2016

Investigation of a Ubiquitous Spore Surface Protein

Justyna Kordzikowska
Loyola University Chicago

LOYOLA UNIVERSITY CHICAGO

INVESTIGATION OF A UBIQUITOUS SPORE SURFACE PROTEIN

A THESIS SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

MASTER OF SCIENCE

PROGRAM IN MICROBIOLOGY AND IMMUNOLOGY

BY

JUSTYNA KORDZIKOWSKA

MAYWOOD, IL

AUGUST 2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

LITERATURE OVERVIEW

Spore Structure and Environment of Spores

In response to a lack of nutrients or environmental stress, a variety of species belonging to the phylum *Firmicutes* produce endospores (or spores). All known spore-forming species are members of *Bacilli*, *Clostridia*, *Erysipelotrichia*, and *Negativicutes* classes [1]. The best studied species are members of *Bacillaceae.* A spore is a dormant cell that protects the bacterial genome. The basic design of a spore is similar across bacterial species. A spore is made out of concentrically arranged layers [2]. The inner most compartment is the core which houses the bacterial DNA. The core is surrounded by an inner forespore membrane, then a germ cell wall which is enclosed by a peptidoglycan layer, called the cortex. The cortex is encompassed by an outer membrane which is encircled by the coat. The coat is composed of a few layers and the number of layers varies depending on the species. Core, cortex and coat are three essential structures that provide protection [3]. Interestingly, the coat is not always the outer most shell. In fact, some species have an additional layer called the exosporium. This layer is separated from the coat by the interspace. The exosporium is made out of the inner basal layer and the outer layer. The outer layer is made out of nap of hair-like projections [4]. Examples of bacteria that possess the exosporium are: *Bacillus thuringiensis*, *Bacillus megaterium*

and *Bacillus anthracis.* On the other hand, species of bacteria whose spore lack the exosporium are *Bacillus subtilis, Bacillus licheniformis* and *Bacillus clausii.* This striking dissimilarity among bacteria raises the question of why the microorganisms would differ in their outer appearance. It is interesting because species with and without the exosporium can be found in the same environmental niches. For example, in one study, it was shown that species with the exosporium (*Paenibacillus polymyxa, Bacillus megaterium* and *Bacillus cereus*) as well as species without the exosporium (*Bacillus pumilus*) are commonly found in the rhizosphere of wild barley [5]. An even a more complex mixture of both types of spores was shown by Barbosa et al; where spores with the exosporium (*Bacillus megaterium, Paenibacillus alvei, Bacillus cereus, Brevibacillus brevi* and *Brevibacillus laterosporus*) and without the exosporium (*Bacillus pumilus, Bacillus subtilis, Bacillus licheniformis* and *Bacillus amyloliquefaciens*) were rescued from the feces of broilers [6]. Additionally, *Bacillus licheniformis*, *Bacillus pumilus*, *Bacillus circulans*, and *Bacillus subtilis* were found together with *Bacillus cereus* group members in multiple samples of pasteurized milk [7]. The observation that various spores are found in the same environment suggests that all dormant cells have to face the same external challenges. For example, many spore-forming bacteria reside in the soil. In this environment, spores are in contact with, and attach to, the soil components. Moreover, spores found in the same environment are subject to water and humidity changes. Since the coat and the exosporium are the outer most layers, it is intuitive to speculate that both of them have similar structures to perform related functions. This reasoning led to the question of my research: despite the differences in the outer layers among the bacterial species, is there a protein that commonly appears on the outside of a spore?

BclA: Domain Architecture and Function

To answer the question whether there is a universal spore surface marker, we chose a well characterized protein called BclA. This protein is present in *B. anthracis*, *B. cereus*, *B. thuringiensis* and other closely related species (so called *B. cereus* group). BclA is the immunodominant protein in *B. anthracis* [8, 9]. BclA (bacillus collagen-like protein of anthracis) is a glycoprotein that was first discovered on the surface of *B. anthracis* spores [8]. It has been shown that BclA is absent in vegetative cells and only found on spores. BclA is the major component of the hair like nap that protrudes from the basal layer of the exosporium in *B. cereus* group species. Recently, BclA has been also found on the spore surface of *Clostridium difficile*. This finding is surprising because *C. difficile* does not possess the exosporium. Furthermore, *C. difficile* is not related to *B. cereus* group. In fact, *C. difficile* belongs to *Clostridiaceae*. The discovery that BclA is found on the spore surface of species from two different families made it plausible to think BclA might be a common spore protein. BclA contains three major regions: an N- terminal domain that anchors the protein to the basal layer [10] a region of GXX collagen-like repeats (CLR), and a C-terminal domain which is the immunodominant part of the protein (Figure 1). The N- and C-termini are relatively conserved whereas the collagen-like region varies across *B. anthracis* strains [11]. The crystal structure of the C- terminus resembles the C1q domain from the complement, which is a member of the TNF superfamily [12]. It is worth noting the resemblance is at the level of the three dimensional structure, not the amino acid sequence. The function of BclA is still unclear. Sylvestre et al. showed BclA is not required for resistance to lysozyme or proteinase K [8]. Moreover, spores that lacked BclA were similarly resistant to treatments with 100 mM hydrogen peroxide, 0.5

M hydrochloric acid, 0.5 M sodium hydroxide, 10% toluene, and 100% methanol [8]. Also, Bozue et al. showed the lack of BclA did not affect the virulence of the spores in animal models using several challenge methods [13]. However, Oliva et al. showed, the recognition of BclA promotes spore uptake by phagocytes which facilitates transport of the spores to the site where they germinate [14]. This interaction contributed to the mortality of mice as the lack of the receptor that recognized BclA corresponded to the increase survival of the animal. Gu et al. reported BclA is important for the classical complement pathway activation where it recruits C1q to the spore surface. This is followed by spore phagocytosis by macrophages [15]. The possible function of BclA may be linked to its interaction with the environment. It has been noted that spores lacking the protein germinate faster and are less hydrophobic than the wild type spores [16]. A different result was presented by Lequette et al. who looked into the role of BclA in *B. cereus* spores. Loss of BclA made the spores more hydrophobic [17]. It was also shown the spores lacking BclA on their surface had a reduced interaction with stainless steel. The author claims that the adherence properties of *B. cereus* which lacks BclA are lowered and consequently spores get detached easier from the surface. In another study, Chen et al, reported spores lacking BclA on their surface were more adhesive than the wild type spores [18].

Differences between the Coat and the Exosporium

It is important to consider the difference between the coat and the exosporium when determining whether BclA is present in both of these layers. The coat is composed of approximately 70 different proteins in *B. subtilis* [3]. The exosporium is composed of about 20 proteins and glycoproteins as well as lipids and carbohydrates [9, 19 and 20].

Most of the proteins found in the exosporium are specific to that layer except for three proteins whose orthologues are found in the coat of *B. subtilis* –CotB, CotY and ExsY [20, 21]. Because the coat and exosporium are two distinct structures, one would expect BclA would be incorporated differently in spores that have the exosporium. We have some knowledge of how BclA is incorporated into the exosporium of *B. anthracis*. In order for BclA to assemble around the entire spore surface, BxpB (also called ExsFA) protein is required [22, 23, and 24]. A conserved motif in the N-terminal domain of BclA is required as well. Based on BLAST, BxpB is only present in species belonging to the *B. cereus* group. If BclA is present in spores lacking an exosporium, then it will be anchored to the coat in a different way than in species that have an exosporium. Consequently, I expect to see novel N-terminal domain sequences in various species. The support for my expectation comes from the knowledge that BclA is present on the spore surface of *C. difficile* which seem to lack BxpB protein (BLAST). This means there must be an alternative mechanism which anchors BclA to the spore surface. In *B. subtilis*, I expect BclA to interact with one or more of the outer coat protein or crust proteins. There are 24 proteins that are found in the outer coat of *B. subtilis* [25]. It is possible that BclA would interact with the most abundant proteins of the outer coat – CotB, CotG or CotC [21]. Additionally, we need to consider that BclA might interact with CotX/CotY/CotZ or CotW which are the crust proteins [McKenney 26]. If we find BclA in *B. subtilis*, we will consider looking for the protein in closely related species. *B. subtilis*, *B. amyloliquefaciens*, *B. methylotrophicus*, and *B. atrophaeus* belong to the Subtilis clade [27]. I would hypothesize that closely related species have homologous proteins that make up the spore.

The fact BclA is found on the spore surface of *C. difficile* lets us speculate the protein may be found on the spore surface of species belonging to other families harboring spore-forming bacteria besides *Bacillaceae* and *Clostridiaceae*. We know members of *Paenibacillaceae* include spore-forming bacteria. Spore formers are found in *Aneurinibacillus*, *Paenibacillus* and *Brevibacillus* genera. I am curious to explore whether BclA is present in any of the species from *Paenibacillaceae.*

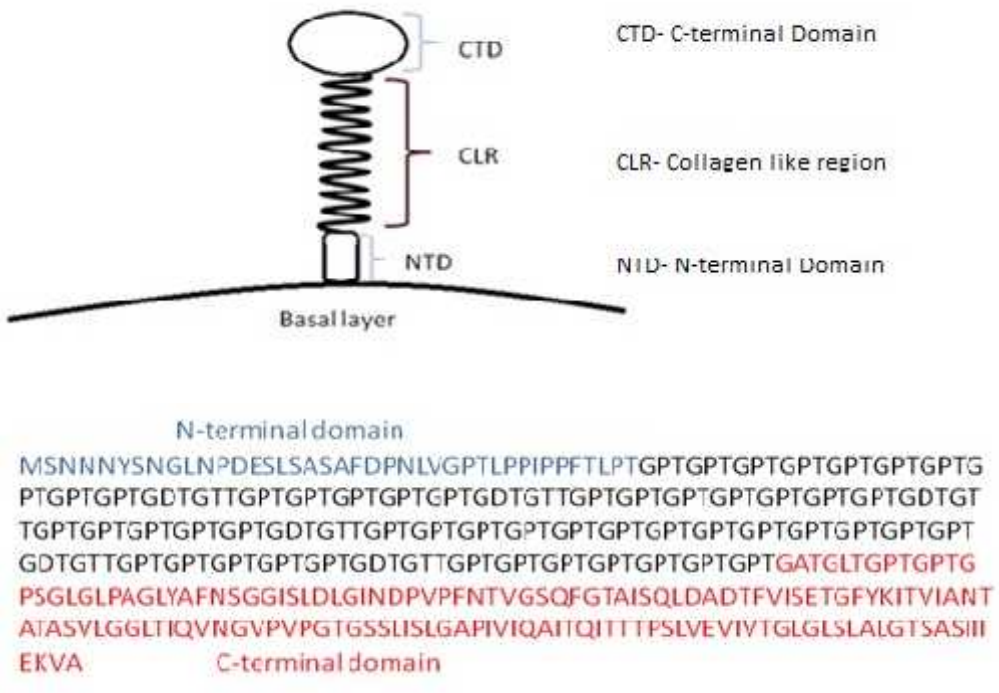

Figure 1. Domain organization of BclA protein.

CHAPTER TWO

METHODS AND MATERIALS

Strains and Media

Bacterial species and strains used in this study are listed in Table 1. Spores were prepared by exhaustion in Difco Sporulation Medium (DSM) [28]. A single colony was picked from a Luria Broth (LB) plate and suspended in 200 µl of DSM. The bacterial suspension was spread onto a DSM plate and incubated at 37 °C for 8 to 9 hours. After the incubation period, the lawn was collected by suspension in 5 ml of DSM. 1ml of the lawn was transferred into the 35 ml of DSM in a 250 ml flask. The flask was shaken at 225 rpm / 37 °C overnight. The next day, 5 µl of the sample was placed onto a glass slide and the sample was checked for the presence of spores under the phase-contrast microscope. Spores were spun down and pellets were washed 3 times with Milli-Q water. Spores were stored in water in 50 ml tubes at 4°C.

Escherichia coli strains were cultured in LB medium. Antibiotics were added when appropriate: 100 µg/ml spectinomycin, 100 µg/ml erythromycin, 100 µg/ml ampicillin, and 15µg/ml chloramphenicol. Plasmid DNA was isolated using the ThermoScientific Gene JET Plasmid Miniprep Kit (Lithuania), and genomic DNA was isolated with the Promega Wizard Genomic DNA Purification Kit (Madison, WI).

Construction of Mutant Strains

Plasmids were introduced into bacterial cells by conjugation [29]. On day 1, the donor strain, the recipient strain and the helper strain were streaked out onto LB +100 µg/ml spectinomycin, BHI, and LB +100 µg/ml ampicillin respectively. The donor strain contained the allelic exchange construct cloned into pRP1028. The recipient strain was either *B. atrophaeus* 1942, *B. subtilis* BSn5, *B. subtilis* PY79 or *B. anthracis*. Plates were incubated at 37°C overnight. On day 2, each strain was scraped off the plate, plated onto a BHI plate, and mixed together. The plate was incubated at the room temperature. On day 3, the mixed strains were streaked out onto a selective plate that contained BHI + 250 µg/ml spectinomycin + 60 units/ml polymixin B. The plates were left at the room temperature for 48 hours.

Bacteria were transformed by electroporation [29]. A single colony was suspended in 3 ml of BHI containing 0.5% glycerol and incubated at 37°C with shaking at 225 rpm overnight. The next day, 0.2 ml of an overnight culture was transferred to 25 ml of BHI containing 0.5% glycerol in a 250 ml flask and incubated at 37°C with shaking at 225 rpm. At OD_{600} of 0.8, cells were harvested by spinning down at 3000 rpm or alternatively by filtering, washed twice with 25 ml of ice-cold electroporation buffer (1 mM HEPES, 10% glycerol, pH 7.0). The cells were resuspended in electroporation buffer to 1/20 of the original volume. 5 µl of plasmid DNA was mixed with 0.2 ml of the cell suspension on ice in a cooled 0.2-cm-gap electroporation cuvette. The cells were exposed to a single pulse at 2.5 kV, 25 uF, and 200 Ohm (time constant of 4-5 msec). After the pulse, the cells were transferred to a sterile tube containing 1 ml of BGGM (BHI with 10% glycerol, 0.4% glucose, and 10 mM $MgCl_2$). Samples were incubated with shaking at 225

rpm for 2-3 hours at 28 °C. The samples were plated on LB plates with 100 µg/ml spectinomycin and incubated at the room temperature for 48 hours.

An alternative electroporation protocol was used based on Xue's work [30]. A single colony was suspended in 5 ml of LB and incubated at 37°C at 225 rpm overnight. 0.6 ml of the overnight culture was transferred to 10 ml of LB + 0.5 M sorbitol in a 250 ml flask and incubated with shaking at 225 rpm. At OD_{600} of 0.9, the cells were harvested by spinning down at 5000 rpm and washed four times with 10 ml of ice-cold electroporation buffer (0.5 M sorbitol, 0.5 M mannitol and 10% glycerol). The cells were resuspended in electroporation buffer to 1/20 of the original volume. 5 µl of plasmid DNA was mixed with 0.2 ml of the cell suspension on ice in a cooled 0.2-cm-gap electroporation cuvette. The cells were exposed to a single pulse at 2.5 kV, 25 uF, and 200 Ohm (time constant of 4-5 msec). After the pulse, the cells were transferred to a sterile tube containing 1 ml of the recovery medium (LB + 0.5 M sorbitol + 0.38 M mannitol). Samples were incubated with aeration for 2-3 h at 28 °C. The samples were plated on LB plates with 100 µg/ml spectinomycin and incubated at the room temperature for 48 hours.

The third electroporation protocol was based on Zhang's study [31]. The adjusted protocol was based on Xue et al. work with some modifications: at OD_{600} of 0.5, 1% threonine and 0.07% Tween 80 were added to the growing cells. Additionally, the field strength was increased to 20,000 KV/cm by using 0.1-cm-gap electroporation cuvette at a single pulse at 2.5 kV, 25 uF, and 200 Ohm.

Bioinformatics

Amino acid sequences of *bclA* genes were obtained from the National Center for Biotechnology Information (NCBI) database.

Molecular and Genetic Techniques

First, *gfp* carrying plasmid was built. The *gfp* was amplified from pUTE-*gfp* using High Fidelity Platinum Taq polymerase (Invitrogen). Amplified PCR product had A-overhangs and was ligated into pGEMT (Promega). pGEMT-*gfp* was digested with KpnI + SacI and ligated into pRP1028. pRP1028-*gfp* was digested with SacI and PmeI and *bclA* fragment with SacI and EcoRV. *bclA* alleles were amplified by Polymerase Chain Reaction (PCR). Primers are listed in Table 2. The PCR products were digested with Eco321 and SacI and cloned into pRP1028 which was digested with SacI and PmeI. The pRP1028 construct was used to transform *E. coli* DH5α. This plasmid was passed through *E. coli* JM110 as well.

Spore Surface Protein Extraction and Western Blot

Chemical extraction of coat proteins was performed. Proteins were extracted from each *bacillus* species listed in Table 1 [32]. 50 ml of spores were pelleted down and resuspended in 10 ml of water. 5 µl of spore suspension was mixed with 495 µl of water and OD_{600} was measured and recorded. The formula 0.037/ OD_{600} was used to estimate the amount of spore suspension needed for each well. The appropriate amount of spore suspension was pelleted for 5 min/6000 rpm and the supernatant was removed. 13 µl of Laemmli buffer (0.63 ml 1M Tris pH 6.8, 1 ml 100% glycerol, 2ml 10% SDS and 6.37 ml MQ water) and 1.5 µl of 1M DTT were added to the pellets. The samples were vigorously shaken for 1 minute and spun down for 2 seconds. The samples were boiled for 5 minutes at 100 °C. The samples were shaken again for 45 seconds and boiled at 100°C for an additional 5 minutes. The samples were vigorously shaken for 30 seconds and spun down for 5 minutes at 13000 rpm. The supernatants were collected and the

protein concentration was measured with Nanodrop. Proteins were resolved on 15% SDS-PAGE (10% 29:1 acrylamide: N, N'-methylene-bis-acrylamide, 375 mM Tris pH 8.6, 0.1% SDS), at100 µg/100 µl concentration, and transferred to polyvinylidene difluoride membranes or nitrocellulose membranes [33]. The membranes were incubated with 2% bovine serum albumin (BSA), washed three times with Tris-buffered saline with Tween 20 (1x TBST: 1.21 g Tris, 8.76 g NaCl, 0.05% Tween 20 in 1L of water) then incubated with monoclonal anti BclA antibody (BA-MAB 5; Critical Reagents Program, Department of Defense) or polyclonal antibodies (anti-BclA antiserum, non-immune serum or serum coming from a mouse injected with PBS, Livermore, CA). Goat anti mouse was the secondary antibody (Sigma).

Immunofluorescence Microscopy

10 µl of spore suspension was placed into each well of a multiwell slide. The slides were pretreated with 0.01% (wt/vol) poly-L-lysine, washed twice with water and air-dried. 10 µl of phosphate-buffered saline (PBS) was then placed into each well and replaced with 2% (wt/vol) BSA in phosphate-buffered saline prior to the addition of primary antibody. Monoclonal anti BclA antibody was used at a 1:5000 dilution. Polyclonal antibodies were used at a 1:50 dilution. Polyclonal anti-BclA antibody raised in rabbits was used at a 1:100 dilution (BEI Resources). Secondary antibody was used at a 1:300 dilution. Alexa Fluor 488 goat anti mouse IgG or goat anti rabbit IgG were used as the secondary antibody (Molecular probes, Life Technologies ^{TM}).

Spore Hydrophobicity Measurement

BATH (Bacterial Adherence to Hydrocarbons) assay was performed to measure the hydrophobicity of spores that either have or lack BclA on their surface [16, 34]. Spore suspensions in sterile water were prepared at OD_{440} 0.4-0.6. The prepared samples were mixed with various (25, 125 or 250 µl) amounts of n-hexadecane and mixed by vortexing for 1 minute. The samples were left for 30 minutes to allow the aqueous and nonaqueous phases to separate. The aqueous layer was carefully removed and the OD_{440} was measured again. The hydrophobicity was calculated by using the equation: 100% - (OD before / OD after) x 100%. To study whether or not heat has an impact on the hydrophobicity of the spores, the samples were treated with heat before mixing with the n-hexadecane. Spores adjusted to OD_{440} 0.4-0.6 were treated with heat (37, 65 or 100 °C) for 10 minutes and afterward left in the fridge overnight. As a control, a sample was left at room temperature (around 25 °C) for the duration of the heat treatment. The next day, the spore suspensions were mixed with 250 µl of n-hexadecane and the hydrophobicity as measured.

Clumping Assay

Spores suspended in distilled water were placed on a vortex for 1 minute. OD_{580} was adjusted to 0.5-0.6 for each 500 µl samples. The samples were placed in the spectrophotometer and the OD_{580} was measured every 10 minutes for 90 minutes [35]. Clumping assay was performed in duplicates and the average was recorded as the percentage of spores that did not clump at any given time.

SyntTax

SyntTax was used to study gene preservation in bclA locus in various species.

Electron Microscopy

A pellet from 1 ml of spore stock was prepared. The pellet was mixed 1 ml of a 2.5% glutaraldehyde, 0.1 M sodium cacodylate solution and 0.1% of ruthenium red. The sample was incubated for 1 hour at 37 °C. Next, the sample was spun down and washed in 1 ml of PBS. The pellet was mixed 2% osmium tetroxide, 0.1 M sodium cacodylate solution and 0.1% ruthenium red. The pellet was incubated for 3 hours at room temperature. Next, the pellet was washed twice with 1ml of PBS and resuspended in 100 µl of water. 300 µl of 3% melted agarose was mixed with the spore pellet and transferred onto an agarose cushion. The sample was spun down for 30 seconds. After the sample solidified, the pellet was cut with a razor blade into small pieces and placed into a scintillation vial. The sample was subjected to dehydration steps in 4 ml of 30%, 50%, 70% and 100% ethanol with rocking for 1 hour in-between. Next, the samples were mixed with 1:1 mixture of 100% ethanol and resin, followed by 100% resin. The sample in resin was left rocking overnight at room temperature. The next day, the samples were fished out and placed into a resin mold. Fresh resin was poured onto the sample and the sample was left baking overnight.

Species/Strains	**Genotype or description**	**Reference or source**
B. subtilis BSn5	Wild type	BGSC
B. amyloliquefaciens NRRL BD-553	Wild type	USDA
B. amyloliquefaciens BGSC 10A1	Wild type	BGSC
B. amyloliquefaciens NRRL BD-599	Wild type	USDA
B. atrophaeus SB512	Wild type	USDA
B. *atrophaeus* 1942	Wild type	BGSC
Br. laterosporus ATCC9141	Wild type	BGSC
B. methylotrophicus FZB42	Wild type	BGSC
P. chitinolyticus NBRC 15660	Wild type	BGSC
B. anthracis 34F2	Wild type	Laboratory collection
B. anthracis 34F2	pXO1+ pXO2- *bclA*::kan	Laboratory collection
B. subtilis PY79	Lab strain	Laboratory collection
E. coli RG7	GM1684 (*dam*-)	Laboratory collection
E. coli DH5α	Cloning host	Laboratory collection
E. coli C2925H	Cloning host, *dam-/ dcm*-	NEB
E. coli SS1827	Helper strain	Stibitz and Carbonetti 1994

Table 1. Species and strains used in BclA study

Plasmids	Description	Reference
pGEM-T	Cloning vector	Promega
pUTE29-gfp	Plasmid carried green fluorescence protein	Laboratory collection
pRP1028	Temperature sensitive vector in Gram positive bacteria, $Spec^R$	Plaut & Stibitz 2015
pBKJ236	Temperature sensitive vector in Gram positive bacteria, Erm^R	Plaut & Stibitz 2006
pEO-3	Shuttle vector, Erm^R	Mendelson & Friedlander 2004
pIMAY	Temperature sensitive vector in Gram positive bacteria, Cam^R	Monk & Foster 2012

Table 2. Plasmids used in BclA study

Species	Gene	Sequence (5' -3')	Primer
B. subtilis	03520	ttt ttt GAG CTC CTT AAC GCA TGT GGA GGT AGT AGT AGA	FW
B. subtilis	03520	ttt ttt GAT ATC GAG ATA ATA ACG TCC TGC CAC TGG	REV
B. subtilis	20885	ttt ttt GAG CTC GAT GTG ATT GTA AAT GGA GGT	FW
B. subtilis	20885	ttt ttt GAT ATC TCC CCC AGC AGA CTC TAT TAA	REV
B. atrophaeus	00385	ttt ttt GAG CTC TGC TGC GTA AGG GGA GTA	FW
B. atrophaeus	00385	ttt ttt GAT ATC AAT GAC ATC AGC CTC TAT AGC TAC CGT	REV
B. atrophaeus	01385	ttt ttt GAG CTC AAT CTC ATT GTA AAC GGA GGG	FW
B. atrophaeus	01385	ttt ttt GAT ATC GAT TCC ATT CAC AAA CTC AAC	REV
B. atrophaeus	04295	ttt ttt GAG CTC AAT TTA CCT AAT ATT ACA CCG GTC	FW
B. atrophaeus	04295	ttt ttt GAT ATC GGT ATA ATC AGC AGA AGC GTC	REV
GFP-rev		ttt ttt GGT ACC TTA TTT GTA TAG TTC ATC CAT GCC	REV
2GFP-FW-nostartcodon		AA GAGCTC AAA GTT TAA ACT CGG AGG CGG TGG GGG AGG GAG TAA AGG AGA AGA ACTT TTC	FW

Table 3. Oligonucleotides used in BclA study

CHAPTER THREE

EXPERIMENTAL RESULTS

Identification of Spore-Forming Bacteria Species with *bclA* Homologues

Identification of Genes Homologous to *bclA* So far, BclA was only found on the spore surface of *B. anthracis*, species closely related to it and *C. difficile*. It is possible the protein is actually present in other species as well. If BclA is a protein in various organisms, I should be able to find genes that are homologous to *bclA* in the genomes of those organisms. To answer the question whether there are *bclA*-like genes in multiple spore-forming bacteria, I utilized BLAST. I used a *bclA* sequence from *B. anthracis* as my query. In a preliminary search, I did not look for similarity to the entire amino acid sequence of BclA. As mentioned previously, the exosporium and the coat are two different structures and their compositions are not analogous. The way BclA anchors to the exosporium of *B. anthracis* is most likely not conserved across species with or without exosporia. I expect that the NTD will differ among these two classes of species (and perhaps within classes as well) and it will contain a sequence that targets this protein to the coat or the exosporium. For this reason I decided to split my search for BclA orthologues into two parts. First, I looked for proteins that have similar NTD. Second, I looked for collagen-like proteins with CTD similar to BclA. Once I established which organisms have sequences homologous to BclA, I focused on spore-forming

bacteria from Bacilli. Multiple species from *Bacillaceae* and *Paenibacillaceae* showed *bclA* homology. Within those species, I found multiple sequences homologous to *bclA*. After I created a list of *bclA*-like amino acid sequences, I looked for significant similarity between entire sequences of *B. anthracis bclA* and *bclA*-like genes, similarities in NTD only and CTD only. There are three domains in the BclA protein: the N-terminal domain, collagen-like region and C-terminal domain. In my analysis, I want to look for the sequence similarity in those domains between *B. anthracis bclA* and *bclA*-like sequences in diverse genomes. Identifying similarities in N-terminal domain sequences could help recognize a conserved region responsible for BclA anchoring to the surface of a spore. Comparing the collagen-like region will help me determine if the GXX triples are the same across the species or if there is a variation in the composition. I will be able to establish, in particular, if the $(GPT)_5$GDTGTT region is found across the species or if it is *B. anthracis*-specific. Lastly, the C-terminal domain is especially immunogenic. I will be able to see whether this domain is preserved and to what extent.

The BLAST analysis revealed many more species with *bclA*-like genes than previously noted. Identity to the sequence was not only found in genomes of all members of *B. cereus* group (as already noted), but also in many other species belonging to the *Bacillaceae* family (Table 4). In addition, members of the families *Paenibacillaceae, Streptococcaceae* and *Clostridiaceae* also showed some identity with *B. anthracis bclA*. The *Streptococcaceae* is not known to harbor spore-forming species. The majority of species with *bclA* homologues came from spore-forming bacteria. The highest sequence similarity was found in the collagen-like region. Collagen-like regions were not identical to the one found in *B. anthracis*. The signature region $(GPT)_5$GDTGTT was only found

in *B. anthracis*. The CLR in most of the species had GXT repeats. The C-terminal domain is not conserved across the species. Species belonging to the *B. cereus* group had high similarity in C-terminal domains. The remaining species had either some similarity to CTD or none. N-terminal domain sequence is only conserved among *B. cereus* group species (Table 5). This was expected as this protein region anchors BclA to the basal layer, which is likely to be similar among member of *B. cereus* group. Although Table 4 and Table 5 list only the results on the species level, for many of the organisms, I was able to identify multiple strains with *bclA* homologues. I narrowed down my research to the *Bacillus*, *Paenibacillus*, *Brevibacillus* and *Lysinibacillus* genera and analyzed the genomes of multiple strains that showed identity to *bclA* (Table 6). I found that many of the strains have multiple collagen-like proteins. As previously mentioned, the similarity was mostly coming from the CLR. There were a few collagen-like proteins that had high E values when compared to BclA from *B. anthracis*. That might suggest that, those proteins are not BclA homologues. I was not able to show that any of the N-termini present in *bclA*-like sequences was significantly similar to N-terminus of BclA in *B. anthracis*. There was no N-terminus similar in length to the N-terminus in BclA *of B. anthracis*. Moreover, the length of the N-termini varied greatly among the sequences ranging from 2 to 240 amino acids. In some cases, it was difficult to establish how big the N-terminus is. In those cases, there would be a few GXT triples present followed by non-collagen region, which in turn, was followed by GXT triplets. Only 10 out of 94 sequences have significant similarity in C-terminal domain. This could suggest that C-termini are divergent across the species.

Species	Query covered (%)	E value	Identity (%)	Similarity to CTD
Bacillus anthracis	100	0	100	yes
Bacillus thuringiensis	100	4e -147	82	yes
Bacillus cytotoxicus	89	3e -110	79	yes
Bacillus cereus	97	9e -109	77	yes
Streptococcus pneumoniae	97	9e -92	91	yes
Bacillus weihenstephanensis	96	1e -86	74	yes
Bacillus mycoides	95	1e -86	84	yes
Syntrophobotulus glycolicus	99	7e -75	60	yes
Clostridium aerotolerant	99	3e -67	56	yes
Clostridium sordellii	99	4e -63	54	yes
Clostridium aceticum	99	4e -62	47	yes
Brevibacillus laterosporus	100	2e -61	48	yes
Clostridum diolis	99	7e -61	55	yes
Clostridium celerecresens	100	3e -60	54	yes
Bacillus pumilus	100	3e -60	51	yes
Bacillus bombysepticus	96	2e -58	55	yes
Clostridium argentinensis	99	1e -57	47	yes
Clostridium difficle	77	2e -57	60	yes
Paenibacillus chitinolyticus	99	3e -57	50	yes
Kangiella koreensis	60	4e -57	65	no
Clostridium beijerinckii	61	7e -57	76	no
Bacillus endophyticus	60	5e -56	75	yes
Clostridium sacchardyticum	99	9e -56	50	yes
Peptoclostridium difficle	77	1e -55	62	yes
Clostridium autoethanogenum	88	1e -55	77	yes
Clostridium methoxybenzovorans	99	2e -55	51	yes
Blautia producta	97	3e -55	53	yes
Bacillus invictae	97	5e -55	77	yes
Clostridium ljungdahlii	90	5e -55	77	yes
Bacillus altitudinis	100	6e -55	49	yes
Haemophilus parasuis	90	1e -54	52	yes
Bacillus licheniformis	98	1e -53	79	yes
Bacillus safensis	100	9e -53	51	yes
Kangiella aquimarine	70	1e -52	62	no
Bacillus amyloliquefaciens	61	1e -52	73	no
Parachlamydiaceae bacterium	90	3e -52	55	yes

Paenibacillus pinihumi	61	3e -52	55	yes
Waddlia chondrophila	94	7e -52	49	yes
Acinetobacter quillouiae	61	9e -52	67	no
Erysipelatoclostridium ramosum	100	1e -51	72	yes
Bacillus aerophlius	64	1e -51	67	no
Alkaliphilus metalliredigens	93	2e -51	47	yes
Bacillus subtilis	94	2e -51	53	yes
Desulfotomatulum quttoideum	94	7e -51	79	yes
Desulfitobacterium hafniense	59	9e -51	65	no
Escherichia coli	62	1e -49	56	no
Fictibacillus gelatini	99	2e -49	73	yes
Hungatella hathewayi	98	5e -49	49	yes
Bacillus licheniformis	62	1e -48	71	yes
Bacillus methylotrophicus	61	2e -48	73	no
Bacillus atrophaeus	61	3e -48	70	no
Paenibacillus mucilaginosus	62	4e -48	69	no
Gottschalkia acidurici	100	7e -48	43	yes
Lysinibacillus varians	96	2e -47	46	yes
Mesorhizobium loti	60	3e -47	68	no
Clostridium tyrobutyricum	99	4e -46	58	yes
Parachlamydia acanthamoebae	60	7e -46	57	no
Bacillus gaemokensis	60	9e -46	78	no
Bacillus amyloliquefaciens subsp. Plantarum	62	2e -45	65	yes
Bacillus mojavensis	99	4e -45	67	yes
Pandoravirus salinus	62	4e -45	48	no
Streptococcus pyogenes	60	5e -45	50	no
Clostridium scatologenes	64	6e -45	62	no
Paenibacillus borealis	57	4e -44	68	no
Brevibacacillus brevis	62	2e -43	61	no
Clostridium indolis	96	5e -43	73	yes
Paenibacillus polymyxa	60	2e -42	70	no
Streptococcus equi	61	2e -42	50	no
Streptosporangium roseum	80	3e -42	58	yes
Robinsoniella peoriensis	61	1e -41	59	no
Aneurinibacillus migulanus	62	2e -41	63	no
Clostridium baratii	92	3e -41	47	yes
Paenibacillus assamensis	60	6e -41	71	no

Mycobacterium phage Tiffany	71	1e -40	47	yes
Cellulophaga baltica	62	3e -40	50	yes
Muricauda ruestringensis	75	5e -40	45	yes
Pandoravirus inopitanum	61	5e -40	46	no
Paenibacillus massiliensis	93	9e -39	56	yes
Kangiella geojedonensis	62	1e -38	48	no
Pithovirus sibericum	59	2e -38	64	no
Alkaliphilus oremlandii	49	2e -38	70	no
Dyadobacter crusticola	61	7e -35	47	no
Methylobacterium aquaticum	61	1e -34	51	no
Bacillus marisflavis	62	1e -34	51	no
Streptacidiphilus albus	61	2e -34	56	no
Burkholderia phymatum	60	4e -34	54	no
Anaerotruncus colihominis	60	1e -33	44	no
Legionella pneumophila	69	1e -33	44	no
Bacillus clausii	62	2e -29	67	no
Paenibacillus alvei	60	8e -18	62	no

Table 4.Results of BLAST search where CLR and CTD of BclA were used together as the query.

Species	Query covered (%)	E value	Identity (%)
Bacillus anthracis	100%	8e -19	100%
Bacillus cereus	100%	3e -18	100%
Bacillus thuringiensis	100%	4e -18	100%
Bacillus weihenstephanensis	100%	2e -17	100%
Bacillus mycoides	95%	1e -14	95%
Bacillus cytotoxicus	97%	2e -06	67%

Table 5.Results of BLAST search where NTD of BclA was used as the query

			Query: BclA from *B. anthracis* Accession number: CAD56869.1			Query: C-terminus of BclA from B. anthracis			Query: N-terminus of BclA from *B. anthracis*			
species/ strain name	Length of BclA-like protein	Accession number	Query covered**	E value	Identity ***	Query cover **	E value	Identity ***	Query covered**	E value	Identity ***	length of N-terminus
Bacillus licheniformis ATCC 14580	955	WP_011197647	48%	2.00E-27	53%	96%	1.00E-13	45%	NSS*			162aa
	1259	WP_011197646	88%	2.00E-54	65%	3%	3.9	60%	NSS*			163aa
Bacillus licheniformis 5-2-D	420	WP_016886284	46%	6.00E-07	80%	9%	4.6	42%	NSS*			160aa
	1054	WP_016886285	90%	8.00E-63	78%	19%	2.7	80%	NSS*			156aa
Bacillus licheniformis 9945A	534	WP_020450548	68%	4.00E-38	80%	21%	0.35	38%	NSS*			160aa
	1874	WP_051143316	85%	9.00E-62	79%	3%	4.4	60%	NSS*			154aa
Bacillus amyloliquefaciens DSM 7	2083	CBI41846	91%	3.00E-53	65%	35%	0.088	27%	47%	0.2	42%	152aa
	621	CBI41847	71%	6.00E-26	52%	31%	1.9	31%	57%	0.092	26%	70aa
	622	CBI42530	46%	6.00E-20	74%	25%	0.043	32%	50%	0.61	40%	101aa
Bacillus amyloliquefaciens LL3	2200	AEB62317	91%	2.00E-52	65%	47%	0.21	42%	47%	0.21	42%	152aa
	586	AEB62960	51%	2.00E-07	73%	36%	0.045	32%	50%	0.54	40%	101aa
Bacillus amyloliquefaciens KHG19	662	WP_042634957	82%	1.00E-54	48%	12%	1.6	25%	NSS*			158 aa
	353	WP_052484216	9%	0.22	36%	40%	0.042	36%	NSS*			160aa
	948	WP_042634958	65%	9.00E-50	54%	31%	7.3	26%	NSS*			72aa
	371	WP_052484221	26%	2.00E-08	36%	12%	1.2	60%	NSS*			2aa
	365	WP_052484219	54%	4.00E-08	44%	28%	0.024	35%	70%	0.24	40%	101 aa
Bacillus atropheus 1942	1335	ADP31235	63%	6.00E-53	69%	NSS			10%	4	100%	158aa
	511	ADP31037	NSS*			30%	0.42	47%	NSS*			25aa
	513	ADP31813	46%	6.00E-08	51%	10%	4.6	43%	75%	0.16	58%	101aa
Bacillus atrophaeus NRS 1221a	513	AJF84609	46%	6.00E-08	51%	10%	4.6	43%	75%	0.16	58%	101aa
	266	AJF83915	NSS*			5%	3.4	57%	NSS*			11aa
	1692	AJF84093	65%	3.00E-55	69%	NSS*			10%	4.8	100%	158aa
Bacillus subtilis strain T30	488	WP_045590038	18%	3.00E-09	62%	5%	3.4	57%	NSS*			42aa
	1086	WP_052673360	76%	3.00E-46	52%	NSS			NSS*			154aa
Bacillus subtilis BSn5	650	ADV92637	91%	3.00E-60	54%	39%	0.017	28%	NSS*			7aa
	240	ADV93334	41%	3.00E-09	72%	18%	0.41	38%	NSS*			13aa
	238	ADV96781	57%	2.00E-04	70%	NO C-TERMINUS			50%	1.1	35%	154aa
Bacillus subtilis strain Bs-916	428	WP_038462142	62%	4.00E-07	73%	40%	0.042	36%	NSS*			160aa
	581	WP_038462140	82%	4.00E-58	51%	12%	1.6	25%	NSS*			158aa

	407	WP_007407226	25%	5.00E-06	35%	8%	0.83	35%	70%	0.26	40%	101aa
	533	WP_038462387	90%	7.00E-11	43%	12%	1.2	60%	NSS*			20aa
Bacillus subtilis SG6	1541	WP_038429006	62%	2.00E-54	67%	39%	0.014	28%	NSS*			154aa
Bacillus subtilis ATCC 13952	2155	AIW35839	91%	5.00E-49	68%	NSS			47%	0.2	42%	152aa
	627	AIW32755	52%	9.00E-23	59%	31%	1.7	31%	57%	0.088	26%	72aa
	622	AIW33350	46%	6.00E-20	74%	NSS*			50%	0.61	40%	101aa
Bacillus methylotrophicus str. FZB42	459	WP_012117058	64%	1.00E-37	49%	40%	4.4	29%	NSS*			72aa
	665	WP_012117056	82%	2.00E-42	51%	12%	1.7	25%	NSS*			158aa
	365	WP_012117390	56%	1.00E-12	45%	9%	1.1	38%	70%	0.28	40%	101aa
	416	WP_012117057	9%	5.1	57%	NSS*			NSS*			160aa
Bacillus methylotrophicus NAU-B3	633	WP_022553453	61%	6.00E-14	75%	28%	0.027	35%	50%	0.59	40%	101aa
	687	WP_022552743	70%	9.00E-55	59%	40%	4.4	29%	NSS*			72aa
Bacillus methylotrophicus JS25R	687	WP_022552743	70%	9.00E-55	59%	40%	4.4	29%	NSS*			72aa
	709	WP_052110583	64%	7.00E-57	72%	28%	0.024	35%	42%	1.6	41%	21aa or 244
Bacillus pumilus ATCC 7061	342	WP_034620927	68%	2.00E-51	69%	N/A			N/A			N/A
	903	WP_003213888	57%	1.00E-56	69%	14%	2.8	37%	30%	0.53	58%	176aa
	917	WP_003211344	92%	1.00E-58	55%	86%	0.045	32%	32%	0.52	54%	50 aa
Bacillus pumilus SAFR-032	1865	ABV62475	82%	4.00E-50	65%	45%	6.00E-04	34%	52%	0.2	43%	140 aa
	345	ABV62341	5%	1	42%	37%	0.091	32%	NSS*			9aa
Bacillus pumilus strain LK21	420	KML12809	89%	9.00E-68	50%	59%	3.00E-05	35%	N/A			N/A
	468	KML10775	56%	4.00E-54	66%	N/A			65%	0.014	43%	205AA
	179	KML10797	92%	6.00E-14	64%	40%	0.001	37%	NSS*			2aa
Bacillus safensis RIT372	348	KIZ49468	57%	3.00E-43	72%	N/A			N/A			N/A
	489	KIZ54948	90%	3.00E-59	50%	59%	2.00E-05	35%	NSS*			N/A
Bacillus safensis strain JPL_MERTA8	1876	WP_046311377	97%	4.00E-56	52%	45%	0.001	34%	52%	0.26	43%	204aa
Bacillus invictae DSM 26896	292	KJF45741	70%	3.00E-53	65%	N/A			N/A			N/A
	304	KJF46841	73%	3.00E-18	74%	N/A			80%	0.005	43%	205aa
	378	KJF45766	87%	1.00E-62	52%	59%	2.00E-05	35%	N/A			N/A
	740	KJF47812	59%	4.00E-62	67%	14%	3.3	42%	NSS*			5aa
	192	KJF46970	73%	8.00E-07	60%	28%	4.00E-04	39%	N/A			N/A
Bacillus altitudinis RIT380	199	KLV13925	69%	3.00E-54	76%	N/A			N/A			N/A
	491	KLV21725	92%	2.00E-62	49%	30%	2.00E-05	41%	N/A			N/A
	224	KLV13929	71%	5.00E-56	73%	N/A			N/A			N/A
	492	KLV13937	82%	1.00E-55	67%	N/A			N/A			N/A
Bacillus xiamenensis strain HYC-10	171	EKF33836	52%	2.00E-09	58%	24%	0.65	33%	N/A			3 aa

Bacillus *endophyticus* 2102	267	WP_019395434	88%	2.00E-32	50%	91%	2.00E-07	29%	NSS*			15aa
	194	WP_019393702	69%	1.00E-17	46%	15%	0.43	52%	12%	0.53	80%	N/A
Paenibacillus daejeonensis DSM 15419	1659	WP_020619414	69%	4.00E-23	57%	10%	0.14	36%	NSS*			101aa
	696	WP_020620726	71%	2.00E-44	57%	65%	5.00E-04	34%	NSS*			101aa
	700	WP_020620725	70%	1.00E-40	54%	56%	0.48	28%	NSS*			101aa
	693	WP_020620723	70%	4.00E-43	44%	NSS*			NSS*			101 or 221 aa
	693	WP_020620724	85%	4.00E-38	40%	76%	0.007	32%	57%	2.2	57%	101 or 225 aa
	700	WP_020620722	64%	1.00E-36	42%	53%	3.00E-04	33%	NSS*			101 or 221aa
	693	WP_020620727	96%	4.00E-40	40%	37%	0.007	27%	NSS*			101 or 222aa
	553	WP_020620713	56%	4.00E-31	50%	29%	0.006	38%	NSS*			240
Paenibacillus borealis DSM 13188	501	WP_052429824	20%	0.002	68%	15%	3.9	45%	NSS*			5aa
	514	WP_052429797	58%	1.00E-52	68%	25%	1.8	35%	NSS*			24aa
	774	WP_052429702	61%	6.00E-31	57%	59%	0.097	30%	NSS*			341aa
Paenibacillus chitinolyticus NBRC 15660 DNA	210	WP_053228787.1	83%	1.00E-33	45%	93%	5.00E-17	39%	NSS*			2aa
Paenibacillus assamensis DSM 18201	338	WP_028595677	77%	2.00E-05	40%	62%	4.00E-08	33%	25%	3.3	50%	101aa
	395	WP_051217543	75%	1.00E-40	66%	42%	0.34	28%	NSS*			88aa
	770	WP_036605366	95%	4.00E-49	71%	22%	0.55	32%	NSS*			101aa
	410	WP_028595663	88%	2.00E-34	46%	31%	8.00E-06	54%	20%	0.13	63%	104aa
	405	WP_051217544	44%	6.00E-17	66%	43%	0.93	34%	NSS*			95aa
Paenibacillus polymyxa WLY78	376	WP_029518313	84%	1.00E-46	67%	N/A			32%	1.5	56%	162 aa
Brevibacillus laterosporus LMG 15441	372	AIG25670	82%	7.00E-26	78%	26%	0.12	53%	NSS*			165aa
	620	AIG26331	92%	2.00E-69	49%	62%	6.00E-06	28%	NSS*			158aa
	467	AIG24974	11%	3.00E-04	43%	31%	5.00E-06	44%	NSS*			54aa
Brevibacillus brevis NBRC 100599 DNA	312	WP_041749728	41%	0.22	69%	33%	0.045	25%	55%	0.92	36%	102aa
	1003	WP_012684061	59%	1.00E-51	61%	12%	0.38	53%	NSS*			31aa
Brevibacillus reuszeri	1876	KNB69934	84%	2.00E-29	50%	26%	0.029	37%	NSS*			100aa
Lysinibacillus varians GY32	407	WP_038510599	88%	3.00E-55	46%	43%	2	27%	N/A			1aa
Lysinibacillus fusiformis RB-21	763	WP_053833723	62%	2.00E-46	55%	15%	3.2	40%	NSS*			45aa
Lysinibacillus boronitolerans JCM 21713	521	WP_052125391	68%	3.00E-43	46%	15%	3.1	40%	95%	2.1	56%	44aa
Lysinibacillus sphaericus 1987	422	WP_051998216	54%	2.00E-33	56%	15%	2.8	40%	N/A			N/A

Table 6. Results of BLAST search where 382 amino acid sequence of BclA was used as the query. NSS*- no significant similarity. **Query covered- percent of the *BclA* sequence that overlaps the subject sequence. *** Identity- percent similarity between the *BclA* sequence and subject sequences over the length of overlapped region. N/A – not applicable meaning no NTD present. Light gray color-species whose *bclA* loci were established by SyntTax

Analysis of *bclA* Loci in Chosen Species The *bclA* locus is conserved in the genomes of *B. cereus* group members. The order in which glycosyltransferases and methyltransferases surround *bclA* is preserved (Figure 2). I asked if the *bclA* locus is conserved in other spore-forming species. Specifically, I wanted to know if the *bclA* locus was universal to all species or if it was distinct and only preserved in closely related species (like *B. subtilis* group). I also wanted to know if glycosyltransferases and methyltransferases genes surrounded *bclA*. To look for conservation in the *bclA* locus across species and presence of glycosyltransferases and methyltransferases, I used SyntTax. SyntTax is a web service that enables studying the conservation of gene order in chosen organisms. Because SyntTax only operates on fully sequenced genomes, I was not able to investigate *bclA* loci in all of the species listed in Table 6. Therefore, I focused my analysis on a group of species: *B. subtilis* BSn5, *B. atrophaeus* 1942, *B. atrophaeus* NRS 1221a, *B. methylotrophicus* FZB42, *B. licheniformis* ATCC 14580 and *B. pumilus* SAFR-032.

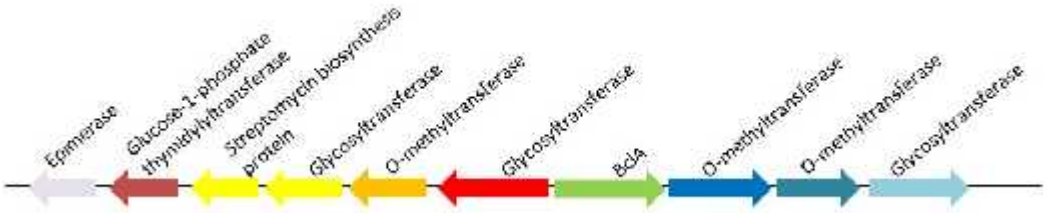

Figure 2: A cartoon depicting *bclA* locus in *B. cereus* group species.

B. subtilis BSn5 genome possesses three *bclA*-like genes in two loci (Figure 3 and Figure 4). Two interesting observations can be made about these loci. First, both of the loci are surrounded by phage-like genes suggesting an integration event. Second, these loci are absent from common lab strains like *B. subtilis* PY79 or *B. subtilis* 168 (Figure 5 and Figure 6). Additionally, locus 1, which has two *bclA*-like genes, is also present in *B. subtilis subsp. subtilis* str. OH_131_1, *B. subtilis* strain RO_NN_1 and B. *subtilis spizizenii* strain W23 (Figure 7). My analysis shows the gene order in *bclA* locus of *B. subtilis subsp. subtilis* str. OH_131_1 is almost identical to that of *B. subtilis* BSn5. There are two *bclA*-like genes in both strains surrounded by glycosyltransferases and methyltransferases. Moreover, both loci contain phage-like genes downstream of *bclA* genes. However, phage-like gene upstream of *bclA* is only conserved in *B. subtilis* Bsn5 whereas *B. subtilis subsp. subtilis* str. OH_131_1 has a noncoding region. Interestingly, *bclA* is missing in *B. subtilis* RO_NN_1. Only methyltransferases and glycosyltransferases are present. It is difficult to conclude whether this locus is related to BSn5 *bclA* locus as there are only four genes that seem to be preserved. *B. subtilis spizizenii* strain W23 has one *bclA*-like gene surrounded by methyltransferases and glycosyltransferases as well. Locus 2 bearing *bclA* in *B. subtilis* BSn5 is not present in

any other *B. subtilis* strain (Figure 8). Intriguingly, a similar locus was found in *B. pumilus* SAFR-032. It is important to note, however, that only four genes seem to be conserved. No methyltransferases or glycosyltransferases are present in either *B. subtilis* BSn5 or *B. pumilus* SAFR-032 loci. The second *bclA* locus in *B. pumilus* SAFR-032 is only found in *B. pumilus* strains (Figure 9). No glycosyltransferases or methyltransferases are found in this locus. *B. atrophaeus* 1942 has three *bclA* genes found in three different loci. Locus 1 is present in *B. atrophaeus* NRS_1221A (AJF84093) and another *B. atrophaeus* strain (Figure 10). This locus is also present in *B. methylotrophicus* (WP_012117056, WP_012117057 and WP_012117058), *B. subtilis*, and *B. amyloliquefaciens*. There are no methyltransferases or glycosyltransferases present. Locus 2 is preserved in *B. atrophaeus* and *B. subtilis* strains (Figure 11). There is one glycosyltransferase upstream of *bclA* gene. Locus 3 is present in *B. atrophaeus* strains, *B. methylotrophicus*, *B. amyloliquefaciens* and *B. subtilis* (Figure 12). There is a methyltransferases and a glycosyltransferase present. Both *bclA* genes identified in *B. licheniformis* 14580 are present in the same locus (Figure 13). This locus is conserved only in *B. licheniformis* strains. There are no methyltransferases or glycosyltransferases present.

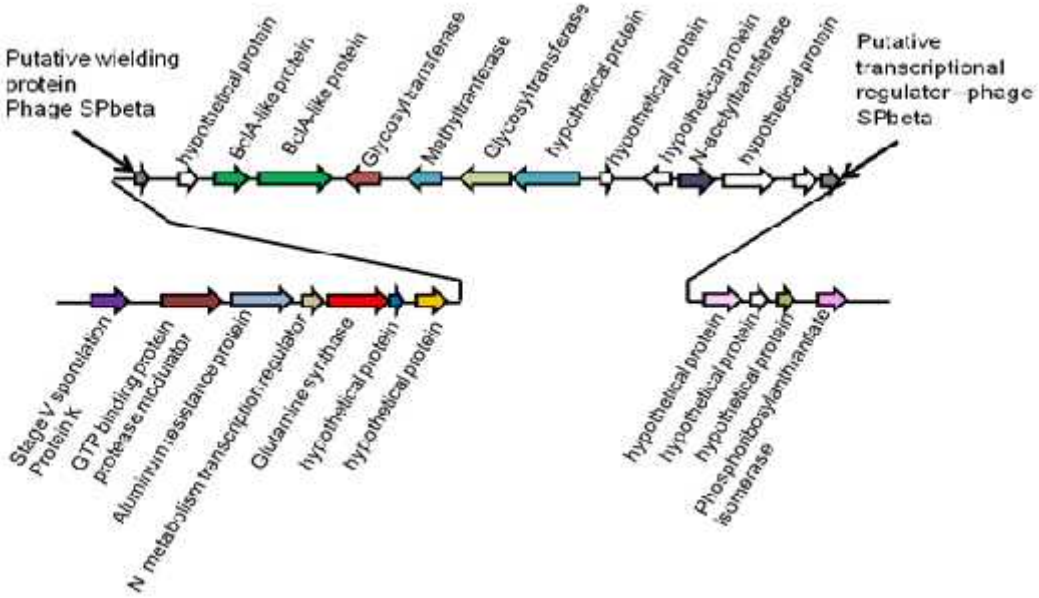

Figure 3. A cartoon depicting the first *bclA* locus in *B. subtilis* BSn5.This locus contains two *bclA*-like genes.

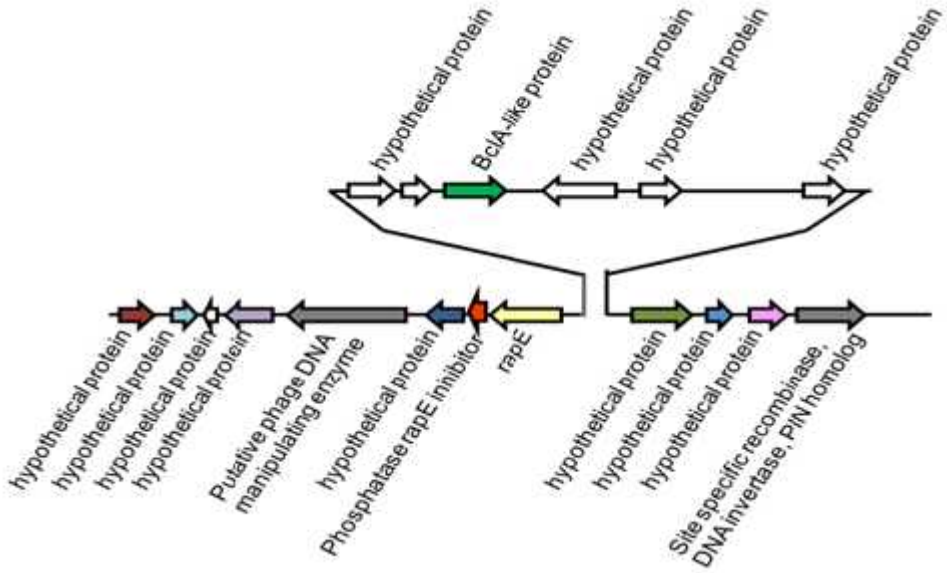

Figure 4. A cartoon depicting the second *bclA* locus in *B. subtilis* BSn5.This locus contains one *bclA*-like gene.

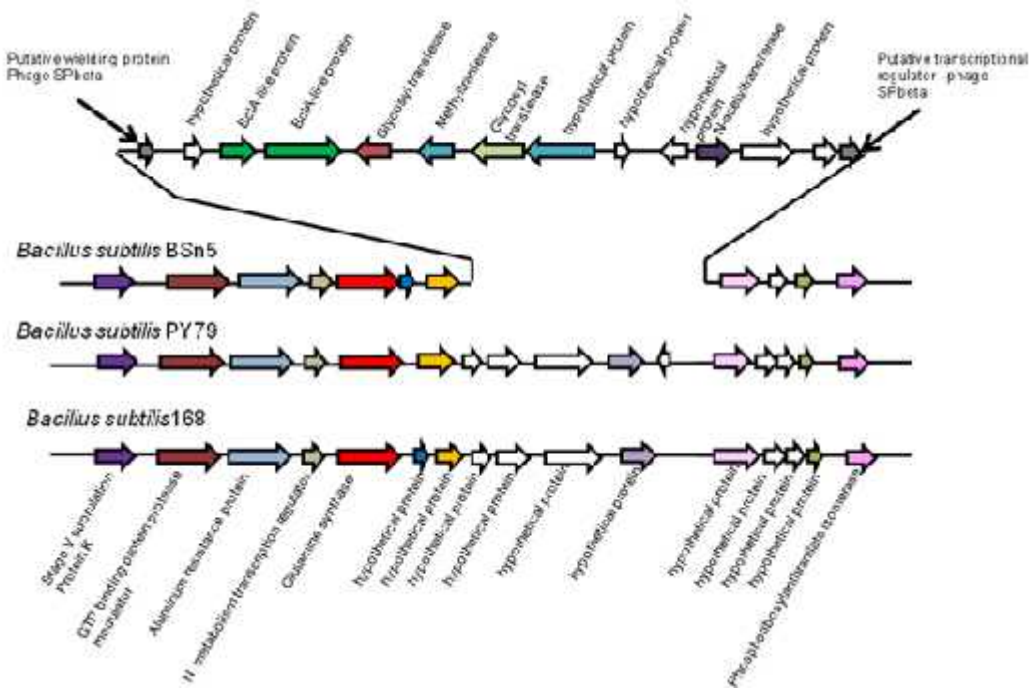

Figure 5. A cartoon comparing the first *bclA* locus found in *B. subtilis* BSn5 to the lab strains.

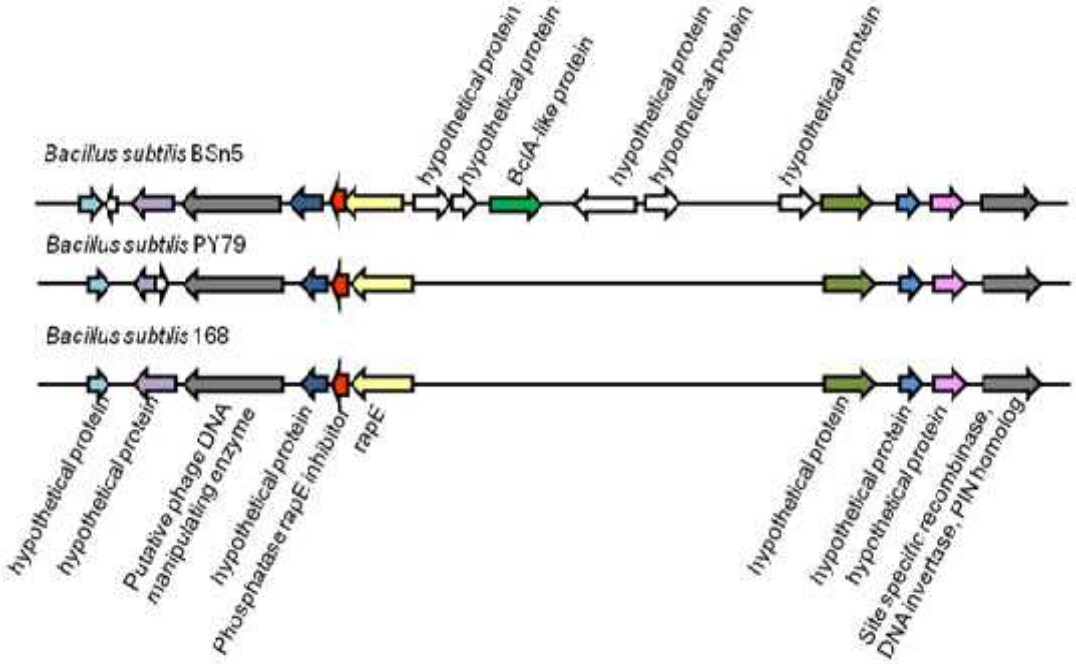

Figure 6. A cartoon comparing the second *bclA* locus found in *B. subtilis* BSn5 to the lab strains.

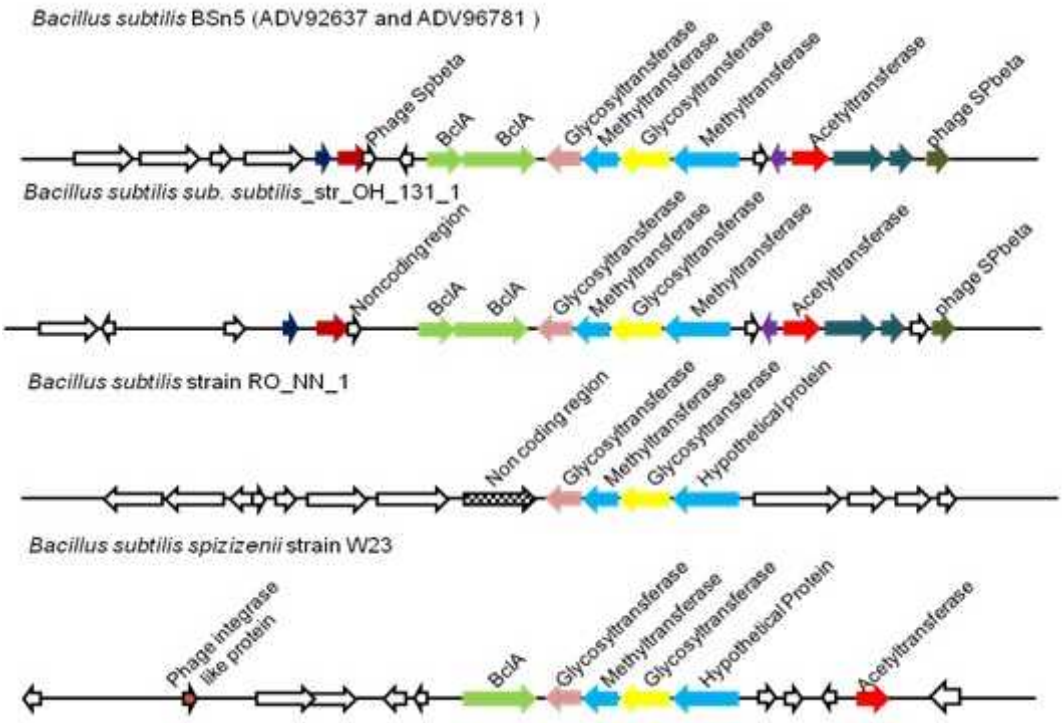

Figure 7: A cartoon depicting similarity between one of two *bclA* loci in *B. subtilis* BSn5 and other *B. subtilis* strains.

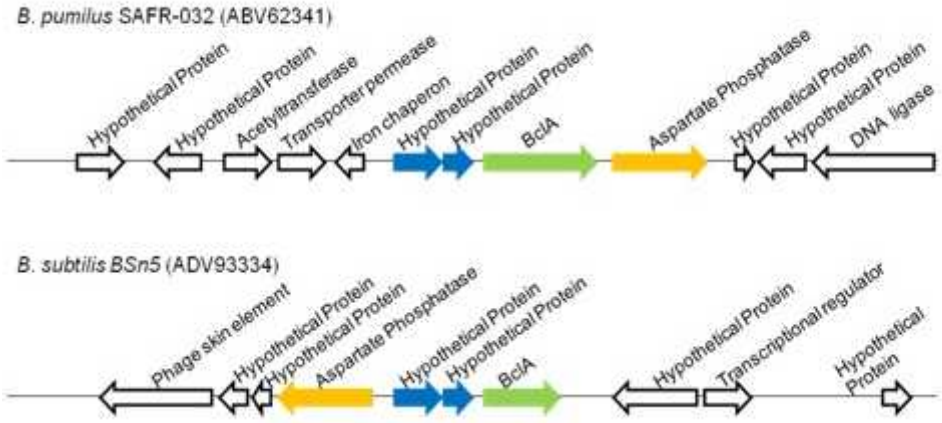

Figure 8: A cartoon depicting similarity between one of two *bclA* loci in *B. subtilis* BSn5 and *B. pumilus* SAFR-032.

Figure 9: A cartoon depicting one of two *bclA* loci in *B. pumilus* SAFR-032.

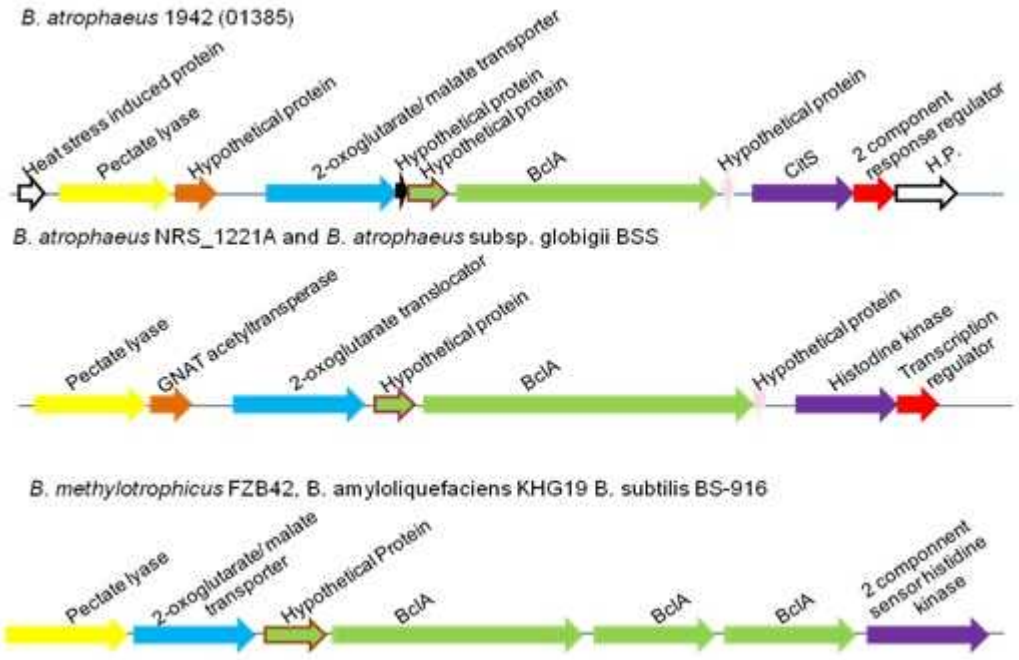

Figure 10: A cartoon depicting similarity between one of three *bclA* loci in *B. atrophaeus* 1942 and other *B. atrophaeus* strains.

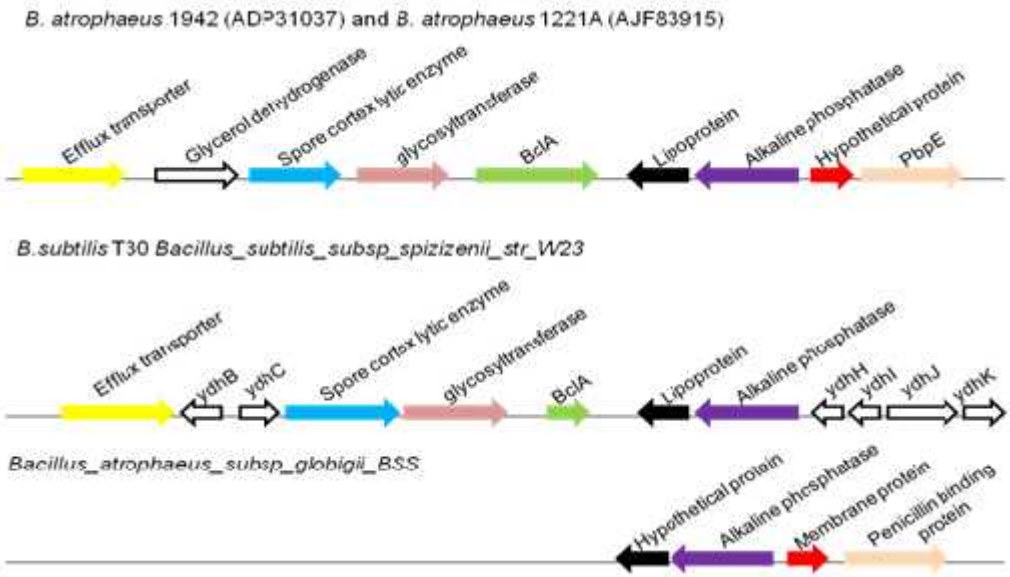

Figure 11: A cartoon depicting similarity between one of three *bclA* loci in *B. atrophaeus* 1942 and other *B. atrophaeus* strains.

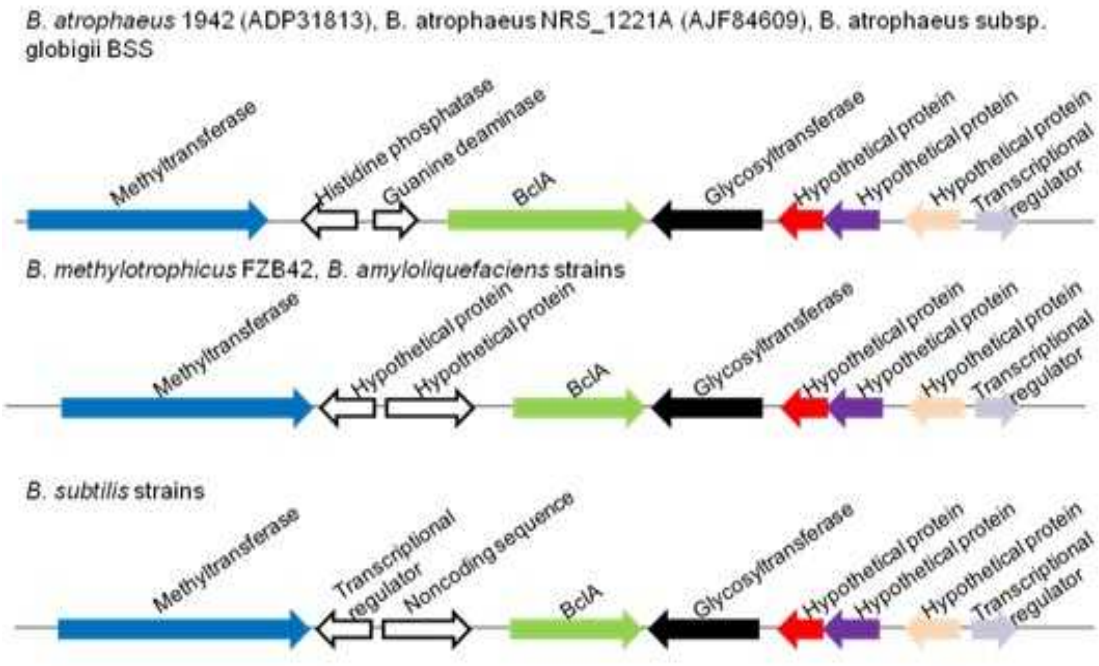

Figure 12: A cartoon depicting similarity between one of three *bclA* loci in *B. atrophaeus* 1942 and other *B. atrophaeus* strains.

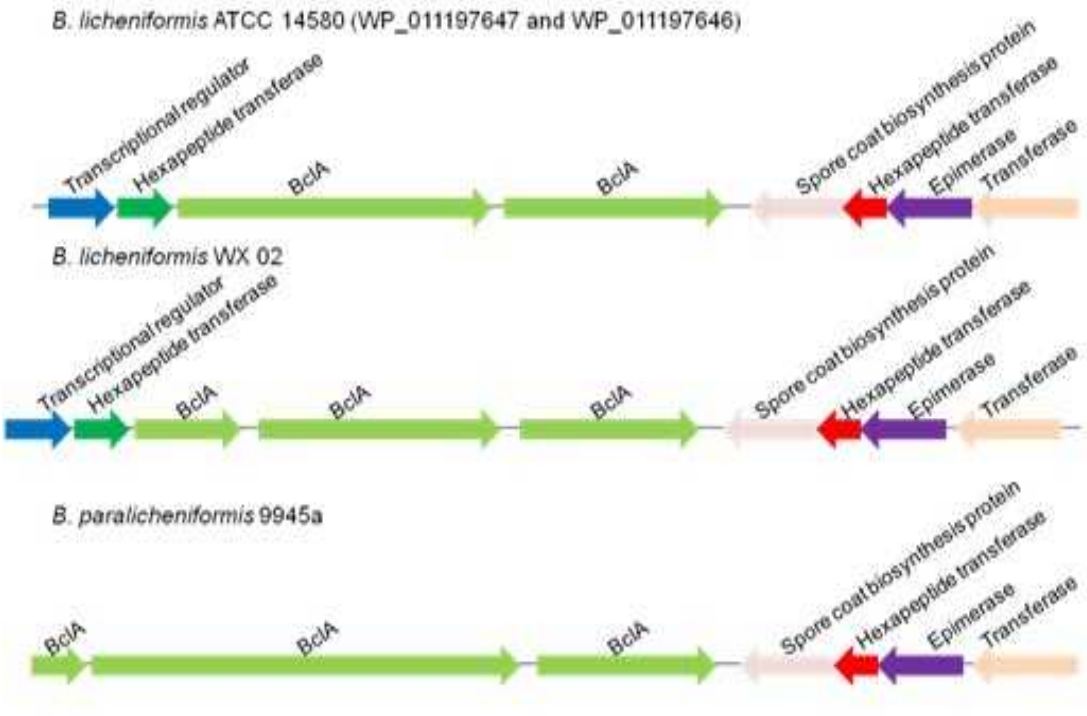

Figure 13: A cartoon depicting similarity between *bclA* loci in *B. licheniformis* 14580 and other *B. licheniformis* strains.

Conclusion Based on my limited analysis I can conclude *bclA* loci vary across species. The *bclA* locus found in *B. anthracis* is only conserved in *B. cereus* group. The locus contains one *bclA* gene surrounded by multiple methyltransferases and glycosyltransferases. This is different from *bclA* loci found in other species (*B. subtilis* BSn5, *B. methylotrophicus* FZB42 or *B. licheniformis* ATCC 14580) where two or three *bclA* genes are present close to each other. Glycosyltransferases and methyltransferases were not present in the proximity of all *bclA*-like genes. In fact, only few *B. subtilis* strains had multiple glycosyltransferases and methyltransferases in their *bclA* loci. Most of the *bclA* loci lack the transferases or have just one or two. In the future, it would be important to extend the *bclA* locus analysis to other species. For now, we only know the composition of the *bclA* loci in species closely related to *B. anthracis* and *B. subtilis*. It would be interesting to know, what type of genes surround *bclA* in genomes of species that represent *Paenibacillaceae* and *Clostridiaceae.* Are glycosyltransferases and methyltransferases close to *bclA* in the genomes of those species or is this feature only preserved in *B. cereus* group? A greater analysis of *bclA* loci would allow us to tell if there is any common feature to all of them.

Presence of BclA in Species that have *bclA* Homologous

Introduction Based on my genomic analysis, I know that *bclA* genes are present in the genomes of many spore-forming bacteria. This finding is not sufficient to conclude BclA is a spore protein in species harboring *bclA* gene as it is not known if the genes are being transcribed and translated during the sporulation. In order to explore whether BclA is present in spores that harbor *bclA*, I performed Western Blot and IFM. Additionally, I prepared thin-section TEM of unpurified spores.

Bacillus subtilis I identified three *bclA*-like genes in the genome of *B. subtilis* BSn5 (Table 6). This finding was surprising as the spores of *B. subtilis* do not possess an exosporium. To learn whether spores of *B. subtilis* BSn5 have hair-like projections on the surface I performed EM. I was able to visualize the crust of the spores. I was not able to detect hair-like projections on the spores (Figure 14). This finding could mean that either BclA does not produce hair-like structures on the spore surface of *B. subtilis* BSn5 or we cannot detect them when we stain with the ruthenium red. I questioned whether BclA is a spore protein in this species. First, I investigated whether I can detect BclA by Western blot. This experiment allowed determining if BclA is present in the coat of the *B. subtilis* BSn5 spore. I chemically extracted coat proteins by SDS and DTT (denaturing agents) and heat treatment. The chemical extraction was followed by SDS-PAGE and immunodetection. Two antibodies are used in this study to detect the presence of BclA*:* anti-BclA monoclonal antibody and anti-BclA antiserum (polyclonal antibody). The monoclonal antibody is specific for the C-terminus of *B. anthracis* BclA. Anti-BclA antiserum was raised against the entire *B. anthracis* BclA and we do not know what region of BclA it detects. The negative control for this study is the serum from a mouse injected with PBS or pre-immune mouse serum. The monoclonal antibody detected a high molecular species (>170kDa) in *B. subtilis* BSn5 (Figure 15 lane 3). Additionally, anti-BclA antiserum also detected a high molecular species (Figure 15 lane 5). This protein was not detected by the negative control (Figure 15 lane 4). Western blot results suggest that BclA is a spore protein in *B. subtilis* BSn5. Interestingly, according to the bioinformatics analysis, BclA is expected to be either 26 kDa or 71 kDa. The fact that it is detected as a high molecular weight species suggests BclA in *B. subtilis* BSn5 is

glycosylated as BclA in *B. anthracis* is. Alternatively, it might mean BclA in *B. subtilis* BSn5 is present in complexes that do not disassociate upon DTT and SDS treatment.

Knowing that BclA is a spore protein in *B. subtilis* BSn5, I asked whether I can detect the protein on the spore surface. To answer this question, I used IFM. 10 µl of spore suspension was transferred to a multiwall glass slide. Spores were first incubated with primary antibodies and afterwards with fluorescently labeled secondary antibodies. Only a subset of spores was fluorescently labeled when anti-BclA antiserum was applied (Figure 16, d-f). No fluorescence was observed with the PBS induced serum (Figure 16, a-c). This suggests the signal I detected on a subset of spores was indeed BclA. No fluorescence was detected when the monoclonal antibody was applied to the spores (Figure 16, g-i). This finding suggests the C-terminus of BclA is not accessible to the antibody on the spore surface of *B. subtilis* BSn5. The finding that only a subset of spores had BclA on the spore surface could explain why we did not see hair-like structures on the spores. Possibly, in our EM analysis, we did not look at enough spores to find hair-like projections.

Bacillus amyloliquefaciens *B. amyloliquefaciens* is another example of bacterium whose spores do not produce exosporia. BLAST analysis showed there are at least three strains of *B. amyloliquefaciens* that have *bclA*-like genes in their genomes (Table 6). The Driks lab has a collection of 6 different strains of *B. amyloliquefaciens*. However, their sequences are not present in NCBI, preventing me from determining whether they encode BclA-like proteins. Nevertheless, investigating our strains could tell me something about the frequency of BclA in nature. First, I questioned if I can visualize hair-like projections of the spore surface of one of the strains, *B. amyloliquefaciens* NRRL BD-553. I was not

able to see any hair-like structures on the spore surface (Figure 17). This could mean either BclA is not present in that strain or alternatively it does not form hair-like structures on the spore surface. Next, I wanted to know if I can detect BclA by immunological methods. I investigated three strains: *B. amyloliquefaciens* NRRL BD-553, *B. amyloliquefaciens* BGSC 10A1 and *B. amyloliquefaciens* NRRL BD-599. I used Western blot to answer the question whether BclA is a spore coat protein. The monoclonal antibody detected proteins only in *B. amyloliquefaciens* NRRL BD-553 (Figure 18, lane 1). Interestingly, multiple bands (>70 kDa, >40 kDa and <40 kDa) were detected. Sylvestre et al. observed a similar pattern, in which anti-BclA monoclonal antibody detected multiple bands in two *B. anthracis* strains [11]. She suggested that this multiple band pattern was due to post-translational modification and/or stability of the glycosylation of BclA. The monoclonal antibody did not bind to any proteins extracted from *B. amyloliquefaciens* BGSC 10A1 or *B. amyloliquefaciens* NRRL BD 599 (Figure 18, lane 2 and lane 3). This result suggests BclA is not a spore protein in those two strains. Alternatively, it could mean that BclA is present but its C-terminal domain is different than in *B. anthracis*. The anti-BclA antiserum detects multiple bands (> 40 kDa, < 40 kDa, around 31 kDa, and around 20 kDa) in *B. amyloliquefaciens* NRRL BD-553 (Figure 18, lane 4). The same sized bands are detected in the negative control (Figure 18, lane 7). This suggests the anti-BclA antiserum does not detect BclA in the coat protein extraction. The anti-BclA antiserum does not detect any proteins in *B. amyloliquefaciens* BGSC 10A1 (Figure 18, lane 5). In the spore protein extract from *B. amyloliquefaciens* NRRL BD 599, the same proteins are detected both by the polyclonal antibody as well as the negative control (Figure 18, lanes 6 and 9). Overall, the results suggest BclA (with

preserved C-terminus) is only present in *B. amyloliquefaciens* NRRL BD-553. Based on the monoclonal and the polyclonal antibody results, either BclA in not present in the coat of *B. amyloliquefaciens* BGSC 10A1 spores or BclA is not detected with our antibodies. BclA is not detected in *B. amyloliquefaciens* NRRL BD 599 with our antibodies.

I performed IFM to answer the question whether BclA is a surface protein on the *B. amyloliquefaciens* spores. IFM was performed on two strains: *B. amyloliquefaciens* NRRL BD-553 and *B. amyloliquefaciens* BGSC 10A1. Based on the Western blot results, I predicted BclA will only be found in *B. amyloliquefaciens* NRRL BD-553. Both the polyclonal antibody (Figure 19, a-c) and the monoclonal antibody (Figure 19, g-i) recognized proteins on the spore surface of *B. amyloliquefaciens* NRRL BD-553. Interestingly, the signal from the monoclonal antibody was only detected at the pole of spores. This result might indicate the C-terminal domain of BclA is only accessible to the antibody at the pole. Neither polyclonal antibody (Figure 19, d-f) nor monoclonal antibody (Figure 19, j-l) recognized any proteins on the surface of *B. amyloliquefaciens* BGSC 10A1 spores. No proteins were detected on the surface of spores in either strain by the negative control (Figure 19, m-o and p-r). Based on these results, only *B. amyloliquefaciens* NRRL BD-553 has BclA on the spore surface. Interestingly, it seems BclA does not form hair-like structures on the spore surface. Alternatively, the hair-like projections are present but I cannot detect with our method.

Bacillus atrophaeus *B. atrophaeus* is another example of species that produces spores without exosporia. I chose two strains of *B. atrophaeus:* SB512 and 1942. *B. atrophaeus* SB512 is a strain from the Driks lab collection that has not been sequenced. *B. atrophaeus* 1942 has been sequenced. Based on the bioinformatics analysis (Table 6),

there are three genes encoding collagen-like proteins in *B. atrophaeus* 1942 and two of them are homologous to *bclA*. The predicted molecular weight of the proteins is 56 kDa, 56 kDa and 147 kDa. First, I asked if I can visualize hair-like projections of the spore surface of *B. atrophaeus* 1942. I was not able to see any hair-like structures on the spore surface (Figure 20). I used Western blot to investigate if BclA is a spore protein in *B. atrophaeus* SB512 and 1942. Western blot analysis showed there is a high molecular species present in *B. atrophaeus* SB512 (Figure 21 lane 1). The high molecular species detected by monoclonal antibody was not detected by anti-BclA antiserum in Western blot (Figure 21, lane 2). This could indicate BclA protein is different from *B. anthracis* BclA in a way that makes it undetectable by the anti-BclA antiserum. Another possibility is that there is a different protein than BclA on the spore surface that is recognized by the monoclonal antibody. High molecular species were also detected by the monoclonal antibody in *B. atrophaeus* 1942 (Figure 22, lane 1). There were multiple proteins detected ranging from 55 to 100 kDa. The 55 kDa band could indicate a monomer form of one of the BclA proteins in *B. atrophaeus* 1942. This could mean that this protein is not glycosylated. Higher molecular weight species could indicate that the protein is glycosylated or in a complex with other proteins. This complex would have to be stable even after SDS-DTT and heat treatment. Species detected by anti-BclA antiserum were also detected by the negative control (Figure 22, lane 2 and 3). To determine if BclA is a spore surface protein in *B. atrophaeus*, I used IFM. Monoclonal antibody detects proteins at the poles of *B. atrophaeus* SB512 spores (Figure 23, a-c). Based on this result we can infer either BclA is present only at the poles of spores or BclA is present on the entire surface but the C-terminal domain is accessible to the antibody only at the poles. Neither

anti-BclA antiserum nor the negative control detected any proteins on the surface of the spores (Figure 23, g-1 and d-f, respectively). A similar pattern of BclA localization was observed in *B. atrophaeus* 1942. The monoclonal antibody detected BclA at the poles of the spores (Figure 24, a-c). Neither anti-BclA antiserum nor the negative control detects any proteins on the surface of the spores (Figure 24, d-f and g-i).

Bacillus methylotrophicus *B. methylotrophicus* FZB42 has four *bclA*-like genes in its genome (Table 6). I performed Western blot to investigate whether BclA is a coat protein in *B. methylotrophicus* FZB42. The monoclonal antibody detected multiple high molecular weight species (Figure 25). This finding suggests BclA is a coat protein in *B. methylotrophicus* FZB42. I speculate either BclA is present in protein complexes that resist the denaturating treatment or the protein is subject to degradation and hence the multiple bands. Alternatively, BclA could be glycosylated differently. My third speculation is the monoclonal antibody detects BclA products from multiple *bclA*-like genes. IFM was done with the monoclonal antibody to see whether BclA is surface exposed in *B. methylotrophicus* FZB42. Fluorescence was only present at the poles of the spores (Figure 26). These data lead to the conclusion BclA is localized at the poles of the *B. methylotrophicus* FZB42 spores. An alternative view is BclA is only accessible to the antibody at the poles of the spores.

Paenibacillus chitinolyticus and *Brevibacillus laterosporus* are exosporium-positive species that belong to *Paenibacillaceae* family. If I find BclA on the surface of spores of these two species I will conclude *B. cereus* group members are not the only ones that produce spores whose exosporia are decorated with BclA.

Brevibacillus laterosporus I used EM to see whether there are hair-like projections on the spore surface of *B. laterosporus*. I was able to make out hair-like structures on the entire surface of the spore, including the asymmetrical structure present at one side (Figure 27). Western blot analysis was used to ask whether BclA is present in the outer layer of *B. laterosporus* spores. The monoclonal antibody detected a >55 kDa species (Figure 28). This suggests BclA is a spore protein in *B. laterosporus* 9141 and the C-terminal domain is preserved in the BclA. It is difficult to guess which *bclA*-like gene is potentially responsible for the signal as the estimated size of the BclA proteins range from 40 to 68 kDa. We cannot exclude some glycosylation of the proteins and/or a cleavage event before the protein was incorporated into the spore surface. I performed IFM to answer the question whether BclA is present on the spore surface of *B. laterosporus* 9141. As IFM images show, the entire surface of spores is fluorescently labeled (Figure 29). This suggests BclA is present on the entire surface of the spores.

Paenibacillus chitinolyticus Finally, I wanted to establish whether BclA is present on the spore surface of *Paenibacillus chitinolyticus* NBRC 15660. I used polyclonal anti BclA antibody. I detected fluorescence at the poles of the spores which suggest the location of BclA (Figure 30). I did not perform the Western analysis on *P. chitinolyticus* NBRC 15660.

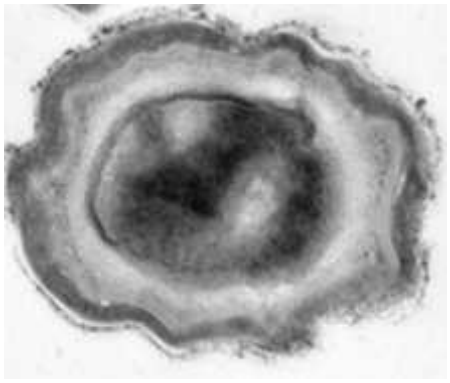

Figure 14. TEM image of a *B. subtilis* BSn5 spore

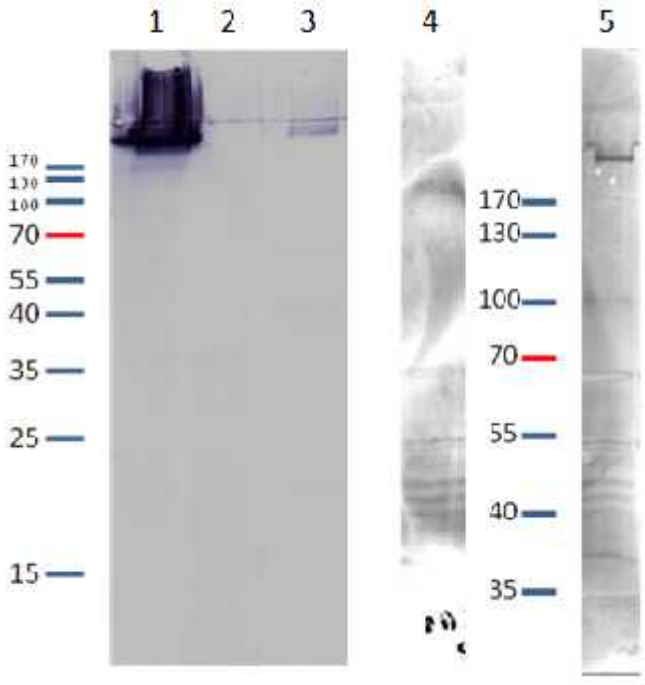

Figure 15. Western blot analysis of spore coat proteins from *B. subtilis* BSn5 and *B. anthracis*. Lanes 1: *B. anthracis*, lane 2: *B. anthracis bclA*, lanes 3-5*: B. subtilis* BSn5. Proteins were separated on 15% (lanes 1-3) or 8% (lanes 4-5) polyacrylamide gels and transferred to a membrane. Anti-BclA monoclonal antibody was used to detect BclA (lanes 1-3). Serum from a mouse injected with PBS was used as a negative control (lane 4) to anti-BclA antiserum (lane 5). Antibody concentration: 1:10, 000 for anti-BclA monoclonal antibody and 1:500 for both sera. The size of markers is given in kDa.

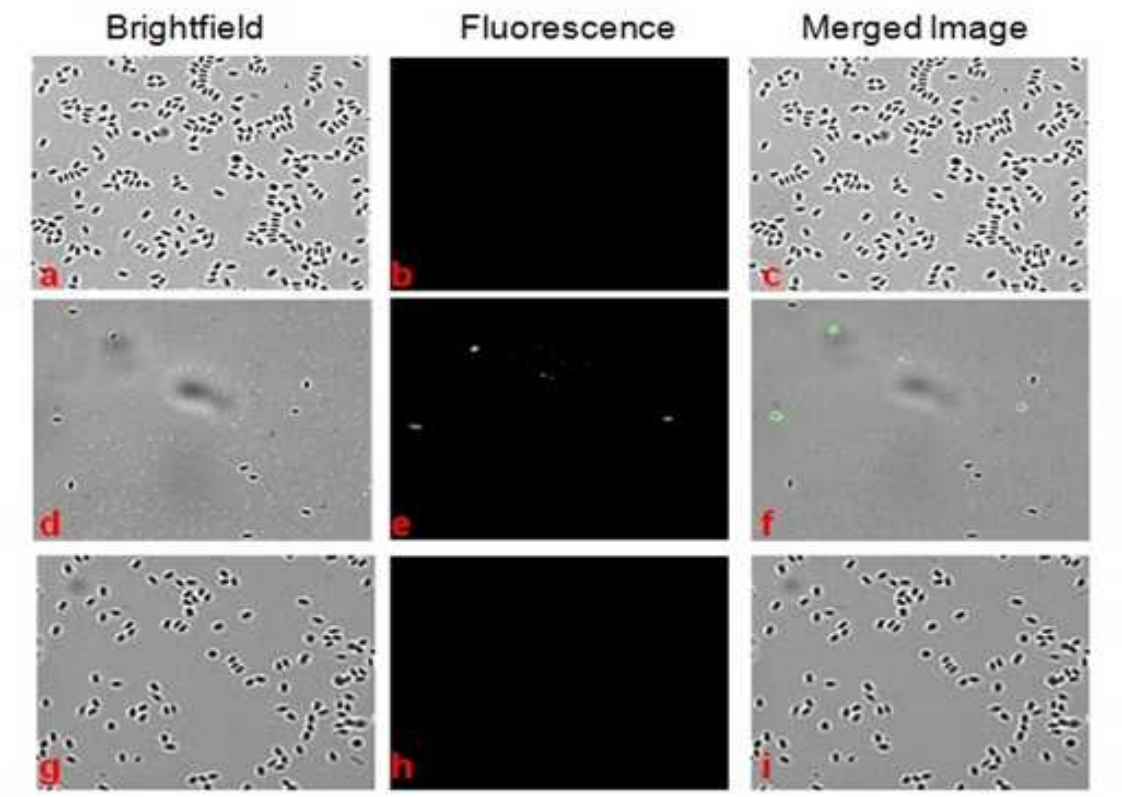

Figure 16. IFM analysis of *B. subtilis* BSn5 spores. Spores were fixed onto a slide and treated with serum from mouse injected with PBS (a-c), anti-BclA antiserum (d-f) or anti-BclA monoclonal antibody (g-i). Antibody concentration: 1:50 negative control, 1:50 anti-BclA antiserum and 1:10 000 anti-BclA monoclonal antibody.

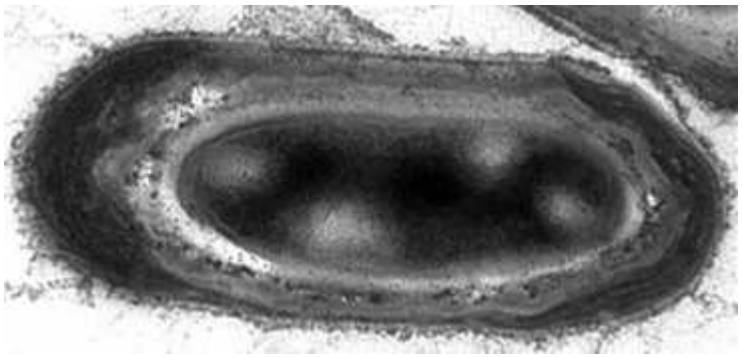

Figure 17. TEM image of a *B. amyloliquefaciens* NRRL BD-553 spore.

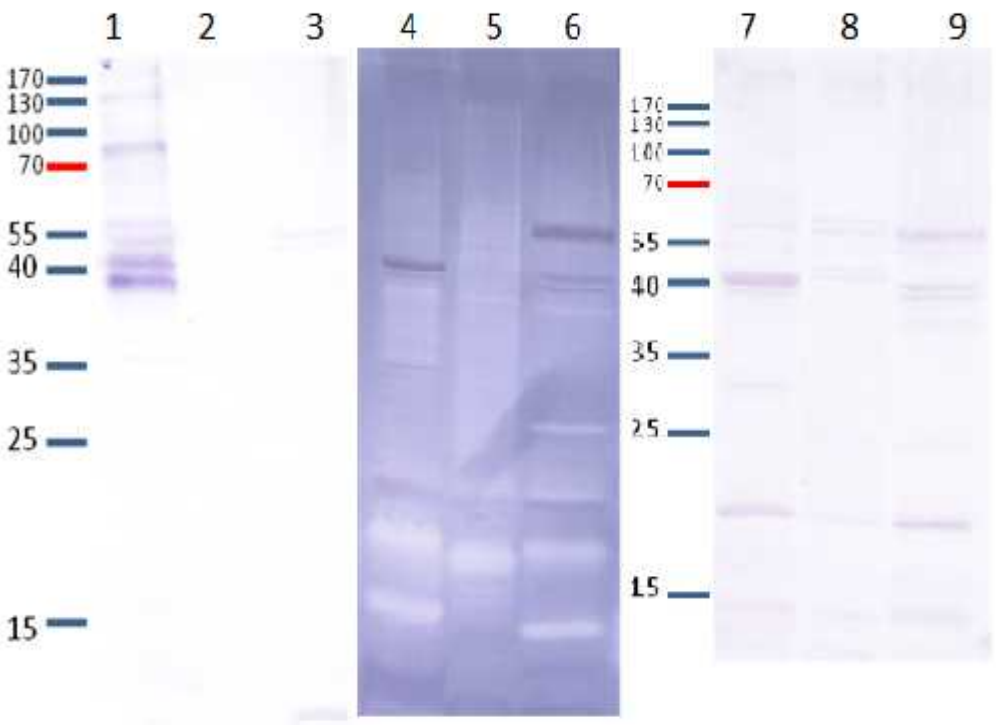

Figure 18. Western blot analysis of spore coat proteins from *B. amyloliquefaciens* strains. Proteins were separated on 15% polyacrylamide gels and transferred to a membrane. Anti-BclA monoclonal antibody was used to detect BclA (lanes 1-3). Anti-BclA antiserum was used to detect BclA (lanes 4-6). Serum from a mouse injected with PBS was used as a negative control (lanes 7-9). The figure shows *B. amyloliquefaciens* NRRL BD-553 (lanes 1, 4 and 7), *B. amyloliquefaciens* BGSC 10A1 (lanes 2,5 and 8) and *B. amyloliquefaciens* NRRL BD-599 (lanes 3, 6 and 9). Antibody concentration: anti-BclA monoclonal antibody: 1:8,000, sera: 1:350. The size of markers is given in kDa.

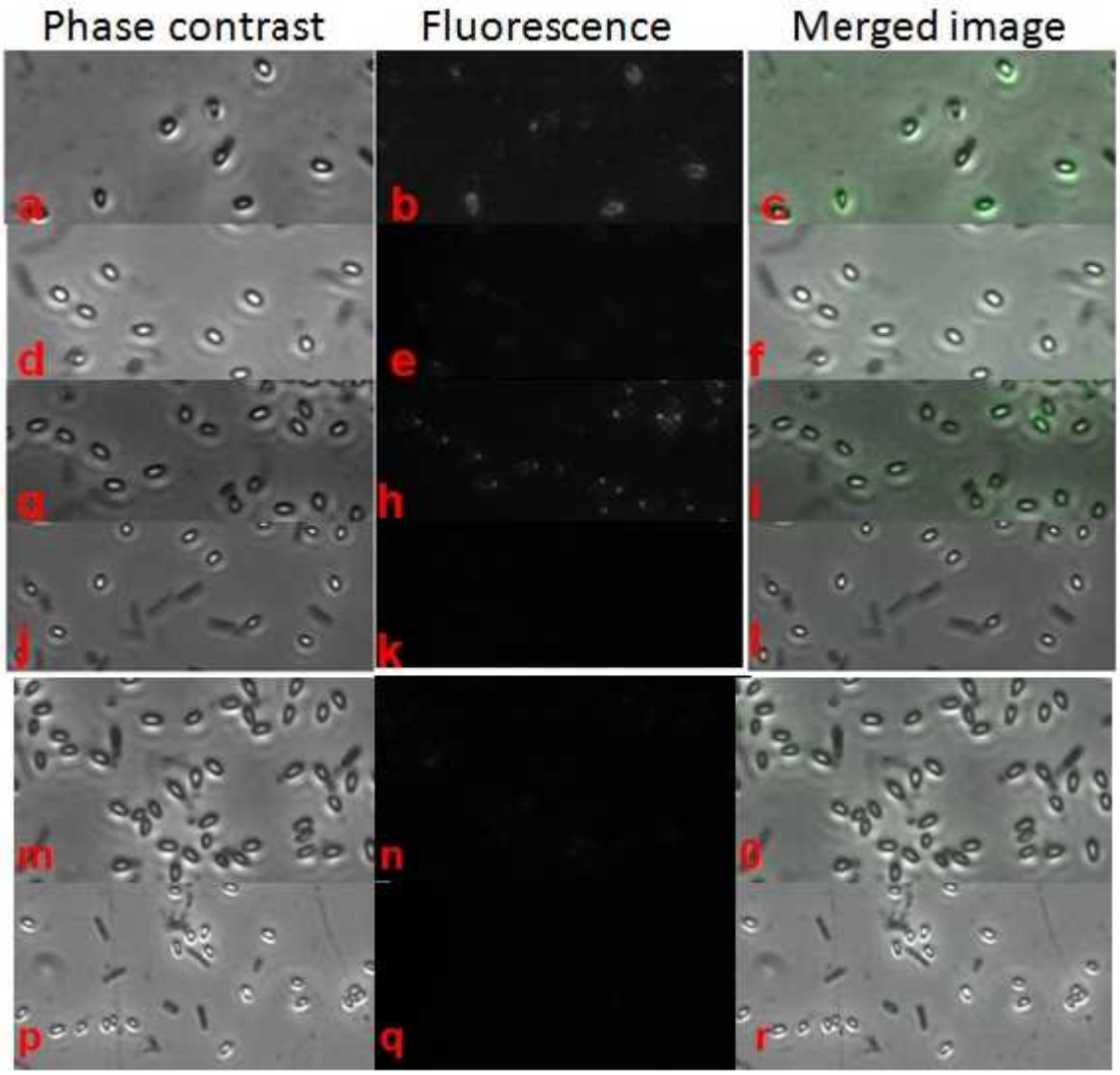

Figure 19. IFM analysis of *B. amyloliquefaciens* NRRL BD-553 and *B. amyloliquefaciens* BGSC 10A1 spores. Spores were fixed onto a slide and treated with anti-BclA antiserum (a-f), PBS serum (m-r) or anti-BclA monoclonal antibody (g-l). The images show *B. amyloliquefaciens* NRRL BD-553 (a-c, g-I and m-o) and *B. amyloliquefaciens* BGSC 10A1 (d-f, j-l and p-r). Antibody concentration: 1:50 anti-BclA antiserum, 1:50 PBS serum, and 1:5 000 anti-BclA monoclonal antibody.

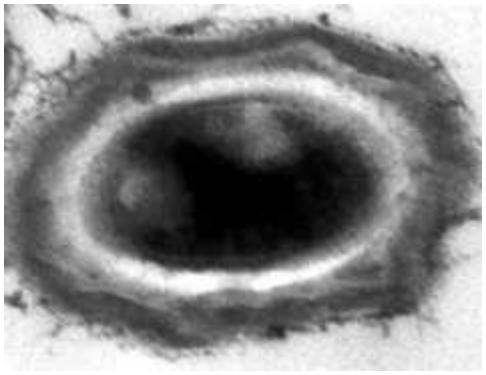

Figure 20. TEM image of a *B. atrophaeus* 1942 spore.

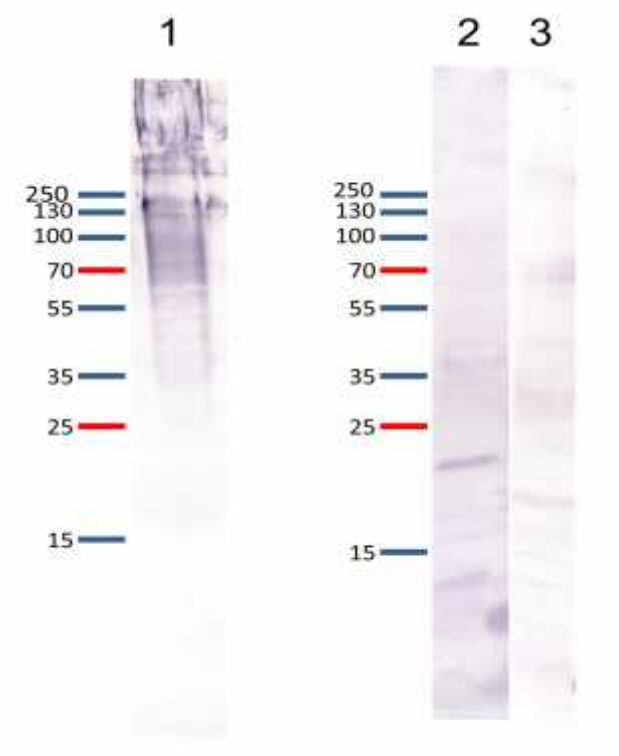

Figure 21. Western blot analysis of spore coat proteins from *B. atrophaeus* SB512. Proteins were separated on 15% polyacrylamide gels and transferred to a membrane. Anti-BclA monoclonal antibody and anti-BclA antiserum were used to detect BclA (lanes 1 and 2 respectively). Serum from a mouse injected with PBS was used as a negative control (lane 3). Monoclonal antibody concentration: 1:8000. Serum and antiserum concentration: 1: 350. The size of markers is given in kDa

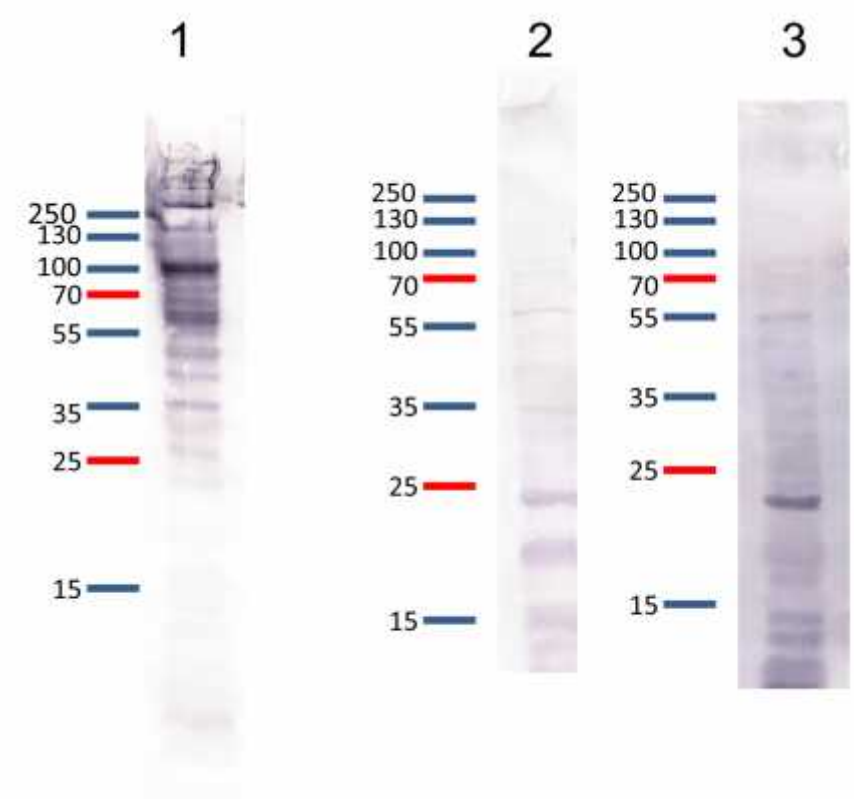

Figure 22. Western blot analysis of spore coat proteins from *B. atrophaeus* 1942. Proteins were separated on 15% polyacrylamide gels and transferred to a membrane. Anti-BclA monoclonal antibody and anti-BclA antiserum were used to detect BclA (lanes 1 and 3 respectively). Serum from a mouse injected with PBS was used as a negative control (lane 2). Monoclonal antibody concentration: 1:8000. Serum and antiserum concentration: 1: 350. The size of markers is given in kDa

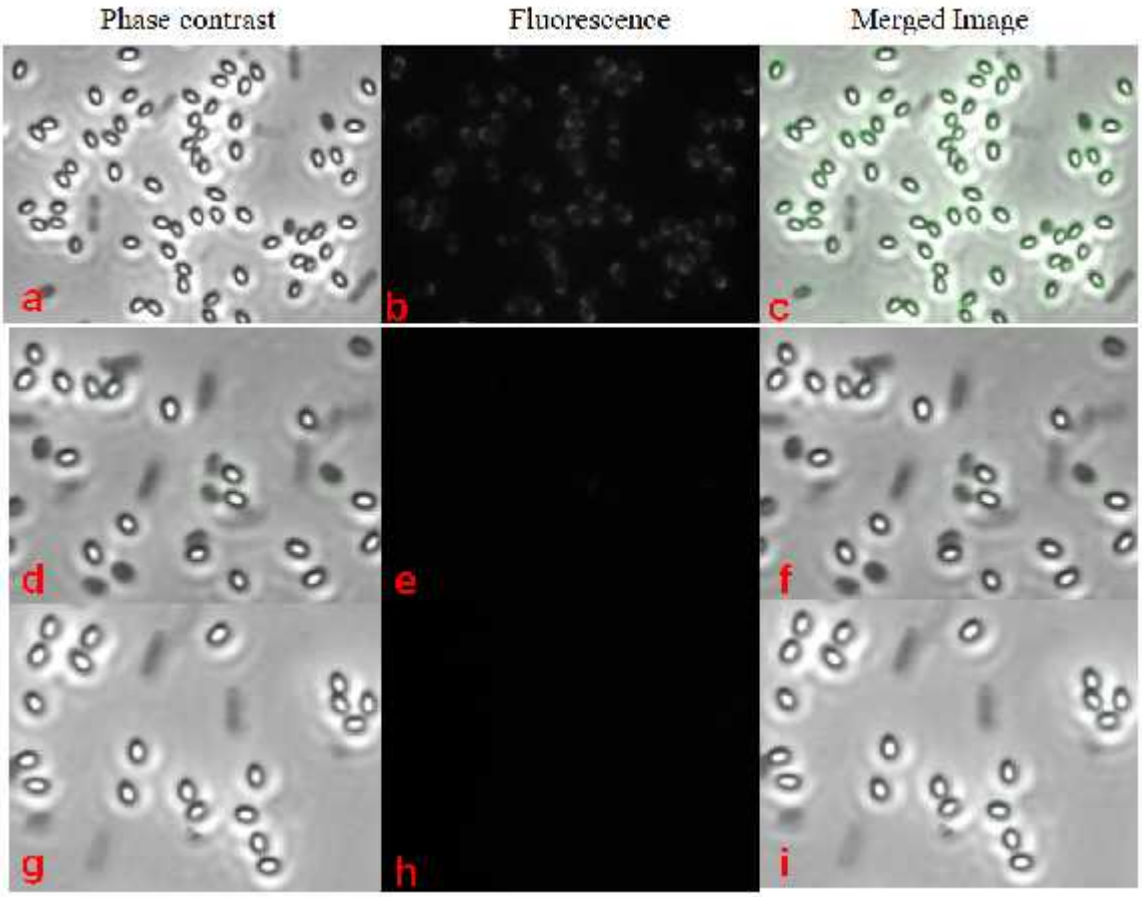

Figure 23. IFM analysis of *B. atrophaeus* SB512 spores. Spores were fixed onto a slide and treated with anti-BclA monoclonal antibody (a-c), serum from mouse injected with PBS (d-f) or anti-BclA antiserum (g-i) or anti-BclA Antibody concentration: 1:50 negative control, 1:50 anti-BclA antiserum and 1:8 000 anti-BclA monoclonal antibody

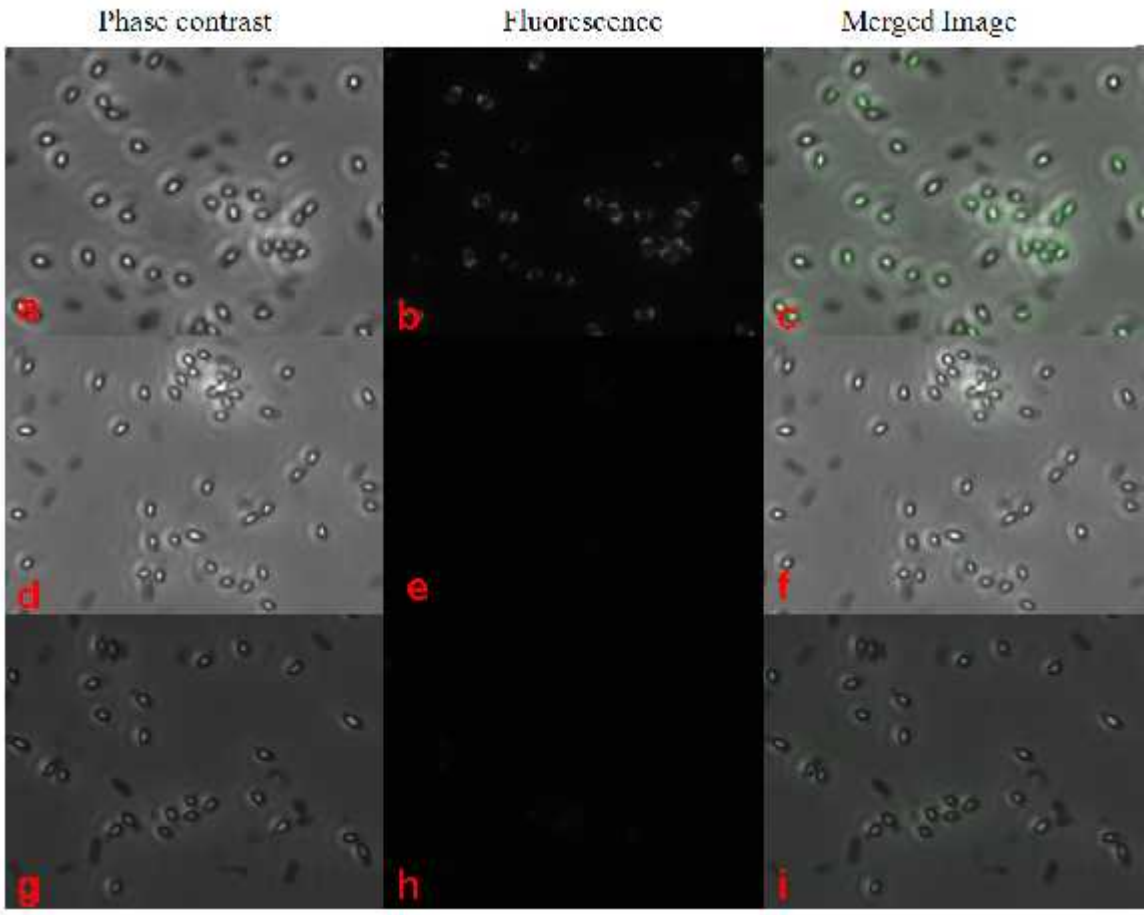

Figure 24. IFM analysis of *B. atrophaeus* 1942 spores. Spores were fixed onto a slide and treated with anti-BclA monoclonal antibody (a-c), serum from mouse injected with PBS (d-f) or anti-BclA antiserum (g-i) or anti-BclA Antibody concentration: 1:50 negative control, 1:50 anti-BclA antiserum and 1:8 000 anti-BclA monoclonal antibody

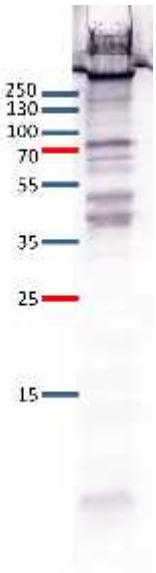

Figure 25. Western blot analysis of *B. methylotrophicus* FZB42 spore coat proteins extraction. Proteins were separated on 15% polyacrylamide gel and transferred to a membrane. Anti BclA monoclonal antibody was used to detect BclA. Antibody concentration: 1: 8000. The size of markers is given in kDa.

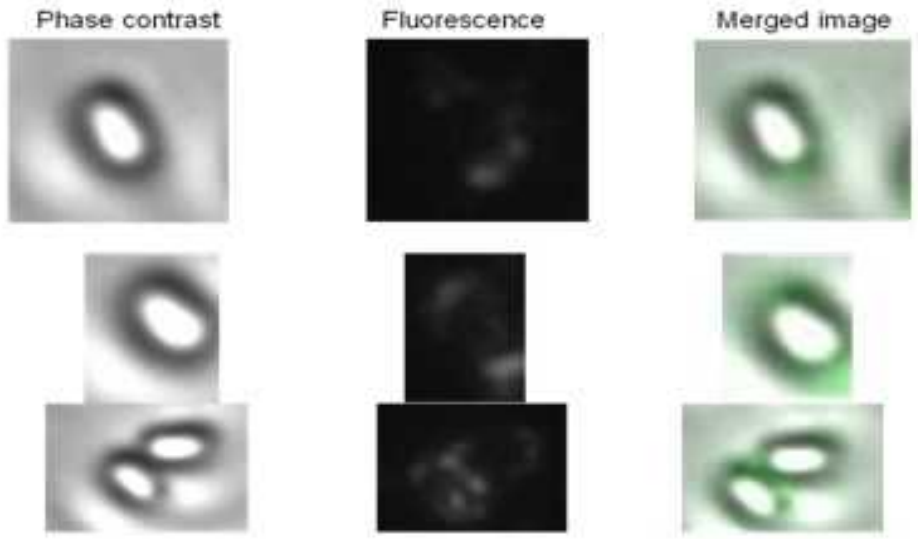

Figure 26. IFM analysis of *B. methylotrophicus* FZB42 spores. Spores were fixed onto a slide and treated with anti BclA monoclonal antibody. Antibody concentration 1:5000.

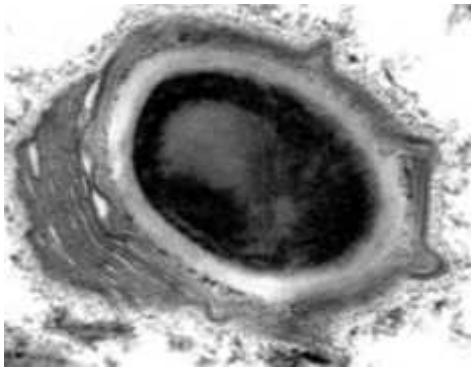

Figure 27. TEM image of a *Br. laterosporus* 9141 spore.

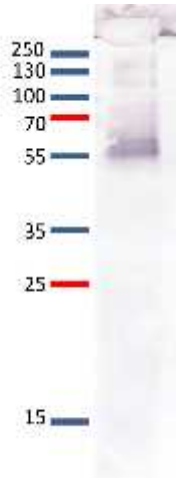

Figure 28. Western blot analysis of *B. laterosporus* 9141 spore outer layer proteins extraction. Proteins were separated on 15% polyacrylamide gel and transferred to a membrane. Anti BclA monoclonal antibody was used to detect BclA. Antibody concentration: 1: 8000. The size of markers is given in kDa.

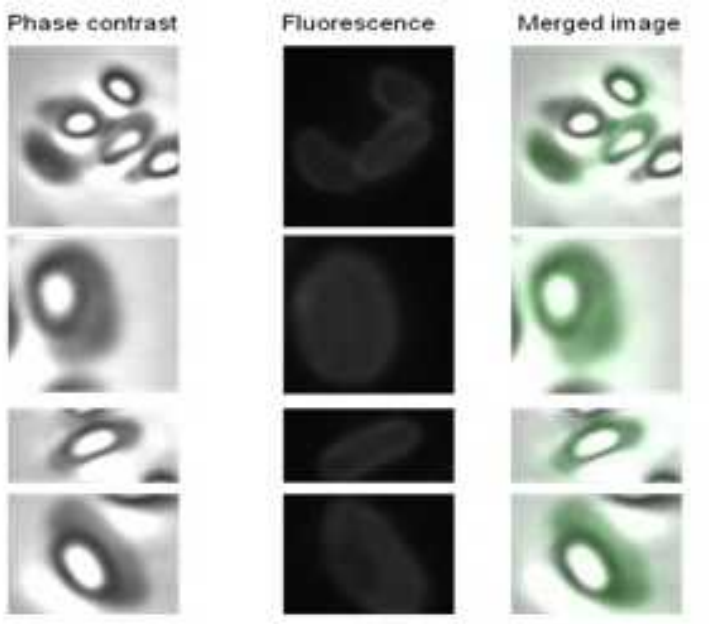

Figure 29. IFM analysis of *B. laterosporus* 9141 spores. Spores were fixed onto a slide and treated with anti BclA monoclonal antibody. Antibody concentration 1:5000

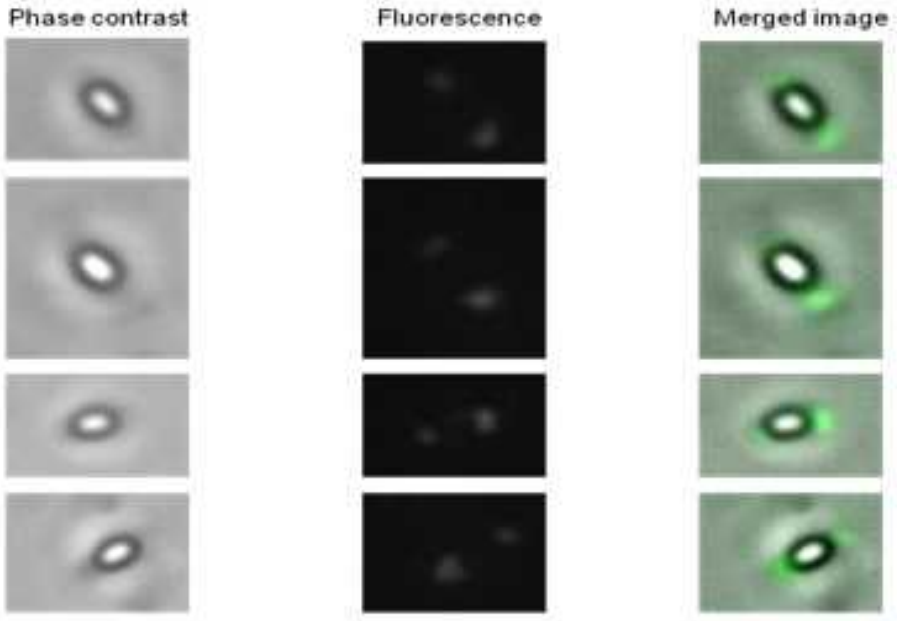

Figure 30. IFM analysis of *Paenibacillus chitinolyticus* NBRC 15660 spores. Spores were fixed onto a slide and treated with anti BclA monoclonal antibody. Antibody concentration 1:100.

Mutations in *bclA* Genes

Introduction BclA was detected on the spore surface of *B. subtilis* BSn5 and *B. atrophaeus* 1942. I was also able to identify *bclA* candidate genes in both of the species. We want to know whether the proteins we detect are indeed products of *bclA*-like gene. Because there are multiple *bclA* genes in both species, I do not know which one(s) contribute to the signal. I will inactivate each of them individually by disrupting the corresponding gene. I decided to disrupt *bclA* like genes using homologous recombination via a single cross over event (Figure 31). By integrating the plasmid into the gene of interest, I will disrupt its function. The plasmid that I am going to use has the spectinomycin resistance gene which will facilitate the screening process. Because the N-terminal domain is responsible for anchoring BclA to the spore surface in *B. anthracis* [24], I expect other BclA-like proteins will preserve this feature. Because I want to establish whether BclA is a spore surface protein in other species, I will preserve the sequence for the N-terminal domain of the native gene, and link it to the *gfp* gene (Figure 32). By preserving the N-terminus, followed by GFP, I will see where the protein localizes. If the protein I detected with anti-BclA antibodies is not a product of *bclA* genes, I will not detect GFP protein on the surface of spores. To complement my mutant, I will express the wild type gene from another plasmid.

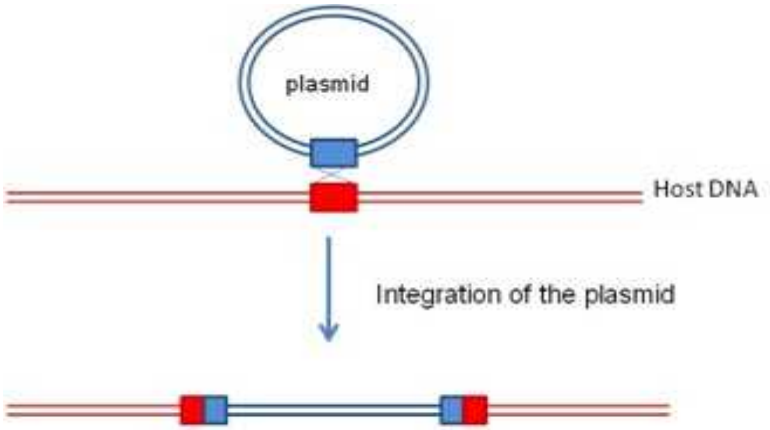

Figure 31: *bclA* gene inactivation in *B. subtilis* BSn5 and *B. atrophaeus* 1942-A. Gene inactivation by plasmid integration.

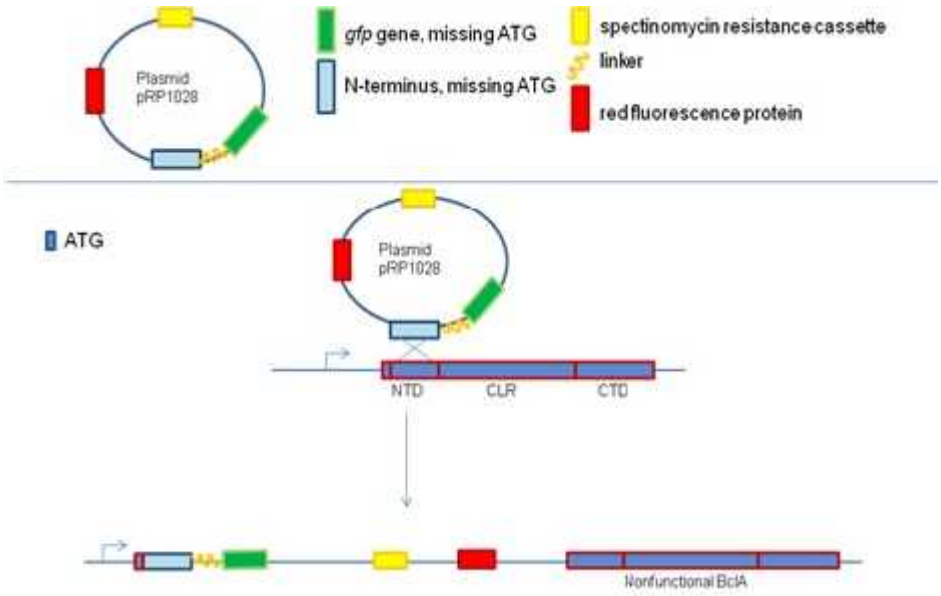

Figure 32. *bclA* gene inactivation in *B. subtilis* BSn5 and *B. atrophaeus* 1942-B. *Gfp* gene is inserted into the pRP1028 plasmid and linked to the truncated N-terminus of *bclA* gene. Plasmid will insert into the genome and disrupt the *bclA* gene.

Inactivation of *bclA* Genes in *Bacillus subtilis* BSn5 I successfully built two knockout plasmids. We decided to inactive two out of three genes in *B. subtilis* BSn5 (accession number ADV93334 and ADV96781). The third *bclA* gene has a very short N-terminal domain (7 amino acids long) and we believed it would not suffice to localize the protein to the spore surface. Both plasmids were sequenced to confirm a proper insertion of *bclA* fragment. First, I attempted to create mutants by following the protocol for three parental mating which involves a donor strain, a helper strain and a recipient strain. This protocol has been successful in introducing pRP1028 into *B. anthracis*. No *B. subtilis* colonies grew on the selective medium after the presumed conjugation. I concluded I was not able to introduce either of the two plasmids into *B. subtilis* BSn5. I performed the procedure again and this time I set up two controls: I used *B. anthracis* as my positive control as I knew the protocol works for this species. I used *B. subtilis* PY79 which is our lab strain. I wanted to see whether I am able to introduce pRP1028 into an organism which is the same species as BSn5. In this round of experiments, I used an empty pRP1028 plasmid meaning there was no *bclA* or *gfp* insert. I successfully introduced pRP1028 into *B. anthracis*. Since no *B. subtilis* colonies grew on the selective medium, I concluded I cannot introduce pRP1028 into either *B. subtilis* PY79 or BSn5 by this technique. Additionally, I investigated whether pRP1028 was the problem and therefore I could not insert in into *B. subtilis*. I used the same procedure to introduce another plasmid into B. *anthracis, B. subtilis* BSn5 and PY79. pBKJ236 is the first generation plasmid from which pRP1028 was derived. Similarly as with pRP1028, I successfully introduced the plasmid into *B. anthracis*, but failed with *B. subtilis* BSn5 and PY79. There is a report in the literature suggesting that passing plasmids through *E. coli* strains deficient in

adenine and cytosine methylation increases the efficiently of plasmid transfer into *B. anthracis* [41]. Although introducing the plasmid into *B. anthracis* was not an issue, I decided to test whether I was able to introduce the demethylated pRP1028 or pBKJ236 into either *B. subtilis* BSn5 or PY79. Again, I did not observe any bacillus colonies in either strain with either pRP1028 or pBKJ236.I concluded three parental mating is not a good method to introduce either pRP1028 or pBKJ236 into *B. subtilis* strains. I moved on to the eletroporation protocol. I tried to transform both *B. subtilis* BSn5 and PY79 with pRP1028 and pBKJ236. Again, I was not able to create mutants. Because I started suspecting pRP1028 is the problem per se, I chose additional two plasmids to test. I chose pEO-3 and pIMAY as both these plasmids have temperature sensitive origin of replication. This feature is important in my study as I try to force the plasmid into the genome of *B. subtilis* BSn5. pEO-3 is used in *B. anthracis* studies and pIMAY is used in *S. aureus* studies. I successfully introduced both of the plasmids into *B. subtilis* PY79. However, I was not able to introduce either of the plasmids into *B. subtilis* BSn5. These results led to the conclusion I do not have a proper protocol to introduce a plasmid into a *B. subtilis* BSn5.

A study by Xue et al investigated electroporation efficiency in various media [30]. Specifically, the group investigated the presence of sorbitol and mannitol in eletroporation, growth and recovery media and its effect on electroporation efficiency. The group established LB +0.5 M sorbitol to be the best growth medium, LB + 0.5 M sorbitol + 0.38 M mannitol to be the best recovery medium. The electroporation medium contained 0.5 M mannitol and 0.5 M sorbitol. I followed our protocol with the updated

media. I was not successful in obtaining any *B. subtilis* BSn5 colonies on the selective medium and I concluded this protocol was not suitable for my strain.

A study by Zhang et al. explored several methods of electroporation to successfully introduce a plasmid into *B. subtilis* ZK [31]. The researchers claimed that two additional steps have to be taken in order to successfully introduce a plasmid into a wild *B. subtilis*. First, the cells have to be grown in medium with wall-weakening agents. Ampicillin, glycine, threonine and Tween 80 were used in the study and showed to increase eletroporation efficiency. Additionally, the researchers demonstrated electroporation efficiency significantly increases when field strength is increased. They reached the highest electroporation efficiency with the field strength of 20,000 KV/cm. The field strength routinely used in our electroporation protocol is 12,500 KV/cm. I updated my protocol by adding 1% threonine and 0.07% Tween 80 into my growing cells one hour prior to the cell harvest (as stated in the article). Additionally, I increased the field strength to 20,000 KV/cm. I did not obtain any *B. subtilis* colonies on the selective medium. I concluded this protocol was not successful in introducing pRP1028 into the cell.

My failed attempts suggest at this time we do not have a proper protocol to transform *B. subtilis* BSn5. I also conclude that pRP1028 might not be the right choice for my study as I was not able to introduce this plasmid into *B. subtilis* PY79 but succeeded with pEO-3 and pIMAY.

Inactivation of *bclA* Genes in *Bacillus atrophaeus* 1942 I successfully built the knockout plasmids. Plasmids were sequenced to confirm a proper insertion of *bclA*

fragment. I followed all the procedures mentioned in the *B. subtilis* BSn5 section with the same results. None of the procedures were suitable for creating the mutants.

Conclusion At this time, I do not have a clear idea why I cannot create mutants in either *B. subtilis* BSn5 or *B. atrophaeus* 1942. I believe I do not have an appropriate protocol for transformation of either of the species. Additionally, based on a failed attempt to introduce pRP1028 into *B. subtilis* PY79, I speculate this plasmid is not appropriate for *B. subtilis* group species (or at least under the conditions I tested). One of the ways mutants are created in *B. subtilis* is by inducing competency in the cells. Competent *B. subtilis* cells will pick up exogenous DNA (like plasmids). Recently, Zhao et al. adapted a protocol that induces competence in *B. subtilis* to *B. amyloliquefaciens* and succeeded in creating mutants [43]. The plasmid used in the study was pMUTIN4. The next step in this study is to adopt Zhao's protocol with pMUTIN4 plasmid and attempt to introduce this plasmid into *B. subtilis* BSn5 and *B. atrophaeus* 1942.

The Effect of BclA on Spore Surface Properties

Measuring Hydrophobicity of Spores The effect of BclA on spore surface properties in not known. We wanted to establish whether lack of BclA on the spore surface will cause changes in spore surface properties such as hydrophobicity. Studies by Brahmbhatt et al. show spores are less hydrophobic when BclA is absent [16]. If spores are treated with heat beforehand, *bclA* spores become more hydrophobic but the heat treatment has little effect on the hydrophobicity of the wild type spores. First of all, I wanted to establish whether I can get similar result to the published one when I use our *B. anthracis* wild type spores and *bclA* spores. First, I investigated whether the amount of hexadecane has an effect on the hydrophobicity of the spores. I mixed spore suspensions with three

different amounts of hexadecane and calculated the hydrophobicity of wild type and *bclA* spores. My results show that no matter the amount of hexadecane, both the wild type and *bclA* spores gave similar results (Figure 33). The hydrophobicity values range from 79% to 91% for wild type spores and 81% to 92% for the *bclA* spores. These results tell us both BclA-positive and BclA-negative spores are equally hydrophobic. Under my conditions, I was not able to detect differences in the spore surface properties in *B. anthracis* spores. Because it has been shown that heat treatment has an effect on spore hydrophobicity of *bclA* mutants I asked if heat treatment changes hydrophobicity of my spores. I pretreated the spores with various temperatures (37, 65 and 100 °C) prior to the BATH assay. I found heat treatment at 65 and 100 °C slightly decreased wild type spore hydrophobicity. There was also a slight decrease in hydrophobicity in *bclA* mutant spores when prior heat of 65 and 100 °C was applied. Overall, the spores of both strains had similar hydrophobicity and I concluded that heat treatment does not cause noticeable changes to the spore surface hydrophobicity (Figure 34).

I examined if various amount of hexadecane have an effect on spore hydrophobicity in *B. subtilis* BSn5 and *B. atrophaeus* 1942. Similar to the *B. anthracis* assay, I mixed spore suspensions with three different amounts of hexadecane and calculated the hydrophobicity of the spores. Because I did not create *bclA* mutants in either *B. subtilis* BSn5 or *B. atrophaeus* 1942, I was only able to evaluate the hydrophobicity of the wild type spores. The percent of hydrophobicity increase with the increase amount of hexadecane added to the spore suspension of *B. subtilis* Bsn5 (Figure 35). The values ranged from 40 to 80%. There was only a slight increase in hydrophobicity of *B.*

atrophaeus 1942 in response to increase in hexadecane concentration. I concluded *B. atrophaeus* 1942 spores are less hydrophobic than spores of *B. subtilis* BSn5.

Measuring Clumping of Spores A study on *B. subtilis* lacking polysaccharides on the spore surface shows spores clump more readily than the wild type [35]. BclA is a highly glycosylated protein and therefore I expected spores lacking BclA to clump faster. I asked whether lack of BclA on the spore surface would cause spores to clump. To answer this question I performed a clumping assay where the change in OD over-time should reflected the rate at which the spores clump. I performed the assay with *B. anthracis* wild type and *bclA* mutant spores at the same time. Wild type served as my positive control of how fast spores clump when BclA is present. There were no striking differences in the OD change (Figure 36). Wild type spores clumped at almost identical rate and the final OD was 91% and 93% of the starting one meaning within 90 minutes of the assay only 9% and 7% of spores clumped. The final OD reading for *bclA* mutant spores was 98% and 87% meaning within 90 minutes 2% and 13% of spores clumped, respectively. These data suggest spores of both strains clump at a similar rate.

I asked whether there is a difference in clumping rate in *B. subtilis* BSn5 and *B. atrophaeus* 1942. I performed the same assay mentioned above. The final OD reading was 95% of the starting OD for *B. atrophaeus* meaning only 5% clump in 90 minutes (Figure 37). The final OD reading was 96% of the starting OD for *B. subtilis* meaning only 4% clump in 90 minutes. I concluded spores of *B. subtilis* BSn5 and *B. atrophaeus* 1942 clump at a similar rate. The results were similar to wild type and *bclA* spores of *B. anthracis*.

Conclusion Based on my results, I conclude spores of *B. anthracis* with and without BclA on their surface are similarly hydrophobic. Moreover, the hydrophobicity does not change with heat treatment. My findings suggest the lack of BclA does not affect the spore surface hydrophobicity. My results are different from the published results. The discrepancy could come from a different spore preparation. Brahmbhatt et al. used modified G medium to make the spores and purified them through a Hypaque-76 gradient prior to the experiment. I used DMS medium and my spores did not require prior purification. An alternative explanation for the data inconsistency is my experiment failed.

I showed the spores of *B. anthracis* with or without BclA clump at the same rate. Because the clumping effect was shown to be due to the lack of sugars on the spore surface of *B. subtilis*, I conclude the sugars present on BclA are not sufficient to change the clumping dynamics of spores lacking BclA. Alternately, spores lacking BclA might still be glycosylated to the same extent as the wild type by some alternative mechanism.

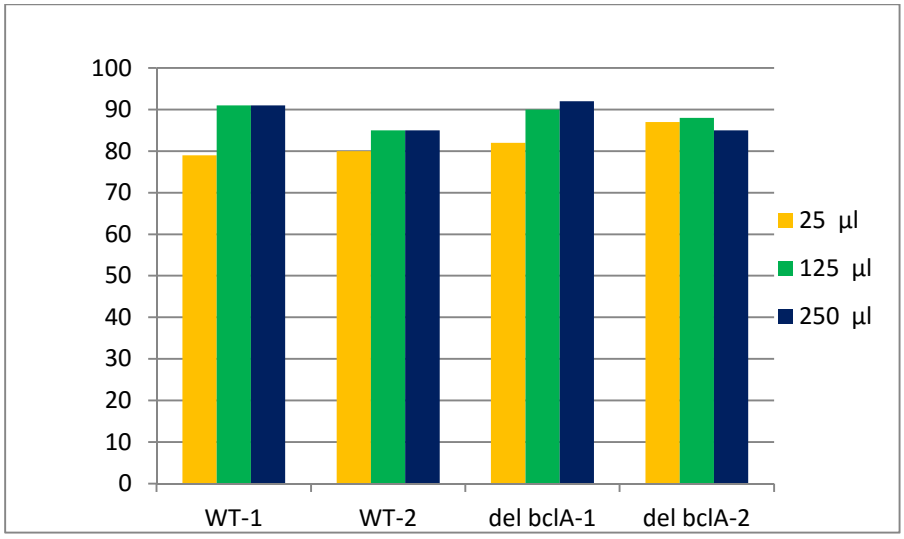

Figure 33. Measuring hydrophobicity of *B. anthracis* spores. Spores with or without BclA were mixed with various amount of hexadecane and the change in the OD_{440} of the spore suspension was used as readout for the hydrophobicity.

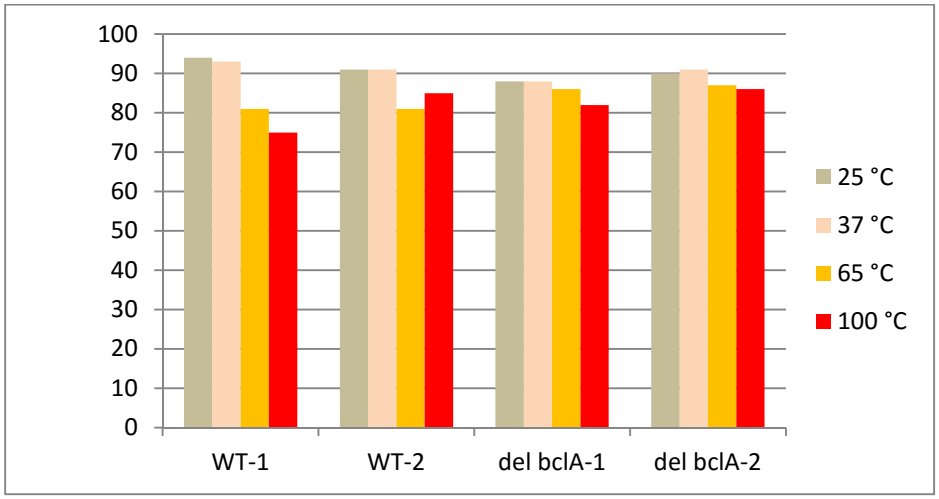

Figure 34. Measuring hydrophobicity of *B. anthracis* spores. Spores with or without BclA were pretreated with heat and incubated overnight. Next day, spores were mixed with the hexadecane and the change in the OD_{440} of the spore suspension was used as readout for the hydrophobicity.

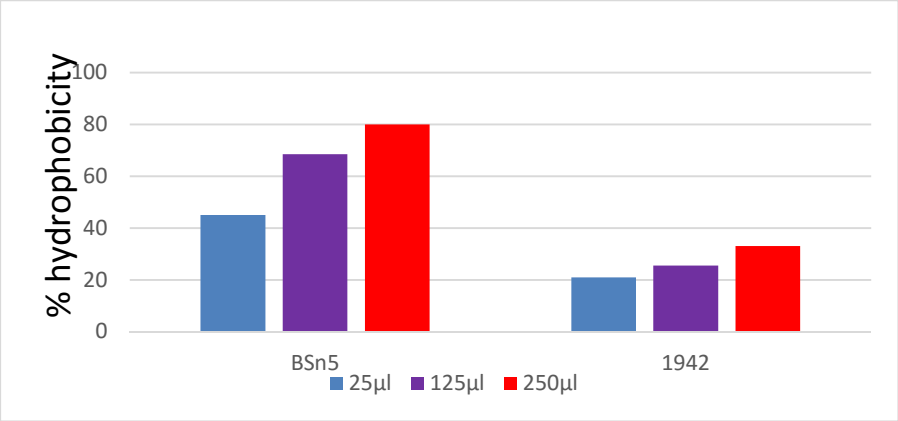

Figure 35. Measuring hydrophobicity of *B. subtilis* BSn5 and *B. atrophaeus* 1942 spores. Spores were mixed with various amount of hexadecane and the change in the OD_{440} of the spore suspension was used as readout for the hydrophobicity.

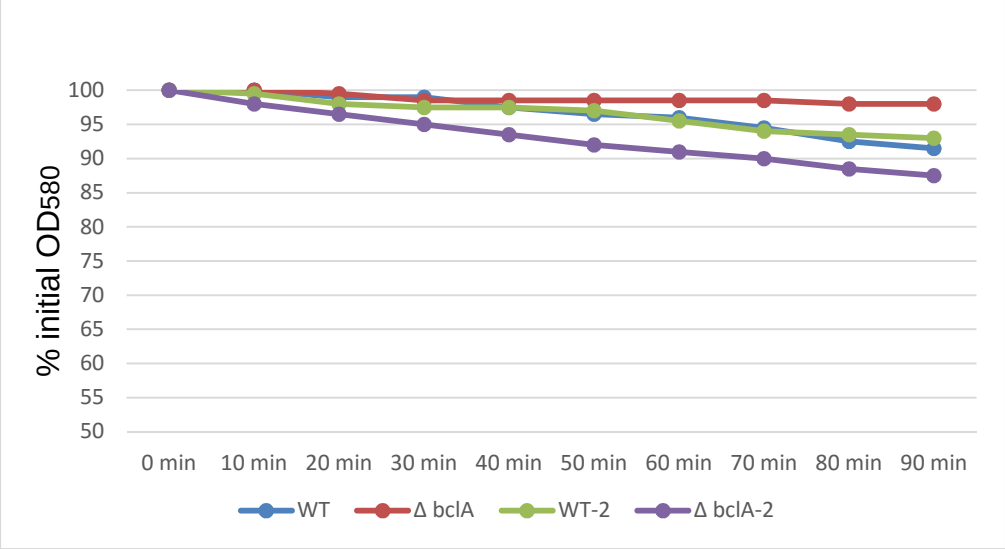

Figure 36. Clumping assay of WT and *bclA* spores of *B. anthracis*. Spores were vigorously shaken and the change in OD was measured every 10 minutes.

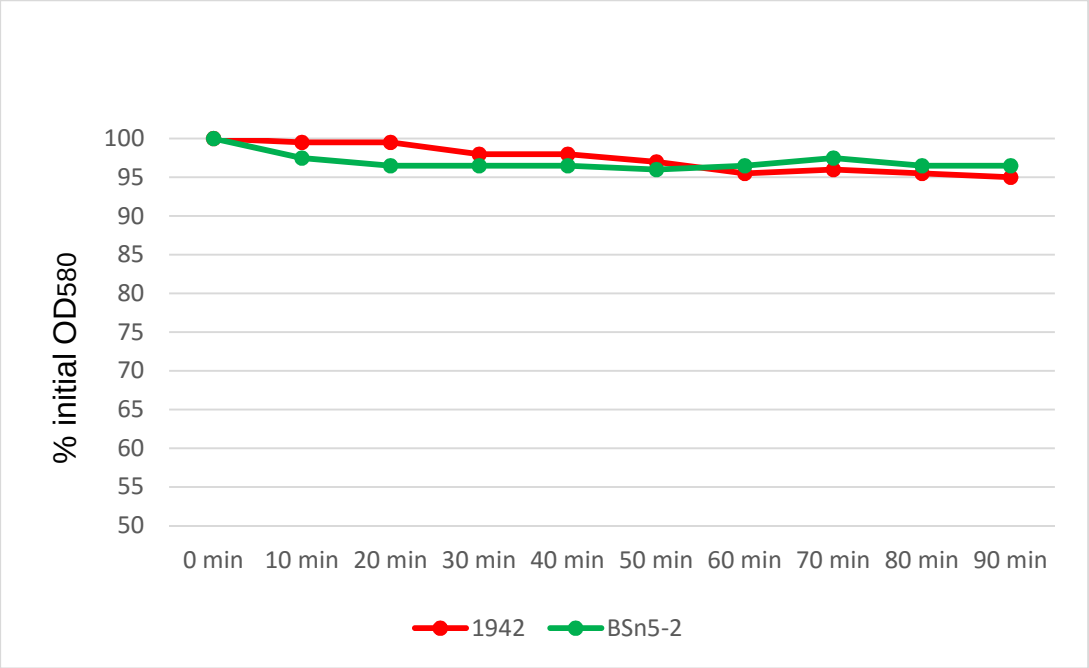

Figure 37. Clumping assay of *B. subtilis* BSn5 and *B. atrophaeus* 1942 spores. Spores were vigorously shaken and the change in OD was measured every 10 minutes.

CHAPTER FOUR

DISCUSSION

Until recently, BclA was only thought to be a protein present on the spore surface of *B. anthracis* and its close relatives. In 2014, Pizarro-Guajardo et al. published a surprising finding that BclA is a spore surface protein in *C. difficile* [36]. This was a remarkable discovery as the spores of *C. difficile* do not possess exosporia. Furthermore, *C. difficile* is not remotely related to the *B. cereus* group as it belongs to a different class of bacteria, called Clostridia. This discovery led to our hypothesis that BclA is a common spore surface protein.

bclA Gene

My genomic analysis revealed that *bclA*-like genes are widely present in the genomes of spore-forming bacteria. *bclA*-like genes are found in genomes of species from *Bacillaceae, Paenibacillaceae,* and *Clostridiaceae.* Surprisingly, there are multiple *bclA*-like genes that can be found in each species and the number of genes varies even across the strains. We discovered the *B. subtilis* PY79 strain that is routinely used in the lab, does not have *bclA*-like genes in its genome. Previously, we tried to find BclA on the spore surface of that species and we failed. It would be interesting to determine which strains of spore-forming bacteria do have *bclA* gene and which do not. An ideal approach would be to focus on a group of fully sequenced species and strains to establish which strains have genes homologous to *bclA.* By building a phylogenetic tree and establishing which strains do have *bclA* we could trace the loss of the gene. It would be important to

know where the strains have been isolated from to pinpoint an environmental factor that drives the the preservation of *bclA*.

BclA Domain Organization

BclA in *B. anthracis* has three domains: N-terminal domain, collagen-like region and C-terminal domain. Based on the amino acid analysis, I conclude that the domains of BclA are preserved in other spore-forming bacteria. The greatest variation is found in the N-terminal domain of BclA. It may be as short as 2 amino acids and as long as 240 amino acids. We know from the *C. difficile* study, that the N-terminal domain as short as 5 amino acids is sufficient to localize the protein to the spore surface [36]. The collagen-like region varies in its length but the unifying feature is an abundance of GXT triplets. This feature might be the hallmark of spore-associated collagen-like proteins. There are other collagen-like proteins in nature; some of them are present in human pathogens. Streptococcal collagen-like (Scl) protein is present at the surface of *Streptococcus pyogenes* cells [37]. Lack of Scl on the cell surface decreases the adherence of the cells to human fibroblast. Additionally, when *scl S. pyogenes* is injected subcutaneously into mice, the virulence is attenuated. Another collagen-like protein, Lcl, is found in *Legionella pneumophilia* [38]. Cell lacking Lcl are deficient in autoaggregation. Interestingly, both Scl and Lcl are rich in GXX triples but rarely is there a GXT triple (BLAST). More work is needed to establish if the difference in the collagen triplets composition is a good way of distinguishing between bacterial protein and spore surface collagen-like proteins. Finally, the C-terminal domain of BclA ranges from 130 to 160 amino acids. Based on BLAST analysis, the sequence of C-terminal domain is mostly preserved in *B. cereus* group, and there is some amino acid preservation in other species.

It would be interesting to solve the three dimensional structure of the C-terminal domain for some nonpathogenic spore-forming bacteria. We know that in BclA of *B. anthracis*, the C-terminal domain resembles the C1q. Do the C-terminal domains of nonpathogenic species have the C1q resemblance or are they completely different in their structure? If the domains are similar, could they also activate the classical complement pathway as in the case of *B. anthracis BclA*?

Localization of BclA

BclA in *B. anthracis* is present on the entire surface of the spore. In our study, BclA distribution was not uniform. I learned BclA is only localized at the poles of *B. atrophaeus* 1942, *B. methylotrophicus* FZB42 and *P. chitinolyticus* NBRC 15660. The simple explanation is BclA is only present at the poles of spores of these three species. Alternatively, BclA might actually be present on the entire spore surface but might be occluded by other proteins. If indeed BclA is only present at the poles of the protein, what is the significance of this localization? Is this localization dictated by the environmental pressure?

BclA was localized on the entire surface of *Br. laterosporus* 9141, *B. amyloliquefaciens* NRRL BD-553 and *B. subtilis* BSn5. Interestingly, only a subset of spores of *B. subtilis* had BclA on their spore surface. I wonder whether *bclA*-like genes in *B. subtilis* BSn5 are flanked by phage-like regions could explain this variation in protein distribution. We know from *B. subtilis* literature that there are prophage regions in the genome that excise during sporulation. An example of this phenomenon is the excision of the *skin* element that disrupts the *sigK* gene which codes for sigma K [39]. Another example is the SP beta prophage which excises from a sporulating cells and reconstitutes

a polysaccharide synthesis gene, spsM [40]. Further studies are needed to evaluate whether the phage elements are responsible for BclA variation in *B. subtilis* BSn5.

BclA Function

The function of BclA is still unknown. There have been studies suggesting BclA has a role in the infection process as it interacts with the complement system [14]. This prompts the question of why is *B. anthracis* BclA capable of interacting with the immune system. Precisely, how is BclA of *B. anthracis* different from BclA of other nonpathogenic species? If BclA has a role in the soil environment, is it possible *B. anthracis* BclA is multifunctional where is has both a purpose in the soil as well as in host and microbe interaction? There are many questions that need to be asked and comprehensive studies to be performed before one can say what the function of BclA is.

Mutations in Wild Spore-Forming Species

In this study, I attempted to create *bclA* mutants in *B. subtilis* BSn5 and *B. atrophaeus* 1942. There are multiple ways one can introduce exogenous DNA into *B. subtilis*: by natural transformation, phage transduction, electroporation and protoplast. In this study, I mainly focused on electroporation as it is easy to perform. Although I explored multiple protocols, none of them seemed to be suitable for the strains I wanted to mutate. Multiple studies have been done to address manipulation of wild spore-forming bacteria. Most of them point out that each species might require customized protocols for transformation. Kolek et al. showed that well established protocols for *C. difficile* and *C. cellulolyticum* were not suitable to transform *C. pasteurianum* [41]. The group recognized the major obstacle of their study was the methylation status of plasmids they used. Only after the plasmid was passed through an *E.coli* strain deficient in *dam* and *dcm* genes, was the

group able to transform *C. pasteurianum* with a modified protocol. The methylation issue was addressed in my study. I passed pRP1028 through a *dam dcm* deficient *E.coli* but this did not allow me to transform either *B. subtilis* or *B. atrophaeus*. Interestingly, what my study revealed is pRP1028 might not be a suitable plasmid to transform *B. subtilis*-like species, at least under our lab conditions. This conclusion was made based on the ability to introduce two different plasmids (pIMAY and PEO-3) into *B. subtilis* PY79, our lab strain, but not pRO1028. I demonstrated the electroporation protocol routinely used is suitable for *B. subtilis* PY79 but not for *B. subtilis* BSn5 or *B. atrophaeus* 1942. The study by Zhang et al. explores multiple factors affecting electroporation efficiency. The group shows that the pretreatment of cells with cell-wall weakening agents and an increase in electric field increase the electroporation efficiency in *B. subtilis* ZK. In my study these two additional steps were not sufficient to allow plasmid uptake. I conclude that *B. subtilis* BSn5 as well as *B. atrophaeus* 1942 need customized protocols for an efficient transformation. Although in this study, I only tried to mutate the strains through conjugation and electroporation, there are other methods that were not explored. An induction of competence is another way to mutate *B. subtilis* [42]. Recently, Zhao et al. were able to mutate *B. amyloliquefaciens* by adopting a *B. subtilis* protocol [43]. Because *B. amyloliquefaciens* is closely related to *B. subtilis*, it is plausible to think, the same protocol could work for *B. subtilis* and *B. atrophaeus.* An alternative way to mutate wild isolates is to use a newer technique that recently emerged. CRISPR/Cas9 method might be the answer to manipulating wild species that do not subject to lab protocols. CRISPR-based genome engineering has been mostly done in eukaryotic cells. Recently, the CRISP/Cas9 system has been successfully used to mutate genes in *Streptococcus*

pneumoniae and *E. coli* [44]. In the future, the CRISPR system should be tried on the wild spore-forming organisms in other to mutate *bclA* gene. This mutation would prove what is detected on the spore surface of various species is indeed BclA and it would also facilitate studies addressing the question of BclA's function.

CHAPTER FIVE

P5303

Introduction

Anthrax is a disease caused by a spore-forming bacterium, *B. anthracis*. Anthrax can be acquired by three routes: by inhalation, through a skin lesion, or by ingestion. Ingestion is the natural route of *B. anthracis* infection in grazing animals [46]. Humans can be exposed to *B. anthracis* usually by working with infected animals. Cutaneous infection is predominant and is often a result of open skin contact with the sick animal. Gastrointestinal infection can result from eating contaminated meat. Inhalational anthrax is a result of breathing in spores usually by working with wool, hides or hair coming from the sick animal. Inhalation of *B. anthracis* spores may result in the most severe disease outcome [47]. The pulmonary and gastrointestinal anthrax are of great concern as the disease symptoms are nonspecific. When anthrax is caught in its early stages, it is treatable. In some cases it is difficult to diagnose anthrax; it is important to focus on preventive measures at least for a population that is at highest risk of *B. anthracis* exposure. One preventive measure used is a vaccine. There are indeed two anthrax vaccines approved for humans: human anthrax vaccine adsorbed (AVA) and human anthrax vaccine precipitated (AVP). The protection of AVA vaccine comes mainly from a single protein, protective antigen (PA).

PA is the cell binding component of both edema and lethal toxins [48]. PA protein is produced after a spore germinates which means the vaccine targets a later stage of infection. More recent approaches target an early stage of anthrax infection – the acquisition of spores. Therefore the current vaccines studies are focused on adding spore surface proteins to the vaccines [49, 50]. There are 30 proteins that are localized to the exosporium of *B. anthracis* [49]. One of the proteins is P5303. P5303 has already been used in vaccine studies. Cybulski et al. showed anti spore antiserum recognizes a recombinant P5303 and therefore it is believed P5303 is a spore protein. However, there is no published study that shows P5303 is a spore surface protein. Moreover, we do not know the localization of the protein within the exosporium. As P5303 is a promising candidate for a vaccine, it is important to establish where the protein is present, what effects does it have on the spore morphology, and what is its function.

Methods

Spore Preparation Spores were prepared by exhaustion in Difco Sporulation Medium (DSM) [28]. A single colony was picked from a Luria Broth (LB) plate and suspended in 200 µl of DSM. The bacterial suspension was spread onto a DSM plate and incubated at 37 °C for 8 to 9 hours. After the incubation period, the lawn was collected by suspension in 5 ml of DSM. 1 ml of the lawn was transferred into the 35 ml of DSM in a 250 ml flask. The flask was shaken at 225 rpm at 37 °C overnight. The next day, 5 µl of the sample was placed onto a glass slide and the sample was checked for the presence of spores under the phase-contrast microscope. The spores were spun down and the

pellets were washed 3 times with MQ water. The spores were stored in water in a 50 ml tube at 4°C.

Creation of *B. anthracis* Mutant *p5303* was deleted by following a protocol published by Warrens et al. [51]. Upstream and downstream regions of *p5303* gene were amplified (Primers 1and 2 for upstream region, Primers 3 and 4 for downstream region) and fused by splicing using overlap extension PCR. The fused fragment was cut with HindIII and BamHI and ligated into pGEM-T plasmid. pGEM-T was transformed into *E. coli* DH5α. pGEMT was isolated with a kit (Thermofisher). The plasmid was digested with HindIII and BamHI. DNA was separated on a 1% agarose and the upstream-downstream region was cut out and purified. The fragment was ligated into a pRP1028 plasmid. pRP1028 was then transformed into *E. coli* DH5α.

Plasmids were introduced into bacterial cells by conjugation [29]. On day 1, the donor strain, the recipient strain and the helper strain were struck out onto LB +100 µg/ml spectinomycin, BHI and LB +100 µg/ml ampicillin, respectively. The donor strain contained the allelic exchange construct cloned into pRP1028. The recipient strain was *B. anthracis*. Plates were incubated at 37°C overnight. On day 2, each strain was scraped off the plate, plated onto a BHI plate and mixed together. The BHI plate was incubated at the room temperature. On day 3, the mixed strains were streaked out onto a selective plate that contained BHI + 250 µg/ml spectinomycin + 60 units/ml polymixin B. The plate was left at the room temperature for 24 hours. *B. anthracis* colonies were picked and restreaked for isolation two times. On day 6, the helper strain and strain containing facilitator plasmid were struck out onto LB-amp and LB-kan (20 µg/ml kanamycin), respectively. On day 7, three strains were mixed together on a BHI plate and incubated at

37°C for 8 hours. Later that day, the mixture was scraped off and streaked out on a BHI-KanPmx plate (BHI + 20 µg/ml kanamycin + 60 units/ml polymixin B). The plate was incubated at 37°C overnight. On day 8, *B. anthracis* colonies were picked and restreaked for isolation. Colonies were patched onto the BHI-SpecPmx and the BHI-KanPmx plates and incubated at 37°C overnight. Spectinomycin sensitive colonies were picked and PCR was done to confirm the gene deletion.

Spore Surface Protein Extraction and Western Blot 50 ml of spores was pelleted down and resuspended in 10 ml of water. 5 µl of spore suspension was mixed with 495 µl of water and OD_{600} was recorded. The formula 0.037/ OD_{600} was used to estimate the amount of spore suspension needed for each well. The appropriate amount of spore suspension was pelleted for 5 min/6000 rpm and the supernatant was removed. 13 µl of Laemmli buffer (0.63 ml 1M Tris pH 6.8, 1 ml 100% glycerol, 2ml 10% SDS and 6.37 ml MQ water) and 1.5 µl of 1M DTT were added to the pellets. Samples were vigorously shaken for 1 minute and spun down for 2 seconds. Samples were boiled for 5 minutes at 100 °C. Samples were vigorously shaken again for 45 seconds and boiled at 100°C for an additional 5 minutes. Samples were vigorously shaken for 30 seconds and spun down for 5 minutes at 13000 rpm. The supernatants were collected and the protein concentration was measured with Nanodrop. Proteins were resolved on 15% SDS-PAGE (10% 29:1 acrylamide: N, N'-methylene-bis-acrylamide, 375 mM Tris pH 8.6, 0.1% SDS), and transferred to polyvinylidene difluoride membranes or nitrocellulose membranes [29]. The membranes were incubated with 2% bovine serum albumin (BSA), washed three times with Tris-buffered saline with Tween 20 (1x TBST: 1.21 g Tris, 8.76 g NaCl, 0.05% Tween 20 in 1L of water) then incubated with anti-P5303 polyclonal antibody

(BEI Resources). Goat anti rabbit IgG was used as the secondary antibody (Sigma).

Immunofluorescence Microscopy 10 µl of culture was placed into each well of a multiwell slide. The slides were pretreated with 0.01% (wt/vol) poly-L-lysine, washed twice with water and air-dried. 10 µl of phosphate-buffered saline (PBS) was then placed into each well and was replaced with 2% (wt/vol) BSA in phosphate-buffered saline prior to the addition of primary antibody. Anti-P5303 antibody was used at 1:300 dilutions. Secondary antibody was used at a 1:300 dilution. Alexa Fluor 594 goat anti rabbit IgG was used as the secondary antibody (Molecular probes, Life Technologies ™).

Electron Microscopy A pellet from 1 ml of the *p5303* mutant spore stock was prepared. The pellet was mixed with 840 µl of 10mM phosphate buffer (pH 7) and 160 µl of 8% gluteraldehyde. The sample was incubated overnight at 4 °C. The next day, the sample was spun down and washed in 1 ml of 0.5 M NH_4Cl twice. The pellet was mixed with 300 µl 10mM phosphate buffer (pH 7) and 100 µl of 4% osmium tetraoxide. The pellet was incubated overnight at 4 °C. The next day, the pellet was washed twice with 1ml of 0.5 M NH_4Cl and resuspended in 100 µl of water. 300 µl of 3% melted agarose was mixed with the spore pellet and transferred onto an agarose cushion. The sample was spun down for 30 seconds. After the sample solidified, the pellet was cut with a razor blade into small pieces and placed into a scintillation vial. The sample was subjected to dehydration steps in 4 ml of 30%, 50%, 70% and 100% ethanol with rocking for 1 hour in-between. Next, the samples were mixed with 1:1 mixture of 100% ethanol and resin, followed by 100% resin. The sample in resin was left rocking overnight at room temperature. The next day, the samples were fished out and placed into a resin mold. Fresh resin was poured onto the sample and the sample was left baking overnight.

Species/strains	Genotype or description	Reference or source
B. anthracis 34F2	Wild type	Laboratory collection
B. anthracis 34F2	pXO1+ pXO2- *bclA*::kan	Laboratory collection
B. anthracis 34F2	*ΔbclA*	Laboratory collection
E. coli DH5α	Cloning host	Laboratory collection
E. coli SS4332	Facilitator strain	Stibitz and Carbonetti 1994
E. coli SS1827	Helper strain	Stibitz and Carbonetti 1994
Plasmids	**Description**	**Reference**
pGEM-T	Cloning vector	Promega
pRP1028	Temperature sensitive vector in Gram positive bacteria, $Spec^R$	Plaut & Stibitz 2015

Table 7. List of strains and plasmids used in P5303 study.

Primer name	Sequence (5' -3')	#
BAS5303-Up-FW-HindIII	ttt ttt AAG CTT TAC AAA ACC ACC CTA GAC C	1
BAS5303-Up-Rev	TAG AGA AAA GAA CCT AAA TAT CAG ACC TTT CTA ATT TAA TAT G	2
BAS5303-Dw-Fw	AAG GTC TGA TAT TTA GGT TCT TTT CTC TAT TCT CAA	3
BAS5303-Dw-Rev-BamHI	ttt ttt GGA TCC CGT ATG GAC AAA CAA AAT TAA	4
BAS5303-Up-Rev-Seq	GTT ATA TAA ATT GAG AAT AGA GAA AAG AAC CT	5
BAS5303-Dw-Fw-Seq	GTA CAA TCA TAT TAA ATT AGA AAG GTC TGA TAT TT	6
BAS5303-Rev-PCR-Screeen	TGA CAA ATC CGT ATG GAC AAA C	7
BAS-Fw-PCR-Screen	CCA TAT AAA ATA AAG CAA TAA TTG CTA CC	8

Table 8. List of primers used in P5303 study.

Results

Previously, a Western blot analysis on wild type spore surface proteins has been performed to establish the electrophoretic pattern on P5303 migration. Multiple molecular weight species were detected with the polyclonal antibody. Having a clean mutation in the *p5303* gene, enabled us to exclude the non-specific bands from the genuine P5303 signal via Western blot. To determine which of the proteins detected by the polyclonal antibody is P5303 I performed a Western blot both with wild type *B. anthracis* and *p5303* strain. Three molecular species were detected in the *p5303* mutant (Figure 38). I concluded those particular bands were non specific bands. A single band migrating at <15kDa was absent from *p5303* mutant. I believe this is the monomeric form of P5303 protein. Additionally, high molecular weight species were only detected in the wild type spores. This suggests P5303 is present in protein complexes that are resistant to DTT-SDS treatment. Next, I wanted to know whether P5303 is a spore surface protein. To find out whether P5303 is present on the surface of *B. anthracis* spores, I performed IFM. Anti-P5303 polyclonal antibody did not detect proteins on the surface of wild type *B. anthracis* spores (Figure 39). As expected, the antibody did not bind to any proteins on the surface of *B. anthracis* lacking P5303. Cybulski et al. suggested P5303 is a basal layer protein located beneath BclA [49]. I questioned whether BclA occludes the epitopes to which our antibody binds and hence we cannot detect P5303 on the wild type spores. I performed IFM on spores that lack BclA on their surface. Interestingly, I detected fluorescence in *bclA* mutant spores (Figure 40). The ring of fluorescence surrounded the entire spore. This suggests P5303 is indeed a spore surface protein. However, the protein is undetectable in wild type spores possibly due to BclA presence. This result could also

suggest that in absence of BclA the proteins in the exosporium are arranged differently leaving P5303 accessible to the antibodies. Next, I wanted to know whether the lack of P5303 has an effect on the spore structure. To answer this question, I performed EM. Spores of P5303 looked indistinguishable from the spores of wild type *B. anthracis* (Figure 41). This suggests P5303 does not have a role in the assembly of *B. anthracis* spore that can be easily detected.

Conclusion and Future Direction

P5303 protein is an anthrax vaccine candidate. Previously, it has been shown P5303 is a spore protein to which anti spore antiserum reacts. At that time, we did not know exactly where the protein resides or what its role is. Additionally, we still do not understand why the addition of this protein to the vaccine preparation makes the vaccine work better. In this study, I was able to show P5303 can be extracted from the spore both as a monomer as well as in a high molecular complex that resist the DTT-SDS treatment. Moreover, my IFM study shown P5303 is indeed a spore surface protein. Its detection is inhibited by the presence of BclA on the spore surface. In spores lacking BclA, we can detect P5303 on the entire spore surface. We can conclude from this observation at least, P5303 is not BclA dependent. What I was able to demonstrate is lack of P5303 does not alter the spore structure. If it does, it is not detectable under EM. Further studies should be done on the *p5303* mutant. One of the questions that should be addressed is when does P5303 appears on the spore surface? This could be answered by a time course experiment. Additionally, we could investigate whether lack of P5303 on the spore surface affects spore germination. We should also explore whether lack of P5303 makes the spore less resistant. This could be addressed by performing toluene sensitivity assay

and lysozyme sensitivity assay.

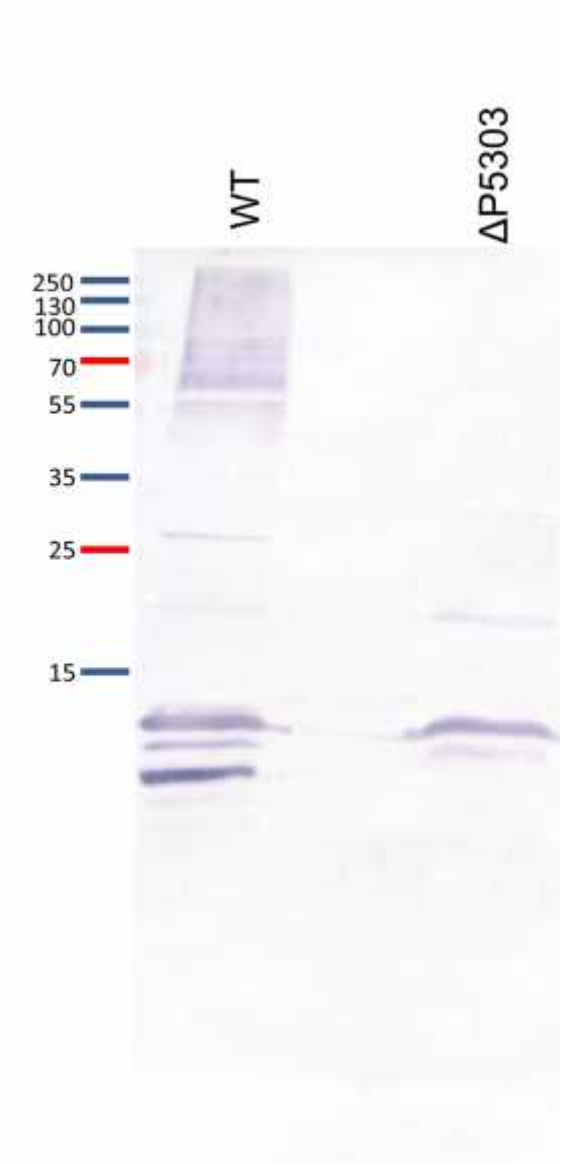

Figure 38. Western blot analysis of spore proteins from wild type *B. anthracis* and *p5303* mutant spores. Proteins were separated on 15% polyacrylamide gel and transferred to a membrane. Anti-P5303 polyclonal antibody was used to detect P5303. Antibody concentration: 1:5000. The size of markers is given in kDa

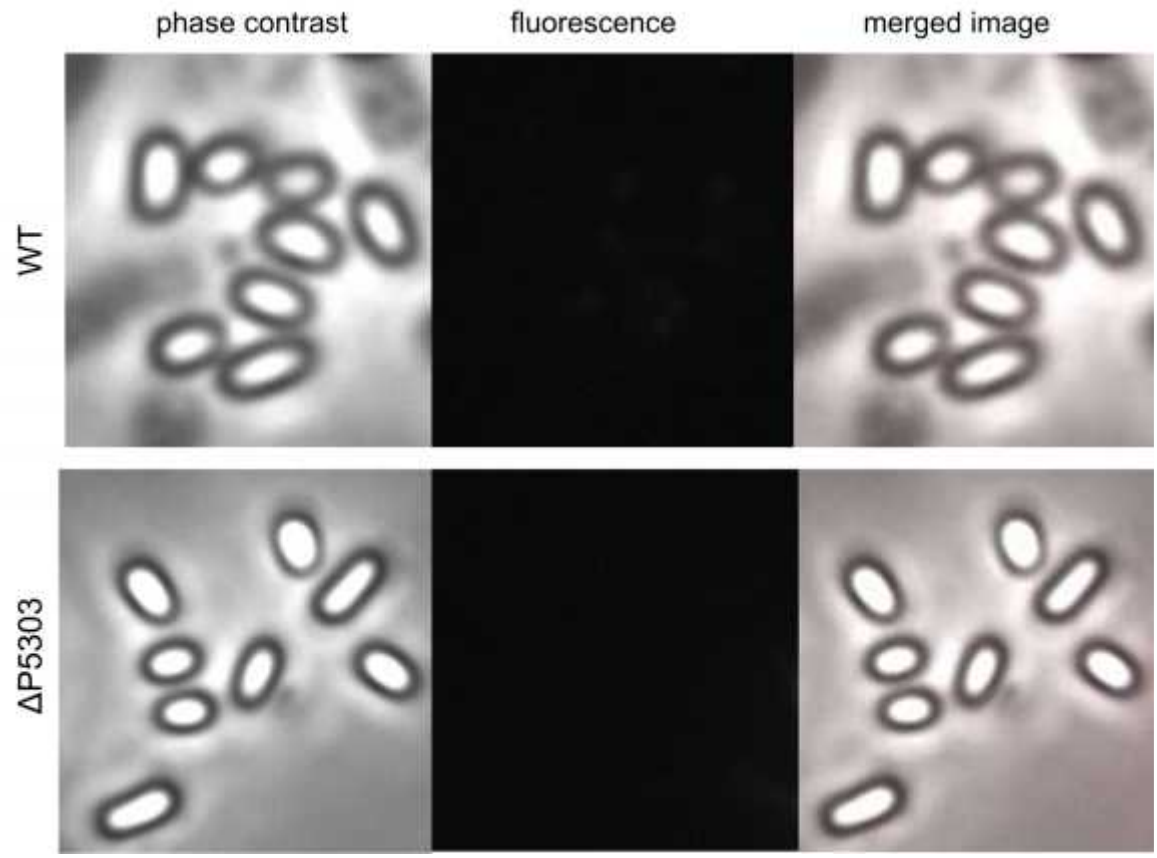

Figure 39. IFM analysis of wild type *B. anthracis* and *p5303* mutant spores. Spores were fixed onto a slide and treated with anti-P5303 antibody. Antibody concentration: 1:300

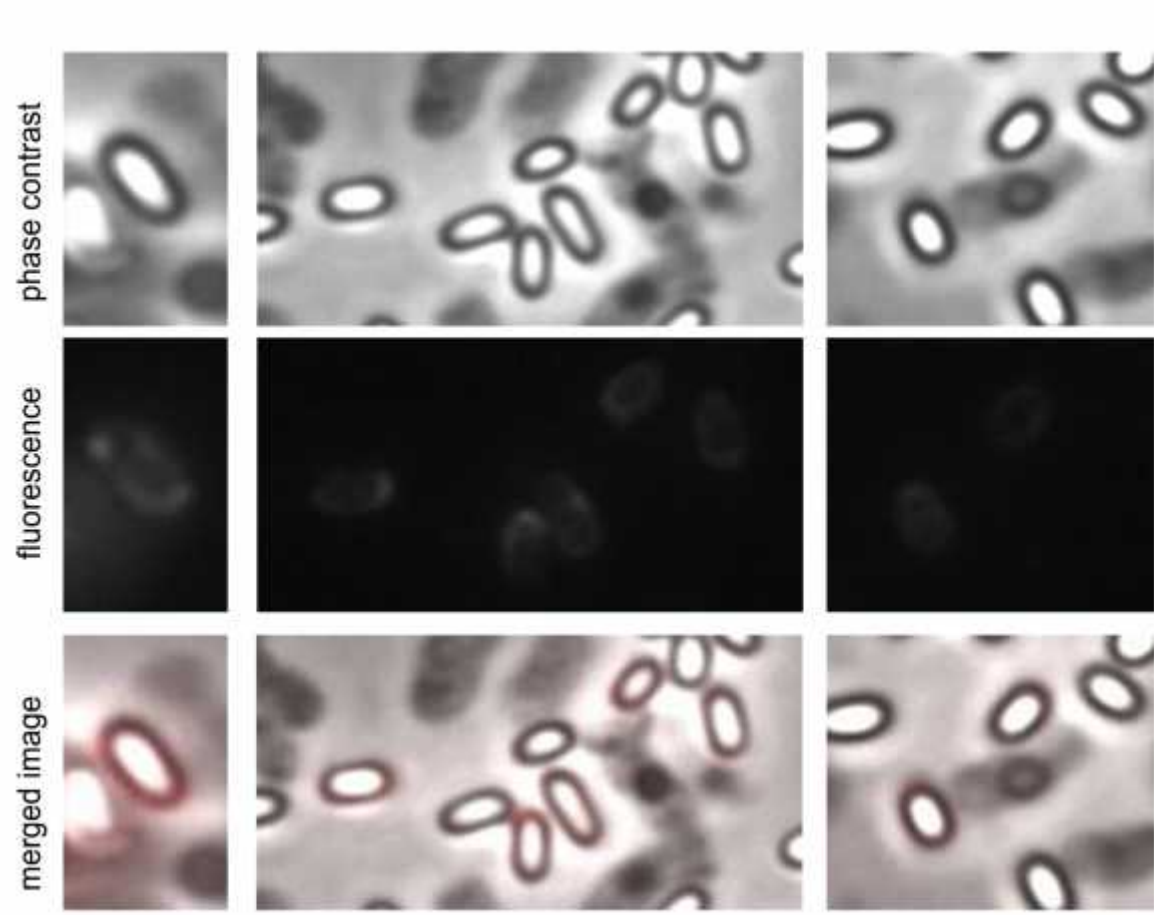

Figure 40. IFM analysis of *bclA B. anthracis* spores. Spores were fixed onto a slide and treated with anti-P5303 antibody. Antibody concentration: 1:300

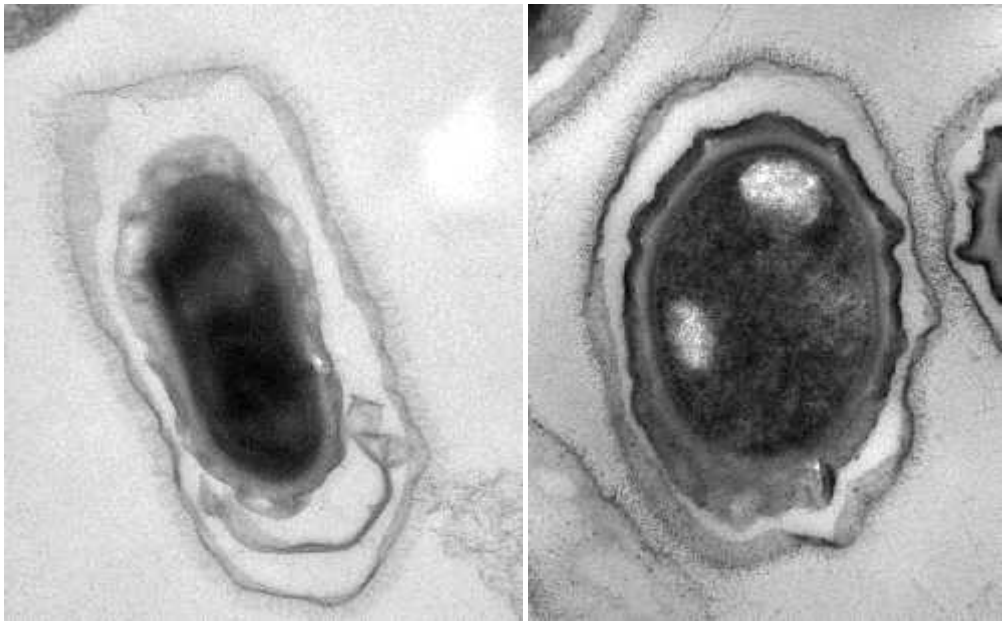

Figure 41. TEM images of wild type and *p5303 B. anthracis* spores. Left: *p5303* mutant spore. Right: wild type spore (image provided by T. Boone)

BIBLIOGRAPHY AND RESOURCES

1. Galperin, M.Y, *Genome diversity of spore-forming Firmicutes.* Microbiol Spect. 2013. **1**(2)

2. Driks A, *The Bacillus anthracis spore*. Mol Aspects Med. 2009. **30**(6): p 368-373.

3. McKenney, PT., A. Driks, and P. Eichenberger, *The Bacillus subtilis endospores: assembly and functions of the multilayered coat*. Nat Rev Microbiol. 2013. **11**(1): p 33-44.

4. Gerhardt, P. and E. Ribi, *Ultrastructure of the exosporium enveloping spores of Bacillus cereus.* J. Bacteriol. 1964. 88: p 1774-1789.

5. Timmusk, S., V. Paalme, T. Pavlicek, J. Bergquist, A. Vangala, T. Danilas, and E. Nevo, *Bacterial distribution in the rhizosphere of wild barley under contrasting microclimates.* PLoS One. 2001. **6**(30)

6. Barbosa, T.M., C.R. Serra, R.M. La Ragione, M.J. Woodward, and A. O. Henriques, *Screening for Bacillus isolates in the broiler gastrointestinal tract*. Appl EnviroN Microbiol. 2005. **71**(2): P 968–978.

7. Buehner, K.P., S. Anand, G.D. Djira, and A. Garcia, *Prevalence of thermoduric bacteria and spores on 10 Midwest dairy farms.* J Diary Sci. 2014. **97**(11): P 6777-6784.

8. Sylvestre, P., E. Couture-Tosi, and M. Mock, *A collagen-like surface glycoprotein is a structural component of the Bacillus anthracis exosporium.* Mol Microb. 2002. **45**(1): p 169-178.

9. Steichen, C., P. Chen, J.F. Kearney, and C.L. Jr. Turnbough, *Identification of the immunodominant protein and other proteins of the Bacillus anthracis exosporium*. J Bacteriol. 2003. **185**(6): p 1903-1910.

10. Tan, L. and C.L. Jr Turnbough, *Sequence motifs and proteolytic cleavage of the collagen-like glycoprotein BclA required for its attachment to the exosporium of Bacillus anthracis.* J Bacteriol. 2010. **192**(50): p 1259-1268.

11. Sylvestre P., E. Couture-Tosi, and M. Mock, *Polymorphism in the collagen-like region of the Bacillus anthracis BclA protein leads to variation in exosporium filament length.* J Bacteriol 2003. **185**(5): p. 1555-1563.

12. Réty, S., S. Salamitou, I. Garcia-Verdugo, D.J. Hulmes, F. Le Hégarat, R. Chaby, and A. Lewit-Bentley, *The crystal structure of the Bacillus anthracis spore surface protein BclA shows remarkable similarity to mammalian proteins*. J Biol Chem. 2005. **280**(52): p 43073-43078.

13. Bozue, J., C.K. Cote, K.L. Moody, and S.L. Welkos, *Fully virulent Bacillus anthracis does not require the immunodominant protein BclA for pathogenesis*. Infect Immun. 2007. **75**(1): p 508-511.

14. Oliva C., C.L. Turnbough, and J.F. Kearney, *CD14-Mac-1 interactions in Bacillus anthracis spore internalization by macrophages.* PNAS 2009. **106**(33): p 13957-13962

15. Gu, C., S.A. Jenkins, Q. Xue, and Y. Xu, *Activation of the classical complement pathway by Bacillus anthracis is the primary mechanism for spore phagocytosis and involves the spore surface protein BclA*. J Immun. 2012. **188**(9): p 4421-4431.

16. Brahmbhatt, TN., B.K. Janes, S. Stibitz, S.C. Darnell, P. Sanz, S.B. Rasmussen, and A.D. O'Brien, *Bacillus anthracis exosporium protein BclA affects spore germination, interaction with extracellular matrix proteins, and hydrophobicity.* Infect Immun. 2007. **75**(11): p 5233-5239.

17. Lequette,Y., E. Gare´naux, T. Tauveron, S. Dumez, S. Perchat, C. Slomianny, D. Lereclus, Y. Gue´rardel, and C. Faille, *Role played by exosporium glycoproteins in the surface properties of Bacillus cereus spores and in their adhesion to stainless steel.* Appl Environ Microbiol. 2011. **77**(14): p 4905-4911.

18. Chen, G., A. Driks, K. Tawfiq, M. Mallozzi, and S. Patil, *Bacillus anthracis and Bacillus subtilis spore surface properties and transport.* Colloids and Surf B Biointerfaces. 2009. **76**(2010) p 512-518.

19. Matz, L.L., T.C. Beaman, and P. Gerhardt, *Chemical composition of exosporium from spores of Bacillus cereus.* J Bacteriol. 1970. **101**(1): p 196-201.

20. Redmond, C., L.W.J. Baillie, S. Hibbs, A.J.G. Moir, and A. Moir, *Identification of proteins in the exosporium of Bacillus anthracis.* Microbiol 2004:150, p 355–363

21.Henriques, A.O. and C.P. Jr. Moran, *Structure, assembly, and function of the spore surface layers.* Annu Rev Microbiol. 2007. 61: p 555-88.

22. Sylvestre, P., E. Couture-Tosi, and M. Mock, *Contribution of ExsFA and ExsFB proteins to the localization of BclA on the spore surface and to the stability of the bacillus anthracis exosporium.* J Bacteriol. 2005. **187**(15): p 5122-5128.

23. Steichen, C.T., J.F. Kearney, and C.L. Jr. Turnbough, *Characterization of the exosporium basal layer protein BxpB of Bacillus anthracis.* J Bacteriol. 2005. **187**(17): p 5868-5876.

24. Thompson, B.M., and G.C. Stewart, *Targeting of the BclA and BclB proteins to the Bacillus anthracis spore surface*. Mol Microbiol. 2008. **70**(2): p 421-434.

25. Kim, H., M. Hahn, P. Grabowski, D.C. McPherson, M.M. Otte, R. Wang, C.C. Ferguson, P. Eichenberger, and A. Driks, *The Bacillus subtilis spore coat protein interaction network.* Mol Microbiol. 2006. **59**(2): p 487-502.

26. McKenney, P.T., A. Driks, H. A. Eskandarian, P. Grabowski, J. Guberman, K.H. Wang, Z.G. Gitai, and P. Eichenberger, *A Distance-Weighted Interaction Map Reveals a Previously Uncharacterized Layer of the B. subtilis Spore Coat*. Current Biology. 2010. **20**(10): p 934–938.

27. Bhandari,V., N.Z. Ahmod, H.N. Shah, and R.S. Gupta, *Molecular signatures for Bacillus species: demarcation Molecular signatures for Bacillus species: demarcation of the Bacillus subtilis and Bacillus cereus clades in molecular terms and proposal to limit the placement of new species into the genus Bacillus.* Int J Syst Evol Microbiol. 2013. 63: p 2712–2726.

28. Lai, E.M., N.D. Phadke, M.T. Kachman, R. Giorno, S. Vazquez, J.A. Vazquez, J.R. Maddock, and A. Driks, *Proteomic analysis of the spore coats of Bacillus subtilis and Bacillus anthracis*. J Bacteriol. 2003. **185**(4): p 1443-1454.

29. Janes, B., and S. Stibitz, *Routine Markerless gene replacement in Bacillus anthracis*. Infect Immun. 2006. **74** (3): p 1949–1953.

30. Xue, G.P., J.S. Johnson, and B.P. Dalrymple, *High osmolarity improves the electro-transformation efficiency of the gram-positive bacteria Bacillus subtilis and Bacillus licheniformis.* J Microbiol Methods. 1999. 34: p 183-191.

31. Zhang, Z., Z.T. Ding, D. Shu, D. Luo, and H. Tan, *Development of an efficient electroporation method for iturin A-Producing Bacillus subtilis ZK*. Int J Mol Sci. 2015. 16: p 7334-7351.

32. Bauer, T., S. Little, A. Stoever, and A. Driks, *Functional Regions of the Bacillus subtilis Spore Coat Morphogenetic Protein CotE.* J Bacteriol. 1999. **181** (22): p 7043–7051.

33. Sambrook, J., E. F. Fritsch, and T. Maniatis. *Molecular cloning: a laboratory manual, 2nd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989.

34. Rosenberg, M., D. Gutnick, and E. Rosenberg, *Adherence of bacteria to hydrocarbons: a simple method for measuring cell surface hydrophobicity*. FEMS Microbiol. 1980, 9: p 29-33.

35. Cangiano, G., T. Sirec, C. Panarella, R. Isticato, L. Baccigalupi, M. De Felice, and E. Ricca, *The sps Gene Products Affect the Germination, Hydrophobicity, and Protein Adsorption of Bacillus subtilis Spores*. Appl Environ Microbiol. 2014. **80**(23): p 7293-302.

36. Pizarro-Guajardo, M., V. Olguín-Araneda, J. Barra-Carrasco, C. Brito-Silva, M.R. Sarker, and D. Paredes-Sabja, *Characterization of the collagen-like exosporium protein, BclA1, of Clostridium difficile spores*. Anaerobes. 2014. 25: p 18-30.

37. Lokomski, S.,K. Nakashima, I. Abdi, V.J. Cipriano, B.J**.** Shelvin, E.A. Graviss, and J.M. Musser, *Identification and characterization of a second extracellular collagen-like protein made by group A Streptococcus: control of production at the level of translation*. Infection and Immunity. 2001. **69**(3): p 1729-1738.

38. Abdel-Nour, M., C. Duncan, A. Prashar, C. Rao, C. Ginevra, S. Jarraud, D.E. Low, A.W. Ensminger, M.R. Terebiznik, and C. Guyard, *The Legionella pneumophila collagen-like protein mediates sedimentation, autoaggregation, and pathogen-phagocyte interactions.* Appl Environ Microbiol. 2014. **80**(4): p 1441-1454.

39. Popham, D.L., and P. Stragier, *Binding of the Bacillus subtilis spoIVCA product to the recombination sites of the element interrupting the sigma K-encoding gene.* PNAS. 1992. **89**(13): p 5991-5995

40. Abe, K., Y. Kawano, K. Iwamoto, K. Arai, Y. Maruyama, P. Eichenberger, and T. Sato, *Developmentally-regulated excision of the SP β prophage reconstitutes a gene required for spore envelope maturation in Bacillus subtilis*. PloS One. 2014. **10**(10).

41. Kolek, J., K. Sedlar, I. Provaznik, and P. Patakova, *Dam and Dcm methylations prevent gene transfer into Clostridium pasteurianum NRRL B-598: development of methods for electrotransformation, conjugation, and sonoporation*. Biotechnol Biofuels. 2016. **9**(4).

42. Kunst, F. and G. Rapoport, *Salt stress is an environmental signal affecting degradative enzyme synthesis in Bacillus subtilis*. J Bacteriol. 1995. **177**(9): p 2403–2407.

43. Zhao,X., Y. Wang, Q. Shang, Y. Li, H. Hao, Y. Zhang, Z. Guo, G. Yang, Z. Xie, and R. Wang, *Collagen-Like Proteins (ClpA, ClpB, ClpC, and ClpD) Are Required for Biofilm Formation and Adhesion to Plant Roots by Bacillus amyloliquefaciens FZB42*. PLoS One. 2015. **10**(2)

44. Jiang, W.,D. Bikard, D. Cox, F. Zhang, and L.A. Marraffini, *RNA-guided editing of bacterial genomes using CRISPR-Cas systems.* Nat Biotechnol. 2013. **31**(3): p 233-239

45. Turnbull P.C., M. Diekmann, J.W. Kilian, W. Versfeld, V. De Vos, L. K. Wolter, P. Bartels, and A. Kotze, *Naturally acquired antibodies to Bacillus anthracis protective antigen in vultures of southern Africa*. Onderstepoort J Vet Res. 2008. **75**(2): p 95-102.

46. Mock, M. and A. Fouet. *Anthrax.* Annu. Rev. Microbiol. 2001. 55: p 647–71

47. Baillie, L., T. Townend, N. Walker, U. Eriksson, and D. Williamson, *Characterization of the human immune response to the UK anthrax vaccine.* FEMS Immunol Med Microbiol. 2004. **42**(2): p 267-70.

48. Cybulski, R.J. Jr, P. Sanz, D. McDaniel, S. Darnell, R.L. Bull, and A.D. O'Brien, *Recombinant Bacillus anthracis spore proteins enhance protection of mice primed with suboptimal amounts of protective antigen.* Vaccine. 2008. **26**(38): P 4927-4939

49. Cote, C.K.,L. Kaatz, J. Reinhardt, J. Bozue, S.A. Tobery, A.D. Bassett, P. Sanz, S.C. Darnell, F. Alem, A.D. O'Brien, and S.L. Welkos, *Characterization of a multi-component anthrax vaccine designed to target the initial stages of infection as well as toxaemia*. J Med Microbiol. 2012. **61**(10):p 1380-1392

50. Warrens, A., M.D. Jones, and R. I. Lechler, *Splicing by overlap extension by PCR using asymmetric amplification: an improved technique for the generation of hybrid proteins of immunological interest*. Gene. 1997. 186: p 29–35.

VITA

Justyna Kordzikowska was born and raised in Rzeszow, Poland. After graduating from high school, she and her family moved to the United States. She attended the College of DuPage, where she earned an Associate Degree in 2008. Following graduation from community college, she attended the University of Illinois at Chicago where she earned her Bachelor Degree in 2010. After finishing her studies, she started a job as a microbiology technician. While working, she became interested in microbiology and decided to expand her knowledge in this field. She applied to the Microbiology and Immunology program at Loyola University and joined the program in August 2014.

THESIS APPROVAL SHEET

The thesis submitted by Justyna Kordzikowska has been read and approved by the following committee:

Adam Driks, Ph.D., Director

Department of Microbiology and Immunology

Loyola University Chicago

Alan Wolfe, Ph.D.

Professor and Committee Chair

Department of Microbiology and Immunology

Loyola University Chicago

Francis Alonzo, Ph.D.

Associate Professor

Department of Microbiology and Immunology

Loyola University Chicago

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Science

_____ _____

Date Director's Signature