Investigation of *clpP* Expression During *Bacillus subtilis* Sporulation

by

Audrey Baxter Womer

A Paper Presented to the

Faculty of Mount Holyoke College in

Partial Fulfillment of the Requirements for

the Degree of Bachelors of Arts with Honor

Department of Biological Sciences

South Hadley, MA 01075

May 2026

This paper was prepared

under the direction of

Professor Amy Camp

for eight credits.

ACKNOWLEDGEMENTS

The past year has probably been one of the most insane and challenging years of my life. It pushed me to grow and be strong in ways I did not think were possible, but it also made the support I received even more meaningful.

I am proud of myself, first and foremost, for sticking through it all and keeping with my dreams and goals of becoming a scientist. This is not the path I originally saw for myself when I first arrived at Mount Holyoke, and I am endlessly grateful to the amazing science departments here and the professors who inspired me with the passion and inclusivity they brought to science. A special shoutout to Darren Hamilton, my Gen Chem and Orgo professor, who first suggested to me that I should apply to an independent research lab.

Each step in this journey has felt surreal, from my first meeting with Amy as an excited and nervous sophomore to presenting at Senior Symposium just last week. Each time it's hard for me to believe that I can do it, but I have, and here we are. I am so appreciative of my mentor and advisor Amy Camp for allowing for all of this to be possible and guiding and supporting me throughout this process, especially as a non-traditional thesis student. Being a part of the Camp Lab has truly been one of my favorite experiences at Mount Holyoke, and I am so grateful to you for creating this supportive and uplifting research environment.

To all the past senior members of the lab, Jen, Jade, Khushi and Kim for giving me people to look up to and bringing so much enthusiasm, and to my cohort BBC (Becks and Chelsea). Returning as a senior, I also want to shout out the newer members, especially EJ, who, despite us often only being able to catch up in passing, is always so funny and supportive.

I am also so grateful to other important figures, Sarah Bacon, my academic advisor and Borden Lacy who gave me endless advice, support and inspiration as I have navigated my scientific journey. Navigating a space full of PhD students and post-docs was incredibly scary, so I am so grateful to Bruno and other PhD students for making me feel like I belonged and encouraging my scientific thinking, it truly meant the world to me.

To my amazing, incredible and supportive friends, you all have truly been my rock in this tumultuous time. Rose, doing this thesis alongside you is something I don't know if I could have done this without. Aisha, Imari and Cindy, thank you for always being understanding even when you didn't know what I was doing. I love all of you guys so much.

Last, but not least, to my family. To Mama and Dad, Grandma and Grandpa and Opa for being my biggest supporters and making it possible for me to have this experience. None of this would be possible without you and the love you have for me. I love you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Sporulation in *Bacillus subtilis* is highly regulated by stage-specific sigma factors that control gene expression as the cell transitions to a metabolically dormant state. The AAA+ protease complex ClpCP, activated in the forespore by the adaptor protein MdfA, plays an important role in degrading metabolites during metabolic shutdown. Recent findings in our lab and others demonstrate that *clpC* is turned on by the forespore-specific sigma factor, sigF, as is *mdfA,* raising the possibility that *clpP* is similarly regulated. Additionally, bioinformatic analysis of the region upstream of *clpP* reveal matching to consensus sequences for both sigF and the stress-activated sigB, in line with other dual-activated promoters in *B. subtilis.* We also identified a match to a sigE consensus sequence nearby. My project aimed to test the hypothesis that *clpP* expression during sporulation is controlled by a sigF-activated promoter. In order to evaluate this, we constructed a P_{clpP}-*lacZ* fusion and transformed it into IPTG-inducible sigF, sigB or sigE strains and measured beta-galactosidase activity. We found that the P_{clpP} promoter is activated by sigF, but not by sigE. This project helped advance our understanding of sporulation in *B. subtilis* and provided insight into conserved regulatory mechanisms in pathogenic bacteria.

CHAPTER 1: INTRODUCTION

Spore formation is an important feature of reproduction and survival for many different organisms including bacteria, algae, fungi and some plants. Specifically, the existence of fungal spores often elicits a fear response, as we think of the spores that spread *Cordyceps* in The Last of Us. And it is true that spores are extraordinary, in fact, viable bacterial spores have been found preserved in amber for tens of millions of years and culturable fungal spores have been found preserved in ocean sediments and arctic ice for hundreds of thousands of years (Huang & Hull, 2017). Spores are built to be incredibly resistant to environmental and nutritional stressors, able to survive harsh conditions, including the vacuum of space. At its core, spore formation is about reproduction and colonization. In plants and fungi, spores are often an efficient means of spreading as they are small and durable and can be easily dispersed to areas where they can then germinate and grow if the environmental conditions are favorable. While there is not as much research available on eukaryotic spore formation that occurs in plants and fungi, we have much more information from bacterial spore formers. Increasing our understanding of this process in bacteria and other microbes will help us determine the mechanisms for survival and resilience underlying this process.

1-1 Sporulation in Bacteria

Bacteria form a part of our body and organ systems as humans as well as the environment around us. While many provide a harmless or even beneficial effect, some bacteria are pathogenic, or disease causing. One of the survival mechanisms that allow bacteria to evade harsh conditions and targeted treatments is the process of sporulation, which results in the creation of a dormant spore protected by a thick spore coat (Tan & Ramamurthi, 2014).

Bacterial spores, or endospores, are one of the most resilient cell types, resisting extreme environmental, chemical and UV conditions, allowing the bacterial cell to outlast suboptimal conditions.

1-1-1 *Bacillus subtilis* as a Model Spore Former

Endospore-forming bacteria species mainly include the *Bacillus* and *Clostridia* families. The *Bacillus* and *Clostridia* families are the species of bacteria that we have the most detailed molecular information about. *Bacillus subtilis* is related to the pathogenic *Bacillus anthracis* and further to pathogenic *Clostridia* species. *Bacillus subtilis* is a gram-positive bacterium that has commonly been identified as a soil-dweller, but also forms a part of the human gut microbiome (Koopman et al., 2022). It is a model organism that we can use to study this process, as extensive work and sequencing has been done on *B. subtilis,* and thus we have extensive knowledge on its regulatory pathways. Additionally, it has natural competence, or the ability to take up exogenous DNA, which makes it easy to manipulate genetically (Harwood, 2007). In studying sporulation specifically, *B. subtilis* genes involved in sporulation are usually not necessary for normal growth which makes it easier to create mutant strains and identify novel, sporulation-specific factors (Tan & Ramamurthi, 2014).

1-1-2 Sporulation in *Bacillus subtilis*

Sporulation is initiated when bacterial cells sense certain stressful environmental signals such as heat stress, UV radiation, chemical stress and starvation conditions. It is tightly regulated by the cell as it involves the transition from a normally growing vegetative condition to a metabolically dormant state. The single, rod-shaped bacterium undergoing sporulation first divides asymmetrically to form a larger “mother cell” and a smaller forespore, a process known

as polar septation. The cell then enters the engulfment phase, where the forespore is taken up by the mother cell, making it doubly membrane-bound in the mother cell cytosol. The mother cell also synthesizes a peptidoglycan cortex and a multilayered protein coat, making the spore highly resistant when it is released via mother cell lysis (Tan & Ramamurthi, 2014).

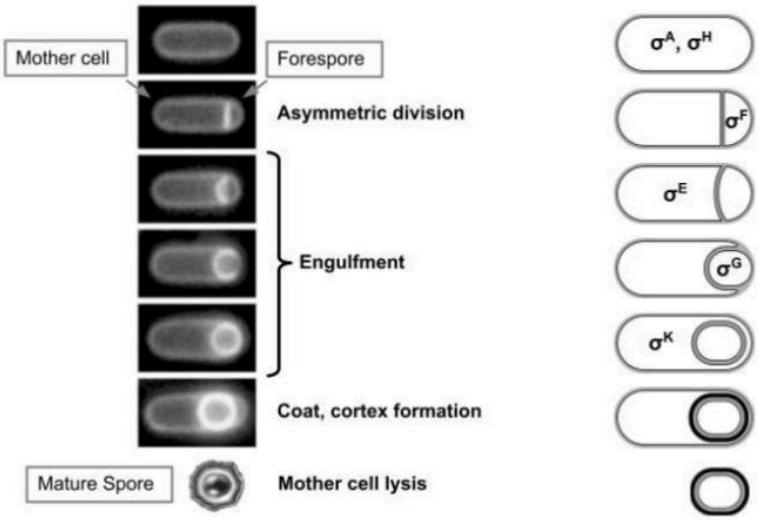

Figure 1. The stages of sporulation and the sigma factors involved. On the left we have a fluorescence microscopy visualization of this process, matched with the schematic on the right that indicates which sigma factors control each stage. Before polar septation, sigA, the housekeeping sigma factor, and sigH, control gene expression. After polar septation, sporulation-specific sigma factor control expression in the different compartments.

1-2 Sigma Factors Control Gene Expression

Sigma factors are transcription factors that regulate gene expression in bacteria. Transcription is the molecular process of making RNA copies of a DNA template, which are then translated into protein. Sequences in the DNA determine the start and end of transcription, with each gene having a promoter region, a start site upstream of the gene itself and a terminator sequence downstream. Transcription factors bind to specific promoter regions, directing RNA

polymerase to that promoter. They then form a part of the holoenzyme that makes up RNA polymerase, a complex of alpha and beta subunits that assemble at the promoter sequence of a gene. Once transcription has been initiated, the sigma factor dissociates and RNA polymerase continues, catalyzing the phosphodiester bonds of the resulting RNA strand. Specifically, at the promoter site, the sigma factor directs the RNA polymerase to the -35 and -10 regions of the promoter sequence. Each sigma prefers a specific combination of nucleotides, known as its consensus sequence. The promoter must provide a good match to this consensus sequence. This is what determines the binding specificity of the sigma factor, allowing it to regulate specific genes at specific times (Haldenwang, 1995).

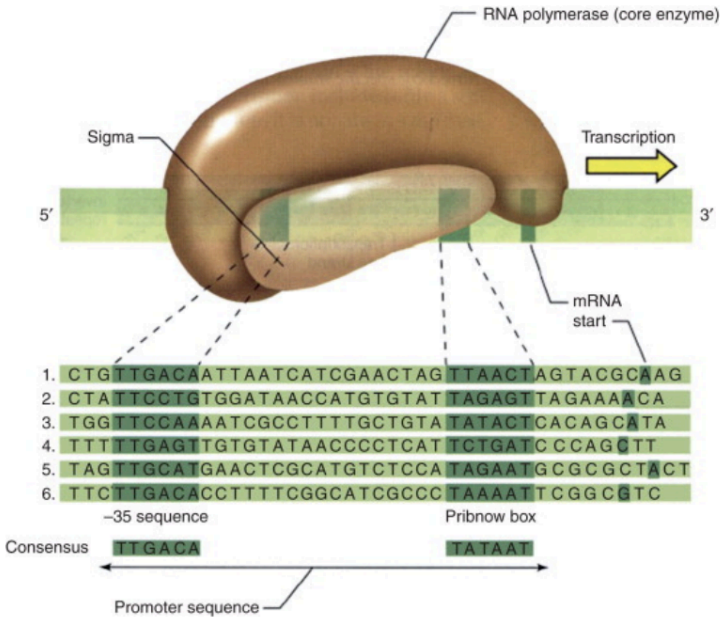

Figure 2. A representation of the RNAP and sigma factor complex. The complex binds the promoter region upstream of the transcription start site. The sigma factor recognizes these two regions, the -35 and the -10, known as the sigma factor consensus sequence. Different consensus sequences exist for different sigma factors (from *Encyclopedia of Microbiology*, 3rd Edition, 2009).

1-2-1 General Response Sigma Factors

Different sigma factors have been identified to turn on genes during different phases of a bacterium's development. In *Bacillus subtilis,* sigA is the sigma factor responsible for regulating the "housekeeping" genes, or genes necessary for normal, vegetative growth. Another sigma factor, sigH, controls postexponential phase gene expression and competence and early sporulation genes (Haldenwang, 1995). SigB is the sigma factor responsible for turning on stress-response genes, activated when the bacterium is exposed to environmental, chemical stress or other unfavorable conditions. Specifically, sigB activates a regulon of over 100 genes that form part of a usually transient response to stressful conditions (Carniol et al., 2004). SigB controls the GSR (general stress response), a short-lived induction of general stress proteins

(GSP), separate from the time-consuming sporulation response, although both can be activated by similar environmental signals. SigB is induced after a particular stressor halts or inhibits the rate of growth. GSPs under sigB control include catalases like KatX, heat stress resistance proteins (ClpC and ClpP), protein synthesis under stress (Ctc), and many more (Rodriguez Ayala et al., 2020).

1-2-2 Sporulation-specific Sigma Factors

As discussed earlier, sporulation is another response to unfavorable conditions, however as this process is irreversible and more complex, different sigma factors are involved in initiating this transition and guiding the cell through the process. The sigma factors involved in sporulation are known as alternative sigma factors and include sigF, sigE, sigG and sigK that are activated in either the forespore or the mother cell during different developmental stages. SigF is the first sigma factor to be activated in the forespore after asymmetric division which then creates an intercompartmental signalling cascade activating sigE in the mother cell. After the engulfment stage, sigG replaces sigF in the forespore which in turn activates sigK in the mother cell during late-stage sporulation (Carniol et al., 2004; Pan & Losick, 2003; Piggot & Hilbert, 2004; Riley, Schwarz, et al., 2021).

1-3 Metabolic Shutdown During Sporulation

The mature spore is much more metabolically limited, and thus its development involves a breakdown of enzymes and proteins involved in metabolic pathways it no longer needs. Spores do not rely on the same metabolic pathways and thus do not make their own metabolites via glycolysis. An important part of a cell's metabolic shutdown as it enters sporulation is proteolysis, wherein these metabolites and enzymes involved in metabolism (e.g. the Krebs

cycle) are degraded. Because the metabolically dormant spore does not have the enzymes needed to carry out the TCA cycle and generate ATP or synthesize amino acids and ribonucleotides, it was assumed that they were being degraded (Kornberg et al., 1968; Setlow, 2025; Swarge et al., 2020).

1-4 Proteolysis in Bacteria

Proteolysis, or the degradation of proteins, is a process that is essential for all living organisms and in both eukaryotic and prokaryotic cells. Proteins are degraded for a variety of reasons including during cell death, targeting misfolded proteins and as steps in cellular signaling processes where polypeptides need to be broken down into smaller peptides. In bacteria, regulated proteolysis is used to destroy and activate regulatory proteins and to produce signaling molecules, while general proteolysis involves the degradation of damaged proteins to maintain homeostasis (Konovalova et al., 2014). It is also a good method of regulation when a quick response is needed such as under stress conditions, as transcription and translation are not involved.

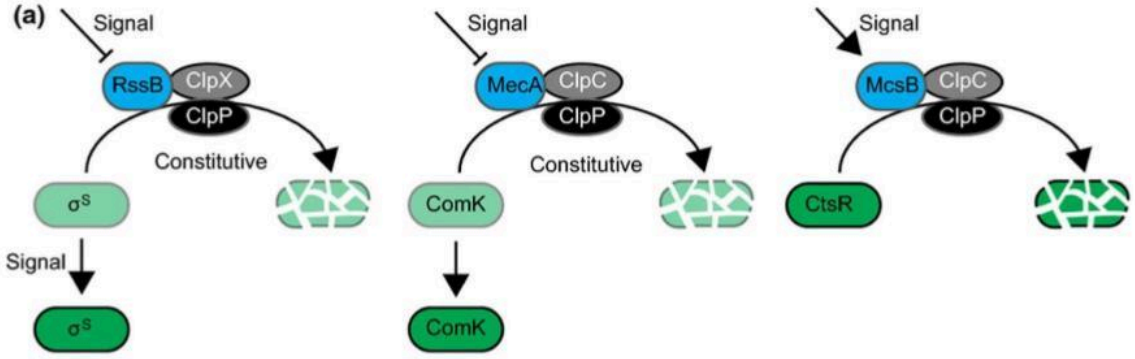

Figure 3. Principles of regulated proteolysis from stress response studies. In all three examples, the level of green indicates the accumulation level of the substrate. Adaptors are indicated in blue, proteases in black, and AAA+ proteins in gray. ComK and σ^S are synthesized under nonstress conditions due to inhibition of the protease complex while CtsR is degraded under nonstress conditions due to activation of the protease complex (from *FEMS Microbiol Rev*, Volume 38, Issue 3, May 2014).

Regulated proteolysis is an important part of a cell's metabolic shutdown as it enters sporulation. It allows the cell to modulate this developmental process by controlling the amount of enzymes and substrates, which can then have a cascading effect. As this is an irreversible process, proteases, or the enzymes that catalyze this breakdown, are tightly controlled to select substrates with specificity. Specifically, in bacteria, regulated proteolysis is carried out by AAA+ proteases (ATPases associated with cellular activities) where activity is regulated by ATP hydrolysis. They are composed of an unfoldase domain with an ATP-hydrolysis active site that recognizes and unfolds the target protein, which is then translocated through a central pore to the chambered peptidase domain where it is degraded (Mahmoud & Chien, 2018).

1-4-1 Clp-family Proteases

There are several conserved AAA+ protease complexes found amongst bacteria. Of these, one of the most common is the Clp family. Clp family proteases are multi-subunit

complexes where the unfoldase and peptidase domains are not encoded on the same polypeptide chain, but rather exist as separate proteins that then assemble to form the complex (Konovalova et al., 2014; Mahmoud & Chien, 2018). Clp ATPases are hexameric and include ClpA and ClpX in gram-negative bacteria and ClpX, ClpE, and ClpC in gram-positive bacteria. They associate with the 14-subunit peptidase ClpP, a barrel-shaped structure with stacked heptameric rings forming a core that contains the proteolytic active sites which degrade the substrate once it has been translocated. ClpP can individually degrade small peptides but must associate with the ATP-dependent unfoldase in order to degrade larger, more stable proteins, creating the chaperone-protease AAA+ complex (Konovalova et al., 2014). Clp complexes are often involved in degradation during sporulation in *Bacillus subtilis,* most especially ClpXP and ClpCP.

1-4-2 Adaptor Protein-dependent Complexes

While these complexes do have intrinsic specificity for their substrates via recognition of short sequences known as degrons, they can be further specified by adaptor proteins that increase selectivity of substrates and promote or inhibit activity. These adaptor proteins can be grouped into two functional groups that include adaptor proteins that activate the protease and those that deliver substrates to an already active protease (Kuhlmann & Chien, 2017). For example, ClpXP is a protease complex found in both *E. coli* and *B. subtilis*. In *E. coli,* it both degrades incomplete peptides resulting from stalled ribosomes and regulates the stress response in this bacterium. It utilizes an adaptor protein, SspB, that forms a homodimer to act as a tether attaching substrates to ClpX, increasing the effective concentration and making the protease more efficient (Mahmoud & Chien, 2018). In this case, SspB acts as an adaptor protein that delivers substrates to the protease. ClpXP also plays a role in degradation during sporulation of *B. subtilis* with adaptor protein CmpA. Specifically, it plays a role in proteolysis to regulate sporulation by

degrading the structural protein SpoIVA in sporulation-defective cells (Mahmoud & Chien, 2018).

1-4-3 The ClpCP Complex

The ClpCP complex is one of these adaptor protein-dependent AAA+ protease complexes found in *Bacillus subtilis.* The ClpCP complex, which can also associate with multiple adaptor proteins, requires an adaptor protein in order to activate ClpC ATPase function (Kuhlmann & Chien, 2017). ClpC rests in a decamer state until binding of an adaptor induces hexamerization, the active state (Carroni et al., 2017). Adaptor proteins associated with ClpC include MecA, McsB and MdfA. McsB is an arginine kinase whose phosphorylated products can both activate ClpC and target substrates for degradation by ClpCP, as phosphorylated arginine residues are what target substrates for recognition by ClpC (Kirstein et al., 2007; Kuhlmann & Chien, 2017; Massoni et al., 2025; Trentini et al., 2016). MecA regulates ComK, which regulates genetic competence, or the uptake of exogenous DNA by activating genes required for this process. MecA binds to both ComK and ClpC, targeting ComK for degradation when these genes are not needed. However, when they are, MecA is inactivated, allowing ComK to accumulate (Konovalova et al., 2014). MdfA, which has been recently identified, specifically binds to and activates ClpC in the forespore, in order to induce protein degradation during sporulation (Massoni et al., 2025).

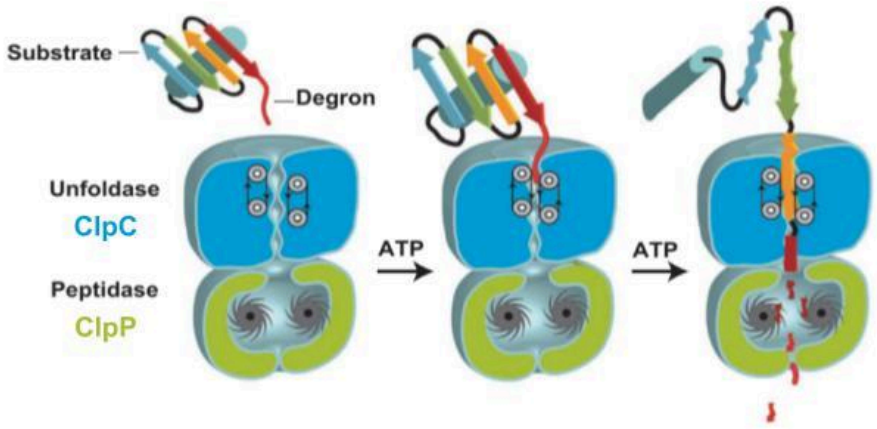

Figure 4. ClpCP. The ClpCP proteolytic complex is made up of ClpC, an ATPase with an active site that recognizes and unfolds the target protein via ATP-hydrolysis, and ClpP, a protease/peptidase that degrades the polypeptide (from Sauer & Baker, 2011).

1-5 Role of the ClpCP Complex During Sporulation

ClpCP, and other proteolytic complexes play an important role in degradation to control signals and concentrations of metabolites when guiding the bacterium through the process of sporulation. The transition into sporulation is ultimately governed by the transcriptional master regulator Spo0A. A phosphoryl group is transferred to Spo0A via phosphotransferases Spo0F and Spo0B from histidine kinases KinA-KinE which are activated in response to different environmental stressors (Tan & Ramamurthi, 2014). Official entry into sporulation is characterized by the activation of sigF and sigE by Spo0A~P prior to asymmetric division. SigF is held inactive by the anti-sigma factor SpoIIAB but after asymmetric division, SpoIIE dephosphorylates the anti-anti-sigma factor SpoIIAA (previously phosphorylated by SpoIIAB). This allows it to act on SpoIIAB, thereby getting rid of the inactivation of sigF. SigE is also initially produced in both compartments as an inactive pro-sigE precursor, but is specifically activated only in the mother cell by forming a multimer with SpoIIGA and SpoIIR. SpoIIR is produced in the forespore under control of sigF and is secreted into the intermembrane space of

the septum where it binds to and activates SpoIIGA. SpoIIR needs to be acetylated in order to activate SpoIIGA which can only occur on the mother cell side. Only SpoIIR localized there will be able to activate sigE, thereby selectively activating sigE only on the mother cell side (Tan & Ramamurthi, 2014).

The process of transitioning a vegetatively growing cell to a dormant spore also involves metabolic reprogramming of the forespore in order for it to reach a metabolically dormant state. This process is aided in large part by the mother cell, specifically through a phenomenon known as the "feeding tube" model. This involves a gap junction-like feeding tube that connects the mother cell and the forespore, multimeric protein channels formed by mother cell proteins encoded by the spoIIIA operon and the forespore protein encoded by spoIIQ. Through this channel the mother cell is able to provide small molecules to sustain the forespore as it enters metabolic dormancy (Riley, Lopez-Garrido, et al., 2021).

This is part of how spores reach metabolic shutdown- helped by the degradation of proteins by proteases like ClpCP. Recent data confirms that ClpCP, along with adaptor protein MdfA, are required for degrading metabolic enzymes and other proteins such as transcriptional regulators Spo0A and SpoIIAB, as well as sigF itself (Kain et al., 2008; Massoni et al., 2025; Riley et al., 2025). Spo0A controls entry into sporulation itself while SpoIIAB inhibits sigF. This makes ClpCP an important regulator in the timing of both the initiation of sporulation and processes undertaken by the forespore in early sporulation. It continues to play a role during metabolic shutdown. For example, a single amino acid mutation in the *clpC* gene was sufficient to prevent the synthesis of the TCA enzyme citrate synthase (CitZ) in the forespore (Riley et al., 2025).

1-6 Sigma Factor-specific Regulation of the *clpC* Operon

If ClpCP plays a role in regulating entry into sporulation and processes in the forespore to further this pathway, we would assume that it would be somehow specifically expressed during these stages. Transcriptional regulation of the *clpC* and *clpP* genes is a question of interest because if we understand what sigma factors control these genes, we can understand more about where and when they are expressed. Previous thesis work in the lab has focused on the transcriptional regulation of the *clpC* operon.The *clpC* operon is made up of 4 genes: *ctsR, mcsA, mcsB* and *clpC. ctsR* is controlled by an autoregulatory feedback loop with the product of *ctsR* binding to the CtsR binding site (shown by a black box in Figure 5) repressing the transcription of the operon under normal conditions. Under stressful conditions, the repression is removed. There are four sigma factor-dependent promoters that control the operon: two sigA-dependent promoters, one sigB promoter and one sigH-dependent promoter located in the *mcsB* gene (Berg, 2023; Mayangsari, 2018; Nkomboni, 2017; Quigley, 2023).

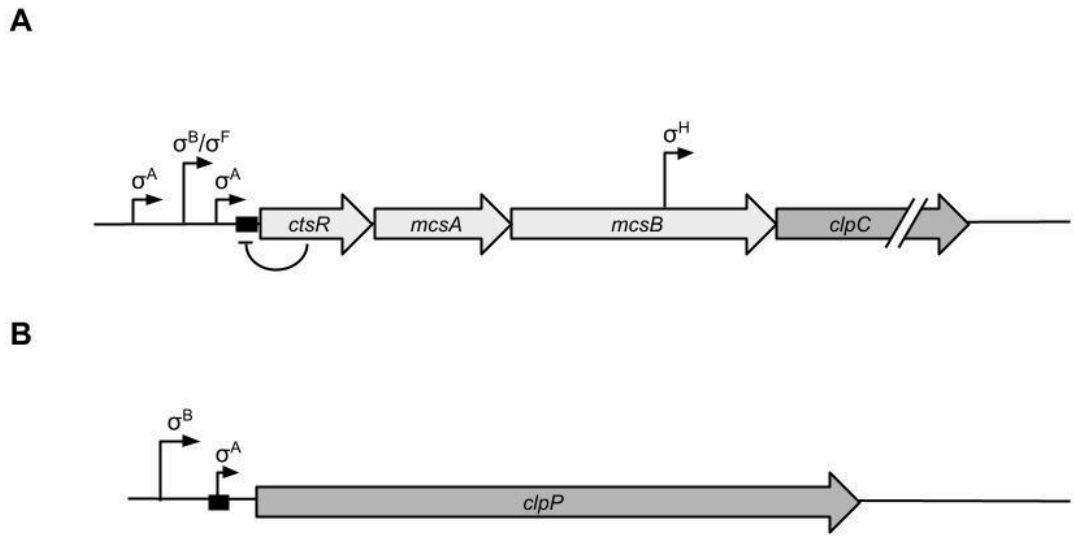

Figure 5. The *clpC* (A) and *clpP* (B) promoter regions. *clpC* and *clpP* are located at different (not adjacent) locations on the genome. Although *clpP* is not an operon because it contains one gene (*clpP*), while the *clpC* operon contains four, it is also controlled by CtsR and there is a CtsR binding site located upstream of the *clpP* gene.

Additional consensus sequence analysis of this sigB-dependent promoter indicated that it could also be a match for sigF. Previous work by myself and other colleagues showed that this promoter is specifically activated by both sigB and sigF, as well as specifically during sporulation (Nkomboni et al., 2025). This sigF/sigB promoter was also shown to be responsible for the increase in *clpC* expression in the forespore, as a sigB/sigF promoter knockout diminished *clpC* expression (Panchal, 2025). This gives strong evidence that not only does ClpC play a role in protein degradation during sporulation, but its gene, *clpC,* is also specifically expressed during the forespore by sigF, the forespore-specific sigma factor. We also know that sigF activates *mdfA* and that ClpCP and MdfA interact in the developing forespore to play a necessary role. This suggests that all components of this complex, including *clpP*, would need to be expressed during this stage.

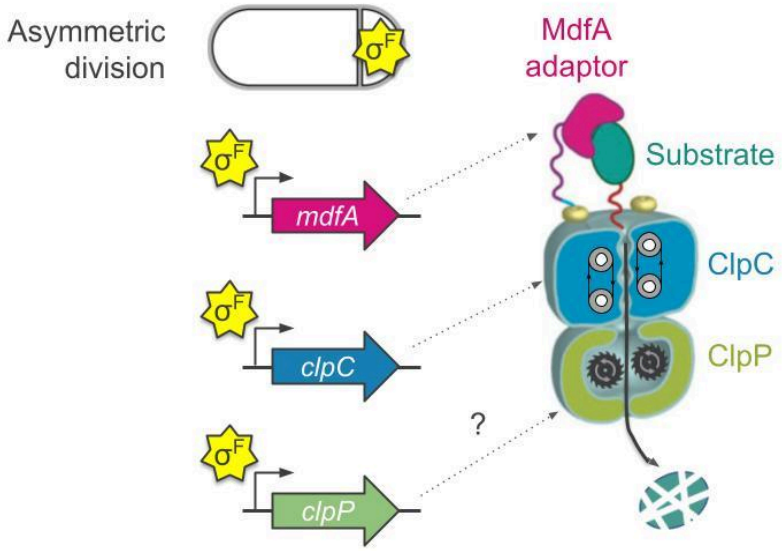

Figure 6. A schematic diagram linking the production of sigF in the forespore to the ClpCP-MdfA adaptor complex. The knowledge that *mdfA* is activated by sigF in the forespore, as well as unpublished data that *clpC* is also activated by sigF in the forespore leads to the hypothesis that *clpP* must be as well (adapted from Nkomboni et al., 2025).

1-6-1 Sigma Factor-specific Regulation of *clpP*

The *clpP* promoter sequence contains only one gene, in contrast to the *clpC* operon. However, *clpP* is also regulated by *ctsR,* a transcriptional repressor that inhibits the expression of both *clpC* and *clpP* under non-stress conditions. Modifications to CtsR during stress conditions by McsA and McsB lead to derepression of the operons as it now cannot bind to the CtsR binding site. This helps enhance stress induced expression of these genes (Krüger & Hecker, 1998). During sporulation, ClpP plays a necessary role by interacting with ClpC and MdfA. Previous studies also confirmed that mutants lacking *clpP* had a sporulation-deficient phenotype (Gerth et al., 1998). Additionally, sequence analysis of the *clpP* operon suggests the presence of three different sigma factor-dependent promoters, including sigA, sigE and an overlapping

sigB/sigF, akin to *clpC.* This suggests that *clpP* could also be activated by sigF in the forespore, and also by sigE in the mother cell.

1-6-2 Other Dual-Activated sigB/sigF Promoter Sequences

Furthermore, the dual-activation of known sigB-dependent promoters by sigF is a phenomenon that has been documented in other *B. subtilis* genes. Many of these sigB-induced proteins are also under dual transcriptional control by other regulatory proteins. For example, the sigB-dependent promoter of the *ctc* gene and the sigB-dependent promoter of the *katX* gene are both dual-activated (Carniol et al., 2004; Petersohn, Bernhardt, et al., 1999). KatX is a catalase responsible for resistance to hydrogen peroxide in germinating and outgrowing spores and so is recognized by sigF. However, it was also demonstrated to be sigB-transcribed which makes sense, since it is induced by heat, salt and ethanol stress as well as energy depletion (Petersohn, Engelmann, et al., 1999). This hypothesis was developed with the realization that the *katX* promoter had overlap with both the sigB and sigF-dependent consensus sequences. As mentioned above, recent data from the Camp lab shows that the *clpC* gene is activated by both sigB and sigF, which emphasizes its importance as both a GSP involved in heat shock and as important during sporulation (Nkomboni et al., 2025). These overlaps suggest that the sigF regulon expressed in the forespore and the sigB regulon expressed under stressed and starving conditions do not exist independently from each other. However, the nature of their relationship is still unknown.

1-6-3 Assessing the Transcriptional Control of *clpP* and Other sigB-dependent Promoters

My aim was to test the hypothesis that *clpP* is expressed under the control of sigF and sigE during sporulation. The sporulation-specific expression of ClpCP's peptidase domain, *clpP*, is an open question. Through analysis of the *clpP* promoter region, we found that the sigB-dependent promoter also matched the consensus sequence for sigF, raising the possibility of forespore-specific expression of *clpP* as well. We also found a region that matches a sigE consensus sequence, a sigma factor expressed in the mother cell during sporulation. We aim to test these questions in order to increase our understanding of ClpCP's role during *Bacillus subtilis* sporulation. In order to do this, I characterized both the sigB/sigF promoter and the sigE promoter.

We also are interested in investigating the question of dual-activated sigB and sigF promoters to understand the mechanism by which they are related, and to answer the question why would a promoter need to be activated by both sigB and sigF. I characterized two known sigB-dependent promoters with overlapping consensus sequences for sigF, *sigB* and *gspA,* furthering work by another lab member.

CHAPTER 2: MATERIALS AND METHODS

2-1 General Methods

Bacterial strains were streaked and plated on LB agar plates and incubated at 37°C to grow overnight and LB plates containing appropriate antibiotics were used as needed for strain selection. Strains were inoculated in liquid LB or other specified media and grown on a roller drum in a 37°C incubator. Strains were stored as glycerol stocks at -80°C in cryo tubes. Glycerol stocks were made by mixing 900ul of culture with 600ul sterile 50% glycerol.

2-2 Plasmid Construction: pBW3 Assembly

The plasmid pBW3 was constructed from plasmid backbone pAH124 by transforming it with a "gblock" DNA fragment (gBW1) using *E. coli.* This fragment of DNA was designed using the FastCloning method described by Li et al. (2011) to contain the sigB promoter sequence, the upstream sigE promoter sequence, the ribosome binding site (RBS), *clpP* start codon and additional overhangs of 15-17 bases overlapping with the vector ends. The plasmid pAH124 was digested with EcoRI and HindIII to cut the DNA at the insert regions. NEBioCalculator was used to determine the amount of gBW1 needed to get a plasmid:insert ratio of 1:5. This 1:5 mixture of digested pAH124 (4 ul) and P_{clpP} fragment (1 ul) was transformed into *Escherichia coli.*

2-2-1 Transformation into *E. coli*

NEB 10-beta competent *E. coli* cells were thawed on ice to which the 5 ul of plasmid-insert mixture was added. The cells were then incubated on ice for 30 minutes, heat shocked at 42°C for 30 seconds and incubated on ice again for 5 minutes. The cells were

recovered by adding NEB 10-beta Stable Outgrowth Medium and incubating on a roller drum at 37°C for 1 hour. 100 ul of the mixture was plated onto an LB agar plate supplemented with ampicillin (LB/amp). The remainder was pelleted and resuspended in 200 ul and this more concentrated mixture was plated on another LB/amp plate. The plates were incubated at 37°C overnight. Colonies appeared on both plates, with more on the more concentrated one, indicating that the *E. coli* cells took up the plasmid containing an ampicillin resistance gene.

2-2-2 Plasmid Isolation and Verification

Colonies from selection plates were cultured overnight in 3 ml of LB/ampicillin. Plasmid DNA was isolated from colonies using the NucleoSpin Plasmid Miniprep kit. Prepped samples were sent for sequencing to confirm uptake of the *clpP* promoter region. The plasmid that did successfully take up this promoter was sent for whole-plasmid sequences to confirm that there were no unintended mutations. This plasmid was named pBW3.

Table 1. Plasmids used in this study

Name	Genotype	Construction/Source
pBW3	gBW1→pAH124 *amyE::PclpP-lacZ cat, amp*	Constructed by transforming "gblock" DNA fragment-gBW1 containing *clpP* sigF/B and sigE promoter regions

2-3 Isolating Chromosomal DNA

Chromosomal DNA was isolated following a modified version of Protocol III.G "Isolating Genomic DNA from Gram Positive and Gram Negative Bacteria" from Promega Wizard Genomic DNA Purification Kit. Single colonies were inoculated in 3 ml of LB and incubated on a roller drum at 37°C for 3-4 hours. 1.5 ml were transferred to a micro centrifuge

tube and centrifuged at 13000xg for 2 minutes, the supernatant discarded and the remaining 1.5 ml spun down on top. The cell pellet was resuspended in 50 mM EDTA to which 20 mg/ml of lysozyme was added. The tubes were incubated at 37°C for 30-60 minutes. Next, Nuclei Lysis Solution was added and the tubes were incubated at 80°C for 5 minutes to lyse cells, then cooled to room temperature on the bench. 1.2 ul of 10 mg/ml RNase solution was added and tubes were incubated once again at 37°C for 15-60 minutes then cooled on the bench. Next, 200 ul of Protein Precipitation was added and tubes were inverted vigorously and then vortexed until a white precipitate was formed and the mixture became less viscous. The mixture was incubated on ice for 5 minutes and centrifuged at 13000xg for 3 minutes. 900 ul of this supernatant was transferred to a micro centrifuge tube containing 600 ul of room temperature isopropanol. This mixture was inverted gently until thread-like strands of DNA appeared, then centrifuged at 13000xg for 2 minutes to collect the DNA at the bottom of the tube. The supernatant was discarded and the tube was spun again for 1 minute at the same speed. The remaining liquid was aspirated then the pellet was left to air dry overnight in the laminar flow hood. The pellet was finally left to rehydrate in 100 ul of elution buffer for 10 hours and vortexed to mix. DNA samples were stored at -20°C.

2-4 *Bacillus subtilis* Strain Construction

The strains constructed for this study involved creating a *clpP-lacZ* construct with the *clpP* promoter region inserted upstream of the *lacZ* gene in order for the *lacZ* translated to be quantified later. Some strains also contained specific IPTG-induced sigma factors, including IPTG-inducible sigB, IPTG-inducible sigF and IPTG-inducible sigE.

2-4-1 Transformation into *B. subtilis*

DNA was transformed into *B. subtilis* following the "One-Step Bacillus subtilis Transformation" protocol adapted by the Losick lab and based on the method described by Wilson and Bott (1968). Single colonies of the recipient strains were inoculated in 1x Modified Competence (MC) Medium containing 1 M MgSO4 and incubated for 4-5 hours on the roller drum at 37°C. 200 ul of cells were aliquoted for each transformation to which 10 ul DNA was added in 1:100 and 1:1000 concentrations. The transformations were incubated on a roller drum for 60-90 minutes then plated onto LB agar plates containing the appropriate antibiotic. The plates were incubated overnight and subsequent colony growth indicated successful uptake of the plasmid containing an antibiotic resistance gene.

2-4-2 *clpP-lacZ* Strain Construction

Strains containing the *clpP-lacZ* promoter region were selected on LB agar plates containing chloramphenicol as the pBW3 plasmid contains a chloramphenicol resistance gene. The plasmid should have integrated at the *amyE* site which would render the strains unable to digest starch. Thus, the strains were patched onto LB/starch agar plates and stained with iodine. Some parent strains also contained an "alternative" *amyE* site containing a spectinomycin resistance gene. We want integration at the native *amyE* site so we also patched onto LB agar plates supplemented with spectinomycin to ensure the strains were still spectinomycin resistant. Multiple isolates fitting these checkmarks were stored as glycerol stocks.

2-4-3 Sigma Factor Knockout Strain Construction

From the original *clpP-lacZ* promoter-containing *B. subtilis* strains constructed, a sigma factor knockout mutants were constructed by isolating chromosomal DNA from a ΔsigB::Kan

strain and transforming into a strain containing the promoter region. After transformation, ΔsigB strains were plated onto LB/kanamycin agar plates with colony growth indicating uptake of the novel DNA. Multiple isolates of each strain were stored as glycerol stocks.

Table 2. *B. subtilis* strains used in this study

Name	Genotype	Construction/Source
BWB1 (AHB1)	Wild type	Camp Strain Collection
BWB5 (MHB11)	*ywrK::Tn917::amyE::Pspank-sigF lacI spc amyE::PspoIIQ-lacZ cat*	Mayang Hasibuan Strain Collection
BWB13	*amyE::P_{clpP}-lacZ cm*	pBW3→AHB6337 (Isolate 1)
BWB15	*P_{IPTG}-rsbV-rsbWΔ1-sigB-rsbX lacI erm amyE::P_{clpP}-lacZ cm*	pBW3→BWB22 (Isolate 1)
BWB17	*ywrK::amyE::P_{IPTG}-sigF lacI spc, amyE::P_{clpP}-lacZ cm*	pBW3→AHB6081 (Isolate 1)
BWB20 (AHB6275)	*ywrk::Tn917::amyE::Pspank-sigF lacI spc amyE::P_{13clpC}-lacZ cat*	Camp Strain Collection
BWB21 (AHB6289)	*P_{spac}-rsbV-rsbWΔ1-sigB-rsbX lacI erm amyE::P_{13clpC}-lacZ cat*	Camp Strain Collection
BWB24 (AHB6348)	*amyE::P_{spoIID}-lacZ cat, ywrK::amyE::$P_{hyperspank}$-sigE[17-239] lacI spc*	Camp Strain Collection
BWB30	*amyE::P_{clpP}-lacZ cm ywrK::amyE::P_{IPTG}-sigE lacI spc*	pBW3→BWB23 (Isolate 1)
BWB34	*ywrK::amyE::P_{IPTG}-sigF lacI spc, amyE::P_{clpP}-lacZ cm ΔsigB operon::kan*	AHB1129 chDNA→BWB17 (Isolate 1)
BWB38 (HBB3)	*ywrK::amyE::P_{IPTG}-sigF lacI spc, amyE::P_{sigB}-lacZ cat*	Hailey Brooks Strain Collection
BWB39 (HBB5)	*P_{IPTG}-rsbV-rsbWΔ1-sigB-rsbX*	Hailey Brooks Strain

	lacI erm, amyE::P_{sigB}-lacZ cat	Collection
BWB40 (HBB9)	*ywrK::amyE::P_{IPTG}-sigF lacI spc, amyE::P_{gspA}-lacZ cat*	Hailey Brooks Strain Collection
BWB41 (HBB11)	*P_{IPTG}-rsbV-rsbWΔ1-sigB-rsbX lacI erm, amyE::P_{gspA}-lacZ cat*	Hailey Brooks Strain Collection

2-5 IPTG Induction Assays

IPTG (isopropyl beta-D-1-thiogalactopyranoside) was used to induce expression of specific sigma factors in order to test specific activation of the *clpP* promoter by this sigma factor. This was only done on strains that had an IPTG-inducible sigma factor construct. Colonies were inoculated in 3 ml of LB and incubated on a roller drum at 37°C for 3 hours. The cultures were then back-diluted into 10 ml of LB without IPTG or 10 ml of LB with 1 mM IPTG. These cultures were incubated in a water bath shaker at 37°C for 2 hours when two 1 ml samples were collected from each flask. The samples were centrifuged at maximum speed for 1 minute and the supernatant aspirated from each. The pellets were stored at -80°C.

2-6 Kinetic Beta-galactosidase Assay

Bacterial cell pellets from both the IPTG induction assay and the sporulation by resuspension assay had their beta-galactosidase output quantified via a kinetic beta-galactosidase assay (Bgal). The protocol followed comes from Camp and Losick (2009). Cell pellets were thawed and resuspended in Z-buffer with BME (beta-mercaptoethanol). The resuspensions were added to a clear, flat-bottomed 96-well plate and OD600 values were recorded. For processing pellets from IPTG-induction assays, a 10-fold serial dilution series was carried out (from 1:2 to 1:64). For processing timepoints from sporulation assays, only a 1:2 dilution was made for each

sample. Next, 50 ul of each dilution was transferred to a new 96-well plate with 50 ul of Z-buffer and 20 mg/ml lysozyme added. The plate was incubated for 20-30 minutes at 37°C to allow the cells to fully lyse. Pre-warmed 4 mg/ml ONPG, a substrate used to detect the enzyme beta-galactosidase, was added to each well at a volume of 20 ul. The plates were then immediately transferred to the plate reader where OD420 values were taken.

2-6-1 Data Analysis

For each well, the slope of the curve (mean velocity) during the linear phase of the reaction was calculated. Timepoints before 15 minutes and after 45 minutes were discarded, as well as timepoints with an OD value above 0.6. This mean velocity divided by the OD600 of the well taken previously was used to calculate the beta-galactosidase activity, correcting for light and dilutions as needed.

CHAPTER 3: RESULTS

3-1 Promoter Sequences Reveal Overlap with Sigma Factor Consensus Sequences

The first step of investigating the transcriptional regulation of the *clpP* operon was examining the promoter sequence upstream of the coding region of the *clpP* gene. Two promoters had already been identified and verified upstream of the *clpP* coding region, a sigA-dependent promoter overlapping with the CtsR binding site, and a sigB-dependent promoter. Each sigma factor has a preferred consensus sequence and any given promoter has a specific sequence that may or may not be a good "match" to the consensus for a particular sigma factor. For example, the *clpC* promoter region harbors an experimentally verified sigB promoter. We were able to look at the promoter sequence and see that it matched the sigB consensus. Interestingly, that same region of DNA also matched the sigF consensus sequence.

We performed a similar analysis on the previously verified sigB promoter of *clpP* by running the sequence through the DBTBS database. This database contains a collection of experimentally validated gene regulatory relations and the corresponding transcription factor binding sites upstream of *B. subtilis* genes (Sierro et al., 2008). Through this genomic analysis, we determined that this sigB promoter also matched the consensus sequence for sigF. The region upstream of this promoter also matched the consensus sequence for sigE. We wanted to isolate these two promoters when creating the "gblock" fragment as well as include the ribosome binding site and *clpP* start codon to ensure effective translation.

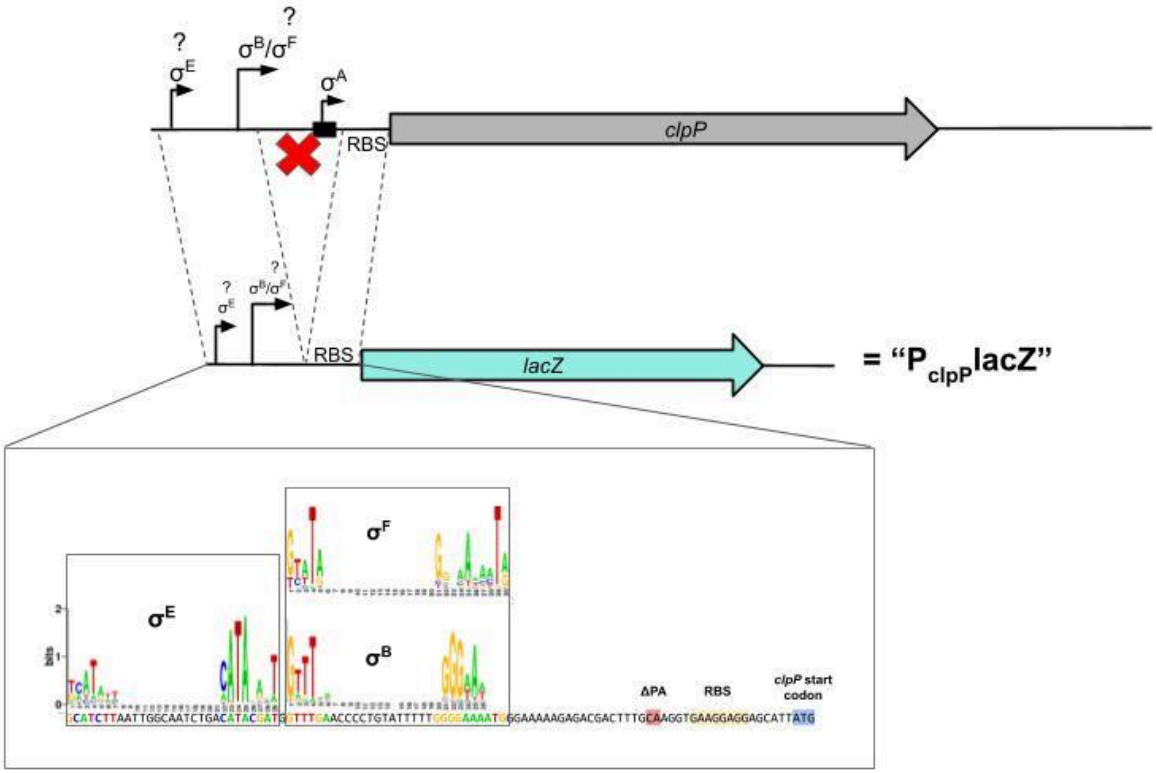

Figure 7. Sequence analysis of the *clpP* promoter regions. The gBW3 fragment used to make the PclpP-lacZ fusion contains both the regions that match the sigE and sigB/sigF consensus sequences, but does not include the sigA promoter region (ΔPA). It also includes the ribosome binding site (RBS) and the *clpP* start codon.

3-2 P_{clpP}-*lacZ* Fusion is Dual-Activated by Both SigB and SigF

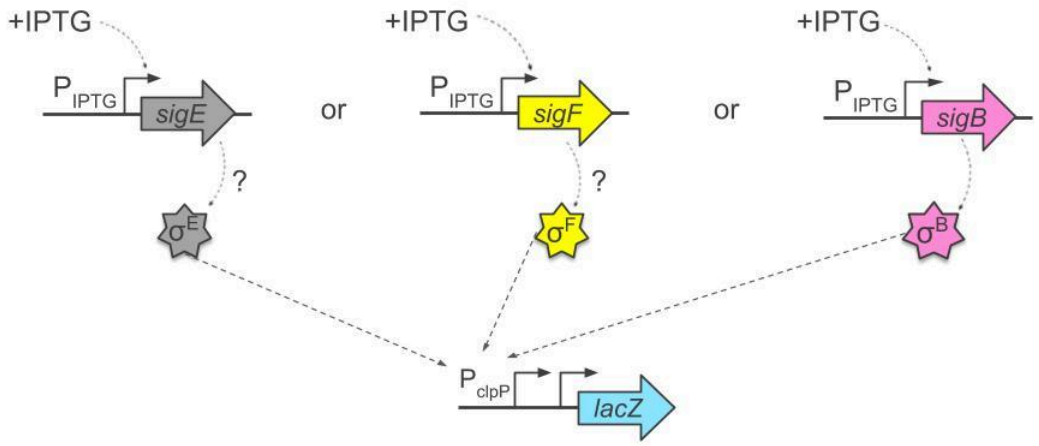

Figure 8. Constructing and testing IPTG-inducible sigma factor strains. IPTG induces the promoter of a specific sigma factor (sigB, sigF or sigE) to express that sigma factor. Since we have specifically turned on the expression of that sigma factor, we tested to see if it binds our P_{clpP} promoter and turns on the expression of *lacZ.*

In order to test the promoters of *clpP,* I fused the upstream region containing the sigB-dependent promoter (that also matches the sigF consensus sequence) and the putative sigE-dependent promoter to a *lacZ* reporter gene. This way, activation of these promoters would induce expression of this gene, which codes for the protein beta-galactosidase. This *lacZ* plasmid was transformed into various *Bacillus subtilis* strains with IPTG-inducible sigma factors. This way, we could induce the expression of a specific sigma factor by growing the strains in the presence of IPTG and then quantify the amount of beta-galactosidase produced. I transformed into an IPTG-inducible sigB strain, an IPTG-inducible sigF strain and an IPTG-inducible sigE strain and compared beta-galactosidase activity with or without IPTG. This allowed us to

determine if the promoter region could be specifically recognized by that sigma factor, since we were inducing the production of that sigma factor in excess, making it available in the cells.

As we would expect, the promoter fusion in a strain without an IPTG-inducible sigma factor ("no sigma factor") shows negligible beta-galactidosidase activity (Figure 9). However, we do see beta-galactosidase activity when the promoter is put in a strain with IPTG-inducible sigB, which we expected based on previous work showing that transcription initiation occurred at this sigB-dependent promoter (Figure 9; Gerth et al., 1998). Notably, we also see beta-galactosidase activity when the promoter is integrated in a strain with IPTG-inducible sigF, showing that this sigB-dependent promoter is also activated by sigF, as we hypothesized after examining the overlapping consensus sequences at this region.

However, inserting this promoter region into different IPTG-inducible sigma factor strains does not isolate expression of only that sigma factor. It simply creates an overexpression of that specific sigma factor. As we can see from the background beta-galactosidase present in the no-IPTG controls, we can assume that the promoter is being activated by other means. Therefore, I created a sigB knockout (ΔsigB) of my IPTG-inducible sigF strain. Interestingly, this strain exhibited higher beta-galactosidase activity than the non-ΔsigB version ("IPTG-inducible sigF"). It also shows decreased beta-galactosidase activity in the no IPTG control, both of which may be due to the lack of sigB (Figure 9).

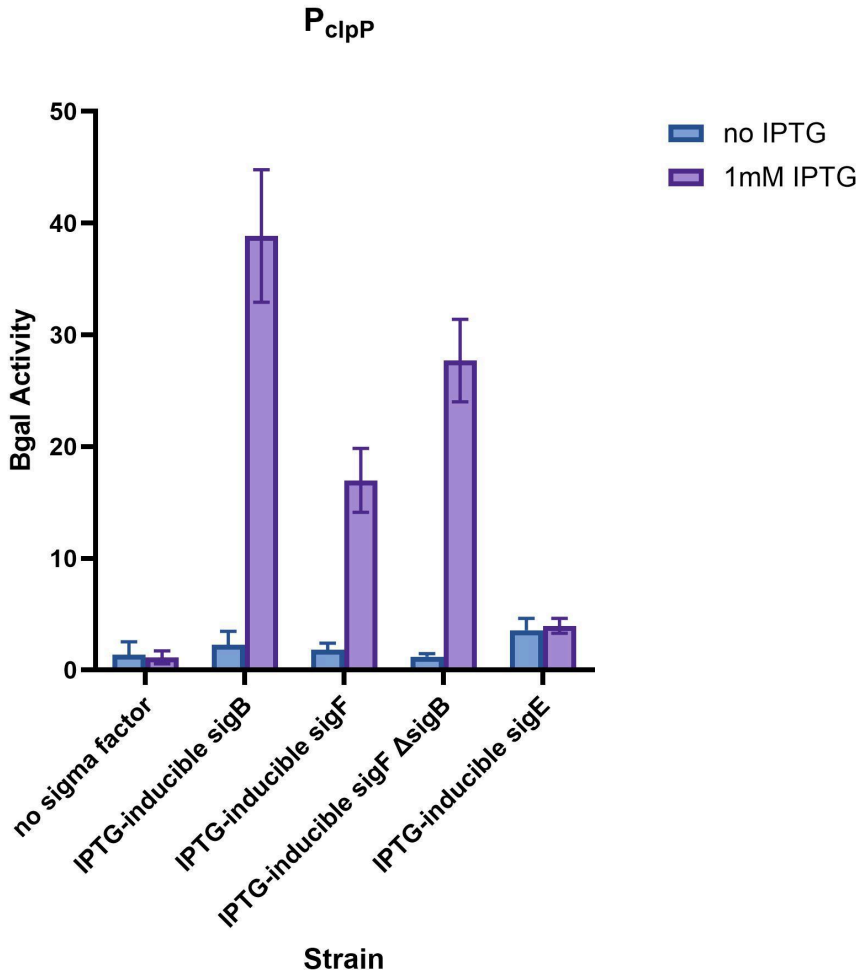

Figure 9. Beta-galactosidase activity of P_{clpP} promoter in different IPTG-inducible sigma factor strains. The P_{clpP} region was transformed into a control strain with no IPTG-inducible sigma factor, an IPTG-inducible sigB strain, an IPTG-inducible sigF strain, an IPTG-inducible sigF strain with a sigB knockout, and an IPTG-inducible sigE strain. Samples were grown with 1mM IPTG or no IPTG and beta-galactosidase activity was quantified by measuring OD420. Error bars were calculated from triplicate runs. Strains used from left to right include: BWB13, BWB15, BWB17, BWB34, and BWB30.

3-3 P_{clpP}-*lacZ* Fusion is Not Activated by SigE

Besides transforming the *clpP* promoter region into IPTG-inducible sigF and sigB strains, we also transformed it into an IPTG-inducible sigE strain, due to the identification of a putative sigE-dependent promoter upstream of the sigB/sigF promoter. However, we can see no difference between the no-IPTG and IPTG samples, which suggests no specific activation of this region by sigE, thus it is likely not a promoter region (Figure 9; Figure 10).

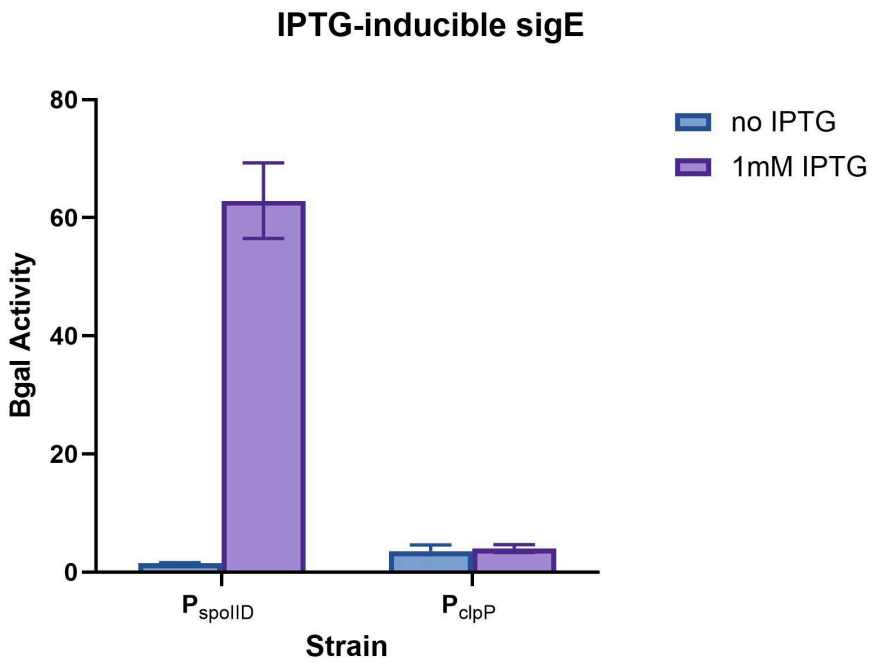

Figure 10. Specifically isolating IPTG-inducible sigE strains. The same P_{clpP} region in an IPTG-inducible sigE strain from Figure 9 is compared next to P_{spoIID}, a known sigE-dependent promoter. Error bars were calculated from triplicate runs. Strains used from left to right include: BWB24 and BWB30.

3-4 Other SigB-dependent Promoters Show Similar Activation by SigF

This phenomenon of dual-activated promoters by both sigB and sigF is something that we can see in other promoters of stress-response genes in *Bacillus subtilis.* For example, the P_{clpC} promoter region contains a promoter that is dually activated by both sigB and sigF (Figure 11). These results were confirmed previously by other lab members (Nkomboni et al., 2025), but my previous work was building this "P_{13clpC}" promoter. Similar to my construction of these P_{clpP} strains, the P_{13clpC} promoter region was transformed into various IPTG-inducible sigma factor strains. It showed more beta-galactosidase activity from both sigma factors than the P_{clpP} promoter, however, both are more activated by sigB than sigF (Figure 11). This is further supported by the results we see from other sigB-dependent promoters, unrelated to the ClpCP construct. These promoter *lacZ* constructs were made by a colleague, Hailey Brooks, by transforming "gblock" DNA fragments into the *lacZ* plasmid, pAH124. She then transformed these plasmids into IPTG-inducible sigB and sigF strains.

The genes that were chosen were *sigB,* the gene that encodes for sigB itself, and *gspA,* a general stress protein. Based on consensus sequence analysis, both of these genes have a promoter region that is recognized by sigB, but could also be a match for recognition by sigF. As expected, both promoters were significantly activated by sigB, however, they showed minimal activation by sigF (Figure 11). Whether or not P_{sigB} and P_{gspA} show specific activation by sigF is hard to conclude due to the insignificant differences in sigF-induced beta-galactosidase activity for both.

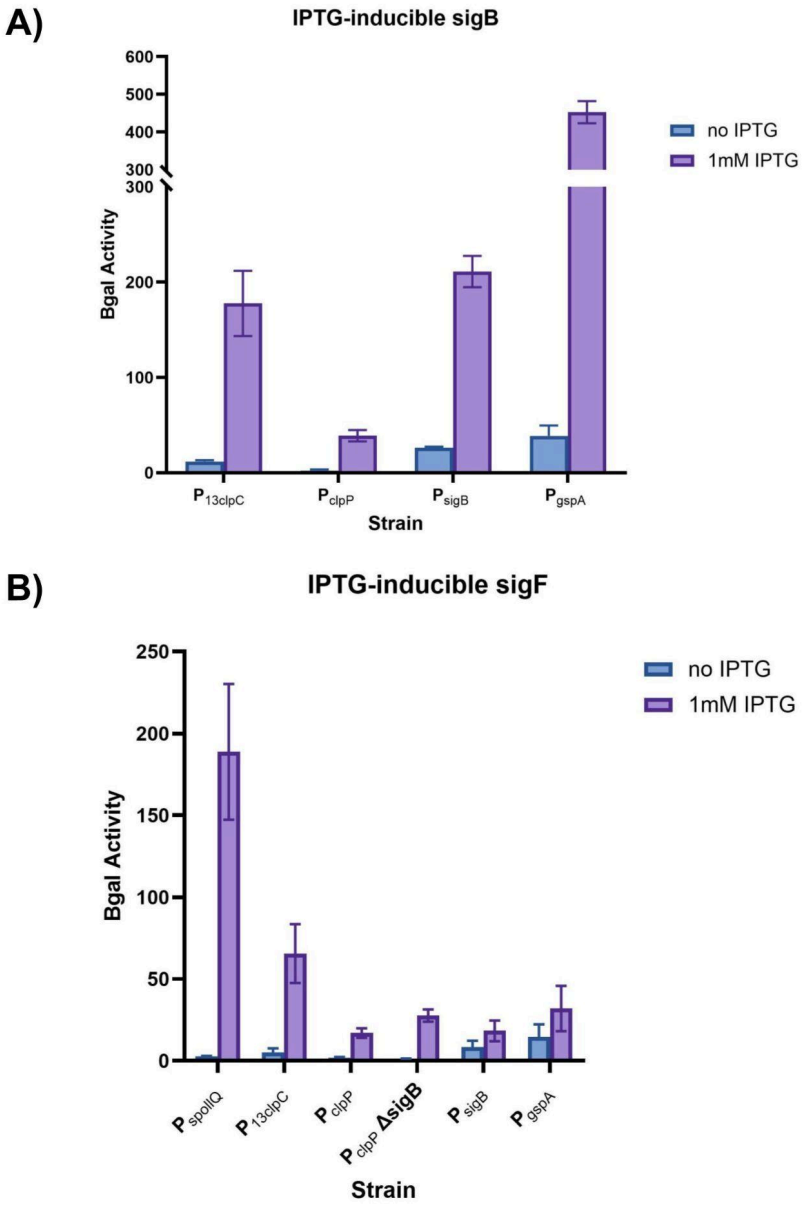

Figure 11. Comparison of activation of promoters by sigF versus sigB. The top graph shows the P_{13clpC} and P_{clpP} promoters as well as the P_{sigB} and P_{gspA} promoters in IPTG-inducible sigB strains. The gap and dashed line on the y-axis of A) represents a scale gap to accommodate for the much larger activation of P_{gspA}. The bottom graph B) includes these same promoters, along with P_{spoIIQ}, a known sigF-dependent promoter and the P_{clpP} sigB knockout, in IPTG-inducible sigF strains. Error bars were calculated from triplicate runs. Strains used from left to right include: A) BWB21, BWB15, BWB39, BWB41 and B) BWB5, BWB20, BWB17, BWB34, BWB38, and BWB40.

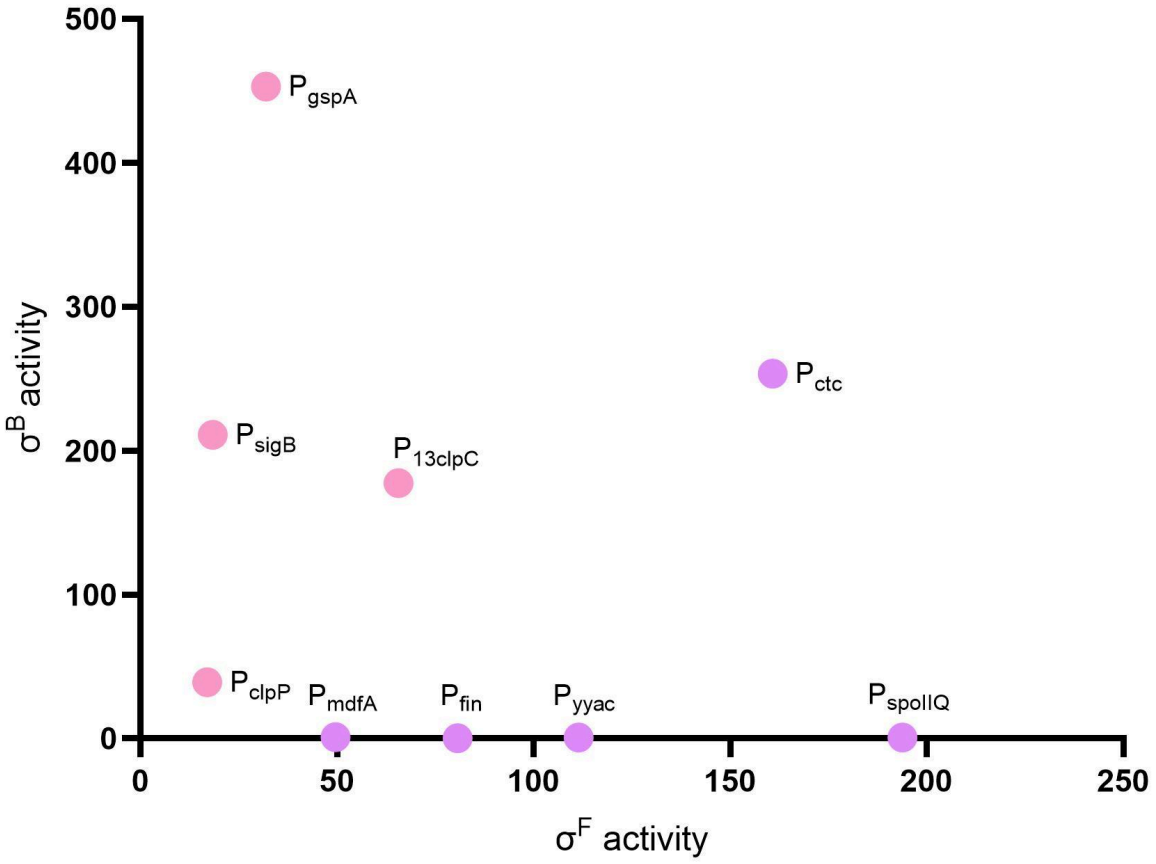

Figure 12. A scatter plot comparing sigB and sigF activity for different promoters. Data points in pink represent averaged data from triplicate beta-galactosidase assays from this study. Purple data points represent data from single runs conducted outside of this study.

In Figure 12, instead of comparing no IPTG and +IPTG for each strain, we are directly comparing the +IPTG results for promoters in IPTG-inducible sigB strains and the +IPTG results for promoters in IPTG-inducible sigF strains. In this format, we are able to compare the magnitude of activation of these promoters that match consensus sequences for both sigB and sigF. Notably, the pattern suggests that confirmed sigB promoters (P_{clpC}, P_{clpP}, P_{sigB}, P_{gspA} and P_{ctc}) are also able to be activated by sigF, as we see some sigF activity, with P_{ctc} showing more consistent levels of activation by both sigma factors. Meanwhile, promoters with previously

confirmed sigF-dependent activation (P_{mdfA}, P_{fin}, P_{yyac}, P_{spoIIQ}) do not show dual-activation by sigB.

CHAPTER 4: DISCUSSION

4-1 P_{clpP} is Recognized by Both SigF and SigB

The overarching results of this study show sigF-dependent activity of the P_{clpP} promoter, alongside the expected sigB-dependent activity. We are able to see this in the significant increase in beta-galactisodase activity when the promoter was inserted into an IPTG-inducible sigma factor strain, and the sample was induced with IPTG, turning on specific expression of that sigma factor. This ultimately supports our hypothesis that the P_{clpP} promoter, like the P_{clpC} promoter, is dual-activated by both sigB and sigF, supporting our model of expression of all three components of this MdfA-ClpCP protease complex in the developing forespore. However, our previous prediction that there is a sigE-dependent promoter upstream of this sigB/sigF promoter is not supported by our results, as there is no specific activation of the promoter when placed into an IPTG-inducible sigE strain and induced with IPTG to specifically express this sigma factor. This suggests that *clpP* is not specifically expressed in the mother cell during sporulation. However, this does not necessarily mean that ClpP-dependent proteolysis does not play a role in the mother cell. For example, ClpP, in complex with ClpX to form ClpXP, is shown to preferentially act in the mother cell during sporulation (Kain et al., 2008).

4-2 Some SigB-dependent Promoters Show Dual-Activation by SigB and SigF

The findings of this study demonstrate that there is an overlap between sigB-dependent promoters and sigF-dependent promoters, not just in the *clpC* and *clpP* promoters, but in other known sigB-dependent promoters as well. These known sigB-dependent promoters, P_{clpC}, P_{clpP}, P_{ctc}, P_{sigB} and P_{gspA}, all show additional activation by sigF. However, the same conclusion cannot be made for previously confirmed sigF-dependent promoters when testing their activation by

sigB. These promoters, P_{mdfA}, P_{fin}, P_{yyac} and P_{spoIIQ}, do not show dual-activation by sigB. This suggests that matching both consensus sequences does not guarantee activation by both sigma factors.

In order to discuss the nuances involved for the matching of a sequence of DNA to a sigma factor consensus sequence, I will use bioinformatic analysis of P_{clpP}, a dual-activated sigB/sigF promoter, P_{gspA}, a "mostly" sigB-dependent promoter, and P_{spoIIQ}, a strictly sigF-dependent promoter.

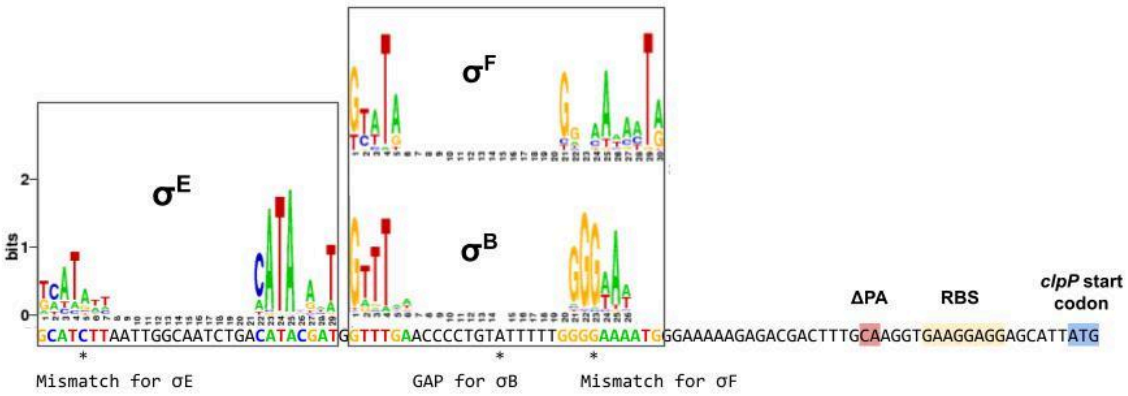

Figure 13. Overlay of sigB, sigF and sigE consensus sequences on P_{clpP} promoter sequence. The sequence of nucleotides for the verified sigB-dependent promoter of *clpP* is also a good match to the sigF consensus sequence, with only one mismatch. Additionally, the proposed sigE-dependent promoter region matches well with the sigE consensus sequence, with only one mismatch. When input into the DBTBS database, this sequence (above) matched the consensus sequences for all three sigma factors with a p-value < 0.01.

It is tempting to attribute the dual-activation of this promoter by both sigB and sigF to the significance of their matching to the consensus sequences of these sigma factors (Figure 13). However, the same phenomenon was not observed for the sequence matching the sigE consensus

sequence with the same amount of specificity, as we did not see any Bgal activity when we expressed sigE (Figure 9; Figure 10). This indeed suggests that there are other factors at play when it comes to the ability of a sigma factor to recognize and bind a specific promoter sequence.

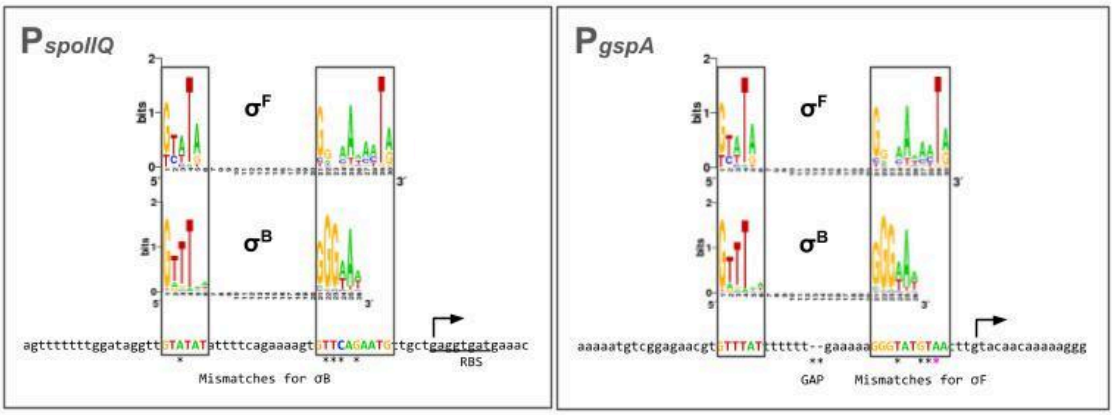

Figure 14. Overlay of sigB and sigF consensus sequences on P_{spoIIQ} and P_{gspA} promoters sequences. P_{spoIIQ}, a known sigF-dependent promoter, is a strong match to the sigF consensus sequence. However, it shows several mismatches with the sigB consensus sequence. Meanwhile, P_{gspA}, a known sigB-dependent promoter, is a strong match to the sigB consensus sequence ($p<0.01$). It also shows several mismatches with the sigF consensus sequence. (from Amy Camp).

P_{spoIIQ} and P_{gspA} further demonstrate the differences in actual recognition by a sigma factor compared to the extent to which a sequence matches the consensus sequence for that sigma factor. Both promoters have one sigma factor they have been experimentally determined to be recognized by and seem to be "equally" a match for the other (Figure 14). However, experimentally, they do not show "equal" levels of recognition by this sigma factor (Figure 12). P_{spoIIQ} is a strong sigF-dependent promoter but shows negligible levels of sigB-dependent activity while P_{gspA} is a strong sigB-dependent promoter and still shows some level of sigF-dependent

activity (Figure 12). These observations cannot necessarily be explained by differences in consensus sequence matching. Instead, it opens the possibility that binding of the sigma factor to the promoter regions favors some nucleotides over others. We still need to fully uncover the biochemical differences between sigB and sigF and their binding specificities in order to understand what is going on.

We can also discuss these differences on a larger, evolutionary scale, looking at the roles both sigma factors play in *Bacillus subtilis.* Both sigB and sigF are involved in regulating pathways as responses to unfavorable conditions. Activation of sigB is what induces a transient stress response in the bacterium in response to unfavorable conditions and activation of sigF is what controls the entry of a cell into the intricate and highly regulated process of sporulation, also in response to stressful or unfavorable conditions. We find that general stress response proteins, like ClpC and GspA, are also able to be activated by sigF, the forespore-specific sigma factor. While the general stress response and the sporulation pathway differ, stress response proteins could also play a useful role in an environmentally similar response, sporulation, as the cell must also cope with similar conditions such as nutrient deprivation during this process. However, sporulation-specific proteins like SpoIIQ, which is involved in forming a "feeding tube" between the forespore and the mother cell, do not make sense to be able to be turned on during a general stress response (Riley, Lopez-Garrido, et al., 2021).

There is also a possibility of sigF-induced expression of sigB, as demonstrated by the recognition of the P_{sigB} promoter by sigF (Figure 12). Therefore, I made and tested the P_{clpP} promoter in an IPTG-inducible sigF strain with a sigB knockout. If sigF activity went away, that would suggest that the sigF activity we saw in the IPTG-inducible sigF strain without the knockout could also include sigB activity, potentially from stressful conditions. However,

instead, we saw an increase in sigF activity in the sigB knockout, which may suggest that sigB is having some interaction with sigF, however, the nature of that interaction is not clear. We tested whether sigF could activate *sigB,* the gene that encodes for sigB, as bioinformatic analysis also showed that the verified sigB-dependent promoter upstream of *sigB* also matches the consensus sequence for sigF. While *sigB* is able to be activated by sigF, it does not seem to be significant enough when compared to its level of sigB activation or when compared to known sigF-dependent promoters (Figure 12). Further work needs to be done to determine the nature of the relationship between sigB and sigF, as these results seem to contradict one another.

4-3 Limitations and Future Directions

One large limitation to this study is being able to quantify the differences in sigma factor-dependent expression of different promoters. We are comparing the beta-galactosidase activity in response to activation of the P_{clpP} promoter by sigB, sigF or sigE. While we see different levels of beta-gal activity in response to different promoters, this does not necessarily mean that one sigma factor activates the promoter "more" than another, because the different sigma factors are transcribed and translated from different genetic pathways. There are many factors at play in order to get from the *lacZ* gene to the beta-galactosidase protein output, with recognition of the promoter by a sigma factor being only the first step. Since we are measuring the protein, or translation output, rather than the mRNA, the transcription output, we have to consider other factors like translation efficiency and availability of ribosome binding sites and can't quantify differences in beta-galactosidase activity across IPTG-inducible sigma factor strains and different *lacZ* reporters. Additionally, the competition of multiple mRNAs transcripts of different genes for a limited number of ribosomes affects translation efficiency and is not something that our study takes into account (Mauri & Klumpp, 2014).

Another limitation to the scope of this study is the lack of sporulation-specific data. While this study demonstrates that the P_{clpP} promoter is recognized by sigF and not by sigE, we cannot conclude that it is recognized by sigF in the forespore during sporulation (Figure 9). I began collecting preliminary data by transforming P_{clpP}-*lacZ* into the wild type strain and measuring beta-galactosidase activity after inducing sporulation. By putting the promoter construct into a strain without an IPTG-inducible sigma factor, we are testing whether or not the natural sigma factors that are produced during sporulation are sufficient to activate this promoter, rather than the artificially induced sigma factor expression when testing via IPTG-induction. We would also need to test various sigma factor knockouts, including a sigB knockout, a sigB/sigF double knockout and a sigB/sigE double knockout to ensure that it is sigF activating the promoter.

We would also like to assess how the sigB/sigF promoter influences *clpP* expression during sporulation. To do this, we would insert a *gfp* reporter downstream of the *clpP* gene with our P_{clpP} promoter, so that whenever *clpP* is transcribed, *gfp* also would be. Green fluorescence then will reflect *clpP* expression levels. This way, we can use fluorescence microscopy of sporulating cells to see differences in expression patterns and test if *clpP* is expressed via activation from this sigF-dependent promoter.

Previous work has measured the sporulation efficiency of a *clpP* knockout compared to the sporulation efficiency of the wildtype. Since ClpP plays an essential role in sporulation ability, a decreased sporulation efficiency of the *clpP* knockout was observed (Msadek et al., 1998). In the future, we should assess the sporulation efficiency of this P_{clpP} promoter alone. Since ClpP plays an essential role in sporulation ability and sigF should be driving *clpP* expression during sporulation, we would expect to see a decreased sporulation efficiency of the

P_{clpP} knockout. Similarly, if this sigF-dependent promoter controls expression of *clpP* during sporulation, we would expect to see a similar sporulation efficiency of the P_{clpP} promoter alone to the wild type.

4-4 Clinical and Public Health Applications

Bacterial spores, or endospores, are one of the most resilient cell types, resisting extreme environmental, chemical and UV conditions, allowing the bacterium cell to outlast suboptimal conditions. This evolutionarily favorable phenomenon is especially ideal for pathogenic types, allowing them to live on surfaces until they come into contact with a suitable host to colonize. The resistance of spores pose risks and threats, especially regarding the persistence and transmission of pathogenic spore formers such as *Clostridiodes difficile.* Spores are what make this disease so persistent and easily spread between hospitalized patients as spores intrinsically resist antibiotics, the immune system and disinfectants on hospital surfaces (Paredes-Sabja et al., 2014). Resistance to conventional antibiotics has also been found in spores of various *Bacillus* strains (Isawumi et al., 2023). Exposure to antimicrobial substances or clinical antibiotics could both select for resistant subpopulations and provide a stressful environment to trigger the formation of endospores. Since spores are easily disseminated they can be a vector for antimicrobial resistance genes (Romero-Rodríguez et al., 2024).

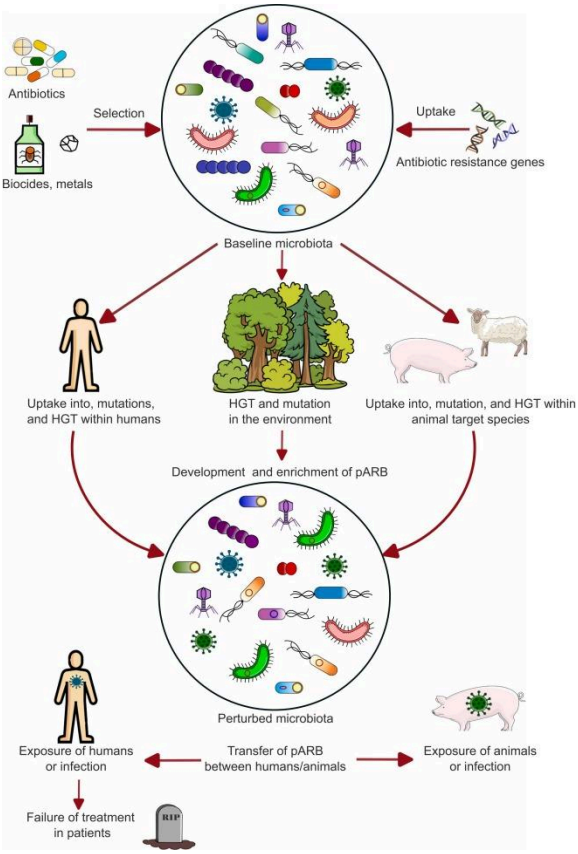

Figure 15. Pathways that contribute to antimicrobial resistance. Sporulation further contributes to this pathway of resistance by providing a resistant vehicle so these genes can be passed to new environments (from Romero-Rodriguez et al., 2024).

The families of *Clostridia* and *Bacillus* make up endospore-forming bacteria species. *Bacillus* contains *Bacillus subtilis,* the focus of this study and a harmless soil bacterium as well as a commensal member of the gut microbiome. *Bacillus* also contains *Bacillus anthracis,* another soil bacteria that causes a serious infection known as anthrax. *Clostridia* are anaerobic spore-formers, meaning that they do not need oxygen to form spores and therefore can germinate and survive within host tissues. *Clostridiodes difficile* causes *Clostridiodes difficile* infection (CDI), the leading cause of hospital-acquired diarrhea. It can also cause colitis, toxic megacolon, sepsis or even death. CDI has a high recurrence rate (20-30%) which is mainly caused by antibiotic-induced gut dysbiosis, or the destruction of commensal gut bacteria by the very

antibiotics used to treat the infection. This destruction of protective intestinal microbiota creates an environment that promotes outgrowth of the bacteria, toxin production and initiates a genetic programme for sporulation (Smits et al., 2016). *C. difficile* spores are able to survive in the intestinal mucosa and resist antibiotic therapies and are therefore the main reason for CDI recurrence (Queraltó et al., 2023). Similar to *B. subtilis,* once favorable environmental conditions resume, i.e. antibiotic therapy ends, the spores are able to germinate, proliferate and cause disease again.

4-4-1 ClpCP in *Clostridiodes difficile*: Possible Avenues for Treatment

Treatment for CDI mainly involves antibiotics such as metronidazole, vancomycin and fidaxomicin and occasionally, faecal microbiota transplantation. However, besides furthering recurrence as mentioned above, clinical antibiotic resistance and treatment failure is becoming increasingly common (Smits et al., 2025). In order to forgo the antibiotic path, new treatment options are being explored. One project that I was a part of during my summer research experiences involved the creation of a subunit vaccine against CDI, that both uses recombinant toxoids in order to prevent symptoms and surface-exposed and secreted proteins involved in binding and colonization. With this strategy, not only are the symptoms of the disease targeted, but also the factors that lead to recurrence. Previous targets have included flagellar basal rod and hook proteins FlgGEK, a cell wall hydrolase protein C40 peptidase 2, a germinant receptor protein CspC and M60 metallopeptidase that degrades complex glycoproteins, like mucins. These candidates were selected from previous research that determined that they were exposed on the bacterial surface, had a high antigenicity value, were conserved among strains and were expressed during infection (Nakjang et al., 2012; Noori Goodarzi et al., 2022). While many of these are common candidates for subunit vaccines, they also targeted key colonization and

persistence factors like motility and sporulation. Recently, a mucosal vaccination strategy including both toxin and non-toxin antigen subunits successfully cleared colonization in a murine model (Thomas et al., 2026).

ClpC, the chaperone protein that associates with Clp proteases, is conserved in *C. difficile.* Recent studies have shown that deletion of the *clpC* gene in *C. difficile* reduces sporulation in a manner analogous to that in *B. subtilis,* by inhibiting SpoIIAB degradation which impairs activation of sigF (Pan et al., 2001; Queraltó et al., 2023). ClpC is also important for motility and toxin production in *C. difficile*. While ClpC could not be a vaccine target as it is not likely to elicit an antibody response in patients, it could be a possible drug target, as its ClpP counterpart has been suggested as a new antimicrobial target. Specifically, there are two ClpP isoforms in *C. difficile,* ClpP1 and ClpP2, however, when and how ClpC interacts with them has not been determined (Lavey et al., 2019).

Based on this recent research by Queraltó et al. (2023) that demonstrates the essential role that ClpC and the *clpC* gene plays in *C. difficile* sporulation, and data from our lab that demonstrates ClpC's role in *Bacillus subtilis* sporulation (Panchal, 2025) and its corresponding association with ClpP during sporulation (Massoni et al., 2025; Riley et al., 2025) suggests that future research should focus on interactions between ClpC and the ClpP isoforms in *C. difficile,* and whether or not together they play a role during sporulation. If so, they could also prove to be a target for new antimicrobial therapies. Targeting sporulation ability is an important factor to consider for both immunotherapies like vaccines as well as targeted drugs.

REFERENCES

Berg, A, (2023). Investigating expression of the AAA+ protease component clpC in the developing Bacillus subtilis spore. Mount Holyoke College.

Carniol, K., Kim, T.-J., Price, C. W., & Losick, R. (2004). Insulation of the σF regulatory system in Bacillus subtilis. *Journal of Bacteriology*, *186*(13), 4390–4394. https://doi.org/10.1128/JB.186.13.4390-4394.2004

Carroni, M., Franke, K. B., Maurer, M., Jäger, J., Hantke, I., Gloge, F., Linder, D., Gremer, S., Turgay, K., Bukau, B., & Mogk, A. (2017). Regulatory coiled-coil domains promote head-to-head assemblies of AAA+ chaperones essential for tunable activity control. *eLife*, *6*, e30120. https://doi.org/10.7554/eLife.30120

Gerth, U., Krüger, E., Derré, I., Msadek, T., & Hecker, M. (1998). Stress induction of the Bacillus subtilis clpP gene encoding a homologue of the proteolytic component of the Clp protease and the involvement of ClpP and ClpX in stress tolerance. *Molecular Microbiology*, *28*(4), 787–802. https://doi.org/10.1046/j.1365-2958.1998.00840.x

Haldenwang, W. (1995). The sigma factors of Bacillus subtilis. *Microbiological Reviews*, *59*(1), 1–30. https://doi.org/10.1128/mr.59.1.1-30.1995

Harwood, C. R. (2007). *Bacillus subtilis* as a Model for Bacterial Systems Biology. In Wiley, *Encyclopedia of Life Sciences* (1st ed.). Wiley. https://doi.org/10.1002/9780470015902.a0002027

Huang, M., & Hull, C. M. (2017). Sporulation: How to survive on planet Earth (and beyond). *Current Genetics*, *63*(5), 831–838. https://doi.org/10.1007/s00294-017-0694-7

Isawumi, A., Ayerakwa, E. A., Abban, M. K., & Mosi, L. (2023). Expression profiles of sporulation genes in multidrug-resistant Bacillus species isolated from intensive care units of Ghanaian hospital. *Experimental Biology and Medicine*, *248*(6), 501–507. https://doi.org/10.1177/15353702231160336

Kain, J., He, G. G., & Losick, R. (2008). Polar Localization and Compartmentalization of ClpP Proteases during Growth and Sporulation in Bacillus subtilis. *Journal of Bacteriology*, *190*(20), 6749–6757. https://doi.org/10.1128/JB.00589-08

Kirstein, J., Dougan, D. A., Gerth, U., Hecker, M., & Turgay, K. (2007). The tyrosine kinase McsB is a regulated adaptor protein for ClpCP. *The EMBO Journal*, *26*(8), 2061–2070. https://doi.org/10.1038/sj.emboj.7601655

Konovalova, A., Søgaard-Andersen, L., & Kroos, L. (2014). Regulated proteolysis in bacterial development. *FEMS Microbiology Reviews*, *38*(3), 493–522. https://doi.org/10.1111/1574-6976.12050

Koopman, N., Remijas, L., Seppen, J., Setlow, P., & Brul, S. (2022). Mechanisms and Applications of Bacterial Sporulation and Germination in the Intestine. *International Journal of Molecular Sciences*, *23*(6), 3405. https://doi.org/10.3390/ijms23063405

Kornberg, A., Spudich, J. A., Nelson, D. L., & Deutscher, M. P. (1968). Origin of Proteins in Sporulation. *Annual Review of Biochemistry*, *37*(1), 51–78. https://doi.org/10.1146/annurev.bi.37.070168.000411

Krüger, E., & Hecker, M. (1998). The First Gene of the *Bacillus subtilis clpC* Operon, *ctsR* , Encodes a Negative Regulator of Its Own Operon and Other Class III Heat Shock Genes. *Journal of Bacteriology*, *180*(24), 6681–6688. https://doi.org/10.1128/JB.180.24.6681-6688.1998

Kuhlmann, N. J., & Chien, P. (2017). Selective adaptor dependent protein degradation in bacteria. *Current Opinion in Microbiology*, *36*, 118–127. https://doi.org/10.1016/j.mib.2017.03.013

Lavey, N. P., Shadid, T., Ballard, J. D., & Duerfeldt, A. S. (2019). *Clostridium difficile* ClpP Homologues are Capable of Uncoupled Activity and Exhibit Different Levels of Susceptibility to Acyldepsipeptide Modulation. *ACS Infectious Diseases*, *5*(1), 79–89. https://doi.org/10.1021/acsinfecdis.8b00199

Li, C., Wen, A., Shen, B., Lu, J., Huang, Y., & Chang, Y. (2011). FastCloning: A highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnology*, *11*(1), 92. https://doi.org/10.1186/1472-6750-11-92

Mahmoud, S. A., & Chien, P. (2018). Regulated proteolysis in bacteria. *Annu Rev Biochem.*, *87*, 677–696. https://doi.org/10.1146/annurev-biochem-062917-012848

Massoni, S. C., Evans, N. J., Hantke, I., Fenton, C., Torpey, J. H., Collins, K. M., Krysztofinska, E. M., Muench, J. H., Thapaliya, A., Martínez-Lumbreras, S., Hart Ferrell, S., Slater, C., Wang, X., Fekade, R., Obwar, S., Yin, S., Vazquez, A., Prior, C. B., Turgay, K., … Camp, A. H. (2025). MdfA is a novel ClpC adaptor protein that functions in the developing Bacillus subtilis spore. *Genes & Development*, *39*, 510–523. http://www.genesdev.org/cgi/doi/10.1101/gad.352498.124

Mauri, M., & Klumpp, S. (2014). A Model for Sigma Factor Competition in Bacterial Cells. *PLoS Computational Biology*, *10*(10), e1003845. https://doi.org/10.1371/journal.pcbi.1003845

Mayangsari, R. (2018). Expression of the ClpC Protease Chaperone Gene During Sporulation in Bacillus subtilis. Mount Holyoke College.

Msadek, T., Dartois, V., Kunst, F., Herbaud, M., Denizot, F., & Rapoport, G. (1998). ClpP of *Bacillus subtilis* is required for competence development, motility, degradative enzyme

synthesis, growth at high temperature and sporulation. *Molecular Microbiology*, *27*(5), 899–914. https://doi.org/10.1046/j.1365-2958.1998.00735.x

Nakjang, S., Ndeh, D. A., Wipat, A., Bolam, D. N., & Hirt, R. P. (2012). A Novel Extracellular Metallopeptidase Domain Shared by Animal Host-Associated Mutualistic and Pathogenic Microbes. *PLoS ONE*, *7*(1), e30287. https://doi.org/10.1371/journal.pone.0030287

Nkomboni, S. (2017). Uncovering the Transcriptional Regulation of clpC during Sporulation in Bacillus Subtilis. Mount Holyoke College.

Nkomboni, S., Mayansari, R., Berg, A., Sawyer-Wheeler, K., Panchal, K., Padrusch, B., Guo, C., Womer, B., Massoni, S. C., & Camp, A. H. A dual SigB/SigF-dependent promoter is required for expression and function of the AAA+ protease chaperone gene clpC in the developing Bacillus subtilis spore. Poster presented at: Bacterial Cell Biology and Development GRC; 2025 June 5-6; Manchester, New Hampshire.

Noori Goodarzi, N., Fereshteh, S., Azizi, O., Rahimi, H., Bolourchi, N., & Badmasti, F. (2022). Subtractive genomic approach toward introduction of novel immunogenic targets against Clostridioides difficile: Thinking out of the box. *Microbial Pathogenesis*, *162*, 105372. https://doi.org/10.1016/j.micpath.2021.105372

Pan, Q., Garsin, D. A., & Losick, R. (2001). Self-Reinforcing Activation of a Cell-Specific Transcription Factor by Proteolysis of an Anti-σ Factor in B. subtilis. *Molecular Cell*, *8*(4), 873–883. https://doi.org/10.1016/S1097-2765(01)00362-8

Pan, Q., & Losick, R. (2003). Unique Degradation Signal for ClpCP in *Bacillus subtilis*. *Journal of Bacteriology*, *185*(17), 5275–5278. https://doi.org/10.1128/JB.185.17.5275-5278.2003

Panchal, K. (2025). Combinatorial regulation of clpC expression during sporulation in Bacillus subtilis. Mount Holyoke College.

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014). Clostridium difficile spore biology: Sporulation, germination, and spore structural proteins. *Trends in Microbiology*, *22*(7), 406–416. https://doi.org/10.1016/j.tim.2014.04.003

Petersohn, A., Bernhardt, J., Gerth, U., Höper, D., Koburger, T., Völker, U., & Hecker, M. (1999). Identification of sigma(B)-dependent genes in Bacillus subtilis using a promoter consensus-directed search and oligonucleotide hybridization. *Journal of Bacteriology*, *181*(18), 5718–5724. https://doi.org/10.1128/jb.181.18.5718-5724.1999

Petersohn, A., Engelmann, S., Setlow, P., & Hecker, M. (1999). The katX gene of Bacillus subtilis is under dual control of σB and σF. *Molecular and General Genetics MGG*, *262*(1), 173–179. https://doi.org/10.1007/s004380051072

Piggot, P. J., & Hilbert, D. W. (2004). Sporulation of Bacillus subtilis. *Current Opinion in Microbiology*, *7*(6), 579–586. https://doi.org/10.1016/j.mib.2004.10.001

Queraltó, C., Ortega, C., Díaz-Yáñez, F., Inostroza, O., Espinoza, G., Álvarez, R., González, R., Parra, F., Paredes-Sabja, D., Acuña, L. G., Calderón, I. L., Fuentes, J. A., & Gil, F. (2023). The chaperone ClpC participates in sporulation, motility, biofilm, and toxin production of Clostridioides difficile. *Journal of Global Antimicrobial Resistance*, *33*, 328–336. https://doi.org/10.1016/j.jgar.2023.05.004

Quigley, F. (2023). Exploring the Role of the Protein-Kinase McsB in the ClpC/MicA Protein Degradation System During Bacillus subtilis Sporulation. Mount Holyoke College.

Riley, E. P., Lopez-Garrido, J., Sugie, J., Liu, R. B., & Pogliano, K. (2021). Metabolic differentiation and intercellular nurturing underpin bacterial endospore formation. *Science Advances*, *7*(4), eabd6385. https://doi.org/10.1126/sciadv.abd6385

Riley, E. P., Lyda, J. A., Reyes-Matte, O., Sugie, J., Kasu, I. R., Enustun, E., Armbruster, E. G., Ravishankar, S., Isaacson, R. L., Camp, A. H., Lopez-Garrido, J., & Pogliano, K. (2025). Developmentally regulated proteolysis by MdfA and ClpCP mediates metabolic differentiation during Bacillus subtilis sporulation. *Genes & Development*, *39*, 524–537.

Riley, E. P., Schwarz, C., Derman, A. I., & Lopez-Garrido, J. (2021). Milestones in Bacillus subtilis sporulation research. *Microbial Cell*, *8*(1), 1–16. https://doi.org/10.15698/mic2021.01.739

Rodriguez Ayala, F., Bartolini, M., & Grau, R. (2020). The Stress-Responsive Alternative Sigma Factor SigB of Bacillus subtilis and Its Relatives: An Old Friend With New Functions. *Frontiers in Microbiology*, *11*, 1761. https://doi.org/10.3389/fmicb.2020.01761

Romero-Rodríguez, A., Ruíz-Villafán, B., Sánchez, S., & Paredes-Sabja, D. (2024). Is there a role for intestinal sporobiota in the antimicrobial resistance crisis? *Microbiological Research*, *288*, 127870. https://doi.org/10.1016/j.micres.2024.127870

Sauer, R. T., & Baker, T. A. (2011). AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. *Annual Review of Biochemistry*, *80*(Volume 80, 2011), 587–612. https://doi.org/10.1146/annurev-biochem-060408-172623

Setlow, P. (2025). Where and why have so many metabolic enzymes gone from developing spores of *Bacillus subtilis* ? *Genes & Development*, *39*(7–8), 426–427. https://doi.org/10.1101/gad.352755.125

Sierro, N., Makita, Y., De Hoon, M., & Nakai, K. (2008). DBTBS: A database of transcriptional regulation in Bacillus subtilis containing upstream intergenic conservation information. *Nucleic Acids Research*, *36*(suppl_1), D93–D96. https://doi.org/10.1093/nar/gkm910

Smits, W. K., Garey, K. W., Riley, T. V., & Johnson, S. (2025). Clostridioides difficile is a bacterial priority pathogen. *Anaerobe*, *93*, 102965. https://doi.org/10.1016/j.anaerobe.2025.102965

Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H., & Kuijper, E. J. (2016). Clostridium difficile infection. *Nature Reviews Disease Primers*, *2*(1), 16020. https://doi.org/10.1038/nrdp.2016.20

Swarge, B., Abhyankar, W., Jonker, M., Hoefsloot, H., Kramer, G., Setlow, P., Brul, S., & De Koning, L. J. (2020). Integrative Analysis of Proteome and Transcriptome Dynamics during Bacillus subtilis Spore Revival. *mSphere*, *5*(4), 10.1128/msphere.00463-20. https://doi.org/10.1128/msphere.00463-20

Tan, I. S., & Ramamurthi, K. S. (2014). Spore formation in Bacillus subtilis. *Environ Microbiol Rep*, *6*(3), 212–225. https://doi.org/10.1111/1758-2229.12130

Thomas, A. K., Peritore-Galve, F. C., Ehni, A. G., Lança, B. B. C., Coggin, J., Brady, E. J., Yoder, S. M., Shrem, R., Rodríguez, R. C., Kroh, H. K., Gibson-Corley, K. N., Kay Washington, M., Olivares-Villagómez, D., Creech, C. B., Nicholson, M. R., Spiller, B. W., & Lacy, D. B. (2026). Mucosal vaccination clears Clostridioides difficile colonization. *Nature*. https://doi.org/10.1038/s41586-026-10138-x

Trentini, D. B., Suskiewicz, M. J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K., & Clausen, T. (2016). Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature*, *539*(7627), 48–53. https://doi.org/10.1038/nature20122