant protein. ArcticExpress cells have been specially engineered to express the GroEL/GroES chaperonin complex for improved protein folding activity at low temperatures. The cells contain the *Oleispira antartica* Cpn10 and Cpn60 chaperonins that have optimal activity at 4-12 ^{o}C. The ArcticExpress strain also contains a T7 polymerase and lacUV5 promoter for protein expression (Agilent Technologies, 2017).

The cell transformation process is similar to the lmo0235 procedure. An important modification is the antibiotics used for the selection process. For the lmo1086 in ArcticExpress (DE3) RIL cells, the antibiotics streptomycin, gentamycin, and ampicillin must be used. Gentamyacin resistant genes are on the plasmid containing the Cpn10 and Cpn60 genes for the

chaperonin proteins (Agilent Technologies, 2017). In order to retain the plasmid containing these genes the cells must be grown in medium containing gentamycin.

Expression of lmo0235 in BL21 (DE3)CodonPlus RIL *E. coli* Cells

Following the overnight incubation of cells that are plated on agar plates, colonies were picked for large-scale protein synthesis. A colony containing pet45b-lmo0235 was picked and placed into an appropriately labeled 14 mL polypropylene tube containing 5 mL of LB medium. The sample was then inoculated with 5 µL chloramphenicol (34 mg/mL) and 5 µL of ampicillin (100 mg/mL). The tube was placed into the shaker at 37 ºC at 250 RPM and incubated for about 5 hours. After the 5 hour growth period the contents of the tube were transferred to 50 mL of LB medium containing 50 µL of chloramphenicol (34 mg/mL) and 50 µL of ampicillin (100 mg/mL), then allowed to incubate in shaker at 37 ºC overnight. Following the overnight incubation, cells were transferred into an Erlenmeyer flask of 1 L LB medium containing 1 mL of the antibiotics chloramphenicol (34 mg/mL) and ampicillin (100 mg/mL). The flask was returned to the shaker and incubated for another 3 hours at 37 ºC. After 3 hours a 1 mL sample was removed from the cell culture into a polypropylene microcentrifuge tube and cells were pelleted at 14,000 RPM for 1 minute using a microcentrifuge (FORCE Micro); this will serve as a control, known as the uninduced sample. The supernatant was discarded and the pellet resuspended in 1X SDS buffer solution then stored at -20 ºC. The 1 L culture was then induced with 0.05 g of IPTG, (Molecular Biologicals International Inc.), and placed back in the shaker for 3 hours. During this final 3 hour time period the cells are forced into a high level of expression of the CMS enzyme. At this point another 1 mL sample was removed and treated like the previous uninduced control sample. The 1 L culture was then centrifuged at 5,000 RPM for 5 minutes at 4 ºC using a J25-I centrifuge and a JLA-10.5 rotor. Following centrifugation, pelleted

cells were resuspended in 20 mL of cell lysis buffer (50 mM Tris, 100 mM NaCl at pH 7.5), frozen at -20 ºC; for analysis, cells were thawed then lysed using a French Press, and used for protein purification.

Expression of lmo1086 in ArcticExpress (DE3) Competent *E. coli* Cells

For high level protein expression, a colony was picked from the agar plate and placed into a 14 mL polypropylene tube containing 5 mL of LB medium with 50 µg/mL ampicillin, 20 µg/mL gentamycin, and 75 µg/mL streptomycin. The culture sample was incubated in the shaker at 37 ºC for 5 hours, and then transferred to a 50 mL flask of LB medium containing the same antibiotics and allowed to incubate overnight. After the overnight incubation, the culture was transferred into 1 L of LB medium that contains 50 µg/mL ampicillin, 20 µg/mL gentamycin, and 75 µg/mL streptomycin and incubated in the shaker for an additional 3 hours at 37 ºC. A sample was removed as the uninduced form of the cells and saved for SDS-PAGE analysis. The 1 L culture flask was then induced with 0.05 g of IPTG and placed in the cold room at 4-12 ºC for 24 hours. Following the 24 hour time period, a sample again was obtained as the induced sample and analyzed using SDS-PAGE to provide evidence of successful protein production. Cells were centrifuged and placed in buffer following the same procedure as used for the lmo0235 cells.

Lmo0235 and lmo1086 Cell Lysis and Purification

Cells were subjected to the French Press (Spectronic Instruments) for cell lysis. A 20 mL sample of cells was thawed and dosed with 500 µL of inhibitor cocktail (Sigma Aldrich) to prevent protein degradation. The cells were placed into the French Press at 10,000 psi to lyse the cells and release the soluble enzymes (lmo0235 and lmo1086) into the buffer. A 1 mL sample was removed, labeled as crude extract, and frozen at -20 ºC for further protein analysis. The

remaining crude extract was then centrifuged at 15,000 RPM for 15 minutes at 4 ºC using a J25-I centrifuge and a JA-17 rotor. Another 1 mL sample was removed and stored as supernatant at -20 ºC, and the pellet was discarded.

A high affinity cobalt TALON column (Clontech) was used to purify CMS from the supernatant after centrifugation. Three milliliters of washed TALON Metal Affinity Resin, stored in 20% ethanol, was poured into a column and 50 mL of lysis buffer was flowed through to equilibrate the column. The supernatant was diluted with buffer to approximately 50 mL for enhanced column flow rate. The diluted supernatant was passed through the Cobalt TALON Metal Affinity column. The CMS enzymes contain a 6x-His tag on the N-terminal end. The histidine tag on the enzymes interacts with the cobalt on the resin beads and the enzymes are retained on the cobalt column. Approximately 40 mL of lysis buffer was passed through the column to remove any unbound proteins and labeled as column flow through (C.F.T). Following the C.F.T, 30 mL of 10 mM imidazole in lysis buffer was passed through the column to remove any weakly bound proteins from the cobalt column. This fraction is labeled as the column wash. Finally, 10 mL of 150 mM imidazole elution buffer in lysis buffer was passed through the column to remove the CMS enzymes, and 1 mL fractions were collected in microcentrifuge tubes. All samples collected were subjected to protein determination.

Protein Determination

A Bio-Rad assay (Bio-Rad Inc.) was used to determine protein concentration in purified samples. The Bio-Rad Protein assay is based on the Bradford method (Bradford, 1976). The assay involves using an acidic dye, that upon binding of protein, shifts the absorbance of the dye. In the Bio-Rad assay, Coomassie Brilliant Blue G-250 dye has a λ_{max} at 465 nm. When protein is added to a solution containing the dye, a shift in λ_{max} is observed from 465 to 595 nm in response

to protein binding. The absorbance at 595 nm is linear with respect to protein concentration. A calibration curve using bovine serum albumin (BSA) from 0-20 μg was utilized.

From each elution fraction 10-100 μL sample was diluted in 200 μL of Bio-Rad assay dye and enough water to achieve a total final volume of 1 mL. Each sample containing the Bio-Rad dye and protein was analyzed using a UV-Vis spectrometer (Evolution 250 BIO by Fisher Scientific). When all fractions are analyzed, the fraction with the highest concentration of protein is chosen for the enzyme assay.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE), is a common method used for the separation of proteins by apparent molecular mass by applying an electric field across the gel (Laemmli, 1970). Lauryl sulfate (SDS) is an anionic detergent, which denatures protein and binds to the proteins to produce an overall negative charge. The negative charges on SDS are strongly attracted toward an anode (positively-charged electrode) in an electric field and cause the migration of protein through the gel. When proteins are denatured into their linear polypeptide chains, SDS molecules interact with each protein similarly so they contain uniform charge to mass ratios. The migration distance of each protein is inversely proportional to the log of its mass. Running a sample of protein standard that contains known molecular mass proteins on the same gel and plotting migration distance vs. the log molecular mass will produce a standard curve that can be used to determine unknown protein masses. A typical SDS-PAGE gel contains two portions, a resolving and a stacking gel portion. The stacking gel is the upper portion of the gel that is responsible for compressing the loaded protein sample into a thin layer prior to reaching the resolving gel. This allows all proteins to enter the

resolving gel at the exact same time. The resolving gel is the portion that is responsible for the separation of the polypeptides by mass.

Resolving and stacking gels were prepared for protein analysis. A 12% (wt/v) resolving gel was prepared by mixing 2.15 mL water, 1.5 mL of 40% acrylamide solution (Fisher), 1.25 mL of 1.5 M pH 8 Tris buffer (USB Corporation), 50 µL of 10% SDS buffer (ACROS), 100 µL of freshly prepared 10% (wt/v) ammonium persulfate solution (APS) (Fisher Scientific), and 4 µL of tetremethylethylenediamine (TEMED) (Fisher BioReagants). The APS is a source of free radicals that enhances the polymerization of acrylamide for gel formation. TEMED is used to stabilize free radicals which aids in the rapid rate of polymerization. The free radicals that are produced by APS react and form free radicals on acrylamide monomers that then react with other acrylamide monomers to polymerize. The acrylamide polymerization forms a porous gel that separates the proteins. The gel is prepared with a pH 8 buffer for more anionic glycine to slow protein migration and enhanced protein separation. The stacking gel was then poured between two glass plates with a layer of isopropanol and allowed to polymerize. The resolving gel is poured to leave approximately 1 cm of space for a stacking gel. A 4% (wt/v) stacking gel was then prepared with 1.5 mL of water, 250 µL of 40% acrylamide solution, 250 µL of 1.0 M pH 6.8 Tris buffer (USB Corp.), 20 µL of 10% SDS buffer (ACROS), 40 µL of 10% APS (Fisher Scientific), and 4 µL of TEMED (Fisher BioReagants). The isopropanol layer was then removed, the stacking gel was poured, and well combs were placed between the plates for sample loading. The lower pH of the stacking gel favors zwitterionic glycine molecules to increase migration.

Protein samples were prepared by diluting pure protein sample with water and 4x SDS dye to a desired protein concentration. The 4x SDS dye is a solution containing 270 mM SDS (ACROS), 6 mM bromophenol blue, 200 mM Tris-HCl (USB), 400 mM dithiothreitol (ACROS),

and 40% (v/v) glycerol (Sigma-Aldrich). The desirable concentrations to load of pure protein is about 0.1 µg/µL and for mixtures of protein is 1.0 µg/µL. For pure protein ideal loading amount is 1-5 µg and for mixtures is 10 µg. Each loading well has the capacity to hold 20 µL of sample. Protein samples diluted for SDS-PAGE analysis were placed into a boiling water bath for 5 minutes to encourage full denaturation and even coating of SDS molecules along the polypeptide. Protein standards and samples were loaded into the wells and electrophoresed at 150 V for an hour or until the dye front runs off the gel. Following electrophoresis, the glass plates were removed and the gel placed into a staining solution. The staining solution contains Coomassie blue dye dissolved in 45% (v/v) methanol (Sigma) and 10% (v/v) acetic acid (Fisher). The staining container was placed on a rocker (Labnet International) at 2 RPM for 30 minutes. The gel was then removed from the staining solution and placed into a destain solution which is 45% (v/v) methanol and 10% (v/v) acetic acid. The destain was then placed on the rocker at 2 RPM to for 1-12 hours to remove any excess dye from the gel. Once destain solution was saturated with dye it was replaced with fresh solution. The gel was analyzed to obtain molecular mass information with target protein migration in comparison to known protein standards.

Enzymatic Assay Procedure CMS 0235

CMS 0235 enzyme was diluted with nanopure water (exception of pH profile) to 0.020 µg/µL, to ensure that when 5.0 µL of enzyme is added, the total enzyme in the assay is 0.10 µg. A previous study showed that using 0.10 µg of enzyme in an assay gives a good signal to noise ratio for measured product (Brault & Friesen, 2016). A stock solution of each of the substrates (MEP and CTP) was prepared to a concentration of 20 mM in H_2O. Cytidine 5'-triphosphate disodium salt and 2-C-methyl-D-erythritol 4-phosphate lithium salt were obtained from Sigma-Aldrich. A solution of 10 mM dithiothreitol (DTT) was prepared for the enzyme assay as well.

The purpose of DTT in an enzyme assay is to keep the protein in a reduced environment. An example would be to keep cysteine amino acids reduced on the protein surface as so to retain enzymatic activity. A final 10x reaction buffer solution was made of 250 mM imidazole (ACROS) and 500 mM magnesium acetate (EM Industries Inc.) adjusted to a pH of 7.4.

The first assay conducted was a time dependent assay. Ten microcentrifuge tubes were used as the reaction vessels for the enzyme assay. For the time dependent assay each tube was loaded with the following: 7 µL of water, 2 µL of 10 mM DTT, 2 µL of 20 mM CTP, 2 µL of 20 mM MEP (Sigma-Aldrich), and 2 µL of 10x reaction buffer. The enzyme and reaction tubes are set to reach room temperature prior to addition of enzyme. Five µL of 0.020 µg/µL of CMS 0235 were added to each tube to begin the reaction, and placed on a hot plate set to 37 ^{o}C. Every 3 minutes a tube was removed and the enzyme reactions were terminated by placing the tubes in boiling water, which will denature the enzyme. Removing tubes every 3 minutes will cover a range of 0-30 minutes of reaction time. After each tube was allowed to boil for 5 minutes, the tubes are centrifuged at 15,000 RPM for 1 minute. This will pellet the denatured protein to the bottom of the microcentrifuge tube. HPLC vials (CG LifeSciences) were prepared by diluting 15 µL of the enzyme reaction in 276 µL of water to have a total of 291 µL samples in each HPLC vial. The remainder of the kinetic assays were conducted for 15 minutes.

An assay was performed with varying concentrations of CTP while keeping the concentration of MEP at 2 mM in every reaction tube. Various concentrations of CTP were prepared from a 20 mM stock solution of CTP. The reaction tubes (in triplicate) were loaded similarly to the previous assay with the exception of CTP as concentration was varied with each tube. Upon addition of CMS 0235, each tube was allowed to incubate for 15 minutes. Reactions

were terminated and handled the same as the previous assay. Varying MEP enzyme assay was performed following the same procedure.

A pH profile of CMS 0235 was obtained. A 10x buffer of 500 mM Tris (pk_a= 8.07), 500 mM glycine (pk_1= 2.4 pk_2= 9.6), 500 mM phosphate (pk_1= 2.2, pk_2= 7.2, pk_3= 12.3) with 250 mM $MgCl_2$ was prepared and 20 mL aliquots were each adjusted to a desirable pH with NaOH or HCl. The buffers were prepared to cover a range of pH 4-12. A minor adjustment to the assay was the enzyme was aliquoted and diluted with 1x pH buffer to 0.020 µg/µL and properly labeled. Reaction tubes were prepared with 7 µL of water, 2 µL of 10 mM DTT, 2 µL of 20 mM CTP, 2 µL of 20 mM MEP, and 2 µL of 10x pH buffer. Reactions were initiated with the addition of 5.0 µL of enzyme. The assay was performed following the same procedure as above.

Various divalent cations were also tested for their effects on enzymatic activity. Various 10x buffers were prepared containing 500 mM imidazole and 250 mM MCl_2 (where M= Mg, Ba, Co, Ni, or Mn) adjusted to pH 7. Reaction tubes were prepared with 7 µL of water, 2 µL of 10 mM DTT, 2 µL of 20 mM CTP, 2 µL of 20 mM MEP, and 2 µL of MCl_2 10x buffer. Reactions were initiated with the addition of 5.0 µL of 0.020 µg/µL CMS 0235 and the tubes were incubated on a hot plate at 37 ^oC for 15 minutes. Reactions were terminated and handled the same as previous assays.

Enzymatic Assay Procedure CMS 1086

An assay was completed to determine if product formation increased with increasing concentrations of CMS1086. The assay mixture still contained the same contents as CMS 0235, 7 µL of water, 2 µL of 10 mM DTT, 2 µL of 20 mM CTP, 2 µL of 20 mM MEP, and 2 µL of 10x reaction buffer. Separate tubes were prepared with the assay mixture and then varying amounts of enzyme (0-1.3 µg CMS 1086) were placed in each tube to begin the reaction. The reaction was

allowed to incubate for 15 minutes and analyzed following procedures of CMS 0235. Product formation was low, but it was determined that 1.0 µg would be the optimal amount of enzyme for further assays.

A time variation assay was performed with 1.0 µg CMS 1086 over the range of 0-285 minutes. For further assays a time of 150 minutes was selected. All further assays were conducted under the same procedure as CMS 0235 with the exception of time and amount of enzyme present in each tube.

HPLC Calibration

A calibration curve was obtained using High Performance Liquid Chromatography (HPLC). Since the product of the reaction is not available for purchase, CDP-choline was used as the standard for the calibration. Using CDP-choline as the standard an assumption will be made that the product CDP-ME will approximately have the same molar absorptivity as CDP-choline. Figure 4 displays the calibrant molecule CDP-choline as well as CDP-ME, which is the product of the enzymatic reaction. The chromophore on both molecules are identical which is the reason CDP-choline was selected to be the calibrant. In separate microcentrifuge tubes were added varying concentrations of CDP-Choline from 0.00100-0.0500 mM and each tube contained 0.200 mM CTP as well. Each tube was boiled for 5 minutes to simulate the enzyme assay procedure. After each tube had been boiled, HPLC vials were prepared containing 200 µL of standard and 200 µL of nanopure water, and placed in the autosampler. HPLC conditions are the same as experimental with a 10 µL injection. A plot with nanomoles of CDP-choline injected vs. peak area is shown in Figure 5.

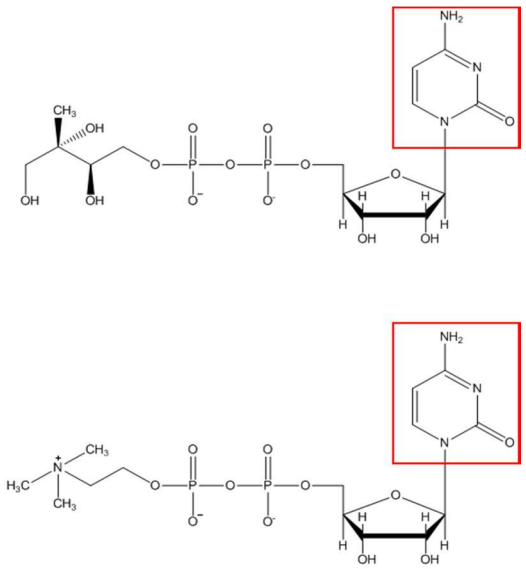

Figure 4. CDP-ME and CDP-choline Structures. Chromophore outlined in red and has an absorbance at 254 nm.

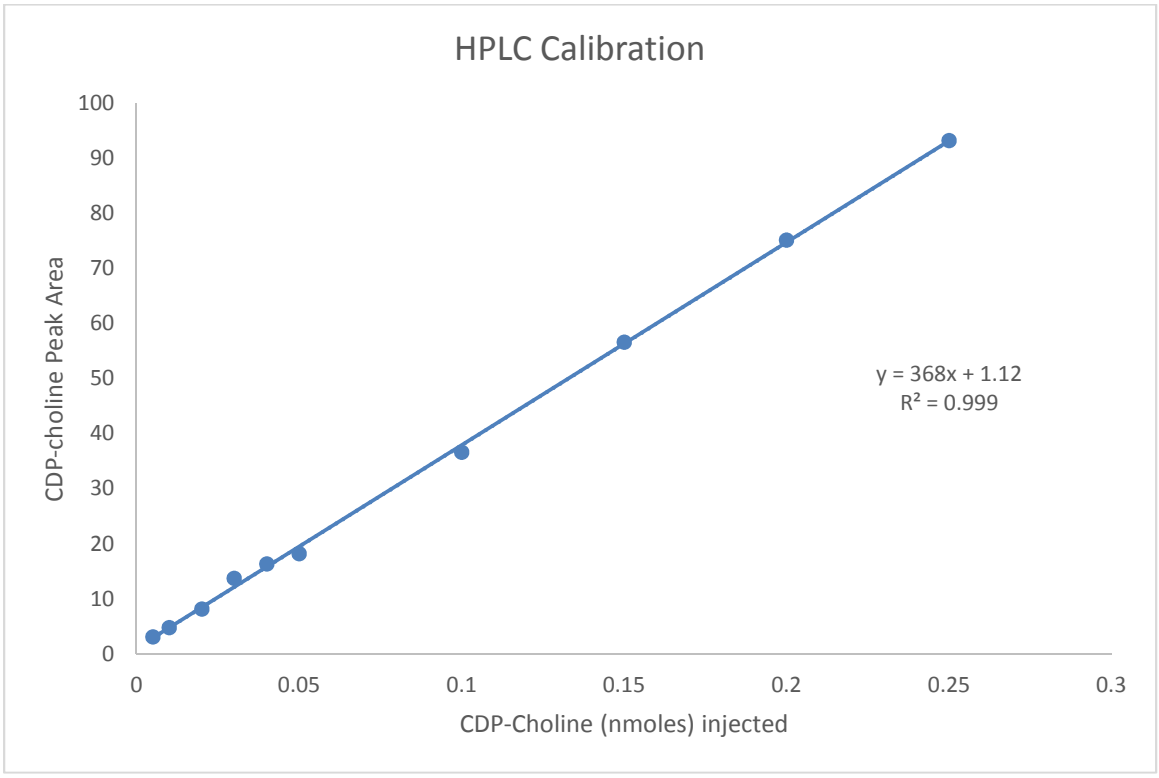

Figure 5. HPLC Standard Curve. Peak area plotted as a function of CDP-Choline. Each point represents a single injection.

HPLC Method Application

A modified version of a recently developed method for enzyme assay analysis using high performance liquid chromatography for cytidylyltransferase enzymes was utilized (Brault & Friesen, 2016). The study showed that the substrate, cytidine triphosphate (CTP) and product, 4-diphosphocytidyl-2-C-methylerythritol (CDP-ME) are able to be separated using HPLC (Brault & Friesen, 2016). Prior assay methods included ^{14}C radioisotope labelling and a colorimetric assay using malachite green as described by Jim Brault (Brault Thesis, 2016). Using HPLC as the preferred method allows detection of substrate and product directly. In order to utilize the malachite green assay, a phosphatase enzyme is required to convert the inorganic pyrophosphate, a byproduct of the enzymatic reaction, to free phosphate (Bernal, Palacin, Boronat, & Imperial, 2005). The free phosphate is then reacted with the malachite green dye to produce a green product, which can then be determined spectrophotometrically. Also, HPLC advantages exist over the isotopic labeling due to more limited waste disposal as well as a more cost effective alternative for the assay (Brault & Friesen, 2016).

The HPLC that was used for analysis was a HP series 1100, containing an autosampler and a PDA Plus detector (HP/Agilent). The column used was a Sunfire™ C18 with measurements of 4.6 mm x 15 cm and 5 μm particle size. A betasil C8 guard column was also assembled to prevent any contamination of the column and increase column lifetime.
An aqueous mobile phase (0.10 M ammonium bicarbonate) was prepared to an adjusted pH of 7.3, essential for column health and optimal separation of CTP and CDP-ME. The final mobile phase was 99% buffer and 1% acetonitrile. The HPLC was allowed to run for 45 minutes at 0.500 mL/min to equilibrate the column with the mobile phase prior to sample injection. The spectra were monitored at 254 nm and the only optically active compounds in the reaction are

CTP and CDP-ME. An auto sampler was used to inject each sample with an injection volume of 10 µL. Typical chromatograms represent a CTP peak eluting at 3.1 minutes and the product peak elutes at 4.7 minutes. Since mobile phase was prepared fresh daily it slight differences in peak elution times (CTP: 3.0-3.3 min. CDP-ME: 4.5-4.9 min.) were observed. The CDP-ME is slightly more nonpolar than CTP due to its loss of phosphate and addition of erythritol, which explains the later elution time on a RP-C18 column.

CHAPTER III: RESULTS AND DISCUSSION

Expression and Purification of CMS 0235 and 1086

CMS 0235 and CMS 1086 were both successfully expressed upon induction with IPTG. CMS 1086 is a 27.5 kDa protein and CMS 0235 is a 25.9 kDa protein each, with a 6x His-tag at the amino terminus. SDS-PAGE (Figure 6) confirms that, upon induction by IPTG, the cells express the proteins, and bands are present in the expected regions of the gel between 25-37 kDa protein standards. Unfortunately, CMS 1086 was not soluble under the expression system of BL21 (DE3)CodonPlus RIL *E. coli* cells as shown by Shilpa Maroju in her thesis work (Maroju Thesis, 2008). Therefore, CMS 1086 was expressed in ArcticExpress (DE3) competent cells and successful expression was observed (Figure 7).

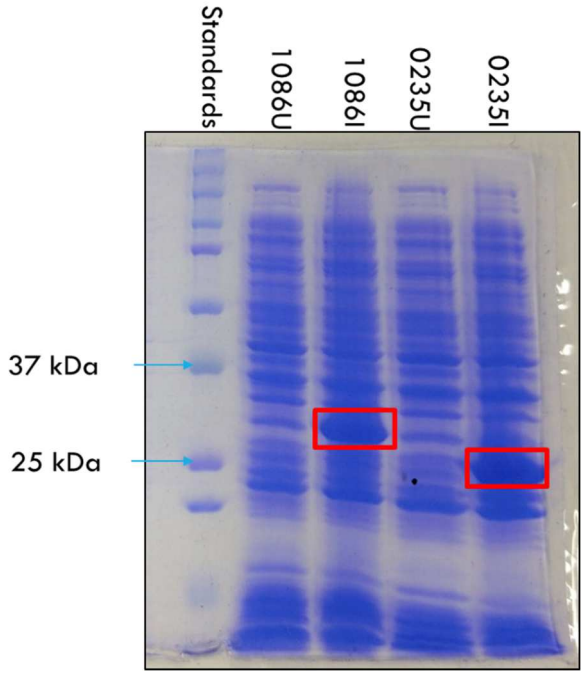

Figure 6. SDS-PAGE of CMS 0235 and 1086 Expression Upon Induction of Cells with IPTG in BL21 (DE3) CodonPlus RIL *E. coli* Cells. "U" is uninduced, "I" is Induced. Red boxes indicate position of CMS 1086 (1086I) and 0235 (0235I) on the gel. CMS 1086 was insoluble using BL21 (DE3) CodonPlus RIL *E. coli* cells.

The ArcticExpress (DE3) RIL competent cells display a large band at about 57 kDa which is the Cpn60 subunit of the chaperonin protein (Figure 7). The Cpn10 subunit is about 10 kDa and not easily identified on the SDS gel since lower molecular mass proteins are more difficult to identify. Purification gels of both CMS isoforms indicate that the elution fractions collected display the presence of a prominent band at the correct molecular mass (Figures 8 and 9). The standards of each purification gel were plotted using the log of the molecular mass as a function of migration distance shown in Figures 10 and 11. Using non-linear analysis, the experimental values for molecular mass of CMS 0235 and 1086 extrapolated from the curve were 26.3 kDa and 28.4 kDa, respectively. These are near the theoretical molecular mass of CMS 1086 and 0235 since both of the proteins on the gel contain a 6x-His tag. The 10 mM imidazole wash removes some weakly bound proteins and a small amount of CMS visible on the gel. The enzymes are also found in the crude extract, supernatant (SUP), and column flow through (C.F.T). The C.E. is the cell extract following the cell lysis procedure which contains all soluble and insoluble materials. The SUP is aqueous sample after centrifugation and the pellet is the insoluble material formed upon centrifugation. The C.F.T is the sample collected after the SUP is flowed through the column, which at this point the CMS should be retained on the column. Each elution number corresponds to the 1 mL elution fraction from the column upon eluting enzyme. By SDS-PAGE analysis, it was determined that both CMS isoforms were successfully expressed and purified for enzyme kinetics.

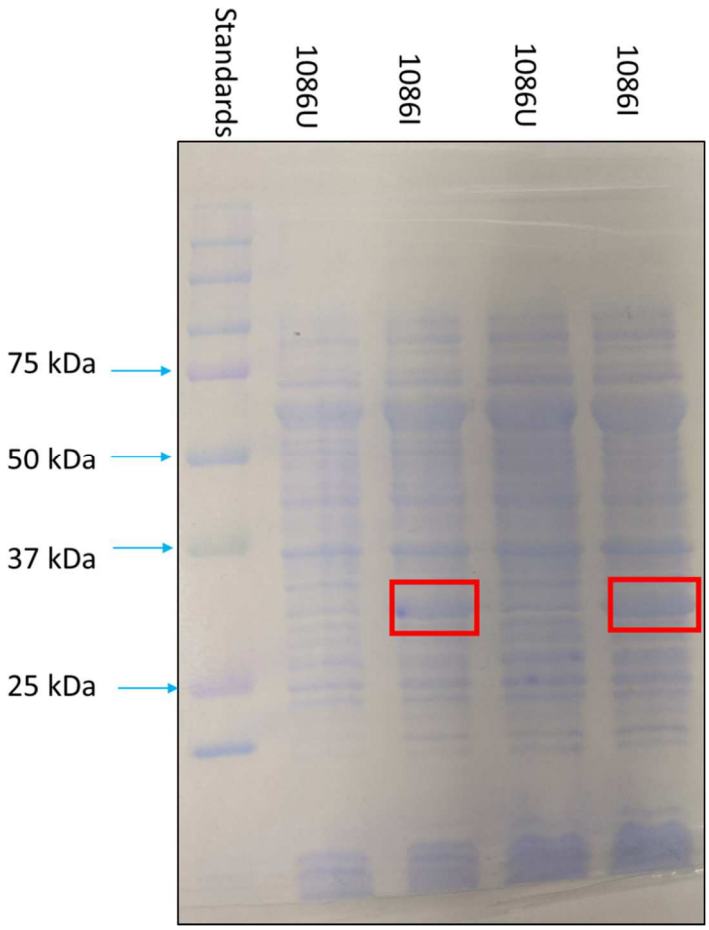

Figure 7. SDS-PAGE of CMS 1086 Expression Upon Induction of Cells with IPTG in ArcticExpress (DE3) RIL *E. coli* Cells. "U" is uninduced, "I" is induced. Induced lanes display a band at around 30 kDa that is outlined in red while the uninduced lanes did not. This is an indication of successful expression of CMS 1086. Samples were loaded twice for clarity.

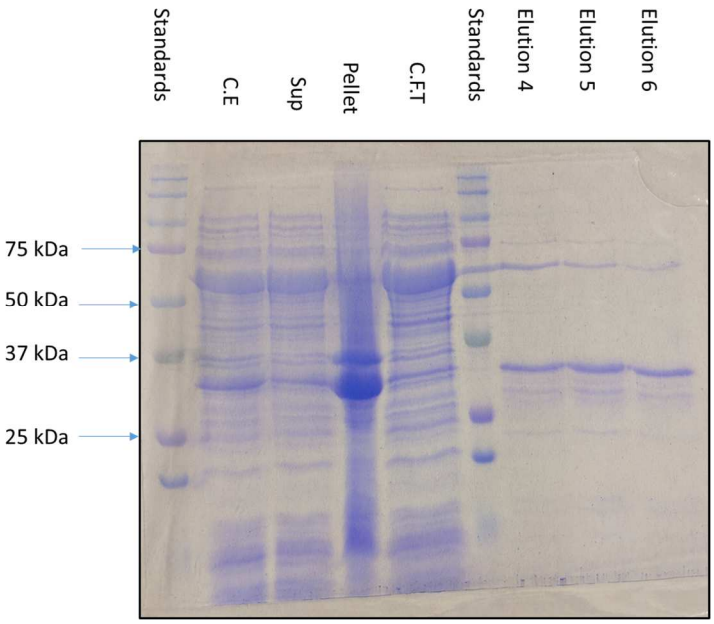

Figure 8. SDS-PAGE Purification Gel of CMS 1086. The large band at 57 kDa is the Cpn60 subunit of the chaperone folding protein expressed in ArcticExpress (DE3) cells. The Cpn10 subunit is not visible on the gel, which is theoretically about 10 kDa.

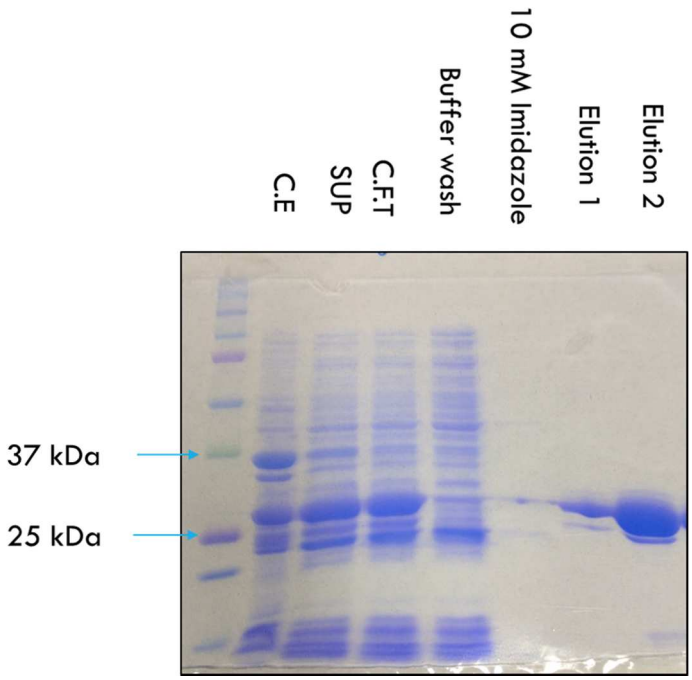

Figure 9. SDS-PAGE Purification Gel of CMS 0235. Elution 1 and Elution 2 are the fractions collected from the column.

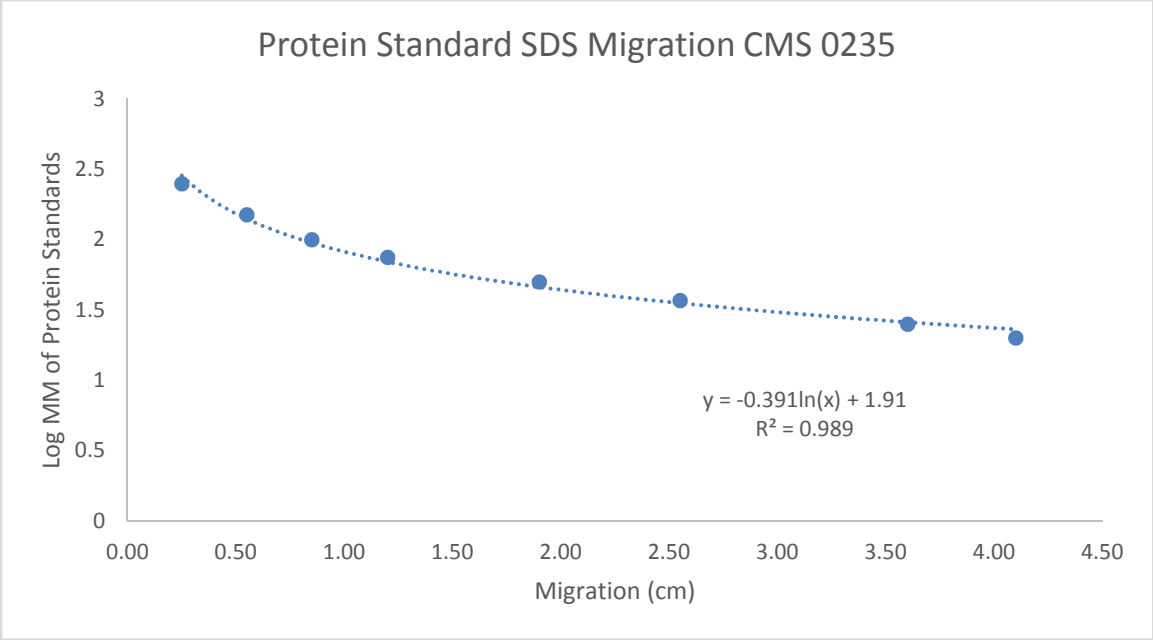

Figure 10. Plot of Log of Molecular Mass vs Migration from CMS 0235 Purification Gel. Experimental molecular mass of CMS 0235 was 26.3 kDa.

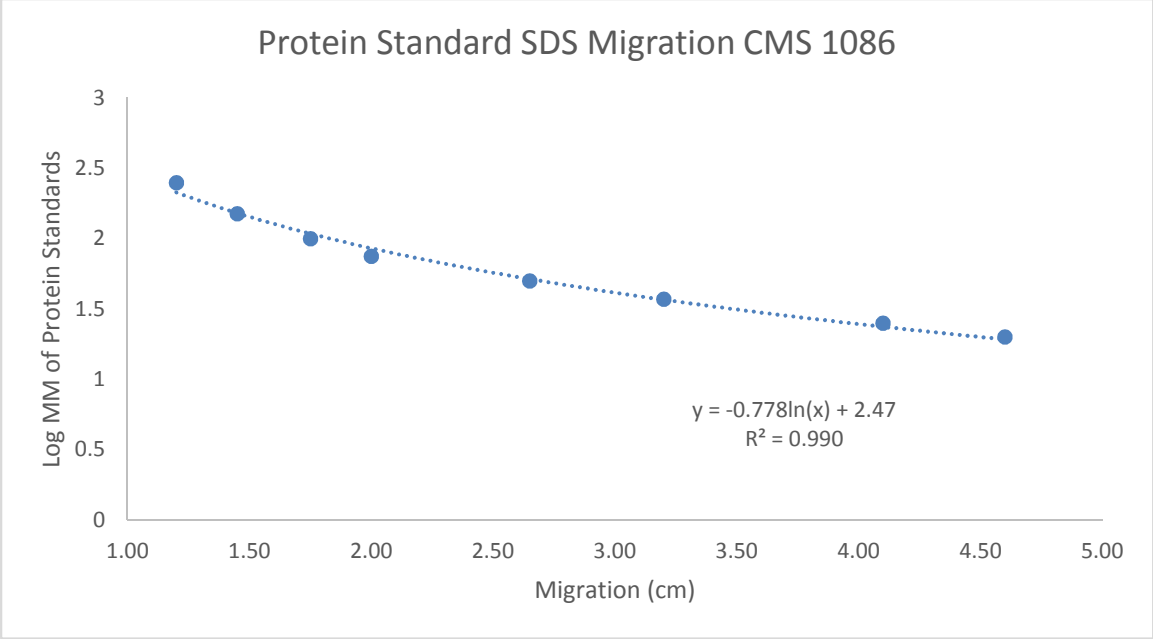

Figure 11. Plot of Log of Molecular Mass vs Migration from CMS 1086 Purification Gel. Experimental molecular mass of CMS 1086 was 28.4 kDa.

Validation of HPLC

A recent study showed a successful separation of CTP and CDP-ME by using HPLC (Brault & Friesen, 2016). In the previous study, substrate CTP eluted at about 2.5 minutes and product CDP-ME eluted at 3.1 minutes. The HPLC conditions used by Jim Brault are outlined in his thesis (Brault thesis, 2016). In this current study, a modified method was used which included a guard column and a reduction of eluent (acetonitrile) from 2% to 1%. In this current work a Sunfire™ C18 column was used with a betasil C8 guard column. A slight drop in pH to 7.3 from 7.4 was used to provide sufficient separation and increase column health. Injection of CTP and CDP-Choline (Figure 12) resulted in sufficient separation with CTP elution at about 3.1 minutes and CDP-Choline elution time at 3.7 minutes. A second chromatogram in Figure 13 is the injection of CTP and CDP-ME (from an enzyme assay) and indicates a larger separation was achieved with CDP-ME eluting at 4.8 minutes. The successful characterization of two CMS enzymes in this thesis further validates this method for the characterization of cytidylyltransferase enzymes without the use of coupled assays or isotope labeling.

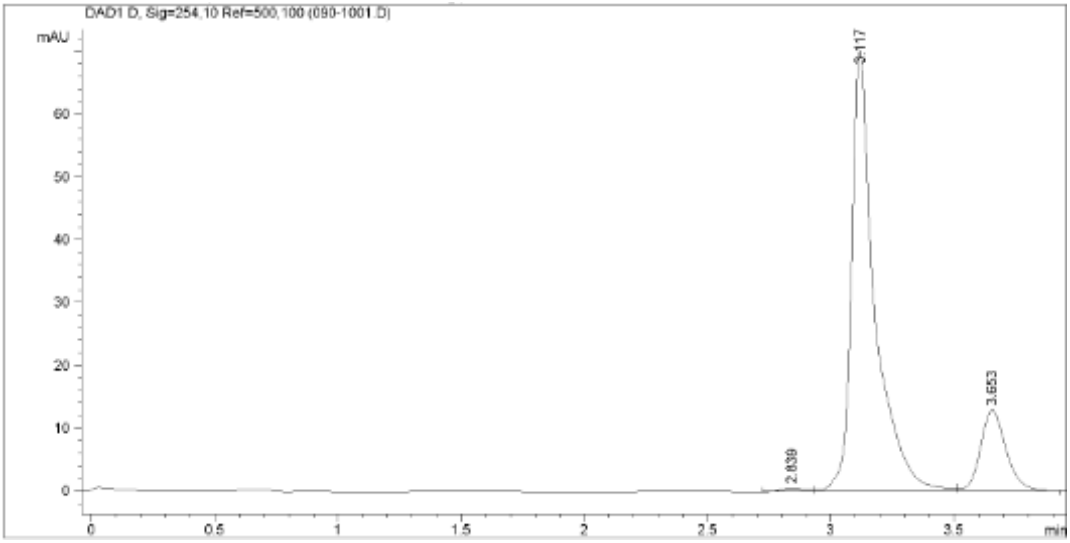

Figure 12. HPLC Chromatogram Injecting CTP and CDP-choline Monitored at 254 Nanometers. Elution of CTP is at 3.117 minutes and CDP-choline elution peak at 3.653 minutes.

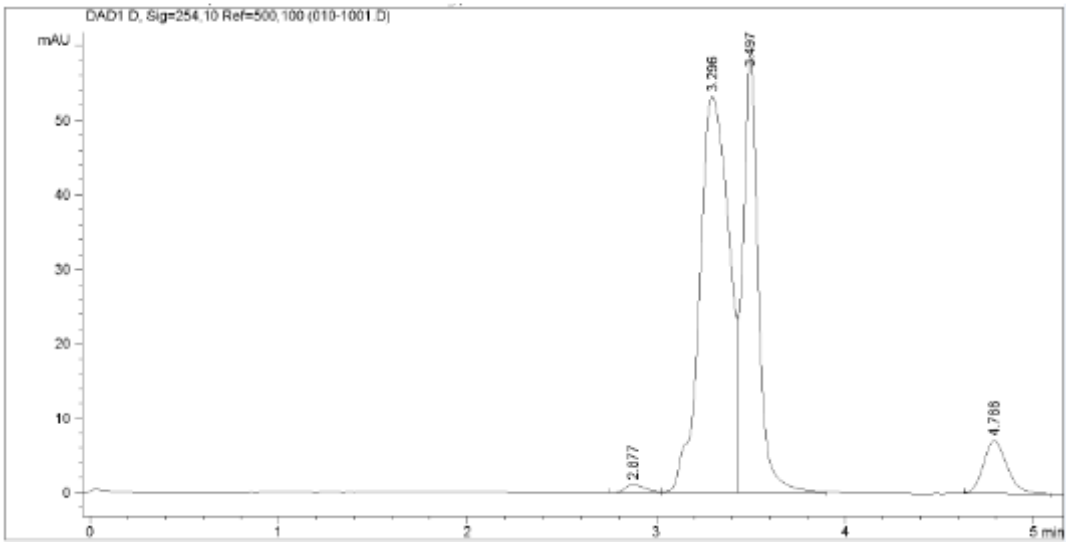

Figure 13. HPLC Chromatogram Injecting CTP and CDP-ME Monitored at 254 Nanometers. Elution of CTP is at 3.296 and 3.497 minutes and CDP-ME elution peak at 4.786 minutes. Split peak for CTP is likely due to multiple protonation states of CTP.

CMS Enzyme Kinetics

Studying enzyme kinetics is important to understanding catalytic mechanisms, drug inhibition, and enzyme regulation in metabolism. One of the hallmarks of enzyme kinetics is the experimental determination of kinetic constants K_m and V_{max}. These values are determined using the Michaelis-Menten equation, which is currently one of the most effective models for investigating enzyme kinetics. Michaelis-Menten kinetics states that the rate of reaction is dependent on substrate concentration, thus if substrate concentration increases, the rate of reaction increases as well. Eventually, increasing substrate concentration to the point of enzyme substrate saturation will reach maximum velocity. The point at which the reaction is considered to have reached a theoretical maximum of enzyme-substrate complex saturation is referred to as the V_{max}. Using this model, K_m is the substrate concentration at which the enzyme is functioning at ½ V_{max}. Michaelis-Menten kinetics fits the model of one substrate and one product, although if the enzymes of interest have multi-substrate molecules this model can still be applied. In order to simplify the reaction and apply the Michaelis-Menten kinetics model, one substrate is held constant at a saturating concentration while the other substrate concentration is varied. Data were best fit using SigmaPlot (SYSTAT) nonlinear regression analysis, Lineweaver-Burk, Eadie Hofstee, and Hanes Woolf plots to determine the kinetic parameters K_m and V_{max}.

Michaelis-Menten: $V_o = \frac{V_{max}[S]}{K_m + [S]}$

Hill Equation: $V_o = \frac{V_{max}[S]^n}{{K_m}^n + [S]^n}$

Nonlinear regression analysis using SigmaPlot employed the Hill equation, which is a powerful tool for the investigation of allosteric regulation. Enzymes that have multiple subunits may undergo allosteric regulation, which is the effect that a substrate molecule binding to one active site has on the binding affinity of other molecules on other active sites. The Hill

coefficient is the "n" term in the Hill equation and provides information on the cooperative binding of each substrate. If the term is >1 it indicates positive cooperativity, thus a value of less than 1 indicates negative cooperativity, and a value equal to 1 reflects no cooperativity. Positive allosteric modulation occurs when the binding of one ligand enhances the attraction between substrate molecules and other binding sites. A negative cooperativity implies the opposite, that binding of one ligand decreases the attraction between substrate molecules and other binding sites, and no cooperativity suggests substrate molecules induce no effect on binding sites.

The conversion of CTP and MEP to CDP-ME as a function of enzyme (µg) for CMS 1086 are reported in Table 1. Product formation was not detected using the HPLC until >0.8 µg of CMS 1086 was added to the reaction. The assay was conducted with a 15-minute incubation time, which was not long enough to detect satisfactory levels of product. CMS 0235 results for varying enzyme were completed in a recently published paper, thus not conducted for this thesis project (Brault & Friesen 2016).

CMS 1086 in reaction (µg)	Product Peak Area
0.10	0
0.40	0
0.60	0
0.80	0
1.0	8.4
1.5	13.4

Table 1. Effects Of CMS 1086 Added To Reaction And Product Peak Formation. Assay incubation period was 15 minutes at 37 ºC.

The outcome of the CMS 0235 (0.10 µg) assay with time vs. product formation are plotted in Figure 14, with the linear portion replotted shown in Figure 15. This assay plot illustrates the reaction incubation time range where product formation is linear with respect to

time, identifying apparent 1st order kinetics. The line eventually reaches the asymptote, meaning with increasing enzyme incubation times, product formation is no longer a function of time. This portion is known to be the 0th order kinetics region. As previously mentioned, a 15 minute incubation period was selected for CMS 0235 as at this time the enzymatic reaction was still under apparent 1st order kinetics. Looking at the time dependent assay plot (Figure 16) for CMS 1086, the x-axis (time in minutes) covers a much larger range (0-285 minutes). The much longer incubation time was necessary for CMS 1086 in order to detect product. Figure 17 is the linear portion where CMS 1086 is also behaving under apparent 1st order kinetics. The time selected to perform further assays was 150 minutes. Table 2 outlines the enzyme parameters chosen to perform further assays. All assays were executed with final reaction volumes of 20 μL.

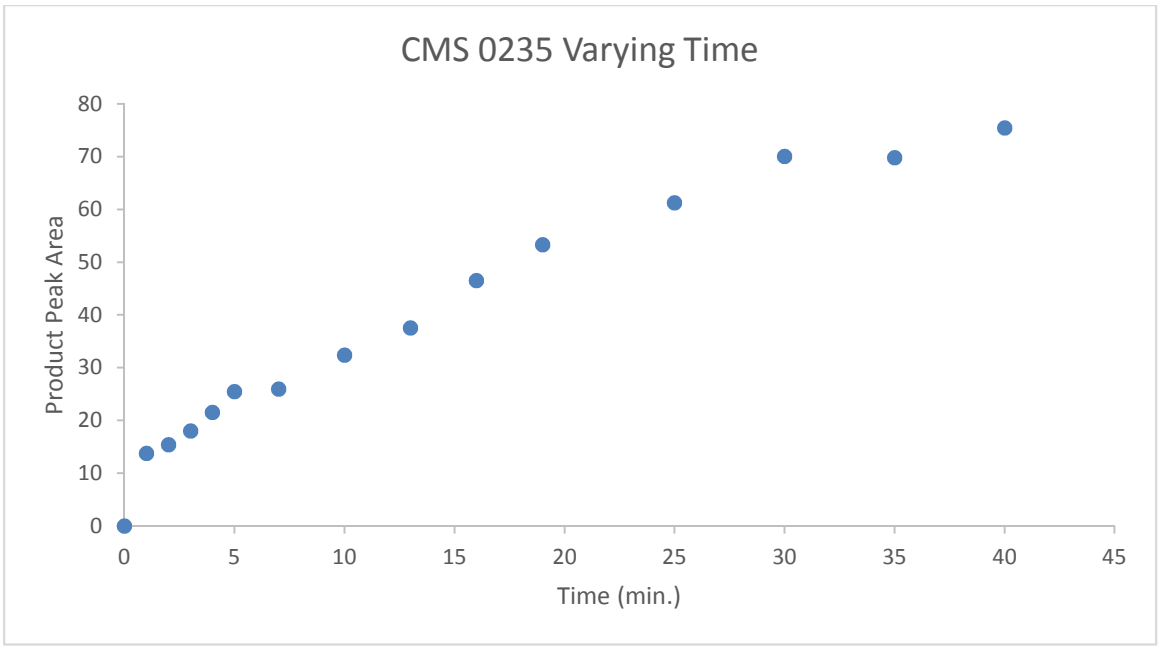

Figure 14. Product Formation vs. Time for CMS 0235. Product Peak Area is plotted as a function of time for CMS 0235

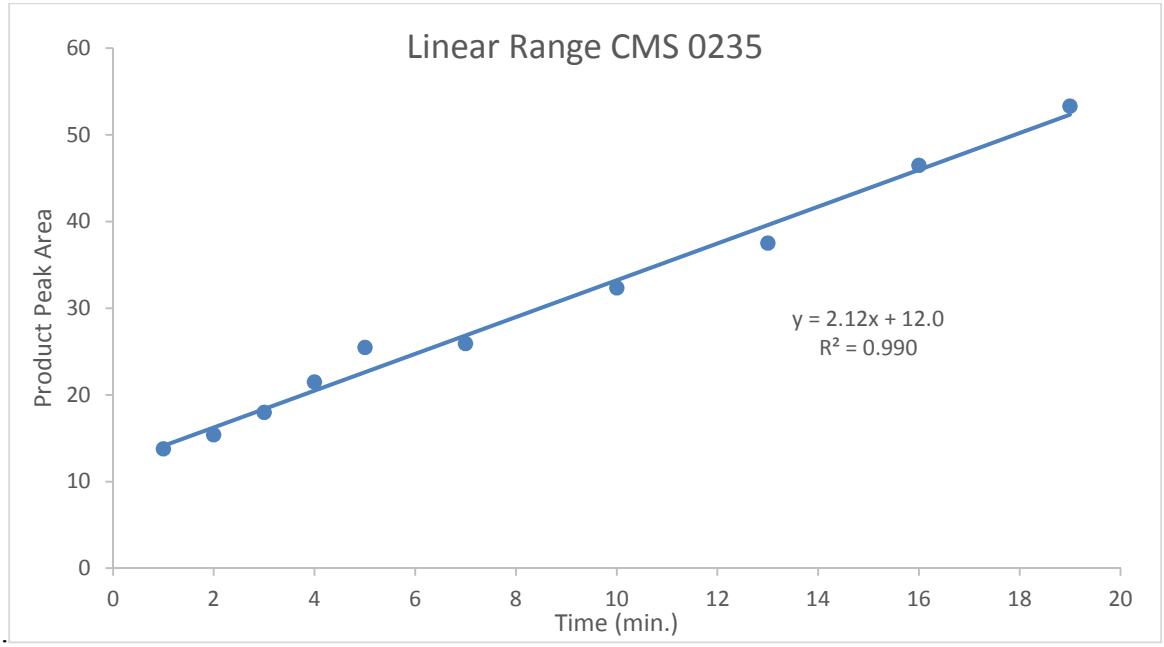

Figure 15. Linear Region of CMS 0235 Varying Time. This is the region where CMS 0235 is displaying 1st order kinetics behavior (time 0-20 minutes).

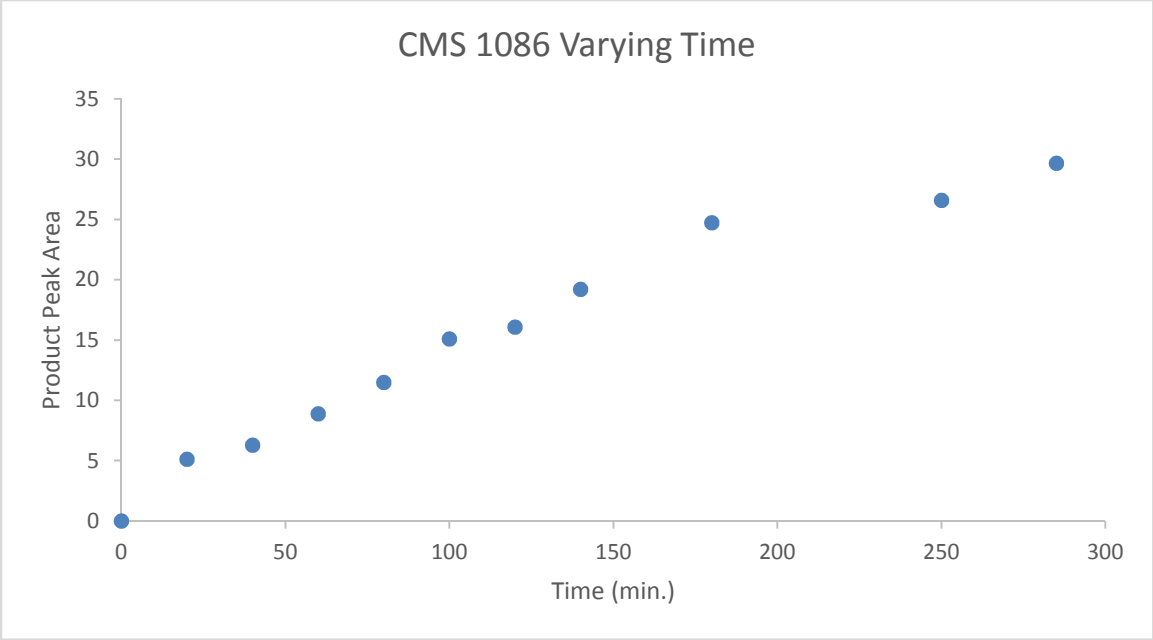

Figure 16. Product Formation vs. Time for CMS 1086. Product Peak Area is plotted as a function of time for CMS 1086.

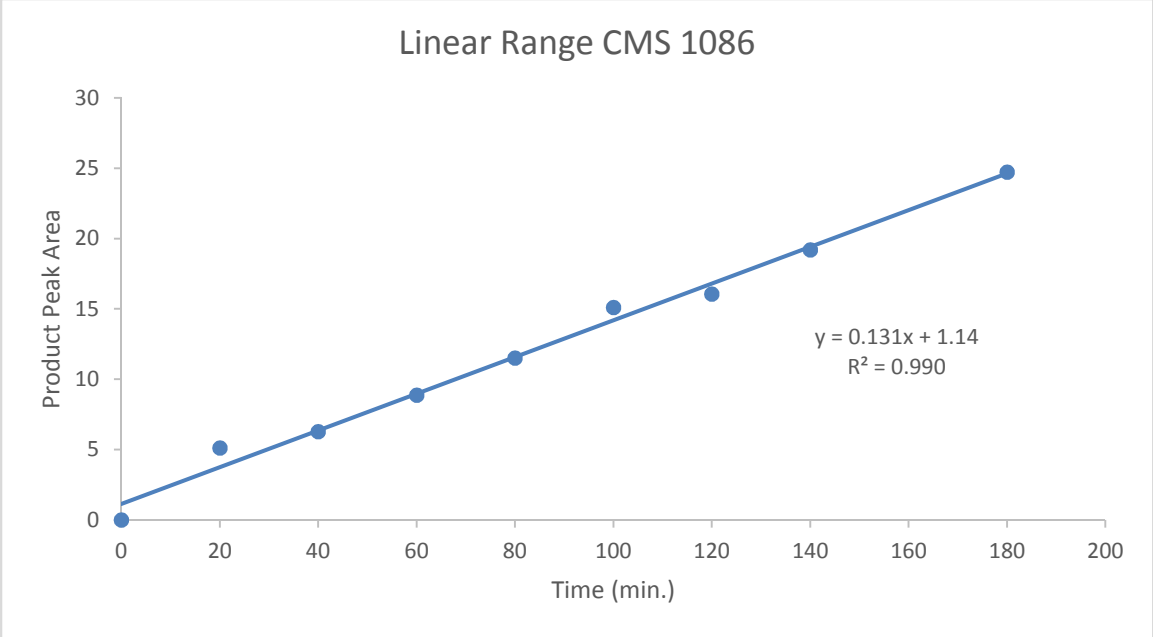

Figure 17. Linear Region of CMS 1086 Varying Time. This is the region where CMS 1086 is displaying 1st order kinetics behavior (time 0-180 minutes).

	Time of 1st order kinetics (min.)	Amount of Enzyme (µg)
CMS 0235	15	0.1
CMS 1086	150	1.0

Table 2: Overview Of Selected Assay Time Periods And Amount Of Enzyme Used. Final reaction volumes were 20 µL assay for CMS 0235 and 1086.

Kinetic assays were completed with the varying substrate in individual assays to determine K_m and V_{max} values for each substrate and enzyme. The first assay was conducted with a constant concentration of MEP (2000 µM) and CMS 0235 (0.1 µg) while varying concentration of CTP in individual reaction tubes (Figure 18). A second assay was conducted, this time with a constant concentration of CTP (2000 µM) and enzyme (0.1 µg) while MEP was varied (Figure 19). Data were collected and analyzed using the Hill equation to determine ${}^{CTP}K_m$ to be 329 ± 66.0 µM and ${}^{CTP}V_{max}$ was 2333 ± 230 nmol/min/mg of CMS0235. The Hill coefficient of the plot for varying CTP was reported to be 1.31, implying positive cooperativity for binding of CTP. For the substrate MEP ${}^{MEP}K_m$ was 187 ± 7.37 µM and ${}^{MEP}V_{max}$ was 2631 ± 48.6 nmol/min/mg of CMS0235. The Hill coefficient for MEP was calculated to be 1.96, thus implying strong evidence of a positive cooperative binding mode of MEP.

Evaluation of CMS 1086 was performed similarly to CMS 0235. Upon data analysis using the Hill equation ${}^{CTP}K_m$ was 950 ± 593 µM and ${}^{CTP}V_{max}$ was 19.2 ± 4.84 nmol/min/mg of CMS1086 (Figure 20). Results with the varying MEP assay determined ${}^{MEP}K_m$ to be 763 ± 337 µM and ${}^{MEP}V_{max}$ was 31 ± 5.42 nmol/min/mg of CMS1086 (Figure 21). Hill coefficients for varying CTP and MEP were 0.878 and 0.877 respectively. Both values are close to 1.0 implying both substrates are behaving in a little to non-cooperative binding mode.

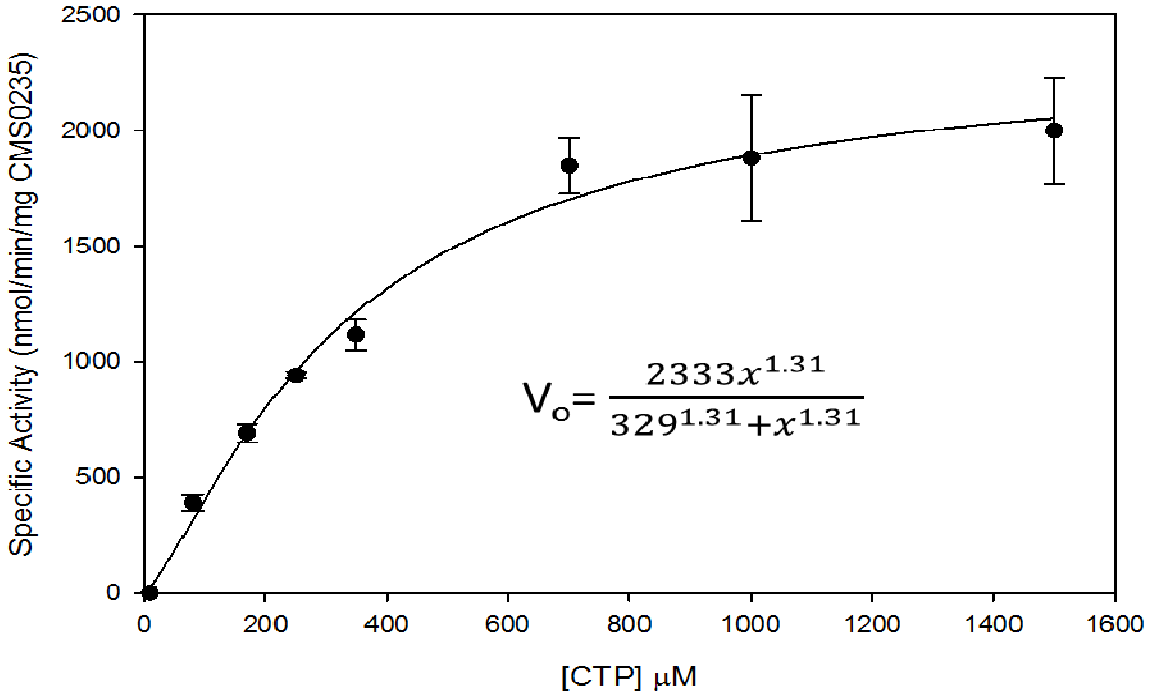

Figure 18. Hill Equation Plot for CMS 0235 with Varying Concentrations of CTP. Each data point is a mean ∓ standard error for n=3 replicates.

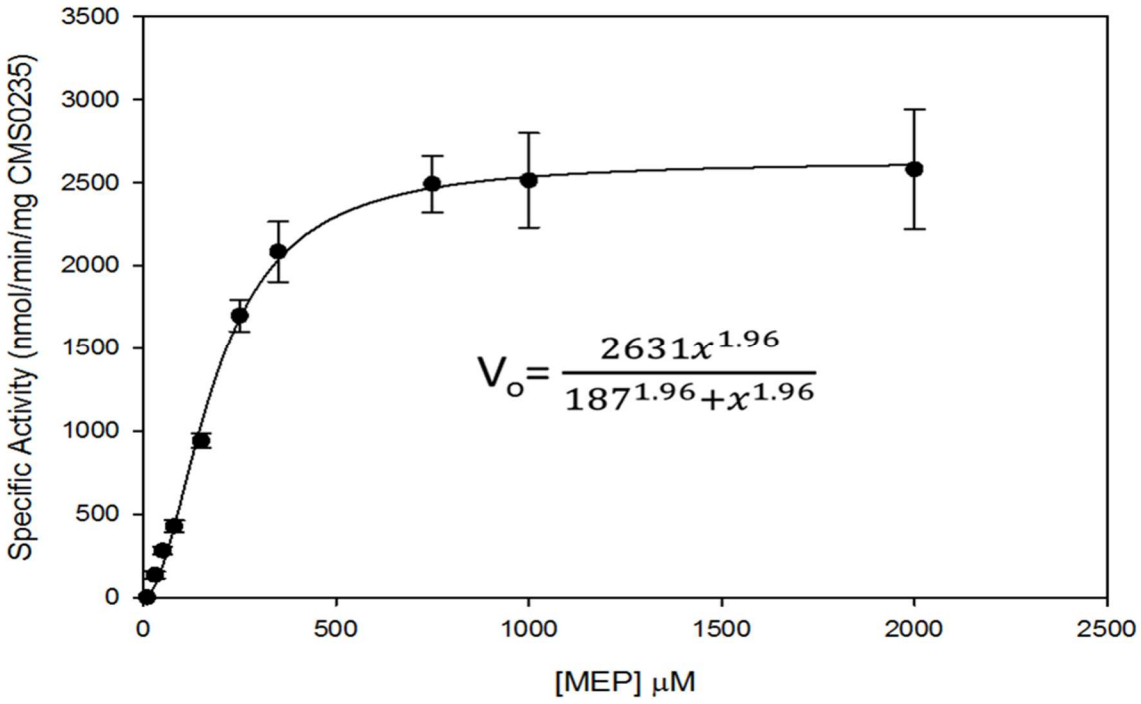

Figure 19. Hill Equation Plot for CMS 0235 with Varying Concentrations of MEP. Each data point is a mean ∓ standard error for n=3 replicates.

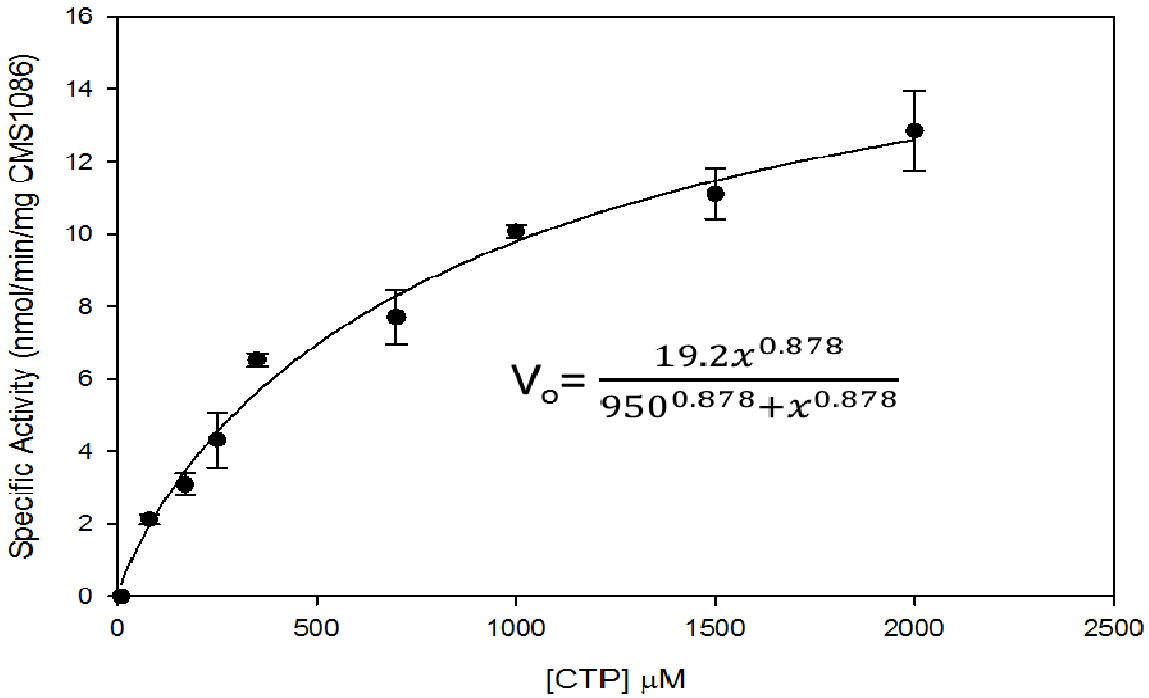

Figure 20. Hill Equation Plot for CMS 1086 with Varying Concentrations of CTP. Each data point is a mean ∓ standard error for n=3 replicates.

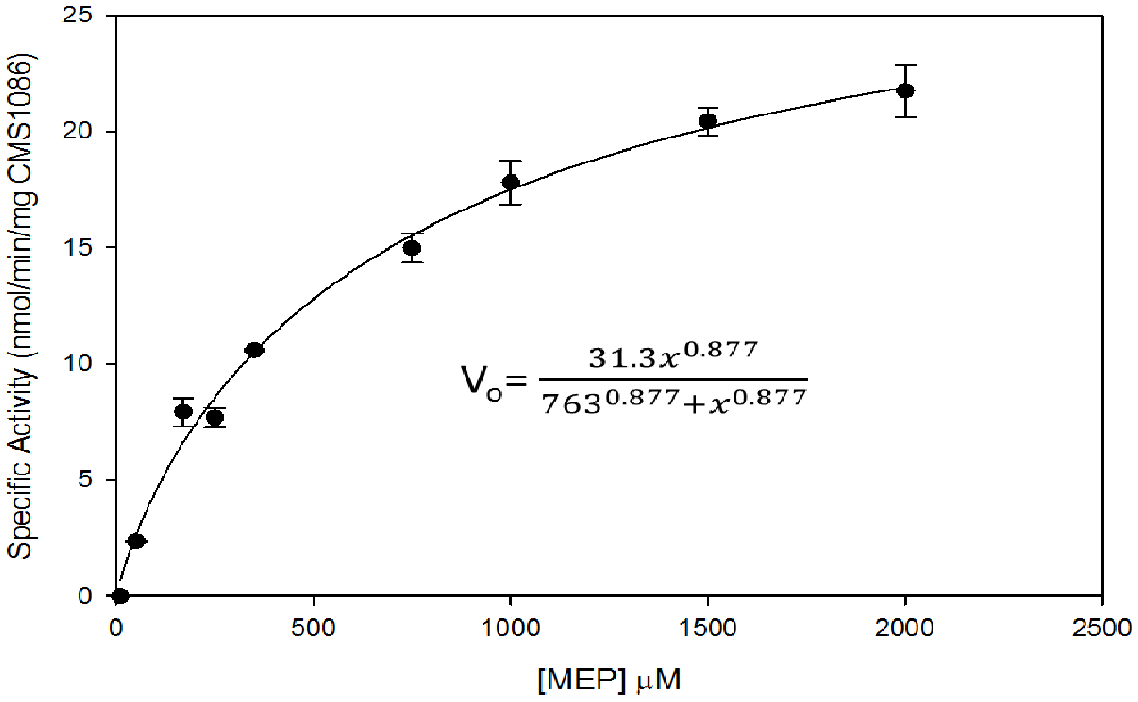

Figure 21. Hill Equation Plot for CMS 1086 with Varying Concentrations of MEP. Each data point is a mean ∓ standard error for n=3 replicates.

Lineweaver-Burk Transformation

Prior to the accessibility of powerful non-linear regression software (SigmaPlot), alternative methods were employed for determination of kinetic constants K_m and V_{max}. These methods are known as the Lineweaver-Burk, Eadie Hofstee, and Hanes Woolf plots, which were simple and useful methods for treatment of enzyme kinetics data. The Lineweaver-Burk plot is simply a double reciprocal plot of the Michaelis-Menten, thus yielding the equation shown below. The Lineweaver-Burk is a linear representation of the initial Michaelis-Menten plot and by using the Lineweaver-Burk (L.W.B) equation in $y = mx + b$ format, kinetic constants of K_m and V_{max} may be easily extrapolated. Figures 22-25 are L.W.B transformations for both enzymes and substrates.

$$\text{Lineweaver-Burk: } \frac{1}{V_o} = \frac{K_m}{V_{max}}\frac{1}{[S]} + \frac{1}{V_{max}}$$

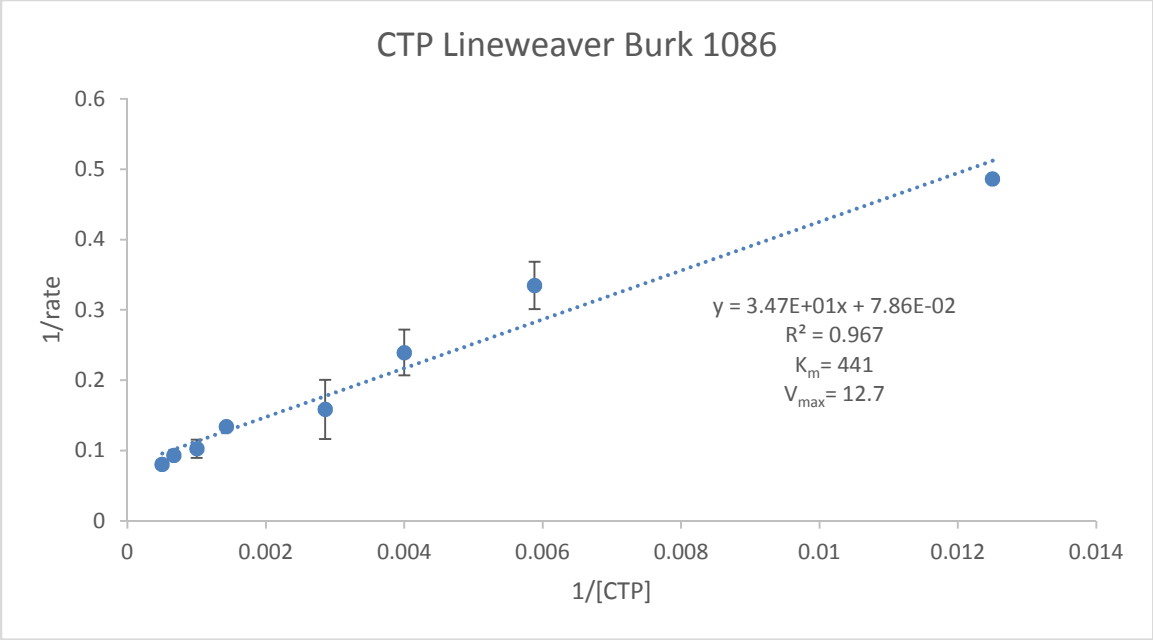

Figure 22. Lineweaver-Burk Transformation Plot for CMS 1086 for Substrate CTP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

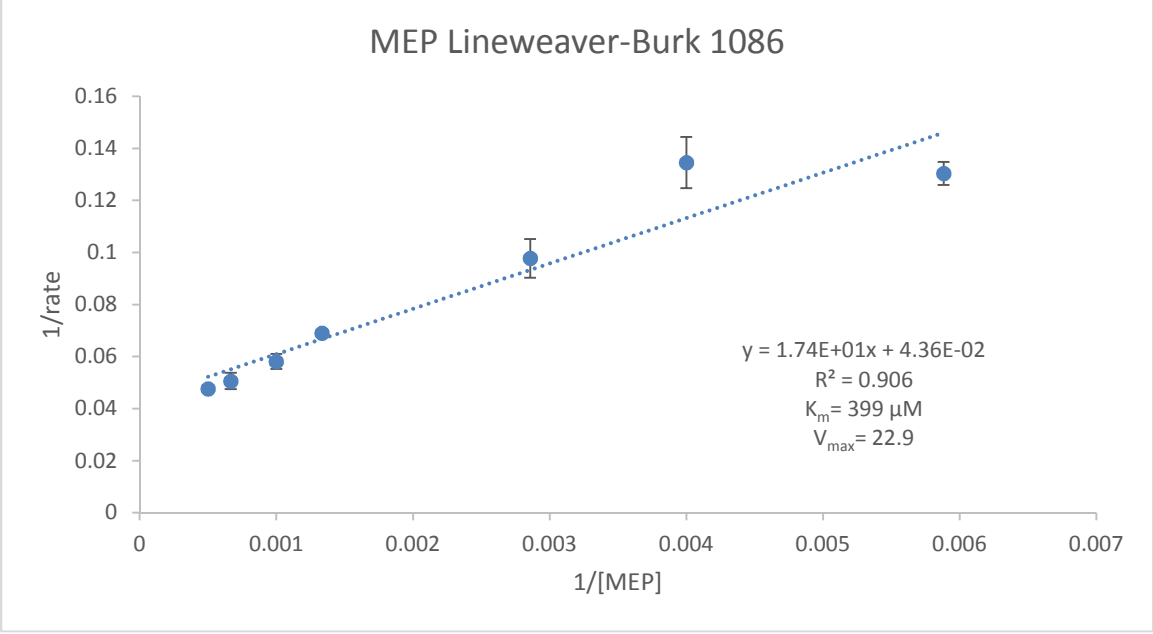

Figure 23. Lineweaver-Burk Transformation Plot for CMS 1086 for Substrate MEP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

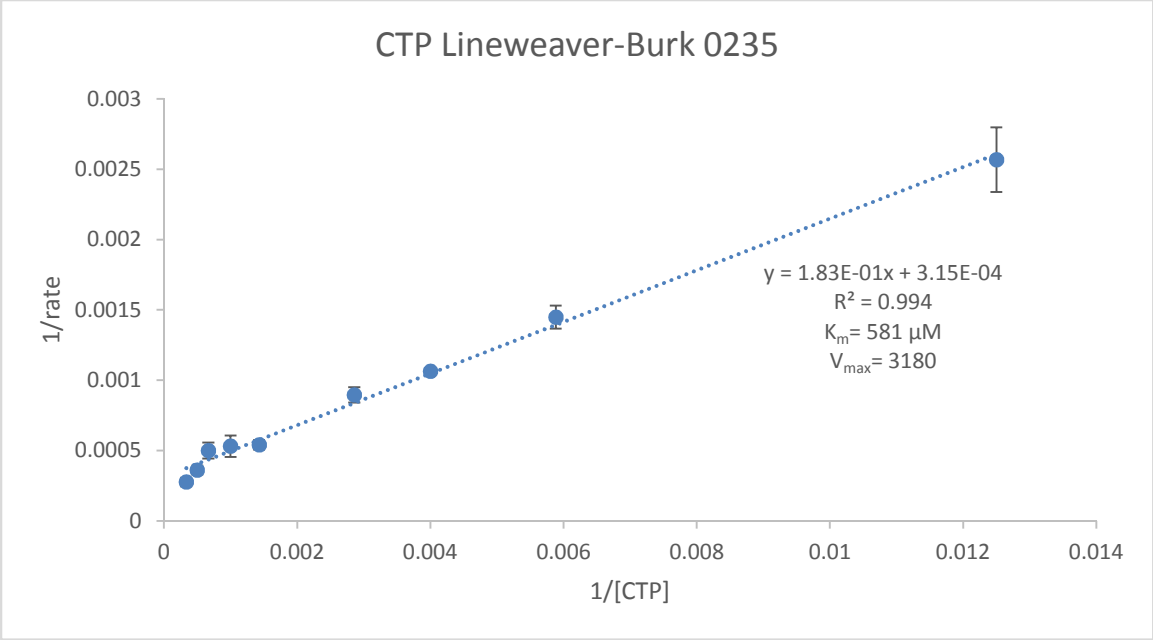

Figure 24. Lineweaver-Burk Transformation Plot for CMS 0235 for Substrate CTP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

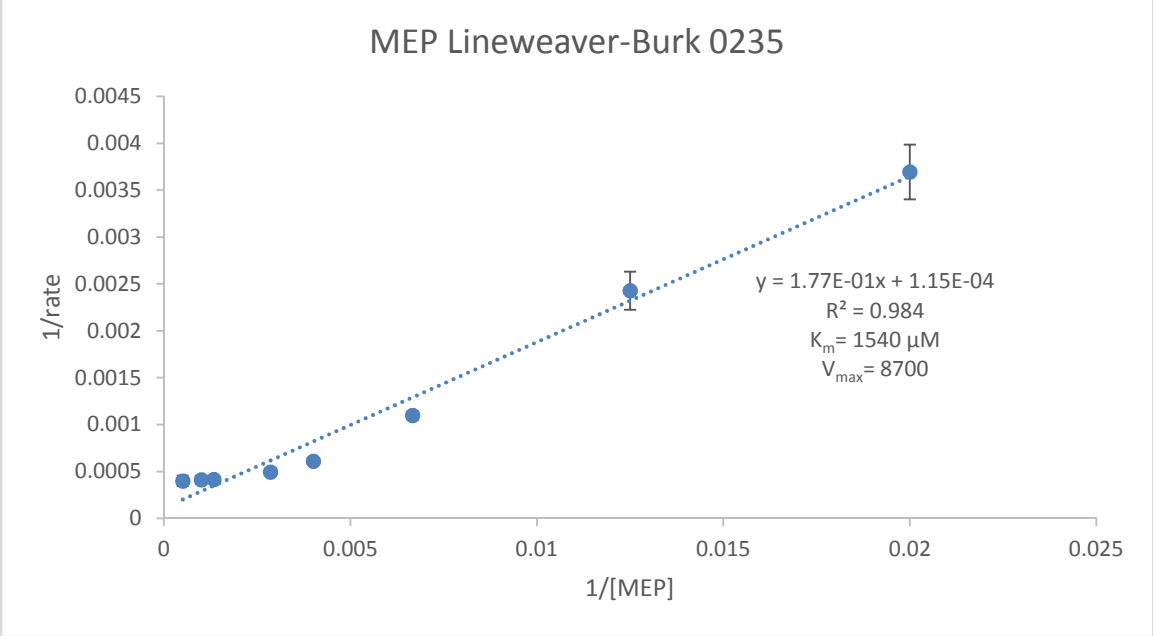

Figure 25. Lineweaver-Burk Transformation Plot for CMS 0235 for Substrate MEP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

Eadie Hofstee Transformation

An alternative option to data analysis can be obtained using another popular linear transformation plot called the Eadie-Hofstee plot. Plotting the reaction rate vs. the ratio of rate and substrate concentration will potentially yield a linear plot. The plot will provide us with K_m as the negative slope of the line and V_{max} is the y-intercept. Calculated kinetic parameters of K_m and V_{max} using the Eadie-Hofstee equation are shown in Table 3. Plots for both enzymes and substrates are shown in Figures 26-29. In Figure 25, the R^2= 0.426 and would not appear to be a good option to extrapolate kinetic constants from, although in this case these constants did seem to agree with other methods shown in Table 3.

$$\text{Eadie-Hofstee} : V_o = -K_m \frac{V_o}{[S]} + V_{max}$$

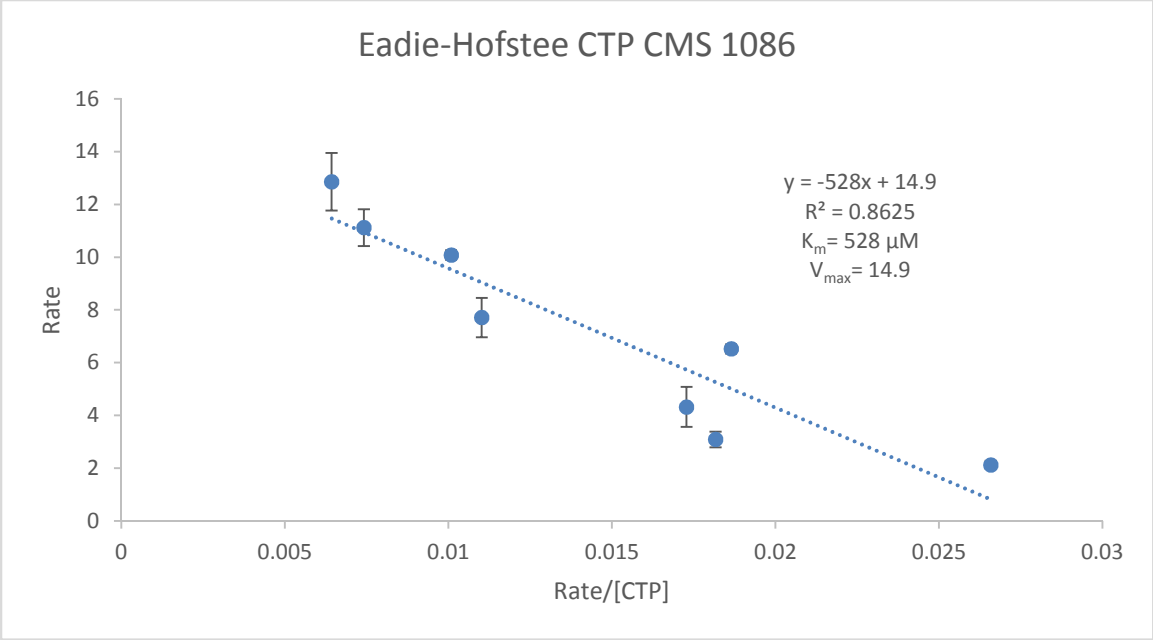

Figure 26. Eadie-Hofstee Transformation Plot for CMS 1086 for Substrate CTP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

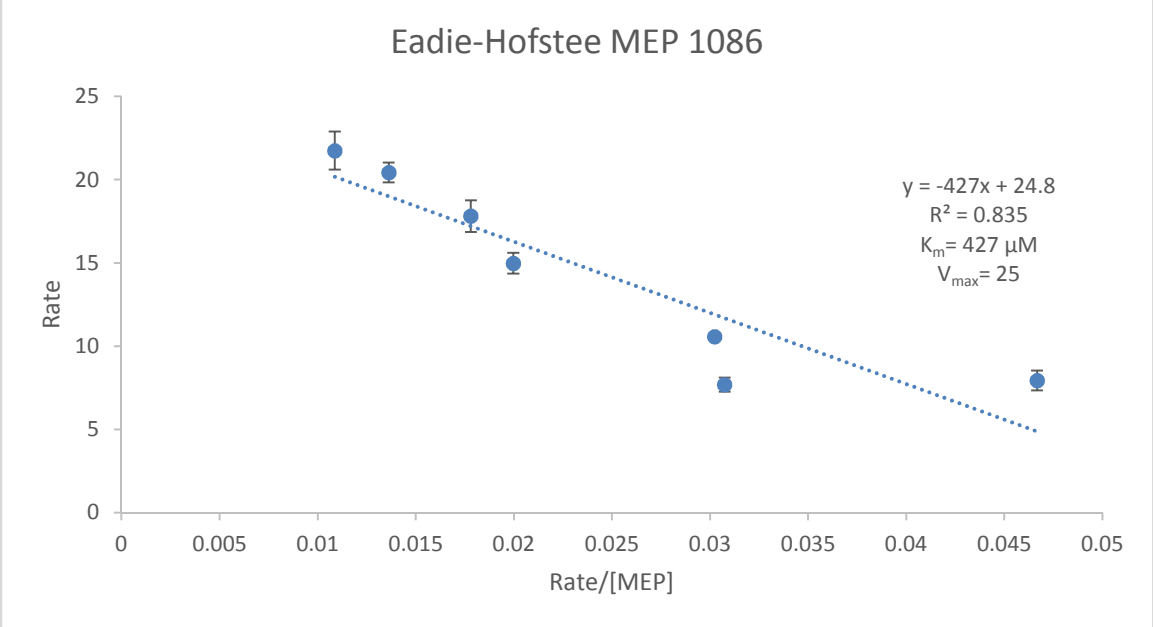

Figure 27. Eadie-Hofstee Transformation Plot for CMS 1086 for Substrate MEP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

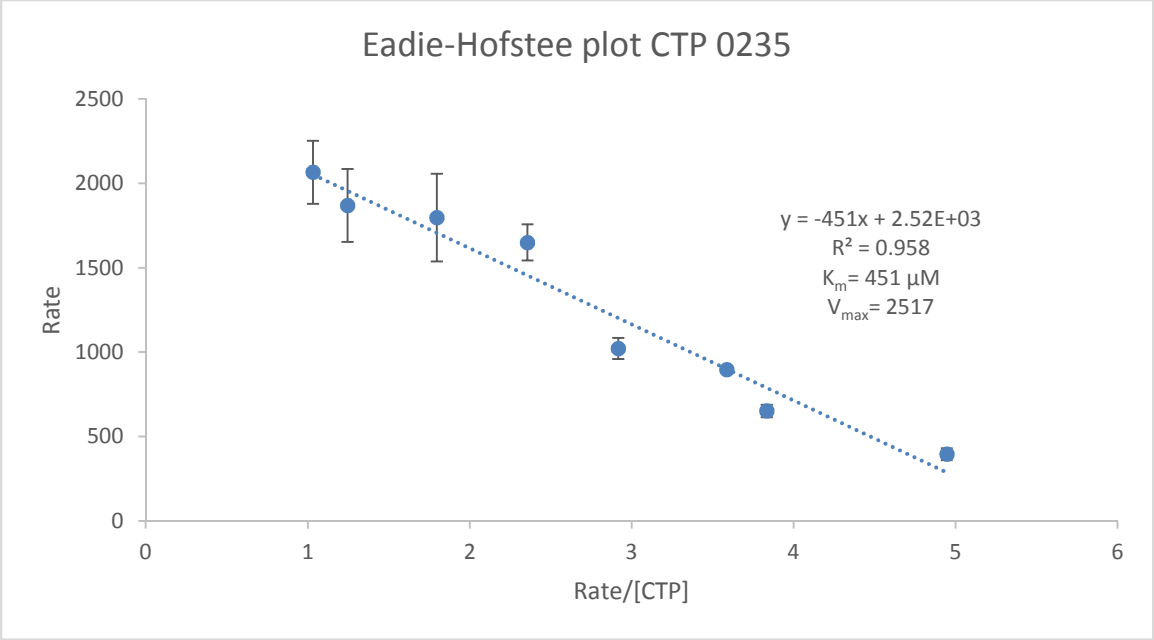

Figure 28. Eadie-Hofstee Transformation Plot for CMS 0235 for Substrate CTP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

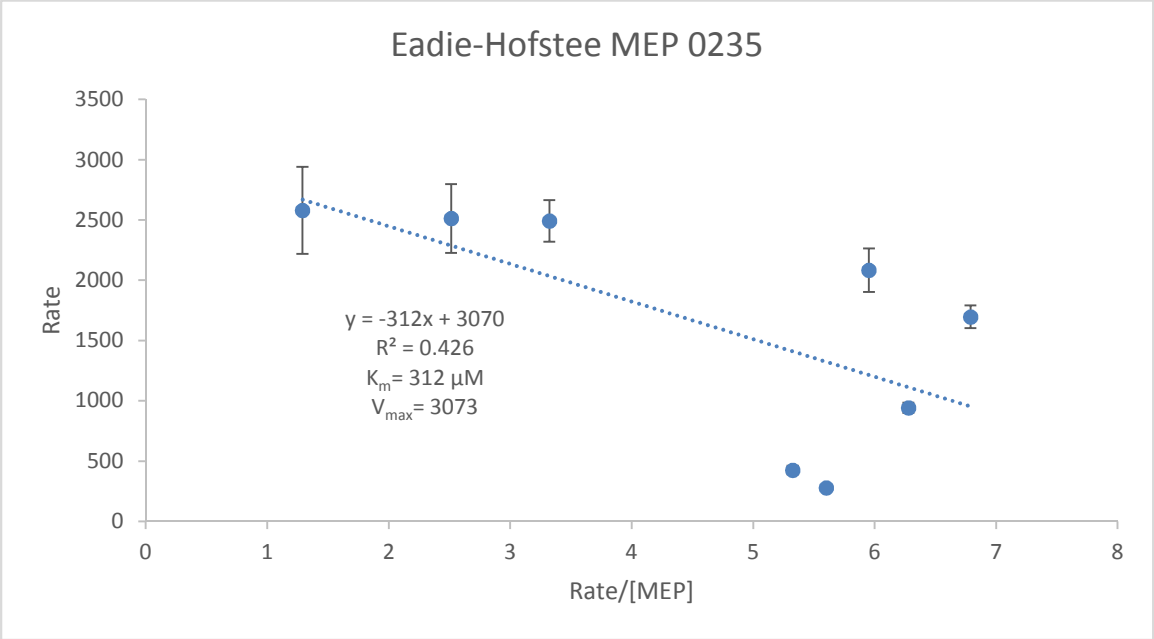

Figure 29. Eadie-Hofstee Transformation Plot for CMS 0235 for Substrate MEP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

Hanes-Woolf Transformation

The Hanes-Woolf plot is another linear transformation plot of the Michaelis-Menten equation for rapid determination of K_m and V_{max}. This plots the ratio of substrate concentration and reaction rate vs. the substrate concentration. V_{max} is equal to $\frac{1}{slope}$ of the line and the x-intercept of the plot is equal to $-K_m$. Refer to Table 3 for the values calculated using this method. Hanes-Woolf plots were created for both enzymes and substrates (Figures 30-33).

$$\text{Hanes-Woolf}: \frac{[S]}{V_o} = \frac{1}{V_{max}}[S] + \frac{K_m}{V_{max}}$$

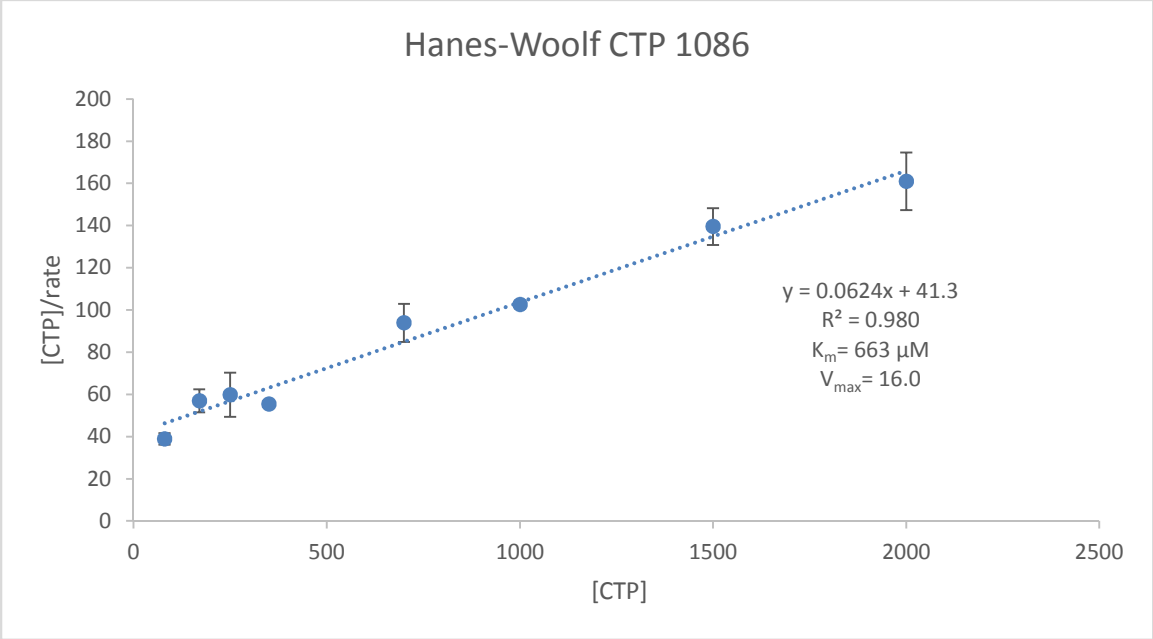

Figure 30. Hanes-Woolf Transformation Plot for CMS 1086 for Substrate CTP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean $\mp$ standard error for n=3 replicates.

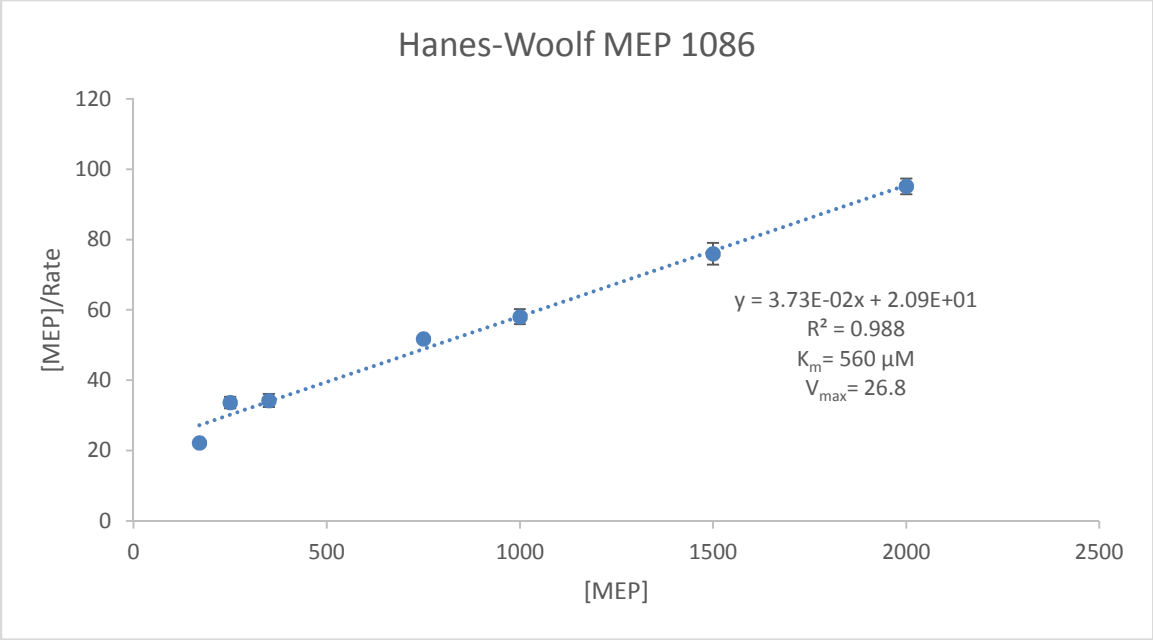

Figure 31. Hanes-Woolf Transformation Plot for CMS 1086 for Substrate MEP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

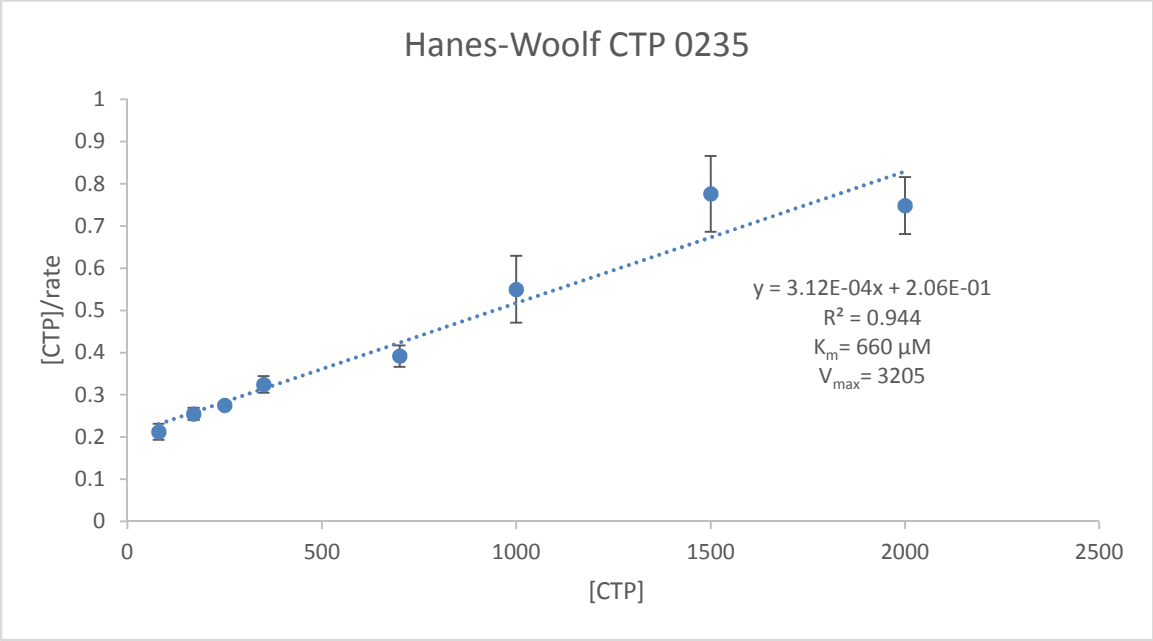

Figure 32. Hanes-Woolf Transformation Plot for CMS 0235 for Substrate CTP. Units of V_{max} are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

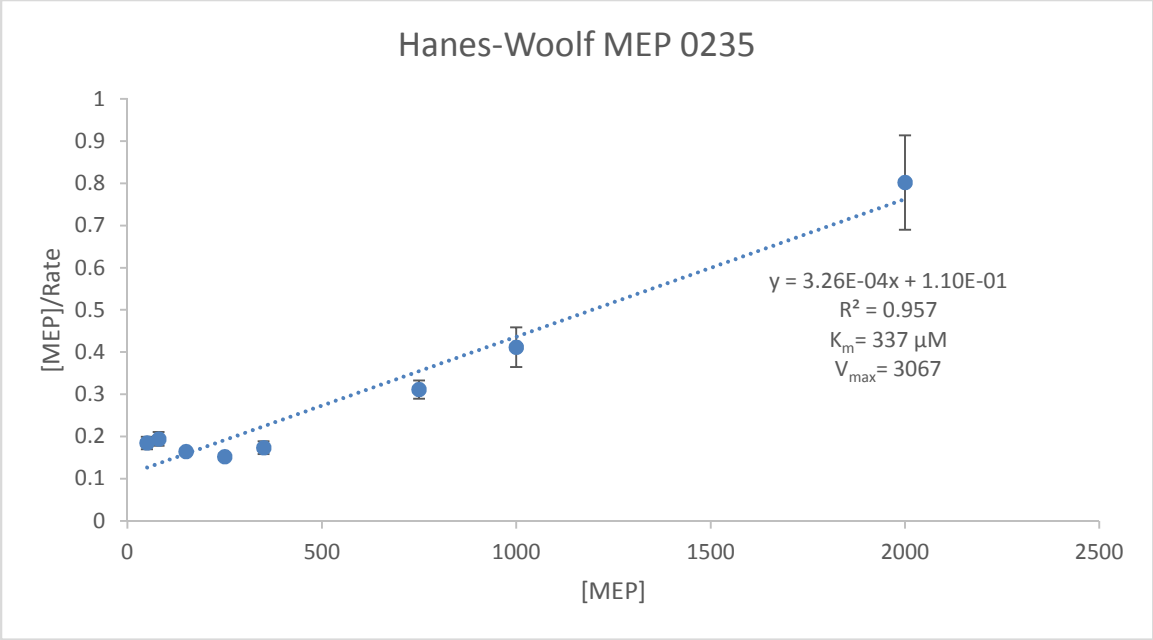

Figure 33. Hanes-Woolf Transformation Plot for CMS 0235 for Substrate MEP. Units of V_{max} and rate are specific activity (nmol/min/mg of enzyme). Each data point is a mean ∓ standard error for n=3 replicates.

The use of linear transformation plots have been superseded by powerful non-linear regression software like SigmaPlot, as they provide far more accurate results. Table 3 represents the various methods used for determining values K_m and V_{max}. Most of the values of K_m and V_{max} for CMS 0235 and 1086 are in agreement with each other despite using multiple methods. An outlier was observed using the Lineweaver-Burk equation in determining $^{MEP}K_m$ and $^{MEP}V_{max}$ for CMS 0235. This is an example of inaccuracies in linear transformation methods compared to non-linear regression analysis. The L.W.B indicates a $^{MEP}K_m$ value of 1540 µM for CMS 0235, and upon inspection of Figure 25, it is certainly obvious this value places $^{MEP}K_m$ at the asymptotic portion of the curve. A feature of the L.W.B plot is the lower concentration points on the Michaelis-Menten curve are heavily weighted in determination of the slope and trend line in the L.W.B plot. Also, CMS 0235 indicates a positive cooperative binding curve with MEP,

which introduces some complexity to accurately determining K_m and V_{max} values using linear transformation methods.

	CMS 0235				CMS 1086			
	$^{MEP}V_{max}$	$^{MEP}K_m$	$^{CTP}V_{max}$	$^{CTP}K_m$	$^{MEP}V_{max}$	$^{MEP}K_m$	$^{CTP}V_{max}$	$^{CTP}K_m$
L.W.B	8700	1540	3180	581	22.9	399	12.7	441
Hanes-Woolf	3070	340	3210	660	26.8	560	16.0	663
Eadie-Hofstee	3070	312	2520	451	24.8	427	14.9	528
Hill Equation	2631	187	2333	329	31.3	763	19.2	950

Table 3. Calculated Values of V_{max} and K_m Using Linear Transformation Methods And The Hill Equation In SigmaPlot. All units of K_m are μM and V_{max} are specific activity (nmol/min/mg of enzyme).

Evaluation of K_m and K_{cat} of CMS 0235 and 1086

To compare CMS 0235 and 1086, these enzymes have very different activity levels as can be seen within the data. K_{cat} values were calculated from V_{max} for each enzyme and their individual substrates. K_{cat} is the turnover rate, number of product molecules formed by the enzyme per unit of time. The CMS 0235 $^{MEP}K_{cat}$ is 1.13 s^{-1} and the CMS 1086 $^{MEP}K_{cat}$ is 0.0143 s^{-1}, which suggests that CMS 0235 converts MEP to product approximately 100x faster than CMS 1086. A similar observation is made with CTP as $^{CTP}K_{cat}$ for CMS 0235 and 1086 are 1.00 s^{-1} and 0.0088 s^{-1} respectively.

Another worthy comparison are the K_m values for each of the enzymes, as they provide information on enzyme substrate binding affinity. Equation 1 shows k_1 as the association constant and k_{-1} is the dissociation constant for typical enzymes. Since k_2 is assumed to be the rate limiting step and thus very small in comparison to k_1 and k_{-1}, the equation is allowed to be simplified by removing k_2. This now shows that K_m is directly related to the association and dissociation constants for an enzyme which correlates to binding affinity.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Equation (1) $$K_m = \frac{(k_{-1}+k_2)}{k_1} = \frac{k_{-1}}{k_1}$$

Since the K_m values for CMS 0235 are lower for both substrates, this suggested that the enzyme may have stronger binding affinity for each substrate than the 1086 isoform. The larger K_m for CMS 1086 is an indication the enzyme is having trouble binding the substrate properly most likely due to poor amino acid to substrate interactions. This is translated into having a larger k_{-1} and/or smaller k_1. In the larger scope of enzymes, CMS 0235 and 1086 are not very efficient enzymes in respect to their turnover rates. Kinetic constants of K_m and K_{cat} can be used to calculate catalytic efficiency. An enzyme's catalytic efficiency (C.E.) is equal to $\frac{K_{cat}}{K_m}$. The ^{CTP}C.E. for CMS 0235 and 1086 were calculated to be 3040 $M^{-1}s^{-1}$ and 9.26 $M^{-1}s^{-1}$, respectively. The ^{MEP}C.E. for CMS 0235 and 1086 were calculated to be 8248 $M^{-1}s^{-1}$ and 18.7 $M^{-1}s^{-1}$, respectively.

A study was done on enzyme kinetics that used the Brenda database to obtain all K_{cat} values for 1942 various enzymes from prokaryotic and eukaryotic organisms (Bar-Even et al., 2011). The median K_{cat} was determined to be 13.7 s^{-1} which is 10x faster than CMS 0235 and is about 1000x faster than CMS 1086 at converting substrate to product. The same study obtained K_m values of over 5000 enzymes and calculated the median K_m of enzymes was 130 µM, and 60% of all enzymes had values ranging from 10-1000 µM (Bar-Even et al., 2011). In comparison to these literature enzymes, CMS 0235 ($^{CTP,\ MEP}K_m$= 329, 187 µM) and CMS 1086 ($^{CTP,\ MEP}K_m$= 950, 763 µM) have "average" values. Catalytic efficiencies were also obtained and the median was determined to be 125,000 $M^{-1}s^{-1}$. This indicates that both CMS 0235 (^{CTP,MEP}C.E= 3040, 8248 $M^{-1}s^{-1}$) and CMS 1086 (^{CTP,MEP}C.E= 9.26, 18.7 $M^{-1}s^{-1}$) are not efficient when compared to

the total enzyme pool. These parameters for an enzyme are limited by many factors such as, the type of metabolic pathway, number of substrates, nature of substrates, and the size of an enzyme. Although some of these parameters are "low" in regards to the average, it may be by a selective mechanism. All enzymes cannot improve kinetic parameters *ad infinitum*, as regulation is essential in maintaining homoeostasis. Overall, these enzymes exhibit lower enzymatic activities, but it may be that no evolutionary pressures have forced CMS enzymes to become more efficient.

pH Profile of CMS 0235 and 1086

Enzymes behave differently under changing pH conditions. pH evaluations of enzymes to obtain optimum activity are important into understanding the mechanistic behavior of substrate conversion to product. In some cases pH is a regulatory factor in enzymes and they may be up or downregulated with changing pH environments. The optimum pH for growth of *Listeria monocytogenes* is 7.0 and the pH tolerance is reported from 4.4-9.4 (Lado & Yousef, 2007). When varying pH, many things may occur, such as changes in protein structure, ionization states of amino acids, substrate stability, and substrate ionization state. Extreme pH conditions will lead to disruptions in higher orders of protein structure and may lead to protein denaturation. By varying the pH environment and modifying the ionization state of amino acid side chains, a correlation may be observed between ionized state and effects on an enzyme substrate affinity. A pH profile of both CMS 0235 and 1086 was obtained with a range of pH 4-12 (Figures 34-35). The pH profile of both enzymes was determined to be similar as they both exhibit pH optimum activity at pH 8.5. Both isoforms exhibit a pH tolerance range of about 7-10, as most of the activity is lost outside this pH range. Some unexpected results occurred in the basic conditions of the pH region. Both enzymes have a Mg^{2+} cofactor; in basic conditions Mg^{2+} is precipitated out

as $Mg(OH)_2$, thus making it inaccessible for the enzyme. Since activity is retained in basic conditions, it may be that the magnesium ion is not in rapid equilibrium with the environment but rather more tightly bound in the active site of the enzyme. A second possibility is the Mg^{2+} does not play a large role in catalysis and may not be very essential to retain enzymatic activity.

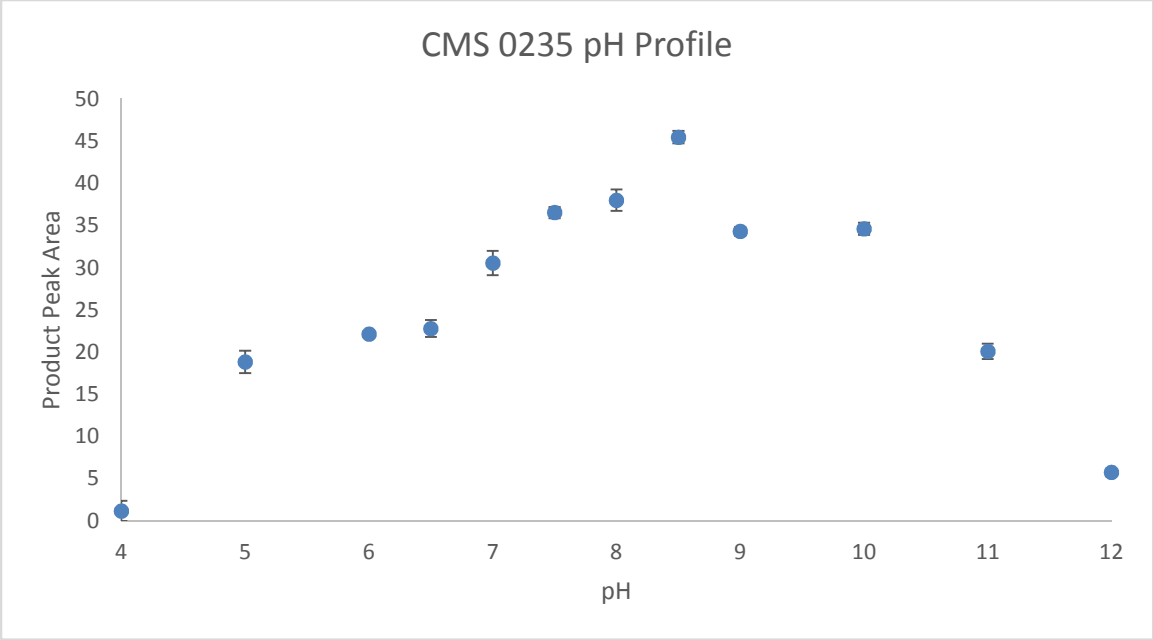

Figure 34. pH Profile for CMS 0235 in Triplicate. Product formation represented as Product Peak Area as a function of pH. Error bars represent standard error n=3 replicates

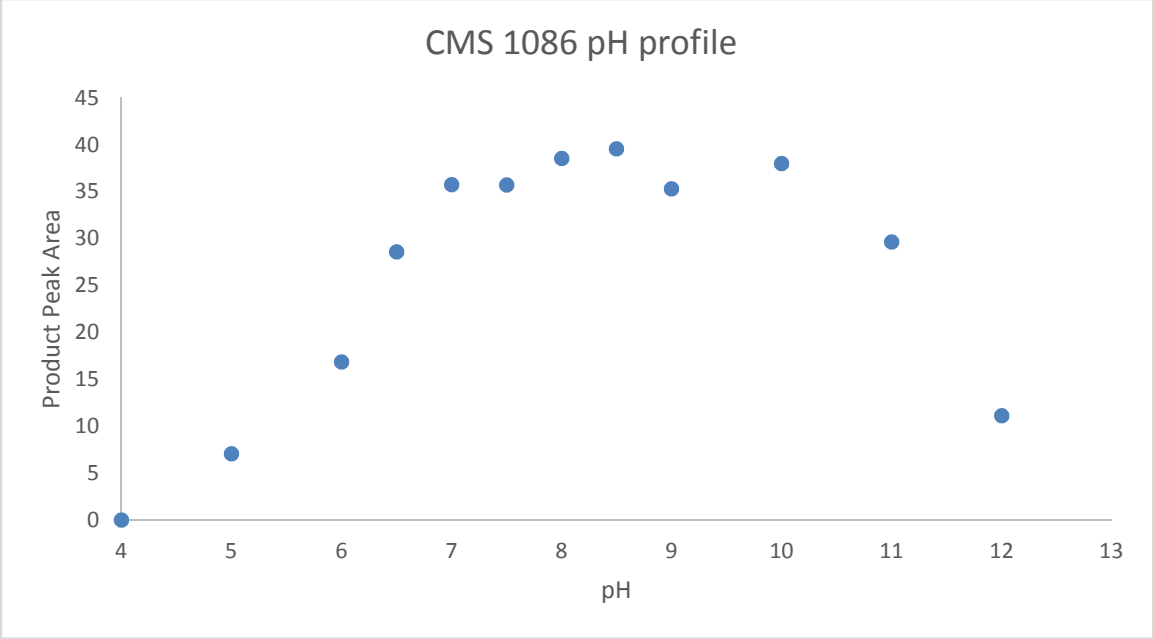

Figure 35. pH Profile for CMS 1086. Product formation represented as Product Peak Area as a function of pH.

Divalent Cation Exchange

Some enzymes require cofactors to perform catalysis. Cofactors can act in a variety of ways by enhancing protein stability, binding substrate, provide platform for oxidative/reduction chemistry and/or stabilizing transition state intermediates. Enzymes without their cofactors are called apoenzymes, which is an enzyme in its inactive form; however, not all enzymes require cofactors. In cases of fluctuating environmental conditions, certain metal cofactors may become scarce which would induce physiological stress on organisms. In some instances, the use of cations in enzymes is a tool for regulation by providing an on off switch for enzymes. When organisms grow in resource depleted environments they are often able to overcome the challenges associated with minimal resources. Some mechanisms in doing so include enzyme cofactor substitution. Cytidylyltransferases typically require a Mg^{2+} metal cofactor. Some isoforms of CMS are not responsive to various divalent cations while others are less sensitive to divalent cation substitution (Tsang et al., 2011; Gabrielsen et al., 2004). In this study we substituted the Mg^{2+} divalent cation in the assay buffer with Mn^{2+}, Ni^{2+}, Ba^{2+}, and Co^{2+} for both isoforms of CMS. CMS 0235 and 1086 both responded similarly with each substituted cation (Figure 36). Exchanging the cofactor for Ba^{2+} responded the best, displaying moderate activity relative to Mg^{2+}. Substitution of Ni^{2+} also revealed that both CMS isoforms are able to function with this metal cofactor. With Co^{2+} both isoforms displayed greatly reduced activities, suggesting this substitution is possible but not ideal. Manganese (II) resulted in the least activity, with only a mild retention of activity for CMS 0235 and no activity for CMS 1086. The results suggest these enzymes are di-cation dependent but still functional with a few other divalent cations. In times of stress without access to Mg^{2+}, *Listeria monocytogenes* is likely able to incorporate other divalent

cations into CMS for survival. This can be an attributing factor for why *L. monocytogenes* has the ability to survive in numerous environments.

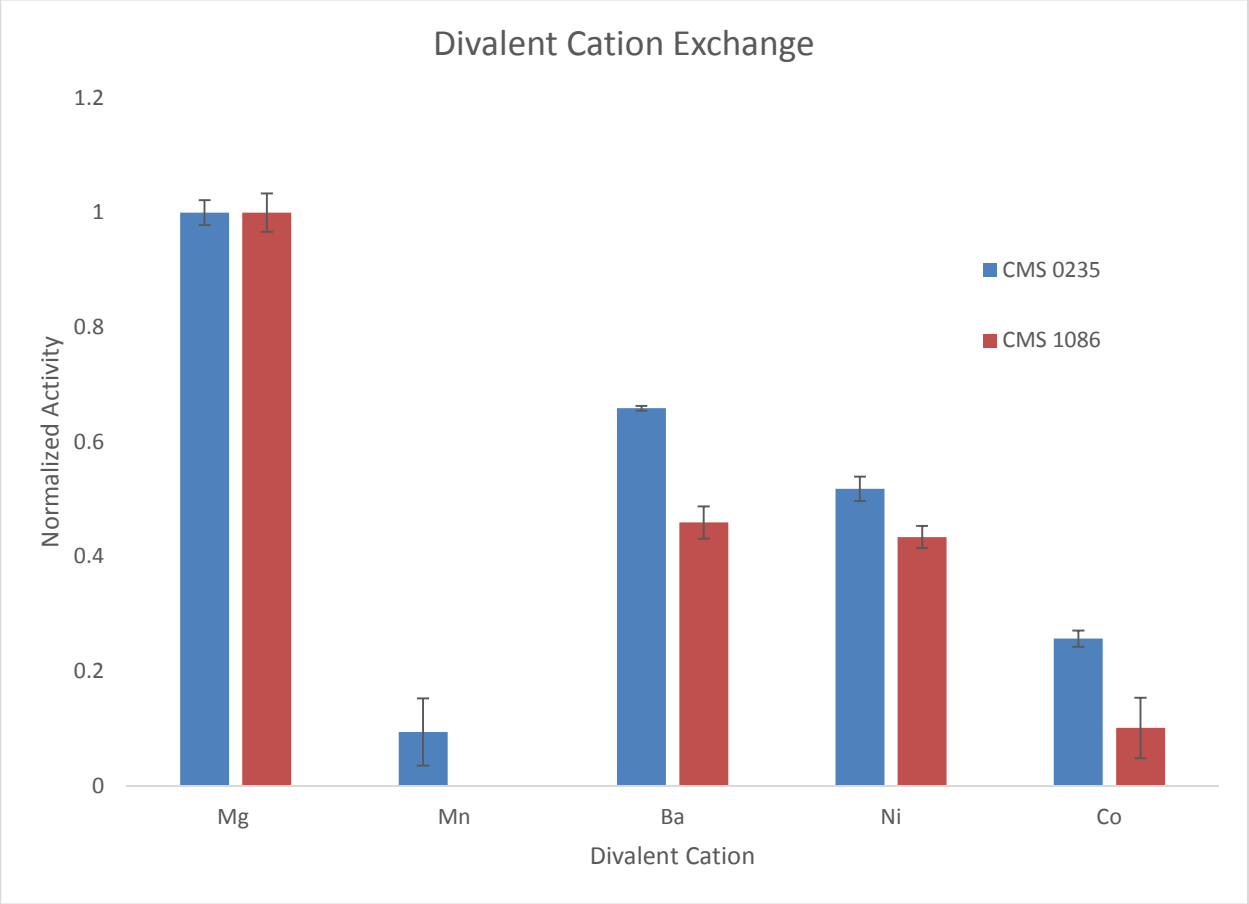

Figure 36. Divalent Cation Assay Results. Normalized activity plotted as a function of divalent cation for CMS 1086 and CMS 0235. Each concentration of cation in the reaction was 25 mM.

Sequence Analysis of CMS Enzyme Isoforms

Kinetic analysis of enzymes, although vital to development of potent inhibitors and helpful in understanding mechanisms, is not the only technique necessary for characterization of enzymes. Enzymes are typically protein molecules that are synthesized by living cells to catalyze biochemical reactions. Not all enzymes are proteins, as there are known catalytic RNA (ribozymes) that are involved in RNA processing. Proteins are composed of amino acids that are linked together by peptide bonds and range from a few amino acids to thousands. Proteins are classified into orders of structure known as primary, secondary, tertiary, and quaternary structures. The simplest form of protein structure is the sequence of amino acids, which is known as the primary structure. A secondary structure of a protein is the next level of protein structure and is a segment of the protein that folds into a set conformation based on favorable near neighbor amino acid interactions. The most common secondary structures are β-sheets (parallel or antiparallel), α-helices, and β-turns. Tertiary structure is simply the three dimensional structure of a protein due to amino acid side chain and backbone interactions. Some proteins contain a fourth level of protein structure, which is when multiple subunits (which may be the same or different) of protein form a larger protein complex recognized as quaternary structure. An enzyme's functionality is contingent on the integrity of the native protein folding conformation. If an enzyme is denatured so that some of the orders of protein structure are disrupted, activity may be lost. Some enzymes also contain cofactors (inorganic ions), coenzymes (vitamins, organic nutrients), and covalent modifications (phosphorylation, glycosylation) to regulate enzyme activity. Enzymatic activity is correlated with protein structure, and sequence analysis was conducted to elucidate possible explanations for varying

activities between isoforms of CMS. Refer to Table 4 for sequence alignment identification and locations.

Sequence alignment ID	PDB Code	Accession Number	Organism	Enzyme	Figure location
lmo0235	-	CAD00762	*L. monocytogenes*	CMS	38,39,40,46
lmo1086	-	CAC99164	*L. monocytogenes*	CMS	38,39,40,46
lmo3F1C	3F1C	CAC99164	*L. monocytogenes*	CMS	38,39,40,46
OUT00427	-	OUT00427	*L. planturm*	RCT	46
Camjej	-	AL139079	*C. jenjuni*	Hypothetical ispDF	40
Myctu		NP_218099	*M. tuberculosis*	CMS	40
Fratul	-	CAJ79964	*F. tularensis*	CMS	40
Athal	-	AAF43207	*A. thaliana*	CMS	40
Ecol, EcoliCM	1INI	1INI_A	*E.coli*	CMS	38,39
Arabdop	1W77	1W77_A	*A. thaliana*	CMS	36
Bsub5HS2	5HS2	5HS2_B	*B. subtilus*	CMS	46
Atal2YC3	2YCS	2YC3_A	*A. thaliana*	CMS	46
Ecol1VGT	1VGT	1VGT_A	*E. coli*	CMS	46
Bsub4JIS	4JIS	4JIS_A	*B. subtilus*	RCT	46
Spn2VSH	2VSH	2VSH_A	*S. pneumonicos*	RCT	46

Table 4. Legend for Sequence Alignments And Their Location. PDB codes are given for known structures and accession numbers are provided. Also provided is the organism the enzyme is from and the enzyme type (CMS or RCT). ispDF is a bifunctional enzyme, a product of gene fusion (ispD + ispF). IspF is 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase or MCS from the MEP pathway.

Structures of CMS 0235 and 1086 from *Listeria monocytogenes* have not been reported, so a multiple sequence alignment was constructed with other known structural isoforms of CMS (Figure 38). Refer to Table 6 for key amino acid equivalents. Amino acids have varying properties, nonpolar, polar, aromatic, positively charged and negatively charged (Figure 37). Many of the amino acids for CMS across the three organisms (*Listeria monocytogenes, E. coli,* and *Arabidopsis thaliana*) are not "conserved", meaning each protein contains the same amino acid at an analogous location. This shows that at many locations on an enzyme an amino acid can be mutated and have little to no effect on enzyme activity. The regions in black from all sequence alignments specify highly conserved amino acids at that specific location, and grey regions represent locations of similar amino acids. These conserved areas indicate amino acids that have a higher probability of being essential in catalysis, which may include binding of substrates (or coenzymes/cofactors) and/or providing protein structural features necessary for proper folding.

Blue Amino Acids

Amino acid locations highlighted in blue from Figure 38 are vital to the enzyme's function and are theorized to be the active site amino acids responsible for catalysis (Richard et al., 2004). A mutant study was conducted by mutating each amino acid highlighted in blue, which resulted in little to no activity for the *E. coli* isoform of CMS (Richard et al., 2004). The study concluded that Arg 20, Lys27, Arg167 and Lys213 are largely involved in the stabilization of the intermediate (Richard et al., 2004). The R20K mutant lost 55% activity while R20A retained full activity. The authors comment that the Mg^{2+} is able to stabilize the intermediate and perform the role of Arg20 when it is absent (Richard et al., 2004). Amino acids Thr140, Asp106, Thr165, and Arg109 contribute to the binding and

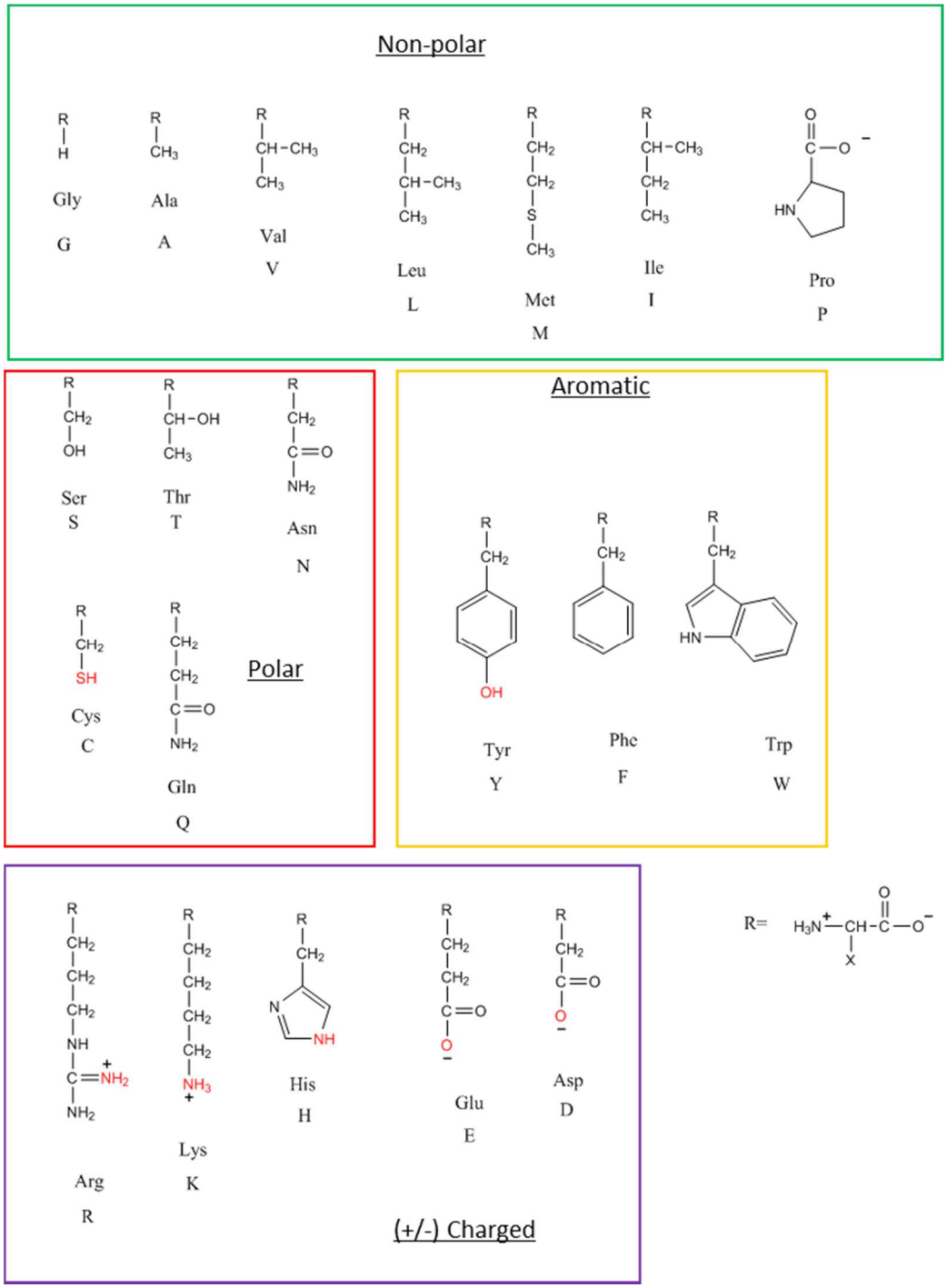

Figure 37. Amino Acid Structures and Their Respective Three and One Letter Codes. Amino acids are displayed as their dominant species at pH 7. Red locations indicate side chains with ionizable functional groups. Structures were made using ChemDraw.

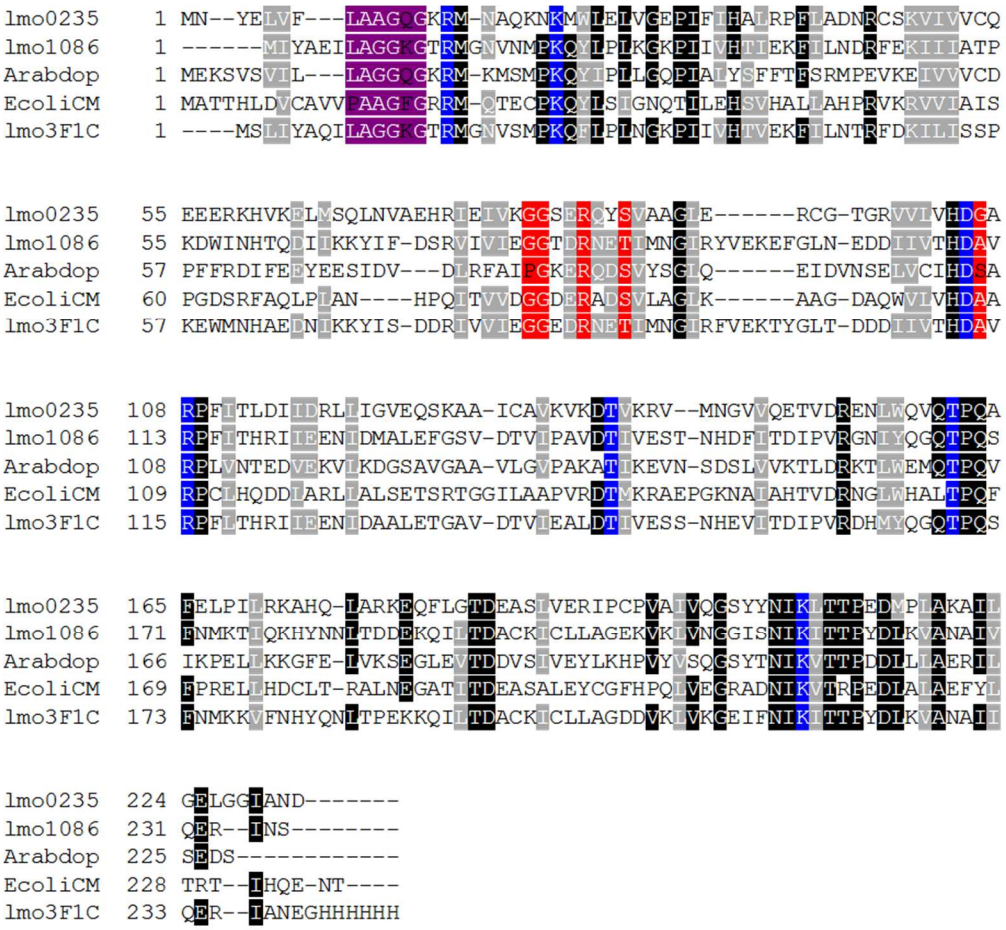

Figure 38. Multiple Sequence Alignment for Various Isoforms of CMS. lmo0235 is CMS 0235, studued in this thesis (CAD00762). lmo1086 is CMS 1086, studied in this thesis (CAC99164). lmo3F1C is a putative CMS structure (PDB: 3F1C). Arabdop is a CMS isoform in *Arabidopsis thaliana* (PDB: 1W77). EcoliCM is a CMS isoform from *E. coli* (PDB: 1INJ). Amino acids highlighted in blue play major roles in catalysis and have been studied by site directed mutagenesis (Richard et al., 2004). Red amino acids are involved in substrate binding of CTP but have not been tested by site directed mutagenesis. Purple amino acids make up the glycine rich loop that also plays a role in CTP binding. Purple amino acids also have not been tested by site directed mutagenesis. Refer to Table 4 for sequence alignment information. Sequence alignments were constructed using kalign (https://www.ebi.ac.uk/Tools/msa/kalign/) and boxshade (https://embnet.vital-it.ch/software/BOX_form.html).

stabilization of the substrate methylerythritol phosphate (MEP). Note that Thr140 is part of the interlocking arm and buries itself in the neighboring subunit.

Red Amino Acids

Amino acids highlighted in red in Figure 39 are amino acids involved in binding of substrate CTP (ribose and cytidine moiety only) (Richard et al., 2004). Amino acids in red have not been studied by site directed mutagenesis. These amino acids provide substrate selectivity for CTP versus other nucleotide triphosphates but are not associated with MEP binding. These selected amino acids require further analysis.

Purple Amino Acids

A glycine rich loop is relatively conserved in CMS enzymes, highlighted in purple in Figure 36. This location is known to play a role in binding substrate through backbone interactions with the cytidine and ribose moiety. In some cases backbone interactions do not need a conserved amino acid as the backbone of a protein is presumably the same regardless of amino acid. In this region it may be an exception since glycine and smaller amino acids lead to secondary structure that other amino acids are unable to provide. This may cause a proper alignment for backbone substrate interactions. In this thesis, in order to gain further details on substrate binding and catalytic function, analysis was performed using multiple isoforms of CMS that have been characterized.

Summary of Results

This section serves the purpose of identifying regions of significant amino acid residues that are involved in CMS enzyme activity. The blue amino acids are found to be critical for enzyme catalysis. Mutagenesis studies revealed that activity is severely compromised when mutating blue amino acids (Richard et al., 2004). Red amino acids have not been tested with

mutagenesis studies, however it has been postulated that they sequester the ribose and cytidine moieties of CTP (Richard et al., 2004). Purple Amino acids is the glycine rich loop also displaying interactions with cytidine and ribose moieties of CTP. Table 5 summarizes an overview of amino acids and their interactions along with color code on sequence alignments.

Amino Acid	Interaction Location	Amino Acid Color	Moiety
Lys27	Side chain	Blue	α-phosphate
Lys213	Side chain	Blue	α-phosphate/MEP
Arg20	Both	Blue	α-phosphate/MEP
Gly82	Backbone	Red	Cytidine
Asp83	Backbone	Red	Cytidine
Arg85	Both	Red	Cytidine
Ser88	Side chain	Red	Cytidine
Ala107	Backbone	Red	Ribose
Asp106	Side chain	Blue	MEP
Arg109	Side chain	Blue	MEP
Thr140	Both	Blue	MEP
Pro13	Backbone	Purple	Ribose
Ala15	Backbone	Purple	Cytidine
Gly16	Backbone	Purple	Ribose
Thr165	Side chain	Blue	MEP

Table 5. Amino Acids From *E. coli* CMS (PDB: 1INJ) And Their Respective Substrate Interactions. Amino acid color is used for rapid identification in sequence alignments used in this chapter.

EcoliCM	lmo0235	lmo1086	arabdop	lmo3F1C
P13	L8	L7	L10	L9
A15	A10	G9	G12	G11
G16	G11	G10	G13	G12
R20	R15	R14	R17	R16
K27	K22	K22	K24	K24
G82	G81	G80	G81	G83
D83	S82	T81	K82	E84
R85	R84	R83	R84	R86
S88	S87	T86	S87	T89
D106	D105	D110	D105	D103
A107	A106	A111	S106	A104
R109	R108	R113	R108	R115
T140	T138	T143	T138	T145
T165	T161	T167	T162	T169
K213	K209	K216	K210	K218

Table 6. Amino Acid Equivalents From Various CMS Isoforms.

Sequence Analysis of *Listeria monocytogenes* CMS Isoforms

A structure of a *Listeria monocytogenes* CMS isoform was deposited into the Protein Data Bank (PDB 3F1C) (Patskovsky et al., 2008). In this thesis the *Listeria* CMS from the PDB will be referred to as CMS 3F1C. A multiple sequence alignment was constructed including the two isoforms of CMS (0235 and 1086) that were characterized in this project, and CMS 3F1C. The sequence alignment reveals a high percentage (78%) of conserved amino acids between CMS 1086 and 3F1C. Due to the high amount of conservation between them, it is likely that both of these enzymes are products of the same gene. However, the *Listeria monocytogenes* culture used to clone the gene used in this study was slightly different from the culture used by the other investigators cloning the genes used to produce the CMS 3F1C protein. As was shown in the results section of this thesis, both CMS 0235 and CMS 1086 displayed CMS enzyme activity, therefore it is expected CMS 3F1C would exhibit CMS activity. Since CMS 3F1C

resembles a more conserved sequence with CMS 1086, the overall activity expected would be similar to that of CMS 1086. So the question is, why is the activity of CMS 0235 much higher than the CMS 1086 for conversion of MEP to CDP-ME?

Blue Amino Acids

According to Figure 39 and Table 7, all of the amino acids speculated to exhibit active roles in catalysis (indicated in blue) are conserved between all 3 CMS isoforms. This is evidence that the variance in enzyme activity of CMS 0235 and 1086 is not from the key active site amino acids.

Red Amino Acids

Possible locations of interest are S82 and S87 of CMS 0235. These locations are essential for substrate binding of the cytidine and ribose moieties of CTP (Richard et al., 2004). The equivalent amino acids of S82 and S87 from CMS 0235 are T81 and T86 from CMS 1086 respectively, indication that serine has been substituted for a threonine at both positions (Table X). Although serine and threonine are similar amino acids, both contain a hydroxyl group and are considered polar; they are different due to a methyl group on threonine. An argument could be that a slight variation in size may lead to the dissimilar levels of enzyme activity due to protein folding issues or steric hindrance caused by the threonine substitutions in CMS 1086. Also, G107 of CMS 0235 is equivalent to a A111 in CMS 1086 and A104 in 3F1C. Since this amino acid is involved in a backbone interaction it may be that the larger side chain of alanine is disrupting enzyme activity. Despite the amino acid differences addressed here, most of the amino acids are not conserved in other regions of the protein, and an in-depth structural analysis must be completed. Further inspection of S82, G106 and S87 of CMS 0235 may be important in evaluating enzyme activity.

Purple Amino Acids

Evaluation of the glycine rich loop from *Listeria monocytogenes* putative CMS enzymes reveals some subtle differences. In Table 7 the sites where amino acids are not conserved between CMS 0235, 1086, and 3F1C have been shaded in purple as they correspond to the purple amino acids in the sequence alignment. The equivalent amino acid of A10 from CMS 0235 is G9 in CMS 1086. Here a substitution of a glycine for an alanine may cause a tremendous change in activity, even though both are considered nonpolar amino acids. Glycine is a special amino acid that contributes to protein structure in ways no other amino acid is capable of doing. Due to its small size and ability to facilitate tight turns it can position the backbone in various orientations. In the literature, this loop is reported to interact with the 2' and 3' hydroxyl groups on the ribose of CTP (Obiol-Pardo, Cordero, Rubio-Martinez, & Imperial, 2010). These interactions are backbone interactions, and this region consists of small amino acids essential for secondary protein structure. In addition, Q12 from CMS 0235, which is the equivalent of K11 in CMS 1086, is not conserved between the isoforms of CMS from *Listeria*. From the literature, this position has not been fully evaluated but seems to be of lesser importance in the glycine rich loop. Further assessment of this region is necessary. Since substrate interactions within the glycine rich loop are facilitated exclusively through backbone interactions, it reduces the probability of this region as the cause of poor activity for CMS 1086.

Summary of Results

A sequence alignment of CMS 0235, 1086, and 3F1C provides some details that may be correlated to enzyme activity (Figure 39). The blue amino acids are conserved in all the potential isoforms, which indicates that the variation in enzymatic activities is not contingent on those locations. The red amino acids were different in CMS 1086 versus CMS 0235 at locations 82,

106 and 87 from CMS 0235, and should be evaluated further. Since 3F1C has not been kinetically characterized, no tangible conclusions can be drawn from this isoform yet. The purple amino acids display variations at 2 locations, where CMS 0235 contains an A10 and a Q12, CMS 1086 contains a G9 and K11. All locations must be further evaluated.

lmo0235	lmo1086	lmo3F1C
L8	L7	L9
A9	A8	A10
A10	G9	G11
G11	G10	G12
Q12	K11	K13
G13	G12	G14
R15	R14	R16
K22	K22	K24
G81	G80	G83
S82	T81	E84
R84	R83	R86
S87	T86	T89
D105	D110	D103
G106	A111	A104
R108	R113	R115
T138	T143	T145
T161	T167	T169
K209	K216	K218

Table 7. Amino Acid Equivalents From CMS Isoforms In *Listeria monocytogenes*. Amino acids that are shaded indicate lack of conservation and match their corresponding colors in the multiple sequence alignment in Figure 39.

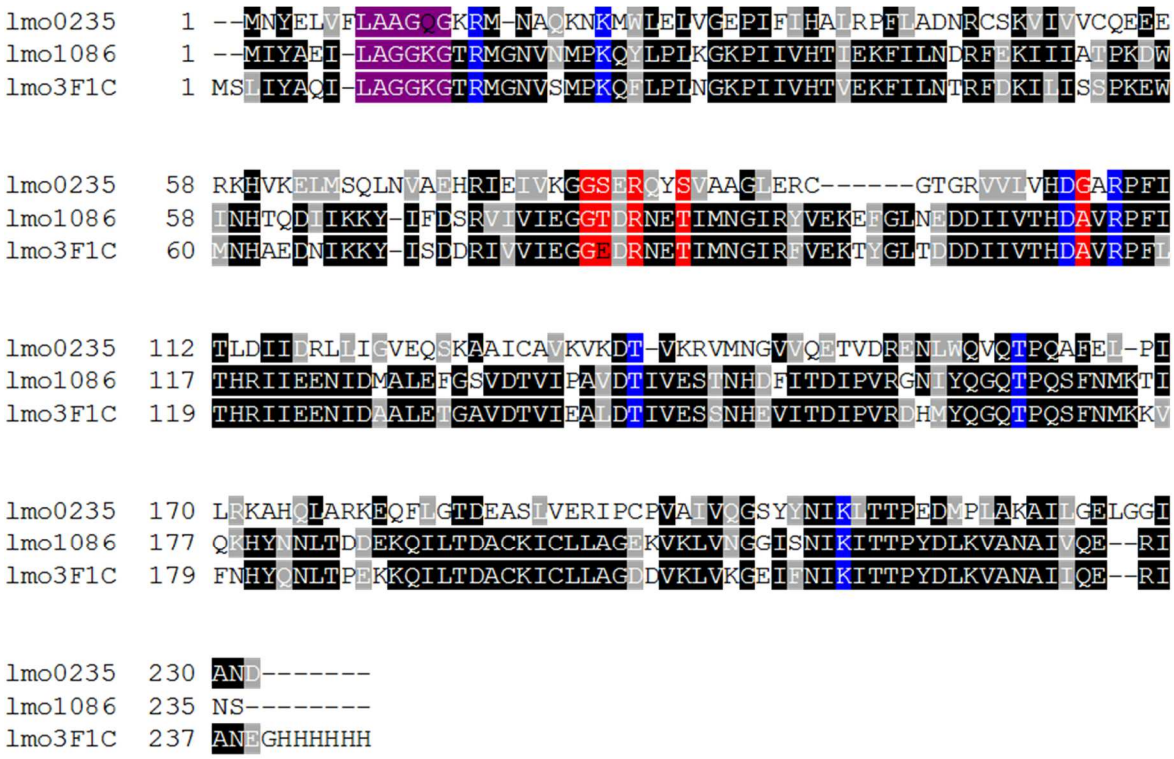

Figure 39. Multiple Sequence Alignment of Putative *Listeria monocytogenes* CMS Isoforms. lmo0235 is CMS 0235 (CAD00762). lmo1086 is CMS 1086 (CAC99164). lmo3F1C is the putative CMS structure (PDB: 3F1C). Blue locations signify amino acids involved in substrate binding (cytidine/ribose) and catalysis. Red locations are amino acids involved in substrate binding only. Purple region is the glycine rich loop. Refer to Table 4 for sequence alignment information. Sequence alignments were constructed using kalign (https://www.ebi.ac.uk/Tools/msa/kalign/) and boxshade (https://embnet.vital-it.ch/software/BOX_form.html).

Sequence Analysis of CMS Isoforms with Known Kinetic Parameters

To further understand how amino acid variations may lead to observed kinetic activities of CMS isoforms, a sequence alignment was constructed using isoforms from various organisms (Figure 40). All CMS isoforms used in the sequence alignment have been characterized, and reported values of K_{cat}, K_m, and assay method are displayed in Table 9. The K_{cat} values of the *E. coli*, *Campylobecter jejuni*, and *Arabidopsis thaliana* enzymes are about 10x higher than CMS 0235 and about 1000x greater than CMS 1086. The differences in these values may be due to key amino acids that are conserved between the highly active enzymes versus less active isoforms. With these data it may be possible to elucidate amino acids that result in reduced activity of CMS 1086.

Blue Amino Acids

The amino acids that are directly involved in catalysis are conserved throughout all the CMS isoforms, suggesting the cause of reduced/activated activity of these enzymes is not due to amino acids involved in catalysis.

Red Amino Acids

Previously, positions S82, S87, and A106 of CMS 0235 were identified for further analysis. The equivalent of amino acid S82 in CMS 0235 is not conserved by CMS isoforms making this position a less likely suspect in reduced enzyme activity for CMS 1086. Variation in the type of amino acid at S82 of CMS 0235 includes polar, positively charged, and negatively charged thus suggests that S82 of CMS 0235 may be less significant. In addition, S82 of CMS 0235 interacts through backbone interactions which makes the identity of the side chain irrelevant. An all-encompassing characteristic is the equivalent of CMS 0235 S82 in all other isoforms being a hydrophilic amino acid. This is an indication that the side chain of the amino

acid is solvent exposed on the exterior of the protein, which upon inspection of *E. coli* CMS (PDB: 1INJ) looks to be true. CMS 0235 S87 is not completely conserved between all the isoforms as well. The S87 of CMS 0235 is equivalent to T86 from CMS 1086 and the *Mycobacterium tuberculosis* CMS isoform (Myctu). Since Myctu displayed better binding affinity ($^{CTP}K_m$= 48 μM) for CTP and similar enzyme activity to CMS 0235 ($^{CTP}K_m$= 329 μM), it can be assured that T86 of CMS 1086 is not an explanation for low enzyme activity. The S87 of CMS 0235 is expected to interact with the cytidine moiety of CTP through its free hydroxyl group on the side chain. Since threonine possess a free hydroxyl group as well, it may be able to facilitate the same interactions as serine. The A106 of CMS 0235 is conserved by all isoforms of CMS except the *Campylobecter jejuni* (S181 equivalent of A107 of CMS 0235) and *Arabidopsis thaliana* (V98 equivalent of A107 of CMS 0235) isoforms of CMS. Isoforms of CMS from *Campylobecter jejuni* and *Arabidopsis thaliana* displayed good enzymatic activity which shows these substitutions are viable. The alanine is only involved in backbone interactions making the exact identity of the amino acid less relevant. CMS 1086 also contained an equivalent alanine at this location rendering this site not the cause of reduced activity. To précis, by applying known kinetic parameters of various CMS isoforms and using multiple sequence alignments we were able to deduce that the S82, S87, and A106 of CMS 0235 equivalents in CMS 1086 were not the primary explanation for poor activity. Nevertheless, further experimental analysis is expected at these locations to verify these results.

Purple Amino Acids

The glycine rich loop displays variation in many of the amino acids within this region including L8, A10, and Q12 from CMS 0235. The first position (L8 in CMS 0235) along the glycine rich loop is not conserved with three isoforms (Ecol, Myctu, and Fratul) containing an

equivalent proline. Using the kinetic data, no trends can be made between the three isoforms that contain a proline and those that do not, implying that this position is likely not responsible for reduced activity of CMS 1086. The next position in the glycine rich loop that is not fully conserved is the A10 of CMS 0235. CMS 1086 and the *Arabidopsis thaliana* (Athal) isoforms contain an equivalent glycine at this location. Since the Athal isoform displays a 1000-fold increase in $^{MEP}K_{cat}$ over CMS 1086 but both contain glycine at this position, this is an indication that this location is not the primary cause of compromised activity in CMS 1086. The final position along the glycine rich loop which is Q12 of CMS 0235 is not conserved between any isoforms except Athal. This location is highly variable amongst isoforms as all types of amino acids are located at the Q12 equivalent of CMS 0235. As mentioned earlier, this location has not been well understood and is visible in protein crystal structures. Other locations in the glycine rich loop are well conserved, ruling out this region as a potential indication for varied activity of CMS 1086. Overall, L8, A10, and Q12 from CMS 0235 are not likely to be interfering with CMS 1086 activity.

Green Amino Acids

A protein modeling and computational experiment was employed to rationalize the increased binding affinity the *Mycobacterium tuberculosis* isoform of CMS has for substrates CTP and MEP (Obiol-Pardo et al., 2010). Using "Myctu" in the alignment, the outcomes suggest that amino acids responsible for increased binding affinity are T141 (T141 remained blue in Figure 40), I142, D193, and T84. D193 and T84 were suspected to coordinate Mg^{2+} while T141 and I142 were expected to help stabilize the dimerization of the protein and improve binding affinity (Obiol-Pardo et al., 2010). Since T141 is conserved with all isoforms this site is not of concern for further evaluation. The *Campylobecter jejuni* CMS has the highest binding affinity

($^{MEP,CTP}K_m$= 20, 3 μM) suggesting it has the best binding affinity of all the CMS isoforms and does not contain an equivalent threonine or isoleucine at T84 and I142 from Myctu. This suggests that the mechanisms of Mg^{2+} binding for these isoforms may in fact be different. At position D193 of Myctu, CMS 1086 contains an equivalent A194. This is a possible location of reduced enzyme activity since the other isoforms contain a negatively charged amino acid while CMS 1086 has a nonpolar alanine. According to structural analysis the *E. coli* isoform does not contain protein to Mg^{2+} interactions, so it may be that Myctu D193 is less relevant in other isoforms (Richard et al., 2004). Without protein structures of CMS 0235 and 1086, these sites becomes more difficult to make confident conclusions from. Without validating the proposed importance D193, I142, and T84 of Myctu by site directed mutagenesis it is not possible to make a conclusion on these amino acids.

Summary of Results

In this section, it has been reinforced that blue amino acids are still conserved and are not the suspected cause of reduced enzyme activity for CMS 1086. This sequence alignment shed light on the relative importance of the red amino acids as well. It was concluded that S82, S87, and A106 of CMS 0235 which are equivalent to T81, T86, and A111 in CMS 1086 are likely not causing poor activity in CMS 1086. The purple amino acids also revealed that L8, A10, and Q12 from CMS 0235 and their respective equivalents in CMS 1086 (L7, G9, and K11) are also not expected to reduce enzyme activity in CMS 1086. This introduces evidence that none of the purple amino acids are liable for the reduced activity in CMS 1086. Now that all blue, red and purple amino acid locations have been ruled out as rationale for poor activity of CMS 1086, this leaves the green amino acids as potential problems. As mentioned the green amino acids were

determined to be important computationally but available structures of *E. coli* CMS (PDB: 1INI) do not support the claims observed computationally.

Using a sequence alignment alone may be inadequate to elucidate the cause of variations in activity for CMS isoforms, therefore site directed mutagenesis studies are of high value to confirm some proposed amino acids involved in dimer stability. In addition, kinetic assay methods are not identical; it is difficult to assess the values properly with variability in assay method. Assay methods such as malachite green are not direct detection of product formation, but a coupled enzymatic assay using a second enzyme to achieve detection. Variations in assay procedures include divalent cation, buffer system, and pH, making a direct comparison of kinetic parameters problematic. Refer to Table 9 and 10 for the experimental conditions of various kinetic assays.

Ecol	lmo0235	lmo1086	Athal	Fratul	Myctu	Camjej
P13	L8	L7	L84	P9	P13	L9
A14	A9	A8	A85	A10	A14	A10
A15	A10	G9	G86	A11	A15	A11
G16	G11	G10	G87	G12	G16	G12
F17	Q12	K11	Q88	I13	S17	N13
G18	G13	G12	G89	G14	G18	G14
R20	R15	R14	R91	R16	R20	R16
K27	K22	K22	K98	K23	K27	K23
G82	G81	G80	G154	G79	G80	G73
D83	S82	T81	K155	E80	S81	D74
R85	R84	R83	R157	R82	R83	R76
A86	Q85	N84	Q158	F83	T84	A77
S88	S87	T86	S160	S85	T86	S79
D106	D105	D110	D180	D106	D107	D97
A107	A106	A111	S181	A107	A108	V98
R109	R108	R113	R183	R109	R110	R100
T140	T138	T143	T212	T141	T141	T129
M141	V139	I144	I213	V142	I142	-
T165	T161	T167	T236	T164	T165	T147
E191	E187	A194	D162	E192	D193	D169
K213	K209	K216	K284	K214	K215	K191

Table 8. Amino Acid Equivalents From CMS Isoforms With Known Kinetic Parameters. Amino acids that are shaded indicate lack of conservation and match their corresponding colors in the multiple sequence alignment in Figure 40.

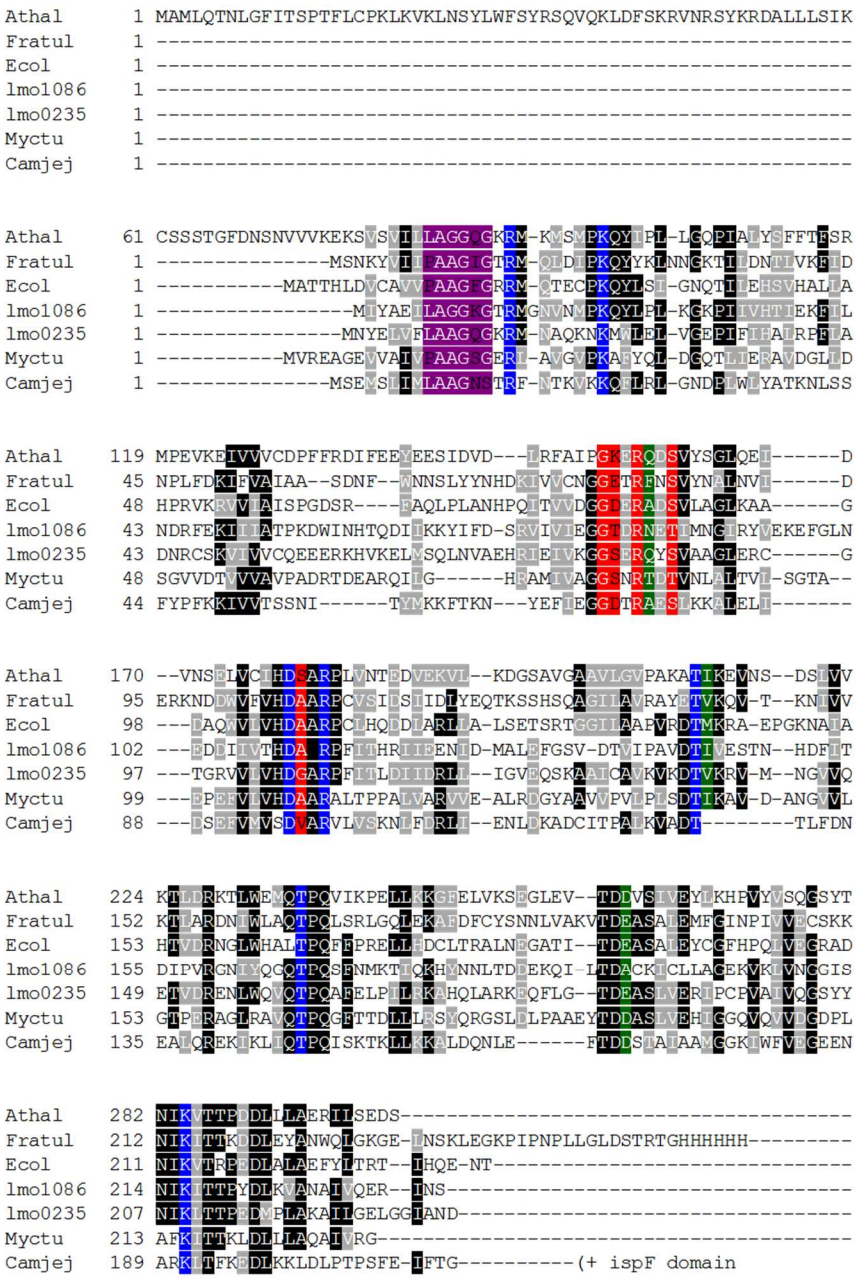

Figure 40. Multiple Sequence Alignment From Kinetically Characterized CMS Enzymes. Blue locations signify amino acids involved in substrate binding (cytidine/ribose) and catalysis. Red locations are amino acids involved in substrate binding only. Purple region is the glycine rich loop. Green location is suspected to coordinate Mg^{2+} and stabilization of interlocking arm in Myctu. Athal is an *Arabidopsis thaliana* CMS. Fratul is the *Francisella tularenisis* CMS isoform. Ecol is the *E. coli* isoform of CMS. lmo 0235 and 1086 are the CMS isoforms studied in this thesis. Myctu is the *Mycobacterium tuberculosis* isoform of CMS. Camjej is a *Camplylobecter jejuni* hypothetical CMS bifunctional enzyme. Refer to Table 4 for sequence alignment information and accession numbers. Sequence alignments were constructed using kalign (https://www.ebi.ac.uk/Tools/msa/kalign/) and boxshade (https://embnet.vital-it.ch/software/BOX_form.html).

Organism	$^{MEP}K_{cat}$	$^{MEP}K_m$ (μM)	$^{CTP}K_{cat}$	$^{CTP}K_m$ (μM)	Assay Method	Reference
E. coli	25.9	760	54.1	370	^{14}C	Richard et al., 2004
Francisella tularenisis	1	73	0.8	178	Malachite Green	Tsang et al., 2011
Mycobacterium tuberculosis	N/A	92	0.8	43	Malachite Green	Shi et al., 2007
Listeria monocytogenes lmo0235	1.13	187	1	329	HPLC	This Study
Listeria monocytogenes lmo1086	0.0143	763	0.0088	950	HPLC	This Study
Campylobecter jejuni	13	20	N/A	3	^{13}C NMR	Gabrielsen et al., 2004
Arabidopsis thaliana	26	500	N/A	114	HPLC	Rohdich et al., 2000

Table 9. K_{cat} and K_m Values of CMS from Multiple Organisms

Organism	pH/Buffer/Cation
E. coli	7.8/Tris-HCl (398 nmol), $MgCl_2$ (40 nmol), NaOH (79 nmol), and inorganic pyrophosphatase 0.7 unit/ Mg^{2+}
Francisella tularenisis	8.0 / 100 mM Tris 1 mM $MgCl_2$ 100 mU/mL inorganic pyrophosphatase / Mg^{2+}
Mycobacterium tuberculosis	8.0 / 50mM Tris, 2 mM $MgCl_2$ / Mg^{2+}
Listeria monocytogenes lmo0235	7.5 / 50 mM imidazole 25 mM $Mg(OAc)_2$ / Mg^{2+}
Listeria monocytogenes lmo1086	7.5 / 50 mM imidazole 25 mM $Mg(OAc)_2$ / Mg^{2+}
Campylobecter jejuni	8.0 / 100 mM Tris 1 mM $ZnCl_2$ / Zn^{2+}
Arabidopsis thaliana	8.0 / 100 mM Tris, 5 mM $MgCl_2$ 0.1 unit inorganic pyrophosphatase / Mg^{2+}

Table 10. Variability In Assay Buffer From Literature Experiments. Buffer system for each enzyme assay used to determine kinetic constants from Table 9.

Structural Homology Modeling of CMS 0235 and 1086

Sequence alignments are useful in the determination of active site amino acids and the effects sequence variability has on kinetic abilities of enzymes. However, structural modeling is another tool that may provide information as to why activities of isoforms are different. As mentioned before, catalytic activity of an enzyme is contingent on protein folding in a proper conformation. Protein modeling was conducted using Biozentrum SWISS-MODEL (Swiss Institute of Bioinformatics). This software takes an inputted amino acid sequence and matches it to amino acid sequences that have known structures deposited into the PDB. The amino acid sequence can then be modeled into 3-D structures using the PDB structures as templates.

Five PDB structures were selected to be used as templates for modeling CMS 0235. Modeled versions of CMS 0235 were analyzed using RasMol (Figure 40). One of the templates used to model the CMS 0235 was a ribitol-5-phosphate cytidylyltransferase (PDB: 4JIS). In comparison, the structures that were modeled all have similar overall secondary and tertiary structure. Table 11 shows the total number of helices, sheets, and turns that were modeled for the CMS 0235 protein. These numbers, although not identical, provide information that some regions of the enzyme that were random coiled in the template structure have defined secondary structure in CMS. These differences are not major, as it is not rare that a short sequence of amino acids is able to form a region of secondary structure depending on favorable amino acid side chain interactions.

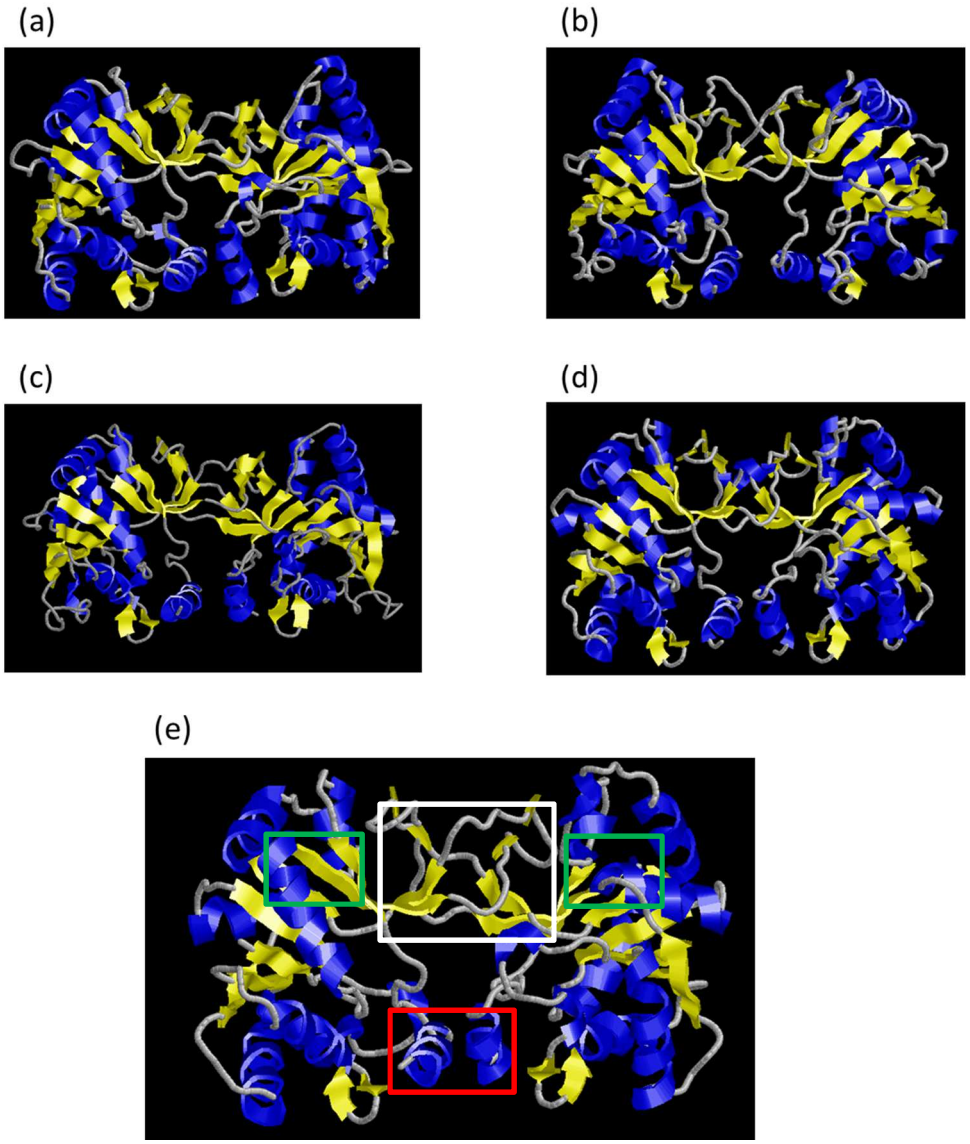

Figure 41. Modeled Protein Folding of CMS 0235 Using SWISS-MODEL. Images created using RasMol. Blue regions indicate α-helices, yellow regions are β-sheets, and grey strings are random coil. (a) CMS 0235 modeled using PDB 5HS2. (b) CMS 0235 modeled using PDB 2YCS. (c) CMS 0235 modeled using PDB 1VGT. (d) CMS 0235 modeled using PDB 3F1C. (e) CMS 0235 modeled on PDB 4JIS, a putative ribitol 5-phosphate cytidylyltransferase. The white box is the region of the interlocking arm, red box indicates the P-loop, and green boxes are the approximate location of the active sites.

When modeling CMS 1086 the template protein that displayed the highest sequence homology was a ribitol-5-phosphate cytidylyltransferase (RCT) from *Bacillus subtilis* (PDB: 4JIS). This result opened up possibilities that the CMS 1086 may be a RCT rather than a MEP cytidylyltransferase. Comparing all structures modeled relative to a RCT, it was still difficult to address locations that would differentiate enzymes structurally. Looking at Figure 42, the RCT model looks to have lower protein order in the interlocking arm region. This difference is slightly visible but 5 models is a small sample size and the difference may not be the cause of poor activity of CMS. Table 12, shows the total number of helices, sheets, and turns that were modeled for the CMS 1086 protein.

CMS 0235	Helices	Sheets	Turns
5HS2	21	28	38
2YC3	18	26	41
4JIS	22	24	45
1VGT	19	27	46
3F1C	22	26	44

Table 11. Secondary Structure Elements For CMS 0235. Number of secondary protein structure elements present for CMS 0235 modeled on proteins using SWISS-MODEL.

CMS 1086	Helices	Sheets	Turns
5DDT	21	26	46
2YC3	20	26	51
4JIS	19	24	43
1VGT	23	28	46
3F1C	22	24	43

Table 12. Secondary Structure Elements For CMS 1086. Number of secondary protein structure elements present for CMS 1086 modeled on proteins using SWISS-MODEL.

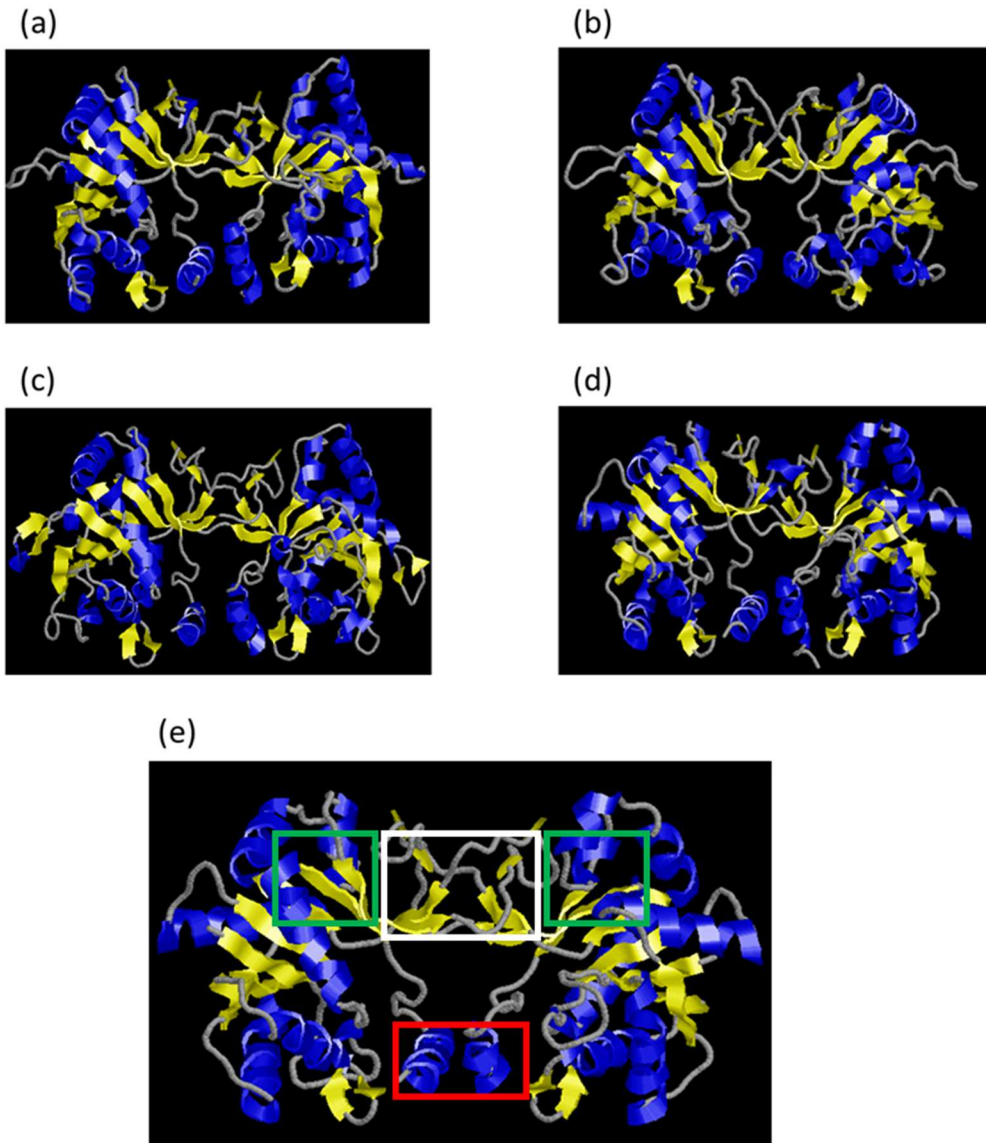

Figure 42. Modeled Protein Folding of CMS 1086 Using SWISS-MODEL. Images created using RasMol. Blue regions indicate α-helices, yellow regions are β-sheets, and grey strings are random coil. (a) CMS 1086 modeled using PDB 5DDT. (b) CMS 1086 modeled using PDB 2YCS. (c) CMS 1086 modeled using PDB 1VGT. (d) CMS 1086 modeled using PDB 3F1C. (e) CMS 1086 modeled on PDB 4JIS, a putative ribitol cytidylyltransferase. The white box is the region of the interlocking arm, red box indicates the P-loop, and green boxes are the approximate location of the active sites

Sequence/Structural Analysis of CMS and RCT Enzymes

A study that specified some key amino acids that are conserved in ribitol cytidylyltransferase enzymes implicated T145, G167, and D137 of Spn2VSH as providing substrate specificity for a ribitol 5-phosphate (Baur et al., 2009). These amino acids are responsible for creating a deeper and more hydrophilic binding pocket (Baur et al., 2009). The reaction catalyzed by RCT alongside that catalyzed by CMS is shown in Figure 43. The substrates ribitol 5-phosphate (RboP) and MEP are similar but the RboP is more polar and longer than MEP, suggesting the need for a deeper and more polar binding pocket. The MEP cytidylyltransferase enzymes do not contain an equivalent of G167 and D137 from Spn2VSH but rather nonpolar amino acids such as Ile, Ala, and Val depending on the isoform in question. Figure 46 shows a multiple sequence alignment with 3 ribitol cytidylyltransferases, and 3 MEP cytidylyltransferases, CMS 0235, 1086, and 3F1C. The T145 of Spn2VSH is conserved by all the enzymes, however, the G167 and D137 of Spn2VSH are conserved between the 3 RboP cytidylyltransferase enzymes, CMS 1086, and 3F1C. Figure 44 displays the potential changes of the active site pocket when glycine and aspartate are substituted for their equivalents, A163 and I132 of *E. coli* CMS (PDB: 1INI), respectively. Figure 45 also illustrates the binding pocket of CMS using a surface map to better visualize the effects of alanine in the binding pocket. A portion of the structure/sequence is missing in all the CMS enzymes in this alignment outlined by a green box (Figure 46). For RCTs this region contains a slightly larger α-helix that the CMS enzymes do not have but the region is in a remote location. This region currently is not well studied although this location is a fingerprint region that may differentiate RCTs and MEP cytidylyltransferases. The study also revealed that the RCT from *S. pneumonia* is active against MEP, although physiologically irrelevant (Baur et al., 2009). *S. pneumonia* does not contain any

enzyme homologs from the MEP pathway but rather contains the genes from the mevalonate dependent pathway. In this section, evidence is provided that CMS 1086 represents better sequence homology with RCTs than MEP cytidylyltransferases.

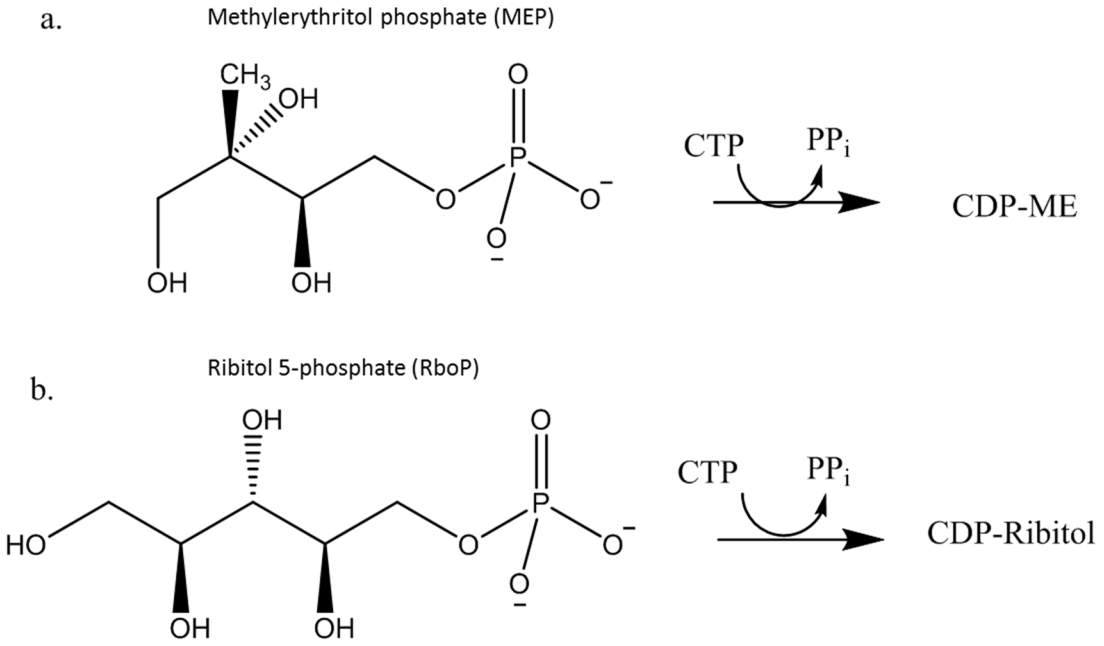

Figure 43. Reaction Catalyzed by Ribitol 5-phosphate Cytidylyltransferase and CMS. (a) Reaction catalyzed by a CMS enzyme. (b) Reaction catalyzed by a RCT. Images were constructed using ChemDraw.

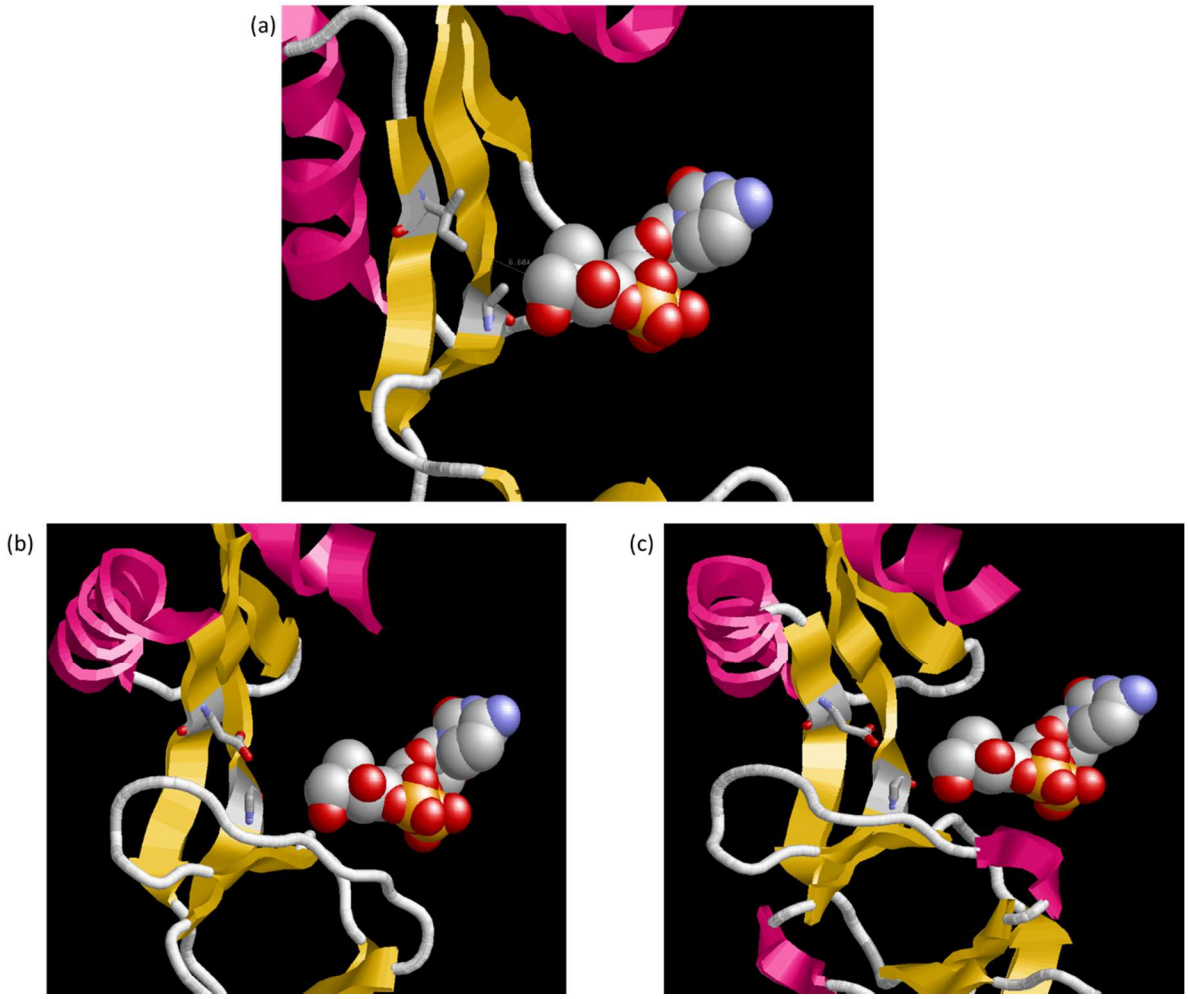

Figure 44. RasMol Models of Partial Active Site Structures From Deposited PDB Files. (a) Structure of *E. coli* CMS complexed with CDP-ME (spacefill) and distance of 6.60 A (PDB 1INI). Alanine and isoleucine are both hydrophobic amino acids that are located within the binding pocket (b) Structure of a known ribitol 5-phosphate cytidylyltransferase from *Streptococcus pneumonia* (PDB 4VSH), CDP-ME artificially for comparison purposes. Binding pocket contains aspartate and glycine allowing selectivity for larger hydrophilic substrates (c) Structure of putative CMS from *Listeria monocytogenes* (PDB 3F1C), product CDP-ME also artificially placed. This structure also contains aspartate and glycine. Comparing the binding region, it is clear that substitution of glycine for alanine creates a larger binding pocket an indication of substrate specificity for RboP. A substitution of aspartate for isoleucine also creates hydrophilic character, also an indication for RboP specificity.

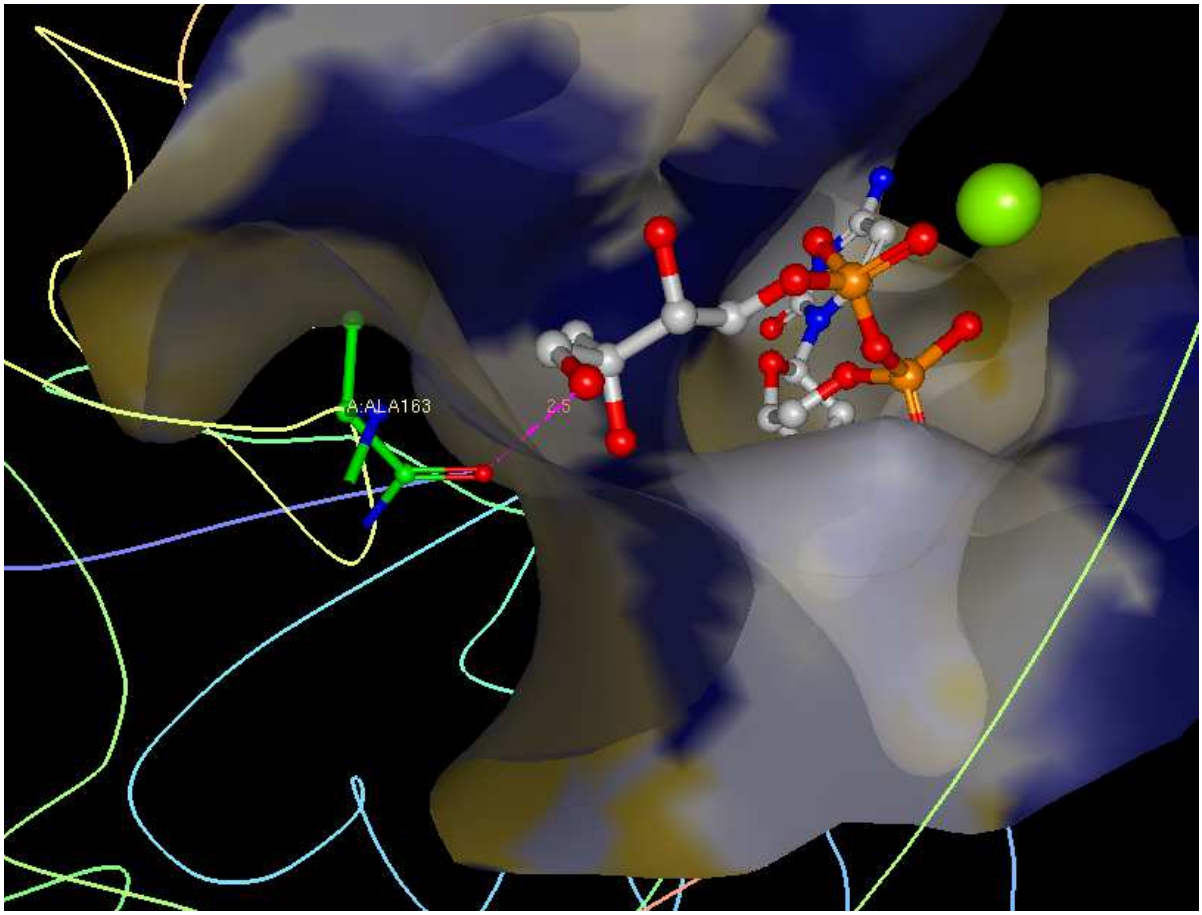

Figure 45. Partial Active Site Surface Map of CMS from *E. coli* (PDB: 1INI). PDB ligand explorer generated picture displaying the short H-bond interaction from the carboxy backbone to the terminal hydroxide of MEP. A surface map is also shown to identify the steric hindrance the methyl group on A163 causes. If a longer more polar RboP was introduced into the binding pocket the methyl group creates a smaller binding pocket. A glycine is conserved at this location for RCTs which would remove the methyl group and open up the binding pocket creating more room to accommodate a larger substrate.

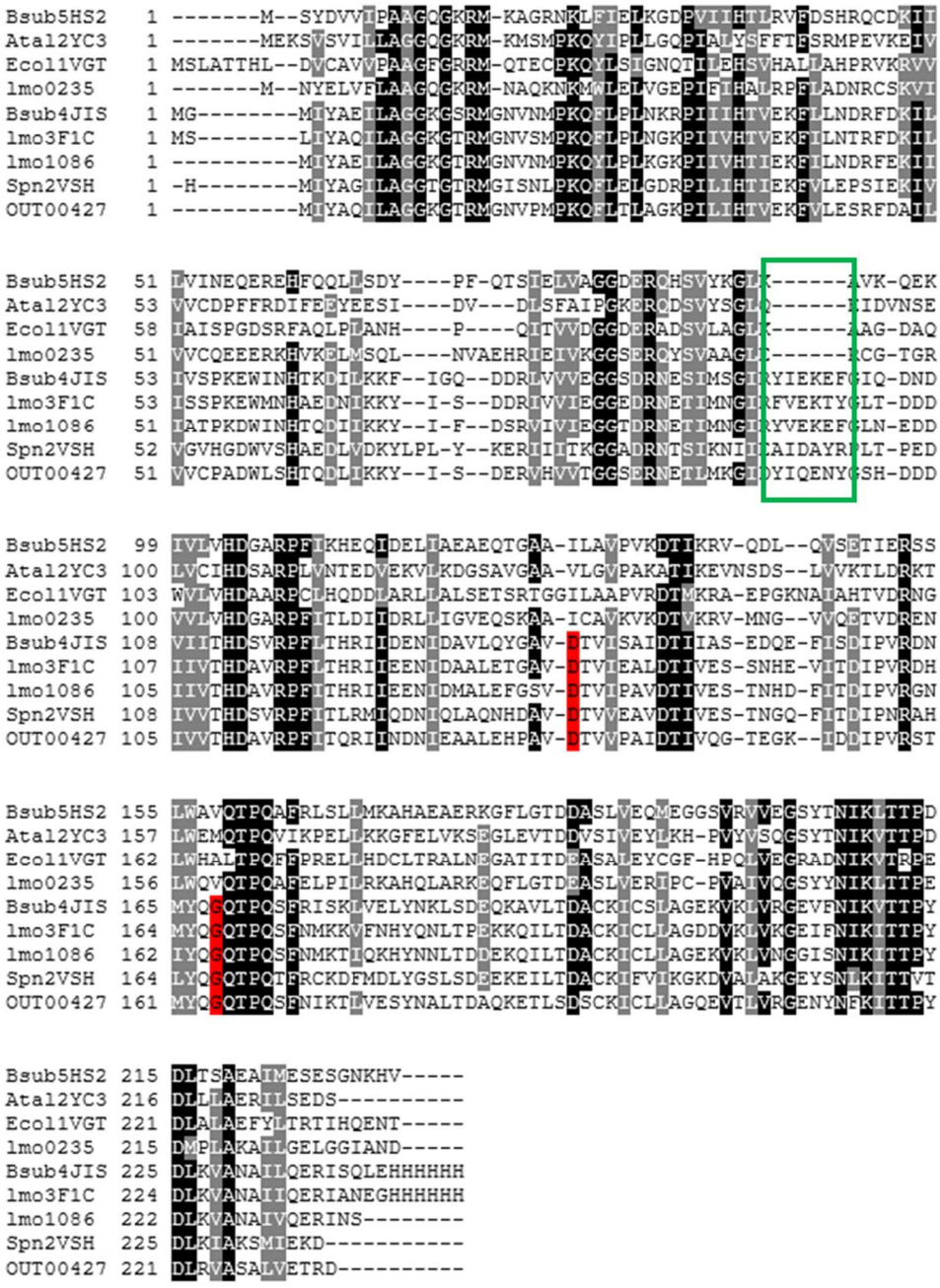

Figure 46. Multiple Sequence Alignment of Ribitol Cytidylyltransferases and MEP Cytidylyltransferases. Multiple sequence alignment including three RboP cytidylyltransferases (Spn2VSH, OUT00427, and Bsub4JIS), three methylerythritol cytidylyltransferases (Bsub5HS2, Atal2YC3, and Ecol1VGT), and the experimental CMS isoforms (lmo0235 and lmo1086). Red locations indicate regions of highly conserved locations for RboP cytidylyltransferases. Sequence alignment indicates CMS 1086 and PDB 3F1C, a putative methylerythritol cytidylyltransferase, contain conserved locations with RboP cytidylyltransferases. Refer to Table 4 for sequence alignment information. Sequence alignments were constructed using kalign (https://www.ebi.ac.uk/Tools/msa/kalign/) and boxshade (https://embnet.vital-it.ch/software/BOX_form.html).

Summary of Results

When using the PDB structures as a template, some of the amino acids involved in binding substrate and in catalysis were not found in the original structure. Similarly, most of the CMS structures that are deposited into the database are found to be missing the same region of protein, which is usually about 15 amino acids following the glycine rich loop. Since these are X-Ray structures using diffractometric techniques, it may be concluded that this portion of the protein is mobile or disordered in a protein crystal.

The underlying theme to all the structures, for both CMS and RCT, is a homodimeric protein that contains the interlocking arm and a P-loop region. This P-loop region is suspected to be an allosteric regulator acting like a seesaw opening and closing the active site upon substrate binding (Jin et al., 2016). Further work must be completed on the P-loop region as it serves an important role and may be the cause of allosteric cooperative binding observed for CMS 0235. Also, all the structures contain a modified version of the Rossmann motif, which is not surprising since it is commonly found in proteins that bind nucleotides. Overall, a structural sequence homology analysis revealed that both putative CMS enzymes (CMS 0235 and 1086) are very well modeled on various CMS isoforms from multiple organisms. However, structural homology modeling coupled with sequence analysis represents evidence that CMS 1086 is a putative RCT and not a CMS.

Evidence using structural modeling and sequence alignments indicate that CMS 1086 and lmo3F1C contain conserved amino acids with known RCTs. These amino acids were Spn2VSH equivalents of G167 and D137 that created a larger binding pocket favorable for the binding of RboP.

CHAPTER IV: CONCLUSIONS

The MEP pathway remains a great source for novel biochemistry and targets for antibiotic drug development. Developing and improving assay techniques such as HPLC contributes innovative ways that allow for the analysis of enzymes. Both enzymes characterized in this project displayed kinetic activity acting as a CMS by converting substrates, MEP and CTP, to product CDP-ME. The CMS 0235 was concluded to be the much more efficient enzyme in comparison to CMS 1086. The catalytic efficiency for CMS 0235 for substrates CTP and MEP were determined to be 3040 $M^{-1}s^{-1}$ and 8248 $M^{-1}s^{-1}$, respectively. Catalytic efficiencies for CMS 1086 with substrate CTP and MEP were calculated to be 9.26 $M^{-1}s^{-1}$ and 18.7 $M^{-1}s^{-1}$, respectively. This is a measure of how efficient enzymes are at catalysis. Data show that the catalytic efficiency for CMS 0235 is about 300x larger than that of CMS 1086. The sequence analysis of CMS isoforms revealed some information on the trends of catalytic activity and conserved amino acids. The structure homology and sequence analysis established some rationale for the poor activity of CMS 1086 by modeling against a ribitol 5-phosphate cytidylyltransferase (RCT). CMS 1086 was best modeled using a RCT, supplying evidence that CMS 1086 is a putative RCT enzyme. Using the model and providing sequence alignments of CMS isoforms and putative RCTs, CMS 1086 displayed similarities with RCTs that the CMS 0235 did not.

CMS 1086: A Ribitol Cytidylyltransferase?

An overall conclusion from sequence alignment and structural modeling is that CMS 1086 is a putative RCT and not a CMS. The National Center for Biotechnology Information (NCBI) reported CMS 1086 as a CMS gene in *Listeria monocytogenes* (Tettelin et al., 2000). Only recently (September 2017) was it identified to be a putative RCT from *Listeria*

monocytogenes after a shotgun genomic project was completed (Casey et al., 2016). If CMS 1086 is a RCT, why would it catalyze the CMS reaction, albeit poorly? In this case, it is not a surprise it was able to behave like a CMS, using MEP and CTP to form CDP-ME. The substrate RboP and MEP have related features as that the CMS 1086 is able to recognize MEP as a substrate as well, and catalyze the reaction using MEP and CTP. During homology modeling and sequence alignments it was determined that these enzymes are structurally similar and this is a contributing factor for CMS 1086 to display CMS activity. Active sites of RCTs are larger and the binding pocket is more hydrophilic due to the smaller glycine and hydrophilic aspartate specific for RboP (Baur et al., 2009). Amino acids in the active site from both RCT and CMS are conserved, entailing that the mechanistic properties are similar if not identical. Further studies must be done to expose more properties of both enzymes to better differentiate them, as it is clear the scientific community is having difficulty differentiating CMS from RCT. Using automated gene prediction methods and protein homology is not sufficient in estimating gene products. In this thesis we have demonstrated that CMS 1086 is active against MEP but displays some structural features consistent with RCTs.

Future Work

The work following the results of this project are to confirm CMS 1086 is a RCT. It would also be useful to clone the gene encoding the putative CMS (PDB 3F1C) from *Listeria monocytogenes* which, according to sequence analysis, may be a RCT as well. Enzyme assays should be completed with CMS 1086 and the 3F1C to validate their identities as a CMS or RCT.

An additional useful study is to perform site directed mutagenesis experiments on purported amino acids directly associated with enzyme function. Computational models are a starting point for analysis but results are based on theoretical calculations and not direct

observations. Evaluation of the enzymes for dimensions of the active site cavity and postulate why RboP or MEP are the preferred substrates. Some interesting questions leading into the future of this project are, is CMS 1086 a RCT, and if so, which amino acids are involved in the catalysis? Is it possible to mutate the amino acid(s) to open up the binding pocket and switch the enzyme into being either a RCT or a CMS? Also a RCT has not yet been characterized using the HPLC method, so the ability to separate product and substrate is still unknown.

The project is also open to developing potent inhibitors of CMS 0235 for the development of antibacterial therapeutics. Obstacles must be overcome in order to design potent inhibitors of CMS. This enzyme utilizes CTP as a substrate and a sugar derivative, which are abundant in all organisms. The dangers of applying a competitive inhibitor approach are evident, as many enzymes utilize CTP as a substrate. Applications of non-competitive, uncompetitive, and mixed competition inhibitors should be explored. This thesis lays a foundation for the development of next generation therapeutics from CMS in *Listeria monocytogenes* as a potential target.

REFERENCES

Ad Heuston, S., Begley, M. I., Gahan, C. G. M., Hill, C., & Ie, G. (2012). Isoprenoid Biosynthesis In Bacterial Pathogens. *Microbiology*, *158*, 1389–1401. http://doi.org/10.1099/mic.0.051599-0

Agilent Technologies. ArcticExpress (DE3) RIL competent cells user manual, 2017. http://www.agilent.com/cs/library/usermanuals/public/230191.pdf (accessed September 2017).

Agilent Technologies. BL21 (DE3) CodonPlus RIL competent cells user manual, 2017. https://www.agilent.com/cs/library/usermanuals/public/230240.pdf (accessed September 2017).

Armstrong, J., Baddiley, J., Buchanan, J., Carss, B., & Greenberg, G. (1958). Isolation and Structure Of Ribitol Phosphate Derivatives (Teichoic Acids) From Bacterial Cell Walls. *Journal of American Chemical Society*, 4344–4354.

Atilano, M. L., Pereira, P. M., Yates, J., Reed, P., Veiga, H., Pinho, M. G., & Filipe, S. R. (2010). Teichoic Acids Are Temporal and Spatial Regulators of Peptidoglycan Cross-inking In *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(44), 18991–6. http://doi.org/10.1073/pnas.1004304107

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., & Milo, R. (2011). The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry*, *50*, 4402–4410. http://doi.org/10.1021/bi2002289

Baur, S., Marles-Wright, J., Buckenmaier, S., Lewis, R. J., & Vollmer, W. (2009). Synthesis of CDP-activated Ribitol for Teichoic Acid Precursors in *Streptococcus pneumoniae*. *Journal of Bacteriology*, *191*(4), 1200–10. http://doi.org/10.1128/JB.01120-08

Bernal, C., Palacin, C., Boronat, A., & Imperial, S. (2005). A Colorimetric Assay For The Determination Of 4-diphosphocytidyl-2-C-methyl-d-erythritol 4-phosphate Synthase activity. *Analytical Biochemistry*, *337*(1), 55–61. http://doi.org/10.1016/j.ab.2004.10.011

Biswas, R., Martinez, R. E., Göhring, N., Schlag, M., Josten, M., Xia, G., Peschel, A. (2012). Proton-binding Capacity of *Staphylococcus aureus* Wall Teichoic Acid And Its Role In Controlling Autolysin Activity. *PloS One*, *7*(7), e41415. http://doi.org/10.1371/journal.pone.0041415

Brault, J. P. (2016) Characterization Of Cytidylyltransferase Enzyme Activity Through High Performance Liquid Chromatography. Illinois State University.

Brault, J. P., & Friesen, J. A. (2016). Characterization of Cytidylyltransferase Enzyme Activity Through High Performance Liquid Chromatography. *Analytical Biochemistry*, *510*, 26–32. http://doi.org/10.1016/j.ab.2016.07.018

Casey, A., McAuliffe, O., Fox, E. M., Leong, D., Gahan, C. G. M., & Jordan, K. (2016). Draft Genome Sequences of *Listeria monocytogenes* Serotype 4b Strains 944 and 2993 and Serotype 1/2c Strains 198 and 2932. *Genome Announcements*, *4*(3). http://doi.org/10.1128/genomeA.00482-16

Chang, M. C. Y., & Keasling, J. D. (2006). Production Of Isoprenoid Pharmaceuticals By Engineered Microbes. *Nature Chemical Biology*, *2*(12), 674–681. http://doi.org/10.1038/nchembio836

Clontech, Talon® Metal Affinity Resin User Manual. Protocol # PT1320-1 (accessed September 2017)

Drawz, S. M., & Bonomo, R. A. (2010). Three Decades of Lactamase Inhibitors. *Clinical Microbiology Reviews*, *23*(1), 160–201. http://doi.org/10.1128/CMR.00037-09

Eugster, M. R., & Loessner, M. J. (2011). Rapid Analysis of *Listeria monocytogenes* Cell Wall Teichoic Acid Carbohydrates by ESI-MS/MS. *PLoS ONE*, *6*(6), e21500. http://doi.org/10.1371/journal.pone.0021500

European Molecular Biology Laboratory. Multiple Sequence Alignment. https://www.ebi.ac.uk/Tools/msa/kalign/ (accessed September 2017).

Gabrielsen, M., Rohdich, F., Eisenreich, W., Grä Wert, T., Hecht, S., Bacher, A., & Hunter, W. N. (2004). Biosynthesis Of Isoprenoids A Bifunctional IspDF Enzyme from *Campylobacter jejuni*. *European Journal of Biochemistry*, *271*, 3028–3035. http://doi.org/10.1111/j.1432-1033.2004.04234.x

Geddes, A., Klugman, K., & Rolinson, G. (2007). Introduction: Historical Perspective And Development Of Amoxicillin/Clavulanate. *International Journal of Antimicrobial Agents*, *30*, 109–112. http://doi.org/10.1016/J.IJANTIMICAG.2007.07.015

Gelmont, D., Stein, R. A., & Mead, J. F. (1981). Isoprene — The Main Hydrocarbon in Human Breath. *Biochemical and Biophysical Research Communications*, *99*(4), 1456–1460. http://doi.org/10.1016/0006-291X(81)90782-8

Gründling, A., Burrack, L. S., Bouwer, H. G. A., & Higgins, D. E. (2004). *Listeria monocytogenes* Regulates Flagellar Motility Gene Expression Through MogR, A Transcriptional Repressor Required For Virulence. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(33), 12318–23. http://doi.org/10.1073/pnas.0404924101

Hanukoglu, I. (2015). Proteopedia: Rossmann fold: A Beta-alpha-beta Fold At Dinucleotide Binding Sites. *Biochemistry and Molecular Biology Education*, *43*(3), 206–209. http://doi.org/10.1002/bmb.20849

Jin, Y., Liu, Z., Li, Y., Liu, W., Tao, Y., & Wang, G. (2016). A Structural And Functional Study On The 2-C-methyl-d-erythritol-4-phosphate Cytidyltransferase (IspD) from *Bacillus subtilis*. *Scientific Reports*. http://doi.org/10.1038/srep36379

King, J., Koc, H., Unterkofler, K., Mochalski, P., Kupferthaler, A., Teschl, G., Amann, A. (2010). Physiological Modeling Of Isoprene Dynamics In Exhaled Breath. *Journal of Theoretical Biology*, *267*(4), 626–637. http://doi.org/10.1016/j.jtbi.2010.09.028

Laemmli, U. K. (1970). Cleavage of Structural Proteins During The Assembly Of The Head Of Bacteriophage T4. *Nature*, *227*(5259), 680–5.

Maroju, S. (2008). Cloning and Partial Characterization of Two Isozymes of Putative 2CMethyl-D-Erythritol 4-Diphosphocytidylyltransferase From *Listeria monocytogenes* EGD-E. Illinois State University.

Neuhaus, F. C., & Baddiley, J. (2003). A Continuum Of Anionic Charge: Structures And Functions Of D-alanyl-teichoic Acids In Gram-positive Bacteria. *Microbiology and Molecular Biology Reviews : MMBR*, *67*(4), 686–723. http://doi.org/10.1128/MMBR.67.4.686-723.2003

Obiol-Pardo, C., Cordero, A., Rubio-Martinez, J., & Imperial, S. (2010). Homology Modeling Of *Mycobacterium tuberculosis* 2C-methyl-d-erythritol-4-phosphate Cytidylyltransferase, The Third Enzyme In The MEP Pathway For Isoprenoid Biosynthesis. *Journal of Molecular Modeling*, *16*(6), 1061–1073. http://doi.org/10.1007/s00894-009-0615-x

Patskovsky, Y., Ho, J., Toro, R., Gilmore, M., Miller, S., Groshong, C., Sauder, J.M., Burley, S.K. New York SGX Research Center for Structural Genomics. http://www.rcsb.org/pdb/explore.do?structureId=3f1c (accessed September 2017)

Portnoy, D. A., Auerbuch, V., & Glomski, I. J. (2002). The Cell Biology Of *Listeria monocytogenes* Infection: The Intersection Of Bacterial Pathogenesis And Cell-mediated Immunity. *The Journal of Cell Biology*, *158*(3), 409–14. http://doi.org/10.1083/jcb.200205009

Ramaswamy, V., Cresence, V. M., Rejitha, J. S., Lekshmi, M. U., Dharsana, K. S., Prasad, S. P., & Vijila, M. (2007). Listeria — Review Of Epidemiology And Pathogenesis. *J Microbiol Immunol Infect*, *40*, 4–13.

Richard, S. B., Lillo, A. M., Tetzlaff, C. N., Bowman, M. E., Joseph P. Noel, A., & Cane, D. E. (2004). Kinetic Analysis of *Escherichia coli* 2-C-Methyl-d-erythritol-4-phosphate Cytidyltransferase, Wild Type and Mutants, Reveals Roles of Active Site Amino Acids†. *Biochemistry*, *43*, 12189–12197. http://doi.org/10.1021/BI0487241

Robbins, J. R., Barth, A. I., Marquis, H., de Hostos, E. L., Nelson, W. J., & Theriot, J. A. (1999). *Listeria monocytogenes* Exploits Normal Host Cell Processes To Spread From Cell To Cell. *The Journal of Cell Biology*, *146*(6), 1333–50.

Rohdich, F., Wungsintaweekul, J., Eisenreich, W., Richter, G., Schuhr, C. A., Hecht, S., Bacher, A. (2000). Biosynthesis Of Terpenoids: 4-Diphosphocytidyl-2C- methyl-D-erythritol Synthase Of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(12), 6451–6456.

Sharma, S., Thulasingam, S., & Nagarajan, S. (2017). Terpenoids As Anti-colon Cancer Agents – A Comprehensive Review On Its Mechanistic Perspectives. *European Journal of Pharmacology*, *795*, 169–178. http://doi.org/10.1016/J.EJPHAR.2016.12.008

Shi, W., Feng, J., Zhang, M., Lai, X., Xu, S., Zhang, X., & Wang, H. (2007). Biosynthesis of Isoprenoids: Characterization of a Functionally Active Recombinant 2-C-methyl-D-erythritol 4-phosphate Cytidyltransferase (IspD) from *Mycobacterium tuberculosis* H37R. *Biochemistr and Molecular Biology*, *40*(6), 911–920.

Simons, K., & Ehehalt, R. (2002). Cholesterol, Lipid Rafts, and Disease. *The Journal of Clinical Investigation*, *110*(5), 597–603. http://doi.org/10.1172/JCI16390

Swiss Institute of Bioinformatics. ExPASy BoxShade. https://embnet.vital-it.ch/software/BOX_form.html (accessed September 2017)

Swiss Institute of Bioinformatics. ExPASy SWISS-MODEL. https://swissmodel.expasy.org/(accessed September 2017)

Swoboda, J. G., Campbell, J., Meredith, T. C., & Walker, S. (2010). Wall Teichoic Acid Function, Biosynthesis, and Inhibition. *Chembiochem : A European Journal of Chemical Biology*, *11*(1), 35–45. http://doi.org/10.1002/cbic.200900557

Tettelin, H., Saunders, N. J., Heidelberg, J., Jeffries, A. C., Nelson, K. E., Eisen, J. A., Cossart, P. (2000). Complete Genome Sequence of *Neisseria meningitidis* Serogroup B Strain MC58. *Science*, *287*(5459), 1809–1815. http://doi.org/10.1126/science.287.5459.1809

Tsang, A., Seidle, H., Jawaid, S., Zhou, W., Smith, C., & Couch, R. D. (2011). *Francisella tularensis* 2-C-methyl-D-erythritol 4-phosphate Cytidylyltransferase: Kinetic Characterization And Phosphoregulation. *PloS One*, *6*(6), e20884. http://doi.org/10.1371/journal.pone.0020884

Weidenmaier, C., Peschel, A., Xiong, Y., Kristian, S. A., Dietz, K., Yeaman, M. R., & Bayer, A. S. (2005). Lack of Wall Teichoic Acids in *Staphylococcus aureus* Leads to Reduced Interactions with Endothelial Cells and to Attenuated Virulence in a Rabbit Model of Endocarditis. *The Journal of Infectious Diseases*, *191*(10), 1771–1777. http://doi.org/10.1086/429692

World Health Organization. (2017). *Tuberculosis* (Fact sheet No. 104). Retrieved from http://www.who.int/mediacentre/factsheets/fs104/en/ (Accessed August 2017)

World Health Organization. (2017). *Malaria* (Fact sheet No. 94). Retrieved from http://www.who.int/mediacentre/factsheets/fs094/en/ (Accessed August 2017)