Katrina Burwash

Advancement of the XylS/Pm Expression System in Gram-positive Bacteria

Evaluated in *Bacillus subtilis*

Master's thesis in Biotechnology (2 year)
Supervisor: Trygve Brautaset

May 2019

NTNU
Norwegian University of
Science and Technology

Katrina Burwash

Advancement of the XylS/Pm Expression System in Gram-positive Bacteria

Evaluated in *Bacillus subtilis*

NTNU

Norwegian University of
Science and Technology

Preface

This Master's Thesis concludes my Master of Science (M.Sc.) in Biotechnology at the Norwegian University of Science and Technology (NTNU) located in Trondheim. This thesis was written at the Department of Biotechnology and Food Science (IBT) in collaboration with Vectron Biosolutions AS under the supervision of Professor Trygve Brautaset (IBT), Dr. Morteza Moghadam (Vectron Biosolutions AS) and Dr. Jostein Malmo (Vectron Biosolutions AS).

Acknowledgements

I would first like to thank my NTNU supervisor, Professor Trygve Brautaset, as well as Vectron Biosolutions AS for giving me the opportunity to work on this project.

Secondly, I would like thank the employees at Vectron Biosolutions AS for everything they have taught me, for their patience with my many questions and for including me in both a great work and social environment. Especially, I would like to thank Dr. Morteza Moghadam for all the guidance and support along the way, I am very grateful for his help.

Finally, I would like to thank my fellow students in the lab, friends, family and boyfriend for all their support and encouragement. They have made this final year unforgettable.

Abstract

Escherichia coli (*E. coli*) is one of the best characterized prokaryotes and therefore one of the most popular hosts used for heterologous protein expression. Today, the XylS/Pm expression system is predominantly used in *E. coli*, however, there are some disadvantages when using *E. coli* as a host for recombinant protein expression; such as poor protein secretion into the extracellular medium. An advantage that gram-positive bacteria have is that they are able to secrete proteins straight into the culture medium. In this study, advancement of the XylS/Pm expression system was evaluated in *B. subtilis*.

In previous work, it was shown that there was no XylS detected from its native promoters, P_{s1} and P_{s2} and was identified as one of the bottlenecks. After changing the original promoters to native *B. subtilis* promoters and codon optimizing XylS, it led to successful XylS production but still no mCherry production from P_m. Currently, this expression system is still not functional in *B. subtilis* and the potential bottleneck is now thought to be the production of the target protein, mCherry, from P_m.

It was hypothesized that the coexpression of *E. coli* sigma factors may facilitate the expression of the target protein, mCherry. Sigma factor production of σ^{32} and σ^{38} that were being expressed from P_{liaI} in constructs pMSM67.4 and pMSM69.4 was verified through western blotting, however, when mCherry fluorescence was measured post induction, there was no detectable signal. A qPCR analysis was performed for the relative quantification of mCherry transcript from P_m, where it was shown that there was a significant fold change between the induced cultures of pMSM67.4 and pMSM67, also between pMSM69 and pMSM67, which showed increased levels of mCherry transcript, however, not between pMSM69.4 and pMSM69. Finally, after running protein sequence alignments it was discovered that there was very little similarity between both the sigma factors and RNAP core enzymes of *B. subtilis* and *E. coli*.

It was concluded that the coexpression of sigma factors had a positive effect on the levels of mCherry transcript, as did expressing higher levels of XylS, however, there was not an additive effect when sigma factors were combined with higher levels of XylS as was seen between pMSM69.4 and pMSM69, with still no detectable mCherry fluorescence signal. Looking at the sequence analyses there could still be bottlenecks at the transcription initiation level due to not only σ^{32} and σ^{38} being unrecognizable by *B. subtilis* RNAP, but also the interactions of XylS with the RNAP α-subunit, which are both required for transcription from P_m.

Nomenclature

α Alpha

β Beta

δ Delta

ϵ Epsilon

Amp   Ampicillin

Bac   Bacitracin

bp   Base-pair

Cm   Chloramphenicol

CTD   C-Terminal Domain

dH$_2$O   Distilled Water

DTT   Dithiothreitol

Kan   Kanamycin

LA   Lysogeny Broth Agar Medium

LB   Lysogeny Broth Medium

NEB   New England Biolabs

OD$_{600}$ Optical Density Measured at 600 nm

rcf   Relative Centrifugal Force

RFU   Relative Fluorescence Unit

RNAP   RNA Polymerase

rpm   Revolutions per Minute

SDS-PAGE   Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Tm   Melting Temperature

TMB   Tetramethylbenzidine

Contents

1   Introduction

1.1   Biotechnology Industry

Biotechnology has been around since before the 19th century, and has developed through three different phases that can be categorized as ancient biotechnology, classical biotechnology and finally, modern biotechnology [1]. In the pre-1800s,even though at the time it was unbeknownst to them, people were forging the future of biotechnology by simply preserving their food with vinegar, taking advantage of yeast fermentation to make bread and beer or by using the enzyme rennet, found in calves' stomachs, which was added to sour milk to produce cheese [1]. The second phase of biotechnology brings us from 1800 to the middle of the 20th century. In this phase, Gregor Mendel presented his laws of inheritance, coining the phrase mendelian inheritance; Charles Darwin published his book, Theory of Evolution; and of course the first antibiotic, penicillin, was accidentally discovered by Alexander Fleming [1]. Shortly after these revolutions, was the start of the modern biotechnology phase, which began with the elucidation of the double helix model of DNA by Watson and Crick in 1953. Gene sequencing was also introduced in this phase, and in 2000, Craig Venter managed to sequence the human genome. Other discoveries include synthesizing recombinant DNA through cloning in various hosts, inserting foreign DNA into another host and also the amplification of DNA in a test tube [1]. This industry continues to grow on a daily basis.

Today, biotechnology can be defined as the application of science and technology to living organisms, as well as parts, products and models thereof, to alter living or non-living materials for the production of knowledge, goods and services [1]. This being such a broad definition, the biotechnology field can be separated into four main branches; blue, green, white and red. The blue biotechnology branch is devoted to marine biotechnology, the green branch deals with agrifood biotechnology and GM plants, the white branch works within the industrial and bioprocessing aspects of biotechnology and finally, the red biotechnology branch deals with the health care industry, genetic testing and gene therapy [2].

Biotechnology is such a large and interdisciplinary field that quite often these four branches of biotechnology intersect, particularly the red and white branches when developing bacterial strains for heterologous protein expression which eventually leads to an industrial-scale production of the desired product. When choosing a bacterial strain to produce a protein of interest, it is important to choose a strain that is genetically stable, safe, easily amendable to genetic manipulation, easily able to harvest protein from and also that it efficiently produces the target product. Unfortunately, it is difficult to choose a bacterial host and promoter system that will be best suited to express any particular protein of interest, when in fact, it often depends on the target protein itself. It is therefore advantageous to explore different promoter systems within different bacterial hosts to create a versatile genetic toolbox. Some commonly used

bacterial hosts include *Escherichia coli*, *Streptomyces*, *Lactobacillus*, *Corynebacterium* and *Bacillus* spp. Other alternative hosts that are also used for heterologous protein expression include yeasts, filamentous fungi and mammalian cell cultures [3].

There are many aspects within a host strain that can hinder or create bottlenecks for the proper production and folding of a protein. To increase production, some factors to consider might include codon optimizing the gene of interest, which promoter system to use, the temperature at which the protein should be expressed at, if signal peptides and/or chaperone proteins should be implemented, if the RNA polymerase of the host requires extra sigma factors and finally the copy number of the plasmid being used [3].

1.2   *E. coli* as an Expression Host

E. coli is a gram-negative bacterium that has played a role as a model organism for decades; and as a result has one of the best characterized genomes amongst prokaryotes. Some of the reasons why it is so attractive to researchers is due to its small genome size which makes it easy to manipulate in a laboratory setting and perform genetic experiments, also its ability to grow rapidly to high densities on inexpensive substrates [4] [5]. This gram-negative prokaryote is an excellent versatile host that is frequently utilized in the biotechnology industry and to date has been used to develop genetic engineering techniques such as molecular cloning, recombinant DNA and heterologous protein expression systems [4]. In the bioprocessing branch, *E. coli* has been used to produce biofuels and industrial chemicals like phenol and ethanol but also therapeutic proteins such as growth hormones, insulin, interleukins and interferons [6] [4].

When it comes to heterologous protein expression, there are unfortunately also a few disadvantages of using *E. coli* as an expression host. Gram-negative bacteria, have a periplasmic space separate from the cytoplasm. The perisplasm is an oxidative environment which allows for protein oxidation, folding and quality control [7]. Sometimes, proteins that are being overexpressed in the cytoplasm that require an oxidative environment for proper folding have difficulty translocating to the periplasm across the inner membrane, which can lead to inclusion body formation [8]. Inclusion body formation can also be the result of sub-optimal codon usage, which would to sub-optimal rate of translation that negatively affects the co-translational folding process which then consequently forms improperly folded proteins leading to the formation of inclusion bodies [9]. When using *E. coli* to produce therapeutic proteins, another disadvantage is that they produce and accumulate lipopolysaccharides (LPS), also known as endotoxins, which are pyrogenic in humans; therefore, making the downstream processing more complicated [3][10]. Scientists have therefore been turning to gram-positive bacteria in hopes for solutions to some of these problems.

1.3   Gram-positive Bacteria as Expression Hosts

One of the main reasons that scientists are now exploring gram-positive bacteria as heterologous protein expression hosts is due to the fact that they have a naturally high secretion capacity and are able to secrete the proteins directly into the culture medium [3]. This is advantageous because it makes it much easier for downstream processing and protein recovery [11]. Another advantage over gram-negative bacteria, is that gram-positive bacteria do not contain LPS on their outer membranes [3].

1.3.1   *B. subtilis*

One of the most popular gram-positive bacterium that is used for heterologous protein expression is *Bacillus subtilis*, mainly due to the above mentioned reasons just stated and as well as the fact that its genome is well characterized. *B. subtilis* is known as a GRAS (generally regarded as safe) organism, which is why this species is highly favored over the use of *E. coli* when it comes to the production of food products [12]. *B. subtilis*, commonly known as grass bacillus, is a rod-shaped bacterium that is found in soil. It is frequently used on an industrial scale for secreted enzyme production and is often considered to be the gram-positive equivalent of *E. coli*. There are of course drawbacks when using *B. subtilis* as a host, and one of the major disadvantages is the fact that it secretes large quantities of proteases into the culture medium, which have the potential to degrade other secreted proteins of interest [3]. Some protease deficient strains have now been produced to circumvent this such as WB600, DB105 and DB104 [13] [14] [15]. Unfortunately, there are other bottlenecks, like poor transportation to the translocase, incorrect protein folding, a lack of suitable expression vectors and plasmid instability [12]. Due to the need for alternate bacterial hosts for heterologous protein expression, there is continuous research being done to find solutions to overcome these bottlenecks.

1.4   Gene Regulation in Bacteria

Prokaryotes reproduce asexually through simple fission and contain a circular chromosomal DNA as well as autonomously replicating mini chromosomes called plasmids [16]. To get to the final point of a functional protein, first a cell must transcribe its DNA or gene of interest, which will get processed into messenger RNA (mRNA). The mRNA then gets translated into a protein, with some posttranslational modifications performed at the very end to make it a functional protein. Gene expression in prokaryotes can be regulated at several different levels throughout this process, which is detailed in the diagram below.

Figure 1.1: Simplified pathway for prokaryotic gene expression with 5 levels of regulation. 1 is transcription, 2 is RNA processing, 3 is RNA stability, 4 is translation and 5 is posttranslational modifications. This image was adapted from [16].

For prokaryotes, the most important level of regulation occurs at the transcriptional level, closely followed by translation. Due to the fact that prokaryotes don't have membrane bound organelles; transcription, translation and mRNA degradation are all coupled in prokaryotes, so therefore the gene products may be produced in varying amounts from one single transcript [16].

1.4.1   Transcription and Translation

Transcription is the process of transcribing the DNA into mRNA and is split up into three steps; initiation, elongation and termination. Transcription is carried out by the core enzyme RNA polymerase (RNAP) that is made up of five subunits (two α, β, β' and ω) and is eventually associated with a sigma (σ) factor. In gram-positive bacteria, especially in the phylum *Firmicutes*, such as *B. subtilis*, the RNA polymerase core enzyme is associated with two additional factors, called δ, encoded by the *rpoE* gene and ϵ, encoded by *rpoY* [17][18]. ϵ is hypothesized to have a role in protection from phage infection [18]. δ facilitates the open complex function formation that leads to rapid synthesis of transcripts as well as recycling of RNAP and finally also functions as a transcription regulator [17]. Artsimovitch *et al.* performed a study that was published in 2000 regarding the recognition of regulatory signals *in vitro* between *B. subtilis* and *E. coli* RNAPs [19]. They found that even though their RNAPs had functional and structural similarities, they were significantly different when it came to promoter utilization patterns, recognition of the hairpin-dependent pause sites and also some arrest and termination signals, and overall were only 50% identical when it came to amino acid sequence [19].

Together, the five subunits with the addition of sigma factors is referred to as the RNA polymerase holoenzyme, where the σ factor assists the RNA polymerase in reading the template DNA to tell it where to begin transcribing [20]. The polymerase holoenzyme slides down the DNA until it reaches the starting point for RNA synthesis, called a promoter. Some point after initiation, the σ factor will dissociate and the RNA polymerase will continue into the elongation step, creating the RNA transcript from ribonucleotides. The RNA transcript continues to elongate until the enzyme reaches a terminator signal, where the polymerase stops and then releases both the newly made single stranded RNA molecule and the DNA template [20].

Translation is the next step in producing a functional protein, and is an important level for gene regulation in prokaryotes. It is the process of translating an mRNA molecule into a protein using ribosomes. As with transcription, translation can also be divided into three steps; initiation, elongation and termination. Initiation of translation in prokaryotes begins at the ribosome binding site, known as a Shine-Dalgarno sequence, which is located a few nucleotides upstream of the start codon [20]. Once all the necessary ribosomal subunits have associated, the translation begins, where the mRNA nucleotides are read in triplicate which are referred to as codons. There are 64 codons that represent 20 different amino acids, meaning that multiple codons encode for the same amino acid. Succeeding the Shine-Dalgarno consensus sequence, is the first codon, AUG, which encodes a methionine. The mRNA is passed through three ribosomal sites, where transfer RNAs (tRNAs) bearing the anticodon complementary to the codon present in the mRNA will attach their amino acid residue to the growing peptide chain. This elongation continues until the ribosome encounters a stop codon [20]. Once a stop codon is reached, release factors bind and release the polypeptide chain and the ribosomal subunits dissociate.

Different bacterial species have their own bias when it comes to the codons used for specific amino acids. When expressing a heterologous protein, it is sometimes beneficial to codon optimize the gene to the expression hosts' preferences because the tRNA population often reflect the codon bias of the mRNA population within a host [21]. Differences in the codon usage between the original host of the heterologous protein and the expression host can therefore impede translation due to the need for one or more tRNAs that might be rare or lacking in the expression host [22]. If these tRNAs are underrepresented, then it can result in translational stalling, premature translation termination, translational frameshift, or misincorporated amino acids [23].

1.4.2 Sigma Factors

As previously mentioned, sigma factors play a fundamental role during transcription initiation. There have been several different types of sigma factors discovered, and they can be categorized into two different families depending on their homology to two sigma factors in *E. coli*; σ^{70} and σ^{54}. In *E. coli*, σ^{70} is responsible for most of the transcription during cell growth and housekeeping genes, where its equivalent in *B. subtilis* is referred to as σ^{A}. The σ^{54} family guides transcription in response to environmental signals, however there is no known representative sigma factor from this family in gram-positive bacteria [24] [25]. *E. coli* have been known to contain seven different sigma factors, where as *B. subtilis* has 18 sigma factors [26]. The σ^{70} family is quite large that it is divided up into four groups based on phylogenetic relatedness [27]. Group 1 includes the σ^{70} and its orthologs, which are responsible for the transcription of housekeeping genes[27]. Group 2 are closely related to group 1, group 3 contain sigma factors that are more divergent in sequence and are organized into clusters with similar functions,

and finally group 4 have distant sequence similarity to the other 3 groups [27]. Sigma factors provide an extra level of transcriptional control as they bind only to distinct sets of promoters [28].

Two important sigma factors are the *E. coli* σ^{38} and σ^{32}. σ^{38}, also known as RpoS, is the main regulator of the general stress response in *E. coli*, and it either directly or indirectly regulates 10% of *E. coli* genes during stationary phase [26]. The second sigma factor, σ^{32}, also referred to as RpoH, regulates the transcription of the heat shock genes [26]. The numbers corresponding to the sigma factors' names also correspond to their protein weight in kDa [29]. *B. subtilis* has a functionally similar sigma factor to both σ^{38} and σ^{32}, known as σ^{B}, that reportedly is involved in the general stress response and heat shock response, where as σ^{H}-dependent promoters in *B. subtilis* are activated during the early stationary phase [30], which is functionally similar to the *E. coli* σ^{32}. All four of these sigma factors belong the the σ^{70} family, where *E. coli* σ^{38} belongs to group 2, and *E. coli* σ^{32}, *B. subtilis* σ^{H} and σ^{B} all belong to group 3 [27]. Structural and biophysical studies have shown that sigma factors interact with the -10 and -35 elements of the promoter region during transcription initiation [31].

1.4.3   Promoters

Promoters are regions of DNA located upstream of the coding sequences of genes and are recognized by the RNA polymerase to initiate transcription of the DNA strand. Promoters can be classified based on their strength; weak, medium or strong, where the strength of a promoter often depends on its nucleotide sequence, and depending on this sequence, the rate of gene transcription to mRNA will be faster or slower [32]. Most prokaryotes contain the -10 and -35 regions that have a consensus sequence of TATAAT and TTGACA respectively, which are essential for the function of bacterial promoters and they also determine the promoter strength by how easily and strongly the RNA polymerase binds [32].

Several studies have found that many functional *E. coli* promoters are not transcribed by the *B. subtilis* promoters, even though the *B. subtilis* promoters are able to function in *E. coli* which could be due to the variation in consensus sequences among promoters. In 1998, Voskuil and Chambliss found a somewhat conserved -16 region among *B. subtilis* promoters containing a TRTG motif. Strangely, they found that this -16 region was not necessary for many promoters and that its utility was context dependent. For example; this -16 region wasn't required among some promoters that had adequate sequence information in other elements like the -10 and -35 regions, and was even detrimental to the full function of the promoter, however on the other hand, this -16 region was found essential for promoters lacking the -35 conserved region or in the "extended -10" class of promoters [33]. The TRTG motif has been found quite common in gram-positive bacterial promoters but rare in gram-negative bacterial promoters like *E. coli* [33]. Both gram-negative and positive bacterial RNA polymerase use the -10

and -35 regions of the promoter for transcription initiation, but gram-positive bacterial RNA polymerases use different combinations of these elements to form functional promoters.

This study makes use of several native *Bacillus* promoters; P_{liaG}, P_{liaI} and P_{met}. The promoter, P_m, is native to *Pseudomonas putida* and is used to express the reporter gene, mCherry as a measure of the XylS/Pm cassette's functionality. P_{liaG} is considered to be a constitutive promoter, P_{liaI} is an inducible promoter by bacitracin and P_{met} promoter can be induced by methanol [34][35][36].

1.5 XylS/Pm Expression Cassette

The XylS/Pm expression cassette originates from the TOL pWWO plasmid which is native to the gram-negative bacteria, *P. putida*. The TOL pWWO plasmid encodes a pathway for the catabolism of toluene and xylenes, where the genes responsible for this are grouped into an upper- and meta-operon [37]. The two operons are positively regulated, where the upper operon is regulated by the transcription factor XylR, and the meta-operon is regulated by the transcription factor XylS [37]. XylS is part of the AraC family of positive transcriptional regulators and is 321 amino acids long. In pWWO, the XylS protein is transcribed from two separate, tandem promoters, P_{s1} which is σ^{54}-dependent and inducible, and the second promoter, P_{s2} is σ^{70}-dependent and provides constitutive, low-level expression of *XylS* [38]. An overview of the Xyls/Pm system is depicted in Figure 1.2.
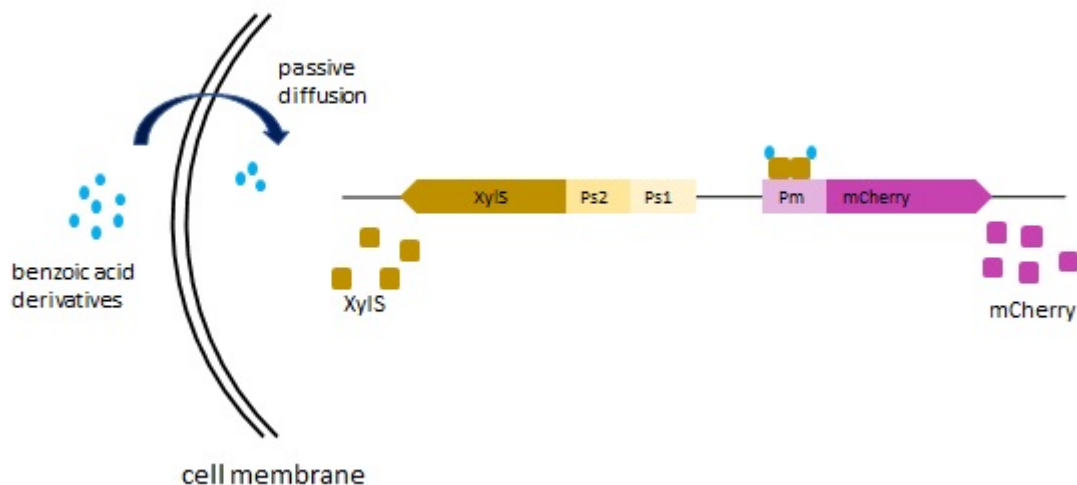
Figure 1.2: XylS/Pm expression system, where XylS is being expressed from its tandem promoters P_{s1} and P_{s2}. The inducer molecules (benzoic acid derivates) cross the cell membrane by passive diffusion, where once inside the cell, they cause XylS monomers to dimerize. Once dimerized, the XylS binds to P_m which then activates expression of mCherry. Figure adapted from [37].

The XylS/Pm expression system can be induced with many different types of benzoic acid derivatives such as acetyl salicylic acid, benzoate, 4-iodobenzoate and m-toluate [39]. These benzoic acid derivatives enter the cell via passive diffusion where they bind to XylS to facilitate its dimerization. The XylS dimer then binds to two 15-bp repeated motifs, one from -70 to -56 and the other from -49 to -35 of the P_m promoter,where the first XylS monomer occupying the proximal binding site interacts with the C-terminal domain of the RNA polymerase α subunit (α-CTD) to activate transcription from P_m [40]. The structure of α-CTD from *E. coli* RNAP has been resolved using NMR by Jeon *et al.* where they found that the α-CTD was composed of four alpha helices; helix 1 (residues 264 to 273), helix 2 (residues 278 to 283), helix 3 (residues 286-292), and helix 4 (residues 297-309) [41]. The residues in helix 3 contain amino acid residue 289 where there have previously been found many interactions with positive regulators [41] [40] [42] [43]. Ruiz and Ramos therefore hypothesized from their study that this amino acid residue may also represent a site of interaction between α-CTD and XylS [40]. The α-CTD is involved in interaction with transcriptional regulators and with upstream promoter elements [43]. P_m is dependent on σ^{32} in the early exponential growth phase and σ^{38} in the early stationary phase and onwards [44]. Figure 1.3 demonstrates where the XylS dimer and RNA polymerase bind along P_m to initiate transcription.
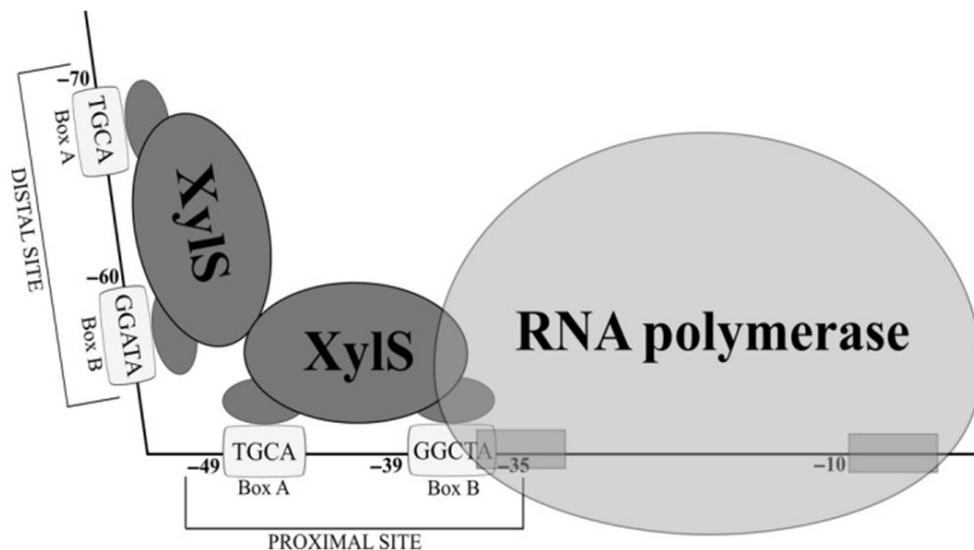
Figure 1.3: XylS dimerizes and interacts with RNA polymerase to initiate transcription from P_m. This image was obtained from [39].

1.6 Previous Work

Previously, growth studies were performed with *B. subtilis* 168 to determine its doubling time and preferred growth medium to optimize the growth conditions. Additionally, inducer diffusion studies were carried out to determine whether m-toluate was able to permeate the cell wall and which concentration was appropriate when inducing the XylS/Pm expression cassette [45]. A series of vectors were made that replaced the original *XylS* promoters with native *B. subtilis* promoters to address the potential bottleneck in XylS production, which once created, was hypothesized to result in mCherry production from P_m. From this 12 vector series, only one construct resulted in mCherry production when expressed in DH5α, which was pMSM67, which had XylS expressed from the P_{liaG} promoter. Unfortunately, there was no mCherry produced from P_m when expressed in *B. subtilis*, however there was a significant increase of XylS expression from the P_{met} promoter, which was construct pMSM69, in both DH5α and *B. subtilis* when compared to the other native *B. subtilis* promoters [45]. From this previous study, it was concluded that there was sufficient amounts of XylS being expressed, and therefore a new possible bottleneck was investigated.

1.7 Aim of Study

E. coli is one of the most popular hosts used for heterologous protein expression, unfortunately, not every protein is able to be functionally expressed within this species due to possible inclusion body formation. Gram-positive prokaryotes could be a potential solution to this due to their different intercellular environments and their ability to secrete proteins directly into the culture medium, which makes downstream processing easier. This study focuses on optimizing heterologous protein expression using the XylS/Pm expression system with *B. subtilis* as a host.

The aim of this study was to further identify and investigate the bottlenecks concerning protein expression from the XylS/Pm expression system, where the current potential bottleneck lies at mCherry production from P_m.

It is hypothesized that coexpressing *E. coli* sigma factors, σ^{32} and σ^{38}, together with the XylS/Pm expression cassette will improve mCherry production from P_m. This will be tested by running expression experiments to measure mCherry fluorescence levels in addition to western blots to verify sigma factor expression. qPCR assays will be performed as well to measure mCherry transcript levels between the relevant constructs. Finally, protein sequence alignments will be run comparing sigma factors and the α-subunit of RNAP between *E. coli* and *B. subtilis*. Hopefully the results of these tests can point towards a solution for the lack of mCherry fluorescence from P_m.

2 Materials and Methods

2.1 Theoretical Introduction to the Methods

2.1.1 Restriction Endonucleases

Restriction endonucleases, also known as restriction enzymes are proteins that recognize specific short nucleotide sequences, that cut the DNA only at that specified site, which is known as the restriction site. Restriction enzymes can either produce blunt ends or sticky ends within the DNA strand; where the sticky end creates a single stranded DNA overhang that is able to overlap with a complementary strand. These enzymes are commonly used when cloning to produce recombinant DNA molecules or plasmids, but are also used when performing a control digest in order to verify a construct already made. When used for cloning purposes, restriction enzymes are used to isolate DNA fragments or a gene of interest, then after using the same restriction enzymes to digest the backbone, the gene of interest can be ligated into the backbone via their compatible ends.

2.1.2 Gel Electrophoresis

Gel electrophoresis is used to separate DNA fragments through an agarose gel within an electric field that forces the negatively charged DNA to migrate towards the anode [46]. The rate of migration of a particular fragment depends on the size of the DNA molecule, the applied voltage, agarose concentration, electrophoresis buffer used and DNA conformation [46]. The DNA fragments show up as bands on the agarose gel and can be viewed by adding a colored dye such as GelGreen or GelRed (Biotium) to the agarose which will fluoresce under UV light. A ladder is a molecular weight standard with fragments of known sizes so that one may compare their unknown or expected fragment sizes [46].

2.1.3 Ligation

Ligation is essentially "pasting" a gene of interest into the backbone using T4 DNA ligase, it can be done with an insert produced from PCR or an insert extracted from a gel. DNA ligase catalyzes the reaction of joining the insert and backbone DNA fragments together, forming a complete circularized vector. The ligase links the 3'OH group of one strand with the 5'PO$_4$ on the neighboring strand [47].

2.1.4 Transformation

There are two types of transformation used in this study; heat shock transformation with chemically competent cells and electroporation with electrocompetent cells. Heat shock transformation uses a calcium rich environment provided by RbCl which counteracts the electrostatic repulsion between the plasmid DNA and bacterial cellular membrane [48]. When the cells are exposed to a sudden increase in temperature, it creates

pores in the plasma membrane that allow for foreign DNA to enter the cell [48].

Electroporation is an alternative method of bacterial transformation. Electroporation is a physical form of transformation that uses an electrical pulse to create temporary pores in the cell membrane, through which foreign DNA can be taken up by the cell [49].

2.1.5   SDS-PAGE and Western Blot

SDS polyacrylamide gel electrophoresis (SDS-PAGE) is a technique used to separate proteins in a similar manner to gel electrophoresis used for DNA separation. SDS is a detergent that binds to the hydrophobic regions of the protein which causes them to denature and gives them a uniform negative charge [16]. To break any disulfide bonds present in the protein, there is the option of adding a reducing agent, such as dithiothreitol (DTT) to the loading dye. An electric current is applied to the gel, and the proteins migrate through the gel. The proteins appear as distinctive bands once the gel is stained in a dye like Coomassie blue.

Once the SDS-PAGE has run, the separated polypeptides can also be transferred from the gel to a nitrocellulose or PVDF membrane when exposed to an electric current. The individual proteins can then be visualized using antibodies. Typically, a primary and a secondary antibody are used. The primary antibody is specific for the protein of interest, and the secondary antibody is specific to the epitopes of the primary antibody. Then the secondary antibody can be conjugated to either a radioactive isotope or an enzyme that produces a visible product when the correct substrate is added [16]. In this study, the secondary antibody is conjugated to horse radish peroxidase (HRP), and it's substrate is tetramethylbenzidine (TMB).

2.1.6   Bioanalyzer

The Agilent 2100 bioanalyzer (Agilent Technologies) is a microfluidics instrument that enables RNA quality assessments. The instrument electrophoretically separates RNA, where a fluorescence detector measures the separation. Based on the separation, the Bioanalyzer software generates an electropherogram and provides the RNA integrity number (RIN) and RNA concentration. An electropherogram is presented as the fluorescence intensity as a function of time. Degraded prokaryotic RNA samples show small or no peaks of the 23S and 16S rRNA, whereas intact samples have higher distinct peaks. The RIN scale goes from 1-10, where 10 represents the highest quality of RNA. In order to use RNA for downstream purposes like qPCR, it is recommended that the RIN value is at least 7 or greater.

2.1.7   qPCR Technology

Quantitative PCR (qPCR) or reverse transcription (RT-PCR) is a variation of the traditional polymerase chain reaction (PCR) with denaturation, annealing and extension steps. In this study, two-step qRT-PCR is used, where the total RNA in the samples are synthesized into cDNA using a reverse transcriptase. After the cDNA is synthesized, it can be used as the template for qPCR. qPCR can be used for a number of different assays such as gene expression profiling, viral titer determination, copy number variation analysis and allelic discrimination [50]. This study used TaqMan chemistry for a gene expression assay. The essential part of TaqMan chemistry is that it incorporates a probe, which is a gene specific sequence that binds to the target gene between the two primers [50]. The 5' end of the probe is attached to a reporter dye such as FAM, which is a fluorescent dye that reports the amplification of the target gene. Then on the 3' end of the probe is the quencher such as TAMRA, that quenches the fluorescence from the reporter dye. [50]. Prior to PCR, the reporter dye is temporarily quenched by the quencher. Then during PCR, the primers and probe anneal to their target, once the DNA polymerase has extended upstream of the probe, the 5' nuclease activity of the polymerase cleaves the probe. Once the probe is cleaved, the quencher and reporter dyes are no longer attracted to each other, meaning that the reporter dye won't be quenched [50]. Eventually, enough probe is cleaved in the reaction, and ends up accumulating, which allows the amplification signal to be detected.

2.2   Solutions and Media

Please refer to Appendix A for a complete list of solutions and media along with their compositions used in this study and Appendix B for the inducer and antibiotic compositions.

2.3   Preparation of Competent Cells

2.3.1   Supercompetent *E. coli*

The procedure of producing chemically competent cells has been adapted from a previously described method [51]. First, an overnight culture of *E. coli* DH5α was prepared from a freezer stock in 5 mL of LB medium and placed at 37°C and 225 rpm. The following day, 1% of the overnight culture was used to inoculate 50 mL of PSI medium and then placed at 37°C and 225 rpm until reaching OD_{600} of 0.4. The cells were transferred to a pre-chilled 250 mL centrifugation flask followed by an incubation on ice for 15 minutes. The cells were harvested at 3000 rpm for 5 minutes at 4°C (Sorvall Lynx 6000 centrifuge). After discarding the supernatant, the cells were gently resuspended in 20 mL cold TFB1, followed by another 15 minutes incubation period on ice. The cells were harvested again at 3000 rpm for 5 minutes at 4°C . The supernatant was discarded and the cells were gently resuspended in 1.5 mL of Cold TFB2. After resuspending,

the cells were distributed into 1.5 mL microfuge tubes in 100 μL aliquots and then immediately frozen with liquid nitrogen. The competent cells were stored at -80°C.

2.3.2   Electrocompetent *E. coli*

This protocol was adapted from the method provided by New England Biolabs with some modifications. [52]. First, an overnight culture of DH5α was prepared in 10 mL LB in a 125 mL shake flask. The following morning, 1% of the overnight culture was inoculated into 50 mL of LB in a 250 mL baffled flask. The culture was placed in an incubator set at 37°C and 225 rpm until the culture reached an OD_{600} between 0.5-0.7. After this incubation step, the culture was transferred to a 50 mL falcon tube and placed on ice for 15 minutes. Then, the culture was centrifuged at 5000 rcf for 10 minutes (Eppendorf 5804R). The supernatant was aspirated off, and the cells were resuspended via vortexing in 20 mL of cold sterile 10% glycerol. This washing step was repeated 2 more times and resuspended via pipetting in a final volume of 500 μL of cold sterile 10% glycerol. The cells were distributed into 100 μL aliquots into pre-chilled eppendorf 1.5 mL tubes. Finally, the cells were frozen with liquid nitrogen and stored at -80°C.

2.3.3   Electrocompetent *B. subtilis*

This protocol was adapted from [53]. An overnight culture of *B. subtilis* was prepared from a freezer stock and grown in five mL of LB medium at 37°C and 225 rpm. The next day, 48 mL of Growth media was re-inoculated with 0.6% of the overnight culture. The fresh culture was grown at 37°C and 225 rpm until reaching an OD_{600} between 0.85-0.95. After reaching the required OD, the fresh culture was incubated on ice for 15 minutes before harvesting the cells. The culture was transferred to a 50 mL falcon tube, and centrifuged at 4°C for five minutes at 10 000 rpm (Eppendorf 5804R). Next, the cells were washed with 24 mL of ice cold washing solution and centrifuged for another five minutes at 4°C and 10 000 rpm. This washing step was repeated four times, between centrifuging, the cells were always kept on ice. Finally, the cells were resuspended in 1.2 mL of washing solution, distributed into 60 μL aliquots in 1.5 mL microfuge tubes and frozen with liquid nitrogen.

2.4   Standard Molecular Biology Techniques

2.4.1   Plasmid Purification

All plasmid purification steps were done using the Monarch® Plasmid Miniprep Kit according to manufacturers instructions. Eppendorf 5804R centrifuge was used to harvest the cells and Eppendorf 5424R centrifuge was used for all other centrifugation steps. The DNA concentration was measured using the Nanodrop One (Thermo Scientific) and then stored at -20°C.

2.4.2 DNA Concentration Measurements

Double stranded plasmid DNA concentrations were measured using the Nanodrop One from Thermo Scientific. The Nanodrop provides a purity 260/280 ratio of absorbance at 260 nm and 280 nm that can be used to get an estimate of the purity of the DNA or RNA in the sample. Typically an absorbance ratio of approximately 1.8 is desired for DNA and a ratio of approximately 2.0 is generally accepted for RNA. The Nanodrop also provides an absorbance at 230 and 285 nm to estimate protein contamination.

2.4.3 Digestion by Restriction Enzymes

All restriction enzymes and associated buffers used in this study were obtained from New England Biolabs (NEB).

A typical set up for a digestion mix can be seen in Table 2.1 below. Sometimes the enzyme alkaline phosphatase, calf intestinal (CIP) was added to the digestion mix containing the backbone. CIP catalyzes the dephosphorylation of the 5' and 3' ends of DNA phosphomonoesters, and when used for cloning, it prevents religation of linearized plasmid DNA. The amount of DNA digested varied depending on if was used for cloning purposes or if it was just a control digest. The digestion mix was also up-scaled to 20 μL or 50 μL when needed.

Table 2.1: Example of a digestion mix set up for a 10 μL reaction.

Component	Volume (μL)
10x Buffer	1
Restriction Enzyme 1	0.5
Restriction Enzyme 2	0.5
CIP	0.5
DNA	200 - 700 ng
dH$_2$O	Up to 10 μL

2.4.4 Gel Electrophoresis

In this study, 0.8% agarose gels were used, which were made by dissolving 0.8% agarose in 1x TAE buffer with either GelGreen or GelRed dye added (8 μL 10 000x per mL). Once the agarose had solidified, the comb was removed and the samples and ladder (Catalog# N3232L, NEB) were loaded into the wells. Each sample also contained 6x purple loading dye (Catalog#B7024S, NEB). Once the gel was loaded, it was run for 45-90 minutes at 70-90 V (Bio-Rad PowerPac Basic Power Supply) until the dye line had migrated approximately 75-80% of the way down the gel. Once it was finished, the gel was viewed using ChemiDoc XRS+ (Bio-Rad).

2.4.5   Gel Extraction

If the samples from gel electrophoresis were to be used for cloning purposes, then the correct fragments were cut from the gel. In this study, the Zymoclean Gel DNA Recovery Kit and protocol were followed according to the manufacturer. The DNA concentration was measured using the Nanodrop and then stored at -20°C.

2.4.6   Ligation

A ligation mix was prepared containing a molar ratio of 1:3 of vector to insert. An online tool called NEBioCalculator provided by NEB was used to calculate the appropriate molar ratios of backbone and insert required for the ligation mix [54]. The online tool takes into account the size of each fragment and the mass of vector DNA available, and then provides the mass of insert DNA to be added. An example of a ligation mix and ligation control is seen in Table 2.2 below. Each ligation was prepared in a 1.5 mL microfuge tube, all components were added to the tube, then it was mixed and briefly spun down before placing at 16°C overnight. The following day, DH5α was transformed with up to 10 μL of the ligation mix. The transformation procedure is described below in section 2.2.7.

Table 2.2: Example of a ligation mix set up for a 10 μL reaction.

Component	Ligation	Re-ligation Control
10x T4 Ligase Buffer	1 μL	1 μL
T4 Ligase	0.5 μL	0.5 μL
Insert	a ng	No insert
Backbone	b ng	b ng
dH$_2$O	Up to 10 μL	Up to 10 μL

2.4.7   Transformation

Cloning attempts were carried out in DH5α. In this study, there were two types of cells used for transformation; chemically competent cells that were transformed using the heat-shock protocol and then electrocompetent cells which were transformed using the electroporation protocol. Both methods are described below.

In this study, chemically competent DH5α cells were used for the heat shock transformation. The heat shock method is done by first thawing the cells on ice. Once thawed, up to 10 uL of a ligation mix or 100 ng of DNA was added to the thawed cells and incubated on ice for 5-30 minutes. Following the incubation, the cells were placed at 42°C for 45 seconds, and then 900 uL of SOC medium was promptly added directly to the cells. The cells were incubated for 1.5 hours at 37°C and 225 rpm to recover. After the recovery period, 100 uL of the transformation mix was plated out onto LA plates containing the appropriate antibiotics. Ligation mixes were up-concentrated when they

were plated out. Plates were incubated overnight at 37°C.

The electroporation protocol used in this study was adapted from Xue *et al.*[53]. In this study, both electrocompetent *E. coli* DH5α and *B. subtilis* cells were used. The first step was thawing the electrocompetent cells on ice. This was followed by adding the appropriate amount of plasmid DNA to the cells, 500 ng was added to competent *B. subtilis* cells, where as only 100 ng was added to DH5α cells. The mixture was gently stirred in the 1.5 mL microfuge tube and placed on ice for another 5 minutes. Following incubation, the transformation mix was transferred to a pre-chilled 2 mm cuvette. Before placing the cuvette into the pod, it was made sure that all condensation and bubbles were removed. The pulse was administered using the Bio-Rad Gene Pulser Xcell; the settings used are seen in Table 2.3 below. The time constants were then recorded, where *B. subtilis* was within a range of 4.7-5 seconds and *E. coli* around 5 seconds. Following the electric pulse, 1 mL of Outgrowth media was added to *B. subtilis*, or alternatively, 1 mL of SOC media for *E. coli*. Next, the cells were transferred to a 13 mL tube for increased aeration and placed in an incubator for recovery at 37°C and 225 rpm for 4 hours or only 1 hour depending on if the species was *B. subtilis* or *E. coli*, respectively. After the recovery period, the cells were spun down for 2 minutes at 10 000 rpm (Eppendorf 5804R), the supernatant was discarded and the cells were resuspended in the remaining volume. Finally, the cells were plated out onto LA plates containing the appropriate antibiotics and incubated overnight at 37°C.

Table 2.3: Settings used for electroporation. These settings were used for both *B. subtilis* and *E. coli*.

Electroporation Settings	
Voltage (V)	3000
Capacitance (μF)	25
Resistance (Ω)	200
Cuvette (mm)	2

2.5   Bacterial Strains and Plasmids

The bacterial strains and the plasmids used in this study are listed in Table 2.4. The synthetic genes that were inserted into constructs pMSM67.2, pMSM67.3 and pMSM67.4 replaced the *mCherry* gene under the control of P$_{liaI}$ in the construct pMSM67.1.

Table 2.4: Bacterial Strains and plasmids used in this study. The genes marked by an asterisk were codon optimized for expression in *B. subtilis* before sending to GenScript to be synthesized.

Strains/Plasmids	Description	Source
E. coli DH5α	General cloning host. Used as a control during expression.	BSL
B. subtilis 168	Expression host for recombinant proteins.	Novozymes
pMSM67	*xylS* under the control of PliaG promoter. Resistance genes: *bla* and *cat*	[45]
pMSM69	pMSM67 in which P_{liaG} was replaced by P_{met}.	[45]
pMSM67.1	P_{liaI}-mCherry* synthetic sequence digested by NheI-SacI (1 kb) inserted into pMSM67 backbone digested with the same enzymes (8.3 kb).	This study
pMSM67.2	Sigma 32* synthetic sequence digested by SpeI-NotI (0.8 kb) inserted into pMSM67.1 backbone digested with the same enzymes (8.6 kb).	This study
pMSM67.3	Sigma 38* synthetic sequence digested by SpeI-NotI (1 kb) inserted into pMSM67.1 backbone digested with the same enzymes (8.6 kb).	This study
pMSM67.4	Sigma 32-38* synthetic sequence digested by SpeI-NotI (1.8 kb) inserted into pMSM67.1 backbone digested with the same enzymes (8.6 kb).	This study
pMSM69.4	P_{met}* synthetic sequence digested by AvrII-BbvCI (1 kb) inserted into pMSM67.4 backbone digested with the same enzymes (10.3 kb).	This study
pVB-1A0B1-mCherry	Medium copy number plasmid. Reporter gene mCherry expressed from Pm. XylS controlled by native promoters P_{s1} and P_{s2}. Resistance gene: *kan*	Vectron Biosolutions As

2.6 Protein Expression

2.6.1 Recombinant Protein Expression

First an overnight culture of each sample was prepared from glycerol stocks by inoculating 10 mL of LB with the appropriate antibiotic in a 125 mL erlenmeyer flask and placed in the incubator at 37°C at 225 rpm. The following morning, OD_{600} measurements were taken, and a 250 mL baffled flask containing 50 mL of LB was re-inoculated to a starting OD_{600} of 0.05. Once the OD_{600} reached 0.5-0.7, the samples were induced. 2 mM of *m*-toluate was used to induce any plasmid containing the P_m promoter and 30 μg/mL of bacitracin was used to induce any plasmid containing the P_{liaI} promoter. Following induction, the cultures were placed at 30°C for the remainder of the expression. Details on how these samples were processed are described in the sections below. At each of these time points, OD_{600} measurements were taken along with mCherry fluorescence readings. mCherry readings were taken with the TECAN Infinite 200 Pro microplate reader.

2.6.2   mCherry Fluorescence Readings

mCherry fluorescence was detected using the TECAN Infinite 200 Pro microplate reader. The TECAN provides one with arbitrary units of RFU, which can then be normalized against the cell density (OD_{600}). These readings needed to be carried out in a black, flat bottomed 96-well plate (Thermo Fisher Scientific), where a 200 μL working volume was used. Each fluorescent protein has differing excitation and emission maximal wavelengths. For mCherry, the excitation wavelength was set at 580 nm and the emission wavelength was set at 615 nm. Specific settings used for the TECAN are found in Appendix D.

2.6.3   Protein Isolation

Note that the samples were always kept on ice until the lysis step.

After withdrawing 5 mL of culture, the cells were harvested by centrifuging for 5 minutes, 4°C and 10 000 rpm (Eppendorf 5804R). The supernatant was discarded and the remaining liquid was removed with a pipette. One milliliter of 0.9% NaCl solution was added per 100 mg of pellet. The pellets were vortexed and then 1 mL of the suspension was transferred to a 1.5 mL microfuge tube which was then centrifuged for 5 minutes at 4°C and 10 000 rpm (Eppendorf 5424R). After the centrifugation step, the supernatant was vacuumed and the cells were lysed for 30 minutes at room temperature at 100 rpm on a rocking table. Lysis of *B. subtilis* and *E. coli* was performed by resuspending the pellets in 500 μL of lysis solution 1 and 2, respectively (Table 2.5).

Table 2.5: Lysis Buffer solutions used for *B. subtilis* (Solution 1) and *E. coli* (Solution 2).

Lysis Solution 1		Lysis Solution 2	
Component	Volume	Component	Volume
Cell Lytic B (mL)	1.0	Cell Lytic B (mL)	5.0
Protease Inhibitors (μL)	10.0	Benzonase Nuclease (μL)	1.0
Benzonase Nuclease (μL)	0.05		
Lysozyme (mg)	0.2		

Following the incubation period, the samples were centrifuged again for another 5 minutes at 4°C and 10 000 rpm. The supernatant from each pellet, which is now referred to as the soluble fraction, was transferred to individual wells in a 800 μL deepwell plate. Then the remaining pellets was resuspended in 500 μL of 1x TGX SDS-running buffer (Bio-Rad). The resuspended pellets, which are now referred to as the insoluble fraction, were also transferred to separate wells into the same 800 μL deepwell plate. The soluble and insoluble fractions were not diluted and were prepared for SDS-PAGE, which is outlined in the protein analysis section below.

2.6.4   Protein Analysis

In this study, western blotting was used to verify that the constructs pMSM67.2, pMSM67.3, pMSM67.4 and pMSM69.4 were producing the sigma factors σ^{32} and σ^{38} from the promoter P_{liaI}.

SDS-PAGE
In a 96-well PCR plate, the samples were subjected to reducing conditions, so the soluble and insoluble fractions were mixed with 2x Laemmli Sample buffer (Bio-Rad) containing 50 mM DTT and then were incubated for 5 minutes at 95°C in a PCR machine (Bio-Rad C1000 Touch Thermal Cycler). Two precast gels (Bio-Rad 12% Criterion TGX) were loaded with 10-15 μL of each sample, and 5 μL of the protein ladder. The gels were run for 40-45 minutes at 200 V. Then, one of the gels was stained in InstantBlue (Expedeon), a Coomassie blue based staining solution, for 1 hour to overnight before being destained in dH$_2$O and imaged using the ChemiDoc XRS+ (Bio-Rad). The second gel was used for western blotting.

Western Blotting
The second SDS-PAGE gel that was run was used for western blotting. After the gel was run, the transfer sandwich was prepared by placing the gel on top of the PVDF membrane, then layering the stack of prepared filter papers on top and finally, placing the completed sandwich in the cassette (Trans-Blot$^{\circledR}$ Turbo^{TM} Midi PVDF Transfer Packs from Bio-Rad). The proteins were transferred to the membrane using the Trans-Blot Transfer System (Bio-Rad) with the settings set to 2.5 A and 25 V for 7 minutes.

Next, the PVDF membrane was transferred to the iBind Flex device (Invitrogen iBind Flex Western System), where the primary and secondary antibodies, along with the iBind Flex solution were added in sequential order according to the accompanying protocol. Compositions of the iBind Flex solution, primary and secondary antibodies are outlined in Tables 2.6, 2.7 and 2.8.

Table 2.6: iBind Flex required for 1 midi membrane.

iBind Flex Solution	(mL)
5x buffer	12
100X additive	0.6
water	47.4
Total	60

The membrane was left in the iBind Flex device for 2.5 hours up to as long as overnight, until the last row was empty of iBind Flex solution. Then, the membrane was rinsed in dH$_2$O before being developed with 2 mL of TMB for 2-10 minutes, the reaction was neutralized by adding dH$_2$O. Finally, the western blot was imaged using the ChemiDoc XRS+ (Bio-Rad).

Table 2.7: Primary Antibodies used for Sigma Factor detection for Western Blot. The antibody was added to 4 mL of iBind Flex solution.

Primary Antibodies	Provider	Dilution
mAb to Sigma 32 subunit of E. coli RNA polymerase	My Biosource	1/1000
Anti-E. coli RNA Sigma S	BioLegend	1/1000

Table 2.8: Secondary Antibody used for Sigma Factor Detection for Western Blot. The antibody was added to 4 mL of iBind Flex Solution.

Secondary Antibody	Provider	Dilution
Polyclonal Rabbit Anti-Mouse Immunoglobulins HRP	Dako	1/400

2.7   RNA Isolation and mRNA Quantification

2.7.1   RNA Isolation

The RNAqueos® Total RNA Isolation Kit from Invitrogen was used for all RNA isolation work, along with the provided protocol with some minor adaptations to the lysis procedure of the *B. subtilis* samples. The *B. subtilis* samples were taken from the -80°C, then thawed at room temperature before adding 40 mM EDTA and 2 mg/mL of lysosyme in a total volume of 600 μL and incubated for 1 hour at 37°C. Following the incubation, the cells were harvested by centrifuging for 3 minutes at 12 000 rcf (Eppendorf 5424R). The supernatant was decanted and the pellet was resuspended by vigorous vortexing in 300 μL of lysis buffer. The remainder of the protocol was followed as recommended by the manufacturers. The RNA concentration was measured using the Nanodrop, then the RNA samples were aliquoted and stored at -80°C.

2.7.2   Bioanalyzer

The 2100 Bioanalyzer Instrument (Agilent) at Sintef was used to analyze the integrity of the RNA samples. The Agilent RNA 6000 Nano Kit Quick Start Guide was followed as recommended with regards to the gel, gel-dye mix and sample preparations as well as loading the RNA 6000 Nano chip for analysis. RNaseZAP (Sigma) was used as required.

At times when the Bioanalyzer was unavailable, the quality of the RNA samples were tested by running them on a 0.8% agrose gel containing Gelred dye. 1 μL of the RNA sample was mixed with 4 μL of 1x loading dye (NEB), a 1 kb ladder was loaded into the first well and then run for 50 minutes at 80 volts. For the RNA to be deemed good quality, two clear bands should appear, which are the 16S and 23S rRNA molecules, and there should be no smearing or fragmentation of the RNA, which would have been indicative of RNA degradation. When RNA is run on an agarose gel, it does not run exactly according to its size because it is single stranded and the secondary structure can be variable. To get a more accurate size, one should use a RNA denaturing gel and

compare it to a RNA standard instead of DNA standard.

2.7.3 cDNA Preparation

The SuperScript^{TM} IV VILO^{TM} Master Mix provided by Invitrogen and protocol were used to make cDNA. The reverse transcription containing ezDNase enzyme protocol was followed. At the end of the incubation periods, the cDNA concentration was measured using the Nanodrop with the dsDNA function, and then was stored at -20°C for up to one week or at -80°C for long term storage.

2.7.4 qPCR and mCherry mRNA Quantification

TaqMan$^®$ reagents were used for qPCR in this study. Please see Appendix E for the sequences of the primers and probes which were designed for use with genes optimized for *B. subtilis*. The passive reference dye used was ROX, FAM was used as the reporter dye and TAMRA was used as the quenching dye. All qPCR reactions were set up on ice.

Primer and Probe Design
The primers were designed by Clone Manager using the guidelines from ThermoFisher Scientific which were to be used with TaqMan$^®$ chemistry. It was recommended that the amplicon length was 50-200 bp to promote efficient amplification. When designing the primers, the optimal length is 20 bp long and the Tm should be between 58-60°C. For both the primers and probe, the G/C content should remain within the range of 30-80%.

Amplification Efficiency
First the function and amplification efficiency of the primers and probes had to be tested. The wt *B. subtilis* strain was used to test the 16 S primers and probe, and *E. coli* harboring pMSM67 was used to test the efficiency of the mCherry primers and probe. A 10-fold serial dilution down to 10^{-6} was made from the cDNA in triplicate of each sample to be used as the template in the qPCR reaction. Next mastermixes were made, one for each primer/probe set. Table 2.9 shows the required component volumes per reaction.

Table 2.9: qPCR mix per reaction. The figures in parentheses under the components refer to the final concentrations.

Component	Volume (μL)
TaqMan Fast Advanced MM (1X)	10
Foward Primer (500nM)	2
Reverse Primer (500nM)	2
Probe (250nM)	2
cDNA template (1pg - 100 ng)	2
Nuclease-free water	2
TOTAL	20

The mastermixes were briefly vortexed and centrifuged to bring the reaction mix to the bottom of the eppendorf tube. Then, 18 μL of the mastermix was aliquoted in the wells of a MicroAmp$^{\text{TM}}$ Optical 96-Well Reaction Plate. Following this, 2 μL of the serial dilutions of cDNA template was added to the appropriate wells. Then 2 μL of nuclease-free water was added to the no template control wells. The reaction plate was sealed with optical adhesive film, then centrifuged briefly before running the qPCR reaction using custom settings in the QuantStudio$^{\text{TM}}$ 5 Real-Time PCR instrument. After the run was finished, a standard curve with dilutions vs. Ct was plotted to calculate the amplification efficiency. The formula is listed as equation 1 below.

$$\text{Efficiency} = (10^{-1/slope} - 1) \times 100 \tag{1}$$

Comparative Ct
Next, the samples were run on the QuantStudio$^{\text{TM}}$ 5 Real-Time PCR instrument for the comparative Ct analysis. Each sample was run in triplicate, testing for mCherry mRNA, where 16 S was used as the endogenous control, which was also included in triplicate for each sample. A no template control (ntc) was also added in the experimental set up. The MicroAmp$^{\text{TM}}$ Optical 96-Well Reaction Plate and mastermixes were prepared in the same manner as when the primers were tested. Based on the amplification efficiency plots, it was decided to use a template concentration of the 10^{-2} dilution.

The fold difference or relative quantification (RQ) was provided from the qPCR experiment set with the default parameters, taken from the gene expression plot. There were three technical replicates per sample, where the uninduced pMSM67 sample was used as the reference sample, and the endogenous control was the housekeeping *B. subtilis* 16S rRNA gene.

$$\text{Fold Difference} = 2^{-\Delta\Delta C_{\text{t}}} \tag{2}$$

3 Results

3.1 Sequence Alignment Comparing Sigma Factors

Even though the *E. coli* sigma factors σ^{32} and σ^{38} have functionally similar sigma factors produced in *B. subtilis*, it was decided to perform a sequence alignment to compare the extent of similarity between the sigma factors in the two bacterial species. These sigma factors are part of group 3 of the σ^{70} family, except σ^{38} which belongs to group 2. Since sigma factor, σ^{32} in *E. coli* was functionally similar to the *B. subtilis* sigma factors σ^{H} and σ^{B}, protein sequence alignments were run amongst them. Then a protein sequence alignment between *E. coli* σ^{38} and *B. subtilis* sigma factor σ^{B} was run since they were also functionally similar. σ^{32} was compared separately to both the *B. subtilis* sigma factors σ^{H} and σ^{B} which is shown in Figure 3.1, panels a) and b) below respectively. The protein sequence alignment between σ^{32} and σ^{H} reported only 17.23% identity, with 51 identical positions and 77 similar positions. The protein sequence alignment between σ^{32} and σ^{B} reported 22.03% identity with 65 identical positions and 87 similar positions. The third protein sequence alignment between σ^{38} and *B. subtilis* sigma factor, σ^{B} (Figure 3.1, panel c)), reported 18.81% sequence identity with 63 identical positions and 103 similar positions. Even though there exists functionally similar sigma factors between *E. coli* and *B. subtilis*, from these results, it can be concluded that they are not similar when it comes to their protein sequences.

```
P17869 RPSH_BACSU    A.    1    --------------MNLQNNKGKFNKEQ-FCQLEDEQVIEKVH-VGDSDALDYLITKYRNF      45
P0AGB3 RPOH_ECOLI          1    MTDKMQSLALAPVGNLDSYIRAANAWPMLSADEERALAEKLHYHGDLEAAKTLILSHLRF      60
                                           **:.     *       :.  *:. : **:*  ** :* . **  .: .*

P17869 RPSH_BACSU         46    VRAKARSYFLIGADREDIVQEGMIGLYKSIRDFKEDKLTSFKAFAELCITRQIITAI---     102
P0AGB3 RPOH_ECOLI         61    VVHIARNYAGYGLPQADLIQEGNIGLMKAVRRFNPEVGVRLVSFAVHWIKAEIHEYVLRN     120
                                *    **.*    *  : *;;*** *** *::* *: :  .: :** *. :*   :

P17869 RPSH_BACSU        103    ------KTATRQKHIPLNSYASLDK-PIFDEESDRTLLDVISGAKTLNPEEMIINQEEFD     155
P0AGB3 RPOH_ECOLI        121    WRIVKVATTKAQRKLFFNLRKTKQRLGWFNQDE-VEMVA----------RELGVTSKDVR     169
                                      *:. *::: :*    :*     *:::.  ::          .*: :..:.

P17869 RPSH_BACSU        156    DIEMKMGE----------LLSDLERKVLVLYLDGRSYQEISDELN------RHVKSIDNA     199
P0AGB3 RPOH_ECOLI        170    EMESRMAAQDMTFDLSSDDDSDSQPMAPVLYLQDKSSNFA-DGIEDDNWEEQAANRLTDA     228
                                ::* :*.           ** :  . ****:.:* :  * ::       : ..: :  :*

P17869 RPSH_BACSU        200    LQRVKRKLEK-----YLEIREISL----------------------------------     218
P0AGB3 RPOH_ECOLI        229    MQGLDERSQDIIRARWLDEDNKSTLQELADRYGVSAERVRQLEKNAMKKLRAAIEA      284
                                :* :..: :.      :*:  : *
```

```
P0AGB3 RPOH_ECOLI    B.    1    MTDKMQSLALAPVGNLDSYIRAANAWPMLSADEERALAEKLHYHGDLEAAKTLILSHLRF      60
P06574 RPSB_BACSU          1    ------------------MTQPSKTTKLTKDEVDRLISDYQTKQDEQAQETLVRVYTNL       41
                                               :  .    *: ** *  .. : * :* :**:  : .:

P0AGB3 RPOH_ECOLI         61    VVHIARNYAGYGLPQADLIQEGNIGLMKAVRRFNPEVGVRLVSFAVHWIKAEIHEYVLR-     119
P06574 RPSB_BACSU         42    VDMLAKKYSKGKSFHEDLRQVGMIGLLGAIKRYDPVVGKSFEAFAIPTIIGEIKRFLRDK     101
                                *  :*::*:    : ** * * ***: *::*::* ** :  :**:..  .**:.::

P0AGB3 RPOH_ECOLI        120    NWRIVKVATTKAQ-RKLFFNLRKTKQRLGWFNQDEVEMVARELGVTSKDVR-EMESRMAA     177
P06574 RPSB_BACSU        102    TWSVHVPRRIKELGPRIKMAVDQL--TTETQRSPKVEEIAEFLDVSEEEVLETMEMGKSY     159
                                .* :    *      :: : : :       .. :** :*. *.*:.::*   **    :

P0AGB3 RPOH_ECOLI        178    QDMTFDLSSDDDSDSQPMAPVLYLQDKSSNFADGIEDDNWEEQAA-NRLTDAMQGLDERS     236
P06574 RPSB_BACSU        160    QALSVDHSIEADSDGSTVTI----LDIV-----GSQEDGYERVNQQLMLQSVLHVLSDRE     210
                                * ::.* * : ***.. ::        *       * ::*.:*.        * ..:: *.:*.

P0AGB3 RPOH_ECOLI        237    QDIIRARWLDEDNKSTLQELADRYGVSAERVRQLEKNAMKKLRAAIEA-------      284
P06574 RPSB_BACSU        211    KQIIDLTYIQNKS---QKETGDILGISQMHVSRLQRKAVKKLREALIEDPSMELM      262
                                ::**  :::::..    :* .* *:* :* :*::::*:**** *:
```

```
P13445 RPOS_ECOLI    C.    1    MSQNTLKVHDLNEDAEFDENGVEVFDEKALVEQEPSDNDLAEEEELLSQGATQRVLDATQL      60
P06574 RPSB_BACSU          1    ------------------------------MTQPSKT-------------------TK-       9
                                                              :**..                     *:

P13445 RPOS_ECOLI         61    YLGEIGYSPLLTAEEEVYFARRALRGDVASRRRMIESNLRLVVKIARRYGNRGLALLDLI     120
P06574 RPSB_BACSU         10    ---------LTKDEVDRLISDYQTKQDEQAQETLVRVYTNLVDMLAKKYSKGKSFHEDLR      60
                                         * . * :  ::      :  *  ::. ::.    .** :*::*.:        **

P13445 RPOS_ECOLI        121    EEGNLGLIRAVEKFDPERGFRFSTYATWWIRQTIERAIMNQTRTIRLPIHIVKELNVYLR     180
P06574 RPSB_BACSU         61    QVGMIGLLGAIKRYDPVVGKSFEAFAIPTIIGEIKRFLRDKTWSVHVPRRIKELGPRIKM     120
                                : * :**: *:::**   *  *.::*     *:* :::* :::::* :* :

P13445 RPOS_ECOLI        181    TARELSHKLDHEPSAEEIAEQLDKPVDDVSRMLRLNERI--TSVDTPLGGDSEKA---LL     235
P06574 RPSB_BACSU        121    AVDQLTTETQRSPKVEEIAEFLDVSEEEVLETMEMGKSYQALSVDHSIEADSDGSTVTIL     180
                                :. :*: : ::.*.,.***** **   ::* . :.:.:     *** : .**: :    :*

P13445 RPOS_ECOLI        236    DILADEKENGPEDTTQDDDMKQSIVKWLFELNAKQREVLARRFGLLGYEAATLEDVGREI     295
P06574 RPSB_BACSU        181    DIVGS-QEDGYERVNQ-QLMLQSVLHV---LSDREKQIID----LTYIQNKSQKETGDIL     231
                                **:.. :*:* *  .* : * **::;         * : : ::.*  :

P13445 RPOS_ECOLI        296    GLTRERVRQIQVEGLRRLREILQTQGLNIEALFRE      330
P06574 RPSB_BACSU        232    GISQMHVSRLQRKAVKKLREALIEDP-SMELM---      262
                                *:::: :*  ::*  :.:::*** *  :  .:* :
```

Figure 3.1: Protein sequence alignment comparing the similarity between *E. coli* and *B. subtilis*. A. Alignment comparing *E. coli* K12 σ^{32} (Accession# P0AGB3) to *B. subtilis* 168 σ^{H} (Accession# P17869). Only 17.23% identity between the two sigma factors. B. Alignment comparing *E. coli* K12 σ^{32} (Accession# P0AGB3) to *B. subtilis* 168 σ^{B} (Accession# P06574). Their protein sequence alignment was only 22.03% identical. C. Protein alignment comparing *E. coli* K12 σ^{38} (Accession# P13445) to *B. subtilis* 168 σ^{B} (Accession# P06574). Their protein sequence alignment reported only 18.81%. Sequence alignment was run on UniProt database using the Clustal Omega Alignment set with the default parameters. Asterisk represents identical residues, single dot represents very similar residue, two dots represent a somewhat similar residue.

24

3.2   Testing Functionality of P*liaI* promoter in *B. subtilis*

The native *B. subtilis* P_{liaI} promoter was used to express sigma factors, so first, the functionality of this bacitracin inducible promoter had to be tested in *B. subtilis* by having the reporter gene, mCherry, fused to it. An expression experiment was run and mCherry fluorescence was measured 18 hours post induction. The culture harboring the construct pMSM67.1 was induced with a final bacitracin concentration of 50 μg/mL, and the wild type *B. subtilis* culture was used as a negative control (3.2).
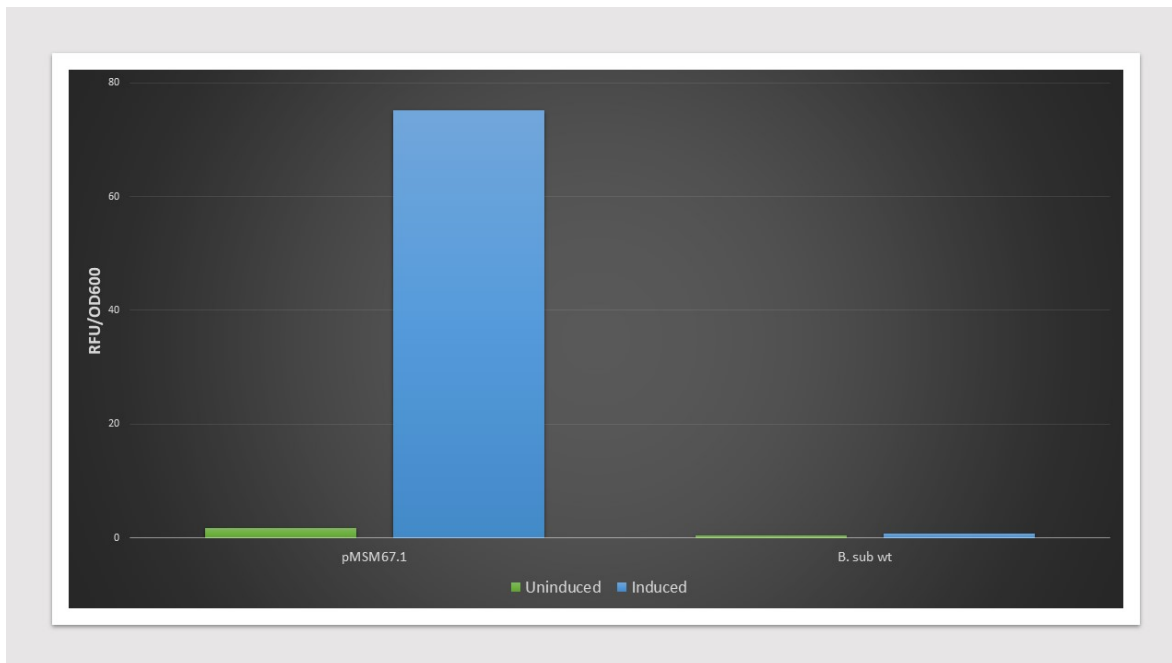

Figure 3.2: The amount of mCherry fluorescence produced from bacitracin inducible promoter P_{liaI} in pMSM67.1 construct in *B. subtilis*. Wild type *B. subtilis* was used as negative control. Each sample contains an induced and uninduced variant. Expression carried out in LB media and the culture was induced with 50 μg/mL and placed at 30°C post induction. The mCherry RFU was normalized against the OD_{600} of the culture.

Even though there was not much mCherry fluorescence detected from pMSM67.1, this expression confirmed that the P_{liaI} bacitracin inducible promoter was functional in *B. subtilis*.

3.3   Expressing Sigma Factors from P*liaI*

B. subtilis was transformed with the vectors pMSM67.2, pMSM67.3, pMSM67.4 and pMSM69.4. Through expression experiments, it was to be verified that they produced sigma factors σ^{32} and σ^{38} from the promoter P_{liaI}. At the same time, any fluorescence that was produced from mCherry expressed from P_m was measured. These results are reported in the sections below.

3.3.1 Coexpression of Sigma Factors in *B. subtilis*

An expression experiment in *B. subtilis* was performed where both promoters, P_m and P_{liaI} were induced, primarily to check if the sigma factors have an effect on the mCherry expression from P_m, and secondly to verify the presence of sigma factors from P_{liaI} in the pMSM67 vector series. In Figure 3.3, there is only mCherry fluorescence visible from the induced positive control, pVB-1A0B1-mCherry. It appears that there is no significant difference between the levels of mCherry expressed at 5 hours post induction and 20 hours post induction. In the three other induced cultures, pMSM67.2, pMSM67.3 and pMSM67.4, there was less than 30 RFU/OD_{600} of mCherry fluorescence detected.
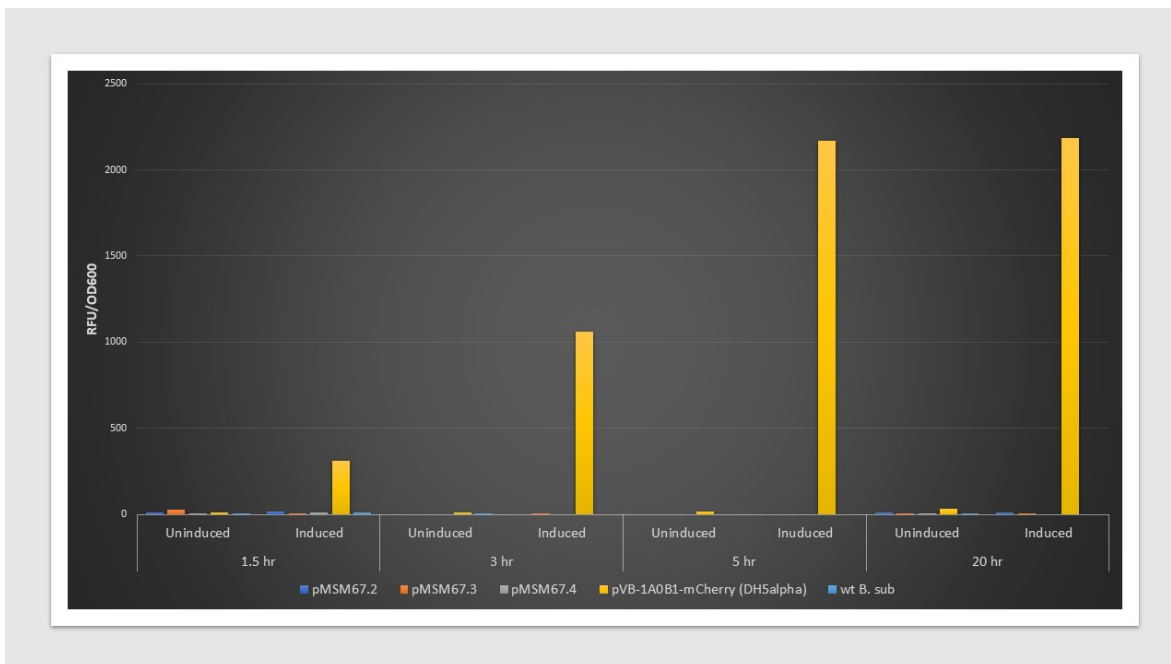

Figure 3.3: Comparing mCherry expression from P_m with sigma factors being expressed simultaneously from P_{liaI} between pMSM67 vector series expressed in *B. subtilis*. pVB-1A0B1-mCherry expressed in DH5α is acting as the positive control and wild type *B. subtilis* is the negative control. Each sample has an uninduced and induced variant. Expression was done in LB media, cultures placed at 30°C post induction. Cultures were induced with 1mM m-toluate and/or 30 μg/mL bacitracin. mCherry fluorescence measurements taken with the TECAN. The RFU of each culture was normalized against its OD.

Another expression experiment was completed once the pMSM69.4 construct was received from Genscript and *B. subtilis* was transformed with it. This expression was performed in *B. subtilis* and evaluated the mCherry expression from P_m between cultures harboring pMSM67.4, pMSM69.4, pMSM67 and pMSM69. The plasmids pMSM67.2 and pMSM67.3 were excluded from this experiment since it was previously verified that both sigma factors were successfully being expressed from pMSM67.4, so there was no need to include them in this round. Due to this reason is also why only pMSM69.4 was constructed and not pMSM69.2 or pMSM69.3. Since vectors pMSM67.4 and pMSM69.4 express different levels of XylS, it was decided to evaluate if *B. subtilis* required a higher level of XylS to initiate transcription from P_m. The three other samples; wild type *B.*

subtilis, pMSM67.1 and pMSM67 in *E. coli* were to be used as controls for qPCR. There were sample measurements taken at 4 different time points post induction, as well as samples collected for RNA isolation and protein expression. mCherry fluorescence that was detected from this experiment is depicted below in Figure 3.4. There is low levels of mCherry fluorescence seen from the induced *B. subtilis* culture harboring pMSM67.1 across the 4 time points post induction, with the highest mCherry fluorescence detected at 3 hours post induction with 113 RFU/OD$_{600}$. Then there are large amounts of mCherry fluorescence seen from the induced sample of pMSM67 in *E. coli*, with the most mCherry fluorescence appearing 20 hours post induction with 8029 RFU/OD$_{600}$. There is little to no mCherry fluorescence detected from the other five constructs, with their RFU/OD$_{600}$ values ranging between zero and four, therefore their data labels were not included in the Figure.

Figure 3.4: Measure of mCherry expression from P$_m$ in *B. subtilis* across four time points post induction. The wild type *B. subtilis*, pMSM67.1 and pMSM67(*E. coli*) were used as controls. Cultures were expressed in LB, induced with 2 mM m-toluate and/or 30 μg/mL bacitracin, then placed at 30°C post induction.mCherry fluorescence measurements taken with TECAN plate reader. The RFU of each culture was normalized against its OD.

3.3.2 Verification of Sigma Factors via Western Blotting

The production of sigma factors from the P$_{liaI}$ promoter in the pMSM67 vectors was confirmed by running an SDS-PAGE gel and then a western blot. Figure 3.5 below shows that σ^{32} is being expressed from the induced cultures harboring pMSM67.2, pMSM67.4 and the positive control, pVB-1A0B1-mCherry. As the time post-induction increases, there is a decrease of the amount of σ^{32} being detected on the western blot. From the 1.5 hour time point post-induction, there is σ^{32} present in both the soluble and insoluble fractions of the induced cultures of pMSM67.2 and pMSM67.4; there are

also two bands noticeable in these samples, one around 32 kDa and one slightly larger than 25 kDa.

Figure 3.5: Western blot of σ^{32} being expressed from pMSM67.2 and pMSM67.4 from P$_{liaI}$ promoter. Samples taken at 4 different time points. After harvesting, lysates were separated into soluble (S) and insoluble (I) fractions and 10 μL of undiluted lysate was loaded onto the gel. pVB-1A0B1-mCherry is positive control. Lane A: pMSM67.2 In/S; lane B: pMSM67.2 In/I; lane C: pMSM67.2 Un/S; lane D: pMSM67.2 Un/I; lane E: pMSM67.4 In/S; lane F: pMSM67.4 In/I; lane G: pMSM67.4 Un/S; lane H: pMSM67.4 Un/I; lane I: pVB-1A0B1-mCherry In/S; lane J: pVB-1A0B1-mCherry In/I; lane K: pVB-1A0B1-mCherry Un/S; lane L: pVB-1A0B1-mCherry Un/I. Ladder used is Precision Plus Protein Standards Dual Color from Bio-Rad (kDa). Induced (In), Uninduced (Un).

A western blot was also run to verify σ^{38} production from constructs pMSM67.3 and pMSM67.4. In Figure 3.6, there is σ^{38} protein detected in the induced cultures of pMSM67.3 and pMSM67.4 in both soluble and insoluble fractions, and in the soluble fraction of the induced culture of the positive control, pVB-1A0B1-mCherry. As noticed with σ^{32}, there is a decreasing trend over the four different time points, where there is little to no σ^{38} detected 20 hours post induction. There are also two bands visible, one at the expected 38 kDa marker and then a second one around the 25 kDa marker.
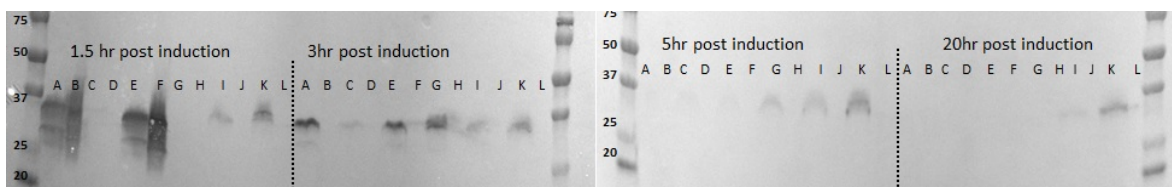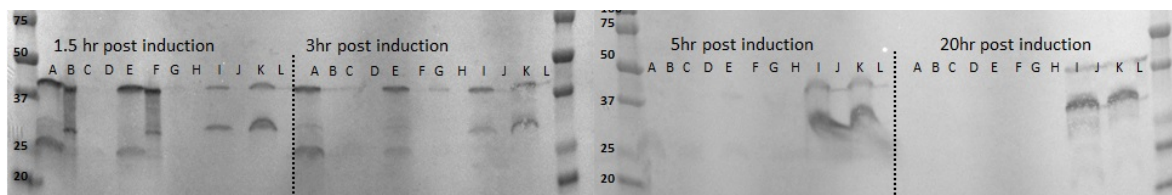

Figure 3.6: Western blot of σ^{38} being expressed from pMSM67.3 and pMSM67.4 from P$_{liaI}$ promoter. Samples taken at 4 different time points. After harvesting, lysates were separated into soluble (S) and insoluble (I) fractions and 10 μL of undiluted lysate was loaded onto the gel. pVB-1A0B1-mCherry is positive control. Lane A: pMSM67.3 In/S; lane B: pMSM67.3 In/I; lane C: pMSM67.3 Un/S; lane D: pMSM67.3 Un/I; lane E: pMSM67.4 In/S; lane F: pMSM67.4 In/I; lane G: pMSM67.4 Un/S; lane H: pMSM67.4 Un/I; lane I: pVB-1A0B1-mCherry In/S; lane J: pVB-1A0B1-mCherry In/I; lane K: pVB-1A0B1-mCherry Un/S; lane L: pVB-1A0B1-mCherry Un/I. Ladder used is Precision Plus Protein Standards Dual Color from Bio-Rad (kDa). Induced (In), Uninduced (Un).

Protein purification and analysis was done with the uninduced and induced *B. subtilis* cultures harboring pMSM69.4 to verify the production of σ^{32} and σ^{38} from P$_{liaI}$ promoter. Below in Figure 3.7, there is σ^{32} being expressed from pMSM69.4 at 1.5 and 3 hours post induction in the induced samples, it is detected in both the soluble and insoluble fractions. There is no σ^{32} protein detected from the other 2 time points, or the positive control, pVB-1A0B1-mCherry. Lane U and V contained the σ^{32} protein standard, 50 ng and 100 ng respectively, which served as positive controls for the antibodies.

Figure 3.8 is the western blot verifying the expression of σ^{38} protein from pMSM69.4. There is σ^{38} expressed mostly in the induced samples, appearing in both the soluble and

insoluble fractions during 1.5 and 3 hours post induction. There are very faint bands visible for the other 2 time points post induction. There is σ^{38} also detected in the soluble fractions of the induced cultures of the positive control, pVB-1A0B1-mCherry.

Figure 3.7: Expression of σ^{32} from pMSM69.4 construct over 4 time points; 1.5, 3, 5 and 20 hours post induction. After harvesting, lysates were separated into soluble (S) and insoluble (I) fractions and 10 μL of undiluted lysate was loaded onto the gel. Lane A: 1.5 hr S/In; lane B: 1.5 hr I/In; lane C: 1.5 hr S/Un; lane D: 1.5 hr I/Un; lane E: 3 hr S/In; lane F: 3 hr I/In; lane G: 3 hr S/Un; lane H: 3 hr I/Un; lane I: 5 hr S/In; lane J: 5 hr I/In; lane K: 5 hr S/Un; lane L: 5 hr I/Un; lane M: 20 hr S/In; lane N: 20 hr I/In; lane O: 20 hr S/Un; lane P: 20 hr I/Un; lane Q: 20 hr pVB-1A0B1-mCherry S/In; lane R: 20 hr pVB-1A0B1-mCherry I/In; lane S: 20 hr pVB-1A0B1-mCherry S/Un; lane T: 20 hr pVB-1A0B1-mCherry I/Un; lane U: σ^{32} standard 50 ng; lane V: σ^{32} standard 100 ng. Ladder used is Precision Plus Protein Standards Dual Color from Bio-Rad (kDa). Induced (In), Uninduced (Un).
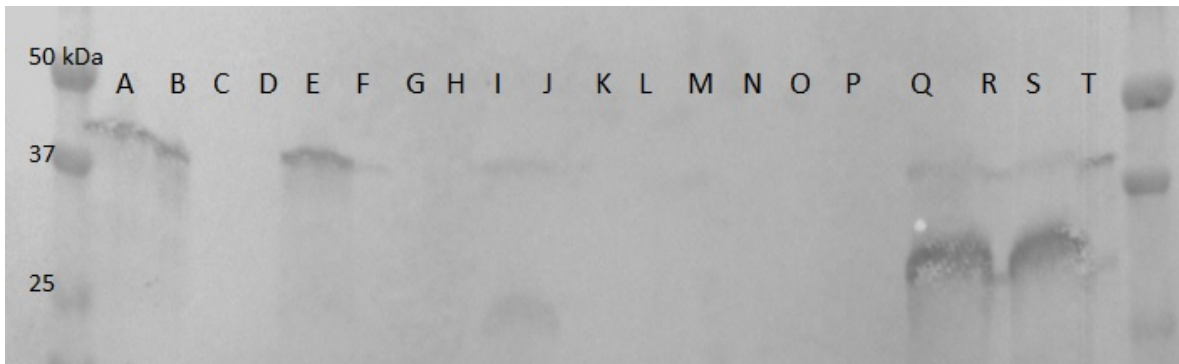

Figure 3.8: Expression of σ^{38} from pMSM69.4 construct over 4 time points; 1.5, 3, 5 and 20 hours post induction. After harvesting, lysates were separated into soluble (S) and insoluble (I) fractions and 10 μL of undiluted lysate was loaded onto the gel. Lane A: 1.5 hr S/In; lane B: 1.5 hr I/In; lane C: 1.5 hr S/Un; lane D: 1.5 hr I/Un; lane E: 3 hr S/In; lane F: 3 hr I/In; lane G: 3 hr S/Un; lane H: 3 hr I/Un; lane I: 5 hr S/In; lane J: 5 hr I/In; lane K: 5 hr S/Un; lane L: 5 hr I/Un; lane M: 20 hr S/In; lane N: 20 hr I/In; lane O: 20 hr S/Un; lane P: 20 hr I/Un; lane Q: 20 hr pVB-1A0B1-mCherry S/In; lane R: 20 hr pVB-1A0B1-mCherry I/In; lane S: 20 hr pVB-1A0B1-mCherry S/Un; lane T: 20 hr pVB-1A0B1-mCherry I/Un. Ladder used is Precision Plus Protein Standards Dual Color from Bio-Rad (kDa). Induced (In), Uninduced (Un).

3.4 Amplification Efficiency

Before setting up the qPCR assay with the different samples, the efficiency of the primers had to be tested. The primer pairs used were specific for the target, *B. subtilis* codon optimized mCherry, and the endogenous controls; *B. subtilis* 16S rRNA and *E. coli* 16S rRNA genes. The amplification efficiency of the three sets of primers were tested by creating a standard curve. From the slope, the amplification efficiency can be calculated with Equation 1. The amplification efficiency for the mCherry primers

was 100%, the *B. subtilis* 16S primer amplification efficiency was 111%, finally, and the amplification efficiency for of *E. coli* 16S primers was 168%. The optimal amplification efficiency ranges between 90% to 110%.
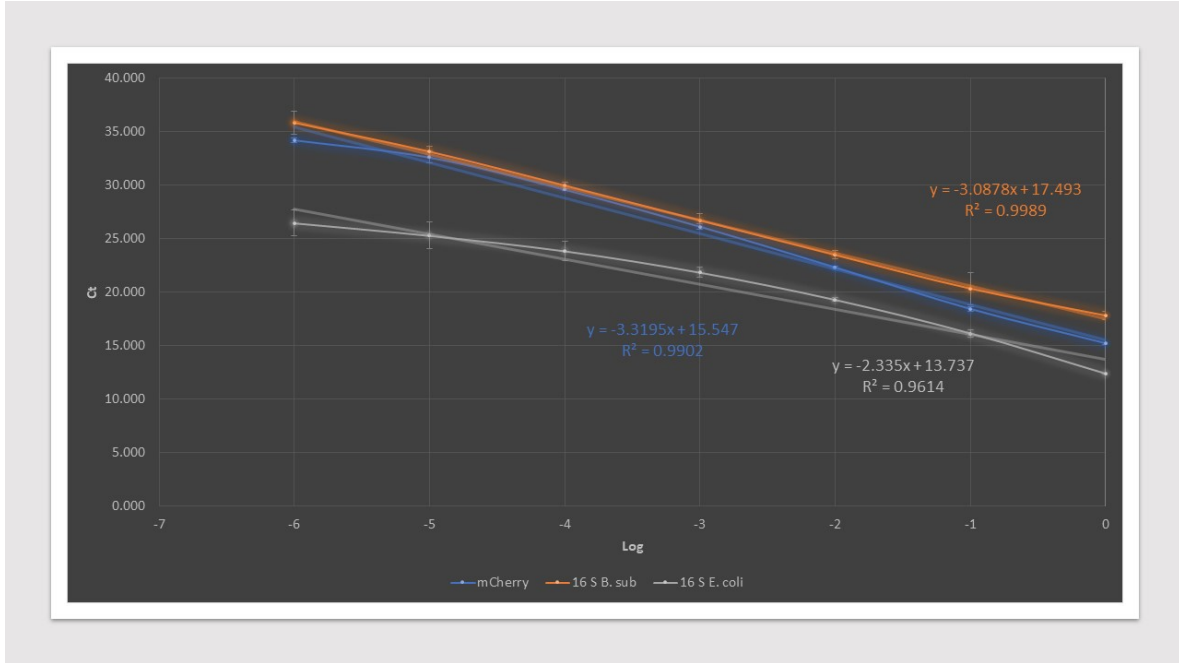

Figure 3.9: Standard curve to evaluate qPCR primer efficiency for the mCherry, *B. subtilis* 16S and *E. coli* 16S primer pairs. A 10-fold serial dilution of the template was prepared in triplicate. Error bars represent standard deviation of measurements for three technical replicates. TaqMan® reagents were used and program run with the Applied Biosystems QuantStudio 5 RT-PCR system.

3.5   Measuring mCherry Transcript Levels from Pm via qPCR

The fold change of mCherry transcript was evaluated comparing each induced or uninduced sample to the uninduced culture of pMSM67. These samples were normalized against the transcript levels of *B. subtilis* 16S rRNA gene, which was the endogenous control. The fold change was calculated for induced and uninduced cultures of pMSM67.4, pMSM69.4, pMSM67 and pMSM69. However, it was not calculated for the samples pMSM67.1, wild type *B. subtilis* and pMSM67 (*E. coli*) since these were used as experimental controls. The pMSM67.1 and pMSM67 (*E. coli*) express codon optimized mCherry for *B. subtilis* and served as positive controls for the mCherry primers. The wild type *B. subtilis* sample served as a negative control to show that the mCherry primers did not detect any endogenous sequences. Looking at Figure 3.10, there is a 3-fold difference between the induced culture harboring pMSM67.4 compared to the induced culture harboring pMSM67. There is also a 3.7 fold increase in the amount of mCherry transcript from the induced culture of pMSM69 compared to the induced culture of pMSM67. However, there was only a 0.5 difference of fold change of the mCherry transcript between the induced cultures of pMSM69 and pMSM69.4.
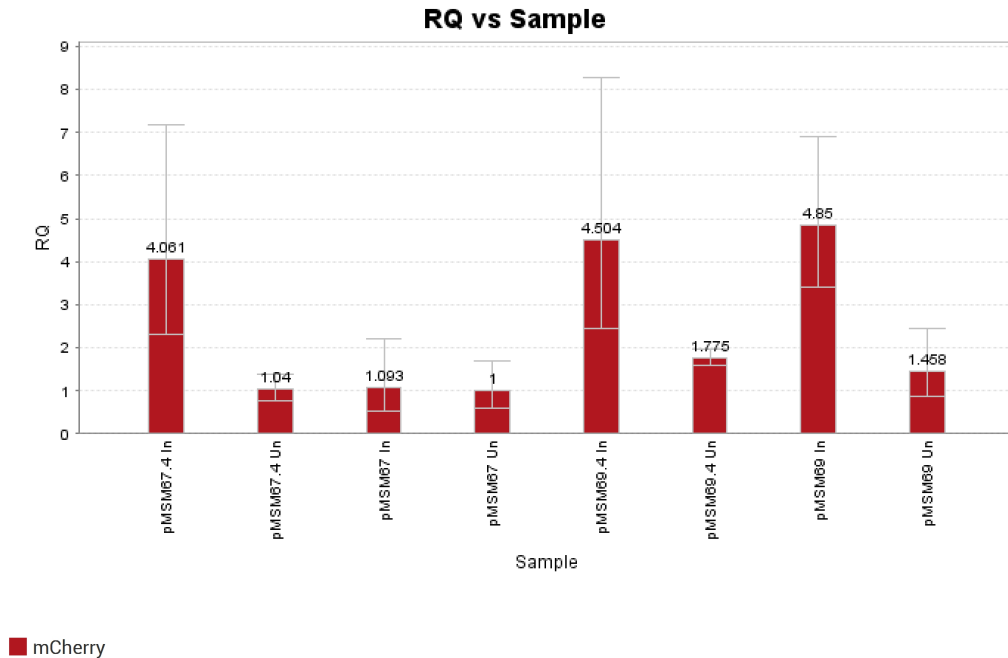
Figure 3.10: Gene Expression plot. The mCherry transcript fold change of each construct expressed in *B. subtilis*, where the uninduced pMSM67 culture with the mCherry target gene was used as the reference and the *B. subtilis* 16S rRNA gene as the endogenous control. Each sample had an Induced (In) and Uninduced (Un) culture. The error bars represent the 95% confidence interval that was defined by ΔC_t SE obtained from the output for two technical replicates. Default threshold and baseline parameters set by the QuantStudio^{TM} Design and Analysis Software v1.5.0 were used for the analysis.

3.6   Sequence Analysis of the Alpha Subunit of RNAP

Since XylS binds with the α subunit in the *E. coli* RNAP core enzyme, it was decided to compare the protein sequences of the alpha subunit between *E. coli* and *B. subtiils*. Using the UniProt database, protein sequences were obtained for the α subunit of RNAP (rpoA) for both *E. coli* K12, the strain that DH5α is derived from, and *B. subtilis* 168. The α subunit of *E. coli* is 329 residues long, and the α subunit of *B. subtilis* is 314 residues long. Using the program's default parameters, a protein sequence alignment was run with the result of only 43% identity between the two sequences. There were 143 identical positions and 107 similar positions. The sequence alignment is displayed in Figure 3.11 below.

```
P0A7Z4 RPOA_ECOLI      1  MQGSVTEFLKPRLVDIEQV-SSTHAKVTLEPLERGFGHTLGNALRRILLSSMPGCAVTEV   59
P20429 RPOA_BACSU      1  ----MIEIEKPKIETVEISDDAKFGKFVVEPLERGYGTTLGNSLRRILLSSLPGAAVTSI   56
                          : *: **:: :*   .:...*..:******:* ****:********;**.***.;

P0A7Z4 RPOA_ECOLI     60  EIDGVLHEYSTKEGVQEDILEILLNLKGLAVRVQGKDEVILTLNKSGIGPVTAADITHDG  119
P20429 RPOA_BACSU     57  QIDGVLHEFSTIEGVVEDVTTIILHIKKLALKIYSDEEKTLEIDVQGEGTVTAADITHDS  116
                          :*******:** *** **:  *;*:;:* **::: ..:*  * :: .* * *********.

P0A7Z4 RPOA_ECOLI    120  DVEIVKPQHVICHLTDENASISMRIKVQRGRGYVPASTRIHSEEDERPIGRLLVDACYSP  179
P20429 RPOA_BACSU    117  DVEILNPDLHIATL-GENASFRVRLTAQRGRGYTPADAN---KRDDQPIGVIPIDSIYTP  172
                          ****::*:  *. * .****; :*:,.******.**.:.  :.*::*** : :*: *:*

P0A7Z4 RPOA_ECOLI    180  VERIAYNVEAARVEQRTDLDKLVIEMETNGTIDPEEAIRRAATILAEQLEAFVDLRDVRQ  239
P20429 RPOA_BACSU    173  VSRVSYQVENTRVGQVANYDKLTLDVWTDGSTGPKEAIALGSKILTEHLNIFVGLTDEAQ  232
                          *.*:;*:** :** * :: ***.::: *;*: .*:*** .:.**;*:*: **.* * *

P0A7Z4 RPOA_ECOLI    240  PE---VKEEKPEFDPILLRPVDDLELTVRSANCLKAEAIHYIGDLVQRTEVELLKTPNLG  296
P20429 RPOA_BACSU    233  HAEIMVEKEEDQKEKVLEMTIEELDLSVRSYNCLKRAGINTVQELANKTEEDMMKVRNLG  292
                          *::*: : : :*   :::*:*;*** **** .*: : :*.::** :::*. ***

P0A7Z4 RPOA_ECOLI    297  KKSLTEIKDVLASRGLSLGMRLENWPPASIADE  329
P20429 RPOA_BACSU    293  RKSLEEVKAKLEELGLGLR--KDD---------  314
                          :*** *:*  *  . **.*   ::
```

Figure 3.11: Protein sequence alignment comparing the alpha subunit of RNAP between *E. coli* K12 (Accession#: P0A7Z4) and *B. subtilis* 168 (Accession#: P20429). Protein sequence alignment provided by UniProt with the Clustal Omega program. Asterisk represents identical residues, single dot represents very similar residue, two dots represent a somewhat similar residue.

A BLAST protein alignment of the CTD of the α subunit of RNAP between *E. coli* and *B. subtilis* was also performed. The α-CTD of *E. coli* is 81 residues long from 249-329, where as *B. subtilis'* is 70 residues long from 245-314. According to the output, it was established that there was only 42% identity within this domain between the two species. The sequence alignment is displayed below in Figure 3.12.

Score	Expect	Method	Identities	Positives	Gaps
55.8 bits(133)	2e-17	Compositional matrix adjust.	27/65(42%)	43/65(66%)	0/65(0%)

```
Query  2  DPILLRPVDDLELTVRSANCLKAEAIHYIGDLVQRTEVELLKTPNLGKKSLTEIKDVLAS  61
          + +L    +++L+L+VRS NCLK   I+ + +L  +TE +++K  NLG+KSL E+K   L
Sbjct  2  EKVLEMTIEELDLSVRSYNCLKRAGINTVQELANKTEEDMMKVRNLGRKSLEEVKAKLEE  61

Query 62  RGLSL  66
            GL L
Sbjct 62  LGLGL  66
```

Figure 3.12: Protein sequence alignment done with BLAST between the α-CTD of *E. coli* and *B. subtilis*. In this alignment, the query is the *E. coli* sequence and the subject is the *B. subtilis* sequence. In between the two sequences, the plus symbol represents residue similarity, and the amino acid letters represent identical residues that are a match.

4   Discussion

4.1   Sigma Factor Protein Sequence Analysis

After running protein sequence alignments between the sigma factors σ^{32} and σ^{38} from *E. coli* with the sigma factors σ^{H} and σ^{B} from *B. subtilis*, it was concluded that their protein sequences were not similar, with only 22.03% identity being the highest of the three alignments. This suggests that the P_m promoter, which is dependent on the sigma factors σ^{32} and σ^{38} from *E. coli*, would probably not be able to use native *B. subtilis* sigma factors. Therefore, it was decided to coexpress these two *E. coli* sigma factors with the XylS/Pm expression cassette. However, considering the protein sequences of the sigma factors between the two bacterial species were so dissimilar,there is the possibility that the *E. coli* sigma factors are not able to interact with the native *B. subtilis* RNAP core enzyme.

4.2   Functionality of P*liaI* Promoter in *B. subtilis*

The main aim of this study was to address the possible bottleneck of mCherry production from the P_m promoter. One of the solutions was to have sigma factors coexpressed along with the XylS/Pm expression cassette. As previously mentioned, P_m relies on sigma factors σ^{32} and σ^{38} during different phases of growth. It was hypothesized that coexpression of these sigma factors may facilitate the expression of the target protein, mCherry, from the XylS/Pm expression system. These sigma factors were expressed under the control of the bacitracin inducible promoter, P_{liaI}. P_{liaI} is native to *B. subtilis* and controls the *liaIH* operon, which is the main target of the envelope stress-inducible two-component system LiaRS [55]. Bacitracin is an antibiotic that inhibits the biosynthesis of the bacterial peptidoglycan wall, and with a high enough concentration will lead to cell lysis [56]. Up to certain concentrations, *B. subtilis* is naturally resistant to bacitracin, however when inducing a culture, concentrations exceeding 30 μg/mL can lead to minor growth defects and even higher concentrations will result in lysis of the culture [55]. The first expression performed in this study was testing the functionality of the P_{liaI} promoter in *B. subtilis* from the pMSM67.1 construct. At this point in time, we were unsure about the correct concentration of bacitracin to induce the culture with, so it was decided to try 50 μg/mL. Referring back to Figure 3.2, there was very little RFU/OD$_{600}$, in hindsight, perhaps this low signal was the result of inducing with too high of a bacitracin concentration which resulted in some cell lysis therefore lowering the OD$_{600}$ measurements.

There was an attempt to construct plasmids with a proven high copy number of 30-50 copies expressing both the XylS/Pm cassette and sigma factors σ^{32} and σ^{38} with the goal of obtaining increased expression when compared to the low copy number plasmids that were currently being used. Unfortunately, the cloning of these constructs was unsuccessful and the low copy number plasmids (pMSM67.1 - pMSM67.4) were

continued with. The copy number of a plasmid refers to the number of plasmids present in a cell. The advantage of a high-copy number plasmid for protein expression is that if there were low yields of protein expression with the low copy number plasmid, then switching to a high copy number plasmid may improve protein yields. However this is not always the case depending on the protein of interest and it may also result in protein aggregation. An expression experiment was done with one of the attempted high copy number plasmids in *E. coli*, to test whether the P_{liaI} promoter was functional in *E. coli*, however, it seemed that *E. coli* had quite a high resistance to bacitracin, which was supported in literature and it was therefore suggested that this promoter was not functional in *E. coli* (data not shown) [57].

4.3 Effect of Sigma Factor Expression on mCherry Fluorescence from P*m*

4.3.1 Expression from P*m* with Coexpression of Sigma Factors

mCherry fluorescence from P_m was measured in each of the pMSM67 vectors; pMSM67.2, pMSM67.3 and pMSM67.4 while sigma factors were being simultaneously expressed to test if their presence made an impact on the expression of mCherry from P_m. As seen in Figure 3.3, there was only mCherry fluorescence detected from the induced DH5α culture of the positive control, pVB-1A0B1-mCherry. There was a second expression also performed to include the vector pMSM69.4, along with pMSM67.4, pMSM67, pMSM69 and the controls; wild type *B. subtilis*, pMSM67.1 and finally pMSM67 (DH5α). The results of this expression, displayed in Figure 3.4, were very similar to the first, in that there was no mCherry fluorescence detected in the induced cultures from P_m in the two vectors coexpressing sigma factors, pMSM67.4 and pMSM69.4. In the preceding study it was found that expression of XylS from the uninduced medium promoter, P_{met} was greater than P_{liaG} and since it was unclear as to how much of the XylS protein was required by *B. subtilis* to possibly initiate transcription from P_m, expression of the pMSM69.4 construct was included in case the XylS level provided by pMSM67/pMSM67.4 was too low.

Unfortunately, the results from these two expression experiments, with constructs expressing low and high levels of XylS, indicates that the hypothesis that the coexpression of sigma factors to overcome the potential bottleneck of mCherry production from P_m may not be correct. To verify sigma factor expression from P_{liaI} in pMSM67.2, pMSM67.3, pMSM67.4 and pMSM69.4, antibodies were used for western blotting to detect their presence, as discussed in the section below.

4.3.2 Verifying Sigma Factor Expression from P*liaI*

As mentioned earlier, PliaI controls the expression of the *liaIH* operon in *B. subtilis*. During normal exponential growth, this promoter is switched off, and is only switched

on in the presence of inducer, in this case, bacitracin [58]. After the promoter is induced, it can result in a more than 100-fold increased activity as soon as 5-10 minutes post induction [58]. However, Radeck *et al.*, performed a quantitative analysis with PliaI, where they cloned it into the luciferase reporter vector pBS3C*lux*, grew the culture in LB to exponential phase and then induced with varying concentrations of bacitracin. What they discovered was that there was a concentration-dependent increase of luciferase activity between 2 to 4 hours post induction, following this window, there was a 100-fold decrease in signal towards stationary phase [55]. In this study, this phenomenon was also observed with the expression of sigma factors that were expressed from PliaI in both pMSM67.4 and pMSM69.4 in *B. subtilis*. Referring to the western blots in Figures 3.5, 3.6, 3.7 and 3.8 there was a noticeable decrease in the amount of sigma factors detected in the induced cultures across the 4 time points, where there was the most detected 1.5 hours post induction, and almost none detected at 20 hours post induction. Toymentseva *et al.* found that the P*liaI* promoter had a poorly conserved SD sequence with suboptimal spacing between the SD sequence and the start codon, which could lead to a decrease in the ribosome's affinity to the mRNA, resulting in inefficient translation initiation [58]. This could be an explanation as to why there is a decrease in the amount of sigma factors being produced once P_{liaI} is induced with bacitracin.

It was also noticed from the western blots that there was σ^{32} and σ^{38} detected in both the soluble and insoluble fractions in approximate equal amounts of the induced cultures post induction. When expressing heterolgous proteins, one of the main goals is to acquire a high degree of the target protein in a soluble and functional form. Unfortunately, sometimes in recombinant systems, there can be an accumulation of the target proteins in the form of insoluble aggregates, also known as inclusion bodies [59]. These inclusion bodies are usually misfolded and non-functional [59]. Even though the expression of sigma factors was not the main target, perhaps if there were a greater percentage recovered in the soluble fraction, it would have had a positive impact in producing mCherry protein from P_m. Strategies that have been used to limit the aggregation of recombinant proteins include expressing the protein at a reduced temperature or enlisting the help of molecular chaperones [60].

4.4 Amplification Efficiency

When setting up a RT-PCR reaction, it is recommended that the normalizer target, in this case the 16 S rRNA gene has a similar amplification efficiency to the experimental target, mCherry. The amplification efficiency was calculated for each of the three primer pairs. Referring back to Figure 3.9, the primer pairs for the 16 S *B. subtilis* gene and the mCherry target gene fell within the optimal range of 90-110%. Unfortunately, the amplification efficiency for the 16 S *E. coli* gene was calculated to be 168% which is much higher than the accepted limit. Due to time constraints, it was not able to be optimized and this primer pair was only used for one of the control samples, which

was not included in the fold change calculations. When an amplification efficiency exceeds 110% it can skew the results which ends up leading to false conclusions and is usually an indication of inhibition [50]. Causes of inhibition include poor RNA quality, high template concentration or carryover from nucleic acid purification [50]. Sometimes these issues can be solved by simply removing the wells with the highest template concentrations from the standard curve and reanalyzing, if the new amplification efficiency is under 110%, then the assay is fine [50]. One can also attempt at repurifying the template [50].

4.5   Evaluation of mCherry Transcript Levels from Pm

qPCR was used to determine the fold change of mCherry transcript from P$_m$ between the induced *B. subtilis* cultures harboring pMSM67.4, pMSM69.4, and pMSM69, where the uninduced culture of pMSM67 was used as the reference. Referring back to Figure 3.10, since there was a 3.7 fold change between the induced culture of pMSM69 and pMSM67, it can be concluded that the higher levels of XylS protein being expressed from pMSM69 has a positive effect on the amount of mCherry transcript. Then comparing the induced culture of pMSM67.4 and pMSM67, there was a 3 fold change, indicating that the sigma factors being expressed also have a positive effect on the amount of mCherry transcript being produced from P$_m$. It was therefore expected that when pMSM69.4 was expressed in *B. subtilis*, that the added effect with sigma factors being expressed together with a higher level of XylS, it would result in a higher level of mCherry transcript. However, this was not the case. There was only a 0.5 fold difference of mCherry transcript between pMSM67.4 and pMSM69.4. These fold changes also do not correlate to the mCherry fluorescence data obtained for these samples. These fold-changes seem large, however it was relative quantification that was performed and not absolute quantification. Hypothetically there could only have been one copy of mCherry transcript in an uninduced culture and two copies of transcript in the induced culture, this still gives a 2-fold increase, but too low of a transcript level to give a mCherry fluorescence signal. This could be the case for these samples.

After researching differences between the RNAP of *E. coli* and *B. subtilis*, it was found that the two enzymes only shared 50% identity and that gram-positive bacteria like *B. subtilis* have two extra subunits, δ and ϵ [17][19].The δ subunit binds to DNA at A-rich sequences that are immediately upstream of the -35 element of the promoter DNA, which happens to be the same region that the XylS dimer binds to [40][17]. The XylS dimer reportedly binds to two 15-bp repeated motifs, one from -70 to -56 and the other from -49 to -35 of the P$_m$ promoter,where the first XylS monomer occupying the proximal binding site interacts with the C-terminal domain of the RNA polymerase α subunit (α-CTD) to activate transcription from P$_m$ [40]. The XylS dimer and δ subunit both act as transcription regulators, and from the information just provided, seems like they may be competing for the same binding sites upstream of the -35 element of the P$_m$ promoter. Taking into consideration that there is very little sequence similarity between

the sigma factors of *E. coli* and *B. subtilis*, as discussed above, there is the possibility that the native RNAP of *B. subtilis* is unable to facilitate the required connections with σ^{32}, σ^{38}, and also the XylS dimer which are all required for transcription from P_m. Due to the differences in the RNAP between the two species, a bottleneck during transcription initiation from P_m seems likely. The section below comparing the alpha subunit of the RNAP core enzyme of *E. coli* and *B. subtilis* further strengthens this hypothesis.

4.6   Protein Sequence Comparison of the Alpha Subunit of RNAP

XylS contains two separate functionally domains; a conserved *C*-terminal domain that is responsible for DNA binding and interactions with the RNAP, and the *N*-terminal domain that is responsible for effector recognition and protein dimerization [61]. The first XylS monomer occupies the proximal binding site and interacts with the C-terminal domain of the RNA polymerase α subunit (α-CTD) to activate transcription from P_m [40]. Therefore, a sequence analysis comparing the alpha subunit in the RNAP core enzyme between *E. coli* and *B. subtilis* and also only of the C-terminal domain was run. Both sequence analyses reported below 50% identity. Due to the fact that there is quite a substantial difference between the alpha subunits, again, it is hypothesized that the *B. subtilis* RNAP core enzyme would have difficulties with properly binding to XylS.

5 Conclusion

The objective of this study was to make advancements in the optimization of the XylS/Pm expression cassette for heterologous protein expression in the gram-positive bacteria, *B. subtilis*. The next step was to investigate the potential bottleneck of mCherry production from the P_m promoter. It was hypothesized that the coexpression of sigma factors σ^{32} and σ^{38} may facilitate the expression of mCherry from the XylS/Pm expression cassette. The sigma factors were expressed from the native *B. subtilis* promoter, P_{liaI}, which is inducible by bacitracin.

The functionality of P_{liaI} was verified via mCherry fluorescence from pMSM67.1 construct upon induction by bacitracin. Expression experiments were carried out with constructs pMSM67.2, pMSM67.3, pMSM67.4, followed by western blotting, where it was verified that sigma factors σ^{32} and σ^{38} were produced from these constructs upon induction by bacitracin. It was concluded that since both sigma factors were successfully produced from pMSM67.4, there was no need to continue with constructs pMSM67.2 and pMSM67.3 for the qPCR analysis. Therefore, the later expression experiment included only pMSM69.4, and not pMSM69.2 or pMSM69.3 variants. It was noted that there was a decrease of both sigma factors across the four time points in both the soluble and insoluble fractions of the induced cultures harboring these four constructs.

From previous work, it was established that the construct pMSM69 led to more XylS production with the P_{met} promoter, than construct pMSM67 [45]. However, the effect of different XylS levels on the mCherry transcript level from P_m was still unknown. After performing the qPCR analysis, it can be concluded that there is a positive effect of increased XylS levels on the amount of mCherry transcript from P_m based on the result of a 3.7 fold difference between the induced cultures of pMSM69 and pMSM67. There was also a positive effect on the amount of mCherry transcript after coexpressing sigma factors, since there was a 3 fold difference between pMSM67.4 and pMSM67. Surprisingly, when it comes to the induced culture with pMSM69.4, there was not an increase in mCherry transcript as expected, even though while expressing both sigma factors and increased levels of XylS. Neither of these combinations of sigma factors and increased XylS levels led to detectable mCherry production. Since it was relative quantification qPCR that was performed, it is still unclear whether there is sufficient transcript to be translated for a detectable mCherry signal.

Protein sequence alignments between the *E. coli* sigma factors *B. subtilis* sigma factors reported to have less than 25% identity between each other which indicates that *E. coli* sigma factors σ^{32} and σ^{38} may not be compatible with the native *B. subtilis* RNAP core enzyme and therefore may have problems with binding. Then,the protein sequence alignments comparing the α-subunit between *E. coli* and *B. subtilis* RNAP core enzyme, concluded that between the two species, there was less than 50% identity between the entire α-subunit, but also between their C-terminal domains indicating

that XylS may have trouble binding with the α-unit of the *B. subtilis* RNAP.

In summary, the results reported from this study have led to the conclusion that there still remains a bottleneck at the production of target protein from P_m, possibly during transcription initiation, that requires further investigation.

6 Further Work

Previously, features of the XylS/Pm system have been modified through random mutagenesis and screening approaches to improve expression. Some of these features include the P_m promoter, the DNA sequence of the 5'-untranslated mRNA region of the P_m transcript, and the *xylS* coding region [39].

Further investigations concerning the possible bottleneck during transcription from P_m are required. From the results that were discussed, it would be beneficial to coexpress the *E. coli* RNAP core enzyme, in addition to the sigma factors as it seems quite likely that transcription from P_m using the native *B. subtilis* RNAP is not possible. Since *E. coli* RNAP is a large enzyme, it would be recommended to coexpress it in a second plasmid that has a compatible origin of replication with the current vectors, that also has a different antibiotic resistance gene. The *E. coli* RNAP can be expressed under the control of another native *B. subtilis* promoter like P_{lepA} that is a strong constitutive promoter or P_{xylA} which is inducible by xylose [62] [63]. A reporter gene can be inserted first to test the functionality of the the second plasmid as well as the promoter. If the expression of *E. coli* RNAP yields detectable mCherry production, then the performance of various target proteins can be compared with *E. coli*.

References

1. Verma, A. S., Agrahari, S., Rastogi, S. & Singh, A. Biotechnology in the realm of history. *Journal of Pharmacy and Bioallied Sciences* **3,** 321 (2011).

2. Kafarski, P. Rainbow code of biotechnology. *Chemik* **66,** 811–6 (2012).

3. Terpe, K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology* **72,** 211 (2006).

4. Blount, Z. D. The natural history of model organisms: The unexhausted potential of E. coli. *Elife* **4,** e05826 (2015).

5. Baneyx, F. Recombinant protein expression in Escherichia coli. *Current opinion in biotechnology* **10,** 411–421 (1999).

6. Doran, P. *Bioprocess Engineering Principles. Second Edition* (Academic Press, 2013).

7. Miller, S. I. & Salama, N. R. The gram-negative bacterial periplasm: Size matters. *PLoS biology* **16,** e2004935 (2018).

8. Baneyx, F. Recombinant protein expression in Escherichia coli. *Current opinion in biotechnology* **10,** 411–421 (1999).

9. Menzella, H. G. Comparison of two codon optimization strategies to enhance recombinant protein production in Escherichia coli. *Microbial cell factories* **10,** 15 (2011).

10. Petsch, D. & Anspach, F. B. Endotoxin removal from protein solutions. *Journal of biotechnology* **76,** 97–119 (2000).

11. Van Mellaert, L. & Anné, J. in *Novel frontiers in the production of compounds for biomedical use* 277–300 (Springer, 2001).

12. Westers, L., Westers, H. & Quax, W. J. Bacillus subtilis as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1694,** 299–310 (2004).

13. Wu, X.-C., Lee, W., Tran, L. & Wong, S. Engineering a Bacillus subtilis expression-secretion system with a strain deficient in six extracellular proteases. *Journal of Bacteriology* **173,** 4952–4958 (1991).

14. Kawamura, F. & Doi, R. H. Construction of a Bacillus subtilis double mutant deficient in extracellular alkaline and neutral proteases. *Journal of Bacteriology* **160,** 442–444 (1984).

15. Stahl, M. L. & Ferrari, E. Replacement of the Bacillus subtilis subtilisin structural gene with an In vitro-derived deletion mutation. *Journal of Bacteriology* **158,** 411–418 (1984).

16. Snustad, D. & Simmons, M. *Genetics. Sixth Edition* (John Wiley and Sons, 2012).

17. Prajapati, R. K., Sengupta, S., Rudra, P. & Mukhopadhyay, J. Bacillus subtilis δ factor functions as a transcriptional regulator by facilitating the open complex formation. *Journal of Biological Chemistry* **291,** 1064–1075 (2016).

18. Keller, A. N. *et al.* ε, a new subunit of RNA polymerase found in Gram-positive bacteria. *Journal of bacteriology* **196,** 3622–3632 (2014).

19. Artsimovitch, I., Svetlov, V., Anthony, L., Burgess, R. R. & Landick, R. RNA Polymerases from Bacillus subtilisand Escherichia coli Differ in Recognition of Regulatory Signals In Vitro. *Journal of Bacteriology* **182,** 6027–6035 (2000).

20. Alberts, B. bibinitperiod. *Molecular Biology of the Cell. Sixth Edition* (Garland Science, 2015).

21. Dong, H., Nilsson, L. & Kurland, C. G. Co-variation of trna abundance and codon usage inescherichia coliat different growth rates. *Journal of molecular biology* **260,** 649–663 (1996).

22. Kane, J. F. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Current opinion in biotechnology* **6,** 494–500 (1995).

23. Kurland, C. & Gallant, J. Errors of heterologous protein expression. *Current opinion in biotechnology* **7,** 489–493 (1996).

24. Studholme, D. J. & Buck, M. The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS microbiology letters* **186,** 1–9 (2000).

25. Paget, M. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* **5,** 1245–1265 (2015).

26. Tripathi, L., Zhang, Y. & Lin, Z. Bacterial Sigma Factors as Targets for Engineered or Synthetic Transcriptional Control. *Frontiers in Bioengineering and Biotechnology* **2,** 33. ISSN: 2296-4185 (2014).

27. Gruber, T. M. & Gross, C. A. Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. *Annual Review of Microbiology* **57.** PMID: 14527287, 441–466 (2003).

28. Clancy, S. RNA transcription by RNA polymerase: prokaryotes vs eukaryotes. *Nature education* **1,** 125 (2008).

29. Helmann, J. D. & Chamberlin, M. J. Structure and function of bacterial sigma factors. *Annual review of biochemistry* **57,** 839–872 (1988).

30. Haldenwang, W. G. The sigma factors of Bacillus subtilis. *Microbiol. Mol. Biol. Rev.* **59,** 1–30 (1995).

31. Paget, M. S. & Helmann, J. D. The σ 70 family of sigma factors. *Genome biology* **4,** 203 (2003).

32. Li, J. & Zhang, Y. Relationship between promoter sequence and its strength in gene expression. *The European physical journal E* **37,** 86 (2014).

33. Voskuil, M. I. & Chambliss, G. H. The- 16 region of Bacillus subtilis and other gram-positive bacterial promoters. *Nucleic acids research* **26,** 3584–3590 (1998).

34. Jordan, S., Junker, A., Helmann, J. D. & Mascher, T. Regulation of LiaRS-dependent gene expression in Bacillus subtilis: identification of inhibitor proteins, regulator binding sites, and target genes of a conserved cell envelope stress-sensing two-component system. *Journal of bacteriology* **188,** 5153–5166 (2006).

35. Mascher, T., Zimmer, S. L., Smith, T.-A. & Helmann, J. D. Antibiotic-inducible promoter regulated by the cell envelope stress-sensing two-component system LiaRS of Bacillus subtilis. *Antimicrobial agents and chemotherapy* **48,** 2888–2896 (2004).

36. De Vries, G., Arfman, N., Terpstra, P. & Dijkhuizen, L. Cloning, expression, and sequence analysis of the Bacillus methanolicus C1 methanol dehydrogenase gene. *Journal of bacteriology* **174,** 5346–5353 (1992).

37. Brautaset, T., Lale, R. & Valla, S. Positively regulated bacterial expression systems. *Microbial biotechnology* **2,** 15–30 (2009).

38. Gallegos, M.-T., Schleif, R., Bairoch, A., Hofmann, K. & Ramos, J. L. Arac/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* **61,** 393–410 (1997).

39. Gawin, A., Valla, S. & Brautaset, T. The XylS/Pm regulator/promoter system and its use in fundamental studies of bacterial gene expression, recombinant protein production and metabolic engineering. *Microbial biotechnology* **10,** 702–718 (2017).

40. Ruiz, R. & Ramos, J. L. Residues 137 and 153 of XylS influence contacts with the C-terminal domain of the RNA polymerase α subunit. *Biochemical and biophysical research communications* **287,** 519–521 (2001).

41. Jeon, Y. H. *et al.* Solution structure of the activator contact domain of the RNA polymerase α subunit. *Science* **270,** 1495–1497 (1995).

42. Gaal, T. *et al.* DNA-binding determinants of the alpha subunit of RNA polymerase: novel DNA-binding domain architecture. *Genes & development* **10,** 16–26 (1996).

43. Igarashi, K. *et al.* Functional map of the alpha subunit of Escherichia coli RNA polymerase: two modes of transcription activation by positive factors. *Proceedings of the National Academy of Sciences* **88,** 8958–8962 (1991).

44. Marqués, S., Manzanera, M., González-Pérez, M.-M., Gallegos, M.-T. & Ramos, J. L. The XylS-dependent Pm promoter is transcribed in vivo by RNA polymerase with $\sigma32$ or $\sigma38$ depending on the growth phase. *Molecular microbiology* **31,** 1105–1113 (1999).

REFERENCES

45. Strausland Istre, M. *Advancements in Construction of XylS/Pm Expression Vectors in Gram-positive Bacteria. Evaluated in Bacillus subtilis and Lactobacillus plantarum* (Norwegian University of Science and Technology, 2018).

46. Lee, P. Y., Costumbrado, J., Hsu, C.-Y. & Kim, Y. H. Agarose Gel Electrophoresis for the Separation of DNA Fragments (2012).

47. Lehnman, I. R. DNA Ligase: Structure, Mechanism, and Function. *Science* **186,** 790–797. ISSN: 0036-8075 (1974).

48. Froger, A. & Hall, J. E. Transformation of plasmid DNA into E. coli using the heat shock method. *Journal of visualized experiments : JoVE* **6,** 253 (2007).

49. Potter, H. & Heller, R. Transfection by Electroporation. *Current Protocols in Immunology* **117,** 10.15.1–10.15.9 (2017).

50. appliedbiosystems. *Real-time PCR handbook* (ThermoFisher Scientific, 2016).

51. Dagert, M. & Ehrlich, S. Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene* **6,** 23–28 (1979).

52. Biolabs, N. E. *Making your own electrocompetent cells* https://www.neb.com/protocols/2012/06/21/making-your-own-electrocompetent-cells (2019).

53. Xue, G.-P., S Johnson, J. & Dalrymple, B. High osmolarity improves the electro-transformation efficiency of the Grampositive bacteria Bacillus subtilis and Bacillus licheniformis. *Journal of Microbiological Methods* **34,** 183–191 (Jan. 1999).

54. Biolabs, N. E. *NEBioCalculator^{TM} Ligation Calculator* http://nebiocalculator.neb.com/#!/ligation (2019).

55. Radeck, J. *et al.* The Bacillus BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with Bacillus subtilis. *Journal of biological engineering* **7,** 29 (2013).

56. Bernard, R., Guiseppi, A., Chippaux, M., Foglino, M. & Denizot, F. Resistance to bacitracin in Bacillus subtilis: unexpected requirement of the BceAB ABC transporter in the control of expression of its own structural genes. *Journal of bacteriology* **189,** 8636–8642 (2007).

57. El Ghachi, M., Derbise, A., Bouhss, A. & Mengin-Lecreulx, D. Identification of multiple genes encoding membrane proteins with undecaprenyl pyrophosphate phosphatase (UppP) activity in Escherichia coli. *Journal of Biological Chemistry* **280,** 18689–18695 (2005).

58. Toymentseva, A. A., Schrecke, K., Sharipova, M. R. & Mascher, T. The LIKE system, a novel protein expression toolbox for Bacillus subtilis based on the liaI promoter. *Microbial cell factories* **11,** 143 (2012).

59. Villaverde, A. & Carrió, M. M. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnology letters* **25,** 1385–1395 (2003).

60. Sørensen, H. P. & Mortensen, K. K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial cell factories* **4,** 1 (2005).

61. Gallegos, M.-T., Schleif, R., Bairoch, A., Hofmann, K. & Ramos, J. L. Arac/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* **61,** 393–410 (1997).

62. Qin, Y. *et al.* The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. *Cell* **127,** 721–733 (2006).

63. Kreuzer, P., Gärtner, D., Allmansberger, R. & Hillen, W. Identification and sequence analysis of the Bacillus subtilis W23 xylR gene and xyl operator. *Journal of bacteriology* **171,** 3840–3845 (1989).

Appendices

A   Growth Media and Solutions

Lysogeny Broth (LB)

5 g Tryptone (OXOID)

2.5 g NaCl (VWR)

2.5 g Yeast extract (OXOID)

Up to 500 mL dH$_2$O and autoclaved

For LB Agar (LA), add 7.5 g Agar (OXOID) and autoclave

Glycerol Solution

50.15 g 100% Glycerol (VWR)

Up to 100 mL tap water and autoclaved.

PSI medium

10 g Tryptone (OXOID)

2.5 g Yeast extract (OXOID)

5.12 g MgSO$_4$·7H$_2$O (VWR)

Up to 500 mL dH$_2$O

pH adjusted to 7.6 using KOH and then autoclaved

TFB1

0.588 g KAc (Merck)

2.42 g RbCl (Acros Organics)

0.389 g CaCl$_2$·2H$_2$O (Merck)

3.146 g MnCl$_2$·4H$_2$O (J.T. Baker)

30 mL 99.5% Glycerol (VWR)

Up to 200 mL dH$_2$O

pH adjusted to 5.8 with acetic acid and then sterile filtered.

TFB2

0.21 g MOPS (Fisher Scientific)

1.1 g CaCl$_2$·2H$_2$O (Merck)

0.121 g RbCl (Acros Organics)

15 mL 99.5% Glcyerol (VWR)

Up to 100 mL dH$_2$O

pH adjusted to 6.5 with NaOH and then sterile filtered.

Super Optimal Broth (SOB) medium

2 g Tryptone (OXOID)

0.5 g Yeast extract (OXOID)

0.058 g NaCl (VWR)

0.019 g KCl (Merck)

0.5 g MgSO$_4$·7H$_2$O (VWR)

Up to 100 mL dH$_2$O and autoclaved

Super Optimal Catabolite-repression (SOC) medium

100 mL SOB medium

40 uL 1 M glucose solution (VWR)

Autoclaved

I

Growth Media

18.2 g Sorbitol (Sigma)

2 g Tryptone (OXOID)

1 g Yeast extract (OXOID)

1 g NaCl (VWR)

Up to 200 mL dH$_2$O and autoclaved

Washing Solution

18.2 g Sorbitol (Sigma)

18.2 g Mannitol (Sigma)

20 g 99.5% Glycerol (VWR)

Up to 200 mL dH$_2$O and autoclaved

Outgrowth Medium

9.1 g Sorbital (Sigma)

6.9 g Mannitol (Sigma)

1 g Tryptone (Sigma)

0.5 g Yeast extract (Sigma)

0.5 g NaCl (VWR)

Up to 100 mL dH$_2$O and autoclaved

B   Inducers and Antibiotics

0.05 M m-toluate (inducer)

3.40 g m-toluate (Aldrich)

50 mL 95% Ethanol (VWR)

50 mg/mL Bacitracin (inducer)

0.5 g Bacitractin

10 mL dH$_2$O and sterile filtered

100 mg/mL Ampicillin Stock

1 g Ampicillin (PanReac AppliChem)

10 mL dH$_2$O and sterile filtered.

50 mg/mL Kanamycin Stock

0.5 g Kanamycin (PanReac AppliChem)

10 mL dH$_2$O and sterile filtered

25 mg/mL Chloramphenicol stock

0.25 g Chloramphenicol (SIGMA)

10 mL 99.9% Ethanol (VWR) and sterile filtered

C   DNA Ladders and Protein Standards
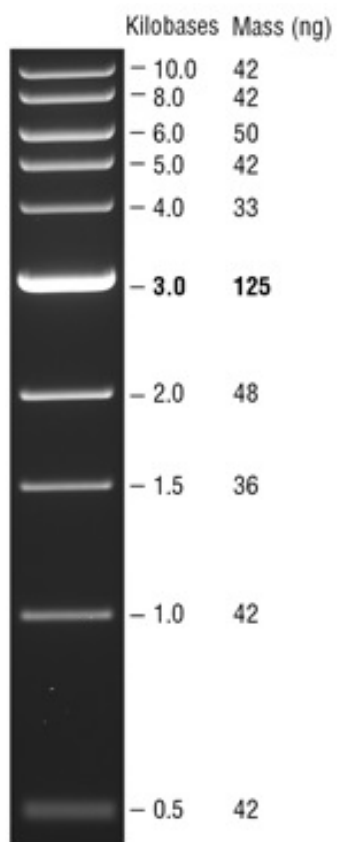
Figure C.1: NEB 1kb ladder used for gel electrophoresis. 3-5 μL of the ladder was loaded into the agarose gel.
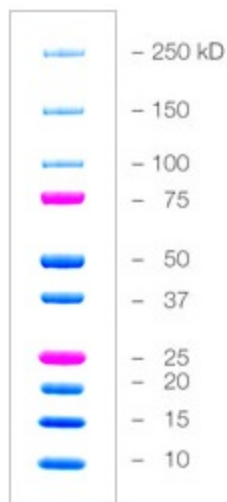
Figure C.2: Bio-Rad Precision Plus Protein Dual Color Standard used for western blotting. 5 μL of the ladder was loaded into the SDS-PAGE gel.
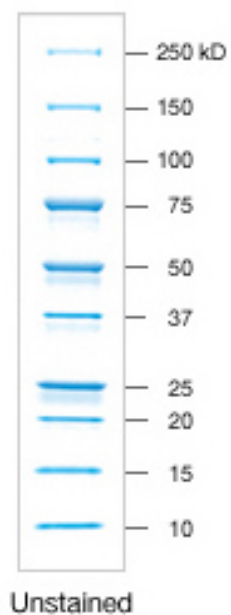
Figure C.3: Bio-Rad Precision Plus Protein Unstained Standard used for SDS-PAGE. 5 μL of the ladder was loaded into the SDS-PAGE gel.

D TECAN Settings

Table D.1: Settings used for the TECAN Infinite 200 Pro Multi-functional plate reader to visualize mCherry fluorescence.

Parameter	Value	Unit
Excitation Wavelength	580	nm
Emission Wavelength	615	nm
Excitation Bandwidth	9	nm
Emission Bandwidth	20	nm
Gain	70	Manual
Number of Flashes	25	
Integration Time	20	µs
Lag Time	0	µs
Settle Time	0	ms
Z-Position (Manual)	20000	µm

E   Primers and Probes

Table E.2: Primer and probe sequences used for qPCR, designed for use with TaqMan® Chemistry. For the probes, FAM is used as the reporter dye and TAMRA is used as the quencher dye. ROX^TM is used as the passive reference dye.

Oligo Name	Sequence (5' - 3')
mCherry Probe	[FAM]TGTCACAACTCCGCCATCTTCAAAGTTC[TAM]
16 S *B. subtilis* Probe	[FAM]TGCCGGTGACAAACCGGAGG[TAM]
16 S *E. coli* Probe	[FAM]CGCAAGCCTGATGCAGCCATGCC[TAM]
mCherry Fw	TGTACGGATCTAAAGCATACG
mCherry Rv	CTGCAGGCTTGAATCTTG
16 S *B. subtilis*Fw	TTGATCTTAGTTGCCAGCATTC
16 S *B. subtilis* Rv	CCCTTTGTTCTGTCCATTGTAG
16 S *E. coli* Fw	CCACACTGGAACTGAGACAC
16 S *E. coli* Rv	ACTTTACTCCCTTCCTCCCC

F   Plasmid Maps

Shown below are the plasmids used in this study.

Figure F.4: Plasmid maps taken from [45] and generated with Benchling.

Figure F.5: Plasmids that were outsourced to Genscript. Generated using Clone Manager software.Vectors have XylS under the control of weak promoter P_{liaG}.
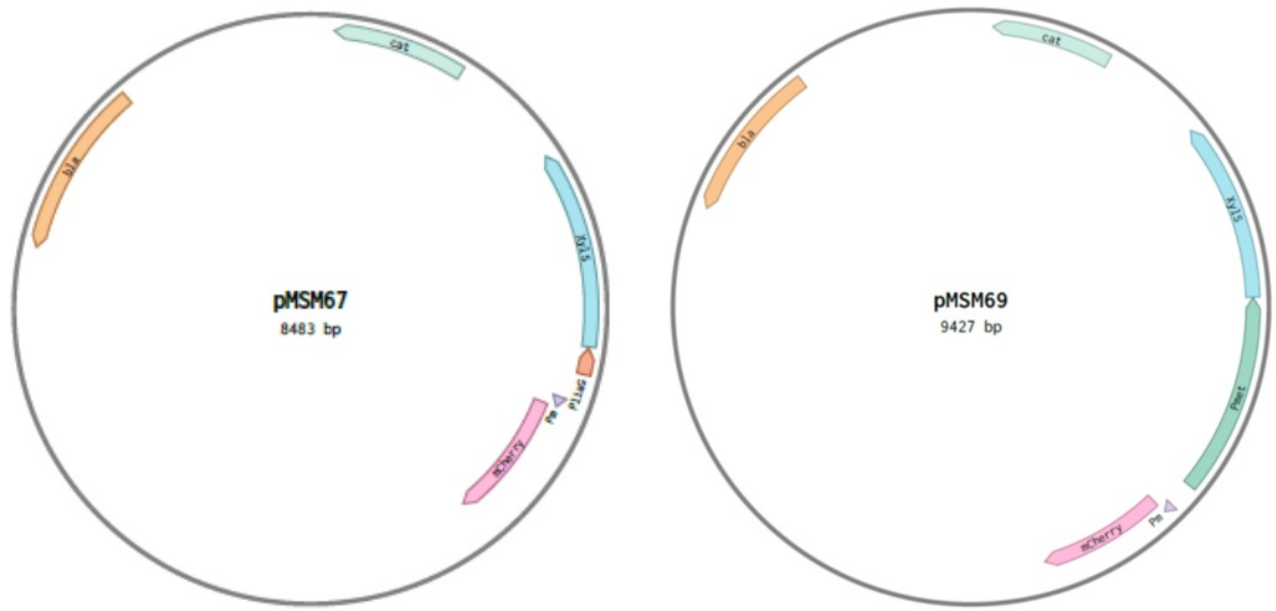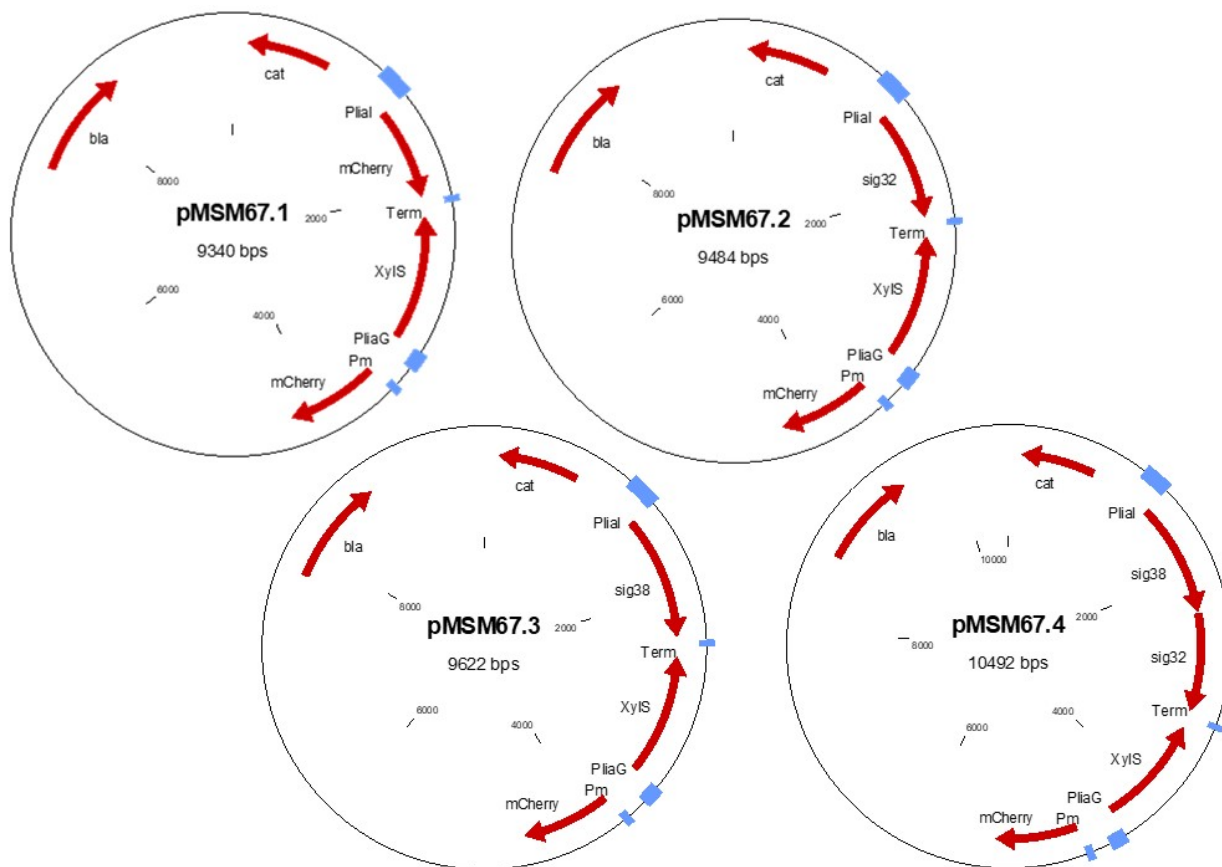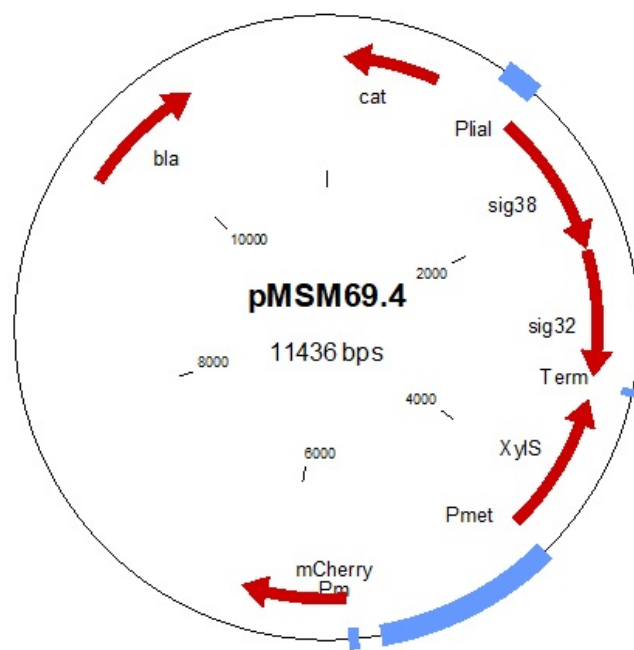
Figure F.6: Plasmid was outsourced to Genscript. Generated using Clone Manager software. Vector pMSM69.4 has XylS gene under control of medium strength promoter P_{met}.

G Codon Optimized Genes

```
GCTAGCGGATCCATTGGCCAAAGCAGAAAGGTCCGACCTAATTAAAGAAAGGGAAGCAAGTGTTCATCTGTAAAGGGTTTTAAAACGCCATGCCTCGTGCATGGCGT
CGATCGCCTAGGTAACCGGTTTCGTCTTTCCAGGCTGGATTAATTTCTTTCCCTTCGTTCACAAGTAGACATTTCCCAAAATTTTGCGGTACGGAGCACGTACCGCA
                20              40              60              80              100

TTTTTTGTGCCAATGGGTCCGGTGCGAGATACGACTCCGGTCTTATATAAAAATCAATCTCTGATTCGTTTTGCATATCTTCCAACTTGTATAAGATGAAGACAAGG
AAAAAACACGGTTACCCAGGCCACGCTCTATGCTGAGGCCAGAATATATTTTTAGTTAGAGACTAAGCAAAACGTATAGAAGGTTGAACATATTCTACTTCTGTTCC
                120             140             160             180             200

AAAACGATAAGGAGGAACTAGTATGGTTTCTAAAGGAGAAGAAGATAACATGGCAATCATCAAAGAATTTATGAGATTTAAAGTCCATATGGAAGGATCAGTTAACG
TTTTGCTATTCCTCCTTGATCATACCAAAGATTTCCTCTTCTTCTATTGTACCGTTAGTAGTTTCTTAAATACTCTAAATTTCAGGTATACCTTCCTAGTCAATTGC
            220             240             260             280             300             320

GCCATGAATTTGAAATTGAAGGCGAAGGAGAAGGCCGCCCGTATGAAGGCACACAAACAGCTAAACTGAAAGTGACAAAAGGCGGACCGCTTCCGTTTGCCTGGGAT
CGGTACTTAAACTTTAACTTCCGCTTCCTCTTCCGGCGGGCATACTTCCGTGTGTTTGTCGATTTGACTTTCACTGTTTTCCGCCTGGCGAAGGCAAACGGACCCTA
                340             360             380             400             420

ATTTTAAGCCCGCAGTTTATGTACGGATCTAAAGCATACGTCAAACATCCGGCCGGATATCCCGGATTATCTGAAACTTAGCTTTCCGGAAGGCTTTAAATGGGAAAG
TAAAATTCGGGCGTCAAATACATGCCTAGATTTCGTATGCAGTTTGTAGGCCGCCTATAGGGCCTAATAGACTTTGAATCGAAAGGCCTTCCGAAATTTACCCTTTC
            440             460             480             500             520

AGTTATGAACTTTGAAGATGGCGGAGTTGTGACAGTGACACAAGATTCAAGCCTGCAGGATGGAGAATTTATCTACAAAGTCAAACTGCGCGGAACAAACTTTCCGT
TCAATACTTGAAACTTCTACCGCCTCAACACTGTCACTGTGTTCTAAGTTCGGACGTCCTACCTCTTAAATAGATGTTTCAGTTTGACGCGCCTTGTTTGAAAGGCA
        540             560             580             600             620             640

CTGATGGCCCGGTTATGCAAAAGAAAACAATGGGCTGGGAAGCATCTTCAGAAAGAATGTATCCGGAAGATGGAGCGCTTAAAGGCGAAATCAAACAACGCTTAAAA
GACTACCGGGCCAATACGTTTTCTTTTGTTACCCGACCCTTCGTAGAAGTCTTTCTTACATAGGCCTTCTACCTCGCGAATTTCCGCTTTAGTTTGTTGCGAATTTT
            660             680             700             720             740

CTGAAAGATGGCGGACATTACGATGCTGAAGTTAAAACAACATACAAAGCCAAAAAACCGGTGCAGCTGCCGGGAGCATACAACGTCAACATCAAACTTGATATCAC
GACTTTCTACCGCCTGTAATGCTACGACTTCAATTTTGTTGTATGTTTCGGTTTTTTGGCCACGTCGACGGCCCTCGTATGTTGCAGTTGTAGTTTGAACTATAGTG
        760             780             800             820             840

ATCACATAACGAAGATTACACAATCGTGGAACAATACGAAAGAGCGGAAGGCCGCCATAGCACAGGCGGAATGGATGAATTATATAAATAAGCGGCCGCCGCCGGCG
TAGTGTATTGCTTCTAATGTGTTAGCACCTTGTTATGCTTTCTCGCCTTCCGGCGGTATCGTGTCCGCCTTACCTACTTAATATATTTATTCGCCGGCGGCGGCCGC
    860             880             900             920             940             960

ACCTGCAGTACTGAAGACAAAAAGGCCCCAAGATGTTGGGGCCTTTTTCTTAATCGTCGGTACCGAGCTC
TGGACGTCATGACTTCTGTTTTTCCGGGGTTCTACAACCCCGGAAAAAGAATTAGCAGCCATGGCTCGAG
    970     980     990     1,000   1,010   1,020   1,030
```

Figure G.7: Codon optimized P*liaI*-mCherry gene for *B. subtilis*.

```
ACTAGTATGACTGACAAAATGCAAAGTTTAGCTTTAGCCCCAGTTGGCAACCTGGATTCCTACATCCGGGCAGCTAACGCGTGGCCGATGTTGTCGGCTGACGAGGA
TGATCATACTGACTGTTTTACGTTTCAAATCGAAATCGGGGTCAACCGTTGGACCTAAGGATGTAGGCCCGTCGATTGCGCACCGGCTACAACAGCCGACTGCTCCT
        20              40              60              80             100

GCGGGCGCTGGCTGAAAAGCTGCATTACCATGGCGATCTGGAAGCAGCTAAAACGCTGATCCTGTCTCACCTGCGGTTTGTTGTTCATATTGCTCGTAATTATGCGG
CGCCCGCGACCGACTTTTCGACGTAATGGTACCGCTAGACCTTCGTCGATTTTGCGACTAGGACAGAGTGGACGCCAAACAACAAGTATAACGAGCATTAATACGCC
        120             140             160             180             200

GCTATGGCCTGCCACAGGCGGATTTGATTCAGGAAGGTAACATCGGCCTGATGAAAGCAGTGCGCCGTTTTAACCCGGAAGTGGGTGTGCGCCTGGTCTCCTTCGCC
CGATACCGGACGGTGTCCGCCTAAACTAAGTCCTTCCATTGTAGCCGGACTACTTTCGTCACGCGGCAAAATTGGGCCTTCACCCACACGCGGACCAGAGGAAGCGG
        220             240             260             280             300             320

GTTCACTGGATCAAAGCAGAGATCCACGAATACGTCCTGCGTAACTGGCGTATCGTCAAAGTTGCGACCACCAAAGCGCAGCGCAAACTGTTCTTCAACCTGCGTAA
CAAGTGACCTAGTTTCGTCTCTAGGTGCTTATGCAGGACGCATTGACCGCATAGCAGTTTCAACGCTGGTGGTTTCGCGTCGCGTTTGACAAGAAGTTGGACGCATT
        340             360             380             400             420

AACCAAGCAGCGTCTGGGCTGGTTTAACCAGGATGAAGTCGAAATGGTGGCCCCGTGAACTGGGCGTAACCAGCAAAGACGTTCGTGAGATGGAATCACGTATGGCGG
TTGGTTCGTCGCAGACCCGACCAAATTGGTCCTACTTCAGCTTTACCACCGGGCACTTGACCCGCATTGGTCGTTTCTGCAAGCACTCTACCTTAGTGCATACCGCC
        440             460             480             500             520

CACAGGACATGACCTTTGACCTGTCTTCCGACGACGATTCCGACAGCCAGCCGATGGCACCGGTGCTCTATCTGCAGGATAAATCATCTAACTTTGCCGACGGCATC
GTGTCCTGTACTGGAAACTGGACAGAAGGCTGCTGCTAAGGCTGTCGGTCGGCTACCGTGGCCACGAGATAGACGTCCTATTTAGTAGATTGAAACGGCTGCCGTAG
        540             560             580             600             620             640

GAAGATGATAACTGGGAAGAGCAGGCGGCAAACCGTCTGACCGACGCGATGCAGGGTCTGGACGAACGCAGCCAGGACATCATCCGTGCGCGCTGGCTGGACGAAGA
CTTCTACTATTGACCCTTCTCGTCCGCCGTTTGGCAGACTGGCTGCGCTACGTCCCAGACCTGCTTGCGTCGGTCCTGTAGTAGGCACGCGCGACCGACCTGCTTCT
        660             680             700             720             740

CAACAAGTCCACGTTGCAGGAACTGGCTGACCGTTACGGTGTTTCCGCTGAACGTGTGCGCCAACTGGAAAAGAACGCGATGAAAAAATTGCGTGCTGCTATAGAAG
GTTGTTCAGGTGCAACGTCCTTGACCGACTGGCAATGCCACAAAGGCGACTTGCACACGCGGTTGACCTTTTCTTGCGCTACTTTTTTAACGCACGACGATATCTTC
        760             780             800             820             840

CGTAAGCGGCCGC
GCATTCGCCGGCG
        860
```

Figure G.8: Codon optimized Sigma 32 gene for *B. subtilis*.

```
ACTAGTATGAGTCAGAATACGCTGAAAGTTCATGATTTAAATGAAGATGCGGAATTTGATGAGAACGGAGTTGAGGTTTTTGACGAAAAGGCCTTAGTAGAAGAGGA
TGATCATACTCAGTCTTATGCGACTTTCAAGTACTAAATTTACTTCTACGCCTTAAACTACTCTTGCCTCAACTCCAAAAACTGCTTTTCCGGAATCATCTTCTCCT
          20              40              60              80              100
```

```
ACCCAGTGATAACGATTTGGCCGAAGAGGAACTGTTATCGCAGGGAGCCACACAGCGTGTGTTGGACGCGACTCAGCTTTACCTTGGTGAGATTGGTTATTCACCAC
TGGGTCACTATTGCTAAACCGGCTTCTCCTTGACAATAGCGTCCCTCGGTGTGTCGCACACAACCTGCGCTGAGTCGAAATGGAACCACTCTAACCAATAAGTGGTG
          120             140             160             180             200
```

```
TGTTAACGGCCGAAGAAGAAGTTTATTTTGCGCGTCGCGCACTGCGTGGAGATGTCGCCTCTCGCCGCCGGATGATCGAGAGTAACTTGCGTCTGGTGGTAAAAATT
ACAATTGCCGGCTTCTTCTTCAAATAAAACGCGCAGCGCGTGACGCACCTCTACAGCGGAGAGCGGCGGCCTACTAGCTCTCATTGAACGCAGACCACCATTTTTAA
          220             240             260             280             300             320
```

```
GCCCGCCGTTATGGCAATCGTGGTCTGGCGTTGCTGGACCTTATCGAAGAGGGCAACCTGGGGCTGATCCGCGCGGTAGAGAAGTTTGACCCGGAACGTGGTTTCCG
CGGGCGGCAATACCGTTAGCACCAGACCGCAACGACCTGGAATAGCTTCTCCCGTTGGACCCCGACTAGGCGCGCCATCTCTTCAAACTGGGCCTTGCACCAAAGGC
          340             360             380             400             420
```

```
CTTCTCAACATACGCAACCTGGTGGATTCGCCAGACGATTGAACGGGCGATTATGAACCAAACCCGTACTATTCGTTTGCCGATTCACATCGTAAAGGAGCTGAACG
GAAGAGTTGTATGCGTTGGACCACCTAAGCGGTCTGCTAACTTGCCCGCTAATACTTGGTTTGGGCATGATAAGCAAACGGCTAAGTGTAGCATTTCCTCGACTTGC
          440             460             480             500             520
```

```
TTTACCTGCGAACCGCACGTGAGTTGTCCCATAAGCTGGACCATGAACCAAGTGCGGAAGAGATCGCAGAGCAACTGGATAAGCCAGTTGATGACGTCAGCCGTATG
AAATGGACGCTTGGCGTGCACTCAACAGGGTATTCGACCTGGTACTTGGTTCACGCCTTCTCTAGCGTCTCGTTGACCTATTCGGTCAACTACTGCAGTCGGCATAC
          540             560             580             600             620             640
```

```
CTTCGTCTTAACGAGCGCATTACCTCGGTAGACACCCCGCTGGGTGGTGATTCCGAAAAAGCGTTGCTGGACATCCTGGCCGATGAAAAAGAGAACGGTCCGGAAGA
GAAGCAGAATTGCTCGCGTAATGGAGCCATCTGTGGGGCGACCCACCACTAAGGCTTTTTCGCAACGACCTGTAGGACCGGCTACTTTTTCTCTTGCCAGGCCTTCT
          660             680             700             720             740
```

```
TACCACGCAAGATGACGATATGAAGCAGAGCATCGTCAAATGGCTGTTCGAGCTGAACGCCAAACAGCGTGAAGTGCTGGCACGTCGATTCGGTTTGCTGGGGTACG
ATGGTGCGTTCTACTGCTATACTTCGTCTCGTAGCAGTTTACCGACAAGCTCGACTTGCGGTTTGTCGCACTTCACGACCGTGCAGCTAAGCCAAACGACCCCATGC
          760             780             800             820             840
```

```
AAGCGGCAACACTGGAAGATGTAGGTCGTGAAATTGGCCTCACCCGTGAACGTGTTCGCCAGATTCAGGTTGAAGGCCTGCGCCGTTTGCGTGAAATCCTGCAAACG
TTCGCCGTTGTGACCTTCTACATCCAGCACTTTAACCGGAGTGGGCACTTGCACAAGCGGTCTAAGTCCAACTTCCGGACGCGGCAAACGCACTTTAGGACGTTTGC
          860             880             900             920             940             960
```

```
CAGGGGCTGAATATCGAAGCGCTGTTCCGCGAGTAAGCGGCCGC
GTCCCCGACTTATAGCTTCGCGACAAGGCGCTCATTCGCCGGCG
     970        980        990       1,000
```

Figure G.9: Codon optimized Sigma 38 gene for *B. subtilis*.

```
ACTAGTATGAGTCAGAATACGCTGAAAGTTCATGATTTAAATGAAGATGCGGAATTTGATGAGAACGGAGTTGAGGTTTTTGACGAAAAGGCCTTAGTAGAAGAGGA
TGATCATACTCAGTCTTATGCGACTTTCAAGTACTAAATTTACTTCTACGCCTTAAACTACTCTTGCCTCAACTCCAAAACTGCTTTTCCGGAATCATCTTCTCCT
        20            40            60            80            100

ACCCAGTGATAACGATTTGGCCGAAGAGGAACTGTTATCGCAGGGAGCCACACAGCGTGTGTTGGACGCGACTCAGCTTTACCTTGGTGAGATTGGTTATTCACCAC
TGGGTCACTATTGCTAAACCGGCTTCTCCTTGACAATAGCGTCCCTCGGTGTGTCGCACACAACCTGCGCTGAGTCGAAATGGAACCACTCTAACCAATAAGTGGTG
        120           140           160           180           200

TGTTAACGGCCGAAGAAGAAGTTTATTTTGCGCGTCGCGCACTGCGTGGAGATGTCGCCTCTCGCCGCCGGATGATCGAGAGTAACTTGCGTCTGGTGGTAAAAATT
ACAATTGCCGGCTTCTTCTTCAAATAAAACGCGCAGCGCGTGACGCACCTCTACAGCGGAGAGCGGCGGCCTACTAGCTCTCATTGAACGCAGACCACCATTTTTAA
        220           240           260           280           300           320

GCCCGCCGTTATGGCAATCGTGGTCTGGCGTTGCTGGACCTTATCGAAGAGGGCAACCTGGGGCTGATCCGCGCGGTAGAGAAGTTTGACCCGGAACGTGGTTTCCG
CGGGCGGCAATACCGTTAGCACCAGACCGCAACGACCTGGAATAGCTTCTCCCGTTGGACCCCGACTAGGCGCGCCATCTCTTCAAACTGGGCCTTGCACCAAAGGC
        340           360           380           400           420

CTTCTCAACATACGCAACCTGGTGGATTCGCCAGACGATTGAACGGGCGATTATGAACCAAACCCGTACTATTCGTTTGCCGATTCACATCGTAAAGGAGCTGAACG
GAAGAGTTGTATGCGTTGGACCACCTAAGCGGTCTGCTAACTTGCCCGCTAATACTTGGTTTGGGCATGATAAGCAAACGGCTAAGTGTAGCATTTCCTCGACTTGC
        440           460           480           500           520

TTTACCTGCGAACCGCACGTGAGTTGTCCCATAAGCTGGACCATGAACCAAGTGCGGAAGAGATCGCAGAGCAACTGGATAAGCCAGTTGATGACGTCAGCCGTATG
AAATGGACGCTTGGCGTGCACTCAACAGGGTATTCGACCTGGTACTTGGTTCACGCCTTCTCTAGCGTCTCGTTGACCTATTCGGTCAACTACTGCAGTCGGCATAC
        540           560           580           600           620           640

CTTCGTCTTAACGAGCGCATTACCTCGGTAGACACCCCGCTGGGTGGTGATTCCGAAAAAGCGTTGCTGGACATCCTGGCCGATGAAAAAGAGAACGGTCCGGAAGA
GAAGCAGAATTGCTCGCGTAATGGAGCCATCTGTGGGGCGACCCACCCACTAAGGCTTTTTCGCAACGACCTGTAGGACCGGCTACTTTTTCTCTTGCCAGGCCTTCT
        660           680           700           720           740

TACCACGCAAGATGACGATATGAAGCAGAGCATCGTCAAATGGCTGTTCGAGCTGAACGCCAAACAGCGTGAAGTGCTGGCACGTCGATTCGGTTTGCTGGGGTACG
ATGGTGCGTTCTACTGCTATACTTCGTCTCGTAGCAGTTTACCGACAAGCTCGACTTGCCGGTTTGTCGCACTTCACGACCGTGCAGCTAAGCCAAACGACCCCATGC
        760           780           800           820           840

AAGCGGCAACACTGGAAGATGTAGGTCGTGAAATTGGCCTCACCCGTGAACGTGTTCGCCAGATTCAGGTTGAAGGCCTGCGCCGTTTGCGTGAAATCCTGCAAACG
TTCGCCGTTGTGACCTTCTACATCCAGCACTTTAACCGGAGTGGGCACTTGCACAAGCGGTCTAAGTCCAACTTCCGGACGCGGCAAACGCACTTTAGGACGTTTGC
        860           880           900           920           960

CAGGGGCTGAATATCGAAGCGCTGTTCCGCGAGTAATAAGGAGGAACTTGTATGACTGACAAAATGCAAAGTTTAGCTTTAGCCCCAGTTGGCAACCTGGATTCCTA
GTCCCCGACTTATAGCTTCGCGACAAGGCGCTCATTATTCCTCCTTGAACATACTGACTGTTTTACGTTTCAAATCGAAATCGGGGTCAACCGTTGGACCTAAGGAT
        980          1,000         1,020         1,040         1,060

CATCCGGGCAGCTAACGCGTGGCCGATGTTGTCGGCTGACGAGGAGCGGGCGCTGGCTGAAAAGCTGCATTACCATGGCGATCTGGAAGCAGCTAAAACGCTGATCC
GTAGGCCCGTCGATTGCGCACCGGCTACAACAGCCGACTGCTCCTCGCCCGCGACCGACTTTTCGACGTAATGGTACCGCTAGACCTTCGTCGATTTTGCGACTAGG
        1,080         1,100         1,120         1,140         1,160

TGTCTCACCTGCGGGTTTGTTGTTCATATTGCTCGTAATTATGCGGGCTATGGCCTGCCACAGGCGGATTTGATTCAGGAAGGTAACTCGGCCTGATGAAAGCAGTG
ACAGAGTGGACGCCAAACAACAAGTATAACGAGCATTAATACGCCCGATACCGGACGGTGTCCGCCTAAACTAAGTCCTTCCATTGTAGCCGGACTACTTTCGTCAC
   1,180         1,200         1,220         1,240         1,260         1,280

CGCCGTTTTAACCCGGAAGTGGGTGTGCGCCTGGTCTCCTTCGCCGTTCACTGGATCAAAGCAGAGATCCACGAATACGTCCTGCGTAACTGGCGTATCGTCAAAGT
GCGGCAAAATTGGGCCTTCACCCACACGCGGACCAGAGGAAGCGGCAAGTGACCTAGTTTCGTCTCTAGGTGCTTATGCAGGACGCATTGACCGCATAGCAGTTTCA
        1,300         1,320         1,340         1,360         1,380

TGCGACCACCAAAGCGCAGCGCAAACTGTTCTTCAACCTGCGTAAAACCAAGCAGCGTCTGGGCTGGTTTAACCAGGATGAAGTCGAAATGGTGGCCCGTGAACTGG
ACGCTGGTGGTTTCGCGTCGCGTTTGACAAGAAGTTGGACGCATTTTGGTTCGTCGCAGACCCGACCAAATTGGTCCTACTTCAGCTTTACCACCGGGCACTTGACC
     1,400         1,420         1,440         1,460         1,480

GCCGTAACCAGCAAAGACGTTCGTGAGATGGAATCACGTATGGCCGGCACAGGACATGACCTTTGACCTGTCTTCCGACGACGATTCCGACAGCCAGCCGATGGCACCG
CGCATTGGTCGTTTCTGCAAGCACTCTACCTTAGTGCATACCGCCGTGTCCTGTACTGGAAACTGGACAGAAGGCTGCTGCTAAGGCTGTCGGTCGGCTACCGTGGC
   1,500         1,520         1,540         1,560         1,580         1,600

GTGCTCTATCTGCAGGATAAATCATCTAACTTTGCCGACGGCATCGAAGATGATAACTGGGAAGAGCAGGCGGCCAAACCGTCTGACCGACGCGATGCAGGGTCTGGA
CACGAGATAGACGTCCTATTTAGTAGATTGAAACGGCTGCCGTAGCTTCTACTATTGACCCTTCTCGTCCGCCGTTTGGCAGACTGGCTGCGCTACGTCCCAGACCT
        1,620         1,640         1,660         1,680         1,700

CGAACGCAGCCAGGACATCATCCGTGCGCGCTGGCTGGACGAAGACAACAAGTCCACGTTGCAGGAACTGGCTGACCGTTACGGTGTTTCCGCTGAACGTGTGCGCC
GCTTGCGTCGGTCCTGTAGTAGGCACGCGCGACCGACCTGCTTCTGTTGTTCAGGTGCAACGTCCTTGACCGACTGGCAATGCCACAAAGGCGACTTGCACACGCGG
        1,720         1,740         1,760         1,780         1,800

AACTGGAAAAGAACGCGATGAAAAAATTGCGTGCTGCTATAGAAGCGTAAGCGGCCGC
TTGACCTTTTCTTGCGCTACTTTTTTAACGCACGACGATATCTTCGCATTCGCCGGCG
1,820   1,830    1,840    1,850    1,860    1,870
```

Figure G.10: Codon optimized Sigma 32-Sigma 38 gene for *B. subtilis*.
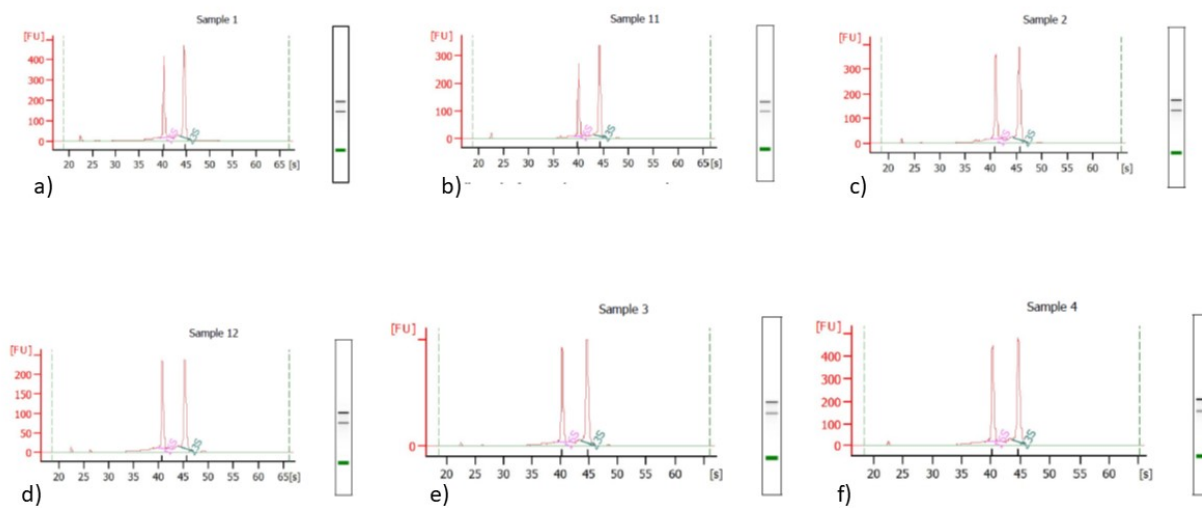
H   Electropherograms and Agarose Gel

Figure H.11: Electropherograms of the samples obtained from the 3 hour time point used in the qPCR analysis. A) pMSM67.1 (RIN 9.6); b) pMSM69.4 Induced (RIN 9.8); c) pMSM69.4 Uninduced (RIN 9.3); d) pMSM69 Induced (RIN 9.1); e) pMSM69 Uninduced (RIN 9.5); f) *B. subtilis* wild type (RIN 9.5).
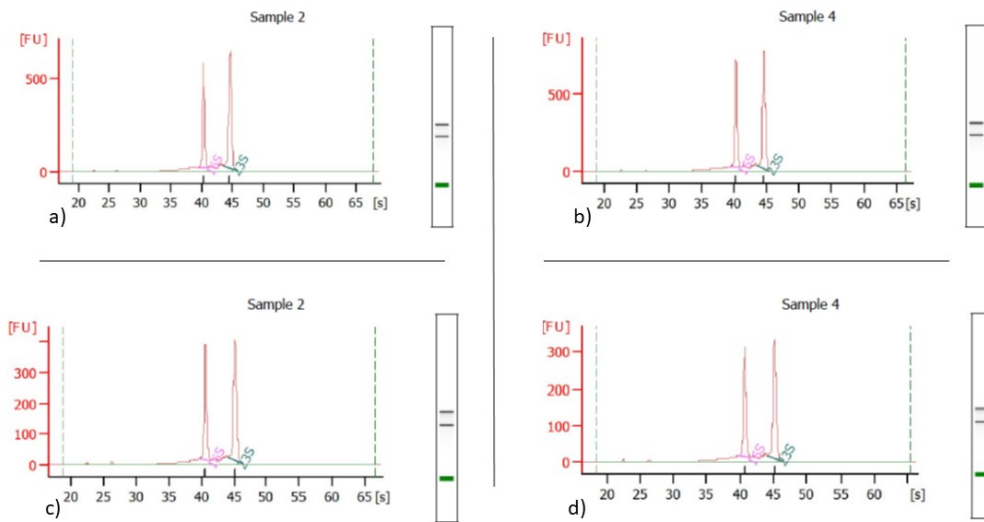
Figure H.12: Electropherograms of the samples obtained from the 5 hour time point used in the qPCR anaysis. A) pMSM67.4 Induced (RIN 9.5); b) pMSM67.4 Uninduced (RIN 9.4); c) pMSM67 Induced (RIN 9.4); d) pMSM67 Uninduced (RIN 9.3).
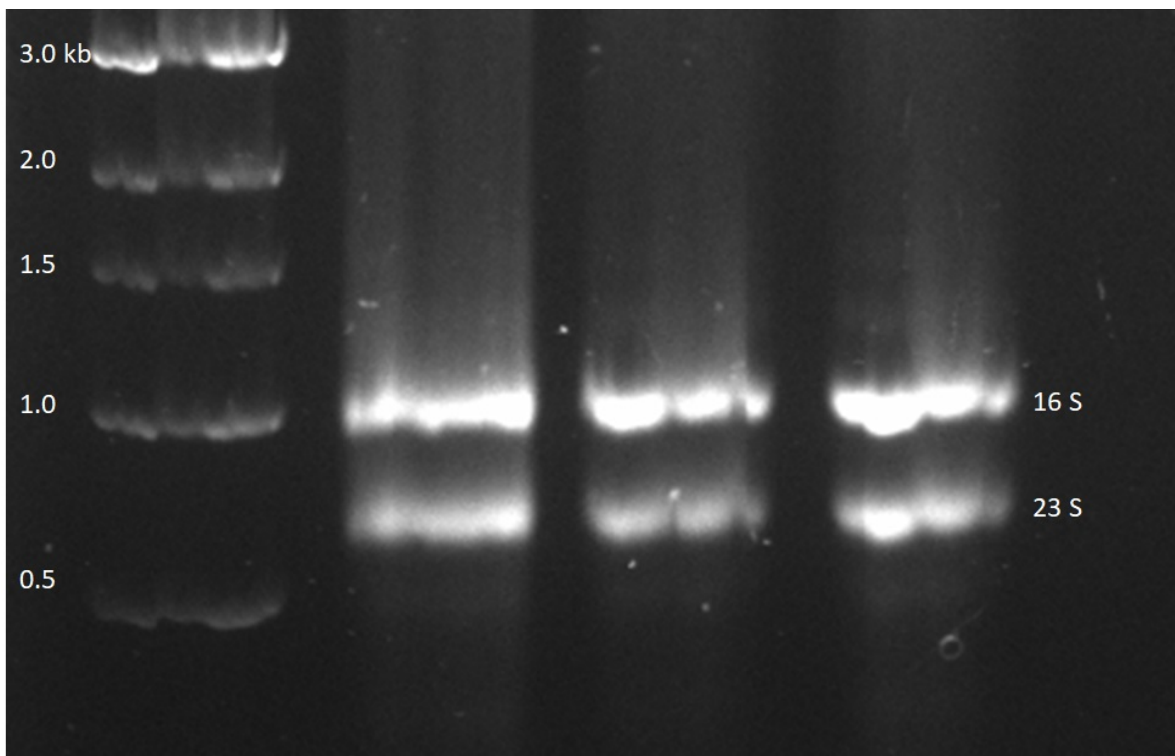

Figure H.13: Clear bands of 16 S and 23 S rRNA from RNA isolations of *E. coli* pMSM67 sample.