VEGUNTA, YOGESH. M.S. Biochemical Characterization of *MmgB,* a Gene Encoding a 3-Hydroxybutyryl-CoA Dehydrogenase from *Bacillus subtilis* 168 *and* Genetic Evidence for the Methylcitric acid Cycle in *Bacillus subtilis* 168. (2011)
Directed by Dr. Jason J. Reddick. 112 pp.

Bacillus subtilis is an industrially important organism because of its ability to produce enzymes and antibiotics on a commercial scale. It is considered the Gram positive counterpart of *E. coli* in terms of its genetic and molecular biological accessibility. Also, it is a model organism for the study of sporulation, an example of prokaryotic cellular differentiation. Sporulation involves several groups of genes that encode apparent metabolic pathways, many of which remain uncharacterized at the biochemical level. The *mmg* (mother cell metabolic gene) operon in *B. subtilis* strain 168 is one of these groups, and is transcribed only during an early stage of sporulation. This operon contains *mmgABC,* which are similar to genes from fatty acid metabolism, and *mmgDE* and *yqiQ*, which encode homologs of enzymes involved in the 2-methylcitric acid cycle.

The focus of this work will be on the biochemical characterization of the *mmgB* protein, which is similar by sequence to 3-hydroxyacyl-coenzyme A (CoA) dehydrogenase from a wide variety of organisms. So far, we successfully cloned and overexpressed *mmgB,* and purified the protein at a yield of 4.0 mg/liter of culture. A spectrophotometric assay and mass spectrometry showed that this enzyme indeed possesses 3-hydroxybutyryl-CoA dehydrogenase activity (E.C. 1.1.1.157) for the production of acetoacetyl-CoA. This oxidation specifically requires $NADP^+$, and has an

optimal pH of 9.8. We will describe these results including the steady-state kinetics that the enzyme follows.

As previously said, the downstream 3 ORF's of the *mmg* operon – *mmgD, mmgE* and *yqiQ* are proposed to encode for the methylcitric acid cycle. Also, as a part of a master's thesis, it has been shown by the Reddick lab that *mmgD* encodes for a citrate/methylcitrate synthase with a substrate preference for propionyl-CoA over acetyl-CoA. This fact encouraged us even more with reference to the involvement of the *mmg* operon in the methylcitric acid cycle, a pathway for propionate metabolism. So, our goal was to create conditional knockout mutants of the *mmg* operon and study the growth characteristics of the organism by feeding studies using propionate as sole carbon source. Eventually, we would also want to perform some NMR analysis of the cultures of the mutants for the intermediates of methylcitric acid cycle/propionate metabolism.

BIOCHEMICAL CHARACTERIZATION OF *MMGB*, A GENE ENCODING A 3-HYDROXYBUTYRYL-COA DEHYDROGENASE FROM *BACILLUS SUBTILIS* 168 *and* GENETIC EVIDENCE FOR THE METHYLCITRIC ACID CYCLE IN *BACILLUS SUBTILIS* 168

by

Yogesh Vegunta

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
Master of Science

Greensboro
2011

Approved by

_____

Committee Chair

To Amma, Nana and Thammudu also Dr. Reddick

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 take this opportunity to whole-heartedly thank my advisor, Dr. Jason J. Reddick, for his valuable teaching, and extending unconditional support and leadership that gave me an amazing, most memorable and cherishable experience as a graduate student. I also would like to thank the other members of my committee, Dr. Gregory Raner for his enormous support during initial days of my career as a graduate student, and Dr. Alice E. Haddy, for her efforts in enlightening me about science. I also acknowledge the department of Chemistry and Biochemistry, The University of North Carolina at Greensboro especially for the generous financial support *via* their teaching assistantship program.

I wish to express my eternal gratitude to my family, for their emotional and financial support in taking initial steps of my journey towards my dream of becoming a scientist. I couldn't have made this long without them. Lastly, I thank my friends and loved ones who have been my rock and my soft place to land. Very special thanks to my dearest friends Balu and Chary.

I gratefully acknowledge generous financial support from National Science Foundation Award # 0817793 for my research and award 0420292 for MALDI-TOF-MS Instrument.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

BACILLUS SUBTILIS

I.A. Description

Bacillus subtilis (*B.subtilis*) is a ubiquitous, aerobic, endospore-forming, rod-shaped bacterium commonly found in soil, water resources and in association with plants[1]. It is the best characterized member among Gram-positive bacteria and is considered the Gram positive counterpart of *Escherichia coli* (*E.coli*). *B.subtilis* was one of the first organisms to be fully sequenced[1]. The *Bacillus subtilis* genome of 4,214,810 base pairs comprises 4,100 protein-coding genes. Among these genes identified only about 1000 have been experimentally characterized[2].

I.B. Significance

B. subtilis and its near relatives are an important source of enzymes such as amylases and proteases. The bacterium is commercially important because of its ability to produce the enzymes at gram/liter concentrations[1]. It has therefore been extensively studied for protein secretion and is under development as a host for the production of heterologous proteins[1]. *B.subtilis* is also employed in the production of a variety of secondary metabolites and chemical agents such as antibiotics and surfactants[1]. Apart from these, it also exhibits some interesting features like sporulation, an example of prokaryotic cell differentiation, and the development of natural competence[3, 4]. Because of the underlying mechanisms governing competence and sporulation, *B.subtilis* has

emerged as a popular prokaryotic model organism. Naturally competent *B.subtilis* cells have the ability to take up exogenous DNA and to integrate it into their genome via homologous recombination. This feature has made *B.subtilis* amenable to detailed genetic and molecular biology analyses[4].

For the above mentioned reasons, the genetics, biochemistry, physiology and molecular biology of *B.subtilis* have been studied extensively. Today, availability of powerful genetic tools facilitates the functional analysis of *B.subtilis*[1].

I.C. Sporulation

B. subtilis grows and divides by a vegetative cycle under normal environmental and physiological conditions[3]. However, upon onset of harsh environmental or physiological conditions such as nutritional limitation and cell cycle perturbation, the cell can undergo sporulation as a survival strategy[3]. Endospores are metabolically dormant and have the ability to survive harsh conditions and most importantly they can germinate back to a vegetative cell in favorable conditions[3].

Usually *B.subtilis* cells enter sporulation towards the end of the exponential phase or about the beginning of stationary phase of the growth period when nutrients become limiting[4, 5]. Under normal conditions, glucose is the preferred carbon source. However during scarcity, it makes use of other carbon sources such as fatty acids[5].

I.C.1. Overview of Sporulation in *B.subtilis*

Sporulation in *B.subtilis* is a primitive developmental process that results in endospores, which are small, dormant cells that exhibit remarkable resistance to heat, desiccation, radiation and chemical insult[3]. The precise identity of the signals inducing

sporulation is still an open question; however, recent studies suggest strategies by which *B.subtilis* integrates diverse signals responsible for the onset of sporulation[6, 7].

Although sporulation is induced because of starvation, it is not initiated immediately when growth slows down due to nutritional scarcity. Alternatives such as chemotaxis, production of antibiotics to destroy potential contenders, and secretion of hydrolytic enzymes to scavenge extracellular proteins and carbohydrates will be explored. It might also undergo induction of competence for uptake of exogenous DNA for consumption. So, sporulation is the last ditch response to starvation and is suppressed until alternative responses are proved inadequate. *B. subtilis* cells also monitor various internal conditions *viz*., chromosome integrity, and the state of chromosomal replication and functioning of the Krebs cycle before initiating sporulation. These controls are presumably to ensure that, once started, sporulation can be successfully completed[7].

Thus, *B.subtilis* decides to sporulate by integrating a wide range of environmental and physiological signals culminating in the activation of Spo0A, a trigger for sporulation[4, 6-8]. It is a response regulator component of the two-component signaling systems typical to bacteria and also a transcription factor. All these signals are channeled into a phosphorelay which is actually responsible for the activation of Spo0A *via* phosphorylation (Spo0A~P)[4, 6, 7].

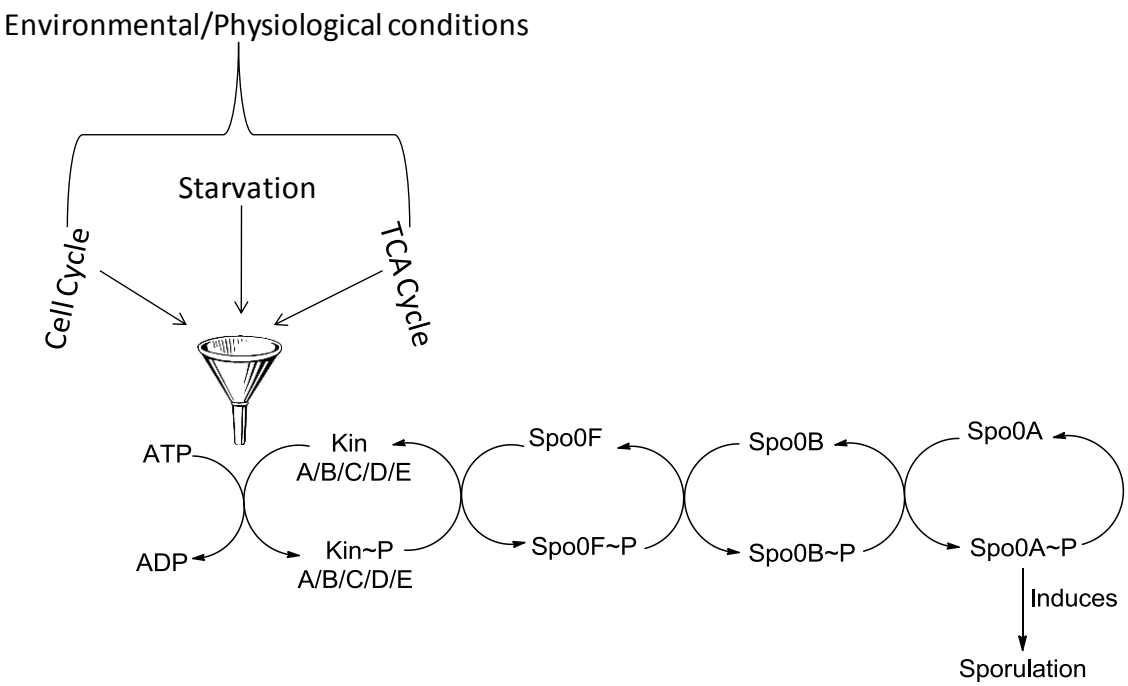


Figure 1: Two-component regulatory systems/Phosphorelay[4] (adapted).

I.C.1.a Two-component regulatory systems

Initiation of sporulation is controlled by proteins belonging to the family of two-component regulatory systems. These signal transduction systems are used in bacteria to control gene expression in response to physiological signals, instead of kinase cascade pathways that play pivotal roles in signal transduction in eukaryotes. Two-component regulatory systems consist of an upstream sensor component which usually is a histidine protein kinase and, the downstream response regulator component, which usually but not always will be a transcription factor[4, 6, 7].

Histidine protein kinases undergo an initial autophosphorylation step in response to signals by phosphorylating a conserved histidine residue with ATP as phosphate donor. The response regulator then interacts with its cognate histidine protein kinase resulting in the transfer of this phosphate form the histidine kinase to the response regulator. This phosphate then enters the phosphorelay[4, 6, 7].

I.C.1.b. Phosphorelay

The phosphate entering the phosphorelay causes activation of Spo0A - Spo0A~P (Figure 1), a key event in controlling sporulation. There are five histidine protein kinases, namely KinA, B, C, D and E, which contribute to the production of Spo0A~P. However, unlike most response regulators, Spo0A does not obtain phosphate directly from cognate histidine protein kinases. Rather, the phosphate is first transferred from kinase(s) to Spo0F, a single domain response regulator, then to Spo0B and finally to Spo0A leading to Spo0A~P. This series of inter-protein phosphotransfer reactions is called a Phosphorelay[4, 6, 7].

I.C.2. Genetics/Biochemistry/Molecular Biology/Physiology

Sporulation is a perturbation of the normal cell cycle. This process in *B. subtilis* proceeds through a well-defined series of morphological stages. It is controlled by an intricate genetic arrangement. The activity of the associated genes is controlled by various DNA (Deoxyribose Nucleic Acid) binding proteins & RNA (Ribose Nucleic Acid) polymerase sigma factors[3, 4].

Sporulation begins with the accumulation of Spo0A~P. It initiates the transcription of genes and operons which are associated with different stages of

sporulation. Throughout the process, different sigma factors of RNA polymerase appear which result in the stepwise transcription of different sets of sporulation genes. This in turn leads to a series of morphological stages for sporulation[3, 4, 6-8].

Series of morphological stages of sporulation (Figure 2):

- Stage O-vegetative cycle.
- Stages I & II - polar division and chromosome segregation followed by cell specific activation of transcriptional factors in both forespore and mother cell.
- Stage III – engulfment of forespore by mother cell.
- Stage IV, V & VI – cortex and coat formation around the endospore.
- Stage VII – Maturation, lysis of mother cell and release of mature spore.

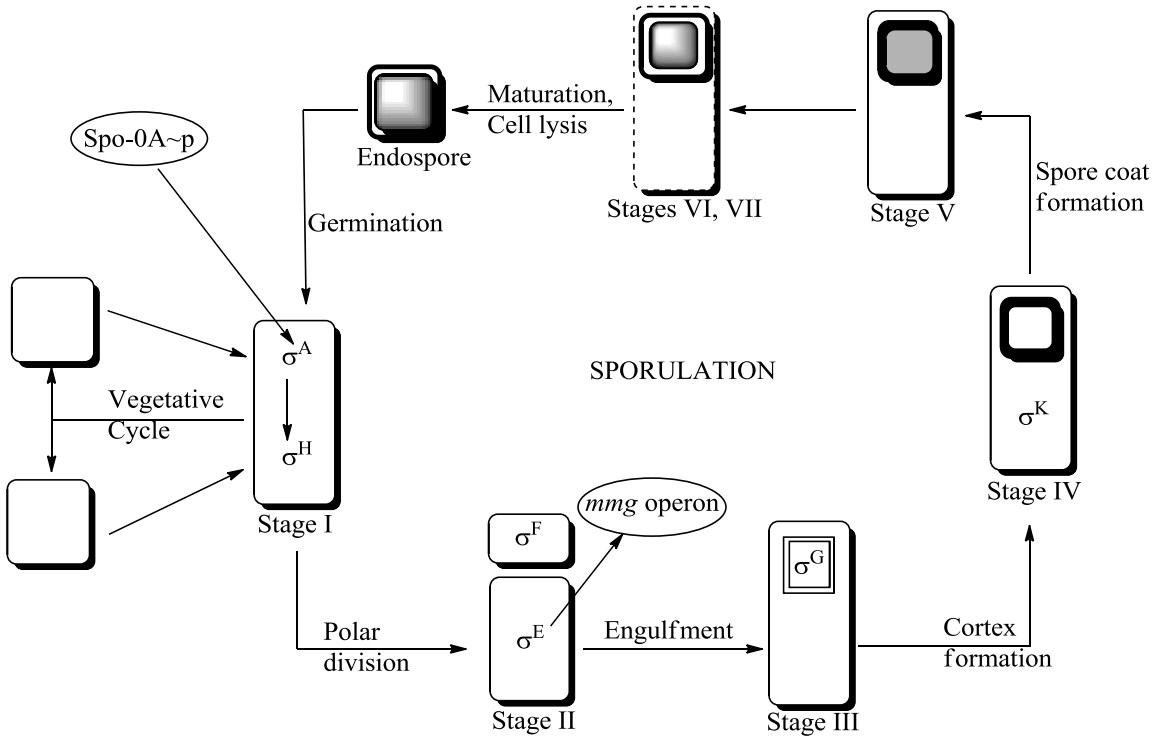


Figure 2: Series of Morphological stages during Sporulation & associated sigma – factors[9] (adapted).

Understanding the regulatory strategies governing sporulation can have broad implications on related signal transduction processes. Consequently, a considerable amount of research has been done in this area. Despite the availability of massive amounts of sequence information in databases and equally powerful computational tools it is not possible to assign precise functional significance to a particular coding sequence based on just sequence information and computational data. Also it is important to relate identified genome structure to intermediary metabolism, since they complement each other. Hence, experimental analysis is the definitive approach as it was before in order to characterize a particular gene[1, 3, 5-7].

I.D. The σE controlled *mmg* operon

Many metabolic enzymes are induced during the various stages of sporulation. Nevertheless, very little is known about their effects. In an effort to identify genes that are induced during the intermediate stages of sporulation, Bryan et al. discovered the *mmg* operon controlled by the σ^E transcription factor. Since σ^E is specific to mother cell, the operon is named "mmg," which stands for **mother cell metabolic gene**[10].

The *mmg* operon consists of six open reading frames (ORF's) namely *mmgA, B, C, D, E* and *yqiQ*. Bryan et al. discovered *mmgA* through part of *mmgE*. The genome sequence in 1997 (Kunst et al) revealed the rest of *mmgE*, and that this operon also has a 6th gene, *yqiQ* (Figure 3)[10].

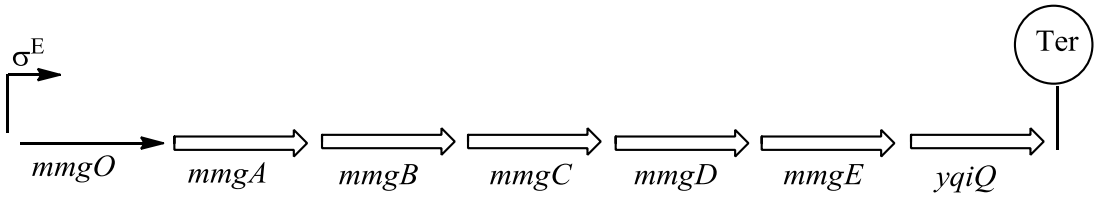


Figure 3: Schematic of *mmg* operon from B.subtilis[10] (adapted).

Sequence homology data is suggestive that the operon is associated with the metabolic activity of the cell. All the ORF's show high similarity to well - established metabolic enzymes. Sequence homology data suggests that *mmgA*, *mmgB* and *mmgC*, are involved in fatty acid metabolism (β-oxidation), while the latter half of the operon, which includes *mmgD*, *mmgE* and *yqiQ,* encodes a part of the methyl citric acid cycle[10-12].

Bryan et al. also discovered that the transcription of the *mmg* operon is under control of σ^E, *mmgO* and CcpA. *MmgO* is a 14 base-pair catabolite responsive element

(CRE) located in the promoter region, downstream of the transcription start site of *mmg* operon. CcpA is a protein which binds to *mmgO* and facilitates the repression of transcription by glucose. This means that the operon is transcribed only by σ^E (early sporulation in the mother cell) and when glucose is absent. These findings reinforce the idea of involvement of *mmg* operon in energy metabolism during sporulation[5, 10].

The Reddick group has thus begun characterizing each of the members of *mmg* operon. The Reddick lab has successfully shown that *mmgA* encodes a degradative acetoacetyl-CoA thiolase[13]. *Spencer Russell, a former Master's student from the Reddick lab has proposed that mmgB encodes a 3-hydroxybutyryl-CoA dehydrogenase and worked towards its characterization.* Also, he showed that *mmgC* encodes for an acyl-CoA dehydrogenase[14]. Rejwi Acharya has shown that *mmgD* is a bifunctional enzyme showing both citrate synthase and methyl citrate synthase activities with slight preference for propionyl-CoA over acetyl-CoA[15]. The gene *mmgE* is proposed to exhibit 2-methyl citrate dehydratase activity based on sequence homology[10, 12]. Currently Grant Hardesty, another Master's student from the Reddick lab is on his way to biochemically characterize this protein. *YqiQ*, the final member of the operon, is proposed to act as a 2-methyl isocitrate lyase. William Booth, a former Master's student from Reddick lab has worked on this system and could partially establish this activity[16]. The two metabolic pathways which involve these types of enzymes are discussed in detail in the next section.

I.D.1. Fatty acids/β-oxidation

Fatty acids occur in all living organisms. They are an important source of carbon and also energy. Fatty acids are classified (Figure 4) into straight and branched chain. Furthermore the branched chain fatty acids are divided into iso, anteiso and ω-acyclic[17].

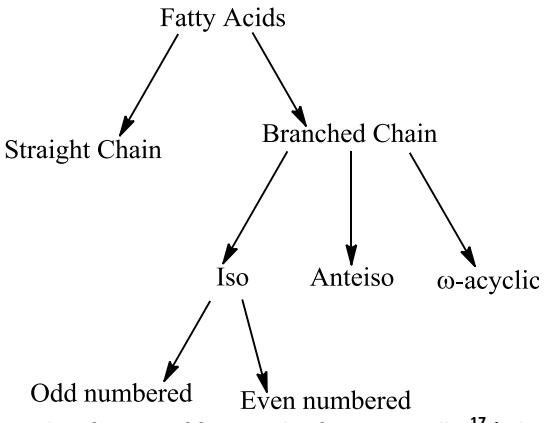


Figure 4: Classification of fatty acids of genus *Bacillus*[17] (adapted).

Fatty acids are significant because they play an important role in phospholipid synthesis and in the functions of the cell membrane. Consequently, fatty acids are important products of the breakdown of cell membrane components[17].

Regardless of the arrangement of their carbon skeleton, fatty acids are subjected to β-oxidation - a means of energy production. β-oxidation is a pathway of enzyme-catalyzed reactions that operate in repetitive fashion, progressively degrading fatty acids by removing two carbon units every cycle. It involves the oxidation of the β carbon.

I.D.1.a Schematic of β-oxidation

At the onset of β-oxidation, fatty acids will be activated by acyl-CoA synthetase/thiokinase, reacting with CoA with the aid of ATP to form a fatty acyl-CoA. This fatty acyl-CoA is then degraded to acetyl-CoA molecules via the β-oxidation in a cycle of four reaction steps (Figure 5).

Step-1: The fatty acyl-CoA is oxidized at the C_α/C_β position to form a trans Δ^2-enoyl-CoA through dehydrogenation by acyl-CoA dehydrogenase which uses FAD (Flavin Adenine Dinucleotide) as a cofactor.

Step-2: The trans Δ^2-enoyl-CoA formed in step-1 is hydrated by enoyl-CoA hydratase, forming a 3-hydroxyacyl-CoA.

Step-3: 3-hydroxyacyl-CoA undergoes oxidation in the presence of a NAD^+ dependent 3-L-hydroxyacyl-CoA dehydrogenase thereby forming the corresponding β-ketoacyl-CoA.

Step-4: The β-ketoacyl-CoA formed in the previous step undergoes C_α/C_β cleavage in presence of CoA catalyzed by β-ketoacyl-CoA thiolase. This leads to formation of acetyl-CoA and a new fatty acyl-CoA, two carbons shorter than the original which began the cycle.

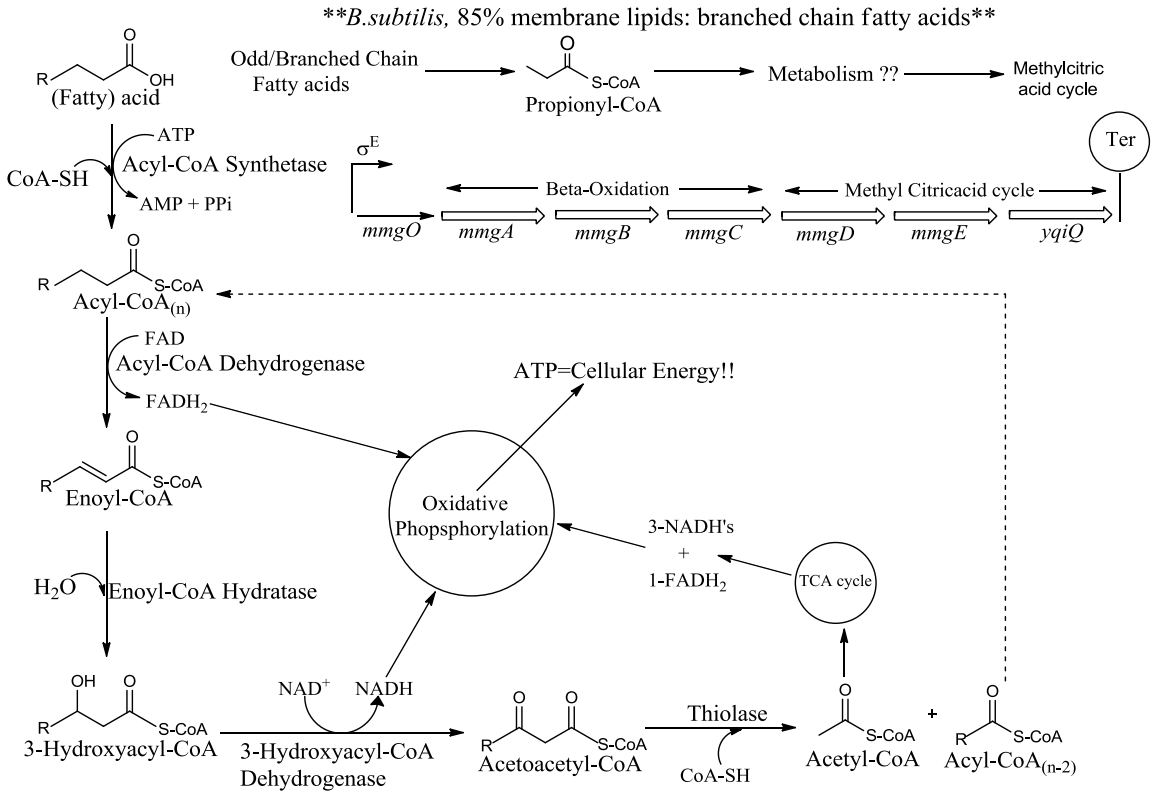


Figure 5: Schematic for β-oxidation of fatty acids and propionate metabolism in *B.subtilis* via methylcitric acid cycle.

Each round of β-oxidation produces one molecule each of $FADH_2$, NADH and acetyl-CoA. Acetyl-CoA formed will be channeled into citric acid cycle where it is oxidized further. This leads to formation of additional three NADH's and one $FADH_2$. Each of these NADH's and $FADH_2$'s will be reoxidized to NAD^+ and FAD respectively through oxidative phosphorylation leading to the formation of ATP. Thus the purpose of β-oxidation is to generate metabolic energy since the complete oxidation of each fatty acid yields many ATP's.

In *B.subtilis,* 85% of the membrane lipids consist of branched chain fatty acids[17]. If the fatty acid is a branched or odd chain fatty acid, the final round of oxidation will

yield propionyl-CoA also alongside acetyl-CoA. This propionyl-CoA is channeled into methyl citric acid cycle (Figure 5).

I.D.2. Propionate Metabolism/ Methylcitric acid Cycle

Odd chain fatty acids are a rich source of energy for many microorganisms that are soil inhabitants. *B.subtilis* is one such example. Branched chain fatty acids comprise 85% of the total fatty acid composition of cell membranes in these organisms[17]. β-oxidation of these fatty acids yields propionyl-CoA. Propionate can be metabolized *via* a wide range of pathways like the methylcitric acid cycle, α-oxidation, β-oxidation, α-carboxylation, reductive carboxylation and Claisen condensation to name a few (Figure 6)[15].

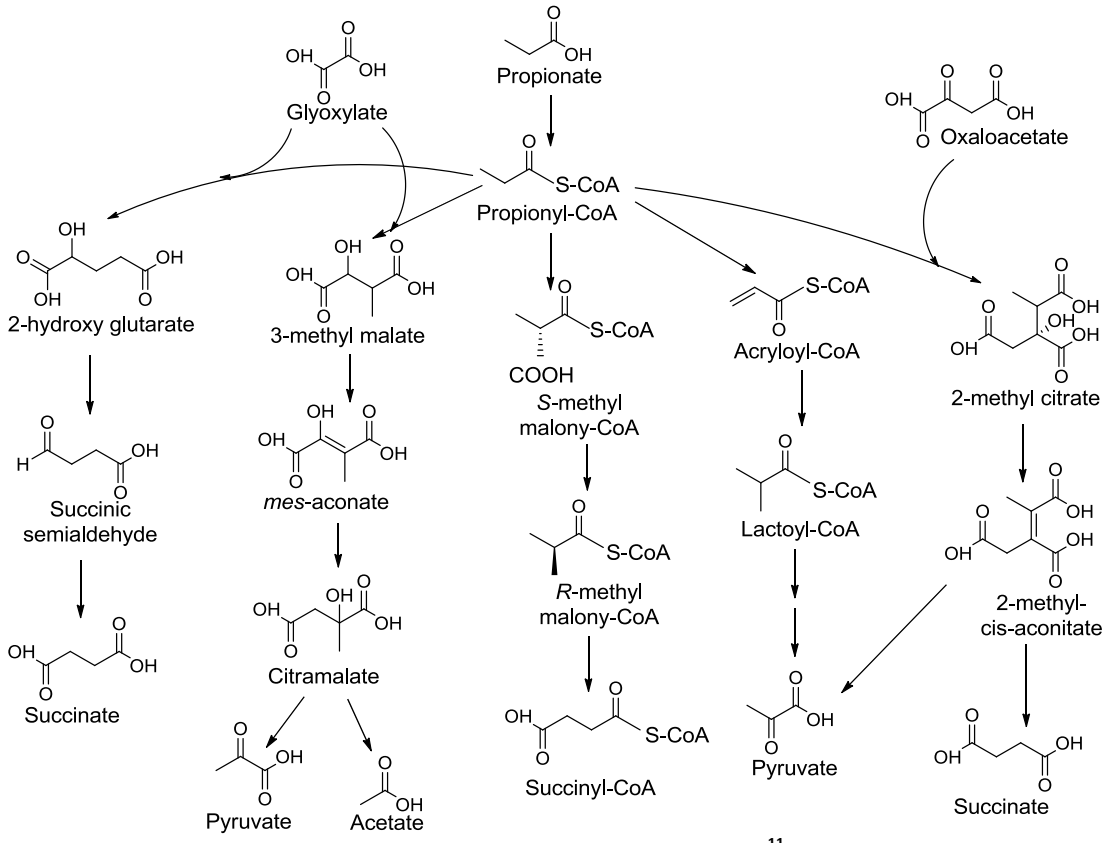


Figure 6: Various pathways for propionate metabolism[11] (adapted).

The *prp* operon from *Salmonella typhimurium* which consists of *prpBCDE* genes was shown to catabolize propionate *via* the methylcitric acid cycle. Horswill *et al*., has shown that *prpE* encodes propionyl-CoA synthetase. *PrpC* encodes 2-methylcitrate synthase. *PrpD* would encode either 2-methylcitrate dehydratase or 2-methylcitrate hydratase. Its function is still unclear. *PrpB* encodes 2-methylisocitrate lyase (Figure 7)[11]. 2-methylcitrate synthase catalyzes the formation of 2-methylcitrate through aldol condensation of oxaloacetate with propionyl-CoA. The 2-methylcitrate formed will be subjected to dehydration by 2-methylcitrate dehydratase thus forming 2-methyl-cis-aconitate which upon subsequent hydration in the presence of aconitase forms 2-

methylisocitrate. Since there is no aconitase analogue in the *prp* operon; it is still unclear if 2-methylcitrate dehydratase itself is also responsible for hydration of 2-methyl-cis-aconitate to 2-methylisocitrate or there is any other enzyme that does this job. 2-methylisocitrate lyase breaks down 2-methylisocitrate into pyruvate and succinate, the last step in the methylcitric acid cycle.

The pyruvate formed goes either into the citric acid cycle or will be directed towards gluconeogenesis. Succinate will be fed into the citric acid cycle where it is subsequently oxidized to oxaloacetate to initiate another round of either the citric acid cycle, the methylcitric acid cycle or both.

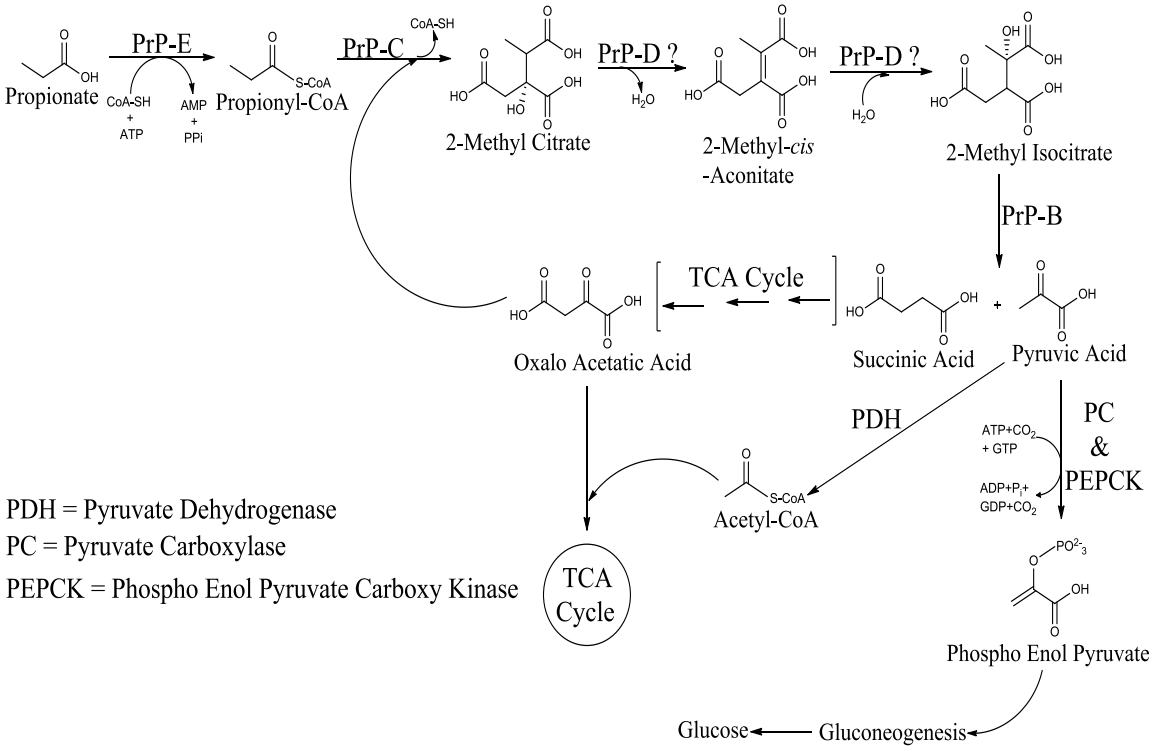


Figure 7: Schematic of propionate metabolism in *S.typhimurium*[11] (adapted).

More recently a similar operon was discovered in *E.coli* and named the *prp* operon. This operon too is associated with propionate metabolism *via* the methylcitric acid cycle (Figure 8)[12].

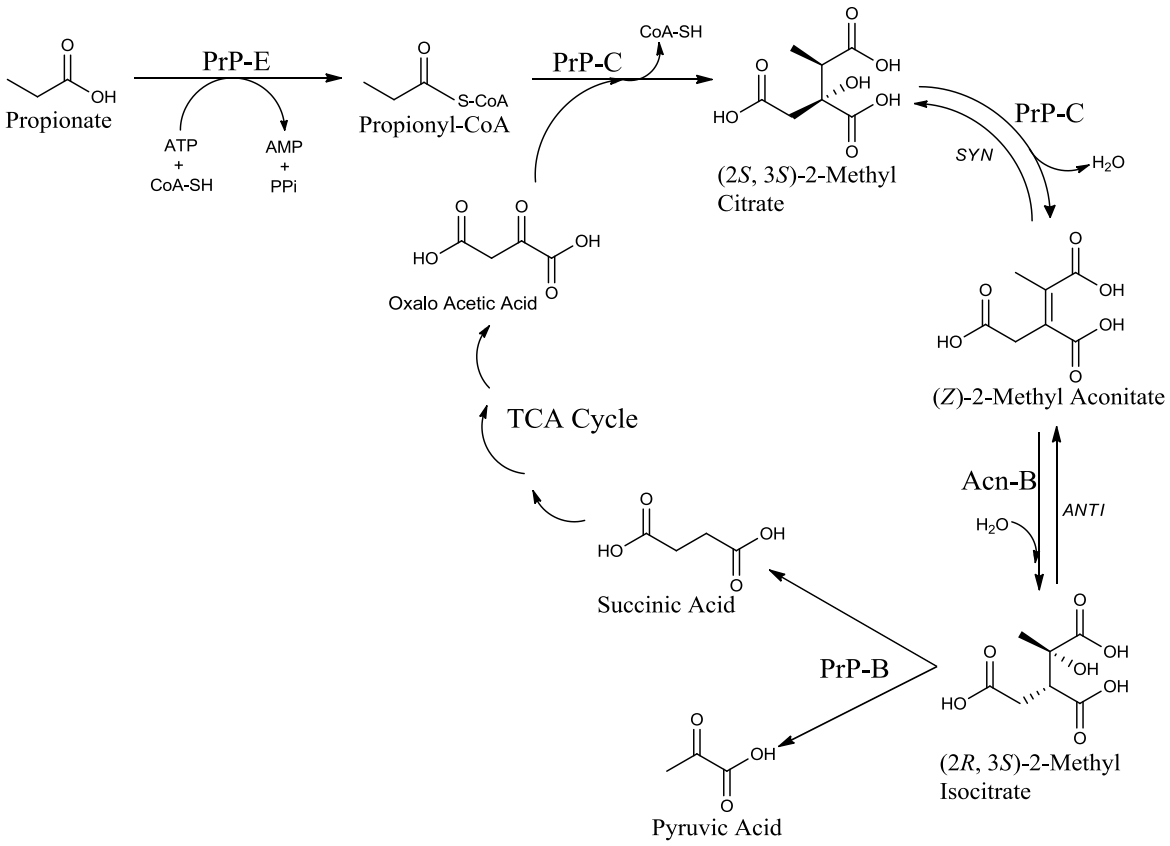


Figure 8: Schematic of propionate metabolism in *E.coli*[12] (adapted).

Sequence comparisons show that a portion of the *mmg* operon is homologous to *prp* operons from *S. typhimurium* and *E.coli*. *B.subtilis* counterparts of *prp* members are *mmgDE* and *yqiQ*: 2-methylcitrate synthase (MCS), 2-methylcitric acid dehydratase (MCD) and 2-methylisocitrate lyase (MCL) respectively. CitB, the only aconitase in

B.subtilis, is considered to play a role in the rehydration of 2-methyl-cis-aconitate to 2-methylisocitrate as does AcnB (aconitase) from *E.coli*[10-12, 18].

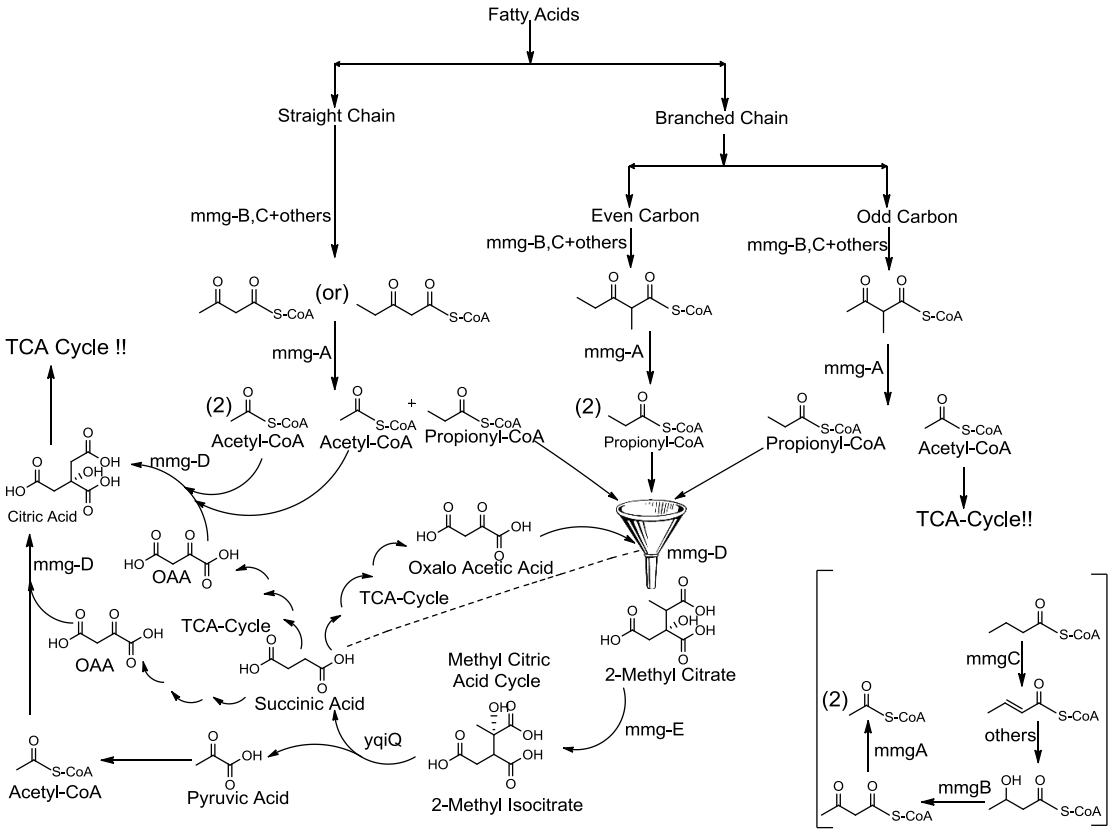


Figure 9: Schematic of proposed functions of *mmg* operon.

To reiterate, one region of the *mmg* (*mmgA, B* and *C*) operon is associated with β-oxidation, which yields acetyl-CoA/ propionyl CoA as end products. These may enter the citric acid cycle or methylcitric acid cycle respectively. Any propionyl CoA is oxidized to pyruvate via the methylcitric acid cycle encoded by remaining portion of the *mmg* (*mmgD, E* and *yqiQ*) operon (Figure 9). Since these processes operate during sporulation, a nutritional deprivation induced state, conservation of every single molecule of metabolite which can yield ATP is really important for the cell.

CHAPTER II

HYPOTHESIS

II.A. *MmgB*

Bacillus subtilis is a model organism for studying sporulation. Sporulation is controlled by several genes and operons in association with various DNA binding proteins and RNA polymerase sigma (σ) factors[3]. The *mmg* operon is one such operon under the control of σ^E factor. The *mmg* operon becomes metabolically active during intermediary stages of sporulation. *MmgB* is one of the members (a gene) of the *mmg* operon[10]. Based on sequence homology, it is proposed to encode for 3-hydroxy-butyryl-CoA dehydrogenase (E.C. 1.1.1.157). Thus, our hypothesis is that the mmgB protein is an oxido-reductase (Figure 10) catalyzing the oxidation of 3-hydroxy-butyryl-CoA (3-HB-CoA) to acetoacetyl-CoA (AcAc-CoA). It would use either NAD^+ or $NADP^+$ as the cofactor[10, 14, 19].

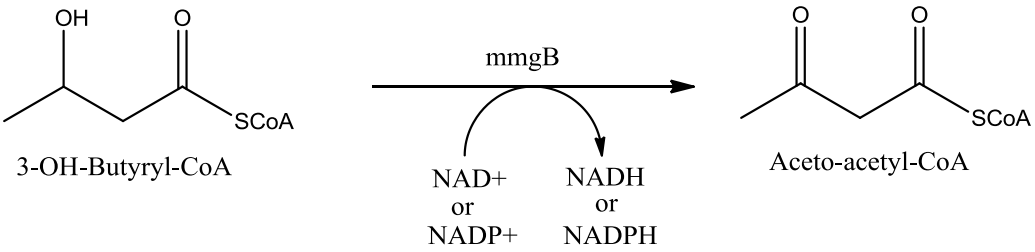


Figure 10: Proposed enzymatic reaction catalyzed by mmgB protein.

II.B. Goals

Biochemical Characterization of *mmgB*: This was initiated by Spencer Russell (2007), but he couldn't establish its enzymatic action. Recently we discovered that the nucleotide sequence of *mmgB* was updated in the genome servers. The new version (864 bp's) of the sequence is 132 base pairs longer than the older version (732 bp's) with which Spencer Russell worked (Figure 11)[14].

```
old                1 ----------------------------------------------------      0
new                1 atggaaatcaaacaaatcatggtagctggcgcaggtcagatggggagcgg     50
old                1 ----------------------------------------------------      0
new               51 aattgctcaaacagccgccgacgcgggcttttatgtgcggatgtatgatg    100
old                1 --------------------------ttgaaacggctgaagaaacagctg     24
                                                |||||||||||||||||||||||
new              101 tgaatccagaggccgcggaggcaggattgaaacggctgaagaaacagctg    150
old               25 gcccgtgatgctgagaaaggaaaaaggaccgagacggaagtgaagagcgt     74
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
new              151 gcccgtgatgctgagaaaggaaaaaggaccgagacggaagtgaagagcgt    200
old               75 aatcaaccgcatttcgatttctcaaacacttgaggaggcagagcatgcgg    124
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
new              201 aatcaaccgcatttcgatttctcaaacacttgaggaggcagagcatgcgg    250
old              125 acattgtgattgaggctatcgcagaaaacatggcggcaaaaactgagatg    174
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
new              251 acattgtgattgaggctatcgcagaaaacatggcggcaaaaactgagatg    300
old              175 tttaaaacacttgatcgcatttgcccgcctcatacgattttggccagcaa    224
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
new              301 tttaaaacacttgatcgcatttgcccgcctcatacgattttggccagcaa    350
old              225 tacatcttccttgcctattacagaaatcgctgctgtaacaaaccggcctc    274
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
new              351 tacatcttccttgcctattacagaaatcgctgctgtaacaaaccggcctc    400
old              275 aacgggttattggcatgcattttatgaatcccgtccctgtaatgaagctg    324
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
new              401 aacgggttattggcatgcattttatgaatcccgtccctgtaatgaagctg    450
```

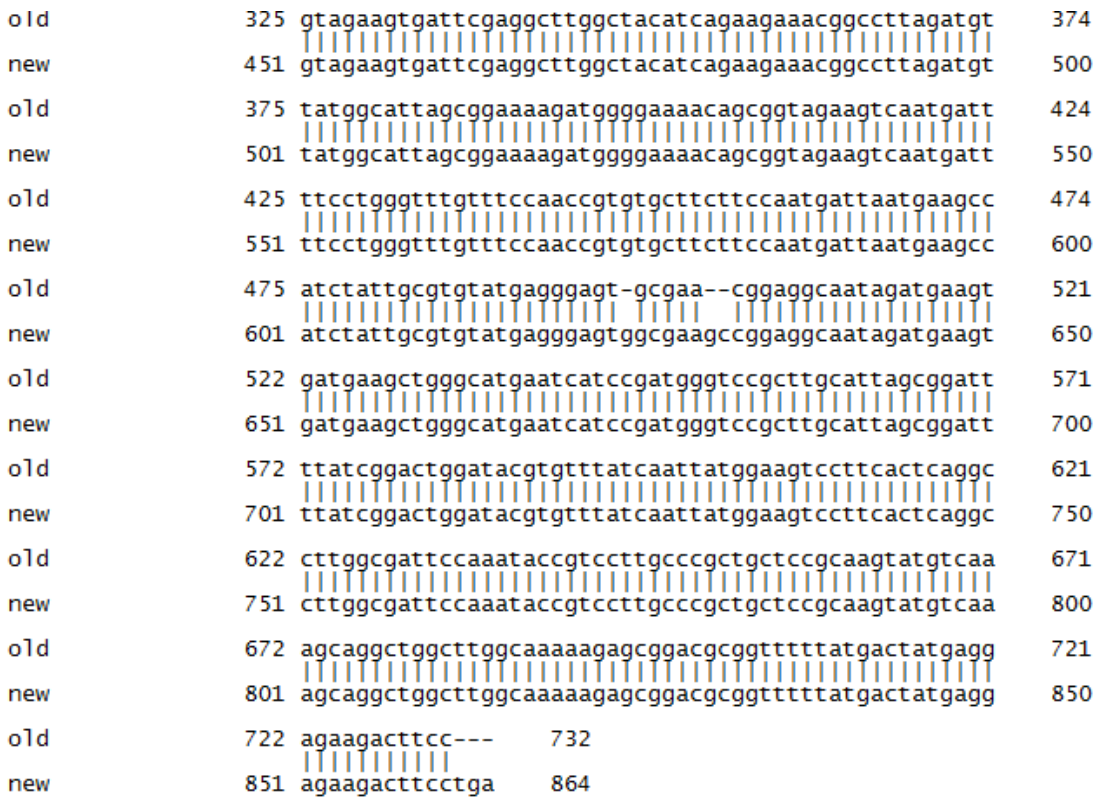


Figure 11: Alignments of old and new *mmgB* gene sequences.

So, the first goal was to clone the new, corrected *mmgB* gene into a suitable expression system (pET-200 plasmid). The second goal was to isolate and purify the mmgB protein. The final goal was to determine the enzymatic activity of this protein.

CHAPTER III

EXPERIMENTAL SECTION

III.A. Cloning of the *mmgB* gene

III.A.1. PCR amplification of *mmgB* gene

Both forward and reverse primers for the *mmgB* gene were designed in order to amplify the gene. Primers were designed following manufacturer instructions for TOPO cloning, into the pET-200 vector from Invitrogen and were obtained from Integrated DNA technologies (Table 1)[20, 21]. Genomic DNA, serve as template for amplification of target gene was isolated from wild type *B.subtilis* 168 (BS 168) cells using the Wizard® Genomic DNA Purification Kit (from Promega) following manufacturer instructions (Cat # A1120).

Table 1: Details about specifications of the *mmgB* gene primers

Parameters	Forward primer	Reverse primer
Sequence	5' - ***cacc*** **atg** gaa atc aaa caa atc atg **g** - 3'	5' - **tca** gga agt ctt ctc ctc ata gtc - 3'
Tm	55.4°C	55.2°C

The *mmgB* gene was amplified by a PCR (Polymerase Chain Reaction). The components of the PCR included: 25 µL PMM (Phusion master mix), 22 µL Molecular Biology Grade Water, 2 µL Genomic (template) DNA, 1.25 µL Forward primer, 1.25 µL Reverse primer. PMM is a stock mixture of High Fidelity Phusion Polymerase + Buffer +

dNTP's + counter ions ($MgCl_2$) stored at -20^{o}C (from Finzymes/New England Biolabs, Cat # F553S).

PCR amplification of the desired *mmgB* gene is carried out using a thermocycler, according to the following parameters: Initial incubation for 1 minute at 98^{o}C followed by 25 cycles of each of the; Denaturation for 5 seconds at 98^{o}C, Annealing for 1 minute at 55^{o}C, Extension for 1 minute at 72^{o}C. After 25 cycles, an extension cycle for 10 minutes at 72^{o}C is carried out and finally the samples are hold at 4^{o}C.

The PCR product was purified by a QIAquick PCR- purification kit from Qiagen (Cat # 28106) following the manufacturer's specifications[22]. The purified PCR product was then analyzed by 1% agarose gel electrophoresis and compared to the 2-log ladder base pair standard from New England Biolabs (Cat # N3200S).

III.A.2. Cloning of the amplified *mmgB* gene into pET-200 vector

After the amplification of *mmgB* was confirmed by agarose gel electrophoresis, the PCR amplified *mmgB* gene was cloned into an expression system. The expression system used in our study was the pET200/D-TOPO® vector from Invitrogen (Cat # K100-01). The procedure of cloning the amplified *mmgB* was done following the manufacturer instructions[21].

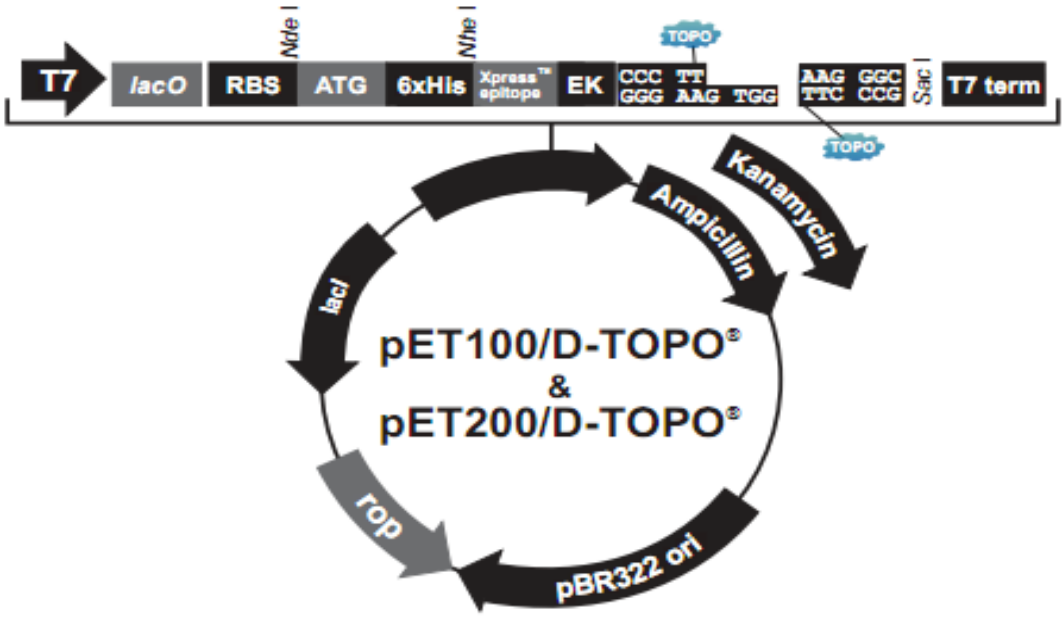


Figure 12: pET 200/D – vector map[21] (adapted).

The plasmid pET200/D-TOPO consists of a kanamycin resistant gene instead of ampicillin resistant gene. The target gene(s) was placed under the control of an IPTG inducible T7 promoter. A nucleotide sequence corresponding to a 6x His-tag (6 consecutive histidines) is present in between the T7 promoter and the target gene. This set up leads to the production of a recombinant fusion protein with an N-terminal His-tag (Figure 12)[21].

III.A.3. Transformation and Retransformation of *E.coli* with cloned pET-200 vector

Transformation of *E.coli* with the cloned pET-200 vector (pET-200_*mmgB*) was carried out following the manufacturer's instructions[21]. Following transformation, the cells were spread on LB (Luria-Bertani) – agar plates with 50µg/mL kanamycin and the transformants were used to inoculate 5 mL LB with 50 µg/mL kanamycin and incubated anaerobically overnight at 37°C and 220 rpm.

Glycerol (10% v/v) stocks of these cultures were made and stored at -80°C for further proceedings. Later, overnight liquid cultures of the transformed *E.coli* cells were made by inoculating the antibiotic resistant colonies/cells in 5 mL LB-broth with 50µg/mL Kanamycin and are grown anaerobically at 37°C /220 rpm. The following morning, the cloned pET-200 vector (pET-200 vector_*mmgB*) was isolated from these liquid cultures by employing the QIAprep Miniprep kit (Cat # 27106) from Qiagen, following the manufacturer instructions[23].

A PCR was performed using cloned pET-200 vector (pET-200 vector_*mmgB*) as template DNA and a 1% agarose gel was run to analyze for the presence of insert. The isolated, cloned pET-200 vector was stored at -20°C for future uses.

Our next target in line was to transform BL21 (DE3) (a strain of commercially available competent *E.coli*) cells. Competent BL21 cells were obtained from Invitrogen and the transformation was achieved following the manufacturer's instructions. The antibiotic selective transformants were picked up and glycerol stocks are made similar to how it was done with TOP10 cells[21].

Again, cloned pET-200 vector (pET-200 vector_*mmgB*) was isolated from the BL21 cells. The isolation procedure involves making overnight liquid cultures of the transformed BL21 *E.coli* cells by inoculating the antibiotic resistant colonies/cells in 5 mL LB-broth with 50µg/mL Kanamycin and growing anaerobically at 37°C /220 rpm. The following morning, cloned pET-200 vector (pET-200_*mmgB*) was isolated from these liquid cultures by employing the QIAprep Miniprep kit from Qiagen, following

manufacturer instructions[23]. The miniprep solution was sent to SeqWright for dye-terminated DNA sequencing.

III.B. Overexpression and Isolation of mmgB protein

For the overexpression of *mmgB* in BL21 (DE3), 1L cultures were grown by inoculating 1L LB-broth (50μg/mL Kanamycin) with 2 mL of an overnight grown starter culture of transformed BL21 cells. The growth was monitored spectrophotometrically at 595nm. When the absorbance (OD_{595}) reached a value between 0.7 – 0.8, IPTG (Isopropyl-β-thiogalactopyranoside) was added to the culture to induce overexpression of *mmgB*. IPTG was added to a final concentration of 1mM. Following IPTG addition, the culture was left to grow overnight at 37°C / 220 rpm allowing for the overexpression of the *mmgB*.

The following morning, cells from the overnight culture were harvested by centrifugation (7500 g, 30 minutes, 4°C). The cells were then resuspended in 40 mL 1x binding buffer (0.5 M NaCl, 20 mM Tris-HCl, 5 mM Imidazole, pH 7.9). This cell suspension was subjected to sonication in order to break open the cells and release cellular components. The lysate was then subjected to centrifugation (14,600 g, 30 minutes, 4°C) to remove the cell debris.

Following centrifugation, the supernatant was further clarified by syringe filtration using a 0.45 μm Corning Incorporated Syringe Filter. The crude extract was applied to a Ni – NTA column chromatography. This was performed in order to isolate the His-tagged mmgB.

III.B.1. Ni-NTA Column Chromatography

The Ni-NTA (nitrolotriacetic acid) column (Cat # 30210) consists of an agarose-nitrolotriacetic acid resin. The resin is charged with hexadentate Ni^{2+} ions, which are held onto the resin by tetra dentate nitrolotriacetic acid molecules (Figure 13)[24].

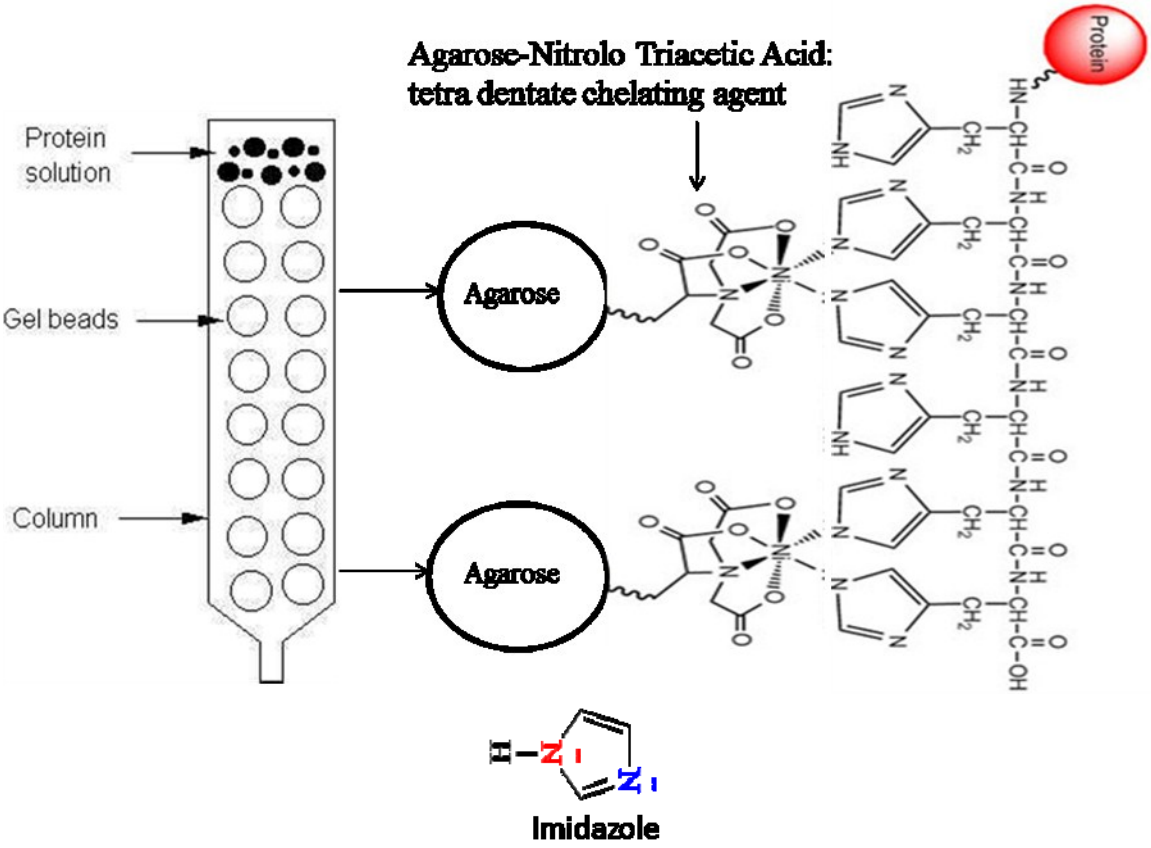


Figure 13: Ni2+-NTA chromatography set-up; his-tag protein interactions with agarose/Ni2+-NTA resin & imidazole structure.

The crude extract (filtrate) was applied onto the column (column bed volume was 2 mL) and was allowed to pass through the resin under the influence of gravity. Once all the crude extract was run through the column, it was then treated with 10 column volumes of 1x binding buffer (0.5 M NaCl, 20 mM Tris-HCl, 5 mM Imidazole, pH 7.9).

This was followed by a washing step with 6 column volumes of 1x wash buffer (0.5 M NaCl, 20 mM Tris-HCl, 60 mM Imidazole, pH 7.9). Later the column was treated with 6 column volumes of 1x elute buffer (0.5 M NaCl, 200 mM Imidazole, 20 mM Tris-HCl, pH 7.9) to elute the His-tagged protein off the column[24].

After isolating the protein to near homogeneity using Ni-NTA column chromatography, it was subjected to dialysis. The MWCO (Molecular Weight Cut Off) value of the dialysis bag used in our experiments was 7000 Da. For dialysis of the protein, we used 4 Liters Tris-HCl, pH 8.0 and the set-up was left undisturbed overnight at 4^{o}C. Following dialysis, glycerol (10% v/v) stocks of the dialyzed protein were made and the aliquots (100µL) were stored at -80^{o}C.

III.C. Characterization of mmgB protein

III.C.1. SDS-PAGE

A 12% SDS-PAGE (Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis) was run to confirm the identity of the mmgB protein obtained after chromatographic separation. All the eluted fractions, including those of binding & washing steps were analyzed. The gel was visualized by Coomassie Blue stain and the destaining was achieved, using methanol (30% v/v) and acetic acid (10% v/v).

III.C.2. Gel-Based Trypsin Digest; MALDI-TOF Mass-Spectrometry

To reinforce the identity of the protein isolated, MALDI-TOF-MS (Matrix Assisted Laser Desorption Ionization – Time of Flight – Mass spectrometry) analyses of the gel based trypsin digests of the mmgB protein bands from SDS-PAGE gel were performed. Gel based trypsin digestion of the SDS-PAGE bands was achieved by

employing the Trypsin Profile IGD kit (Cat # PP0100-1KT) for in-gel digests from Sigma-Aldrich. This was performed following the manufacturer instructions.

III.D. Enzymatic Activity Analyses of mmgB protein

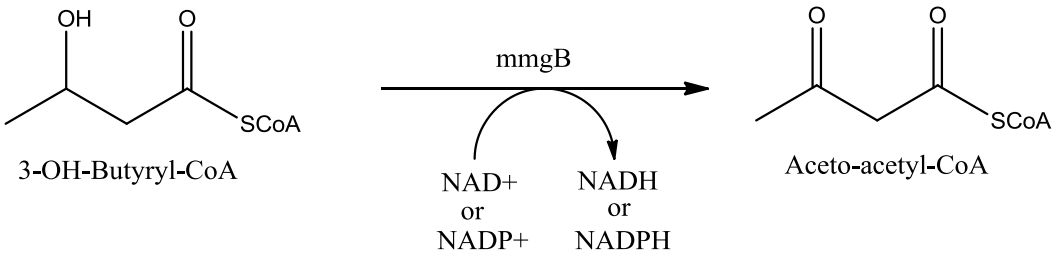


Figure 14: Proposed enzymatic reaction catalyzed by mmgB protein.

III.D.1. Determination of appropriate co-factor

Sets of assays were performed by analyzing the activity with one cofactor (NAD^+ or $NADP^+$) at a time. We set out to begin with NAD^+, since it is the most commonly used one and also is commercially cheaper than $NADP^+$. An initial set of reactions was done where the [substrate] and [NAD^+] or [$NADP^+$] were maintained constant but with increasing [enzyme]. The buffer employed in the assays was 100 mM potassium-phosphate buffer, pH 7.3 and the reaction was spectrophotometrically monitored at 340 nm. The conditions employed are listed in the table below.

Table 2: Details of reactions performed with NAD+/$NADP^+$ as co-factor. D, L-3-HB-CoA = D, L-3-hydroxybutyryl-CoA

D,L-3-HB-CoA (mM)	NAD^+ (mM)	$NADP^+$ (mM)	Enzyme (ng)	Reaction Volume(µL)	Time (min)
0.2	0.2	0.2	8	500	30
0.2	0.2	0.2	16	500	30
0.2	0.2		24	500	30
0.2	0.2		32	500	30
0.2	0.2		80	500	30

III.D.2. Determination of optimal pH

We have done many assays by employing different buffers (both composition and concentration) and thereby testing the effectiveness of the enzyme at different pH values. The different buffers and the pH ranges tested are listed in the table below. The reaction mixture contained $NADP^+$ as co-factor and was spectrophotometrically monitored at 340 nm.

Table 3: Details about buffers/pH range tested

Buffer used	pH tested
50 mM Na-Phosphate	7.3
100 mM Na-Phosphate	7.3
50 mM K-Phosphate	7.3
100 mM K-Phosphate	7.3
50 mM Tris	7.3
100 mM Tris	7.2 7.6
100 mM Tris	8.0 8.5 9.0 9.5
100 mM Glycine-NaOH	10.6 9.8

III.D.3. Determination of enzymatic activity of mmgB protein

We started designing experiments which employed $NADP^+$ - suitable co-factor, Glycine-NaOH buffer pH 9.8 - optimal buffer /pH with varying [substrate], [co-factor] and [enzyme] in order to test the enzymatic activity of mmgB. The progress of the reactions was spectrophotometrically monitored at 340 nm (λ_{max} of NADPH). Later, following spectrophotometric analysis, the reaction mixtures were analyzed by MALDI-TOF-MS (Matrix Assisted Laser Desorption Ionization – Time of Flight - Mass Spectrometry). The table below shows details about the assay conditions employed.

Table 4: Details about enzymatic reaction assay conditions

D,L-3-HB-CoA (mM)	$NADP^+$ (mM)	Enzyme (ng)	Buffer (µL)	Reaction Volume(µL)	Time (min)
0.8	0.5	8	431	500	30
0.8	0.5	16	421	500	30
0.8	0.5	24	411	500	10
0.4	0.25	24	440.5	500	30
0.4	0.25	16	450.5	500	55
0.4	0.25	24	440.5	500	30
0.2	0.2	8	467.2	500	20
0.2	0.2	16	457.2	500	20
0.2	0.2	24	447.2	500	20
0.4	0.2	24	434.4	500	15
0.5	0.2	24	427.9	500	30
0.4	0.2	32	424.4	500	30
0.5	0.2	32	417.9	500	30

Mass spectrometric analyses were performed to characterize the products of the reaction. The MALDI-TOF-MS instrument used was a 4700 Proteomic Analyzer from Applied Biosystems (Triad Mass-Spectrometry Laboratory, Department of Chemistry and Biochemistry, The University of North Carolina, Greensboro). The samples were analyzed with the reflectron in positive ion mode at a fixed laser intensity of 3001. The reaction samples were mixed with matrix (10 mg α-cyano-4-hydroxycinnamic acid butylamine salt dissolved in 1 mL matrix-A diluent: 50% acetonitrile + 50% nano pure H_2O + 0.3% trifluoroacetic acid) and matrix-A diluent. Unless otherwise mentioned, all the samples were prepared by mixing together matrix: diluent: sample in 40:40:20. =

III.D.4. Determination of Enzyme kinetics

Enzymes typically follow Michaelis-Menten kinetics. The kinetic rate constants $K_{m\ app}$ (apparent) of both the substrate (3HBCoA) and co-factor ($NADP^+$) were determined individually. The $V_{max\ app}$ (apparent) of the reactions was also determined. All the reactions for kinetic studies were performed in triplicates.

A set of reactions were performed at constant [$NADP^+$] and [enzyme] of 0.2 mM and 2.0 µM respectively while varying [3-HB-CoA]. The range of [3-HB-CoA] used was 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 2.0, 3.0 mM. The reactions were performed in Glycine-NaOH buffer, pH 9.8 and the reaction volume was 500 µL. The reactions were spectrophotometrically monitored at 340 nm (λ_{max} of NADPH) and initial velocities of the reactions were recorded.

The initial velocities versus [3-HB-CoA] were plotted with the aid of Slide Write software. These plots were then used to calculate $K_{m\ app}$ of the substrate and $V_{max\ app}$ of the reaction.

Another set of reactions were performed at constant [3-HB-CoA] and [enzyme] of 0.4 mM and 24 ng respectively while varying [$NADP^+$]. The range of [$NADP^+$] used was 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 2.0, 3.0 mM. The reactions were performed in Glycine-NaOH buffer, pH 9.8 and the reaction volume was 500 µL. The reactions were spectrophotometrically monitored at 340 nm (λ_{max} of NADPH) and initial velocities of the reactions were recorded.

Later, graphs were plotted for initial velocities versus [$NADP^+$] with the aid of 'slide write' software. These plots were then used to calculate $K_{m\ app}$ of the co-factor ($NADP^+$) and $V_{max\ app}$ of the reaction.

CHAPTER IV

RESULTS / DISCUSSION

IV.A. Cloning of the *mmgB* gene

IV.A.1. PCR-amplification of *mmgB* gene

The *mmgB* gene from *Bacillus subtilis* 168 was amplified by PCR (Polymerase Chain Reaction). In order to test for the success of PCR, the PCR amplified product was first purified by a PCR purification step following the completion of PCR. The purified, amplified *mmgB* gene was then analyzed by 1% agarose gel electrophoresis, which is shown in Figure 15.

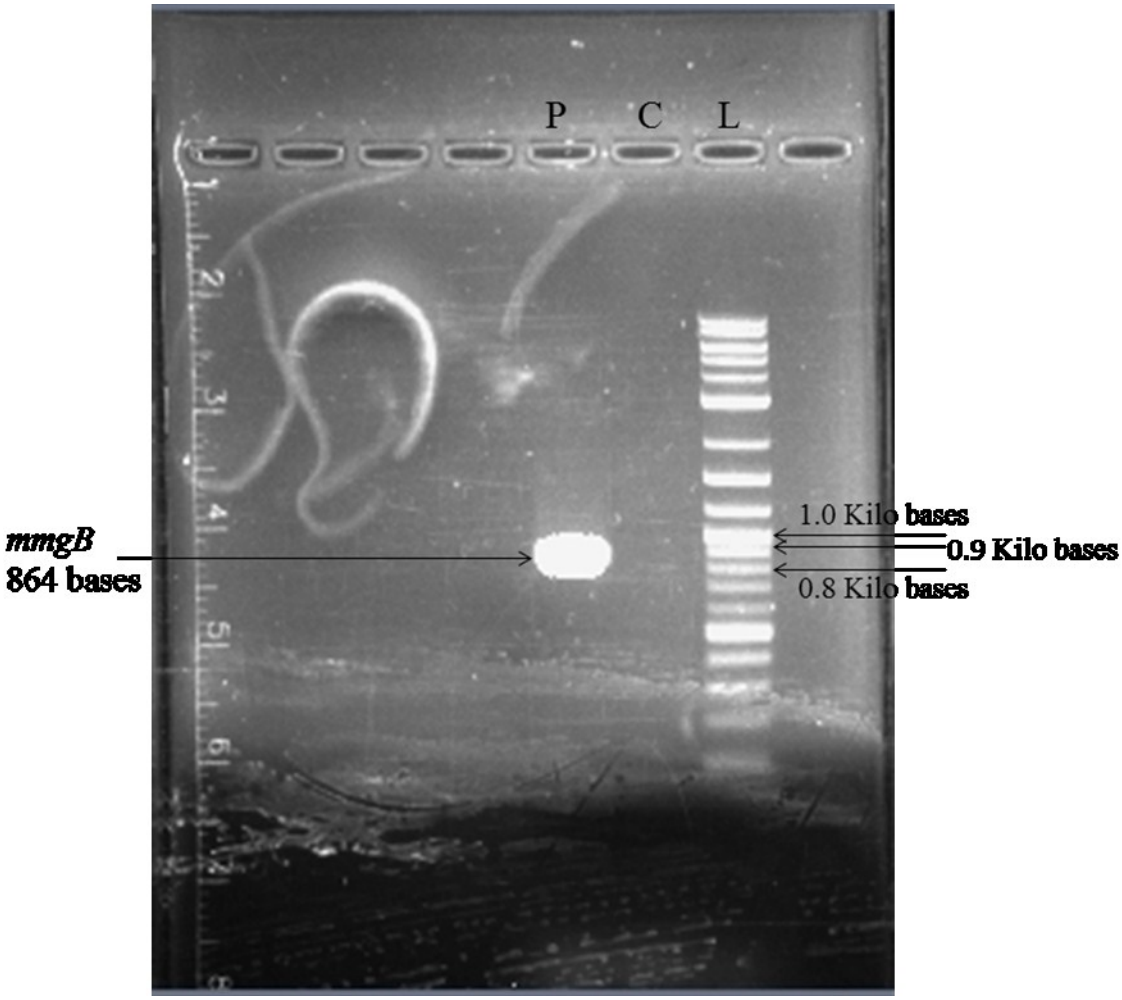


Figure 15: Agarose gel showing molecular weights standard (L), No DNA control (C), PCR product (P).

The agarose gel picture (Figure 15) shows three lanes labeled as L, C, and P. The lane L corresponds to 2-log ladder which serves the purpose of a molecular marker. The bands seen in the lane L are considered as standards in order to characterize the bands of unknown sizes (number of base pairs) appearing in the gel.

The lane C corresponds to a control reaction with respect to the amplification of *mmgB* gene. The control reaction includes all the PCR components (PMM, primers and H_2O) except for template DNA. So, the *mmgB* gene should not amplify as there is no

template to direct its amplification. Hence, there should not be any band seen in the lane C. This is in good agreement, as there is no band visible on the gel in the lane corresponding to C. This can be used as a negative reference for the success of PCR-amplification of *mmgB* gene.

The lane P corresponds to PCR amplified product which in this case is the amplified *mmgB* gene. This reaction involves all the PCR components including the template DNA (genomic DNA from *Bacillus subtilis* 168) which directs the amplification of *mmgB* gene. Hence there should be a band seen on the gel in lane P corresponding to the size of *mmgB* gene in terms of number of base pairs (bp's).

The gel indeed has a band in the lane P which lies at a position in between 0.8 – 0.9 kilo bases (kb) in comparison to molecular marker seen in the lane L. This is in good agreement with the expected size of *mmgB* gene which is 864 bp's. So, the result is suggestive that *mmgB* gene was successfully amplified by a PCR.

IV.A.2. Cloning of amplified *mmgB* gene into pET-200 vector

After confirming the successful PCR amplification of the *mmgB* gene, it was then cloned into the pET-200 vector according to the manufacturer instructions. The map for pET-200 vector (Figure 16) below shows the cloning site of *mmgB* insert into the vector. The pET-200 vector follows topoisomerase-dependent ligation and hence did not require the usual restriction digest and DNA ligase procedure.

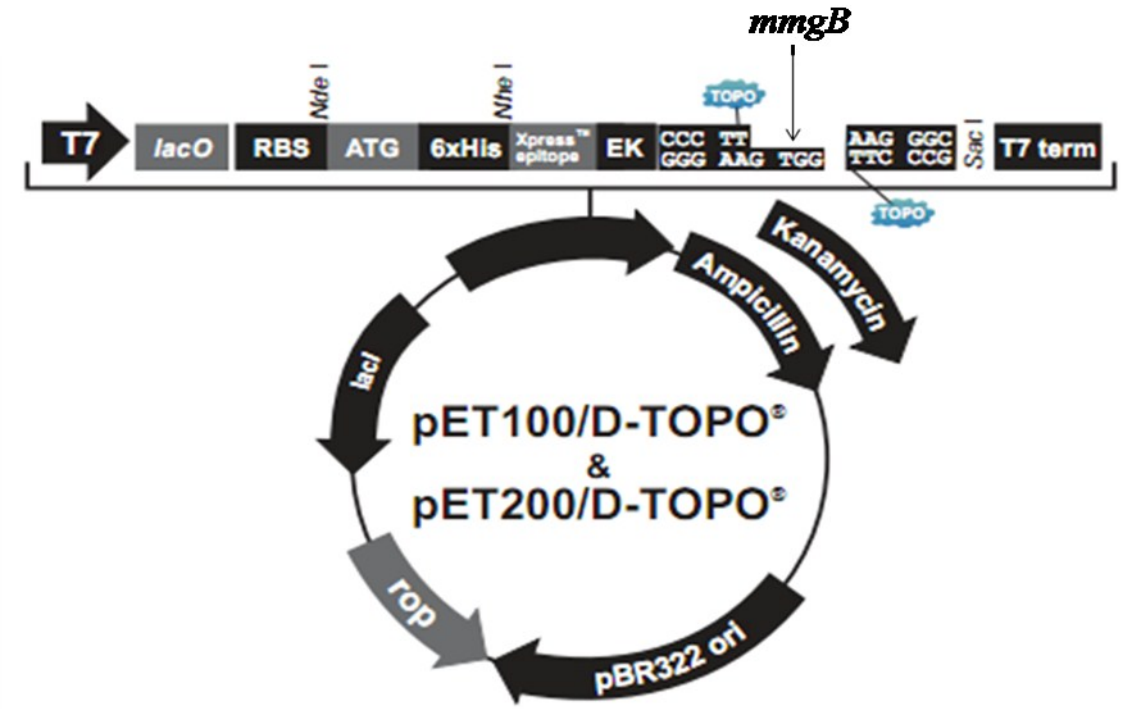


Figure 16: Cloning of *mmgB* insert into pET-200 vector[21] (adapted).

IV.A.3. Transformation and Retransformation of *E.coli* with cloned pET-200 vector

Following the completion of cloning, One Shot® TOP10 Chemically Competent *E.coli* cells were transformed with the TOPO reaction mixture (pET-200 vector_*mmgB*). Later, the transformed TOP10 *E.coli* cells were spread on LB – agar plates with 50µg/mL Kanamycin for antibiotic resistance selection. Later, the plasmid was screened for *mmgB* by PCR.

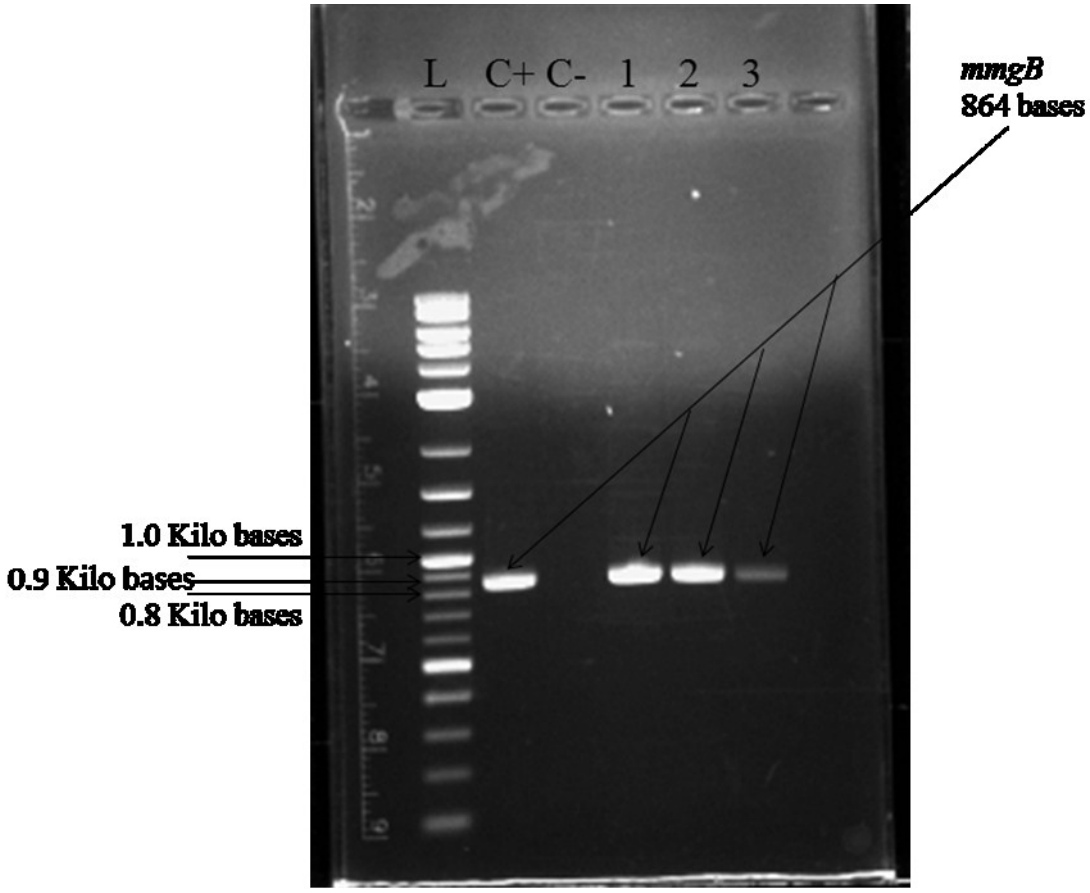


Figure 17: Agarose gel showing molecular weights standard (L), PCR product amplified using BS168 DNA (C+), PCR product amplified without employing BS168 DNA (C-), 1/2/3=PCR product amplified using pET-200 vector_*mmgB* DNA from TOP 10 cells.

The agarose gel picture (Figure 17) contains 6 lanes labeled as L, C+, C-, 1, 2, and 3. The lane L corresponds to the 2-log ladder; it serves the purpose of a molecular marker. The bands seen in the lane L are considered as standards for characterizing the bands of unknown sizes (number of base pairs) present on the gel.

The lane C+ corresponds to a positive control reaction with respective to amplification of the *mmgB* gene. The C+ reaction include all the PCR components (PMM, primers, template DNA and H_2O). The template DNA is genomic DNA from *Bacillus subtilis* 168. This directs the amplification of *mmgB* gene. Hence there should be a band

seen on the gel in lane C+ corresponding to the size of *mmgB* gene with reference to number of base pairs (bp's). This is in good agreement with the expected size of *mmgB* gene; 864 bp's. This was used as a positive reference for amplification of *mmgB* gene.

The lane C- corresponds to a negative control reaction with respective to amplification of the *mmgB* gene. The reaction components of the C- reaction include all the PCR components (PMM, primers and H_2O) except for template DNA. So, the PCR should not amplify anything as there is no available template. Hence there should not be any band seen in the lane C-. This is in good agreement as there is no band visible on the gel in the lane corresponding to C-. This was used as a negative reference for amplification of the *mmgB* gene, and showed that PCR for *mmgB* will be successful only when the *mmgB* is available in the reaction.

The lanes 1, 2 and 3 correspond to PCR reaction with three different plasmids obtained from the TOPO cloning transformants. The components of this reaction include all the PCR components including the template DNA (**pET-200 vector_*mmgB***) instead of genomic DNA, which directs the amplification of *mmgB* gene. If the pET-200 vector possessed the *mmgB* insert, there should be a band seen on the gel in lanes 1, 2 & 3 corresponding to the size of *mmgB* gene with reference to number of bp's.

Looking at the gel, there are bands seen in the lanes 1, 2 and 3 lying at a position in between 0.8 – 0.9 kb in comparison to molecular marker in the lane L. These bands from lanes 1, 2 & 3 align pretty well with the band from the lane C+ in terms of the position. The positions of the bands from lanes C+, 1, 2 and 3 are in good agreement with the expected size of *mmgB* gene (864 bp's). So, the result is suggestive that *mmgB* gene

was successfully cloned into the pET-200 vector and also the cloned pET-200 vector was stable in the TOP10 *E.coli* cells.

Later, BL21 (DE3) cells were transformed using the cloned pET-200 vector. BL21 (DE3) cells were specifically meant for heterologous expression of the target gene, which in our case is *mmgB*. The transformed BL21 cells were spread on LB – agar plates with 50µg/mL Kanamycin to grow. Later, the cloned pET-200 vector was isolated from antibiotic selective transformants. Following this, a PCR was performed using the cloned pET-200 vector (pET-200 vector_ *mmgB*) as template DNA.

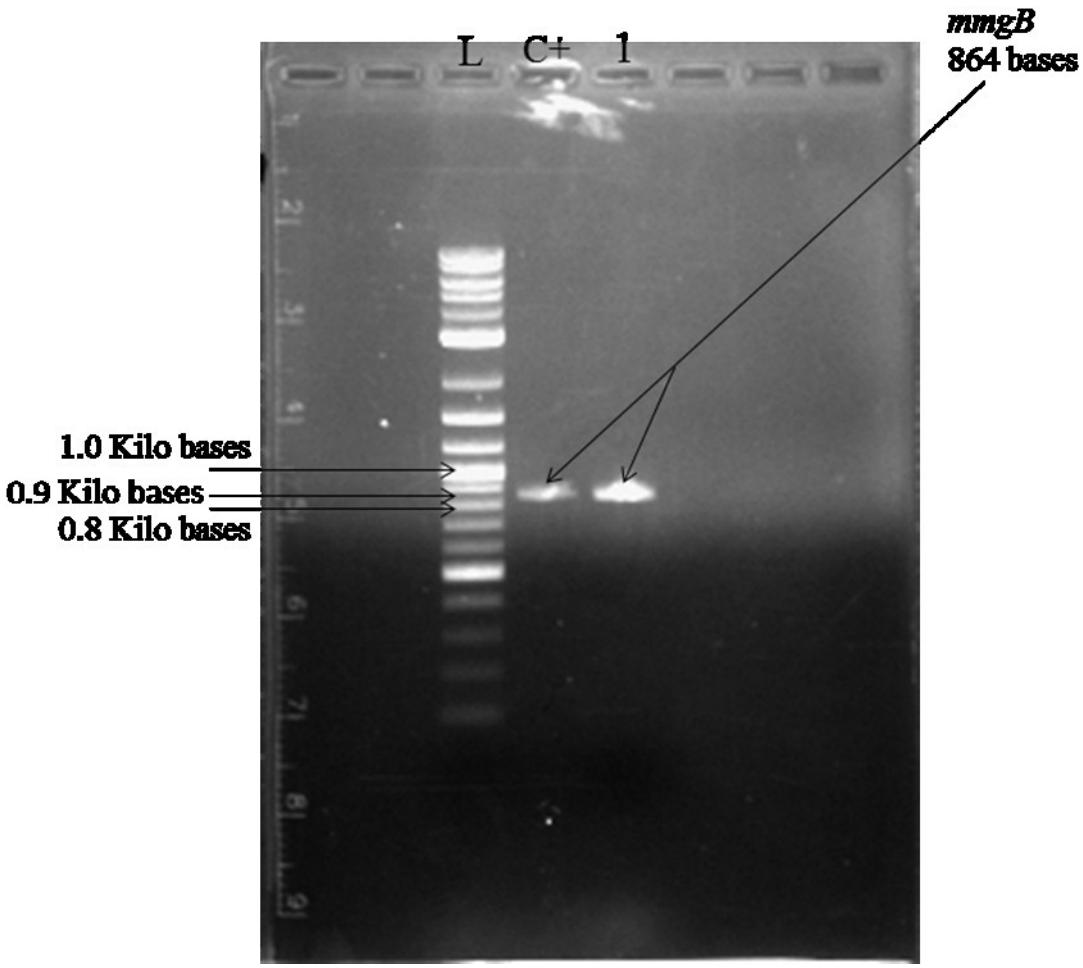


Figure 18: Agarose gel showing molecular weights standard (L), PCR amplified product usingBS168 DNA (C+), PCR amplified product by employing pET-200 vector_*mmgB* from BL21 (1).

In the gel picture above (Figure 18), C+ lane corresponds to a positive control reaction with regard to amplification of the *mmgB* gene. It includes all the PCR components wherein genomic DNA from *Bacillus subtilis* 168 (BS168) was used as the template DNA. The lane 1 corresponds to PCR amplified *mmgB.* It included all the PCR components wherein the cloned pET-200 vector (pET-200 vector_*mmgB*) was used as the template DNA. Looking at the bands in the lanes C+ and 1, they lie at a position in between 0.8 – 0.9 kb in comparison to molecular marker in the lane L. The positions of

these bands are in good agreement with the expected size of *mmgB* gene (864 bp's). Also, these bands align well with bands from previous gels in terms of the number of base pairs. So, the result is suggestive that the cloned pET-200 vector was stable in BL21 cells.

The plasmid was sequenced and below shown (Figure 19) is the alignment of the *mmgB* insert sequenced from the cloned pET-200 vector and the *mmgB* gene sequence from the updated genome database. During sequencing, as the reaction reaches the end, the sensitivity, i.e. the ability to get the sequence information, decreases. Hence, we see some mismatches or dots or sometimes the nucleotides themselves found missing towards the end of the sequencing reaction with both the sequencing primers. However, it should not be a major concern, since the other strand which primes in opposite direction gets the missing sequence information.


mmgB 1 atggaaatcaaacaaatcatggtagctggcgcaggtcagatggggagcgg 50
T7-up 1 ATGGAAATCAAACAAATCATGGTAGCTGGCGCAGGTCAGATGGGGAGCGG 50
mmgB 51 aattgctcaaacagccgccgacgcgggcttttatgtgcggatgtatgatg 100
T7-up 51 AATTGCTCAAACAGCCGCCGACGCGGGCTTTTATGTGCGGATGTATGATG 100
mmgB 101 tgaatccagaggccgcggaggcaggattgaaacggctgaagaaacagctg 150
T7-up 101 TGAATCCAGAGGCCGCGGAGGCAGGATTGAAACGGCTGAAGAAACAGCTG 150
mmgB 151 gcccgtgatgctgagaaaggaaaaaggaccgagacggaagtgaagagcgt 200
T7-up 151 GCCCGTGATGCTGAGAAAGGAAAAAGGACCGAGACGGAAGTGAAGAGCGT 200
mmgB 201 aatcaaccgcatttcgatttctcaaacacttgaggaggcagagcatgcgg 250
T7-up 201 AATCAACCGCATTTCGATTTCTCAAACACTTGAGGAGGCAGAGCATGCGG 250
mmgB 251 acattgtgattgaggctatcgcagaaaaacatggcggcaaaaactgagatg 300
T7-up 251 ACATTGTGATTGAGGCTATCGCAGAAAACATGGCGGCAAAAACTGAGATG 300
mmgB 301 tttaaaacacttgatcgcatttgcccgcctcatacgattttggccagcaa 350
T7-up 301 TTTAAAACACTTGATCGCATTTGCCCGCCTCATACGATTTTGGCCAGCAA 350
mmgB 351 tacatcttccttgcctattacagaaatcgctgctgtaacaaaccggcctc 400
T7-up 351 TACATCTTCCTTGCCTATTACAGAAATCGCTGCTGTAACAAACCGGCCTC 400
mmgB 401 aacgggttattggcatgcattttatgaatcccgtccctgtaatgaagctg 450
T7-up 401 AACGGGTTATTGGCATGCATTTTATGAATCCCGTCCCTGTAATGAAGCTG 450


mmgB 451 gtagaagtgattcgaggcttggctacatcagaagaaacggccttagatgt 500
T7-up 451 GTAGAAGTGATTCGAGGCTTGGCTACATCAGAAGAAACGGCCTTAGATGT 500
mmgB 501 tatggcattagcggaaaagatggggaaaacagcggtagaagtcaatgatt 550
T7-up 501 TATGGCATTAGCGGAAAAGATGGGGAAAACAGCGGTAGAAGTCAATGATT 550
mmgB 551 ttcctgggtttgtttccaaccgtgtgcttcttccaatgattaatgaagcc 600
T7-up 551 TTCCTGGGTTTGTTTCCAACCGTGTGCTTCTTCCAATGATTAATGAAGCC 600
mmgB 601 atctattgcgtgtatgagggagtggcgaagccggaggcaatagatgaagt 650
T7-up 601 ATCTATTGCGTGTATGAGGGAGTGGCGAAGCCGGANGCAATAGATGAAGT 650
mmgB 651 gatgaagctgggcatgaatcatccgatgggtccgcttgcattagcggatt 700
T7-up 651 GATGAAGCTGGGCATGAATCATCCGATGGGTCCGCTTGCATTAGCGGATT 700
mmgB 701 ttatcggactggatacgtgtttatcaattatggaagtccttcactcaggc 750
T7-up 701 TTATCGGACTGGATACGTGTTTATCAATTATGGAAGTCCTTCACTC-NNN 749
mmgB 751 cttggcgatt-ccaaataccgtccttgcccgctgctccgcaagtatgtca 799
T7-up 750 CTTGNCGATTNCCAAATACCGTCCTTGCCCGCTGCTCCGCAAGTATGTCA 799
mmgB 800 aagcaggctggcttggcaaaaagagcggacgcggtttttatgactatgag 849
T7-up 800 AAGCAGGCTGGCTTGG-NNAANNAGCGNANGCGNTTTTTATGACTATGAN 848
mmgB 850 gagaagacttcctga 864
T7-up 849 NAGAANANTTCCTGA 863

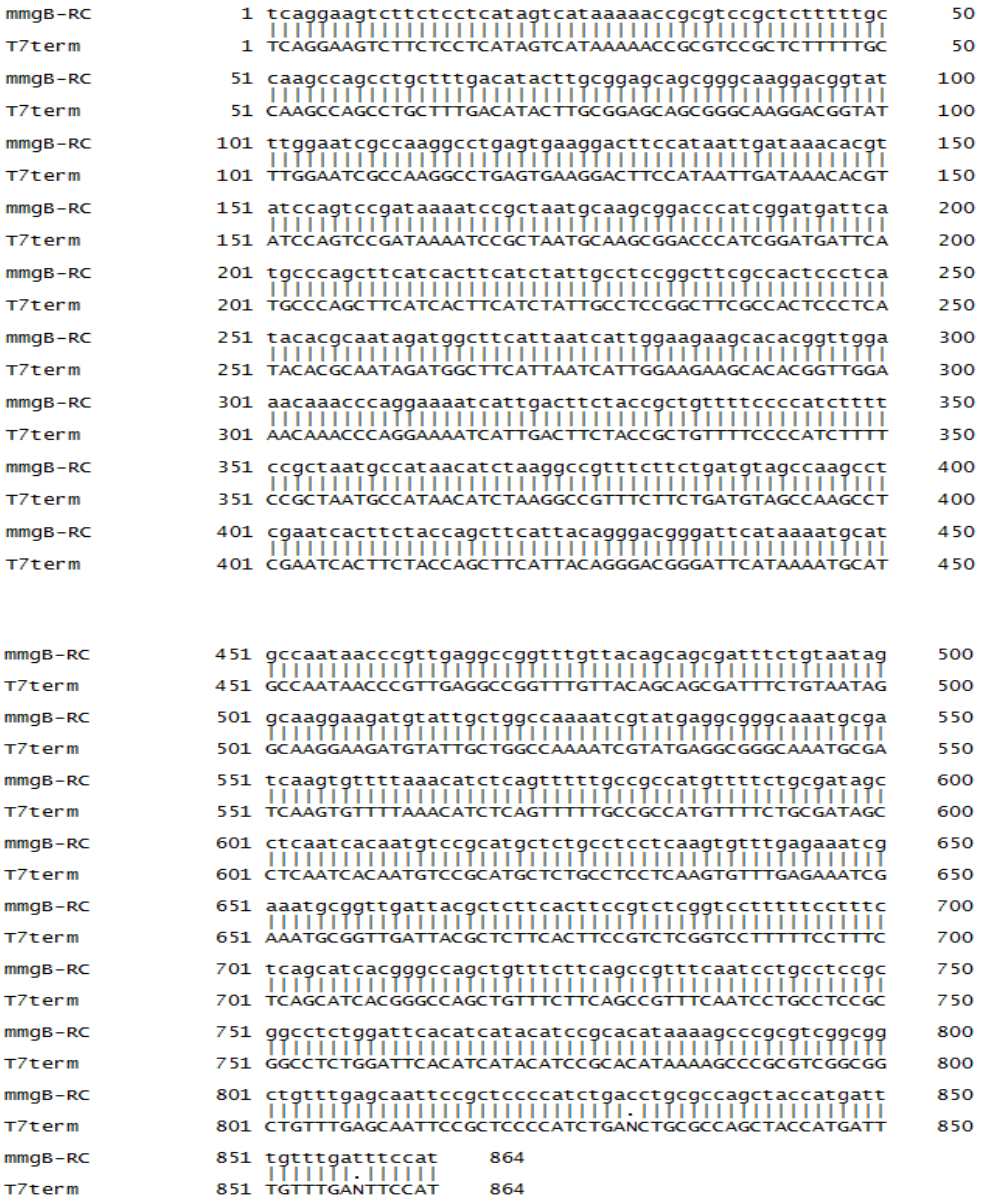


Figure 19: Alignments of *mmgB* insert in pET-200 vector and actual, new *mmgB* gene sequences.

IV. B. Overexpression and Isolation of mmgB protein

With successful completion of cloning, our next goal was to over-express *mmgB* for over-production of the corresponding protein (MmgB). For this purpose, large volumes (1L) of cultures were grown. The presence of antibiotic in the culture allowed for the selective growth of transformants only. Addition of IPTG to the culture induces

overexpression of *mmgB.* This is because the *mmgB* insert in pET-200 vector is under the control of a T7 promoter present in the vector i.e., a transcriptional fusion is constructed (Figure 20). Consequently, the transcriptional activity of the *mmgB* gene is under the control of T7 promoter.

The T7 promoter is highly specific and is recognized only by T7 RNA polymerase. But the pET-200 vector lacks the gene for T7 RNA polymerase. Hence, BL21 (DE3) cells were used for propagation of the vector and subsequent over expression of the *mmgB* (Figure 20).

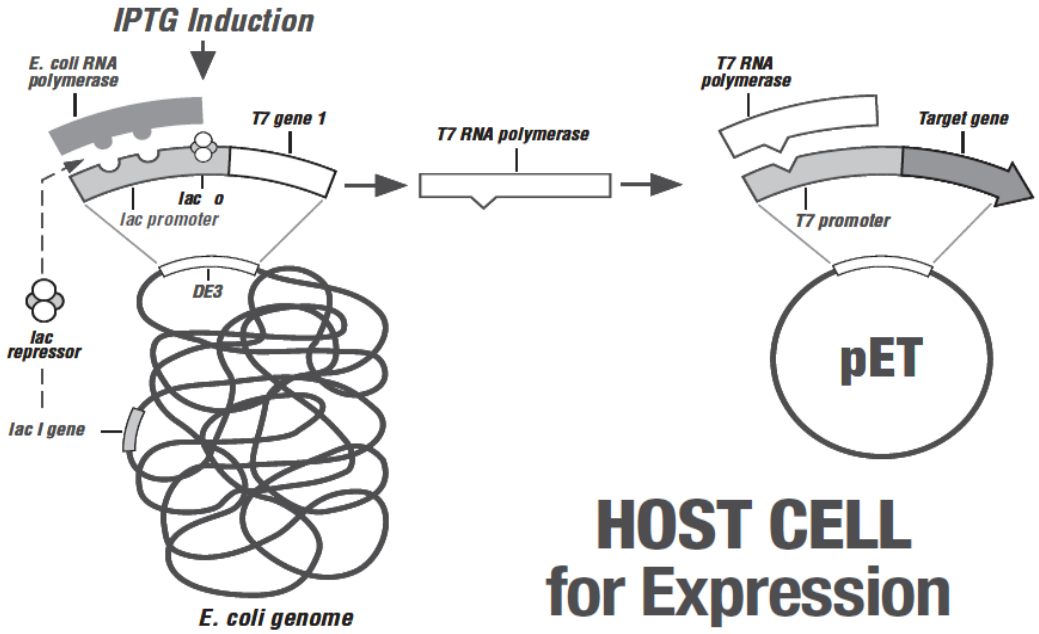


Figure 20: Mechanism of over expression system[21].

BL21 (DE3) is a mutated strain of *E.coli* (lysogen of bacteriophage λDE3), which contains the T7 RNA polymerase encoding gene but, under the control of the *lac* promoter. The BL21 genome contains the *lacI* gene which encodes the lac repressor

protein that binds to the lac operator and inhibits the transcriptional progress of T7 RNA polymerase gene under the control of the *lac* promoter. Hence, to overcome this inhibition due to the lac repressor, IPTG, a molecular mimic of allolactose which is the substrate for *lac* repressor, is added. IPTG binds to the lac repressor protein and inactivates it, thereby allowing for the expression of T7 RNA polymerase. Thus owing to the accessibility of the *lac* promoter, the T7 RNA polymerase produced then binds to the T7 promoter present in the vector and initiates the transcriptional activity of downstream genes, which in our case is *mmgB*. Thus, the over expression of the *mmgB* is achieved *via* the addition of IPTG (Figure 20)[21].

IV.B.1. Ni-NTA Column Chromatography

Following overexpression of the *mmgB*, our next goal was to isolate the protein to near homogeneity. To do so, we employed Ni-NTA column chromatography as a separation technique. The reason for using this chromatographic technique is our protein-mmgB is a recombinant fusion protein with an N-terminal 6x His-tag (6 consecutive histidines) owing to the presence of the corresponding nucleotide sequence from the pET-200 vector. As shown in Figure 13, the Ni-NTA column consists of an agarose-nitrolotriacetic acid resin charged with hexadentate Ni^{2+} ions held onto the resin by tetra dentate nitrolotriacetic acid molecules. When applied, the His-tag protein(s) bind to the remaining two coordination sites of Ni^{2+}. Thus, this method helps in isolating the target protein(s)[21, 24].

After the column is equilibrated, the crude extract (filtrate) was applied onto the column and allowed to pass through it under the influence of gravity. Once all the crude

extract was run through the column, the desired 6x His-tag protein was bound to the Ni^{2+} ions in the column.

The column was then treated with buffers having increasing amounts of imidazole. Binding buffer (5 mM imidazole) allows the His-Tag protein to bind to resin tightly while the wash buffer washes away the undesired proteins. This is possible because, wash buffer (60 mM imidazole) has high concentrations of imidazole which mimics the histidine residues of the proteins and these compete with the His-tag protein for the Ni^{2+} ions present in the agarose bed. This competition causes the proteins without His-tag to get washed off the column due to weak interactions with agarose/Ni^{2+} ions leaving His-tag proteins only[24].

With only His-tagged protein bound to the resin, it was eluted from the column using elute buffer with even higher concentrations of Imidazole (200 mM). This high concentration of Imidazole disrupts the interactions between the agarose/Ni^{2+} and 6x histidines of the His-tag protein and causes the His-tag protein to elute off the agarose bed of the column[24].

IV.C. Characterization of mmgB protein

IV.C.1. SDS-PAGE

Following the isolation process, samples of the protein were analyzed by 12% SDS-PAGE (Figure 21) to confirm the identity of the mmgB protein obtained after chromatographic separation.

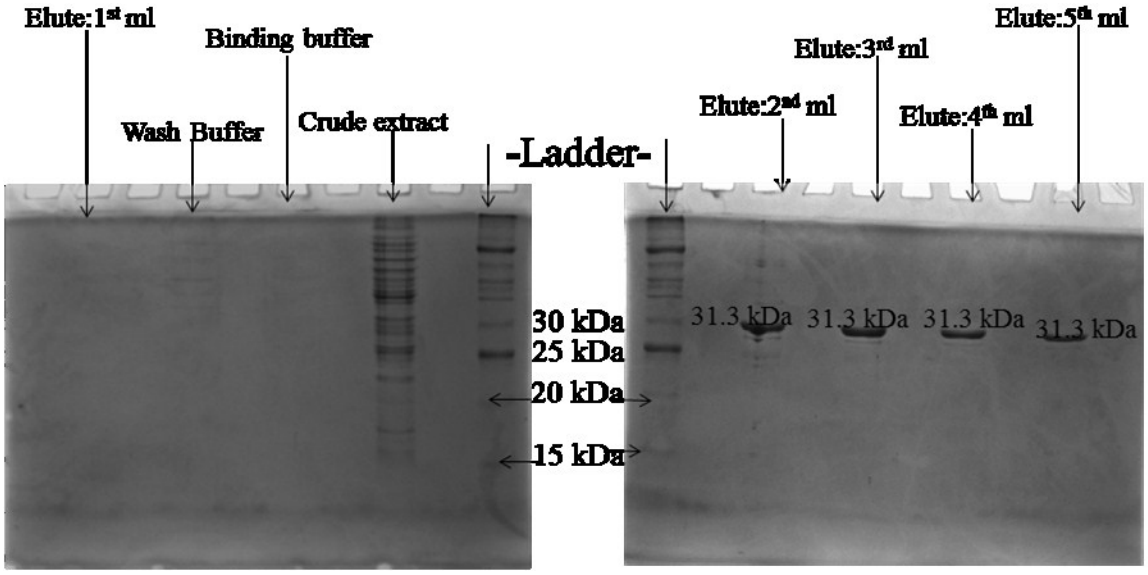


Figure 21: 12% SDS-PAGE gel of mmgB protein before Dialysis.

The SDS-PAGE gel picture (Figure 21) shows lanes labeled as ladder, crude extract, binding buffer, wash buffer and elute: 1st, 2nd mL, etc. The lanes ladder corresponds to the 10-log protein ladder which served as a molecular weights standard. The bands seen in the lanes ladder are considered as standards with reference to characterizing the bands of unknown molecular weights (in kDa) seen on the gel.

The lane crude extract corresponds to the flow through collected after allowing the crude extract to pass through the column. This shows many different bands representing various proteins expressed during the life time of BL21 cells. In the lane crude extract, alongside the various bands, a light band at slightly above 30 kDa was also seen. This is consistent with theoretical mass of the mmgB protein: 31.3 kDa[25]. This shows that, mmgB is also expressed. The lanes binding buffer and wash buffer show no bands at about 30 kDa in reference to ladder and this is suggestive that the actual his-tag protein mmgB was still bound to the column.

Out of the lanes corresponding to elute buffer -1st, 2nd mL, the lane corresponding to 1st mL did not show any band at about 30 kDa corresponding to mmgB protein. However, each of the lanes corresponding to the 2nd through 5th mL of the elute buffer showed a band at slightly above 30 kDa which is consistent with the theoretical mass of the mmgB protein, 31.3 kDa.

Hence, looking at the bands on the gel (Figure 21), it is evident that we achieved the purification of a protein having a mass consistent with the desired MmgB protein. This result is suggestive that over-expression and isolation processes were successfully accomplished.

IV.C.2. Gel-Based Trypsin Digest; MALDI-TOF Mass-Spectrometry

To reinforce the identity of the mmgB protein, MALDI-TOF- MS analyses of in-gel trypsin digests of the bands from SDS-PAGE (Figure 21) were performed. When the gel band pieces were incubated with Trypsin, it cleaved the mmgB protein into smaller fragments. Each of these tryptic fragments have specific mass values based on their amino acid composition. When mass-spectrometric analysis is carried out, the mass spectrum showed peaks at the m/z corresponding to the m/z of each of the tryptic fragments (experimental values). The theoretical m/z of tryptic fragments for any protein can be calculated either manually or can be obtained by making use of online tools available such as Peptide Mass [26].

Below are shown the MALDI-TOF-MS spectra (Figures: 22, 23 and 24) for in-gel trypsin digests of mmgB protein. The peaks in the mass spectra (spectra are obtained from different spots of the MALDI plate but of the same trypsin digest sample) with

numerical values correspond to tryptic fragments of mmgB with their respective m/z. To make it easier to compare the theoretical values with the experimental values, a table listing this data is shown after the MALDI-MS spectra (Table 5).

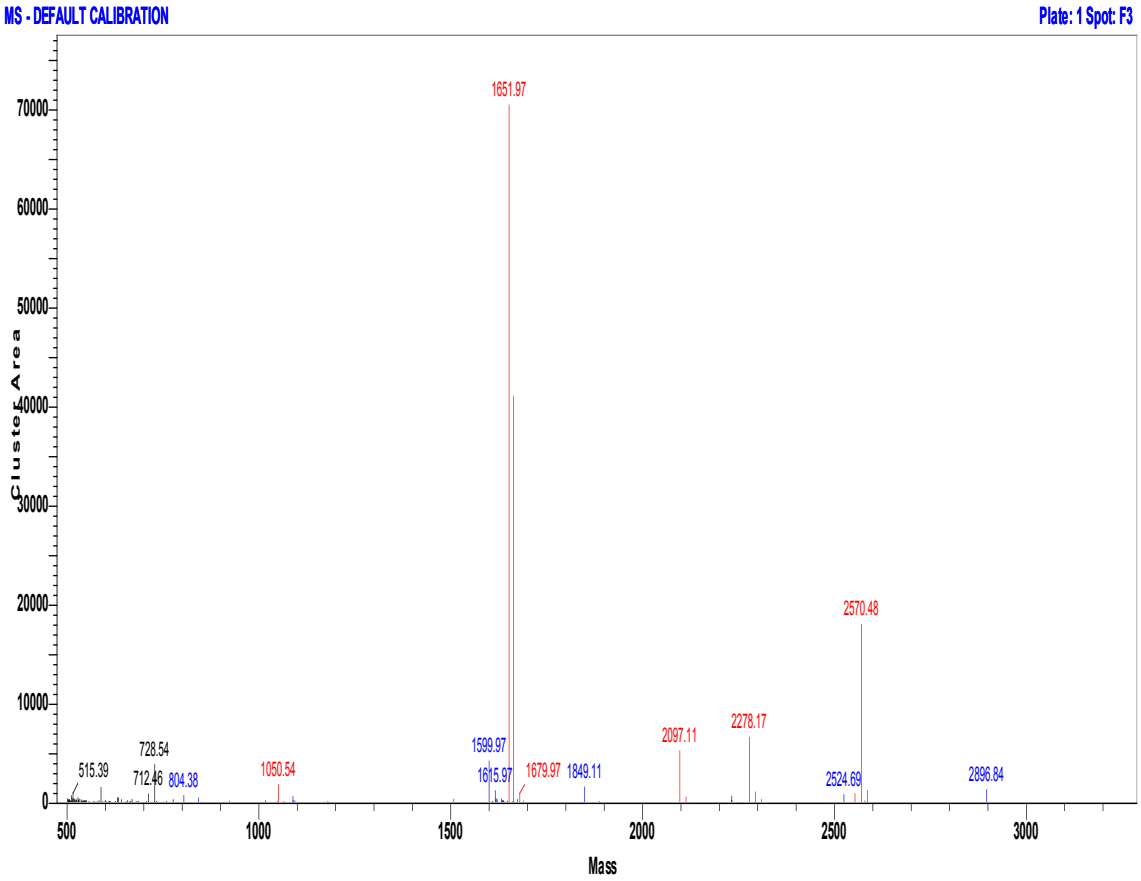


Figure 22: MALDI spectrum (spot-3) of Gel-Based Trypsin Digest of mmgB protein.

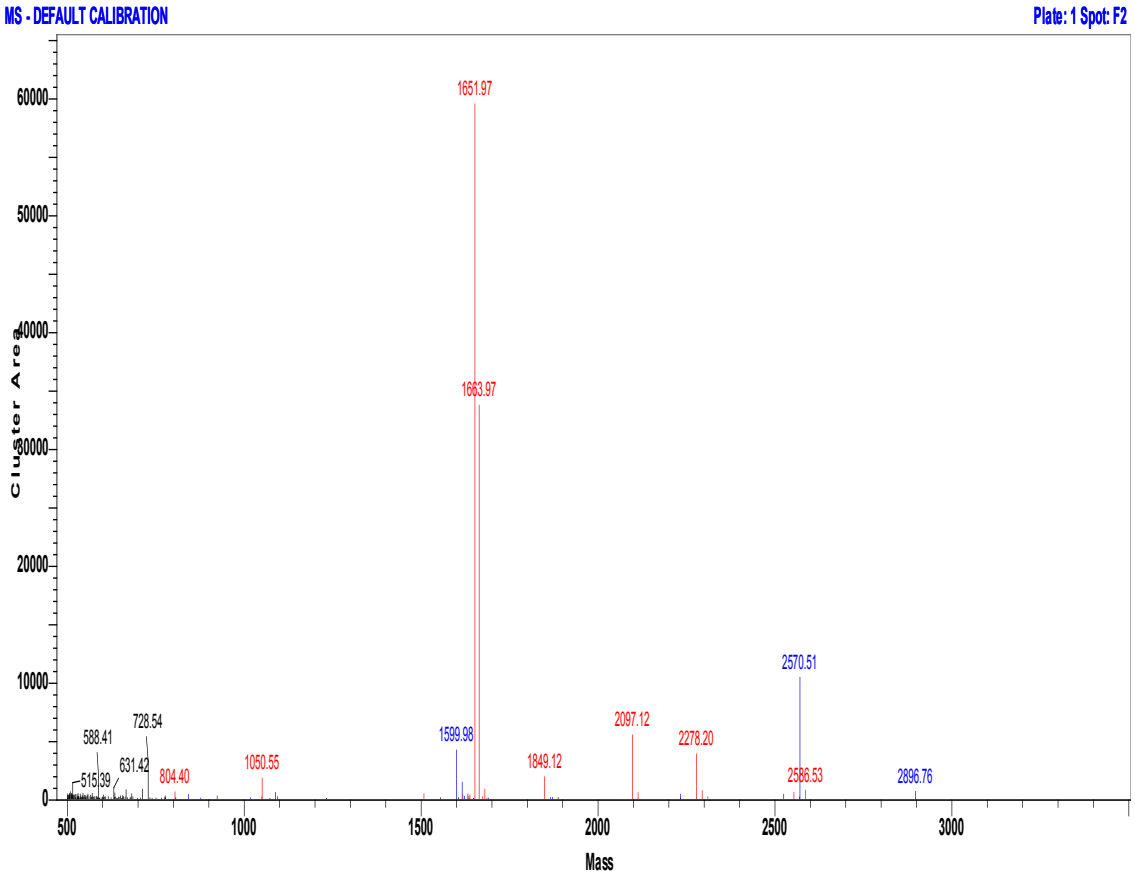


Figure 23: MALDI spectrum (spot-2) of Gel-Based Trypsin Digest of mmgB protein.

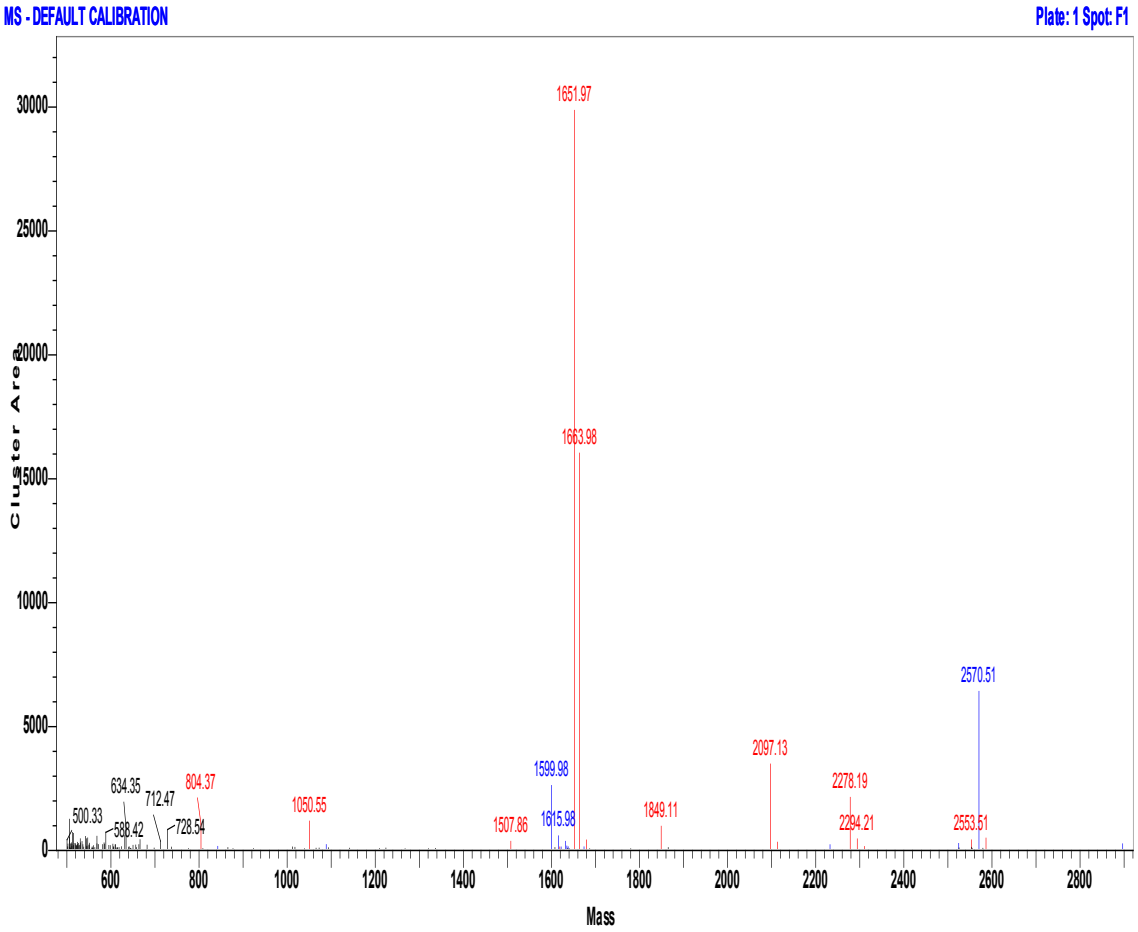


Figure 24: MALDI spectrum (spot-1) of Gel-Based Trypsin Digest of mmgB protein.

Table 5: Comparison between theoretical and experimental values of the masses of trypsin digest peptides

Theoretical values	Experimental values	Theoretical values	Experimental values
503	506	1506	1507
519	515	1598	1599
587	-	1650	1651
630	634		1679
654	-	1847	1849
705	712		2097
727	728		2278
	804	2569	2570
1016	-		2586
1049	1050	2895	2896

Looking at the numbers between theoretical and experimental values from the table above, we can say most of the values are matching. However some of them with lower m/z do not match perfectly. Nevertheless, we have a very good confidence about the identity of the protein, since the nucleotide sequence of the *mmgB* insert in the plasmid (pET-200_*mmgB*) was a 100% match with the known *mmgB* gene sequence.

At this juncture we do not have a proper understanding of why some of the experimental values are off. The possible reasons could be improper calibration of the instrument at low m/z range or post-translational modifications of the protein or may be the very low concentrations of the protein used.

Dialysis was performed to remove the excess salts and imidazole present in the protein solution, coming from chromatography. Since MWCO of dialysis bag is 7000 Da any of the molecules/compounds with MWCO below this value will be separated out of the protein solution.

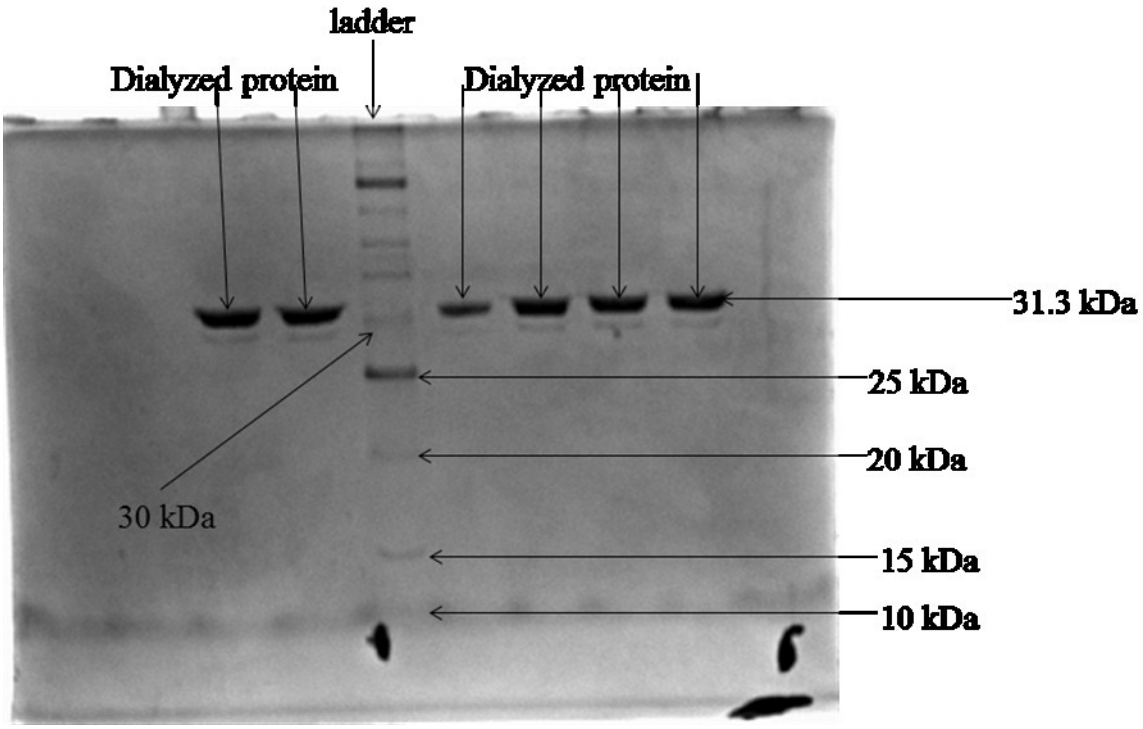


Figure 25: 12% SDS-PAGE gel of mmgB protein after Dialysis.

Above is shown 12% SDS-PAGE gel picture (Figure 25) shows lanes labeled as ladder and dialyzed protein. The lane ladder corresponds to 10-log protein ladder; it serves the purpose of a molecular marker. The bands seen in the lane ladder; are considered as standards in order to characterize the bands of unknown molecular weights (in kDa) seen on the gel.

Each of the lanes corresponding to dialyzed mmgB protein show a band at slightly above 30 kDa which is consistent with the theoretical mass of the mmgB protein (31.3 kDa). Also, the position of the bands on the gel is in good compliance with the previous SDS-PAGE gel (Figure 21). This is suggestive that we have the right protein saved for future experiments, and that dialysis did not result in any undesirable decomposition.

After doing a general Bradford assay (Figure 26), we found that the yield of mmgB protein was 4.0 mg protein produced per liter of the culture of BL21 cells.

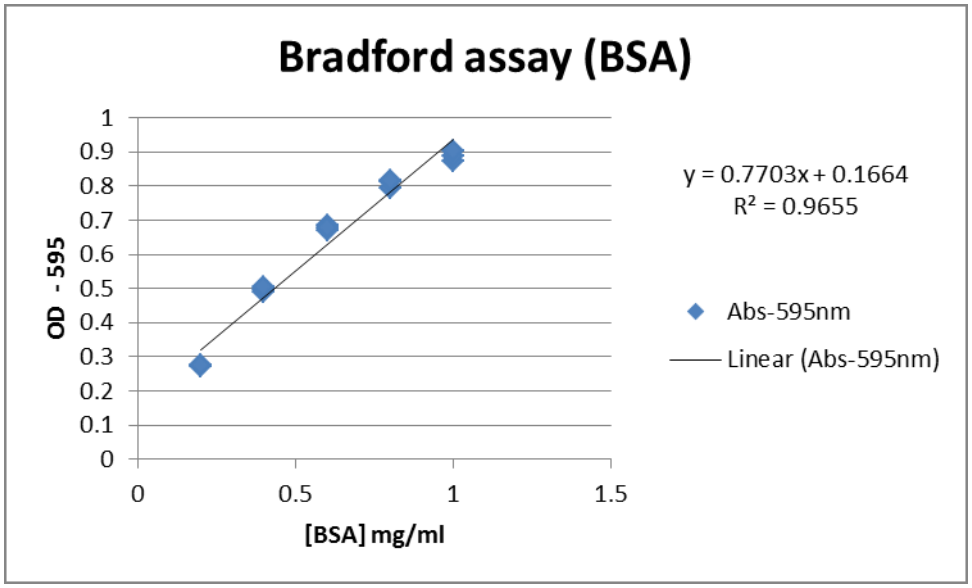


Figure 26: Bradford assay using Bovine Serum Albumin standards.

IV.D. Enzymatic Activity Analyses of mmgB protein

With the protein (mmgB) available in pure form, our next goal was to characterize its biochemical activity.

Based on sequence homology, mmgB is proposed to exhibit 3-Hydroxy-Butyryl-CoA Dehydrogenase (E.C. 1.1.1.157) activity (Figure 26). Thus, the mmgB protein is an Oxido-reductase, catalyzing the oxidation of 3-Hydroxy-Butyryl-CoA (3-HB-CoA) to Acetoacetyl-CoA (AcAc-CoA). Homologs of this enzyme have been previously characterized from a variety of organisms and have been shown to use either NAD^+ or $NADP^+$ as the co-factor[10, 19, 27].

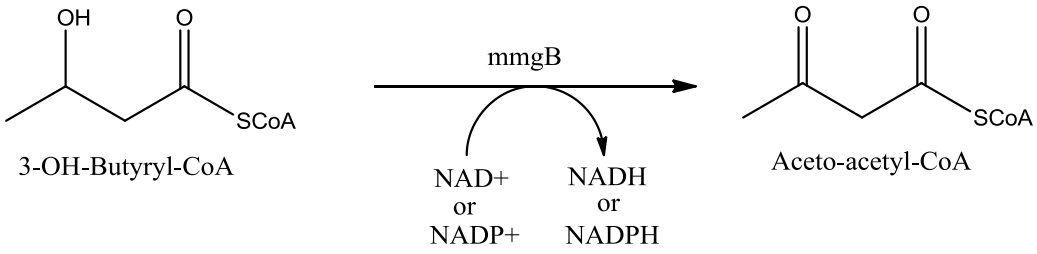


Figure 27: Proposed enzymatic reaction catalyzed by mmgB protein.

One of the ways of determining the activity of the enzyme and also the progress of the reaction is to spectrophotometrically monitor the increase in absorbance owing to reduction of either NAD^+ or $NADP^+$ to NADH or NADPH respectively at 340 nm (λ_{max} of both NADH & NADPH).

IV.D.1. Determination of appropriate co-factor

Since no biochemical data is available for mmgB, we did not know the co-factor preference, if any of the enzyme. So, we put in our initial efforts in determining the co-factor preference of the enzyme between NAD^+ and $NADP^+$.

With NAD^+, we could never see any significant activity of the enzyme; i.e., there was no increase in the absorbance at 340 nm owing to the reduction of NAD^+ to NADH. Hence, the enzyme was not functional with NAD^+ as co-factor. This can be seen in the UV/Vis-time course shown below (Figure 28).

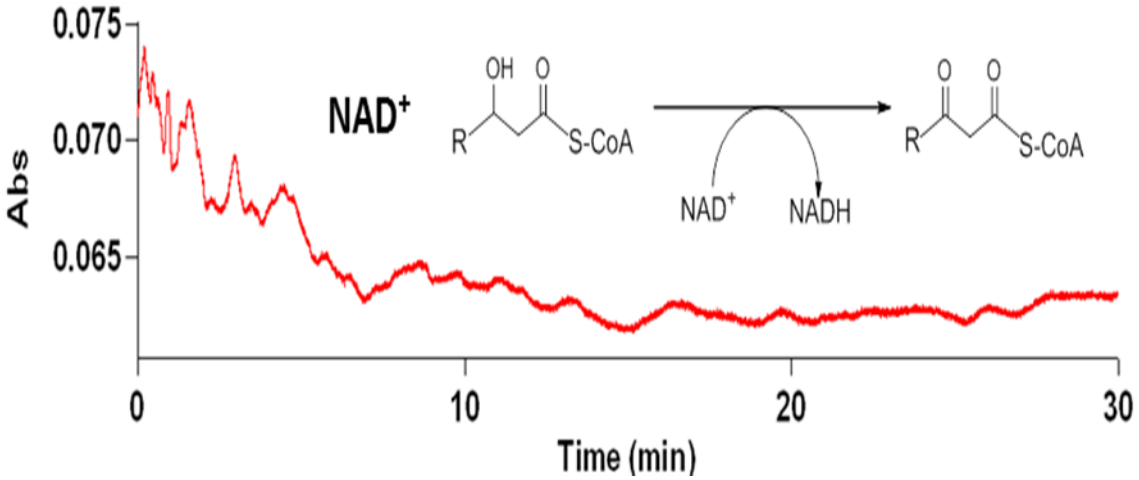


Figure 28: UV/Vis-time course for the reaction with NAD+ as co-factor.

On the other hand, the reaction progressed using $NADP^+$ as cofactor. With $NADP^+$, we could see a significant increase in the absorbance at 340 nm owing to the reduction of $NADP^+$ to NADPH. Hence, the enzyme was specifically functional with $NADP^+$ only as the cofactor. This can be seen in the UV/Vis-time course shown below (Figure 29).

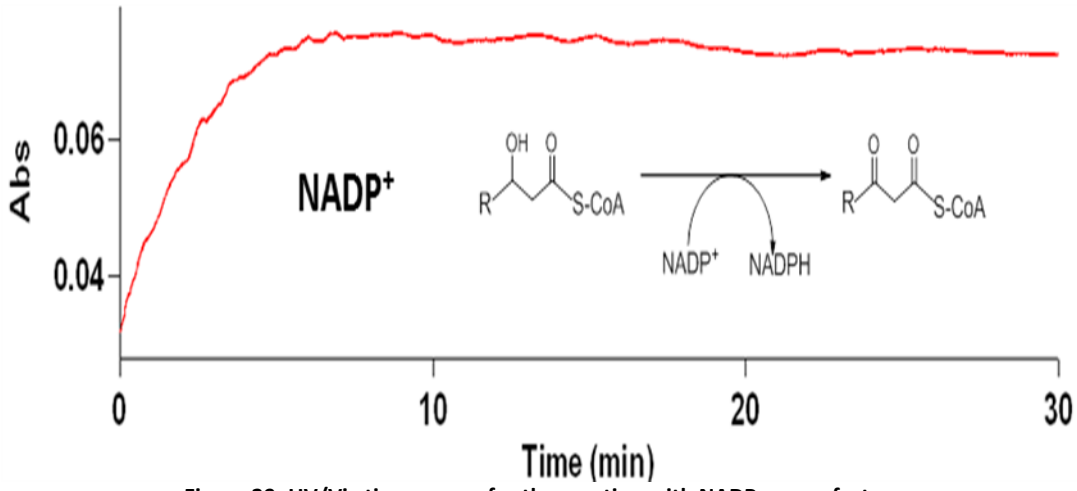


Figure 29: UV/Vis-time course for the reaction with NADP+ as co-factor.

IV.D.2. Determination of optimal pH

With the co-factor preference determined, the next task was to determine the pH optimum for mmgB. Determining the optimal pH of an enzyme is critical, because pH plays a pivotal role in defining the activity of the enzyme. So, one must consider the pH aspect when designing an assay for determining the enzymatic activity of a protein.

The literature shows that, 3-HBDH from *Clostridium kluyveri* in the acetoacetyl-CoA (AcAc-CoA) forming direction exhibits maximal activity at a pH between 9.0 and 10.0 (Figure 30). The enzyme has shown a decreasing trend in terms of its activity out of this pH range (9.0-10.0) – both below 9.0 and above 10.0. Similarly, the same enzyme (3-HBDH) from *Clostridium beijerinckiii* showed maximal activity in AcAc-CoA forming direction at pH 8.0 and the activity decreased when the pH is either increased or decreased away from this value. These studies are suggestive that the enzyme in AcAc-CoA forming direction is active under basic conditions[14, 19, 27].

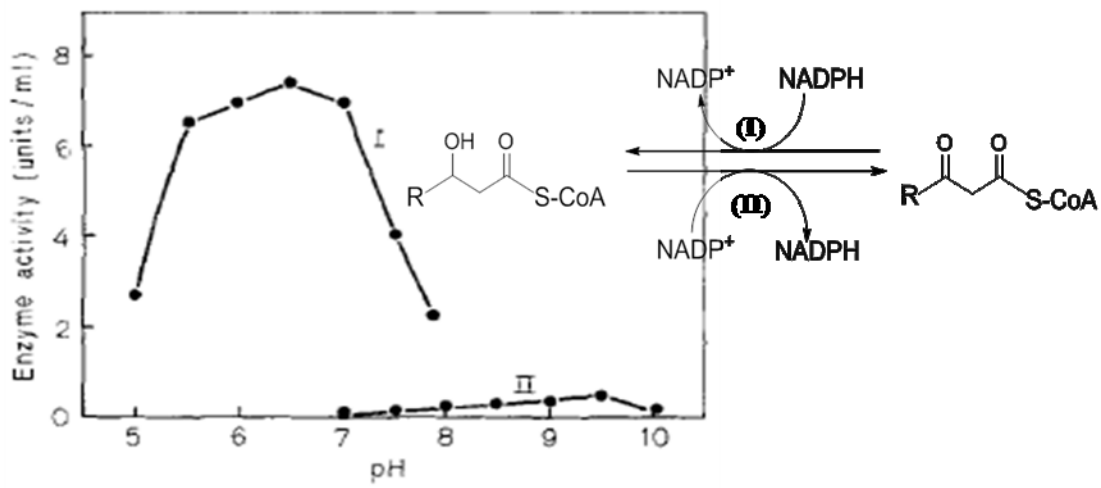


Figure 30: Literature precedent for basic pH dependence of the 3-OH-butyryl-CoA DHase enzyme[27] (adapted).

Based on the precedent from the literature, we looked at buffer solutions whose pKa values fall in the basic buffering range of ~8.0-10.0. With Glycine-NaOH buffer at pH 9.8, the enzyme showed highest activity (Figure 31). This looked more promising upon drawing a comparison between the Figures 29 and 31, since, Figure 29 (pH 7.3) did not show much of an increase in the absorbance whereas the Figure 31 (pH 9.8) showed a prominent increase in absorbance (1.0000) owing to the formation of NADPH. This fact was further supported by our findings about the pH dependence of the enzyme, 3-OH-Butyryl-CoA-DHase (mmgB). For more details about pH dependent behavior of the enzyme, see Figure-55 (supplementary data).

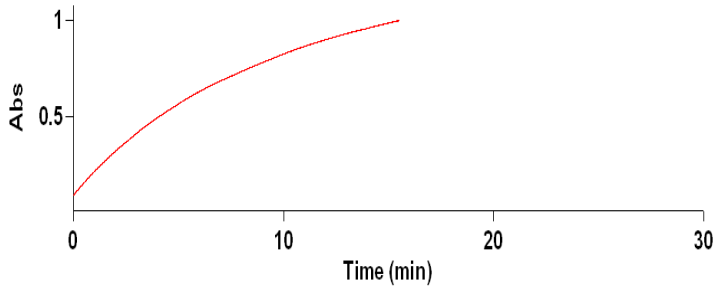


Figure 31: UV/Vis-time course for the enzymatic activity assay with Glycine-NaOH buffer, pH 9.8.

IV.D.3. Determination of enzymatic activity of mmgB protein

The enzyme has shown maximal activity with the following composition of the reaction mixture: 0.4 mM of the substrate - D, L-3-HB-CoA (racemic mixture), 0.2 mM $NADP^+$, 2.0 µM enzyme and Glycine-NaOH buffer, pH 9.8 with the reaction volume of 500 µL.

Upon mixing the components well, we spectrophotometrically monitored the reaction at 340 nm. The increase in absorbance showed a nice rectangular hyperbolic

curve, indicative of the progression of the reaction *via* reduction of $NADP^+$ to NADPH. This can be seen in the UV/Vis-time course shown below (Figure 32).

Later, following spectrophotometric analysis, the reaction mixture was analyzed by MALDI-TOF-MS (Figures 32 and 37). This was performed because the mass spectrum of the reaction mixture can show peaks at the m/z corresponding to the compounds present in the sample. Based on the m/z values seen, we can determine the identity of the compounds present and hence subsequently the identity of the reaction itself.

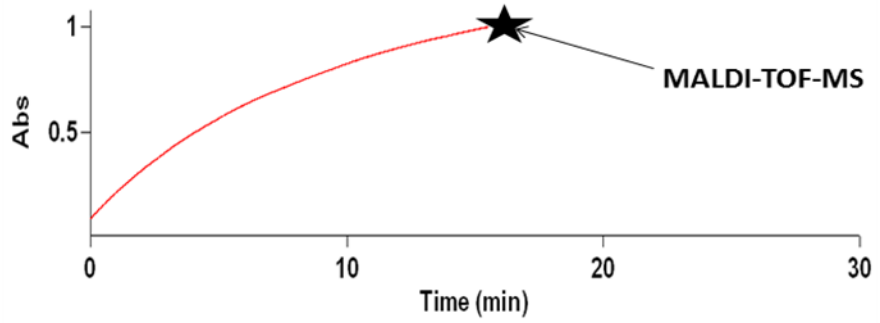


Figure 32: Spectrophotometric monitoring of the reaction followed by MALDI-TOF-MS analysis.

Going with the above mentioned reason, before actually analyzing the reaction mixture by mass spectrometry, we instead analyzed a set of standards and control reactions using MALDI-TOF MS. Below shown are the MALDI-TOF-MS spectra (Figure 33) for standards for the substrate (3-HB-CoA) with m/z of 854, co-factor both oxidized ($NADP^+$) and reduced (NADPH) forms with m/z of 744 and 746, respectively, and the product (AcAc-CoA) with m/z of 852.

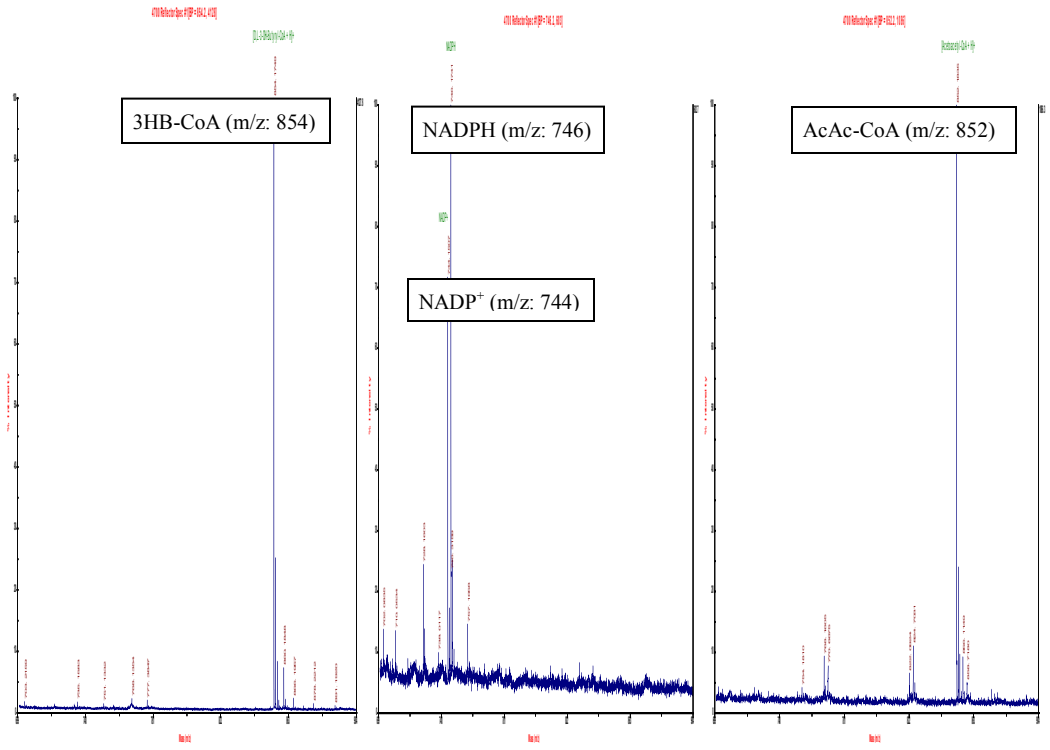


Figure 33: MALDI-TOF-MS spectra for the standards: substrate, NADP+, NADPH, and product.

All the above three spectra pertaining to standards were obtained by employing pure compounds. The spectrum for 3-HB-CoA was obtained by doing MALDI-TOF-MS analysis of the sample containing only 3-HB-CoA along with matrix and diluent. So, the mass spectrum shows a single prominent peak at an m/z of 854 corresponding to the m/z of 3-HB-CoA other than the short peaks corresponding to the components of matrix/diluent. Similarly, spectra for other standards ($NADP^+$, NADPH and AcAc-CoA) were also successfully obtained.

Other short peaks were seen very commonly with no significant abundance value. Of these, the peak corresponding to m/z = 768 is the most commonly seen one. It corresponds to the mass of CoA (Co-enzyme-A). These (including CoA) are seen consistently all through the MALDI-TOF-MS analyses. CoA- thioesters are labile and so

it is not surprising to see it repeatedly in various MS-spectra. Na^+ - adducts will also be seen consistently in almost all the spectra.

The MALDI-TOF-MS spectrum shown below (Figure 34) corresponds to a no-enzyme (3HB DH) control. A reaction was set up with all other components (substrate and co-factor) employed, except for the enzyme. So, there should not be any progress of the reaction and consequently no product (AcAc-CoA; m/z = 852) formation. Also the co-factor $NADP^+$ will not undergo reduction to NADPH (m/z = 746). Thus, the mass-spectrum shows peaks corresponding to only the substrate (m/z = 854) and co-factor ($NADP^+$; m/z = 744). Figure 54 (supplementary data), corresponds to the UV/Vis – time course for no enzyme control.

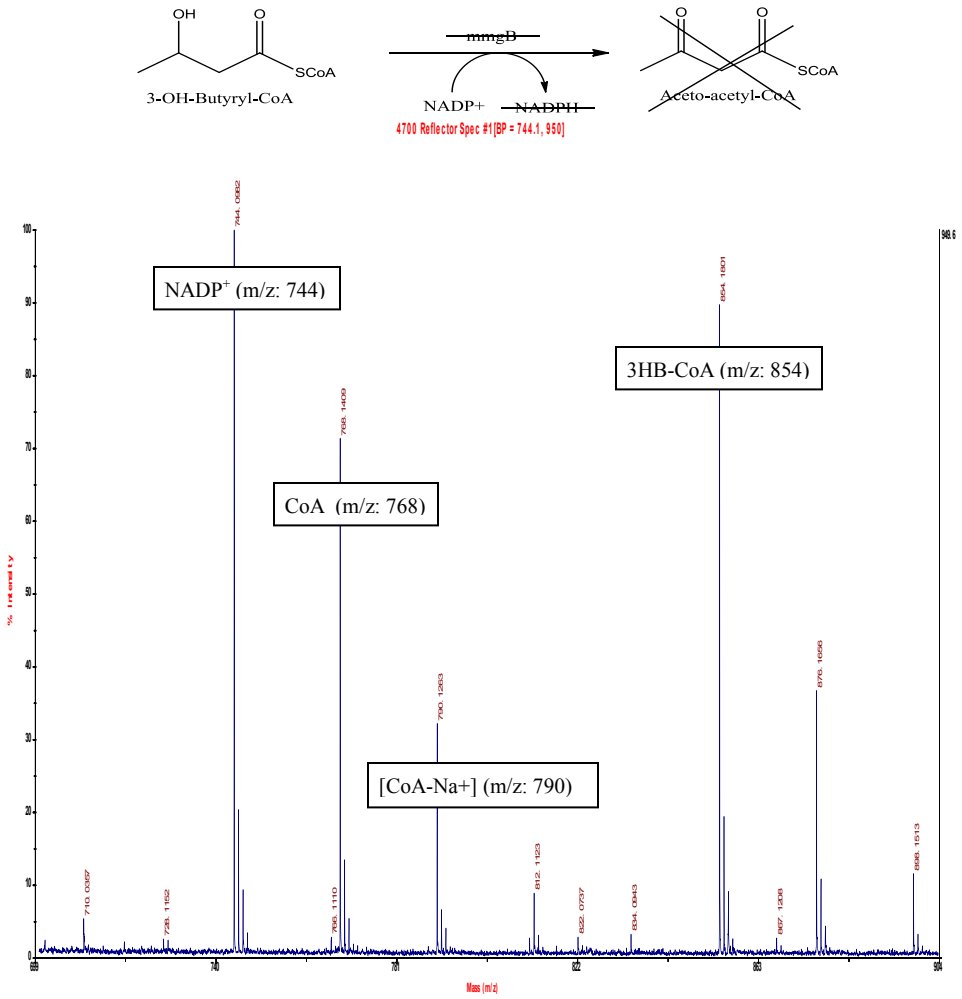


Figure 34: MALDI-TOF-MS spectrum for no enzyme control.

The MALDI-TOF-MS spectrum shown below (Figure 35) corresponds to the no-substrate control (3-HB-CoA; m/z = 854). A reaction was set up with all other components (enzyme and co-factor) employed, except for the substrate. So, the reaction should not progress and hence has zero product (AcAc-CoA; m/z = 852) formation. Also the co-factor $NADP^+$ will not undergo reduction to NADPH (m/z = 746). The mass-spectrum indeed shows only a single prominent peak (m/z =744) corresponding to the co-

factor ($NADP^+$) alone. Figure 53 (supplementary data), corresponds to the UV/Vis – time course for no substrate control.

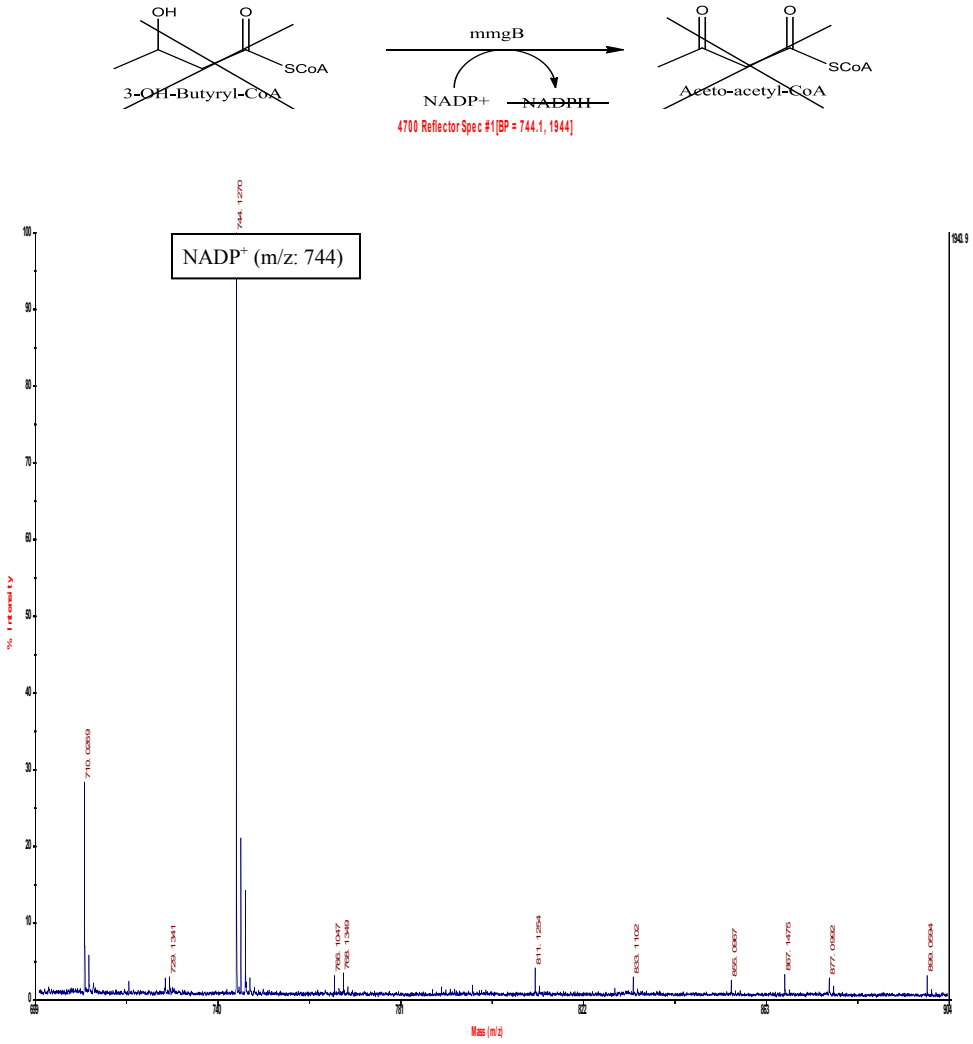


Figure 35: MALDI-TOF-MS spectrum for no substrate control.

The MALDI-TOF-MS spectrum shown below (Figure 36) corresponds to a no-cofactor control ($NADP^+$; m/z = 744). A reaction was set up with all other components (enzyme and substrate) of the reaction except for the co-factor ($NADP^+$). So, the reaction should not progress and hence not form product. Also formation of NADPH (m/z = 746)

will not occur. Subsequently, the mass spectrum shows only a single prominent peak (m/z =854) corresponding to the substrate (3-HB-CoA) alone. Figure 52 (supplementary data), corresponds to the UV/Vis – time course for no cofactor control.

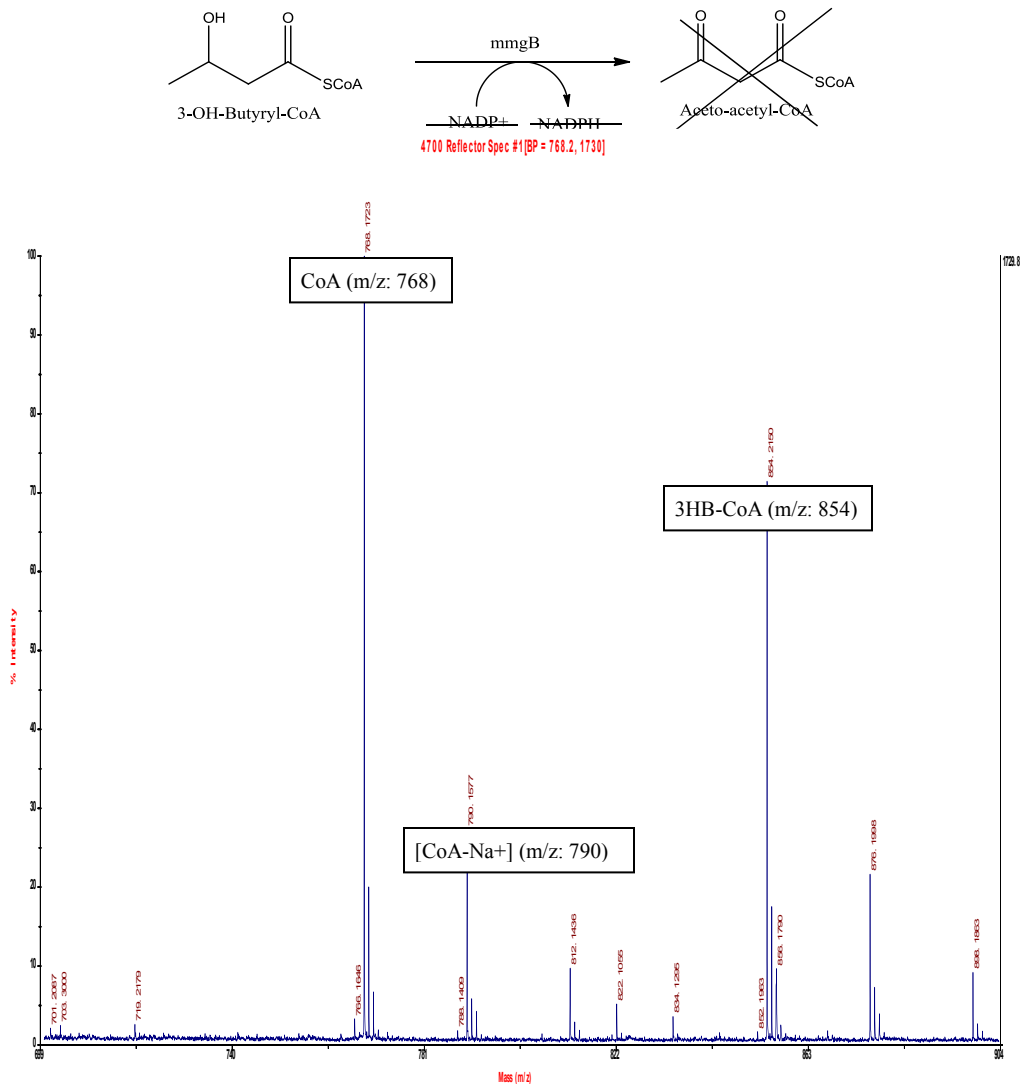


Figure 36: MALDI-TOF-MS spectrum for no co-factor control.

With the standard and control reactions been performed, we now have an idea of peaks (m/z) corresponding to the compounds that we expect to see with the full reaction. So, with full reaction if we see peaks (m/z) corresponding to NADPH and AcAc-CoA, it then explains that *mmgB* indeed encodes for 3-Hydroxybutyryl-CoA-Dehydrogenase (3HBDH).

The MALDI-TOF-MS spectrum shown below (Figure 37) corresponds to the full reaction. A reaction was set up with all the components (enzyme, substrate and co-factor) of the reaction. So, the reaction should progress, leading to the formation of product (AcAc-CoA; m/z = 852) and also NADPH (m/z = 746). The mass-spectrum indeed shows peaks corresponding to the product (AcAc-CoA; m/z = 852) and reduced form of the co-factor (NADPH; m/z = 746) alongside the peaks pertaining to both the substrate (3-HB-CoA; m/z = 854) and the oxidized form of the co-factor ($NADP^+$; m/z = 744) indicating the remnants of the substrates.

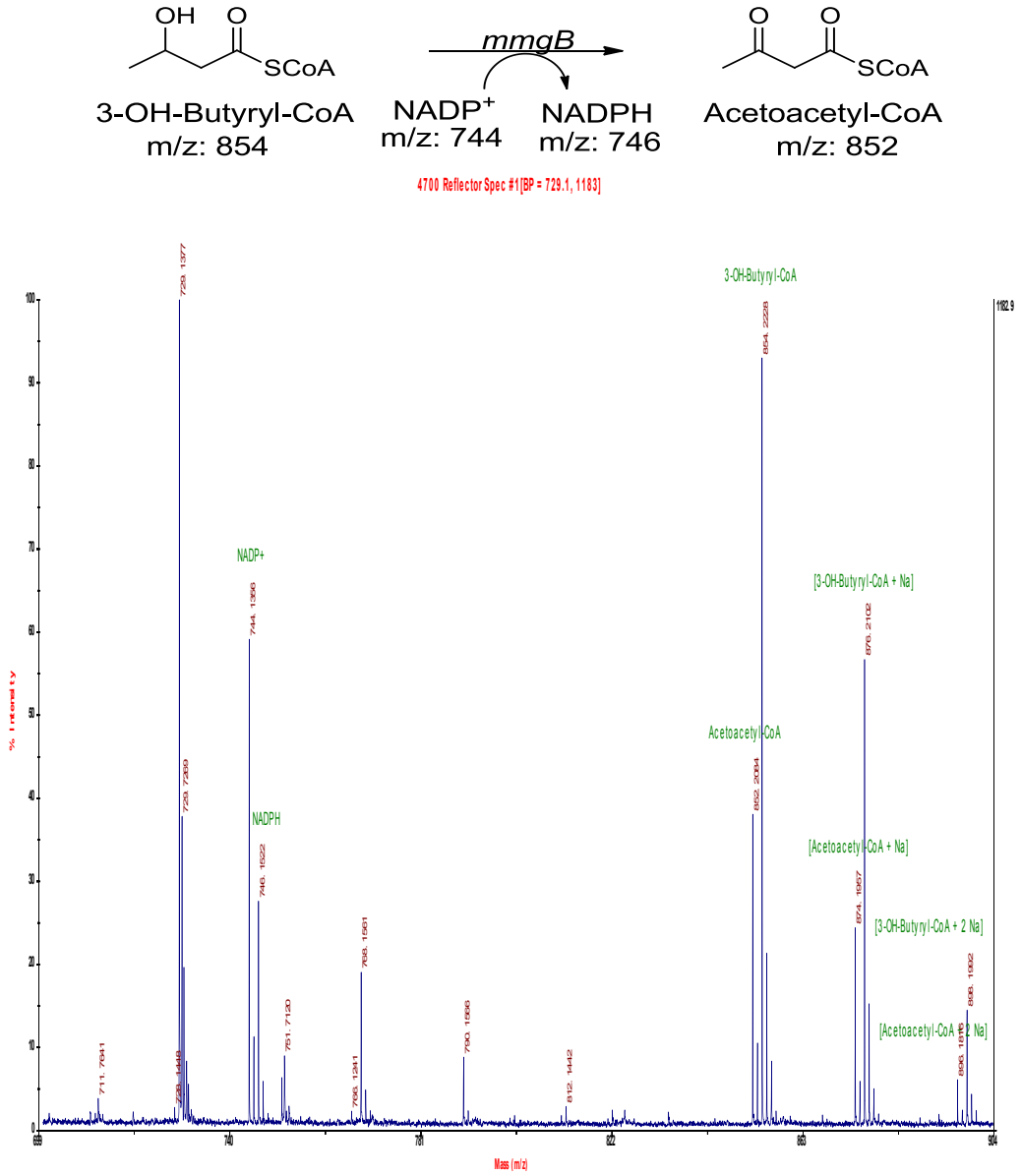


Figure 37: MALDI-TOF-MS spectrum for full reaction.

The above shown MALDI-TOF-MS spectrum (Figure 37) shows peaks corresponding to the product (AcAc-CoA; m/z = 852) and reduced form of the co-factor (NADPH; m/z = 746), it is evident that the reaction has actually progressed and it is because of the catalytic activity of the enzyme with the involvement of both the substrate

and the co-factor. This fact can further be complemented upon overlapping the MALDI-TOF-MS spectra of the controls/standards over that of the full reaction.

IV.D.4. Determination of Enzyme kinetics

After establishing the optimum assay conditions for monitoring the activity of the enzyme, we next moved on determining the kinetic rate constants of the enzymatic reaction using Michaelis-Menten kinetics model. The kinetic rate constants $K_{m\ app}$ (apparent) of both the substrate (3HBCoA) and co-factor ($NADP^+$) were determined individually. The $V_{max\ app}$ (apparent) of the reactions was also determined (Figures 38 and 39).

The $K_{m\ app}$ for substrate (3-HB-CoA) was determined to be 300 μM and the $V_{max\ app}$ of the reaction was shown as 110 μM/s. The turn over number k_{cat} of enzyme was calculated to be 55 s^{-1} of enzyme. The catalytic efficiency of the enzyme k_{cat}/K_m was 0.2 $\mu M^{-1} . s^{-1}$.

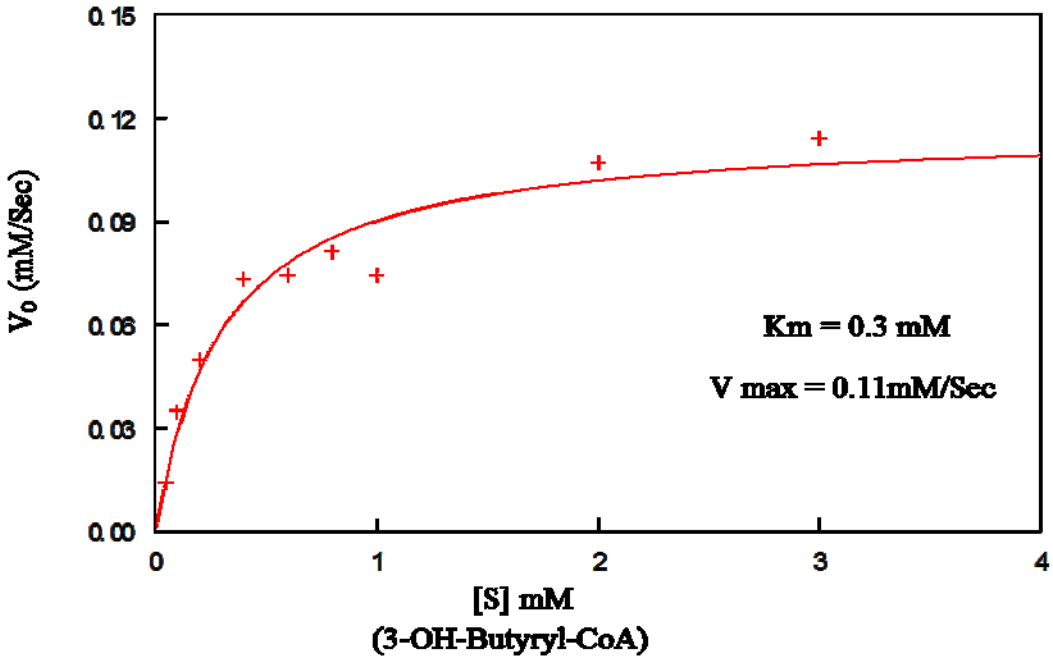


Figure 38: Michaelis-Menten kinetics of the reaction: 3-HB-CoA.

The $K_{m\ app}$ for co-factor ($NADP^+$) was determined to be 200 µM and the $V_{max\ app}$ of the reaction was shown as 800 µM/s. The turn over number k_{cat} of enzyme was calculated to be 400 s^{-1} of enzyme. The catalytic efficiency of the enzyme k_{cat}/K_m was 2.0 $\mu M^{-1} . s^{-1}$.

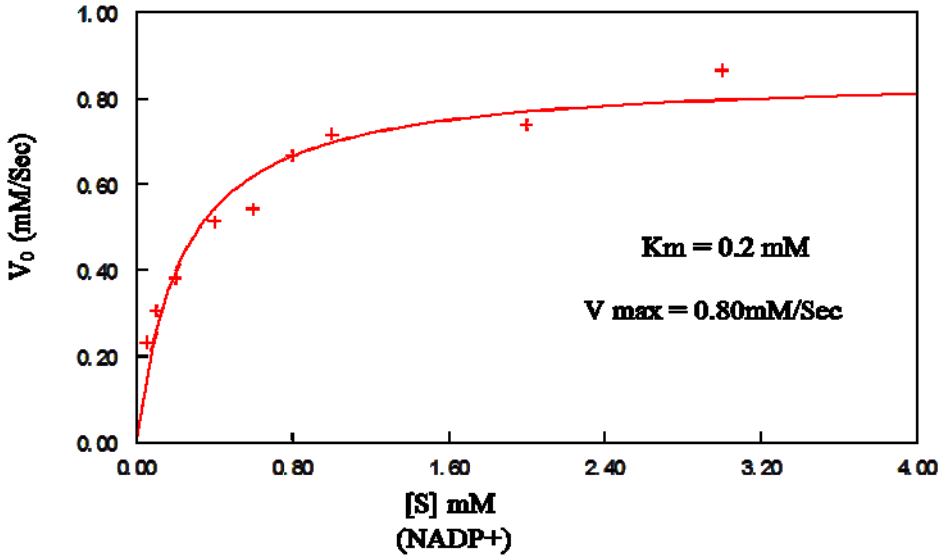


Figure 39: Michaelis-Menten kinetics of the reaction: NADP+.

CHAPTER V

CONCLUSIONS / FUTURE WORK

So far, we successfully cloned the corrected version of the *mmgB* gene (864 bp's) into a suitable expression system using TOPO cloning methodology. Thereafter we successfully over-expressed the mmgB protein by employing BL21 (DE3) strain of *E.coli* cells in an IPTG in dependent manner. Later the N-terminal 6x his-tag recombinant fusion protein - mmgB was successfully isolated by employing Ni^{2+} - NTA column chromatography and dialyzed. With the aid of 12% SDS-PAGE, Trypsin-digest/MALDI-TOF-MS analysis (less confident) and Nucleotide sequencing (100% identical), we were able to prove the identity of the mmgB protein with reference to theoretical details in terms of molecular weight and DNA sequence.

Later by pursuing enzymatic activity assays with the aid of a UV/Vis spectrophotometer and MALDI-TOF-MS, we could confidently conclude that *mmgB* gene encodes for 3-HydroxyButyryl-CoA-Dehydrogenase, which catalyzes the oxidation of the 3-hydroxybutyryl-CoA to acetoacetyl-CoA. The analyses also showed that the optimal pH for the enzyme was 9.8 (Glycine-NaOH buffer) and it is specific for $NADP^+$. Thus the above mentioned analyses conclude the fact that, mmgB protein operates as a $NADP^+$ dependent 3-HydroxyButyryl-CoA-Dehydrogenase.

Coming to the future work: the first and foremost on the priority list is physically characterizing the mmgB protein i.e., to determine its native molecular weight and subunit composition. In the near future, we would want to pursue experiments for characterizing the stereospecificity and the acyl chain length specificity including methyl branching tolerance of the enzyme.

CHAPTER VI

GENETIC EVIDENCE FOR THE METHYLCITRIC ACID CYCLE IN *B.SUBTILIS*

VI.A. Hypothesis/Goals

Bacillus subtilis is a model organism for studying sporulation. Sporulation is controlled by several genes and operons in association with various DNA binding proteins and RNA polymerase sigma (σ) factors[3]. The *mmg* operon is one such operon under the control of σ^E factor. The *mmg* operon becomes metabolically active during intermediary stages of sporulation. There are a total of 6 open reading frames (ORF's) in the *mmg* operon – *mmgA, B, C, D, E* and *yqiQ*[10]. The downstream 3 ORF's of the *mmg* operon – *mmgD, mmgE* and *yqiQ* are proposed to encode for the methylcitric acid cycle. Also, as a part of Master's thesis (Rejwi Acharya), it has been shown by the Reddick lab that *mmgD* encodes for a citrate/methylcitrate synthase with a substrate preference for propionyl-CoA over acetyl-CoA[15]. This fact encouraged us even more with reference to involvement of *mmg* operon in the methylcitric acid cycle, a pathway for propionate metabolism. For more details of this pathway, see Chapter ID.

Our goal was to show physiological evidence for this methylcitric acid cycle. To do this, we sought to create conditional knockout mutants of the *mmg* operon (Figures 40, 41, 42 and 43) and study the growth characteristics of the organism by feeding studies using

propionate as the sole carbon source. To be more specific, we wanted to create 2 sets of mutants for these studies:

1) We wanted to place the *mmgD, mmgE* and *yqiQ* genes under an IPTG inducible *Pspac* promoter. This comprises the first set of mutants (Figures 40 and 41).

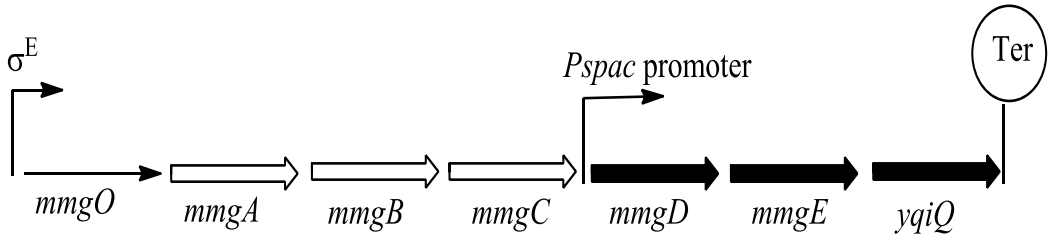


Figure 40: *mmgD* mutant design.

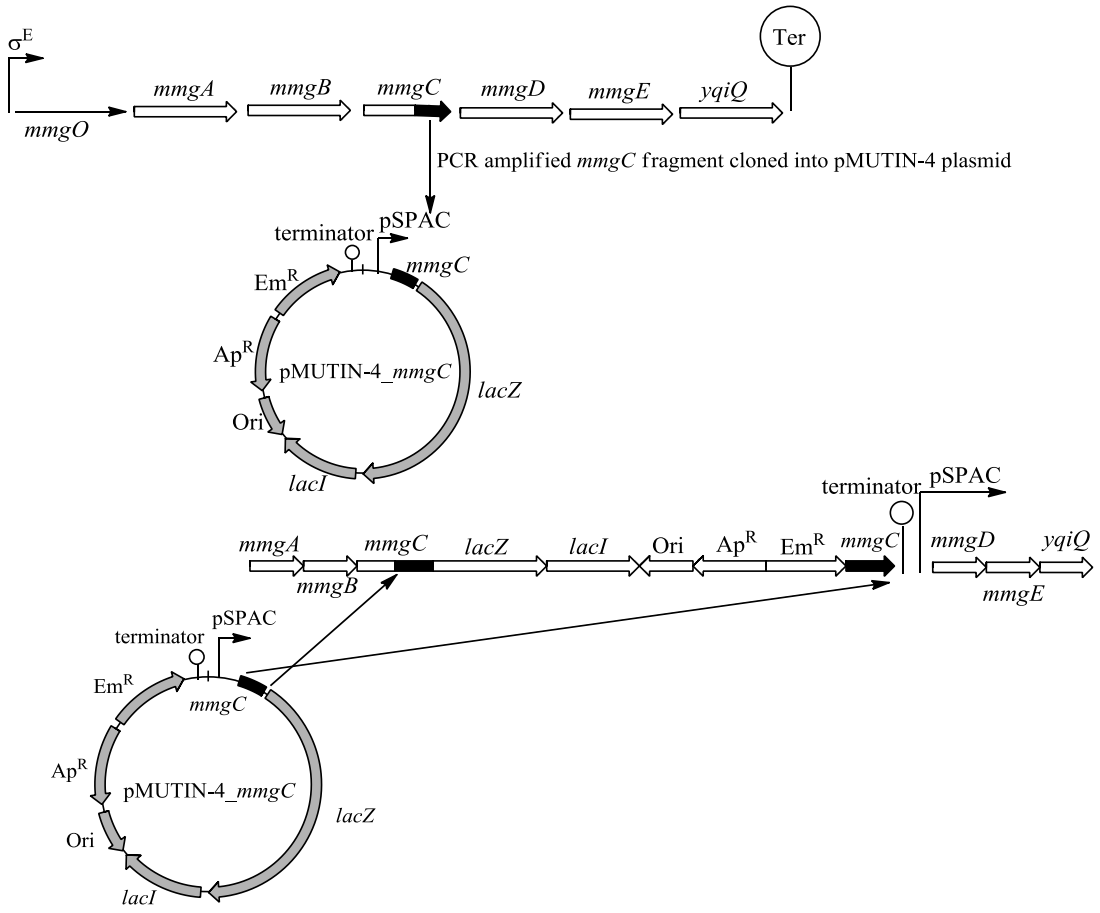


Figure 41: Cloning of *mmgC* insert into pMUTIN-4 and the integration of cloned pMUTIN-4 owing to a homologous recombination event.

2) Whereas with the second set of mutants we wanted to place the whole *mmg* operon (*mmgA, B, C, D, E* and *yqiQ* genes) under the control of IPTG inducible *Pspac* promoter (Figures 42 and 43).

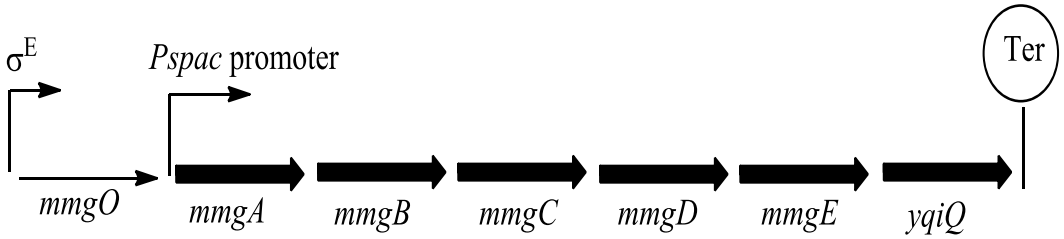


Figure 42: *mmgO* mutant design.

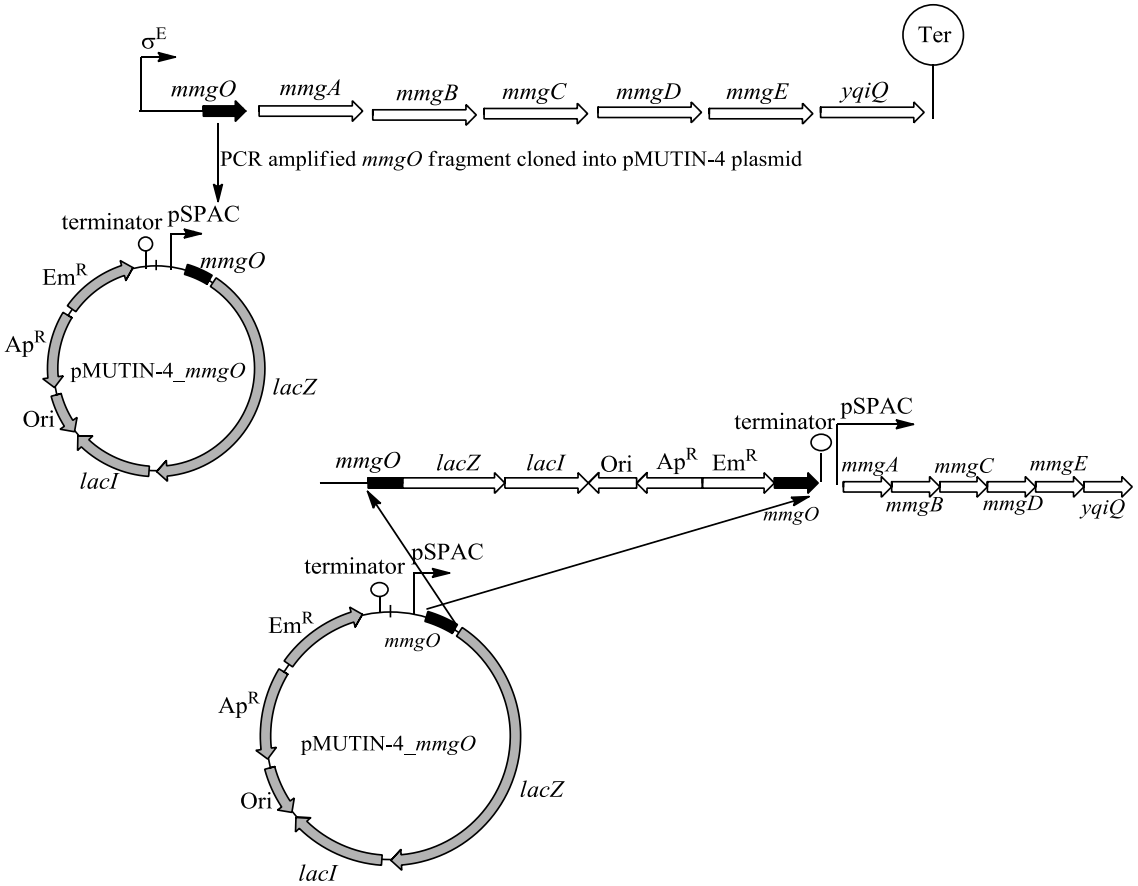


Figure 43: Cloning of *mmgO* insert into pMUTIN-4 and the integration of cloned pMUTIN-4 owing to a homologous recombination event.

The whole idea for creating conditional knock-out mutants (Figures 41 and 43) owing to the insertional mutagenesis is to show that the genes *mmgD, mmgE* and *yqiQ* are responsible for operating the methylcitric acid cycle in *B. subtilis* as a means of propionate metabolism. This fact can be proved by characterizing the enzymatic activities of each of the corresponding genes at the biochemical level. However, to definitively assign the physiological roles of the genes, it is of utmost importance to show the physiological significance of the same genes on the physiology of the associated cell/organism.

Very often, in order to assign physiological roles of the uncharacterized genes, knock-out mutants of the desired genes are created and the result is analyzed. This is a very general and straight-forward method that has been practiced for a very long time. But, this same strategy cannot be applied to our hypothesis about *mmg* operon, because members of the *mmg* operon are normally not turned on until late stages of the growth in the cultures and hence the regular knock-out mutation strategy won't work. So, by creating a conditional knock-out mutant *via* insertional mutagenesis (Figures 41 and 43) that has the members of the *mmg* operon placed under the control of an inducible (IPTG) promoter (*Pspsac*) it will allow us to explore the physiology of these genes (*mmgD, mmgE* and *yqiQ*) without the delays such as waiting till the late growth stages.

The process of insertional mutagenesis to create conditional knock-out mutants of a particular gene or an operon can be performed with the aid of the specially designed plasmids. pMUTIN-4 is one of those plasmids (Figure 44). Hence we used pMUTIN-4.

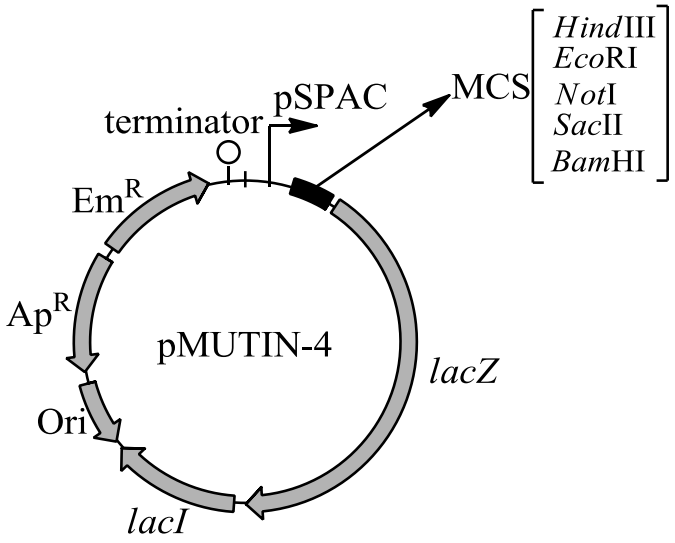


Figure 44: Map of pMUTIN-4[2] (adapted).

pMUTIN-4:

- The intention was to use this for the insertional mutagenesis. Since, pMUTIN-4 is a plasmid created to systematically study uncharacterized genes in *B.subtilis* via insertional mutagenesis (Figure 44)[2].
- pMUTIN-4 has ColE1, an origin of replication specific for *E.coli* with selection for ampicillin resistance gene but lacks an origin of replication for *B.subtilis*. However pMUTIN-4 bears an erythromycin resistance gene that would provide selection if the plasmid integrates into the chromosomal DNA of *B.subtilis*. Hence the only way for *B.subtilis* to benefit from the antibiotic resistance gene in pMUTIN-4 is to integrate the plasmid into its genome by homologous recombination (Figures 41, 43 and 45)[2].

- pMUTIN-4 bears a Multiple Cloning Site (MCS). Also contains an IPTG inducible pSPAC promoter which functions in *B.subtilis* and a *lacI* gene modified to be constitutively express in *B.subtilis* so that pSPAC promoter is turned off when there is no IPTG. pMUTIN-4, it also contains *lacZ* reporter gene and the transcriptional terminators to block the read through transcription (Figure 44)[2].

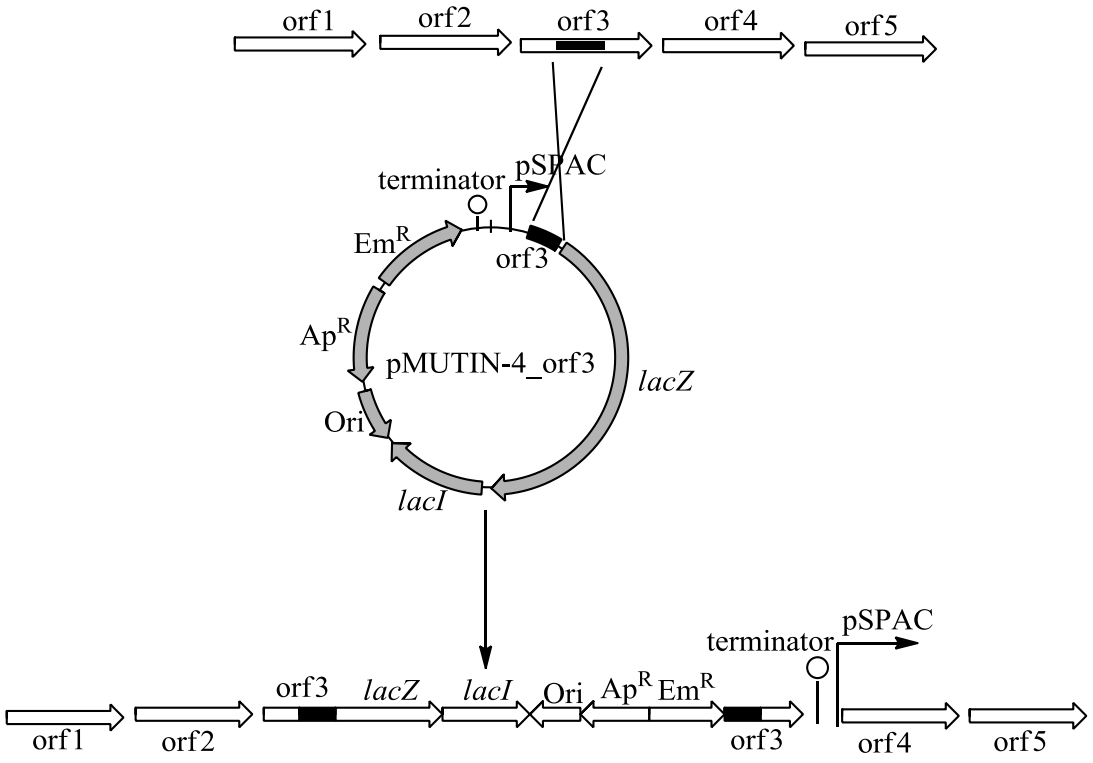


Figure 45: Homologous recombination event[2] (adapted).

Later, after successfully accomplishing the transformation we sought to perform feeding studies of the mutants by feeding them with propionate as sole carbon source and monitoring their growth characteristics. Eventually, we would also want to do some

NMR analysis of the cultures of the mutants for the intermediates of methylcitric acid cycle/propionate metabolism.

VI.B. Experimental section

VI.B.1. PCR-amplification of *mmgC*/*mmgO* inserts

Both forward and reverse primers for *mmgC*/*mmgO* inserts were designed individually in order to amplify the *mmgC*/*mmgO* inserts. The *mmgC* primers were designed in the *mmgC* gene region, ~ 350 bp's upstream of the *mmgD* gene, since, our intention was to keep *mmgD, mmgE* and *yqiQ* genes under the control of an IPTG inducible *Pspac* promoter. However the primers for *mmgO* were designed in the *mmgO* region itself since we wanted to place the complete *mmg* operon under the control of an inducible promoter.

Primers were designed using the software, Bio-Edit from North Carolina State University and were obtained from Integrated DNA technologies[20]. The primers were designed such that the restriction sites were included within them. The forward primer (Up primers) for both *mmgC* and *mmgO* inserts had *HindIII* restriction sites. The reverse primer (Dn primers) for both *mmgC* and *mmgO* inserts had a *SacII* restriction site. For more details, see the table below (Table 6). Genomic DNA which serves as template for amplification of target gene fragments is isolated from wild type *B.subtilis* 168 (BS 168) cells using the Wizard® Genomic DNA Purification Kit (from Promega, Cat # A1120) following manufacturer instructions.

Table 6: PCR primers for *mmgO*/*mmgC* inserts

Primer	Sequence	T_m (^oC)
mmg*O*_forward	5'ACTGAAGCTTATGAGGAACGGTTTGA 3'	60.8
mmg*O*_reverse	5'AGACCGCGGTTACAATTTCTAGTTGT 3'	60.1
mmg*C*_forward	5'GAAAGAAAGCTTGGGCTTTATGGCTC 3'	60.2
mmg*C*_reverse	5'CAGTTTAGGCCGCGGTTATGAAGG 3'	60.0

The *mmgC* and *mmgO* inserts were amplified by a PCR (Polymerase Chain Reaction). The components of the PCR included 25 µL PMM (Phusion master mix), 22 µL Molecular Biology Grade Water, 2 µL Genomic (template) DNA, 1.25 µL Forward primer, 1.25 µL Reverse primer. PMM is a stock mixture of High Fidelity Phusion Polymerase + Buffer + dNTP's + counter ions ($MgCl_2$) stored at -20^{o}C (from Finzymes/*New England Biolabs*, Cat # F553S).

PCR amplification of the desired *mmgC* and *mmgO* inserts was carried out using a thermocycler. Conditions maintained in thermocycler in terms of temperature and time periods as shown in Table 7.

Table 7: Conditions maintained in thermocycler

Parameters	***mmgO***		***mmgC***		
Initial incubation	98^O C	1:00 min	98^O C	1:00 min	
Denaturation	98^O C	0.05 min	98^O C	0.05 min	**25 cycles**
Annealing	63^O C	1:00 min	63^O C	0.30 min	
Extension	72^O C	1:00 min	72^O C	1:00 min	

After 25 cycles, an extension cycle for 10 minutes at 72^{o}C was carried out and finally the samples were held at 4^{o}C.

The PCR amplified *mmgC* and *mmgO* inserts were purified by a QIAquick PCR-purification kit (Cat # 28106) from Qiagen following the manufacturer's specifications[22]. The purified, PCR amplified *mmgC* and *mmgO* inserts were then analyzed by 1% agarose gel electrophoresis.

VI.B.2. Cloning of the *mmgC*/*mmgO* inserts into the pMUTIN-4 plasmid

After validation of the amplification of *mmgC*/*mmgO* inserts following agarose gel electrophoretic analysis, the PCR amplified *mmgC*/*mmgO* inserts were subjected to restriction digestion using HindIII (from *New England Biolabs,* Cat # R0104S) and SacII (from *New England Biolabs*, Cat # R0157S) restriction enzymes. Simultaneously, pMUTIN-4 was also subjected to restriction digestion using the same set of restriction enzymes.

The eppendorf tubes containing the pMUTIN-4, *mmgC* and *mmgO* inserts were incubated for 3.5 hours at 37^{o}C followed by heat inactivation of the restriction enzymes at 65^{o}C for 20 minutes. The two restriction enzymes used were *HindIII* and *SacII*. Each of the three eppendorf tubes contained 2.5 µL New England Biolabs buffer 2, 0.3 µL of BSA (Bovine Serum Albumin) stock from New England Biolabs, 1.0 µL each of both the restriction enzymes. In addition, 6.0 µL of pMUTIN-4 was added to tube-1, 10.0 µL of *mmgC* insert and 10.0 µL of *mmgO* insert were added to the tubes 2 and 3 respectively. Lastly, molecular biology grade water was added to all the three tubes to a final reaction volume of 25 µL.

All the three restriction digest samples were analyzed by a 1% agarose gel electrophoresis. The restriction digested *mmgC/mmgO* insert and the pMUTIN-4 plasmid

bands were cut from the 1% agarose gel and the gel slices were placed in separate eppendorf tubes. The *mmgC/mmgO* insert and pMUTIN-4 plasmid DNA's were recovered from the respective gel slices using QIAquick Gel Extraction Kit (Qiagen, Cat # 28706) following the manufacturer's instructions.

This was followed up by a ligation procedure which involves ligation of the *mmgC* insert (PCR product) with the vector (pMUTIN-4) and also the ligation of *mmgO* insert (PCR product) with the vector (pMUTIN-4), individually. The ligation was achieved using the T4 DNA ligase and ligase buffer from New England Biolabs. Two eppendorf tubes were taken and filled with 1 μL T4 DNA ligase, 2 μL ligase buffer and 11 μL molecular biology grade water. To tube-1, restriction digested *mmgC* insert and pMUTIN-4 plasmid were added in 5:1, whereas to tube-2 restriction digested *mmgO* insert and pMUTIN-4 plasmid were added in 5:1. Finally, the tubes with the ligation samples were placed in a thermocycler under the following conditions: for 30 minutes at 37^{o}C, 2 hours at 23^{o}C and the samples were held at 16^{o}C for overnight.

VI.B.3. Transformation and Re-transformation of *E.coli* with cloned pMUTIN-4 plasmids

Following cloning of the inserts (both *mmgC/mmgO*) into the pMUTIN-4 plasmid (vector) separately, the cloned plasmids were individually used to transform separate batches of NEB-5α strain of competent *E.coli* cells for propagation purposes.

NEB-5α strain of competent *E.coli* cells (Cat # C2992H) were obtained from *New England Biolabs*. Transformation of *E.coli* with cloned pMUTIN-4 vectors was carried out following the manufacturer's instructions[28]. Following transformation, the cells were

spread on LB (Luria-Bertani) – agar plates with 50µg/mL Ampicillin and the transformants were grown in liquid culture supplemented with Ampicillin.

Glycerol (10% v/v) stocks of the antibiotic resistant colonies were made and stored at -80°C for further proceedings. Later, overnight liquid cultures of the transformed NEB-5α *E.coli* cells were made by inoculating the antibiotic resistant colonies/cells in 5 mL LB-broth with 50µg/mL Ampicillin and grown anaerobically at 37°C /220 rpm. The following morning, cloned pMUTIN-4 vectors (pMUTIN-4 with *mmgC* insert and pMUTIN-4 with *mmgO* insert) were isolated from these liquid cultures by employing the QIAprep miniprep kit (Cat # 27106) from Qiagen, following the manufacturer's instructions[23].

A PCR was performed using cloned pMUTIN-4 vectors (pMUTIN-4 with *mmgD* insert and pMUTIN-4 with *mmgO* insert) as template DNA and a 1% agarose gel was run to analyze for the presence of inserts. The isolated, cloned pMUTIN-4 vectors were stored at -20°C for future uses.

Our next target in line was to (re) transform another, fresh batch of NEB-5α cells. The transformation was performed as it was done in previous step. The purpose for retransformation was to get monoclonal transformants for sequencing. The antibiotic selective transformants were picked up and glycerol stocks are made similar to how it was done in previous step[28].

Again, cloned pMUTIN-4 vectors (pMUTIN-4 with *mmgD* insert and pMUTIN-4 with *mmgO* insert) were isolated, but this time, it was from new batch of NEB-5α transformants. The isolation procedure involves making overnight liquid cultures of the

transformed NEB-5α (new batch) *E.coli* cells by inoculating the antibiotic resistant colonies/cells in 5 mL LB-broth with 50µg/mL Ampicillin and growing anaerobically at 37°C /220 rpm. The following morning, cloned pMUTIN-4 vectors (pMUTIN-4 with *mmgD* insert and pMUTIN-4 with *mmgO* insert) were isolated from these liquid cultures by employing the QIAprep miniprep kit (Cat # 27106) from Qiagen, following manufacturer instructions[23]. The miniprep solution from NEB-5α (new batch) cells – cloned pMUTIN-4 vectors (pMUTIN-4 with *mmgD* insert and pMUTIN-4 with *mmgO* insert) were sent to SeqWright for getting the sequence information.

VI.B.4. Transformation of *B.subtilis* 168 with cloned pMUTIN-4 plasmid

Our next target was to transform different batches of *B.subtilis* 168 cells with both the cloned pMUTIN-4 plasmids separately. Initially we tried Spizizen natural competence/transformation procedures[29]. Competent *B.subtilis* 168 cells were prepared following Spizizen competence induction protocol using minimal media (Spizizen salts). Transformation of *B.subtilis* 168 with cloned pMUTIN-4 vectors was carried out following the instructions given in Spizizen procedure. Following the transformation, the cells were spread on glucose + minimal salts media – agar plates with 0.3µg/mL Erythromycin and the transformants were picked up in antibiotic selective fashion[29].

Later we moved on to pursuing various two-step transformation protocols for *B.subtilis* transformation. The transformation of *B.subtilis* 168 with cloned pMUTIN-4 vectors was carried out following the instructions given in two-step transformation protocols for *B.subtilis* transformation. Following transformation, the cells were spread

on LB (Luria-Bertani) – agar plates with 0.3μg/mL Erythromycin and the transformants were grown in antibiotic selective media[30, 31].

VI.C. Results/Discussion

VI.C.1. PCR-amplification of *mmgC*/*mmgO* inserts

The *mmgC* and *mmgO* sections were each amplified from *Bacillus subtilis* 168 genome by PCR (Polymerase Chain Reaction). In order to test for the success of PCR, the amplified products were first purified by a QIAquick PCR purification step (Qiagen, Cat # 28106)[22]. The purified, PCR products were then analyzed by 1% agarose gel electrophoresis, as shown in Figure 46.

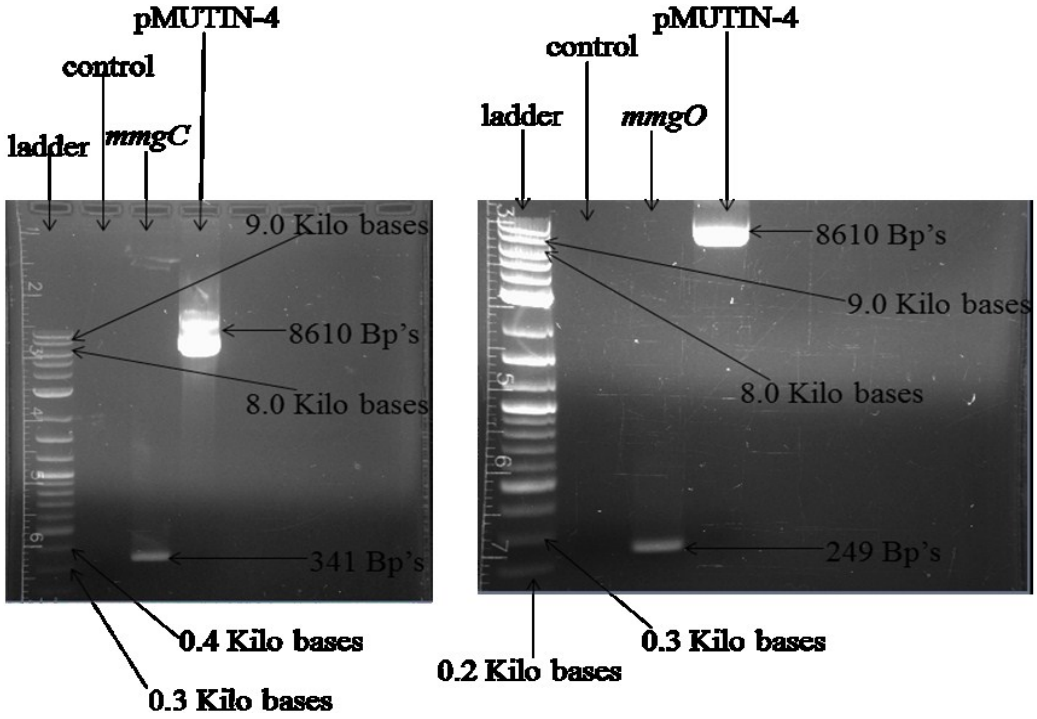


Figure 46: 1% agarose gels describing the successful amplification of both *mmgC*/*mmgO* inserts.

Each of the agarose gel pictures (Figure 46) shows 4 lanes labeled as Ladder, Control, *mmgC* (left gel) or *mmgO* (right gel) and pMUTIN-4. The lane ladder corresponds to 2-log ladder from *New England Biolabs* (Cat # N3200S) and serves the purpose of a molecular marker. The bands in the ladder were used as standards in order to characterize the bands of unknown sizes (number of base pairs) seen on the gel.

The control lane corresponds to a control reaction that includes all the PCR components (PMM, primers and H_2O) except for template DNA. So, the *mmgC*/*mmgO* inserts should not amplify as there is no template to direct their amplification. This is in good agreement, as there were no bands visible on either gel in those lanes. These show that the PCR is dependent on template and that there is no template contaminating the PCR components.

The lane *mmgC* in the left gel corresponds to the PCR amplified product which in this case is amplified *mmgC* insert. This reaction contains all the PCR components including the template DNA (genomic DNA from *Bacillus subtilis* 168) which directs the amplification of *mmgC* insert. Hence there should be a band seen on the gel in lane *mmgC* corresponding to the size of *mmgC* insert in terms of number of base pairs (bp's). This lane indeed have a band which lies at a position in between 0.3 – 0.4 kilo bases (kb) in comparison to molecular marker seen in the lane L. This is in good agreement with the expected size of the *mmgC* insert which is 341 bp's. So, the result is suggestive that *mmgC* insert was successfully amplified by a PCR.

The lane *mmgO* in the right gel (Figure 46) corresponds to the PCR amplified product of the *mmgO* insert. This reaction contains all the PCR components including the

template DNA (genomic DNA from *Bacillus subtilis* 168) which directs the amplification of *mmgO* insert. Hence there should be a band seen on the gel in lane P corresponding to the size of *mmgO* insert in terms of number of base pairs (bp's). There is indeed a band in the lane *mmgO* which lies at a position in between 0.2 – 0.3 kilo bases (kb) in comparison to molecular marker seen in the lane L. This is in good agreement with the expected size of *mmgO* insert which is 249 bp's. So, the result is suggestive that the *mmgO* insert was successfully amplified by a PCR.

The lanes pMUTIN-4 from both the gels (left & right) (Figure 46) corresponds to the pure plasmid after miniprep. No PCR was performed for these samples. In these gel lanes, there are indeed bands which lie at a position in between 8.0 – 9.0 kilo bases (kb) in comparison to molecular marker seen in the lane L. Since pMUTIN-4 consists of 8610 bp_s, the plasmid with which we are working is correct.

VI.C.2. Cloning of amplified *mmgC*/*mmgO* inserts into pMUTIN-4 plasmid

After confirming the successful PCR amplification of the *mmgC* and *mmgO* inserts; they were cloned into the pMUTIN-4 vector. The map for pMUTIN-4 plasmid (Figure 47) below shows the multiple cloning site of the pMUTIN-4 vector, where the *mmgC/mmgO* inserts are to be cloned.

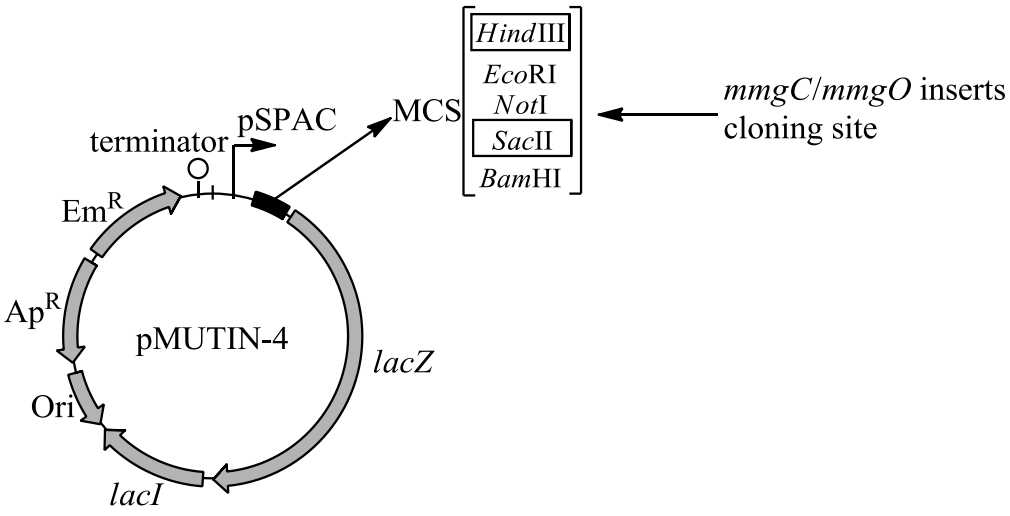


Figure 47: Cloning of *mmgC*/*mmgO* inserts into pMUTIN-4 plasmid[2] (adapted).

VI.C.3. Transformation and Retransformation of *E.coli* with cloned pMUTIN-4 plasmids

Following the completion of cloning, NEB-5α *E.coli* cells were transformed with the cloned pMUTIN-4 vector (pMUTIN-4_*mmgC* and pMUTIN-4_ *mmgO*). Later, the transformed NEB-5α *E.coli* cells were spread on LB – agar plates with 50μg/mL Ampicillin to grow[28]. Due to the presence of antibiotic, only transformants/clones with pMUTIN-4 vector inside them will survive, while others simply die because of the antibiotics in the media. This antibiotic resistance property of the clones owe to the Ampicillin resistance gene present in the plasmid. Later, a PCR was performed to screen the isolated pMUTIN-4 vectors (pMUTIN-4_*mmgC* and pMUTIN-4_*mmgO*) for *mmgC*/*mmgO* inserts.

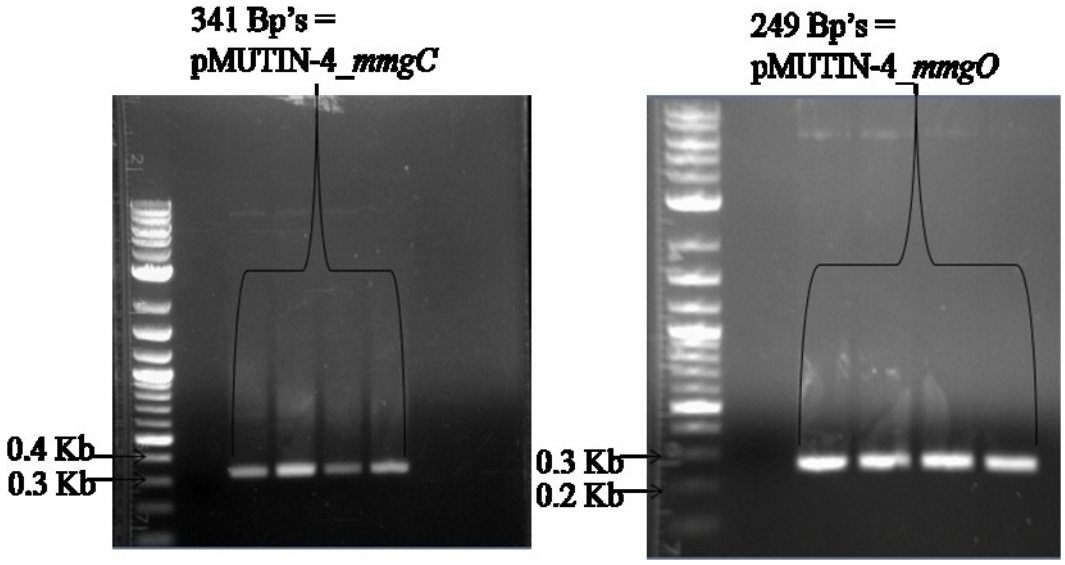


Figure 48: 1% Agarose gel analysis after cloning of the *mmgC*/*mmgO* inserts into the pMUTIN-4 plasmid. The PCR's were performed using the cloned pMUTIN-4 plasmids (pMUTIN-4 _*mmgC*, left gel & pMUTIN-4_*mmgO*, right gel) as template DNA.

Figure 48 shows the gel picture taken of the PCR screening reactions. The agarose gel pictures (Figure 48) contain lanes labeled as 2-log ladder from *New England Biolabs* (Cat # N3200S). This corresponds to the molecular marker. The bands seen in the lane L are considered as standards for characterizing the bands of unknown sizes (number of base pairs) present on the gel.

The agarose gel picture on the left shows four lanes with one band in between 0.3 – 0.4 bp's. All the single bands in those lanes correspond to the amplified *mmgC* insert from the PCR using cloned pMUTIN-4 with *mmgC* insert (pMUTIN-4_*mmgC*) in as the template DNA. The bands are seen at ~350 bp's which corresponds to the size of the *mmgC* insert, which is expected to be 341 bp's. This confirms the cloning of *mmgC* insert into all the four pMUTIN-4 plasmids that were isolated. All these four lanes correspond

to PCR reactions with four different pMUTIN-4_*mmgC* plasmids from different NEB-5α transformants.

The agarose gel picture on the right shows four lanes with one band in between 0.2 – 0.3 bp's. All the single bands in those lanes correspond to the amplified *mmgO* insert from the PCR using cloned pMUTIN-4 with *mmgO* insert (pMUTIN-4_*mmgO*) in as the template DNA. The bands are seen at ~250 bp's which corresponding to the size of *mmgO* insert. This is in good agreement with the expected size of *mmgO* insert which is 249 bp's. This confirms the cloning of *mmgO* insert into each of the four pMUTIN-4 plasmids that were isolated. All these four lanes correspond to PCR reactions with four different pMUTIN-4_*mmgO* plasmids from different NEB-5α transformants.

The bands on the gels (both left & right) were suggestive of the successful cloning of the *mmgC/mmgO* inserts into the pMUTIN-4 plasmid individually (pMUTIN-4_*mmgC* and pMUTIN-4_*mmgO*). The positions of the bands from both the gels were in good agreement with the expected sizes of the *mmgC/mmgO* inserts which were 341/249 bp's respectively.

Later, a new batch of NEB-5α *E.coli* cells was transformed using cloned pMUTIN-4 plasmids (pMUTIN-4_*mmgC* and pMUTIN-4_*mmgO*). The purpose for retransformation was to ensure we obtained monoclonal plasmids for sequencing. The transformed NEB-5α *E.coli* cells were spread on LB – agar plates with 50μg/mL Kanamycin to grow[28]. Later, cloned pMUTIN-4 plasmids were isolated from antibiotic selective transformants using a QIAprep miniprep kit from Qiagen[23]. Following this, a PCR was performed using cloned pMUTIN-4 plasmids (pMUTIN-4_*mmgC* and

pMUTIN-4_ *mmgO*) as template DNA. Both plasmids were sequenced (data not shown), the sequences show that the desired constructs were successfully made.

VI.C.4. Transformation of *B.subtilis* 168 with cloned pMUTIN-4 plasmid

We later moved to transformations of different batches of *B.subtilis* 168 cells with both the cloned pMUTIN-4 plasmids (pMUTIN-4_*mmgC* and pMUTIN-4_*mmgO*) separately. With the Spizizen natural competence/transformation procedures, we got some colonies/cells-transformants in antibiotic selective manner owing to the presence of the erythromycin-resistance gene present in the pMUTIN-4 plasmid[29]. Later, a PCR was used to screen for the mutation (Figures 41 and 43) using the genomic DNA isolated from the transformants (mutants) as template DNA. However, the agarose gel pictures (Figure 49) below do not show conclusive results with respect to successful transformation of *B.subtilis* 168 and the integration of cloned pMUTIN-4 vectors (pMUTIN-4_*mmgC* and pMUTIN-4_*mmgO*) in to the *B.subtilis* 168 chromosomal DNA.

Table 8: Details of the primers for cloning and screening

Primer set	**Screening for cloning** (size of the DNA fragment)	**Screening for insertional mutation** (size of the DNA fragment)
mmgC – Up/Dn (in the tail end of *mmgC* gene)	Yes (341 bp's)	Yes (8610 bp's + 341 bp's)
mmgO - Up/Dn	Yes (249 bp's)	Yes (8610 bp's + 249 bps')
mmgC – Up/Dn (for the entire *mmgC* gene)	No (1134 bp's)	Yes (8610 bp's + 1134 bp's)
mmgO-Up / *yqiQ*-Dn	No (6627 bp's)	Yes (8610 bp's + 6627 bp's)
mmgO-Up / *mmgA* –Dn	No (1428 bp's)	Yes (8610 bp's + 1428 bp's)
mmgD – Up/Dn (for the entire *mmgD* gene)	No (1116 bp's)	Yes (8610 bp's + 1116 bp's)
mmgC –up / *mmgD* – Dn	No (1457)	Yes (8610 bp's + 1147 bp's)
lacI – Up / *mmgA* - Dn	No (PCR product will not be obtained)	Yes (xxxx bp's)
lacI – Up / *mmgC* -Dn	No (PCR product will not be obtained)	Yes (xxxx bp's)

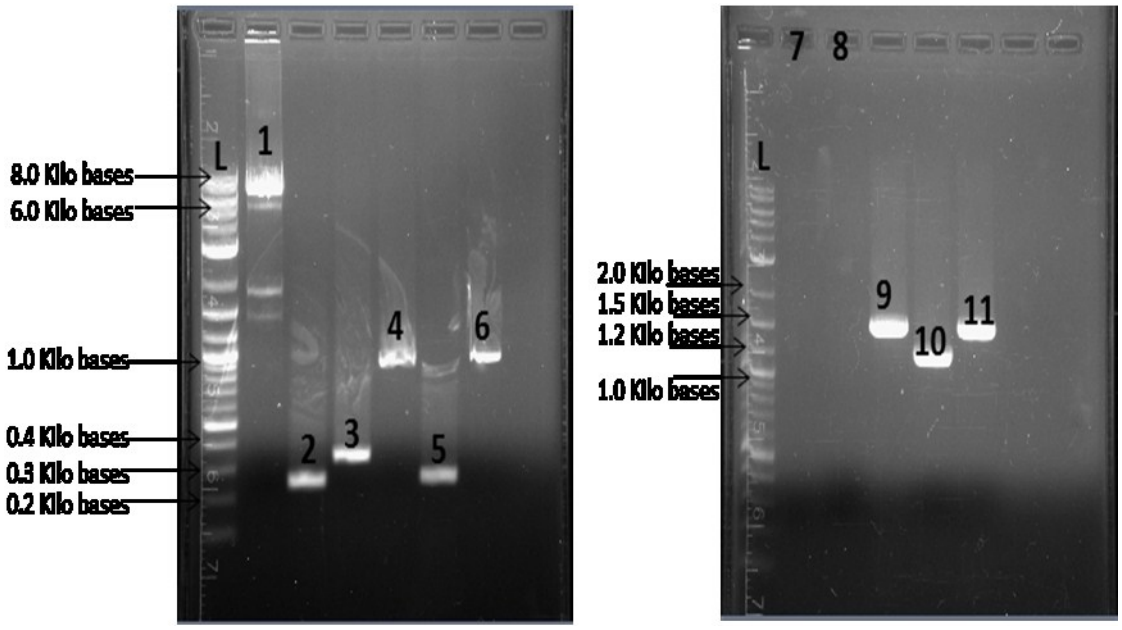


Figure 49: 1% Agarose gels for PCR screening analysis of both *mmgC/mmgO* mutations with Spizizen procedures. For more details see the explanation below.

The agarose gel pictures (Figure 49) contain lanes labeled as L. This corresponds to 2-log ladder, the molecular marker. The bands seen in the lanes L are considered as standards for characterizing the bands of unknown sizes (number of base pairs) present on the gel.

The agarose gel picture on the left shows lanes labeled as 1, 2, 3, 4, 5, and 6 with a single band in each of the lanes. The single band in the lane 1 corresponds to the amplified *mmg* operon (*mmgA, B, C, D, E and yqiQ* genes) after PCR amplification using the primer set: *mmgO*-up and *yqiQ*-dn with BS 168 genomic DNA as template for PCR amplification. The lane 1 shows a band in between 6.0 – 8.0 kb which is consistent with the known length of the *mmg* operon in terms of number of base pairs, 6627 bp's. This can be used as a reference with respect to the size of the *mmg* operon from mutants after PCR amplification (Figure 43).

The single band in the lane 2 corresponds to the amplified *mmgO* insert after PCR amplification using the primer set: *mmgO*-up and *mmgO*-dn with BS 168 genomic DNA as template for PCR amplification. The lane 2 shows a band in between 0.2 – 0.3 kb which is consistent with the theoretical length of the *mmgO* insert in terms of number of base pairs, 249 bp's. This (*mmgO* amplification from BS168 genomic DNA) can be used to compare to the size of the *mmgO* from mutant (*mmgO* amplification from *mmgO* mutant genomic DNA, after integration of pMUTIN-4_*mmgO*) after PCR amplification (Figure 43).

The single band in the lane 3 corresponds to the amplified *mmgC* insert after PCR amplification using the primer set: *mmgC*-up and *mmgC*-dn with BS 168 genomic DNA as template for PCR amplification. The lane 3 shows a band in between 0.3 – 0.4 kb which is consistent with the theoretical length of the *mmgC* insert in terms of number of base pairs, 341 bp's and is in good agreement with the previous gels from screening after PCR amplification. This (*mmgC* amplification from BS168 genomic DNA) can be used to compare to the size of the *mmgC* from mutant (*mmgC* amplification from *mmgC* mutant genomic DNA, after integration of pMUTIN-4_*mmgC*) after PCR amplification (Figure 41).

The single band in the lane 4 corresponds to the amplified *lacI* gene (present in pMUTIN-4) after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with DNA isolated from the *mmgO* mutant (pMUTIN-4_*mmgO* integrated into BS168 chromosome = pMUTIN-4_*mmgO*_BS168) as template for PCR amplification. The lane 4 shows a band at about 1000 bp's which is consistent with the theoretical length of the *lacI* gene in

terms of number of base pairs, 1080 bp's. This is suggestive that the plasmid integrated into the BS 168 genome, since, *B.subtilis* doesn't have the *lacI* gene. But at this point we were not sure of location of the plasmid (pMUTIN-4_*mmgO*) integration into BS 168 chromosomal DNA (Figures 42 and 43).

The single band in the lane 5 corresponds to the amplified *mmgO* insert (cloned into pMUTIN-4) after PCR amplification using the primer set: *mmgO*-up and *mmgO*-dn with DNA isolated from the *mmgC* mutant (pMUTIN-4_*mmgC* integrated into BS168 chromosome = pMUTIN-4_*mmgC*_BS168) as template for PCR amplification. The lane 5 shows a band in between 0.2 – 0.3 kb which is consistent with the theoretical length of the *mmgO* insert in terms of number of base pairs, 249 bp's and is in good agreement (in terms of number of bp's) with the band from lane 2 on the same gel, after PCR amplification. This result is conclusive that the *mmgO* region of the *mmg* operon is not mutated with pMUTIN-4_*mmgC*. This is a good sign, also, insertional mutagenesis due to the cloned plasmid should be site specific (Figures 41 and 43). However the site of mutation is not yet revealed.

The single band in the lane 6 corresponds to the amplified *lacI* gene (present in pMUTIN-4) after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with *mmgC* mutant (pMUTIN-4_*mmgC*_BS168) genomic DNA as template for PCR amplification. The lane 6 shows a band at about 1000 bp's which is consistent with the theoretical length of the *lacI* gene in terms of number of base pairs, 1080 bp's and is in good agreement (in terms of number of bp's) with the band from lane 4 on the same gel, after PCR amplification. This is suggestive that the plasmid is integrated into the BS 168

genome, but does not show the location of the integration. The following PCR reactions were designed to reveal this location (Figures 41 and 43).

The agarose gel picture (Figure 49) on the right shows lanes labeled as 7, 8, 9, 10, and 11with a single band each in the lanes 9, 10 & 11, but no bands are seen in the lanes 7, & 8. The lane 7 corresponds to PCR pertaining to the amplification of *lacI*-up/*mmgA*-dn insert using the primer set: *lacI*-up and *mmgA*-dn with *mmgO* mutant (pMUTIN-4_*mmgO*_BS168) genomic DNA as template for PCR amplification. The lane 7 shows no band at all. This is suggestive that the plasmid did not integrate into the BS 168 chromosomal DNA in the *mmg* operon region and the expected *mmgO* conditional mutant (Figure 43) was not created.

The lane 8 corresponds to PCR pertaining to the amplification of *lacI*-up/*mmgD*-dn insert using the primer set: *lacI*-up and *mmgD*-dn with *mmgC* mutant (pMUTIN-4_*mmgC*_BS168) genomic DNA as template for PCR amplification. The lane 8 showed no band at all. This was suggestive that the plasmid did not integrate into BS 168 chromosomal DNA in the *mmg* operon region and the expected *mmgC* conditional mutant (Figure 41) was not created.

The lane 9 corresponds to PCR pertaining to the amplification of *mmgO* -up/ *mmgA* -dn insert using the primer set: *mmgO*-up and *mmgA*-dn with *mmgO* mutant (pMUTIN-4_*mmgO*_BS168) genomic DNA as template for PCR amplification. The lane 9 shows a band in between 1.2 – 1.5 kb which is consistent with the theoretical length of the *mmgA* gene in terms of number of base pairs; 1182 bp's. This is again suggestive that the plasmid did not get integrated into BS 168 chromosomal DNA in *mmg* operon region

and the expected *mmgO* conditional mutant is not created. If mutation has happened, expected size of the *mmgA* gene would have been 1182 (*mmgA*) + 8610 bp's (pMUTIN-4) (Figure 43).

The lane 10 corresponds to PCR pertaining to the amplification of *mmgD*-up/*mmgD*-dn insert using the primer set: *mmgD*-up and *mmgD*-dn with *mmgC* mutant (pMUTIN-4_*mmgC*_BS168) genomic DNA as template for PCR amplification. The lane 10 shows a band in between 1.0 – 1.2 kb which is consistent with the theoretical length of the *mmgD* gene in terms of number of base pairs; 1137 bp's. This result is conclusive that the *mmgD* region of the *mmg* operon is not mutated with pMUTIN-4_*mmgC*. This is a good sign, also, insertional mutagenesis due to the cloned plasmid should be site specific (Figure 41). However the site of mutation is not yet revealed.

The lane 11 corresponds to PCR pertaining to the amplification of *mmgO* -up/ *mmgA*-dn insert using the primer set: *mmgO*-up and *mmgA*-dn with BS 168 (wild type) genomic DNA as template for PCR amplification. The lane 11 shows a band at a position in between 1.2 – 1.5 kb which is consistent with the theoretical length of the *mmgA* gene in terms of number of base pairs; 1182 bp's. Since, this PCR amplified product is from wild type BS 168; it can therefore be used as a reference with respect to the size in comparison to *mmgO* mutant (Figures 42 and 43).

At this juncture we understood that the transformation experiments we pursued were not successful i.e., the wild type *B.subtilis* 168 cells did not integrate the pMUTIN-4 vectors (pMUTIN-4_*mmgC* and pMUTIN-4*mmgO*) into the desired locations of the genome. Since we were not confident in the Spizizen procedure, which was difficult to

interpret and follow, we moved onto pursuing different two-step transformation procedures[30, 31]. Below are shown gel pictures (Figure 50) corresponding to the results after the two-step transformation experiments.

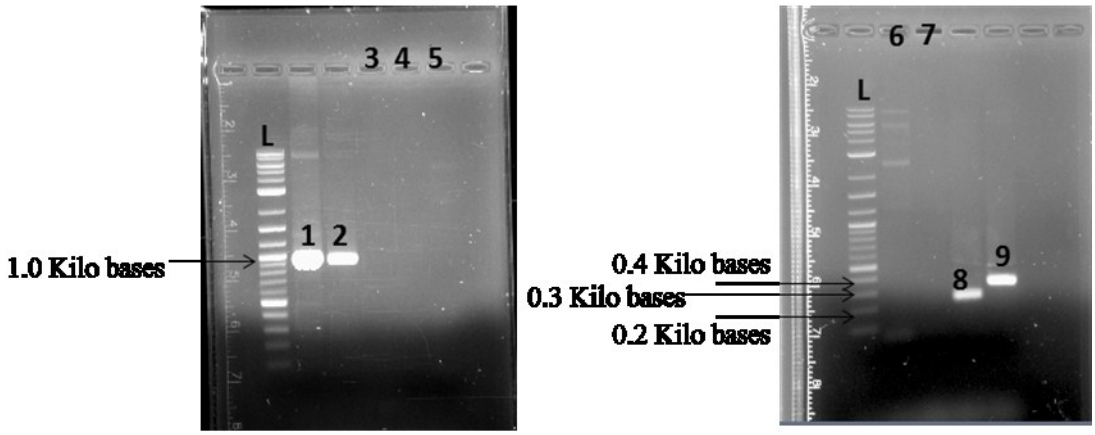


Figure 50: Agarose gels for PCR screening analysis of *mmgC* mutation with two-step transformation procedures. For more details see the explanation below.

The agarose gel pictures (both left & right) shown above (Figure 50) correspond to the PCR amplified products with respect to *mmgC* mutants (pMUTIN-4_*mmgC*_BS168). The agarose gel pictures contain lanes labeled as L. This corresponds to 2-log ladder, the molecular marker. The bands seen in the lanes L are considered as standards for characterizing the bands of unknown sizes (number of base pairs) present on the gel.

The agarose gel picture on the left shows lanes labeled as 1, 2, 3, 4, and 5 with a single band each in the lanes 1 & 2, but no bands are seen in the lanes 3, 4, & 5. The single band in the lane 1 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgO*) DNA as template for PCR amplification. The lane 1 shows a band at about 1.0

kb which is consistent with the theoretical length of the *lacI* gene in terms of number of base pairs, 1080 bp's. This can be used as a reference with respect to integration of the pMUTIN-4_*mmgO* into the BS168 chromosomal DNA (Figures 43).

The single band in the lane 2 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgC* insert) DNA as template for PCR amplification. The lane 2 shows a band at about 1.0 kb which is consistent with the theoretical length of the *lacI* gene in terms of number of base pairs; 1080 bp's. This can also be used as a reference with respect to integration of the pMUTIN-4_*mmgC* into the BS168 chromosomal DNA (Figures 41).

The lane 3 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with BS 168 (wild type) genomic DNA as template for PCR amplification. The lane 3 shows no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of base pairs; 1080 bp's. This can be used as a negative control with respect to the insertional mutation in *mmg* operon. If mutation does occur owing to the integration of the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgC*) into the chromosomal DNA of BS 168, then upon PCR screening, the agarose gel should show a band at exactly the same position (1080 bp's) as it is seen in lanes 1 & 2. But, if mutation does not occur, then no band corresponding to the size of *lacI* gene would be seen on the gel.

The lane 4 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with *mmgC* mutant (pMUTIN-4_*mmgC*_BS168)

genomic DNA as template for PCR amplification. The lane 4 shows no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of base pairs, 1080 bp's. This is suggestive that the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgC*) is not integrated into BS168 chromosomal DNA (Figure 41).

The lane 5 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with another *mmgC* mutant (pMUTIN-4_*mmgC*_BS168) genomic DNA as template for PCR amplification. The lane 5 shows no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of base pairs, 1080 bp's. This is suggestive that the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgC*) is not integrated into BS168 chromosomal DNA (Figure 41).

The agarose gel picture on the right shows lanes labeled as 6, 7, 8, and 9 with a single band each in the lanes 8 & 9, but no bands are seen in the lanes 6 & 7. The lane 6 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with another *mmgC* mutant (pMUTIN-4_*mmgC*_BS168) genomic DNA as template for PCR amplification. The lane 6 shows no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of base pairs, 1080 bp's. Again, this too is suggestive that the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgC*) is not integrated into BS168 chromosomal DNA (Figure 41).

The lane 7 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with another *mmgC* mutant (pMUTIN-4_*mmgC*_BS168) genomic DNA as template for PCR amplification. The lane 7 shows no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of

base pairs 1080 bp's. This too is suggestive that the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgC*) is not integrated into BS168 chromosomal DNA (Figure 41). However, the strains had erythromycin resistance; so, we do not understand why there was no integration.

The single band in the lane 8 corresponds to the amplified *mmgO* insert after PCR amplification using the primer set: *mmgO*-up and *mmgO*-dn with cloned pMUTIN-4 DNA (pMUTIN-4_*mmgO*) as template for PCR amplification. The lane 8 shows a band in between 0.2 – 0.3 kb which is consistent with the theoretical length of the *mmgO* insert in terms of number of base pairs, 249 bp's and is in good agreement with the previous gels from screening after PCR amplification. This is suggestive that the *mmgO* insert cloned into the Pmutin-4 plasmid used for the transformation was still intact with the plasmid (Figure 43).

The single band in the lane 9 corresponds to the amplified *mmgC* insert after PCR amplification using the primer set: *mmgC*-up and *mmgC*-dn with cloned pMUTIN-4 DNA (pMUTIN-4_*mmgD*) as template for PCR amplification. The lane 9 shows a band in between 0.3 – 0.4 kb which is consistent with the theoretical length of the *mmgC* insert in terms of number of base pairs which is 341 bp's and is in good agreement with the previous gels from screening after PCR amplification. This is suggestive that the *mmgC* insert cloned into the Pmutin-4 plasmid used for the transformation was still intact with the plasmid (Figure 41).

The agarose gel pictures (both left & right) shown below (Figure 51) correspond to the PCR amplified products with respect to *mmgO* mutants. The agarose gel pictures

contain lanes labeled as L. This corresponds to 2-log ladder, the molecular marker. The bands seen in the lanes L are considered as standards for characterizing the bands of unknown sizes (number of base pairs) present on the gel.

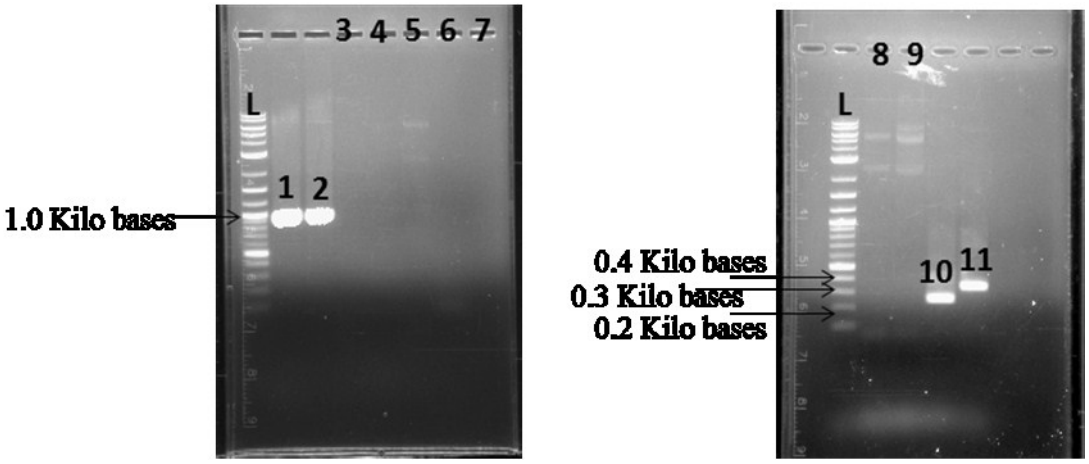


Figure 51: Agarose gels for PCR screening analysis of *mmgO* mutation with two-step transformation procedures. For more details see the explanation below.

The agarose gel picture on the left shows lanes labeled as 1, 2, 3, 4, 5, 6 and 7 with a single band each in the lanes 1 & 2, but no bands are seen in the lanes 3, 4, 5, 6 & 7. The single band in the lane 1 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgO*) DNA as template for PCR amplification. The lane 1 shows a band at about 1.0 kb which is consistent with the theoretical length of the *lacI* gene in terms of number of base pairs 1080 bp's. This can be used as a reference with respect to integration of the pMUTIN-4_*mmgO* into the BS168 chromosomal DNA (Figures 43).

The single band in the lane 2 corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with cloned pMUTIN-4 plasmid

(pMUTIN-4_*mmgC*) DNA as template for PCR amplification. The lane 2 shows a band at about 1.0 kb which is consistent with the theoretical length of the *lacI* gene in terms of number of base pairs 1080 bp's. This can also be used as a reference with respect to integration of the pMUTIN-4_*mmgC* into the BS168 chromosomal DNA (Figure 41).

The lane 3corresponds to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with BS 168 (wild type) genomic DNA as template for PCR amplification. The lane 3 shows no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of base pairs; 1080 bp's. This can be used as a negative control with respect to the insertional mutation in *mmg* operon. If mutation does occur owing to the integration of the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgO*) into the chromosomal DNA of BS 168, then upon PCR screening, the agarose gel should show a band at exactly the same position (1080 bp's) as it is seen in lanes 1 & 2. But, if mutation does not occur, then no band corresponding to the size of *lacI* gene would be seen on the gel.

The lanes 4, 5, 6, 7, 8 & 9 correspond to the amplified *lacI* gene after PCR amplification using the primer set: *lacI*-up and *lacI*-dn with *mmgO* mutant(s)-1, 2, 3, 4, 5, 6 (pMUTIN-4_*mmgO*_BS168) genomic DNA as template for PCR amplification. All these lanes (4, 5, 6, 7, 8 & 9) show no band at about 1.0 kb corresponding to the length of the *lacI* gene in terms of number of base pairs 1080 bp's. This is suggestive that the cloned pMUTIN-4 plasmid (pMUTIN-4_*mmgO*) is not integrated into BS168 chromosomal DNA (Figure 43).

The single band in the lane 10 corresponds to the amplified *mmgO* insert after PCR amplification using the primer set: *mmgO*-up and *mmgO*-dn with cloned pMUTIN-4 (pMUTIN-4_*mmgO*) DNA as template for PCR amplification. The lane 10 shows a band in between 0.2 – 0.3 kb which is consistent with the theoretical length of the *mmgO* insert in terms of number of base pairs 249 bp's and is in good agreement with the previous gels from screening after PCR amplification. This is suggestive that the *mmgO* insert cloned into the Pmutin-4 plasmid used for the transformation was still intact with the plasmid (Figure 43).

The single band in the lane 11 corresponds to the amplified *mmgC* insert after PCR amplification using the primer set: *mmgC*-up and *mmgC*-dn with cloned pMUTIN-4 DNA (pMUTIN-4_*mmgC*) as template for PCR amplification. The lane 11 shows a band in between 0.3 – 0.4 kb which is consistent with the theoretical length of the *mmgC* insert in terms of number of base pairs; 341 bp's and is in good agreement with the previous gels from screening after PCR amplification. This is suggestive that the *mmgC* insert cloned into the pMUTIN-4 plasmid used for the transformation was still intact with the plasmid (Figure 41).

VI.D. Conclusions/Future work

So, in conclusion, we successfully PCR amplified the *mmgO/mmgC* inserts and cloned them into the pMUTIN-4 plasmid. However, we were not able to transform the BS 168 cells with the cloned pMUTIN-4 plasmids, such that this forms the desired mutation.

Coming to the future work, the first thing we would want to achieve is to determine whether BS 168 cells become competent. Using that critical information we would pursue transformation experiments in order to transform BS 168 cells with the cloned pMUTIN-4 plasmids.

REFERENCES

1. al, K. e. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis* 1997, p. 249-256.

2. Vagner, V.; Dervyn, E.; Ehrlich, S. D., A vector for systematic gene inactivation in Bacillus subtilis. *Microbiology (Reading, England)* **1998,** *144 (Pt 11)*, 3097-3104.

3. Stragier, P.; Losick, R. molecular genetics of sporulation in bacillus subtilis *annual reviews of genetics* [Online], 1996, p. 297-341.

4. Grossman, A. D., Genetic networks controlling the initiation of sporulation and the development of genetic competence in Bacillus subtilis. *Annual Review Of Genetics* **1995,** *29*, 477-508.

5. Koburger, T.; Weibezahn, J.; Bernhardt, J.; Homuth, G.; Hecker, M., Genome-wide mRNA profiling in glucose starved Bacillus subtilis cells. *Molecular Genetics And Genomics: MGG* **2005,** *274* (1), 1-12.

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

7. Stephens, C., Bacterial Sporulation: A question of commitment. *current biology* **1998,** *8*, R45-R48.

8. Ireton, K.; Jin, S.; Grossman, A. D.; Sonenshein, A. L., Krebs Cycle Function is Required for Activation of the SpoOA Transcription Factor in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* **1995,** *92* (7), 2845-2849.

9. Errington, J. Regulation of endospore formation in *Bacillus subtilis Nature Reviews Microbiology* [Online], (November 2003), p. 117-126.

10. Bryan, E. M.; Beall, B. W.; Moran, C. P., Jr., A sigma E dependent operon subject to catabolite repression during sporulation in Bacillus subtilis. *Journal Of Bacteriology* **1996,** *178* (16), 4778-4786.

11. Horswill, A. R.; Escalante-Semerena, J. C., Salmonella typhimurium LT2 catabolizes propionate via the 2-methylcitric acid cycle. *Journal of Bacteriology* **1999,** *181* (18), 5615.

12. Brock, M.; Maerker, C.; Schütz, A.; Völker, U.; Buckel, W., Oxidation of propionate to pyruvate in Escherichia coli. *European Journal of Biochemistry* **2002,** *269* (24), 6184-6194.

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

14. Russell, S. A. Over expression, Purification, and Characteization of MmgB and MmgC from *Bacillus subtilis* strain 168. The University of North Carolina, Greensboro, 2008.

15. REJWI, A. Overexpression, Purification, and Characterization of MmgD from Bacillus subtilis Strain 168. The University of North Carolina, Greensboro, 2009.

16. BOOTH, I., WILLIAM T. Characterization of the Biochemical Activity of the Open ReadingFrame “yqiQ” of Bacillus Subtilis. The University of North Carolina, Greensboro., 2011.

17. Kaneda, T., Fatty acids of the genus Bacillus: an example of branched-chain preference. *Bacteriological Reviews* **1977,** *41* (2), 391-418; Kaneda, T., Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiological Reviews* **1991,** *55* (2), 288-302.

18. Jin, S.; Sonenshein, A. L., Identification of two distinct Bacillus subtilis citrate synthase genes. *Journal Of Bacteriology* **1994,** *176* (15), 4669-4679.

19. Chen, J.-S.; Colby, G. D., Purification and properties of 3-hydroxybutyryl-coenzyme A dehydrogenase from Clostridium beijerinckii (""Clostridium butylicum") NRRL B593. *Applied & Environmental Microbiology* **1992,** *58* (10), 3297.

20. Owczarzy R, T. A., Wu Y, Manthey JA, McQuisten KA, Almabrazi HG, Pedersen KF, Lin Y, Garretson J, McEntaggart NO, Sailor CA, Dawson RB, Peek AS. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. 2008, p. W163–W16.

21. Invitrogen™, *Champion™ pET Directional TOPO® Expression Kits User Manual*. Version-H ed.; 2006; p 1-62.

22. Qiagen, QIAquick Spin Handbook. In *For QIAquick PCR purification Kit*, November 2006; pp 1-39.

23. Qiagen, QIAprep Miniprep Handbook. In *For purifcation of molecular biology grade DNA plasmid*, Second edition ed.; December 2006; pp 1-51.

24. Novagen, *His-Bind Kit: Technical Manual.* TB054 ed. Darmstadt, DE: Merck KGaA,. 2006.; Vol. Print. Rev. FO106.

25. Subtilist World-Wide Web Server.

26. Wilkins, M. R., Lindskog, I., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Hochstrasser, D.F., and Appel, R.D.; Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., *Detailed peptide characterisation using PEPTIDEMASS - a World-Wide Web accessible toolProtein Identification and Analysis Tools on the ExPASy Server*. 1997, 2005; Vol. 18(3-4).

27. Madan, V. K.; Hillmer, P.; Gottschalk, G., Purification and Properties of NADP-Dependent L(+)-3-Hydroxybutyryl-CoA Dehydrogenase from Clostridium kluyveri. *European Journal of Biochemistry* **1973,** *32* (1), 51-56.

28. Biolabs, N. E., Transformation protocol (C2988), NEB.

29. SPIZIZEN, C. A. A. J. REQUIREMENTS FOR TRANSFORMATION IN BACILLUS SUBTILIS 1960.

30. The great bacillus competency/transformation assay. In *iGEM*, University of Cambridge: University of cambridge, 2007; Bacillus-, -. B. o. M. B. M. f. *Two-step Bacillus subtilis Transformation Procedure*; Cambridge, i. U. o., Bacillus subtilis SynBio Chassis. In *i GEM 2007, Bacteria online*, University of Cambridge, 2007.

31. Jayakumar, R. Bacillus subtilis transformation methods.

32. DUBNAU, D. Genetic Competence in *Bacillus subtilis Microbiological reviews* [Online], 1991, p. 395-424; Horvath, S. Competence in *Bacillus subtilis* transformation system *Journal of general microbiology* [Online], 1968, p. 85-95.

APPENDIX A
SUPPLEMENTARY DATA

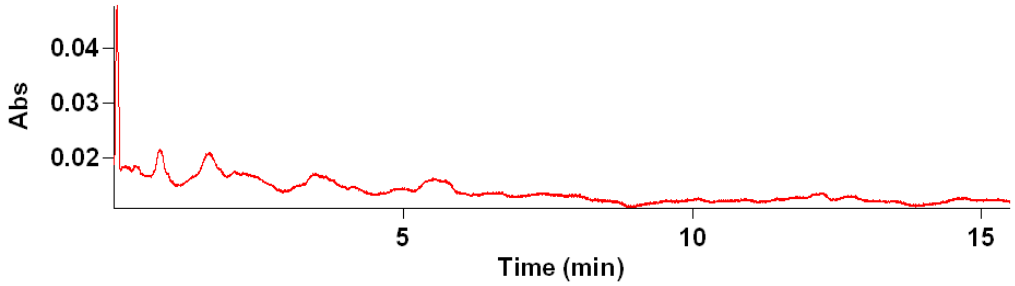


Figure 52: UV/Vis - time course of no cofactor control.

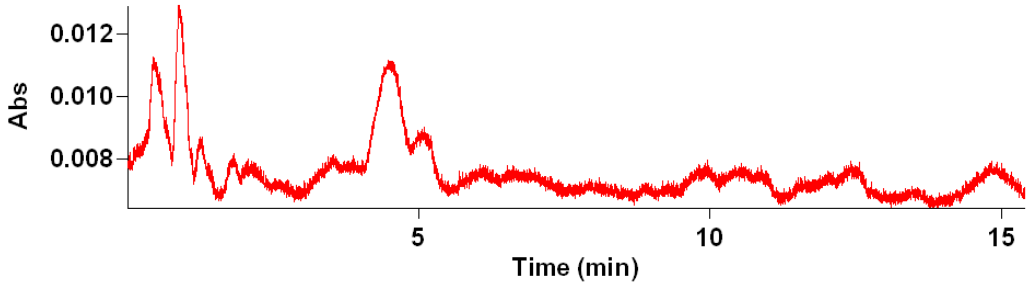


Figure 53: UV/Vis - time course of no substrate control.

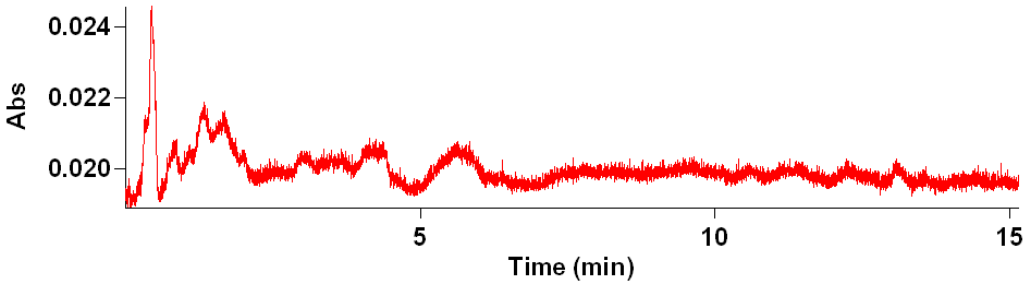


Figure 54: UV/Vis - time course of no enzyme control.

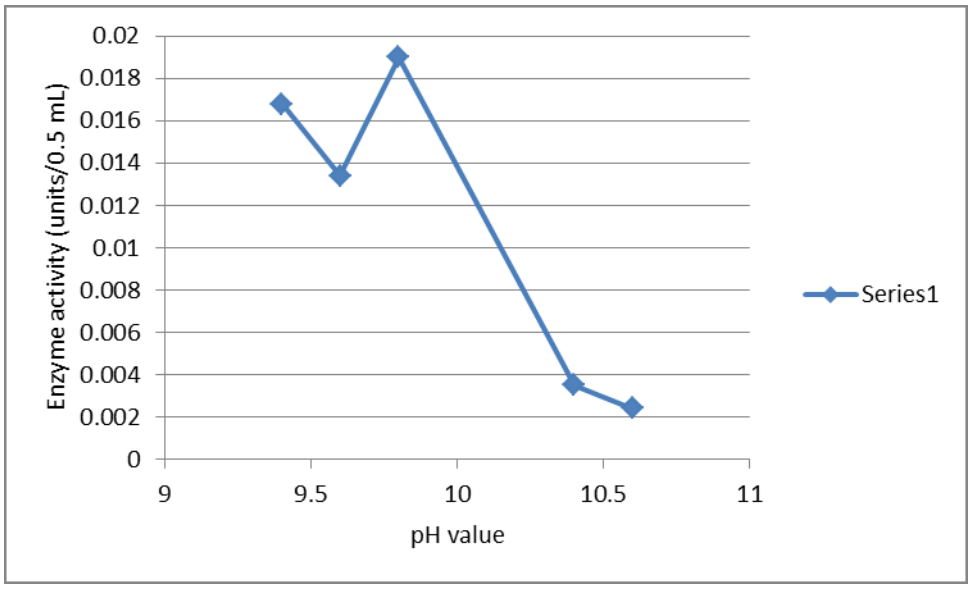


Figure 55: pH dependence curve of 3-OH-Butryl-CoA-DHase enzyme with glycine-NaOH buffer.

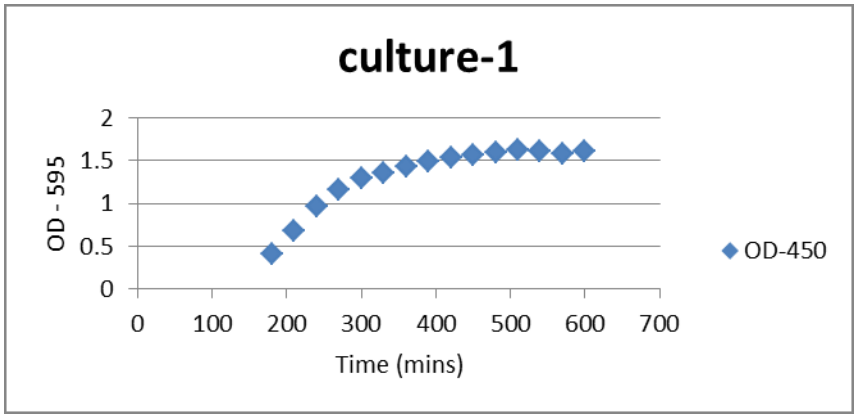
Growth curves of wild type *B.subtilis* 168

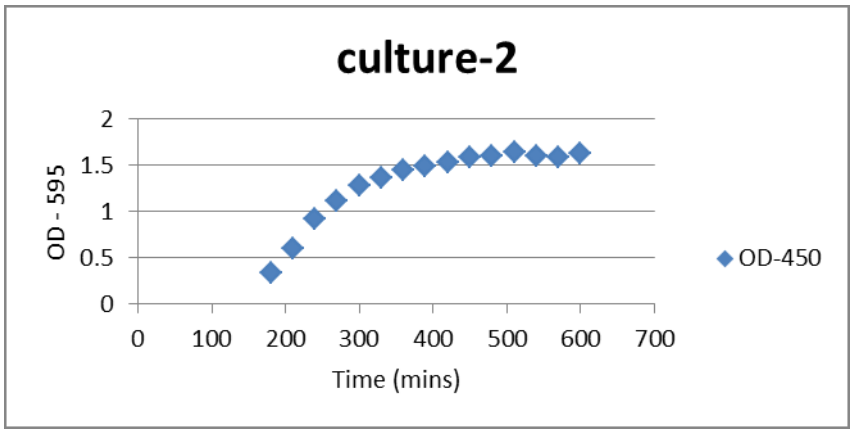
Since our experiments for transformation of BS 168 with the cloned pMUTIN-4 plasmids weren't successful, we were concerned that we were transforming the BS 168 when they were not competent. So, in an effort to sort out a way for making competent cells and subsequently transforming BS 168 cells with the cloned plasmids, we initiated growth curves. Literature precedent show that BS 168 cells become naturally competent during the transition from exponential to stationary phase (i.e., either in late exponential or in early stationary phase)[30, 32].

Hence, we started to measured growth curves of the wild type *B.subtilis* 168. Overnight starter cultures of the wild type *B.subtilis* 168 cells were made. Starter culture was grown by inoculating 5.0 mL of LB-broth with a colony of the wild type *B.subtilis* 168 cells and incubating aerobically at 37^0 C/220 rpm. 2.0 mL of this starter culture was

used to inoculate the bigger batches (50 mls) of the liquid media and the growth was spectrophotometrically monitored at 595 nm. Growth curves were obtained for triplicates of the cultures[30, 31].

Looking athe growth curves (Figure 56), it seems like, BS 168 cells reach the stationary in about 550 minutes. The transistion from exponential to stationary phase seems to take place at about 400 minutes from the time the culutures were initiated.





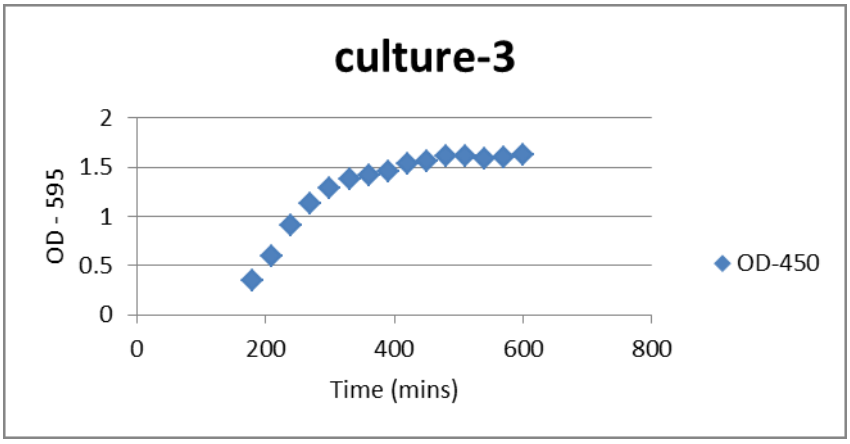


Figure 56: Triplicate of the growth curves of BS 168 cells in LB-broth.