ACHARYA, REJWI, M.S. Overexpression, Purification, and Characterization of MmgD from *Bacillus subtilis* Strain 168. (2009)
Directed by Dr. Jason J. Reddick. 62 pp.

Bacillus subtilis, a rod shaped gram-positive bacterium, is highly studied because of its capability of forming endospores when there is an inadequate supply of nutrients. Despite the lack of nutrients in the environment, this microorganism has to be able to generate energy to support sporulation.

One operon expressed during sporulation is the mother cell metabolic (*mmg*) operon. The six open reading frames (ORFs) of the *mmg* operon are *mmgABCDE* and *yqiQ*. The first three genes are homologs of fatty acid metabolism enzymes and the last three are expected to encode part of the methyl citric acid cycle, which is a metabolic pathway that processes propionyl-CoA. In the absence of glucose, the preferred carbon source, this organism most likely makes use of other carbon sources, such as fatty acids and/ or propionates to fuel sporulation.

Based on sequence homology the function of *mmgD* is similar by sequence to citrate synthase III. *MmgD* has been successfully cloned, overexpressed, and the purified MmgD enzyme from *Bacillus subtilis* strain 168 has demonstrated citrate synthase and methyl citrate synthase characteristics. MmgD showed slight preference for propionyl-CoA over acetyl-CoA. Besides activity determination and substrate preference, product formations of this enzyme have also been validated via ^{1}H-proton NMR.

OVEREXPRESSION, PURIFICATION, AND CHARACTERIZATION OF MmgD FROM *BACILLUS SUBTILIS* STRAIN 168

by

Rejwi Acharya

A Thesis Submitted to
The Faculty of the Graduate School at
The University of North Carolina at Greensboro
In Partial Fulfillment
of the Requirements for the Degree of
Master of Science

Greensboro
2009

Approved by

_____
Committee Chair

To mamu and baba

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Committee Chair _____

Committee Members _____

_____

_____

Date of Acceptance by Committee

_____

Date of Final Oral Examination

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Jason Reddick, for his time, patience, and for much appreciated guidance throughout my graduate years. I would also like to thank my committee members, Dr. Norman Chiu and Dr. Alice Haddy, for their assistance. I also thank Dr. Gregory Raner for his guidance and advice.

To my family and friends, I would not be where I am today if it were not for your unlimited support. I would also like to thank all of those who helped and supported me in any respect during my years here.

This work was supported by National Science Foundation Award #0817793.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

I.A Sporulation and the Sigma Factors Involved in *Bacillus subtilis*

Bacillus subtilis, a rod shaped gram-positive bacterium, has attracted great interest because of its ability to form endospores. Beside this characteristic, *B. subtilis* has other advantageous qualities such as: it is easily manageable in experimental conditions, it has well understood genetics, and its cellular organization is fairly simple (Errington, 1993). These obligate aerobic bacteria interrupt their cell cycle and form endospores as a defense mechanism against harsh internal and external conditions. Some conditions that induce sporulation are cell density, DNA damage, and nutrient depletion (Stragier, 1996).

Of all the conditions mentioned above, nutrient depletion due to an inadequate supply of carbon, nitrogen or phosphorus usually induces sporulation (Errington, 1993) by inducing the phosphorylation of Spo0A, a key transcriptional regulatory protein. Initially, phosphate groups are transferred from at least three histidine protein kinases to Spo0F which are then transferred to Spo0B, and lastly Spo0A receives those phosphate groups and becomes activated (Stragier, 1996). There are several phases of Spo0A activation; of these the ‘transition state’ is associated with several events that precede sporulation. Biofilm formation, protease production, competence for transformation,

motility, and even cannibalism are some of those events (Piggot, 2004). In order for *B. subtilis* to form endospores, enough Spo0A has to be accumulated. Since activated Spo0A controls expression rate of *spo0A* and other phosphate genes, Spo0A is phosphorylated until enough Spo0A are collected (Stragier, 1996).

As mentioned above, sporulation is induced when there are no nutrients; but when there are plenty of nutrients available, vegetative *B. subtilis* double in length and divide into two identical daughter cells. Keeping the sporulation cycle in mind, those vegetative cells are said to be in stage 0 (Errington, 1993). In an environment with nutrient depletion, *B. subtilis* does not divide right away. Instead, the chromosomes form axial filaments by aligning along the long axis of the cell, this process describes stage 1. Next, septum formation occurs at "one of the two polar rings of FtsZ," a cell division protein, (Stragier, 1996) and only then are the cells divided (Piggot, 2004). Stage III is called the engulfment stage; the smaller prespore is engulfed by the mother cell. Cortex formation in stage IV followed by coat formation in stage V leads to maturation in stage VI. In the final stage, stage VII, the mature dormant spore is released as the mother cell lyses, which concludes the cycle of sporulation (Errington, 1993). Figure 1 depicts sporulation cycle in *B. subtilis* along with the various sigma factors involved.

Upon division, 30% of the origin-proximal chromosomes remain in the prespore while the remaining chromosome is transported across the septum. DNA translocase SpoIIIE governs the translocation; without it the proteins produced in the mother cell would create holes in the asymmetric septum. Such pores would be big enough for some regulatory proteins to diffuse across and do some damage. The asymmetric division of

the cell is due to the activation of Spo0A and the sporulation sigma factor σ^H in the predivisional cells (Piggot, 2004).

Transcription factor σ^F is activated only in the prespore after the asymmetric division and before the complete chromosome partition. In the presence of the anti- σ factor SpoIIAB, σ^F is inactive in the pre-divisional cell (Piggot, 2004). The anti-anti-σ factor SpoIIAA, active when dephosphorylated by SpoIIE and inactive when phosphorylated by SpoIIAB, reverses the inhibition by binding to the SpoIIAB-σ^F complex and releasing the active σ^F (Piggot, 2004; Stragier, 1996). Linkage between σ^F activation and asymmetric division along with how σ^F is localized only within prespore is still unclear. SpoIIE "has been proposed to 'sense' asymmetric division and activate σ^F in response" (Piggot, 2004).

Activation of the transcription factor σ^E in the mother cell occurs after the activation of σ^F in the prespore. The sigma factor σ^E is synthesized as an inactive membrane-bound precursor, known as pro- σ^E; but when a 27 residues peptide is removed from the N-terminal by the putative protease SpoIIGA, the precursor converts to its active form (Stragier, 1996; Piggot, 2004). The promoter-distal member, SpoIIGB, of the SpoIIG operon is responsible for producing the pro- σ^E. The regulator Spo0A, whose activity is elevated in the mother cell after the asymmetric division, is known to enhance the expression of SpoIIGB. As a result, the level of pro- σ^E increases in the mother cell (Piggot, 2004). Since σ^E activity is only seen in the mother cell even though both the pro-σ^E and SpoIIGA are present before septation, it implies that the protease SpoIIGA

removes the pro-amino acid sequence of pro- σ^E only after the septation once it receives a signal from the prespore. That signal is spoIIR, a gene controlled by σ^F (Stragier, 1996).

The next step in the sporulation process is engulfment. Various proteins produced by the mother cell slowly degrade the peptidoglycan septum. Once the septum is eliminated, the prespore is engulfed within the cytoplasm of the mother cell (Errington, 1993; Stragier, 1996). A sigma factor σ^G is responsible for transcription of many genes in the engulfed prespore (Stragier, 1996). Prior to engulfment, this factor is synthesized and held inactive (Piggot, 2004).

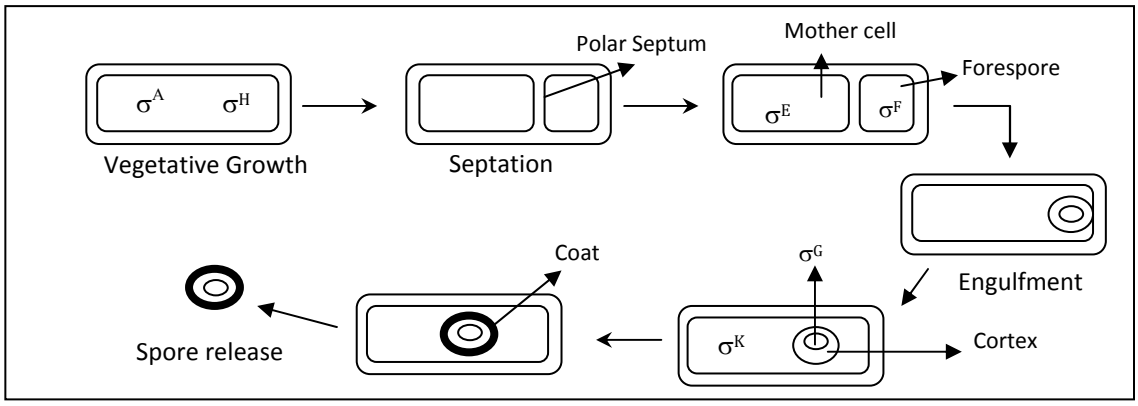

Figure 1: **Sporulation in *Bacillus subtilis*** (Kroos, 2003)

Although the regulation of sporulation is well studied, its biochemistry and metabolism are not well understood. For that reason various transcription factor dependent promoters were screened in order to identify genes that are used at various stages of sporulation.

I.B The Mother Cell Metabolic Genes *(mmg)* Operon

Bryan and coworkers screened for σ^E dependent promoters in order to identify genes expressed at intermediate stages of sporulation. One promoter was shown to drive expression of an operon of with least five open reading frames (ORFs). The first five ORFs of this operon are *mmgA* (393a.a), *mmgB* (287a.a), *mmgC* (379a.a), *mmgD* (371a.a), and *mmgE* (Bryan, 1996); a sixth ORF revealed later by the genome sequence is called *yqiQ* (Kunst, 1997). Since the transcription of the genes was found to be dependent on the mother-cell-specific sigma factor, σ^E, the operon was named *mmg*, which stands for mother cell metabolic genes. An overview of the operon is shown below (figure 2).

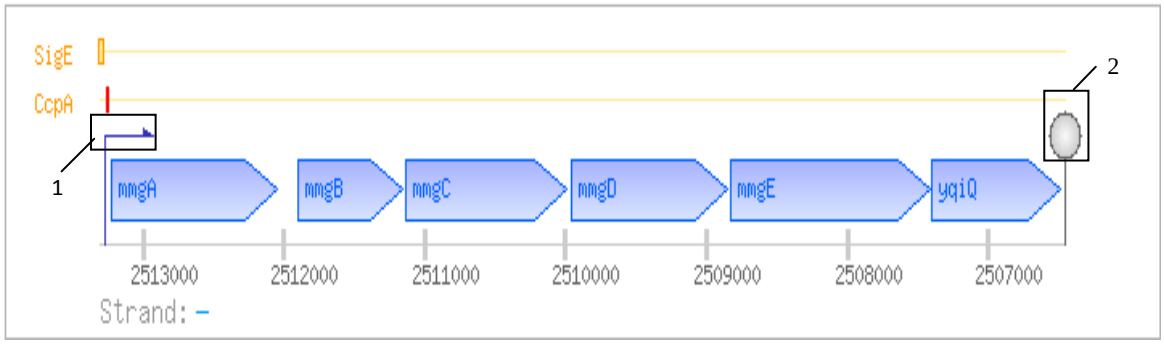

Figure 2: **Overview of the *mmg* Operon. SigE: RNA polymerase; CcpA: regulator; : Genes with their genome positions; 1: promoter; 2: transcription terminator sequence (htt://dbtbs.hgc).**

I.B1 Regulations of the *mmg* Operon

The *mmg* promoter region includes a sequence that resembles the catabolite-responsive element (CRE) sequence (Bryan, 1996). This 14 bp sequence, called *mmgO*, is

22 bp downstream of the transcriptional start site of the *mmg* operon. Bryan et al. successfully demonstrated the role of *mmgO* as a glucose-dependent repressor of *mmg* promoter activity. *CcpA* is a protein that helps to decrease transcription from specific promoters by binding to the CRE. Glucose repression of the *mmg* promoter requires the CRE site throughout sporulation. *CcpA* and *mmgO,* together with σ^E mediate the transcription of *mmg* operon (Bryan, 1996).

I.B2 Functions Based on Sequences of *mmg*

Based on sequence homology, *mmgA, mmgB, and mmgC* are similar to the enzymes involved in fatty acid metabolism. It is known that *mmgA* encodes an acetyl-coenzyme A (CoA) thiolase (Reddick, 2008). The sequences of *mmgB* and *mmgC* are similar to that of 3-Hydroxybutyryl-CoA dehydrogenase and acyl-CoA dehydrogenase, respectively (Eichenberger, 2003). The sequence of *mmgD* is similar to citrate synthase III. The sequence of *mmgE* is similar to 2-methylcitrate dehydratase. Lastly, *yqiQ* is similar to 2-methylisocitrate lyase (Sonenshein, 2002). Based on sequence homology, *mmgD, mmgE* and *yqiQ* are similar to the enzymes involved in propionate metabolism.

I.C Fatty Acids and the β-Oxidation Cycle

Important sources of carbon and energy for many microorganisms are fatty acids, which can be short ($<C_6$), medium (C_6 to C_{12}), and long-chain ($>C_{12}$). While the latter two are breakdown products of cell membrane components, the former is a product of fermentation (Horswill, 1999). Acetate, propionate, and butyrate are some of the products of fermentation. These short chain fatty acids are oxidized to yield CO_2 either aerobically

or anaerobically. Most fatty acids are metabolized via β-oxidative pathway; the exception is propionate, which can be metabolized by several known pathways (Textor, 1997).

Odd chain and even chain fatty acids are degraded by β-oxidation cycle, in which two carbon units of fatty acids are removed after a series of enzyme-catalyzed reactions. Since the β carbon of fatty acids is oxidized when they are broken down, this cycle is named the β-oxidation cycle. An overview of this cycle is shown in figure 3.

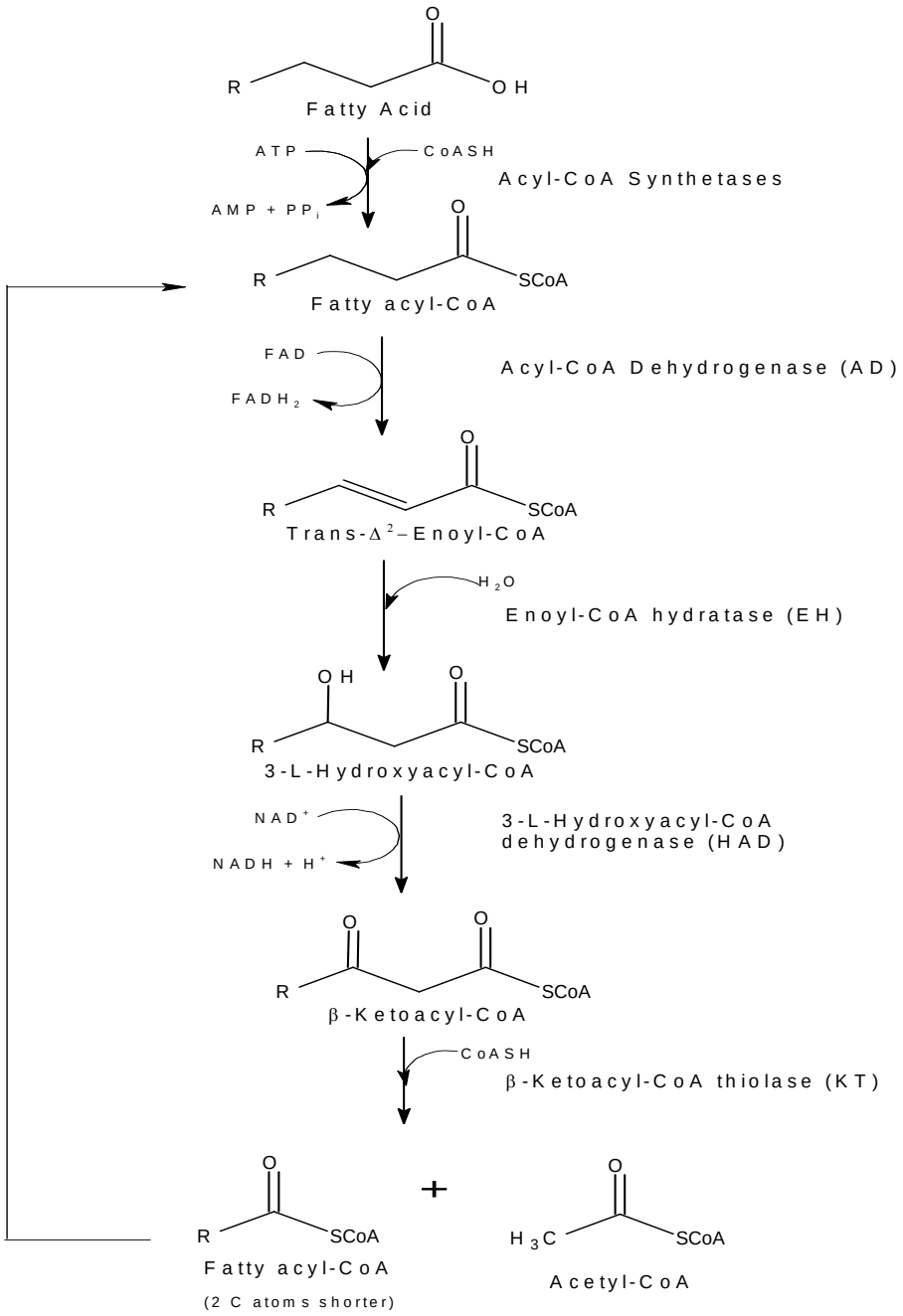

Figure 3: β-Oxidation of Fatty Acids.

Fatty acids can be degraded to generate energy or they can function as precursors for membrane phospholipids. In *Escherichia coli*, the genes of *fad* (fatty acid

degradation) regulon are responsible for β-oxidation of fatty acids. In *E. coli* there is only one acyl-CoA synthetase, *fadD,* which is an inner membrane associated enzyme. Along with *fadL,* an outer membrane protein, *fadD* transports long chain fatty acids across the cell membrane in order to metabolism them (Fujita, 2007).

Prior to the β-oxidation of fatty acids, they are thioesterified with coenzyme A via an ATP-dependent acylation reaction. This step is crucial since fatty acids are activated for β-oxidation in this step. In *E. coli* the first step of β-oxidation is catalyzed by *fadE,* the only acyl-CoA dehydrogenase in this organism. The flavoenzyme acyl-CoA dehydrogenase catalyzes the conversion of fatty acyl-CoA to trans-Δ^2-Enoyl-CoA. A tetrameric complex that has two copies of *fadB* and *fadA* is responsible for catalyzing rest of the steps of β-oxidation. *fadB* functions as enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase while *fadA* functions as 3-ketoacyl-CoA thiolase (Fujita, 2007). Enoyl-CoA hydratase is used to hydrolyze trans-Δ^2-Enoyl-CoA to yield 3-L-Hydroxyacyl-CoA. The NAD^+ dependent enzyme 3-L-Hydroxyacyl-CoA dehydrogenase is responsible for dehydrogenating 3-L-Hydroxyacyl-CoA to produce β -ketoacyl-CoA. In the last step of this cycle, β-ketoacyl-CoA thiolase with CoASH cleaves the C_α–C_β of β -ketoacyl-CoA to produce acetyl-CoA and a fatty acyl-CoA, which after a cycle of the β-oxidation has 2 less carbon atoms.

FAD dependent dehydrogenation, hydration, NAD^+ dependent dehydrogenation, and C_α–C_β cleavage in thiolysis reaction make up the β-oxidation cycle. This cycle continues to break down even chain fatty-acyl-CoA until the final two acetyl-CoA are

generated. When odd chain or branched fatty acids are digested, the final products are propionyl-CoA and acetyl-CoA (Voet, 2006). For additional energy generation, acetyl-CoA enters the citric acid cycle and condenses with oxaloacetate to produce citrate. Propionyl-CoA condenses with oxaloacetate to produce methylcitrate in the methylcitrate cycle (Figure 4).

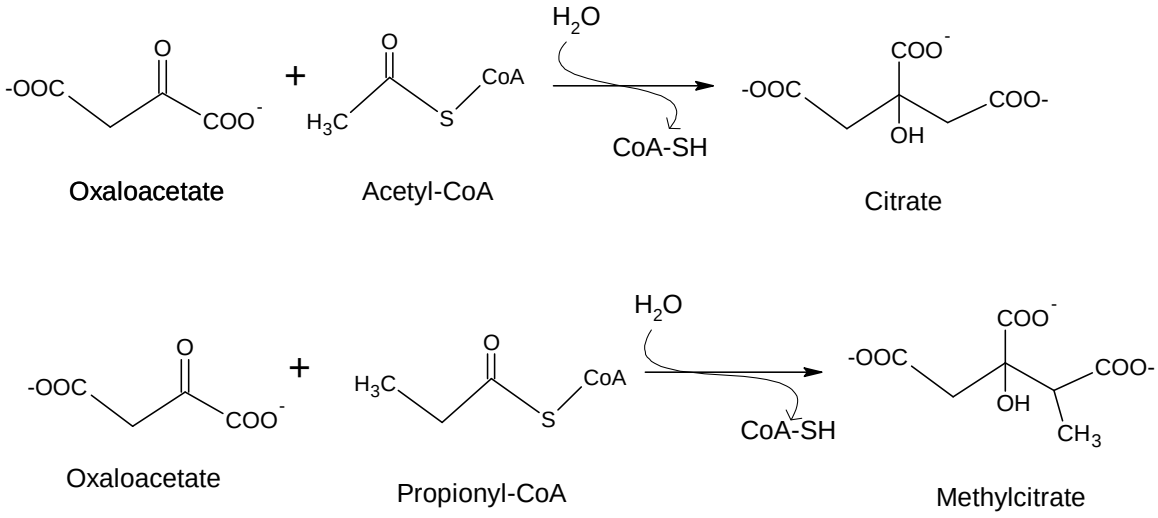

Figure 4: (top) Condensation of Oxaloacetate and Acetyl-CoA to Yield Citrate; (bottom) Condensation of Oxaloacetate and Propionyl-CoA to Yield Methylcitrate.

I.D Propionates and the Methylcitric Acid Cycle

Odd chain length fatty acids are a rich source of energy for many microbes that survive in soil (Upton, 2007). An example of such microbes is *B. subtilis*. Cell membranes of this organism are composed of fatty acids, 80% of which are branched (Klein, 1999). The metabolism of branched-chain fatty acids, branched-chain amino acids, odd-chain length fatty acids, and propionates yield propionyl-CoA (Upton, 2007).

There are several possible propionate metabolism pathways; some of which are: α-oxidation, β-oxidation, α-carboxylation, reductive carboxylation, and Claisen condensation (Textor, 1997).

Textor et al. showed that propionate is oxidized via the methylcitrate cycle in *Escherichia coli*. They have shown that aldol additions of propionyl-CoA with either oxaloacetate or glycoxylate is evident with "the isotope patterns of alanine derived from (2-^{13}C), (3-^{13}C), and (3-$^{2}H_3$) propionate only." Acetate and pyruvate are the products of condensation between glyoxylate and propionyl-CoA. Pyruvate and succinate are the products of condensation between oxaloacetate and propionyl-CoA (Textor, 1997).

Many microbiologists have accepted that *E. coli* is able to generate pyruvate from propionate via acryloyl-CoA. First the propionyl-CoA is dehydrogenated to acryloyl-CoA, which is subject to α-hydration yielding lactoyl-CoA, which eventually yields pyruvate. In *Acidaminococcus fermentans* there is a gene that codes the activator of 2-hydroxyglutaryl-CoA; that gene is similar in sequence to an ORF from *E. coli* called "yijL". Because of the similarity between lactoyl-CoA dehydratase and 2-hydroxyglutaryl-CoA dehydratase, Textor et al. thought that propionate is broken down via generation of acryloyl-CoA, which is hydrated to produce lactoyl-CoA. But when they saw evidence of CoA-SH being released from propionyl-CoA, they suspected methylcitrate synthase (MCS) was present. They partially sequenced the purified enzyme and found a putative operon, which is similar to the *prp* operon from *Salmonella typhimurium* (Textor, 1997). The *Prp* operon is known to catabolize propionate via the methylcitrate cycle (Horswill, 1999).

Existence of the methylcitrate cycle is evident in *Salmonella typhimurium* and *Escherichia coli* (Horswill, 1999; Textor, 1997). In the last few decades, many prokaryotic genomes have been sequenced and as a result an operon from a class of bacterium can be compared with an operon of another class of bacterium. For example, the *E. coli* genome sequencing project helped identify the similarities between a chromosomal region of *E. coli* and the *prpRBCDE* locus of *S. typhimurium*. The five genes of this locus are involved in propionate oxidation. Sequence comparisons of the *Prp* enzymes showed that they aren't similar to the enzymes involved in acryloyl-CoA pathway; but they are similar to the enzymes involved in the 2-methylcitrate cycle of fungi (Horswill, 1999).

To establish the pathway of propionate oxidation in *S. typhimurium* with definitiveness, Horswill et al. focused on identifying the intermediates of the pathway. Changes in the oxidation state of the methylene carbon of propionate were monitored by ^{13}C-NMR since it changes dramatically throughout the pathway. Their data eliminated citramalate and acryloyl-CoA pathways and supported the 2-methylcitrate cycle as a way to catabolize propionate. Some enzymes specific to the 2-methylcitrate cycle (figure 4) are propionyl-CoA synthetase, 2-methylcitrate synthase, and 2-methylisocitrate lyase. The latter two are encoded by *prpC* and *prpB* genes, respectively. Propionyl-CoA synthetase (not shown in the figure) catalyzes the conversion of propionate to propionyl-CoA; this enzyme is the PrpE enzyme in *S. typhimurium* LT2. The exact role of *prpD* is unclear; it could convert 2-methyl-*cis*-aconitate to 2-methylisocitrate and/or it could just

dehydrate 2-methylcitrate to 2-methyl-*cis*-aconitate. Either way, ^{13}C-NMR data have shown that *PrpD* is needed for the conversion of 2-methylcitrate to 2-methylisocitrate.

There are genes in *Bacillus subtilis* that are homologs of *prpC, prpD, and prpB,* (Horswill, 1999). These genes are *mmgD*, *mmgE*, and *yqiQ*, respectively. In this microorganism methylcitrate synthase (MCS), methylcitrate dehydratase (MCD), and methylisocitrate lyase (MCL) are likely encoded by *mmgD*, *mmgE*, and *yqiQ* genes, respectively.

Condensation of propionyl-CoA and oxaloacetate to produce 2-methylcitrate is catalyzed by methylcitrate synthase. 2-methylcitrate is converted to 2-methyl-cis-aconitate, which produces 2-methylisocitrate when catalyzed by aconitase. Upon initiation of aldol cleavage by methylisocitrate lyase, 2-methylisocitrate generates pyruvate and succinate. Pyruvate either enters the citric acid cycle or goes to produce glucose via gluconeogenesis. Succinate in subsequent steps regenerates oxaloacetate. The main function of the methylcitrate cycle is to convert propionyl-CoA to pyruvate on an equimolar basis (Upton, 2007).

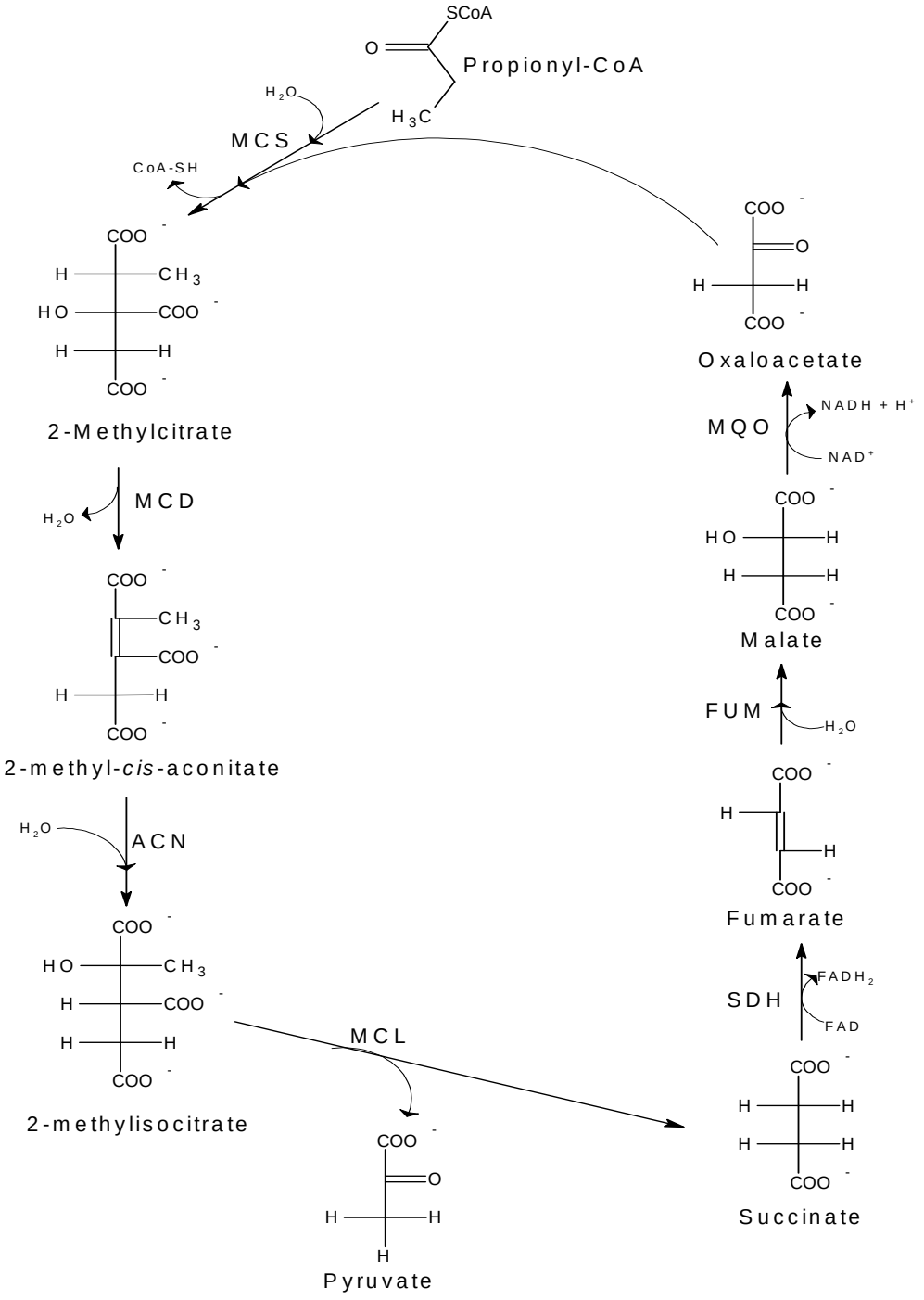

Figure 5: The Methylcitrate Cycle. Enzymes involves are: MCS, methylcitrate synthase. MCD, methylcitrate dehydratase. ACN, Aconitase. MCL, methylisocitrate lyase. SDH, succinate dehydrogenase. FUM, Fumarase. MQO, malate-quinone oxidoreductase.

I.E Introduction to *mmgD*

Based on sequence homology mother cell metabolic gene (mmg) D is similar to citrate synthase of other organisms, such as *S. typhimurium* (Horswill, 1999). The citrate synthase *mmgD* might function with *mmgABC* since they are present near one another; and this possibility of interaction suggests that fatty acids are broken down via the β-oxidation pathway (Bryan, 1996). Acetyl CoA and oxaloacetate in the presence of *mmgD* and water is expected to produce citrate and CoA-SH.

MmgD-like enzymes in *Pseudomonas aeruginosa* and *Salmonella typhimurium* are known to be bifunctional (Watson, 1983; Horswill, 1999). When the bacterium is grown on propionate medium, the enzyme that is highly active is methylcitrate synthase (Watson, 1983). While citrate synthase is involved in the citric acid cycle with acetyl-CoA being the cycle's precursor; methylcitrate synthase is involved in the methylcitric acid cycle with propionyl-CoA being the cycle's precursor. If *mmgD* were a bifunctional enzyme, the expected product of propionyl-CoA and oxaloacetate would be methylcitrate and CoA-SH while citrate and CoA-SH would be produced when acetyl-CoA and oxaloacetate condense (figure 6).

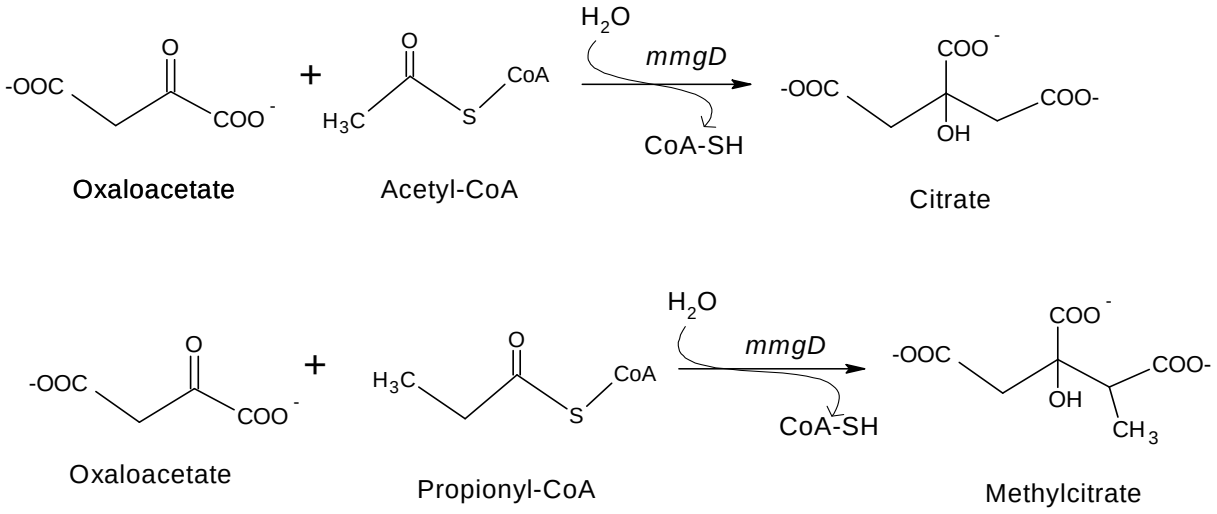

Figure 6: (top) Condensation of Oxaloacetate and Acetyl-CoA to Yield Citrate; (Bottom) Condensation of Oxaloacetate and Propionyl-CoA to yield Methylcitrate.

The presence of methylcitrate synthase is highly evident when bacteria, such as *P. aeruginosa*, are grown on propionate medium. Watson et al. have reported low but detectable activities of MCS when this bacterium was grown on glucose, proline, or succinate. However, in the presence of propionate, n-propanol, or heptanoate, MCS activity was 41-60 fold higher. Methylcitrate synthase (MSC) and citrate synthase (CS) activities in units/mg protein for *P. aeruginosa* when grown in two different substrates are shown below (Watson, 1983).

Table 1: MSC and CS activities in *Pseudomonas aeruginosa*

		Enzyme activity (units/mg protein)	
Strain	Growth Substrate	MSC	CS
Pseudomonas aeruginosa	Succinate	27	106
	Propionate	358	212

Works of Horswill et al. have shown *PrpC* of *S. typhimurium* to be a bifunctional enzyme with a preference for propionyl-CoA over acetyl-CoA. k_{cat}/K_m of acetyl-CoA and propionyl-CoA were 5 X 10^3 M^{-1} s^{-1} and 150 X 10^3 M^{-1} s^{-1}, respectively (Horswill, 1999).

As mentioned earlier, β-oxidation of odd-chain or branched length fatty acids yields propionyl-CoA and acetyl-CoA. While the generated propionyl-CoA enters the methylcitrate cycle, acetyl-CoA enters the citric acid cycle. Asp375 and His274 work together to form a neutral enol of acetyl-CoA. Asp375 functions as a base as it removes the alpha proton and His274 functions as an acid as it donates a hydrogen to the enolate oxygen of acetyl-CoA. In the next step the acetyl-CoA enolate attacks oxaloacetate forming citryl-CoA as His274 and His320 catalyze the condensation step. In the final step water is used to hydrolyze citryl-CoA to release the citrate and CoA (Remington, 1992). Even without a metal ion cofactor, citrate synthase is capable of forming carbon-carbon bond. The reaction of this enzyme proceeds in an ordered sequential kinetic mechanism with oxaloacetate binding before acetyl-CoA (Voet, 2006). Condensation of propionyl-CoA with oxaloacetate proceeds in a similar manner when the alpha proton of propionyl-CoA is removed by a base.

I.E1 Citrate Synthases in *B. subtilis*

Based on amino acid sequence and DNA sequence identity Jin and coworkers concluded that *B. subtilis* has two distinct citrate synthase homologs: *citA and citZ.* Amino acid sequence of *citZ* was shown to have 26 to 43% identity with other known citrate synthase proteins. When comparing *citA* and *citZ* they have 59 and 42% identity

at DNA sequence and at the amino acid sequence level, respectively, with each other (Jin, 1994). On the chromosome of *B. subtilis mmgD, citA*, and *citZ* are located in 217^{o}, 90^{o}, and 250^{o} region, respectively (Bryan, 1996; Jin, 1994). Glucose and glutamate are known to synergistically repress the activity of citrate synthase. While both glucose and glutamate repress the transcription of *citZ*; glucose represses the transcription of *citA* and glutamate in the presence of glucose stimulates the transcription of *citA*. The majority of citrate synthase activities are due to *citZ*; but when *citZ* is repressed, *citA* provides basal activity. Under different metabolic conditions, CitA and CitZ might catalyze the same reaction (Jin, 1994).

Since sequence similarity was noted between *mmgD* and citrate synthase, Bryan et al compared *mmgD* with other genes known to encode citrate synthase. *CtsA*, citrate synthase protein of *B. coagulans,* was found to have 80% similarity and 66% identity to the *mmgD* gene product of *B subtilis*. When compared to the gene products of *citA* and *citZ, mmgD* was 35% identical and 56% similar to *citA*; and with *citZ* it was 42% identical and 61% similar (Bryan, 1996).

In order to demonstrate the role of *mmgD,* Bryan et al. mutated the genes that are known to function as citrate synthase in *B. subtilis*. Strain SJB67 is a mutant lacking both *citA* and *citZ.* Strain EUX9504 is strain SJB67 with pSpacmmgD integrated into the chromosome. Integration of this plasmid places the IPTG inducible P_{spac} promoter just upstream of *mmgD*; thereby designing a conditional knockout mutant. Both strains were grown on liquid and solid medium with and without IPTG after they were subject to

identical treatments. After 36 hours of incubation on TSSA-glucose plates, the following results were seen (Bryan, 1996). SJB67 requires glutamate to grow (Jin, 1994).

Table 2: Results from conditional knockout mutants

Strain	With IPTG	Without IPTG
EUX9504	1 X 10^8 CFU/mL	5 X 10^3 CFU/mL
SJB67	No colonies	No colonies

When IPTG is present there is 20,000 fold increase in CFU (colony forming unit) from strain EUX9504. When grown in liquid media to assess growth defect of SJB67 by *mmgD*, the doubling time of SJB67 was >10 hours while the doubling time of EUX9504 was 2.5 hours. Hence, *mmgD* expression can overcome the growth defect of SJB67 (Bryan, 1996).

Bryan and coworkers showed that *mmgD* can also complement the *E. coli* mutant strain, W620, which lacks *gltA*, citrate synthase. pSpacmmgD was used to transform strain W620 to ampicillin resistance. Then, the tranformants were plated on M9-glucose agar plates with and without IPTG. Although colonies were present on both plates, growth was more abundant on the plates with IPTG. These data show that *mmgD* can complement *citA* and *citZ* in *B. subtilis* and *gltA* in *E. coli* and therefore, it functions as citrate synthase in the citric acid cycle (Bryan, 1996). These data, however, do not indicate that MmgD could function as methylcitrate synthase.

CHAPTER II

OVEREXPRESSION, PURIFICATION, AND CHARACTERIZATION OF MmgD FROM *BACILLUS SUBTILIS* STRAIN 168

II.A Project Goals

The overall goal of this project was to determine the biochemical function of *mmgD.* Our hypothesis is that *mmgD* encodes a 2-methylcitrate synthase and citrate synthase. To study this enzyme we needed purified MmgD protein, therefore, the first step was to clone and over-express the gene encoding the desired enzyme. Our next goal was to purify the protein and begin its characterization. MmgD characterization incorporated its kinetic activity analysis and validation of product formation.

II.A1 Cloning and Overexpression

To amplify the gene sequence a polymerase chain reaction (PCR) is performed. Prior to amplification, design of primers included mutagenesis to place specific restriction sites so that the PCR product can be inserted into a plasmid. The design was such that the gene would be placed in frame with a sequence that codes for a C-terminal His_6-Tag. An ideal method to purify the His-tagged protein is to use Nickel-Nitrilotriacetic Acid (Ni-NTA) column.

After successful amplification the gene is inserted into a vector. pET-28a is an ideal vector because this system is “the most powerful system yet developed for the

cloning and expression of recombinant proteins in *E. coli*" (Novagen, 2006). The cloning and expression region of pET-28a is shown in figure 7. This vector has an N-terminal His-Tag sequence along with a C-terminal His-Tag sequence, both of which are optional (Novagen, 2006). By choosing a specific restriction enzyme, the plasmid could be designed to contain either N-terminal His-Tag sequence or C-terminal His-tag sequence or both. His-tag sequences are included mainly for the purpose of protein purification by using Ni-NTA column.

In this system, transcription and translation signals of bacteriophage T7 controls the cloning of target genes in the plasmid while T7 RNA polymerase source induces gene expression (Novagen, 2006). The T7 promoter sequence is highly specific towards the T7 RNA polymerase. Another important feature in this vector is the presence of the *lac* operator, which can control the transcription of a gene. The presence of an optimized ribosome binding sequence on the mRNA aids in translation of a protein. This plasmid contains the T7 Tag sequence, which allows for protein detection with antibodies via the Western blot analysis. The thrombin cleavage sequence and kanamycin resistance sequence (not shown) are other important features of this vector.

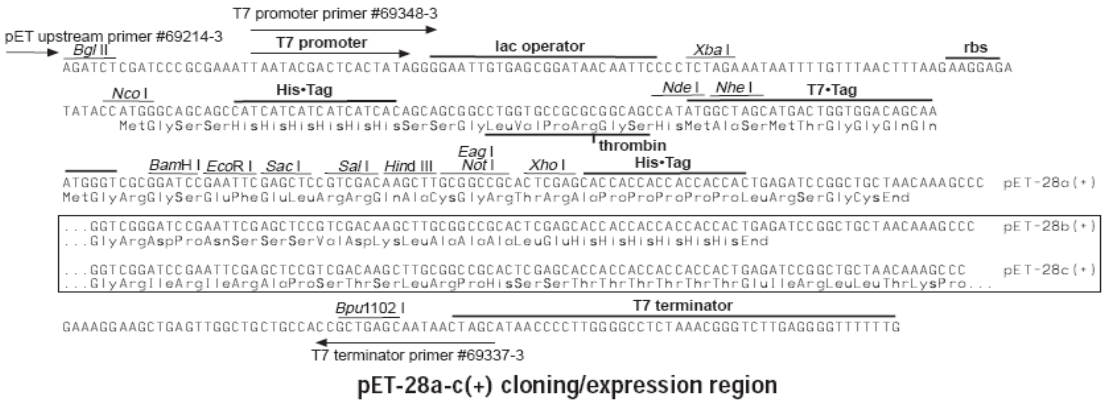

Figure 7: Cloning and Expression Region of pET-28a. (Novagen, 2006)

Specific restriction enzymes cut the genomic DNA and the plasmid. A DNA ligase anneals the PCR product and the vector together then the recombinant plasmid is transformed into a non-expression host, followed by retransformation into an expression host in order to transcribe target genes (Novagen, 2006).

E. coli stains BL21(DE3) are lysogens of bacteriophage λDE3, which contain the *lac*I gene, *lacUV5* gene, and the T7 RNA polymerase gene. The *lac*I gene, which is present in the *E. coli* genome (Novagen, 2006), produces a lac repressor protein via mRNA production. Transcription of lac genes, such as T7 gene 1, by RNA polymerase is inhibited when the lac repressor binds to the lac operator (Pearson). However, upon isopropyl-beta-D-thiogalactopyranoside (IPTG) addition gene transcription continues (Figure 8). IPTG binds to the repressor and inactivates it; and as a result RNA polymerase is able to bind with the promoter and thus initiate T7 gene transcription, which leads to the production of T7 RNA polymerase. The T7 RNA polymerase from *E. coli* binds with the T7 promoter region of the pET plasmid and transcription of target

genes initiates (Novagen, 2006), which ultimately leads to the translation of target proteins.

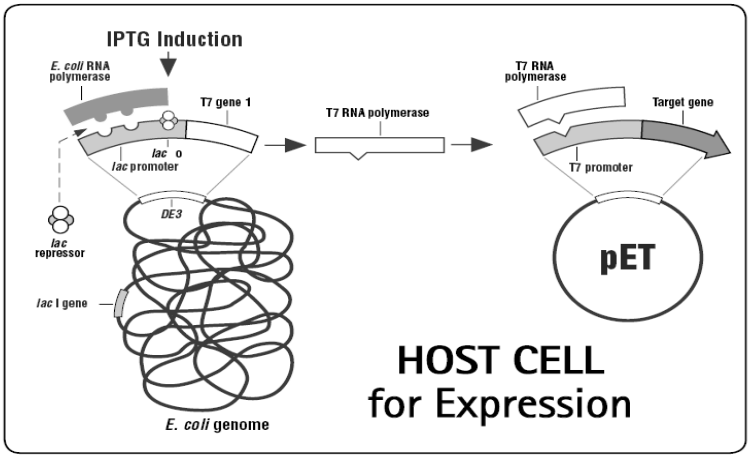

Figure 8: Over-expression System (Novagen, 2006).

II.A2 Protein Purification

Once the gene is successfully cloned and overexpressed, the next step is to purify the His-tagged product. To do so, Ni-NTA (Nickel-Nitrilotriacetic Acid) metal affinity chromatography is an ideal method. The 6xHis tag is uncharged, small, and doesn't affect compartmentalization, secretion, or protein folding within the cell. But most importantly, it securely holds the protein on metal-chelating surfaces. NTA, a tetradentate chelating adsorbent, binds with four ligand binding sites of the nickel ion, whose coordination sphere has six ligand binding sites. The remaining two sites are occupied by the 6xHis tag. Figure 9 shows the interaction between Ni-NTA matrix and 6xHis tag (Qiagen, 2003).

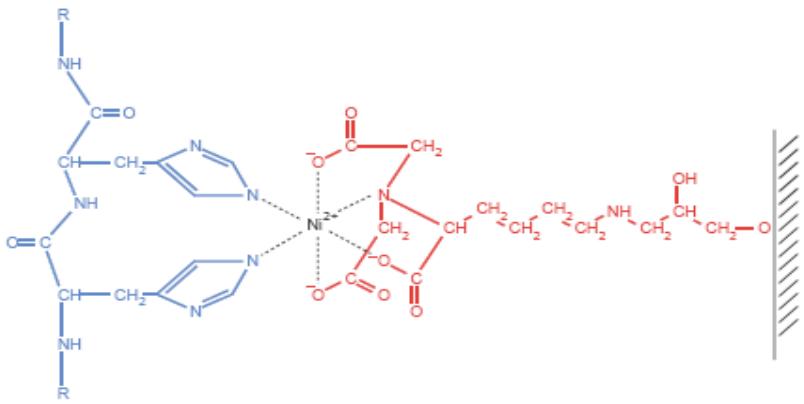

Figure 9: Ni-NTA Matrix and 6xHis Tag Interactions (Qiagen, 2003)

As the protein moves through the Ni-NTA column, it binds with the Ni-NTA matrix. The column is rinsed with low concentration of imidazole to remove unwanted protein, and to elute the desired His-tagged protein, the column is rinsed with a high concentration of imidazole. Due to structural similarity, imidazole competes with the histidine of 6xHis tag protein and binds to the Ni-NTA on the column (Qiagen, 2003).

II.A3 Enzyme Activity and Characterization

The *mmgD* gene was shown to function as a citrate synthase by mutant complementation (Bryan, 1997). However the enzyme has not been characterized at a biochemical level. Unlike *E. coli* and *Salmonella*, *B. subtilis* cannot utilize propionate as the sole carbon source (Bryan, 1996) and therefore, the methylcitrate cycle is a non-essential pathway and is probably used mainly during sporulation because σ^E most likely controls the methylcitrate synthase gene transcription (Sonenshein, 2002). As a part of characterizing this enzyme we investigated its potential dual-activity since basic genetics does not reveal any of that information.

In order to characterize the enzyme we monitored the enzymatic reaction. Citrate and methyl citrate synthase activity can be measured by monitoring the appearance of CoA-SH by using DTNB (5.5'-dithiobis-(2-nitrobenzoate)). The mercaptide ion is released by the reaction of CoA with DTNB and is absorbed at 412 nm while none of the other reagents (substrate or enzyme), involved in this reaction absorbs at this wavelength. The boxed reaction in figure 10 shows the DTNB reaction (Colowick 1969).

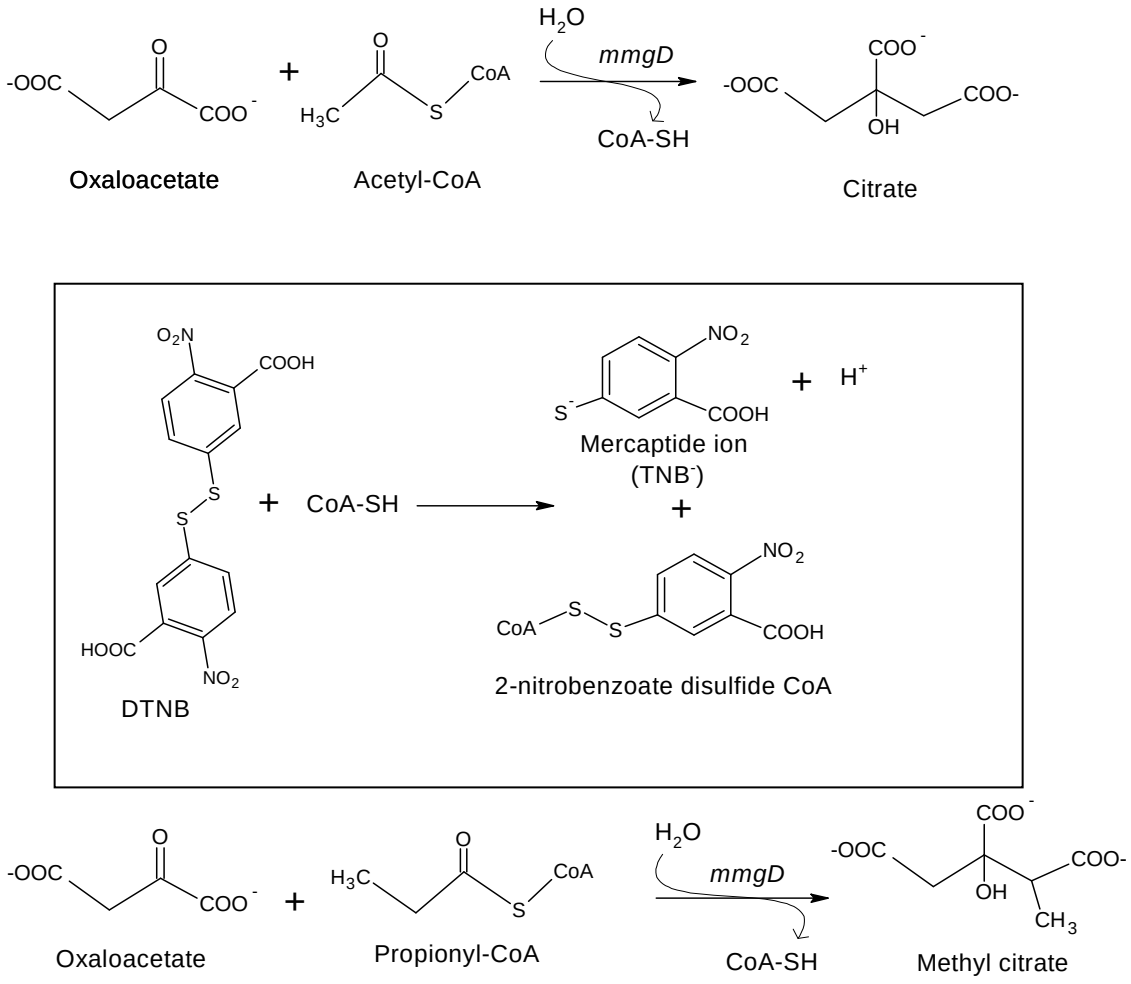

Figure 10: (Top) Formation of Citrate; (Middle) The DTNB Reaction; (Bottom) Formation of Methylcitrate.

Enzymes are pH and temperature dependent. The optimal pH range for citrate synthase reactions are 7.5 to 9.0 (Colowick, 1969). *S. typhimurium* citrate synthase and *E. coli* citrate synthase have an optimal pH of 7.5 and 7.4, respectively (Horswill, 1999; Textor, 1997).

II.A3.i Steady State and Michaelis-Menten Kinetics

Citrate synthase was believed to function by the rapid equilibrium random-order ternary-complex mechanism but the works of Beard and coworkers ruled out that theory in support of the ordered bi-bi mechanism, also called the compulsory-order ternary-complex mechanism (Beard, 2008). This mechanism is also seen in citrate synthase of *Thermoplasma acidophilum*, a thermophile, and pig heart, a mesophile (Kurz, 2000).

Citrate synthase is a multisubstrate system involving two substrates. The rate of reaction which is dependent on substrate concentration is explained by the following rate law, called the ordered bi-bi mechanism:

$$E + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EA + B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EAB \xrightarrow{k_3} E + P + Q$$

First, substrate A binds to an active site of the enzyme (E), then substrate B binds. After turnover in the EAB complex, products P and Q are released in an irreversible manner as the enzyme recovers. The following equation defines the behavior of an enzyme with two substrates for this mechanism.

$$V_o = \frac{V_{max}[A]_o[B]_o}{[A][B] + K_m^B[A] + K_m^A[B] + K_m^{AB}} \quad (1)$$

When one substrate concentration is very high, this equation conveniently reduces to a form similar to the single substrate Michaelis-Menten equation (2).

When initial concentration of one substrate, A, is held constant while varying the initial concentration of another substrate, B, a rectangular hyperbola defines the initial velocity. The horizontal asymptote of such hyperobola indicates the apparent maximum velocity, $V_{max,\ app}$, which is dependent on enzyme concentration. The K_m for the substrates with ordered bi-bi mechanism is defined by equation 3 and generally the K_m value is the concentration of the substrate when initial velocity is half of the apparent maximum velocity (Kyte, 2007). With that in mind, we used equation 2 to calculate initial velocity of MmgD reactions. Kinetic data were fit to this equation by non-linear regression. The apparent V_{max} and the apparent K_m for a set of reactions were generated from the SlideWrite program.

$$V_o = \frac{V_{max}[S]}{K_m + [S]} \quad (2)$$

$$K_m = \frac{k_2 + k_{cat}}{k_1} \quad (3)$$

$$k_{cat} = \frac{V_{max}}{[E]_T} \quad (4)$$

MmgD is thought to be bifunctional; therefore we investigated its preferred substrate. In order to do so, we had to calculate the apparent specificity constant, (k_{cat}/K_m). High specificity constant value indicates the preference of an enzyme for that substrate. We used equation 4 to calculate the apparent catalytic constant, $k_{cat,}$ also known

as the enzyme turnover number (Fersht, 1985). An ideal substrate would have high $k_{cat,app}$ and low $K_{m,\ app.}$

II.B Results and Discussion

II.B1 Amplification, Cloning, and Overexpression

MmgD gene from the *B. subtilis* stain 168 (BS168) was successfully amplified via the polymerase chain reaction (PCR). Figure 11 depicts the result of *mmgD* amplification. This gene has 1116 bps.

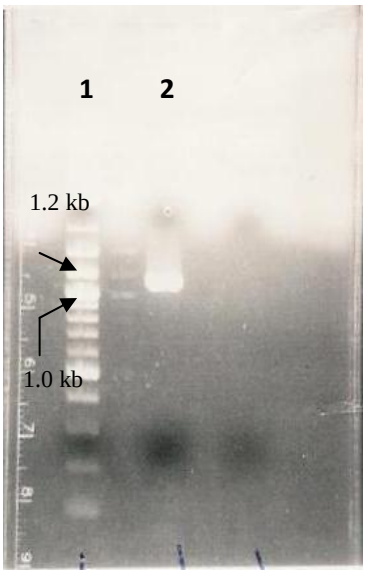

Figure 11: Agarose Gel of the PCR Product. Lane 1: DNA ladder; Lane 2: amplified *mmgD*.

The restriction enzymes *XhoI* and *NcoI* were used to cut the PCR product and the vector, pET-28a. T4 DNA ligase was used to ligate the plasmid and the PCR product at various insert to plasmid ratios. Then the ligation products were transformed with *E. coli*

DH5α, a non-expression host. To avoid toxicity to the cells it is important to have a non-expression host. Since DH5α doesn't encode T7 polymerase, there is no fear of transcribing the target genes at this point. The main purpose of a non-expression host is mostly to reproduce the plasmid.

After transformation the cells were spread on kanamycin plates because the plasmid confers resistance to kanamycin and were incubated overnight. There was a colony on the plate using 5:1 insert to plasmid ligation ratio. There were three colonies on the plate with 1:1 insert to plasmid ratio. Three out of four colonies showed growth on both liquid and solid media. Figure 12 shows PCR screening of the purified plasmid of *mmgD*/pET-28a/ DH5α. Purified BS168 was the positive control and pET-28a was the negative control. Neither controls were seen on the 1% agarose gel. After the first successful transformation, *E. coli* strain BL21(DE3), an expression host, was used for retransformation of the plasmid.

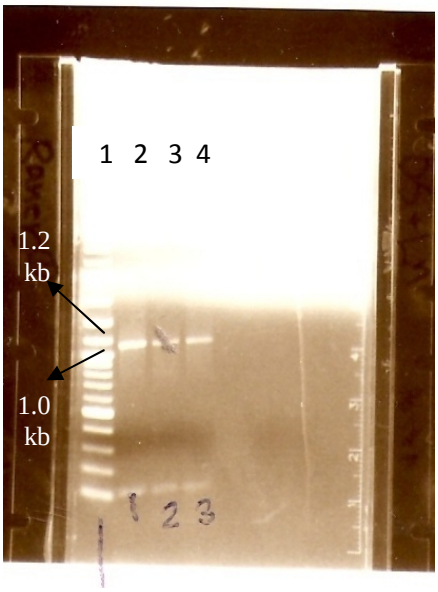

Figure 12: PCR Screening of the Purified Plasmid of *mmgD*/pET-28a/DH5α. Lane 1: standard. Lanes 2, 3, 4: transformants.

II.B2 Protein Purification

Using the pellets from a liter culture of *mmgD*/pET-28a/BL21(DE3) the process of protein purification began by making use of a Ni-NTA column. Several attempts were made to successfully produce soluble and purified protein. The main problem we encountered was precipitation of protein. When purified at 4 ^{o}C and at room temperature the eluted protein was not stable for very long. If the protein did not precipitate while eluting then it precipitated out of solution within minutes of being collected. Our first attempt was to change purification temperature; then we tried using buffer containing 10% glycerol, followed by buffer containing 10% glycerol and 0.1% triton X100. At this point, we tried changing the volume of buffers used. We lowered the volume by half and also tried doubling the volume. The ideal volume of buffers for our purification was ¼ times higher than the standard protocol. Although the eluted protein was not completely

stable at first, after leaving it overnight in the dialysis buffer and centrifuging the precipitates the following morning, we generated stable protein and hence continued its characterization.

When soluble protein was produced, its concentration was calculated based on standard of bovine serum albumin (BSA) using the Bradford assay. Figure 13 shows the standard curve of BSA.

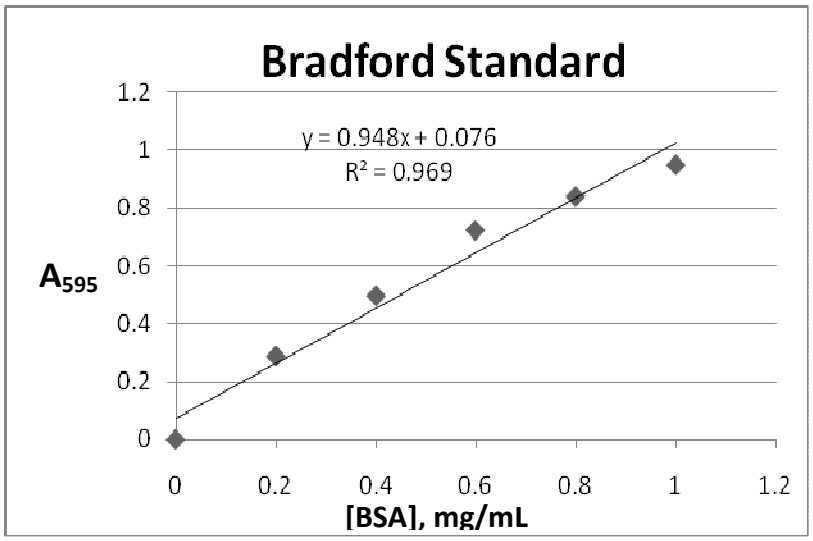

Figure 13: Standard Curve of Bovine Serum Albumin

Approximately 7.3 mg of MmgD was produced per half a liter of culture. After dialysis, the protein size and purity was analyzed via 10% SDS-PAGE (figure 14). Prior to running this gel the protein was stored at –80 ^{o}C for few weeks. The expected molecular weight of *mmgD* with six histidines is 42.88 kDa.

Figure 15 also represents 10% SDS PAGE of MmgD. This is a more recent gel picture with freshly purified protein. Approximately 1.0 mg of purified protein was

produced per half a liter of culture. The difference in the amount of protein generated could be related to when IPTG was added to induce *mmgD* production. IPTG was added when the OD_{595} of *mmgD*/pET-28a/BL21(DE3) was anywhere between 0.5 and 0.6. That could also be the reason why additional bands are seen in MmgD lane in figure 14 but are missing from MmgD lanes in figure 15.

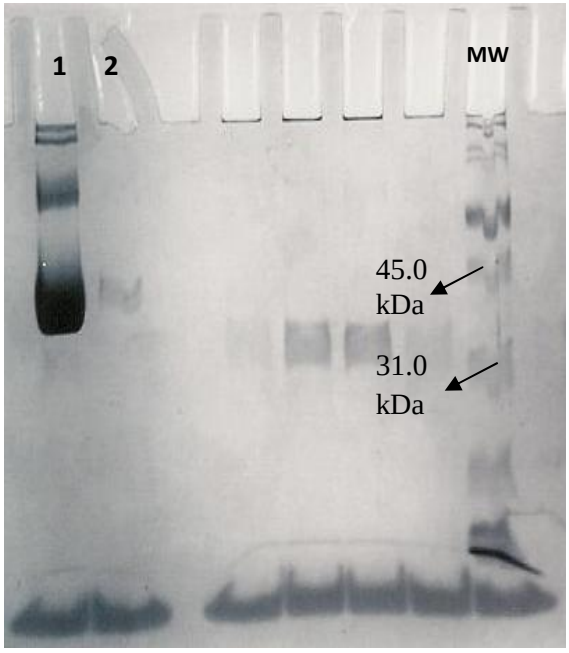

Figure 14: 10% SDS PAGE of MmgD. (1) Purified protein (2) 10% of the purified protein. (MW) standard molecular weight. Bands seen in between lane 2 and MW lane are from an unrelated experiment. Composition of loaded proteins: 75 µL of sample buffer and 25 µL of protein.

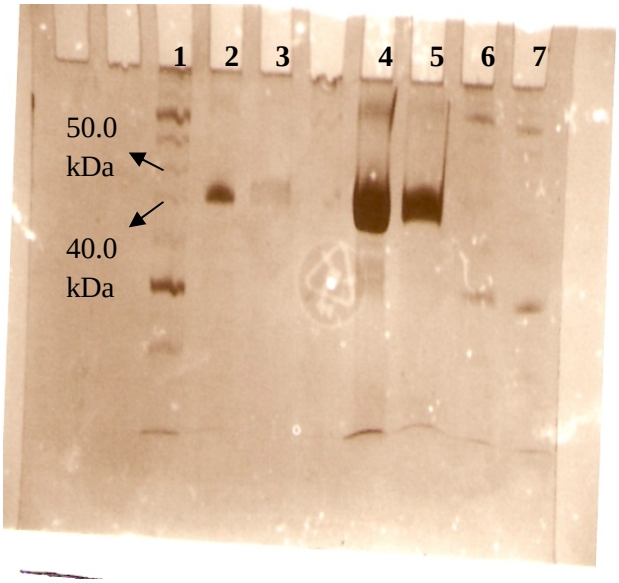

Figure 15: 10% SDS PAGE of MmgD. (1) Standard molecular weight. (2) 25 µL sample buffer and 75 µL of 10 % mmgD. (3) 75 µL sample buffer and 25 µL of 10 % mmgD. (4) 25 µL sample buffer and 75 µL of mmgD. (5) 75 µL sample buffer and 25 µL of mmgD. (6, 7) Traces of standard and the protein.

II.B3 Kinetic Activity Analysis

As mentioned earlier, the MmgD reaction can be analyzed by monitoring citrate and methyl citrate formation, which can be done by monitoring the appearance of CoA-SH by using DTNB (5.5'-dithiobis-(2-nitrobenzoate)). Using the UV Genesis Spectrometer the absorbances were monitored and recorded every 30 seconds. The mercaptide ion is absorbed at 412 nm while none of the other reagents (substrate or enzyme), involved in this reaction are absorbed at this wavelength. Furthermore, yellow color appearance also indicates CoA-SH formation. As soon as oxaloacetate was added to the reaction mixture, it started to turn yellow. When any one of the reactants (acetyl-CoA, buffer, or enzyme) were missing, the yellow color was absent, indicating a lack of reaction progression.

Purified MmgD has shown activity against acetyl-CoA and propionyl-CoA in the presence of oxaloacetate. The Michaelis-Menten (MM) constants for the above-mentioned substrates are shown in table 3 and the Michaelis-Menten curves are represented in figures 16 and 17.

Table 3: The Michaelis-Menten constants of acetyl-CoA, propionyl-CoA, and oxaloacetate.

MM constants	Propionyl-CoA	Acetyl-CoA	Oxaloacetate
Apparent K_m	8.7 μM	16 μM	17 μM
Apparent V_{max}	6.7 μM/min	5.8 μM/min	7.7 μM/min
Apparent k_{cat}	0.36 s^{-1}	0.32 s^{-1}	0.41 s^{-1}
Apparent k_{cat}/K_m	41 x 10^3 M^{-1} s^{-1}	18 x 10^3 M^{-1} s^{-1}	24 x 10^3 M^{-1} s^{-1}

k_{cat}/K_m of acetyl-CoA and propionyl-CoA of MmgD-like enzyme from *S. typhimurium* are 5 x 10^3 M^{-1} s^{-1} and 150 x 10^3 M^{-1} s^{-1}, respectively. Even though MmgD is shown to prefer propionyl-CoA over acetyl-CoA in *B. subtilis*, a reason why the preference is not as strong in this microorganism as it is in *S. typhimurium* could be because of interferences from the presence of unexpected protein. MmgD shown in figure 14 was used to perform all kinetic experiments. An unexpected band in lane 1 could be *E. coli* aconitase that was attached to MmgD during overexpression. This hypothesis of multi-protein complex has not been examined yet.

In figure 16, concentration of oxaloacetate is held constant as either propionyl-CoA or acetyl-CoA concentrations are varied. In figure 17, concentration of acetyl-CoA is held constant as the concentrations on oxaloacetate are varied.

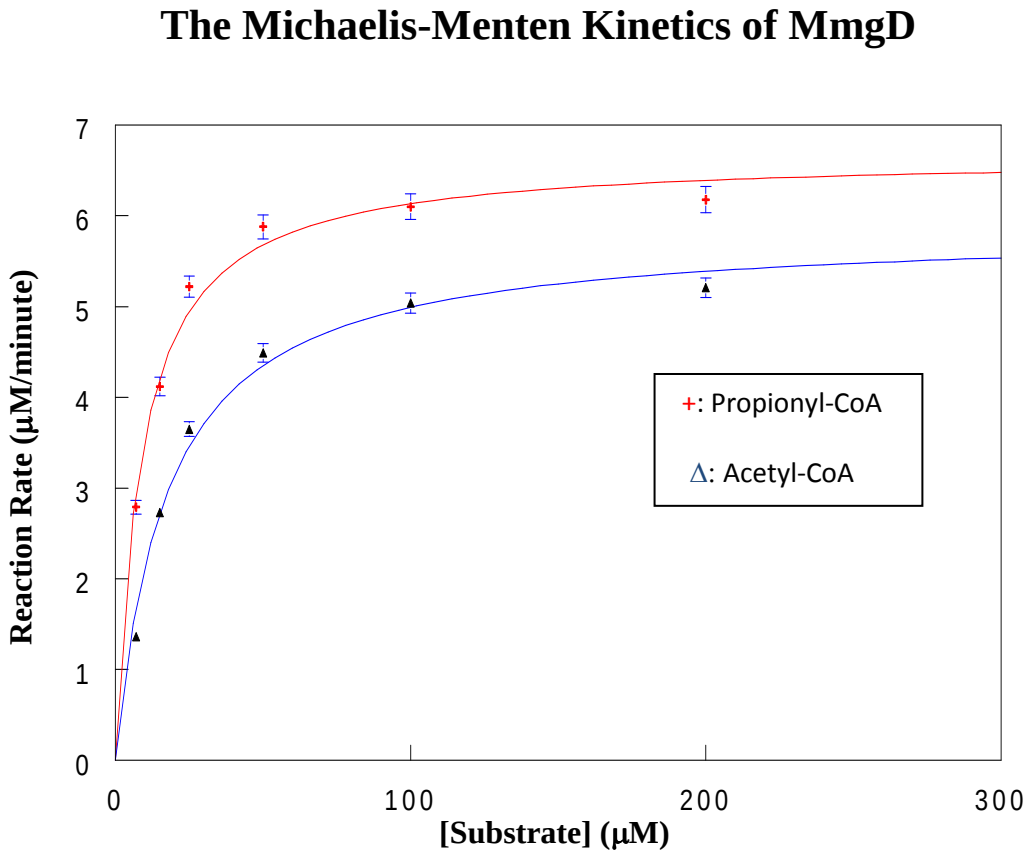

Figure 16: The Michaelis-Menten Curve of MmgD Reactions.

Generally, absorbance and concentration are directly related. When concentration of a substrate increases, absorbance increases. In the following figure as the concentration of oxaloacetate increased, the absorbance of DTNB decreased and as a result concentration of DTNB decreased as well. This type of graph represents a situation where a substrate is inhibited.

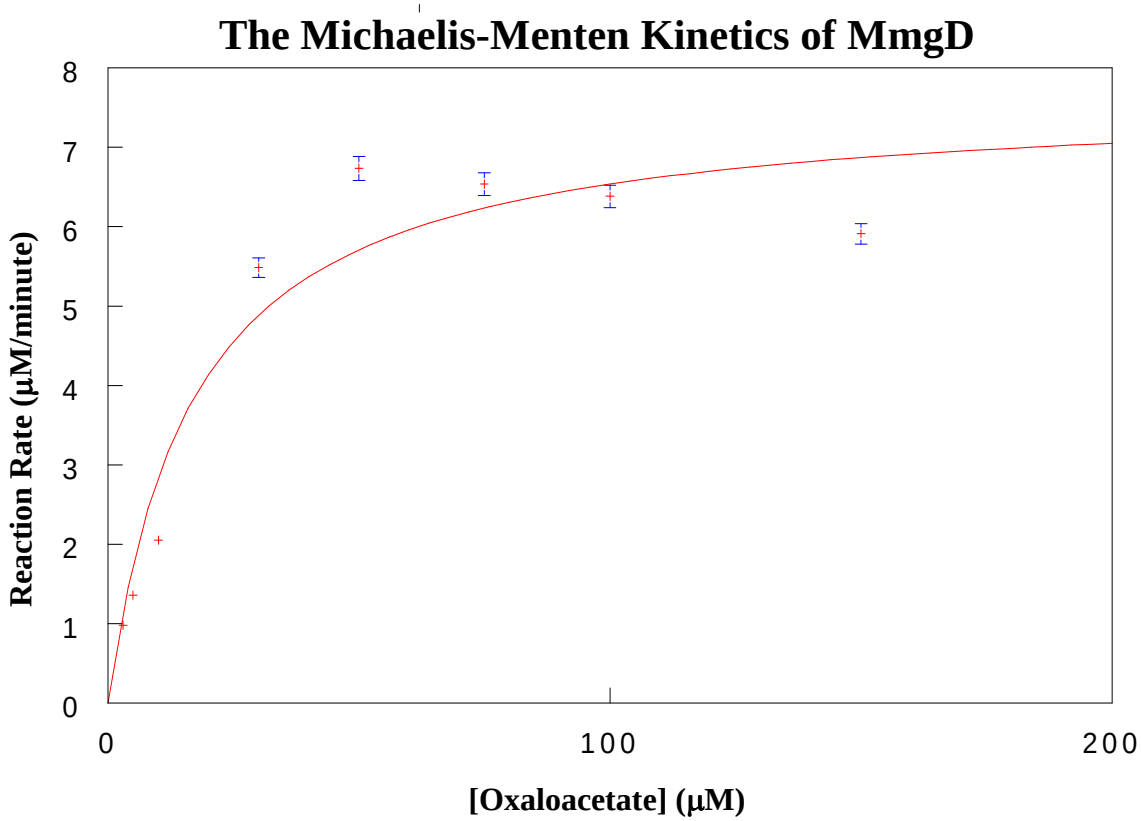

Figure 17: The Michaelis-Menten Curve of MmgD Reactions.

II.B4 Product Characterization

The formation of CoA-SH indicates that the reaction is taking place; in other words, acetyl-CoA or propionyl-CoA is reacting with oxaloacetate to form CoA-SH and possibly citrate or methylcitrate. The formation of the latter products are not indicated by this spectrometer analysis. Therefore, in order to show that citrate and methylcitrate were being produced in the presence of appropriate substrates and the enzyme, NMR and GC/MS analyses were chosen for the validation of product formation.

II.B4.i Nuclear Magnetic Resonance

Figures 18 and 19 show the spectra of ammonium citrate in D_2O. The differences in these two spectra are due to differences in sample preparation. Figure 18 represents 4.4 mM of ammonium citrate prepared in D_2O. The second spectrum represents 5 mM of ammonium citrate prepared in phosphate buffer; this mixture simulated the condition of enzymatic reaction. In other words, after decreasing the pH, the sample was extracted with ethyl ether, the solvent was evaporated, then the residue was mixed in D-chloroform and evaporated again. Finally, the residue was dissolved in D_2O and subjected to NMR analysis. The shifting seen in 19 is due to differences in sample preparations.

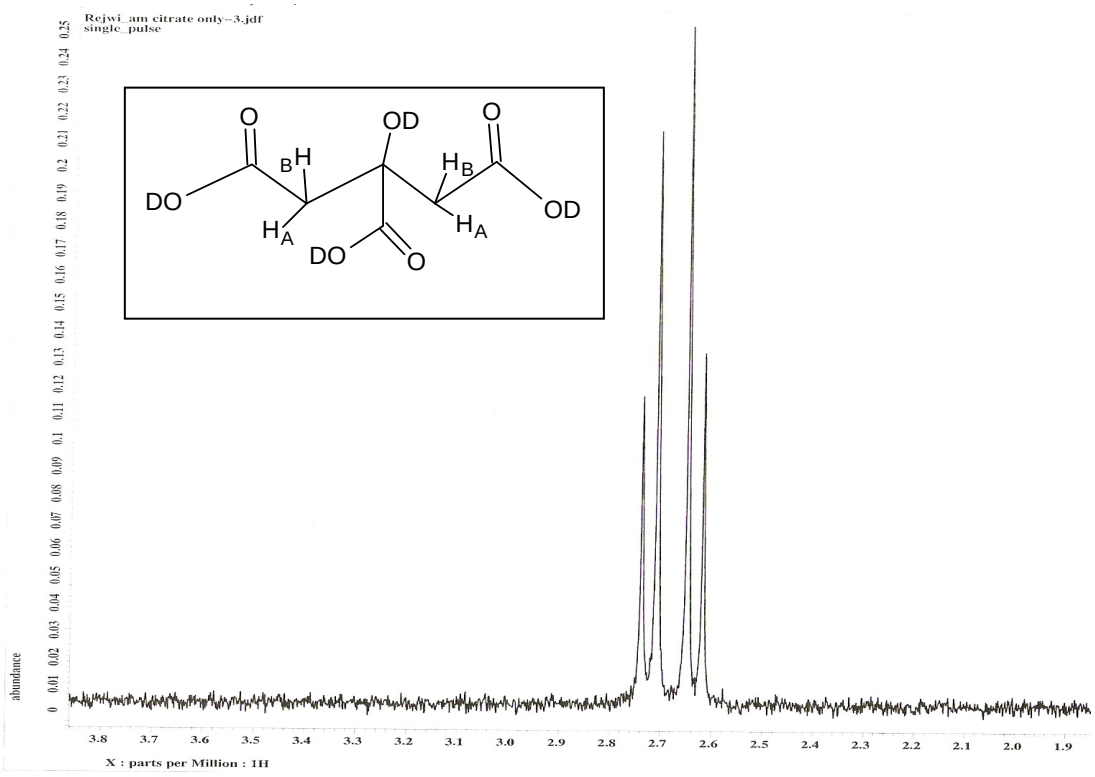

Figure 18: ^{1}H-NMR Spectrum of Ammonium Citrate in D_2O. The peaks are seen approximately between 2.6 and 2.75 ppm; the boxed structure represents citrate.

The above spectrum is consistent with an "AB spin system." Also, H_A and H_B are enantiotopic; if one of the hydrogen atoms is replaced with deuterium, the compound becomes chiral. In an achiral condition enantiotopic protons are equivalent.

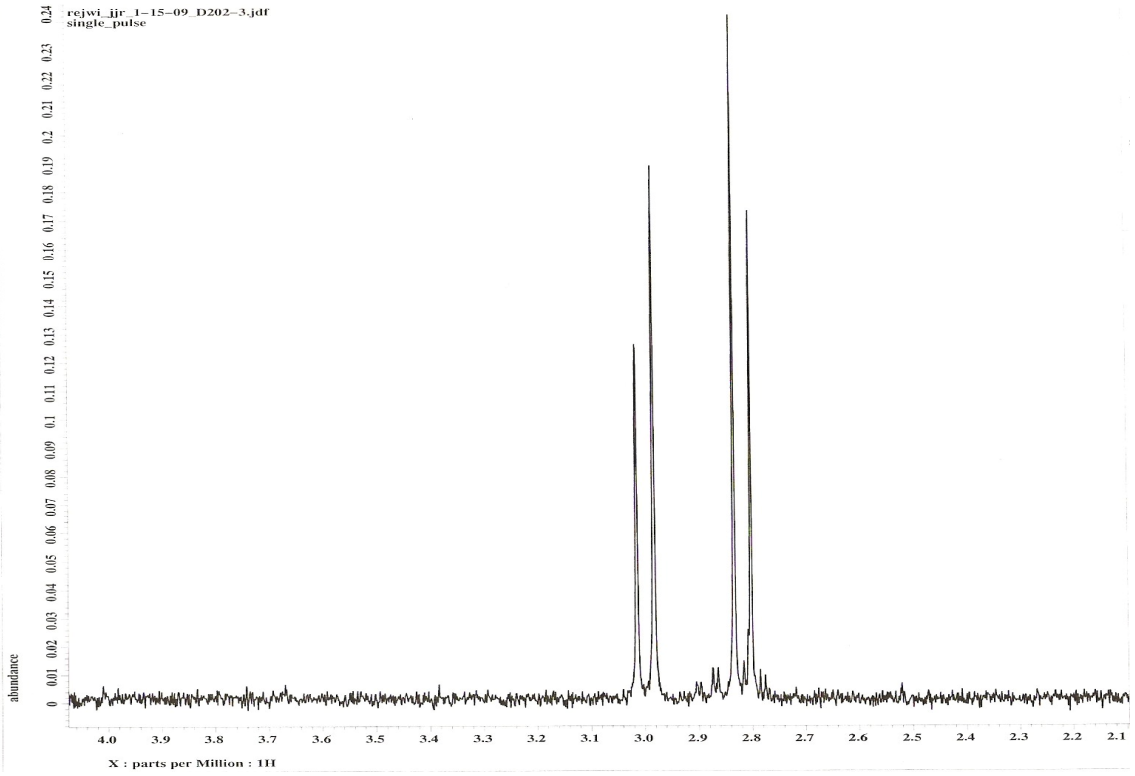

Figure 19: ^{1}H-NMR Spectrum of 5 mM Ammonium Citrate in Phosphate Buffer After Enzymatic Workup. The peaks, seen approximately between 2.8 and 3.02 ppm, are shifted up field.

^{1}H-NMR analysis of triton X100 (figure 20) revealed several peaks between 3.6 and 3.75 ppm and intense singlet at 0.65 ppm. ^{1}H-NMR analysis of glycerol (figure 21) revealed several peaks between 3.5 and 3.8 ppm. The spectrum of triton X100 and glycerol are shown in the following figures.

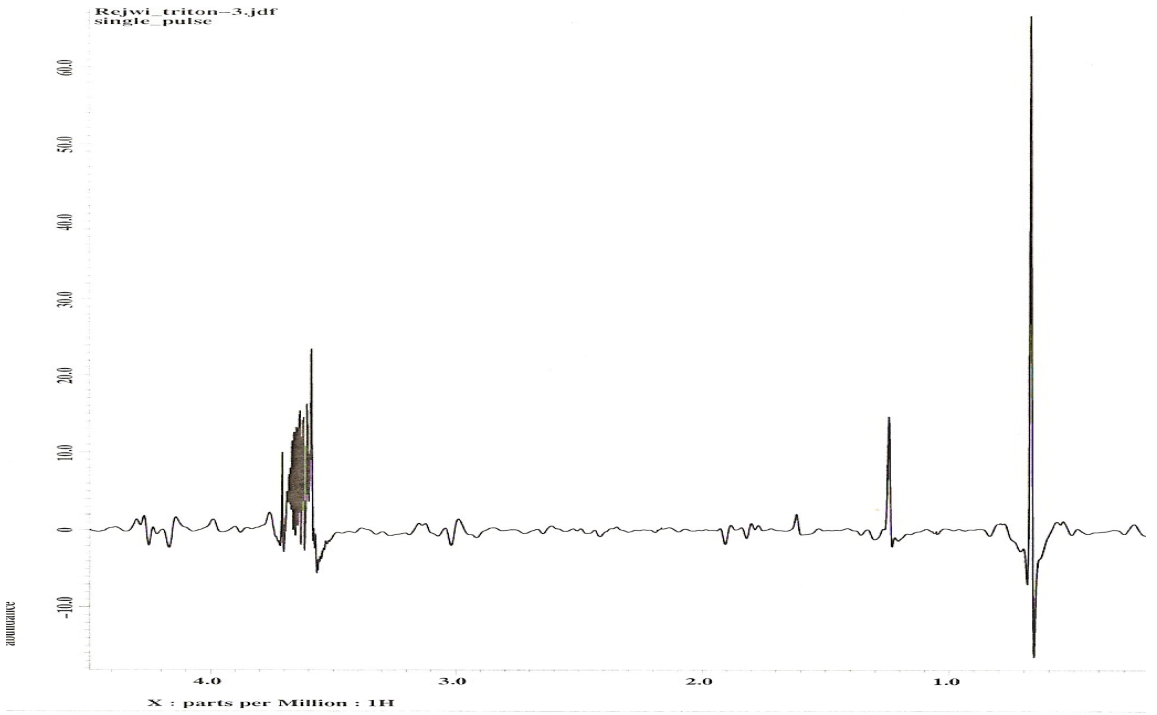

Figure 20: ^{1}H-NMR Spectrum of Triton X100. Singlets are present at 0.65 ppm and 1.32 ppm. Several peaks are also seen between 3.6 and 3.75 ppm.

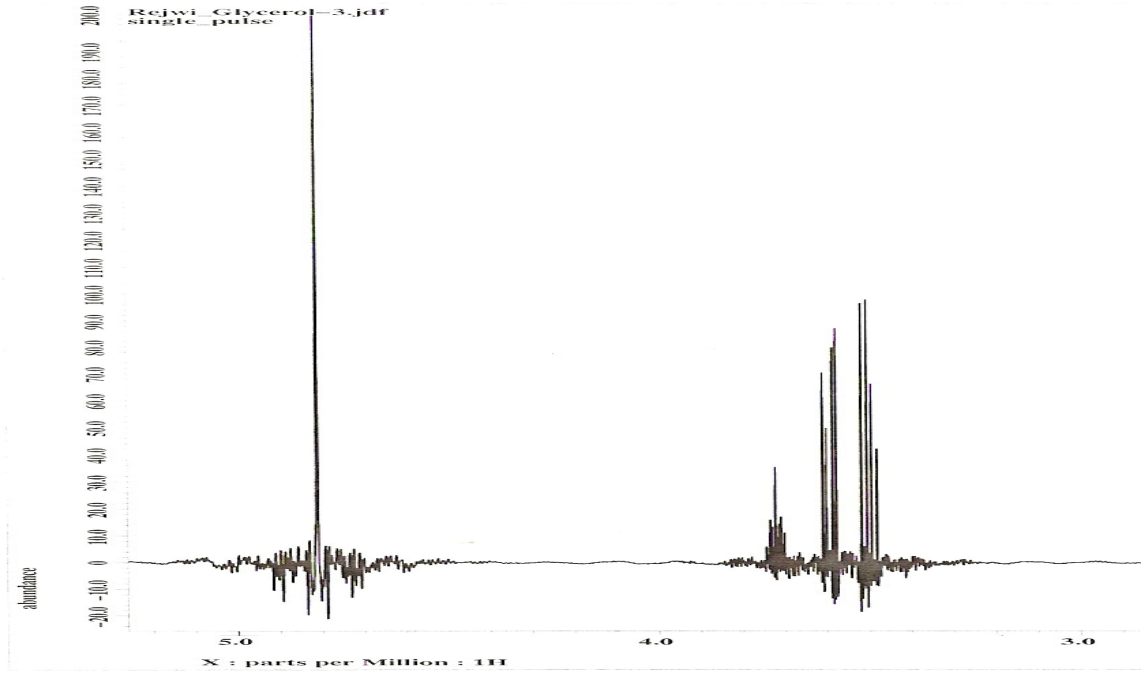

Figure 21: ^{1}H-NMR of Glycerol. D_2O peak is seen at 4.8 ppm. Several peaks of glycerol are seen approximately between 3.5 and 3.78 ppm.

Figure 22 shows the spectrum of oxaloacetate in D_2O. The only peak around 2.2 ppm represents oxaloacetate.

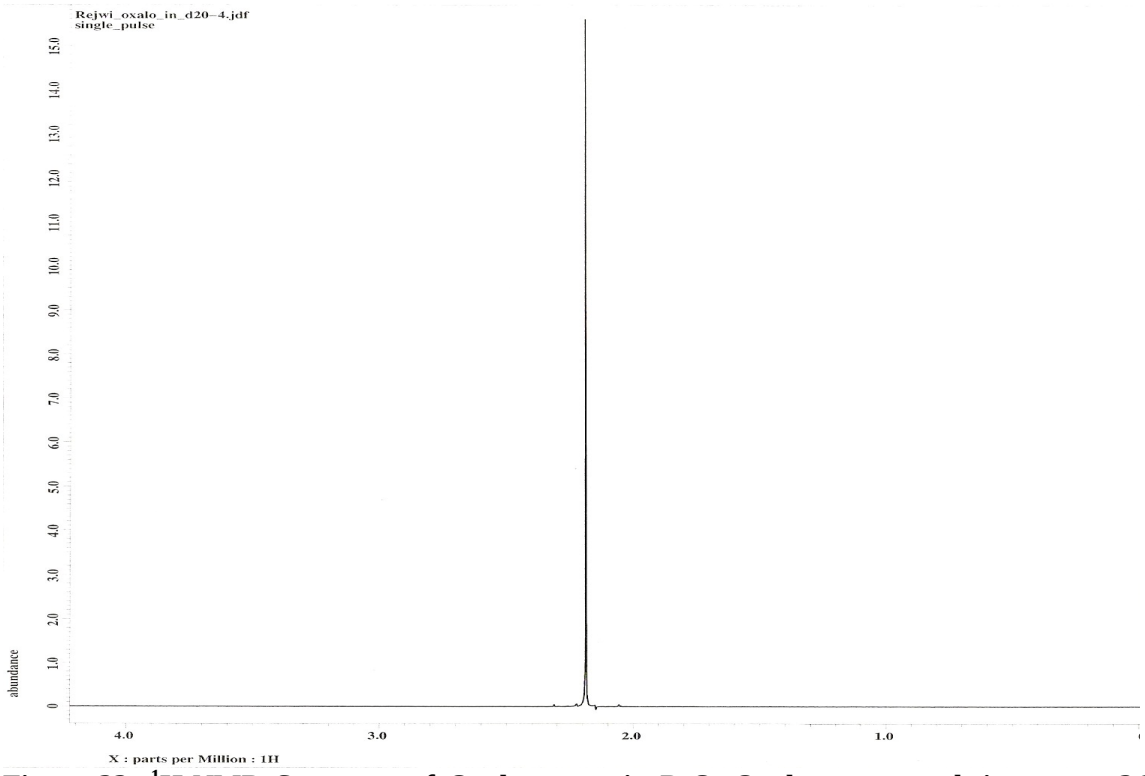

Figure 22: ^{1}H-NMR Spectrum of Oxaloacetate in D_2O. Oxaloacetate peak is seen at 2.2 ppm.

To generate figure 23, 5 mM ammonium citrate in water in the presence of MmgD and buffer underwent enzymatic workup. Notice the presence of several peaks between 3.51 and 3.78 ppm, which were not present in figures 18 and 19. Because MmgD was not used while preparing samples to generate spectrums 18 and 19, it is very likely that the sources of extra peaks seen in spectrum 23 represent materials used during

the protein purification, such as triton and glycerol. Figures 20 and 21 represent the spectra of those solvents.

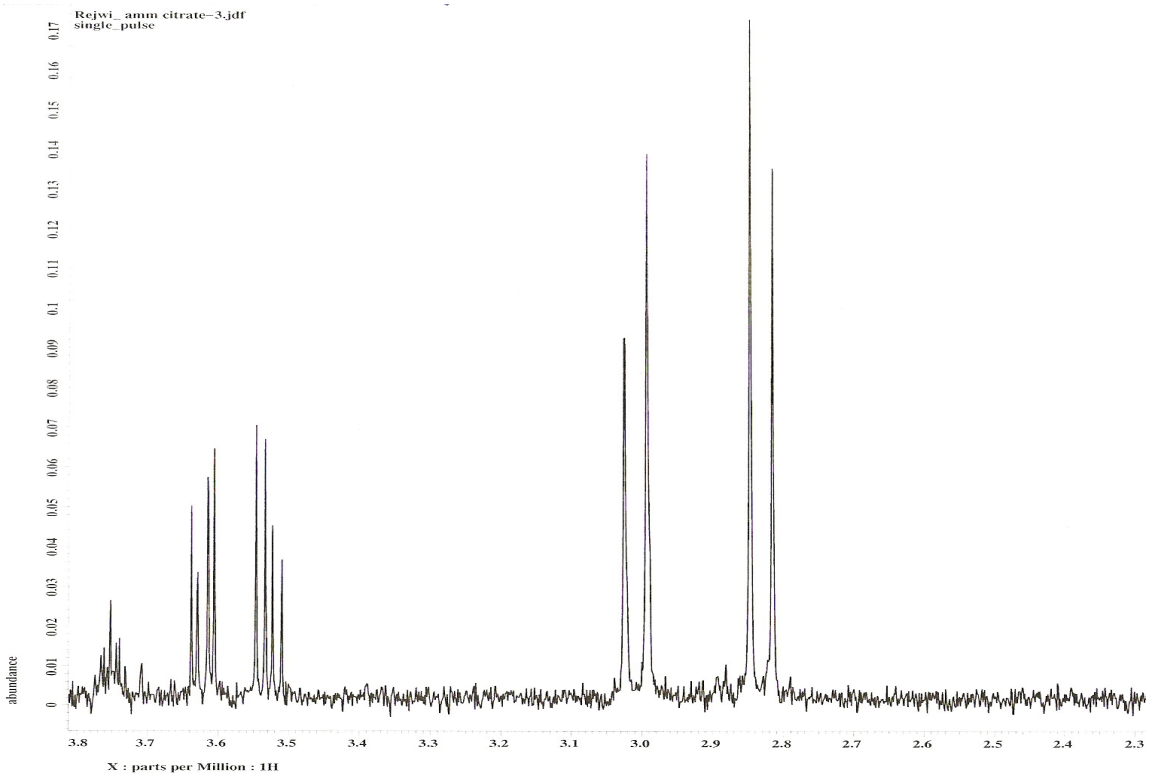

Figure 23: ^{1}H-NMR Spectrum of 5 mM Ammonium Citrate with MmgD and the Buffer. Citrate peaks are seen between 2.81 and 3.02 ppm. Additional peaks are present approximately between 3.51 and 3.78 ppm.

The MmgD reaction with acetyl-CoA and oxaloacetate in the presence of potassium phosphorous buffer (pH, 7.5) and water was incubated for 7.5 hours then quenched and extracted. After solvent evaporation, the residue was dissolved in D_2O and the mixture was subjected to proton NMR analysis. The following spectrum, figure 24, represents MmgD reaction with acetyl-CoA. The expected product, citrate, is seen

approximately between 2.8 and 3.02 ppm, just like the standard ammonium citrate. There are several unidentified peaks present as well. If the reaction does not go to completion, then excess substrate would be present. Suspecting the most intense peak to be oxaloacetate, another set of enzymatic reactions were incubated with decreased oxaloacetate concentration while keeping everything else the same. The generated spectrum (not shown) still contained an intense peak around 2.68 ppm. Oxaloacetate spectrum (figure 22) showed that oxaloacetate shows up around 2.2 ppm.

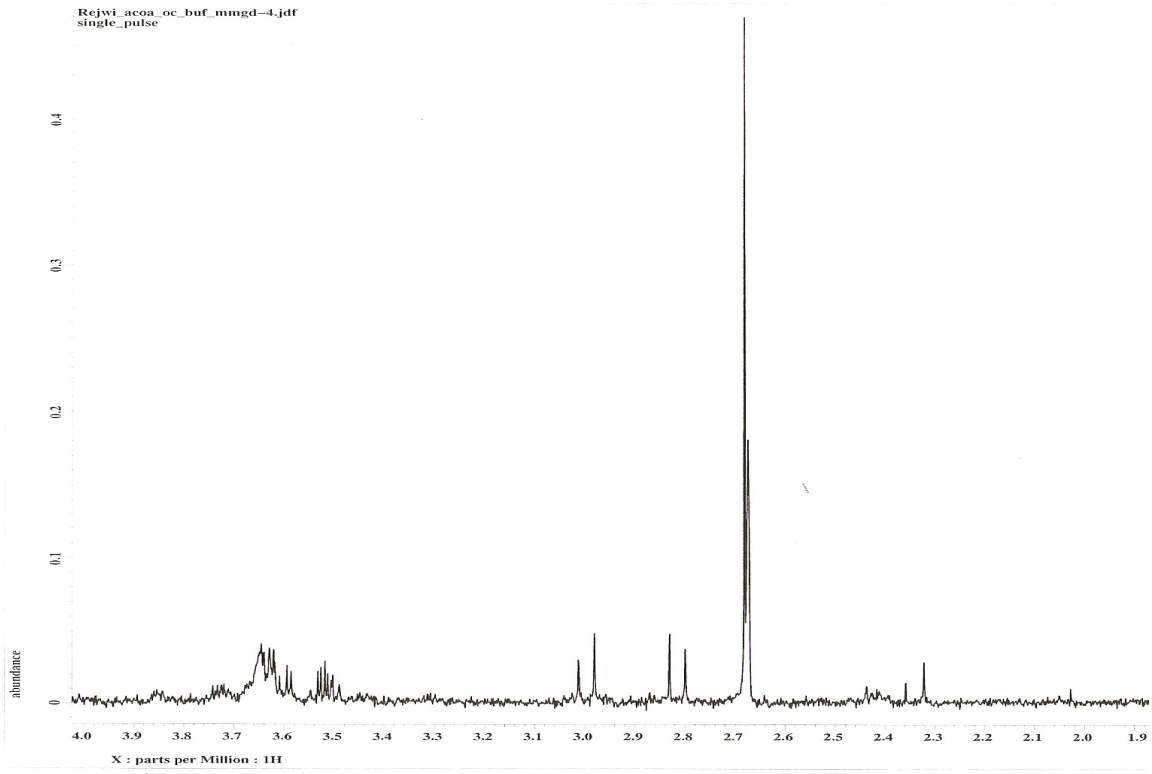

Figure 24: ^{1}H-NMR Spectrum of MmgD Reaction with Acetyl-CoA. Citrate peaks are seen between 2.80 and 3.02 ppm.

Figures 25 and 26 represent MmgD reactions with acetyl-CoA and propionyl-CoA, respectively. The conditions of these reactions are similar to the MmgD reaction with acetyl-CoA (figure 24). The only difference is in the purified MmgD; spectrum in figure 24 was generated using MmgD that was purified and stored in -80 ^{o}C for few months. The following two spectra were generated using freshly purified MmgD (figure 15), which lacks extra bands. MmgD (figure 14), which contained extra bands in lane 1 was used in a reaction that generated figure 24. The final concentration of MmgD was similar in all three sets of reactions. Although there is a peak at 2.68 ppm in the following spectrum, it is not as intense as seen in figure 23.

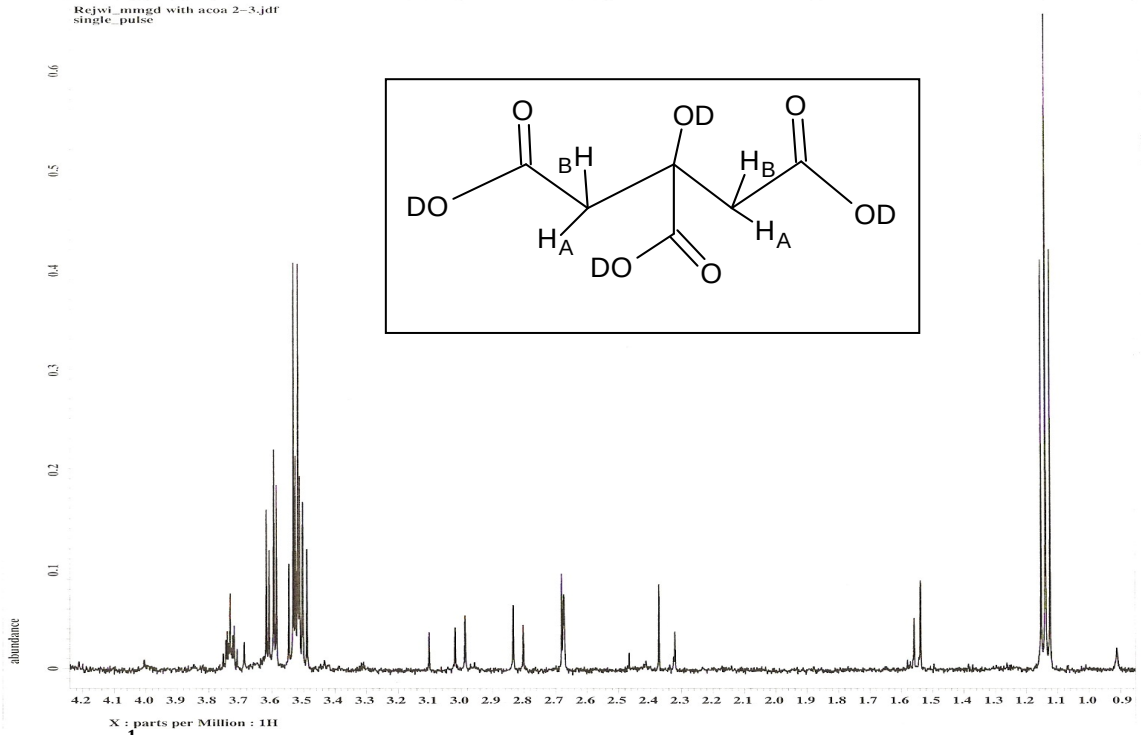

Figure 25: ^{1}H-NMR Spectrum of MmgD Reaction with Acetyl-CoA. Citrate peaks are present approximately between 2.8 and 3.02 ppm. A triplet between 1.14 and 1.17 ppm most likely represents ethyl ether.

Even though citrate is produced, like we suspected, there are several other peaks present, most of which have been identified.

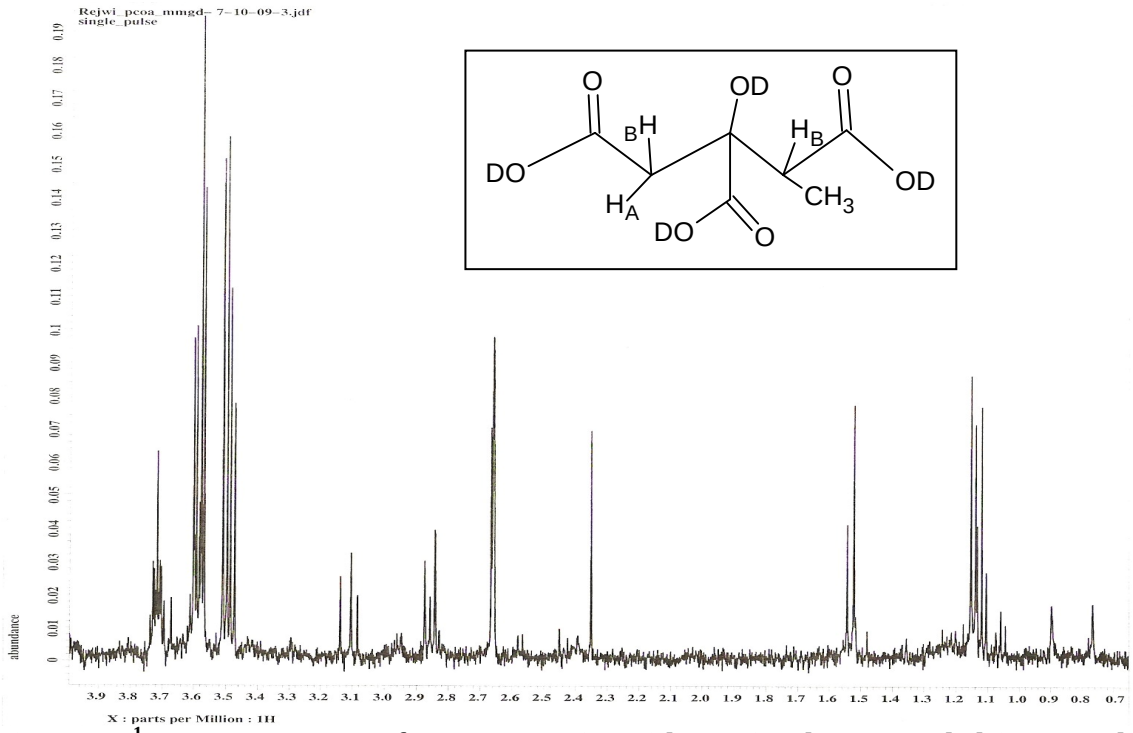

Figure 26: ^{1}H-NMR Spectrum of MmgD Reaction with Propionyl-CoA. Methylcitrate peaks are seen between 2.85 and 3.11 ppm and at 1.15 ppm.

Other peaks present could be from several solvents used during reaction workup. Proton NMR analysis of ethyl ether revealed several peaks: a triplet between 1.14 and 1.17 ppm , a quadruplet between 3.52 and 3.57 ppm, and also a singlet at 2.2 ppm.

II.B4.ii Gas Chromatography/ Mass Spectrometry

Figure 27 represents the spectra of derivatized standard citrate after GC/MS analysis. The first spectrum is a chromatogram showing retention time of citrate to be

33.61 minutes. According to experiments performed by Sun and coworkers the retention time of citric acid under similar condition was 33.00 minutes while characteristic ions had 215, 273, and 211 m/z. Characteristic ions in our experiment were 273 and 207 m/z; 211 m/z was present as well.

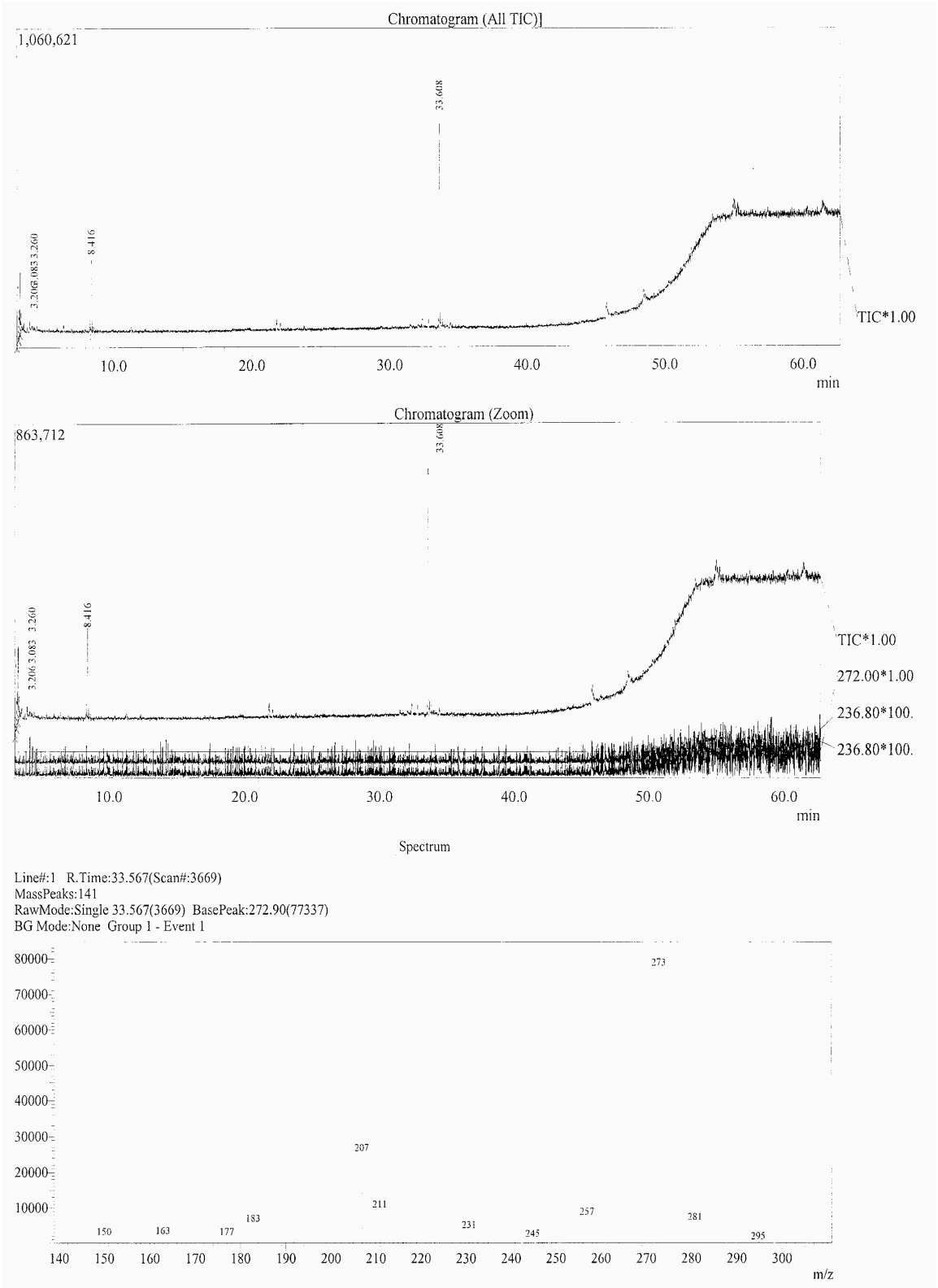

Figure 27: GC/MS Spectrum of Standard Citrate.

Due to time constraints, at this point, we are unable to optimize GC/MS analysis of MmgD reactions with acetyl-CoA and propionyl-CoA.

II.C Experimental

II.C1 Amplification and Cloning

The first step in this research was to amplify the *mmgD* gene using the Polymerase Chain Reaction (PCR) in a thermocycler. The upstream (UP) primer of *mmgD* had the following sequence: 5'- TTG AGG CCA TGG AGG AGA AAC AGC -3'. The sequence of the downstream (DN) primer of *mmgD* was: 5'- ATT CTA CCA GCT CCT CGA GTG ATT TTG TTT G -3'. The UP primer included an *NcoI* restriction site while the DN primer included an *XhoI* restriction site. The primers were specifically designed to place the *mmgD* in frame with a C-terminal His-tag sequence encoded by pET-28a.

The PCR samples were prepared in the following manner: 35.5 μL of ultra pure water, 10.0 μL of 5X high-fidelity (HF) buffer, 1.0 μL of stock dNTPs, 1.0 μL of *mmgD* UP primer, 1.0 μL of *mmgD* DN primer, 1.0 μL of genomic DNA from *Bacillus subtilis* strain 168 (BS 168), and 0.5 μL of stock Phusion polymerase were mixed together in a tube and inserted into the thermocycler. The initial activation step at 98.0 ^{o}C lasted for 30 seconds, then began the 30 cycles of reactions starting at 98.0 ^{o}C for 5 seconds, 63.0 ^{o}C for 10 seconds, and 72.0 ^{o}C for 30 seconds. After the 30 cycles the sample was held at 72.0 ^{o}C for 5 minutes and at 4.0 ^{o}C until the next step in the procedure.

The PCR product was subjected to 1 % agarose gel electrophoresis at 120 V. The gel was stained with Ethidium bromide and visualized with UV light. Upon validation of a successful PCR amplification, the PCR product was purified with the Qiagen PCR purification kit. The Qiagen QIAprep Miniprep kit was used to purify the pET-28a plasmid, which was prepared from *E. coli* strain DH5α.

Restriction digests followed purification of the PCR products and pET-28a. The tubes containing PCR products and pET-28a were incubated for 3.5 hours at 37 ^{o}C followed by heat inactivation of enzymes for 20 minutes at 65 ^{o}C in a thermocycler. The two restriction enzymes used were N*coI* and *XhoI*. Each tube contained 2.5 μL New England Biolabs Buffer 4, 0.3 μL of stock BSA, 14.2 of μL water, 1.0 μL of *NcoI,* and 1.0 μL *XhoI* and in addition, 6.0 μL of pET-28a and 10.0 μl of *mmgD* were added to tubes 1 and 2, respectively. The two samples and a 2-log DNA ladder were applied to 1% low melting agarose gel at 120 V at 4 ^{o}C. The PCR products and pET-28a bands were cut from the gel and placed in a 75 ^{o}C water bath and incubated until the agarose had completely melted. The next step was ligation using the T4 DNA ligase. Three different samples with various insert to vector ratio were prepared. All three sample tubes contained 11 μL of water, 2 μL of ligase buffer, and 1 μL of T4 DNA ligase. In addition to that, first, second, and third tubes contained, 5:1, 1:5, and equal insert to plasmid ratio by volume, respectively. The total volume in each tube was 20 μL. The sample tubes were placed in a thermocycler for 30 minutes at 37 ^{o}C, followed by 2 hours at 23 ^{o}C and then overnight at 16 ^{o}C.

Competent cells

Using 5 mL of Luria-Bertani (LB) a bacterial culture was grown at 37 ^{o}C then centrifuged at maximum speed until the supernatant was clear. Once the supernatant was discarded the cell pellets were suspended in 3 mL of cold 50 mM $CaCl_2$ and incubated in ice for 30 minutes. Once again the mixture was centrifuged at maximum speed for five minutes, then the supernatant was discarded. The DH5α cell pellets were resuspended in 3 mL of cold 50 mM $CaCl_2$ while BL21(DE3) cell pellets were resuspended in 1 mL of cold 50 mM $CaCl_2$.

Transformation

Respective DNA and a host cells were incubated on ice for 30 minutes, followed by incubation at 37 ^{o}C for 2 minutes and at room temperature for 10 minutes. Then 1.0 mL of LB was added to the mixture and incubated for an hour at 37 ^{o}C. After centrifuging for 5 minutes at 8000 rpm, 1 mL of supernatant was discarded and the cell pellets were resuspended in the remaining supernatant. Then the transformed cells were spread on an agar LB plate containing 30 mg/L of kanamycin (KAN) and incubated overnight at 37 ^{o}C.

From here on, composition of liquid LB media, 5 mL LB and 5 μL of KAN from 30 mg/mL stock, will simply be referred as liquid LB media. Solid media (KAN agar LB plate) and liquid media contain 30 mg/L of Kanamycin. Concentration of stock KAN is 30 mg/mL.

Chemically competent cells of DH5α were prepared according to "Competent Cells" protocol. 100 μL of competent cells were mixed with 20 μL of ligation products and "Transformation" protocol was followed.

After the overnight incubation there was one colony on the plate which originated from the 5:1 insert to plasmid ratio ligation. There were three colonies on the plate with 1:1 insert to plasmid ratio. Each of those colonies was grown on a liquid and a solid media. Both the tubes and the plates were incubated overnight at 37 ^{o}C. Generated plasmids were purified as before and were subjected to PCR screening.

II.C2 Overexpression

In order to overexpess the *mmgD* gene, we retransformed the plasmid into *E. coli* stain BL21(DE3). Prior to retransformation competent *E. coli* cells of strain BL21(DE3) were prepared according to "competent cells" protocol. 100 μL of the competent cells were mixed with one μL of plasmid and "transformation" protocol was followed.

One colony of mmgD/pET28a/BL21(DE3) was inoculated in liquid LB media and incubated overnight at 37 ^{o}C. The next morning, 2 mL from the starter culture was used to make a culture with 1 L LB and 1 mL KAN stock and it was shaken at 37 ^{o}C at 200 rpm and the absorbance was checked every hour to make sure the OD_{595} was between 0.5 and 0.6, at which point 1 mM IPTG was added to induce *mmgD* production. After induction, the culture was shaken at 37 ^{o}C at 200 rpm overnight.

The next morning, the culture was spun down at 7480 g for 30 minutes and the supernatant was poured out and only the pellets were stored in the -80 ^{o}C freezer. The next step was to purify the protein.

II.C3 Purification

The stored pellets from a liter culture were resuspended in 20 mL 1X binding buffer (5 mM imidazole, 500 mM NaCl, 20mM Tris-HCl, pH 7.9) and lysed with a sonicator on ice for 3 minutes. Then the lysed cells were centrifuged at 11500 g for 30 minutes. The supernatant was syringe filtered into a Ni-NTA column at 4 ^{o}C using a 0.45 μm filter. 20 mL 1 X binding buffer was poured through the column, then 12 mL of 1X wash buffer (60 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl, pH 7.9) was used to wash out non-His tagged proteins. Finally 12 mL of 1X Eluent buffer (200 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl, pH 7.9) was used to elute 1 mL fraction of *mmgD* at a time.

Each fraction was subjected to color assessment of Bradford assay to check for protein presence. A dark blue color of the mixture is indicative of a positive test and the intensity of blue color could correlate with protein concentration. As evident by color assessment, the first five fractions using the elute buffer contained protein with different concentration levels. However, there were white specks in the fractions, which indicated insoluble protein.

Due to protein precipitation, the next purification was performed at room temperature using the above-mentioned buffer conditions. Unfortunately, this temperature change did not prevent precipitation. Several attempts were made to generate soluble protein. Each attempt was done with freshly prepared 1 L culture using a colony grown from the stock BL21(DE3)/pET-28a/*mmgD*. Each attempt showed high protein concentration based on Bradford color assessment. At times, protein would fall out of

solution as it was being eluted from the column and at other times, it would fall out of solution after being collected. Regardless of insolubility, the solution showed presence of protein, as indicated by present of a very dark blue color when Bradford reagent was added.

To purify soluble protein several attempts were made; with buffers containing 10% glycerol, with lower buffer volumes, with higher buffers volume, and with buffers containing 0.1% Triton X100 and 10% glycerol. Finally, the following buffer conditions did not result in precipitated proteins.

Buffer conditions #1

All the buffers contained 0.1% Triton X100 and 10% glycerol. From a liter culture only half of the generated pellets were used to purify the protein. The same volume of buffers was used as before even though only half the cells were used for purification. The results were promising; as mostly soluble proteins were being produced. Bradford color assessment was positive and 10% SDS PAGE supported MmgD presence. The elute fractions containing protein were collected and left overnight in 4 L of 25 mM Tris-HCl buffer (containing 10 % glycerol and 0.1 % Triton and pH of 8) using snakeskin pleated dialysis tubing with 7,000 MWCO at room temperature. Due to the unstable nature of this protein, each time freshly prepared protein was used to analyze its activity. Each time the MmgD presence was verified by 10% SDS PAGE and Bradford assay. Buffer conditions #1 worked for few times and the protein purified under those conditions were used for several MmgD activity analyses. However, after a month these conditions did not give

the desired results; color assessment of the fractions was too faint, indicating lower protein concentration or the protein could have been too diluted.

Buffer conditions #2

Under this condition the volume of buffers used were lowered. 15 mL of Binding buffer, 9 mL of wash and elute buffer were used to purify the protein generated from half a liter of culture. Once stable proteins were produced, the sample was left in dialysis buffer as before. Not all the eluted fractions were clear; the eluent that was cloudy was excluded from dialysis. The following morning, concentration of the protein was measured by the Bradford assay and the MmgD presence was verified by 10% SDS PAGE.

Once the MmgD was successfully cloned, overexpressed, and purified to yield soluble protein, the next step was to determine its activity.

II.C4 Activity Determination

To determine MmgD activity, one mL reaction containing acetyl CoA (final concentration: 300 μM), DTNB (0.1 mM, in 1 M Tris-HCl, pH 8.1), MmgD (20 μL), water was prepared and used as a blank. As oxaloacetate (500 μM) was added to the reaction, then began the monitoring of absorbance at 412 nm by using a Genesis 10 UV scanning instrument. Next, MmgD activity using propionyl-CoA instead of acetyl-CoA was monitored. The concentrations of all reactants were the same. Enzyme concentrations were not taken in account when performing the above experiments.

To calculate kinetic parameters the absorbances of several reactions that had varying concentrations of oxaloacetate, acetyl-CoA, or propionyl-CoA were monitored and recorded. In a 500 μL reaction 0.1 mM of DTNB (0.1 mM, in 1 M Tris-HCl, pH 8.1),

1.31 X 10^{-2} mg/mL of MmgD and water were mixed. The reactions were monitored for five minutes and a reading was taken every thirty seconds. Concentration of oxaloacetate was held at final concentration of 300 μM as acetyl-CoA concentration varied from 7 μM to 200 μM. The absorbances were plotted against time in minutes and the slope of reaction was noted. Each reaction was performed at least three times with fresh reactants. The average slopes were plotted against respective concentration of a substrate, acetyl-CoA, in SlideWrite program. $V_{max, app}$ and $K_{m, app}$ values were generated from the program by non-linear regression. In order to have compatible units on both axis, concentration of DTNB (μM/min) was calculated using absorbance at certain concentration of a substrate according to the equation $A=\varepsilon bC$. ε of DTNB is 13,600 /M/cm. In this way the Michaelis Menten curve was obtained.

In a similar manner, the concentration of propionyl-CoA varied from 15 μM to 200 μM as the oxaloacetate concentration was held constant at 300 μM. The third set of reactions dealt with constant acetyl-CoA at 300 μM and varying oxaloacetate concentration from 3 μM to 150 μM.

II.C5 Product Characterization

Once catalytic constants for propionyl-CoA and acetyl-CoA along with substrate preference of MmgD were determined, the next step was to verify the product formation. It was expected that a reaction of propionyl-CoA and oxaloacetate would yield 2-methylcitrate while reaction of acetyl-CoA and oxaloacetate would yield citrate. Two different methods were implemented to verify the desired product formation.

II.C5. i Nuclear Magnetic Resonance

At first 5.0 μmol of propionyl-CoA, 5.0 μmol of oxaloacetate, 50 μmol of phosphate buffer (pH, 7.5) and 20 μg of MmgD were mixed in a mL of reaction mixture (Horswill, 1999) and incubated overnight at room temperature. The following morning the reaction was quenched by adding 2 M HCl and making the final pH <2. The reaction mixture was exacted 4 times with 5 mL of ethyl ether (Horswill, 1999). Since methyl citrate becomes protonated at lower pH, upon addition of ethyl ether it mixes with the organic solvent. To evaporate the solvent, the mixture was left in a rotory evaporator at 33 ^{o}C until there were no liquid present. Because even a trace of ethyl either could interfere with methyl citrate peak in NMR spectrum, the remaining residue was mixed with deuterated chloroform and was subjected to evaporation again. The residue was resuspended in 600 μL of 100 % D_2O and subjected to ^{1}H NMR analysis (Horswill, 1999).

Several overnight reactions with propionyl-CoA and acetyl-CoA were tried as the total numbers of reactions were increased to eight. Ethyl ether to mixture ratio was 5:1. Therefore, when 8 reactions were incubated, exaction was done with 40 mL each time. Finally, the buffer concentration and number of reactions were increased. Also, several overnight reactions, 7.5 hours reactions, 3.5 hours reactions and 3 hours reactions with higher number of reactions were attempted. Another attempt included lowering the oxaloacetate concentration to 0.10 mM while keeping rest of the parameters same in order to get better results.

NMR conditions:

The final concentrations of reactant were as follows: 33.5 mM of potassium phosphate buffer (pH, 7.5), 0.30 mM of oxaloacetate and acetyl-CoA and 1.20 μg/mL of *mmgD*. Water was added to make total reaction volume 1 mL. Fifteen reactions were incubated at room temperature for 7.5 hours. All reactions were quenched and extracted in a manner mentioned before. The solvent were evaporated in SpeedVac and the residues were dissolved in 100% D_2O and were subjected to ^{1}H-NMR analysis.

When working with propionyl-CoA, the reactant concentrations were similar to reactions involving acetyl-CoA.

Next, NMR analysis with reactions from kinetic activity analysis were tried. There were 5 different sets of reactions; each set consisted of 10 reactions with 0.5 mL final volume. Set 2 was a full reaction; 200 μM of acetyl-CoA, 1.31 x 10^{-2} mg/mL of mmgD, 300 μM of oxaloacetate, 50 μL of DTNB (3.96 mg per 10 mL of potassium phosphorous buffer, pH 7.5) and water to bring the volume up to half a mL. The reaction set 2 did not have any enzymes; set 3 did not have acetyl-CoA; set 4 did not have oxaloacetate; set 5 did not have the buffer. All reactions were incubated at room temperature for approximately 1.5 hours and the pH was lowered to <2. The reaction was extracted in ethyl ether, solvent was evaporated in speedVac, and the residues were dissolved in D_2O and were subjected to proton NMR.

With a fresh batch of proteins ^{1}H- NMR analysis of MmgD products with acetyl-CoA and propionyl-CoA were attempted. The 'NMR conditions' were same as before.

II.C5.ii Gas Chromatography/ Mass Spectrometry

In order to verify the product formation of MmgD reactions, GC/MS was used since it is more sensitive compared to NMR. Reactions were incubated, quenched, extracted, and evaporated just as mentioned before. The reaction product was derivatized with TMS (trimethylsilyl) before it was injected into GC/MS. The TMS reagent mixture was prepared by mixing MSTFA (*N*-methyl-*N*-trimethylsilyltrifluoroacetamide) and TMS-Cl (trimethylsilyl chloride) at 100: 1 ratio in anoxic environment (Suh, 1997).

The residue was mixed with 50 μL of TMS mixture and was heated in water bath for 15 minutes at 60 ^{o}C. Then, 1 μL samples were injected onto the GC column. The GC/MS protocol is similar to the one performed by Suh et. al. The standard citrate in ether, and MmgD product with acetyl-CoA and propionyl-CoA were analyzed.

II.D Conclusion

MmgD gene of *Bacillus subtilis* strain 168 was successfully cloned into pET-28a vector. The plasmid was transformed into *E. coli* strain DH5α followed by retransformation into *E. coli* strain BL21(DE3). MmgD transcription and translation was induced by IPTG addition. The Ni-NTA column was used to purify the C-terminal His-Tagged product with the best yield of soluble protein being 7.3 mg.

MmgD activity was determined spectrophotometrically by monitoring the appearance of CoA-SH by using DTNB as only the mercaptide ion is absorbed at 412 nm. MmgD was shown to be a bifunctional enzyme by demonstrating a citrate synthase and methylcitrate synthase activity. Citrate synthase catalyzes the first reaction in the

citric acid cycle, condensation between acetyl-CoA and oxaloacetate to produce citrate and CoA. Methylcitrate synthase catalyzes the first reaction in methylcitrate cycle, condensation reaction between propionyl-CoA and oxaloacetate to produce methylcitrate and CoA. Of the two substrates, MmgD slightly favors propionyl-CoA over acetyl-CoA. The apparent k_{cat} for acetyl-CoA is 0.32 s^{-1} and apparent k_{cat}/K_m is 18 x 10^3 M^{-1} s^{-1}. The apparent k_{cat} for propionyl-CoA is 0.36 s^{-1} and apparent k_{cat}/K_m is 41 x 10^3 M^{-1} s^{-1}.

MmgD product formation was verified via proton NMR analysis. Although citrate and methylcitrate peaks are seen when MmgD catalyzes reaction with acetyl-CoA and propionyl-CoA, respectively, there are several other unidentified peaks present as well. Ether peaks are shown to interfere with methylcitrate peaks. Several peaks of triton X100 and glycerol are also noted in the spectra. There are other methods that could give better results. GC/MS is more sensitive compared to NMR; therefore next course of action could be to make use of this instrument to further validate product formation.

In a nutrient depleted environment *B. subtilis* undergoes sporulation. A sigma E dependent promoter was shown to drive expression of *mmg* operon consisting six genes during the intermediate stage of sporulation (Bryan, 1996). MmgA and MmgC are known to function as acetyl-CoA thiolase (Reddick, 2008), and acyl-CoA dehydrogenase, respectively. Based on sequence homology, MmgB is similar to 3-hydroxybutryl-CoA. While MmgABC are involved in fatty acid metabolism, MmgE and YqiQ are thought to be involved in propionate metabolism based on sequence homology. Odd chain and even chain fatty acids in cell membranes of this microorganism are broken down to ultimately produce propionyl-CoA and acetyl-CoA. MmgD has been shown to condense each

substrate with oxaloacetate and enter a specific cycle. *MmgE* and *yqiQ* are further downstream from *mmgD*; therefore they are thought to be involved in the methylcitrate cycle. MmgE most likely functions as a methylcitrate dehydratase to produce methyl-*cis*-aconitate from methylcitrate. YqiQ most likely functions as methylisocitrate lyase by catalyzing the breakdown of methylisocitrate to produce pyruvate and succinate. The main function of the methylcitrate cycle is to convert propionyl-CoA to pyruvate on an eqimolar basis (Upton, 2007). MmgBE and YqiQ are being characterized while MmgC was characterized a year ago in the Reddick lab.

Based on information mentioned here along with previously published information, *mmg* operon is most likely used to fuel sporulation in *Bacillus subtilis*.

REFERENCES

1. Beard, D. A.; K. C. Vinnakota, Wu, F (2008), “Detailed Enzyme Kinetics in Terms of Biochemical Species: Study of Citrate Synthase.” PLoS ONE 3 (3):e1825.

2. Bryan, E. M.; B. W. Beall, et al. (1996). "A sigma(E)-dependent operon subject to catabolite repression during sporulation in Bacillus subtilis." Journal of Bacteriology **178** (16): 4778-4786.

3. Colowick, S. P; J. M. Lowenstein (editor). (1969) "Citrate Synthase." Methods in Enzymology: Citric Acid Cycle New York; Academia Press. pp: 3-5.

4. http://dbtbs.hgc.jp/COG/prom/mmgABCDE-yqiQ.html. Accessed: 02/05/09

5. Eichenberger, P.; S. T. Jensen, et al. (2003). "The sigma(Epsilon) regulon and the identification of additional sporulation genes in Bacillus subtilis." Journal of Molecular Biology **327** (5): 945-972.

6. Errington, J. (1993). "Bacillus-Subtilis Sporulation - Regulation of Gene-Expression and Control of Morphogenesis." Microbiological Reviews **57** (1): 1-33.

7. Fersht, A. (1985) "The Basic Equation of Enzyme Kinetics." Enzyme Structure and Mechanism. (2nd edition). New York, W. H. Freeman and Company, 98-117.

8. Fujita, Y., H. Matsuoka, et al. (2007). "Regulation of Fatty Acid Metabolism in Bacteria." Molecular Microbiology **66** (4): 829-839.

9. Horswill, A. R.; J. C. Escalante-Semerena. (1999). "Salmonella typhimurium LT2 Catabolizes Propionate via the 2-Methylcitruc Acid Cycle." Journal of Bacteriology. **181** (18): 5615-5623.

10. Jin, S. F. and A. L. Sonenshein (1994). "Identification of 2 Distinct Bacillus-Subtilis Citrate Synthase Genes." Journal of Bacteriology **176** (15): 4669-4679.

11. Kroos, L. and J. R. Maddock (2003). "Prokaryotic development: Emerging insights." Journal of Bacteriology **185** (4): 1128-1146.

12. Kunst, F.; N. Ogasawara; I. Moszer; A.M. Albertini; G. Alloni; V. Azevedo; M. G. Bertero; et al. (1997). "The complete genome sequence of the gram-positive bacterium *Bacillus subtilis.*" Nature. ***390*** (6657): 249-256.

13. Kurz, L. C.; G. Drysdale, et al. (2000). "Kinetics and Mechanism of the Citrate Synthase from the Thermophilic Archeon *Thermoplasma acidophilum.*" Biochemistry. **39**: 2283-2296.

14. Kyte, J. (1995) Binding of Substrates to the Active Site. Mechanism in Protein Chemistry New York, Garland Publishing, INC. 149-159.

15. Novagen (2006), pET System Manual. 11th Edition.

16. Pearson Education. "The *lac* Operon in *E. coli.*" BioCoach Activity. http://www.phschool.com/science/biology_place/biocoach/lacoperon/protein.html. Accessed: 06/15/2009.

17. Piggot, P. J. and D. W. Hilbert (2004). "Sporulation of Bacillus subtilis." Current Opinion in Microbiology **7** (6): 579-586.

18. Qiagen (2003). The Quiaexpressionist™. 5th Edition.

19. Reddick, J.J.; J.K. Williams (2008). "The mmgA gene from Bacillus subtilis encoded a degradative acetoacetyl-CoA thiolase." Biotechnology Letters. **30** (6):1045-1050.

20. Remington, S. J. (1992). "Mechanism of citrate synthase and related enzymes (triose phosphate isomerase and mandelate racemase)." Current Opinion in Structural Biology. **2**: 730-735.

21. Stragier, P. and R. Losick (1996). "Molecular genetics of sporulation in Bacillus subtilis." Annual Review of Genetics **30**: 297-341.

22. Sonenshein, A.L (editor). (2002) "*Bacillus subtilis* and its Closest Relative: From Genes to Cells." Washington, D.C., ASM Press, 289-312.

23. Textor, S., V. F. Wendisch, et al. (1997). "Propionate oxidation in Escherichia coli: evidence for operation of a methylcitrate cycle in bacteria." Archives of Microbiology **168** (5): 428-436.

24. Upton, A. M. and J. D. McKinney (2007). "Role of the methylcitrate cycle in propionate metabolism and detoxification in Mycobacterium smegmatis." Microbiology-Sgm **153**: 3973-3982.

25. Voet, D.; J. G. Viet, and C. W. Pratt. (2006) "Citric Acid Cycle" and "Lipid Metabolism." Fundamentals of Biochemistry Life at the Molecular Level. (2nd Edition) Hoboken, NJ. John Wiley & Sons, INC. 525-527; 638.

26. Watson, D., D. L. Lindel, et al. (1983). "Pseudomonas-Aeruginosa Contains an Inducible Methylcitrate Synthase." Current Microbiology **8** (1): 17-21.