COMPLEX ORGANIZATION AND DYNAMIC REGULATION OF THE *pks* GENE CLUSTER IN *Bacillus subtilis*

A Dissertation

by

CAROL MILEIDY VARGAS BAUTISTA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Paul Straight
Committee Members,	Michael Polymenis
	James Sacchettini
	Deborah Siegele
Head of Department,	Gregory Reinhart

December 2014

Major Subject: Biochemistry

ABSTRACT

The *pks* genes are the largest antibiotic- encoding gene cluster in *Bacillus subtilis* and encode the Pks enzymatic complex that produces bacillaene. Bacillaene plays important roles in the fitness of *B. subtilis* during competition with other bacterial species.

In this dissertation, I investigate the regulatory mechanisms used by *B. subtilis* to control the expression of the *pks* genes and the production of bacillaene. First, I focus on understanding the transcriptional regulatory network that coordinates the activation of the *pks* genes. My results indicate that multiple transcriptional regulators, in particular the stationary phase regulators Spo0A and CodY, coordinate the control of the *pks* gene activation. Also, cells dedicated to the formation of biofilms and spores but not motility induce the expression of the *pks* genes. I discuss these findings in light of their roles during bacterial competition.

I also identified multiple regulatory elements along the *pks* genes. Promoters upstream of *pksB*, *pksC* and *pksS* are active during vegetative growth while a promoter upstream of *pksG* is active only during spore formation. The activity of the *pksG* promoter is exclusive to the nascent spores and not the mother cells. In addition to promoters, a *cis*-regulatory element at the intergenic region of *pksC* and *pksD* promotes readthrough of transcription terminators along the *pks* genes.

Finally, I focus on the function of PksA, previously presumed to regulate the *pks* genes. I have found that PksA is not involved in the control of the *pks* gene expression.

Instead, PksA negatively regulates the expression of *ymcC*. My data suggests that YmcC is not involved in bacillaene production but, consistent with structural prediction, I have found that YmcC is a membrane protein produced during sporulation. I hypothesize the function of YmcC during spore maturation or germination and propose experiments to elucidate this role.

In general, this dissertation contributes to the understanding of *pks* gene regulation and its implications in the competitive fitness of *B. subtilis*. This work also provides a model for the activation of Type I *trans*-AT PKSs encoded in gene clusters with similar organization to the *pks* genes.

DEDICATION

To Samuel, my unborn child, who gave me the last strength to complete this journey. To my husband, Manuel, for dreaming my dreams and giving me a shoulder to lean on during all these years. To my mother, brothers and sisters for believing in me my entire life and being my greatest cheerleaders.

ACKNOWLEDGEMENTS

I am very grateful to all the people who contributed to this work. Thanks to the faculty that was directly involved in my research, in particular to my advisor, Paul Straight and the rest of my committee, Michael Polymenis, James Sacchettini and Deborah Siegele. I appreciate the deep interest and all the invested time in providing me direction and keeping track of my progress. Thanks also to Jennifer Herman, Ryland Young and Craig Kaplan for contributing to my scientific thinking, technical skills and courage every time I needed it.

I also want to acknowledge the members of my lab, current and former, and other scientists and collaborators outside this department who closely followed my research and provided me with many insights about this project.

Finally, I want to thank the administrative staff and all my friends and colleagues, senior and younger biochemistry students, who gave me plenty of good moments full of laughs, food and happiness.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Page

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Secondary metabolites were considered for long time as small molecules that are dispensable for growth and reproduction of the organisms that synthesize them, contrasting with the essential primary metabolites (1). Kossel, a German plant physiologist that was trying to classify compounds produced by plants, introduced the terms "primary" and "secondary" metabolites for the first time in 1891. He called "primary" those metabolites that were present in all living organisms and therefore he considered essential for survival. Given that Kossel found an increasing number of compounds that were distinctive to individual species, he decided to group them as "secondary metabolites" or dispensable by-products of central metabolism (2).

Secondary metabolites are widespread and abundant in nature, suggesting that they provide important functions to the organisms that synthesize them. It is estimated that a wide number of plants produce, as a group, over 22,000 different secondary metabolites (3). In addition, rough calculations predict more than 50,000 different secondary metabolites produced by fungi and bacteria. Seventeen percent of known secondary metabolites produced by microbes are synthesized by molds while filamentous bacteria belonging to the group of actinomycetes produce seventy four percent of these natural compounds (3). Therefore, secondary metabolites have gained recognition over time and their beneficial roles for the organisms that produce them are an active area of investigation (4, 5).

Broad evidence indicates that secondary metabolites have a variety of biological roles in nature. The effects of secondary metabolites depend on their chemical structure, concentration and composition of the biotic surroundings. Secondary metabolites can be involved in monitoring cell density, growth coordination, virulence, bacterial communication and competition (3–6). However, the most extensive studies have been done in the role of secondary metabolites as antibiotics, antitumor, antiviral and other properties related with medicinal uses.

Significance of antibiotics

Antibiotics were first discovered in the mid 1930's of the last century with such success that they were considered the cure of the infectious diseases. In 1929, the Scottish scientist Alexander Fleming observed that his cultures of *Staphylococcus aureus* were contaminated with a mold, *Penicillium notatum*, which inhibited the growth of *S. aureus*. Fleming was able to separate the mold cells from the liquid medium that contained the active ingredient, and called this ingredient Penicillin. By 1943, the production of Penicillin was industrialized and the compound was successfully used to treat wound infections during World War II (7). In the same year, the laboratory of Selman Abraham Waksman, at Rutgers University, isolated a compound with antibiotic properties from actinomycetes. The compound was named streptomycin and became the first antibiotic to cure tuberculosis. Waksman research also led to the recognition that the filamentous bacteria that belong to actinomycetes are the most prolific producers of

antibiotics (8). Today, hundreds of antibiotics are approved to treat bacterial infections that affect humans, animals and plants (3).

In spite of the success antibiotics have had to treat infections, pathogenic bacteria have been able to counteract antibiotics by developing resistance against them. By chromosomal mutation or exchange in genetic material, bacteria have been able to overcome antibiotic effects by inactivating of the compound, preventing of access to the target or altering of the antibiotic target site (9). Unfortunately, antibiotic discovery has not kept pace with resistance. The cases of common infections that do not respond to conventional antibiotic treatment are increasing every year. The increasing resistance offsets the growing collection of antibiotics over time. A worldwide alarm has risen to search for new strategies in order to keep fighting this battle (10, 11).

The rate of new antibiotic development has also steeply declined (12). Since the early 1980s, the US Food and Drug Administration has approved fewer antibiotics for clinical trials. For example, nineteen new drug applications were approved in the early 1980s, but just three new ones were approved from 2005-2009, showing a marked decrease in new antibiotic development (13). The technical difficulty in identifying and developing truly novel and clinically useful antibiotics is the most significant factor (12).

According to predictions based on bacterial genome capacity, our current stock of antibiotics is just a minimal fraction of antimicrobial compounds that can be isolated (14, 15). However, the majority of gene clusters that encode for antibiotics remain silent under standard laboratory conditions due, in part, to our insufficient knowledge about the environmental signals and metabolic networks that activate antibiotic synthesis

pathways in bacteria (16). Therefore, it is imperative to increase our understanding of how bacteria respond to their biotic and abiotic surroundings, adapt their general metabolism and couple such metabolism with activation of antibiotic synthesis.

Structure and biosynthesis of antibiotics is broadly diverse

The diversity and abundance of structures in different families of antibiotics is remarkable. These small molecule families differ from each other, not only in their backbone structures, but also in their biosynthesis (17). In the case of non-ribosomal peptides and polyketides, the largest and most studied families of antibiotics, the variety of chemical structures is related to the combinatorial utilization of chemical building blocks that are put together by enzyme complexes.

Peptides, the most abundant family of compounds with antimicrobial activity, are usually cyclic, with highly hydrophobic character, and generally resistant to peptidases and proteases due to the presence of D-amino acids (18, 19). Two different mechanisms may lead to the synthesis of these unconventional small peptides. Ribosomes may synthesize small peptides that are subsequently subjected to post-translation modifications. These small compounds synthesized by ribosomes are usually called bacteriocins. Alternatively, antimicrobial peptides might also be produced by large non-ribosomal peptide synthetases (NRPS). NRPSs coordinate multistep catalytic activities of highly conserved domains that recognize different amino acid substrates to activate them and condense them into the nascent molecules in a series of sequential reactions (20).

The basic catalytic modules of NRPSs contain either condensation (C), adenylation (A) or peptidyl carrier protein (PCP) domains (Fig. 1-1A) (21). The A domain selects its cognate amino acid and enzymatically generates a stabilized aminoacyl adenylate substrate, requiring ATP. The PCP domain contains an invariant serine at the catalytic site that binds a phosphopantotheine cofactor to which the adenylated amino acid substrate is transferred and thioesterified, with release of AMP (22). The C domain is located between the A and the PCP domain and catalyzes the formation of a new peptide bond by condensation. Sequential reactions of the assembly line arrangement of the three domain units ensure the coordinated elongation of the peptide product. Subsequently, the linear peptide chain may be cyclized by the action of a thioesterase present at the C-terminus of the NRPS (19, 22). Due to this modular biochemistry –known as the colinearity rule- the products that are synthesized by NRPS can be predicted from the organization of their synthetic modules (23). Conversely, the primary structure of the peptides can also help to identify the orphan gene clusters responsible for their synthesis (24).

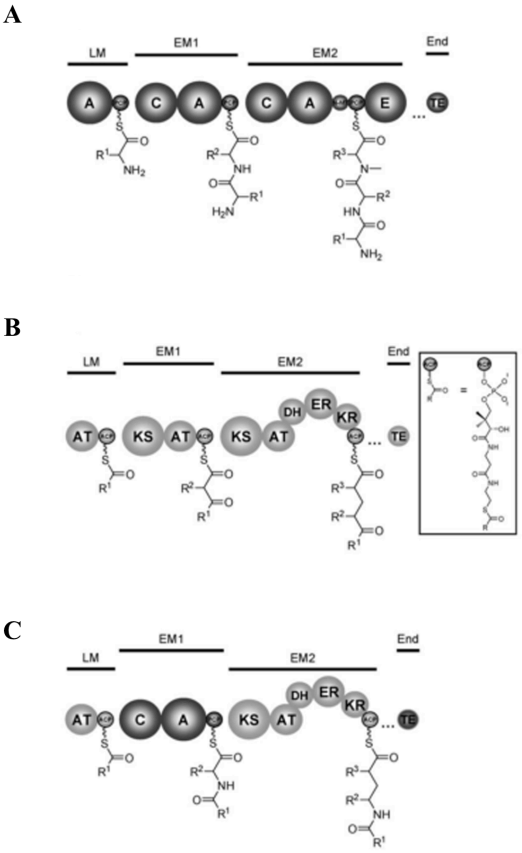


Figure 1- 1. General biosynthetic organization of NRPS and PKS.
(A) In NRPS systems, the loading module (LM) consists of an A domain that transfers the substrate to the PCP domain. The minimal extension module (EM1) contains a C domain that catalyzes a new peptide bond between the PCP and the A domain substrates. Optional processing domains (EM2) are used to incorporate modifications to the nascent molecule. Biosynthesis is terminated by an integral thioesterase (TE) domain. (B) In PKS systems, the typical loading module (LM) contains an AT domain that recruits the acyl monomer and transfer it to the ACP. In addition to the AT and ACP domains, the extension module (EM1) contains a KS domain to perform the Claisen-like condensations. As ilustration, a second extension module (EM2) contain other modifying enzymes. (C) The hybrid PKS-NRPS systems contain a mix of domains and modular organization from PKS and NRPS. Abbreviations are: AT, acyltransferase; ACP, acyl carrier protein; KS, ketosynthase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; TE thioesterase; A, adenylation; PCP, peptidyl carrier protein; C, condensation; N-MTN-methyl transferase; E, epimerase. This figure was adapted from Weissman et al (2008) (21).

Type I polyketide biosynthesis follows the same colinearity rule principle as in the biosynthesis of non-ribosomal peptides (24). Mechanistically, polyketide synthetases are closely related to NRPS, as both modular systems utilize multiple carriers for covalent binding of monomers and growing chains (Fig. 1-1B) (19). In general, PKS contain multiple modules of at least three catalytic domains that act in tandem to modify the nascent molecule. A ketosynthase (KS) domain performs the Claisen-like condensations, an acyl carrier protein (ACP) domain binds covalently to acyl intermediates, and an acyltransferase (AT) domain recruits and transfers the acyl monomers to the ACP. The multidomain structure of type I PKSs contrast with type II and type III PKSs, which carry each of their catalytic sites on different proteins or are homodimeric single-domain proteins, respectively. Type II and type III PKSs are beyond the scope of this thesis. For a further review see (25, 26).

Type I PKSs are present in many species that belong to the taxonomic groups of *Pseudomonas spp*, actino-, cyano- and myxobacteria (27). In particular, a growing class of PKSs known as type I *trans*-Acyl-Transferase Polyketide Synthases (*trans*-AT PKSs) have been found in bacteria during the last decade. *Trans*-AT PKSs had been overlooked in the past since they are rare in actinomycetes, the main source of antibiotics (28). However, an increasing number of proteobacteria, myxobacteria and bacilli have been found to be the main source of *trans*-AT PKSs products (29). For instance, some *Bacillus* species produce macrolactin and bacillaene and their enzymatic machineries serve as models for *trans*-AT PKSs (30).

Trans-AT PKSs generally deviate from the colinearity rule (23, 31). They are characterized by the absence of AT domains in their synthetic modules so that the monomers are incorporated by AT enzymes that act *in trans* (32) (Fig.1-2). In addition, *Trans*-AT PKSs are the product of DNA arrangement mosaics from multiple gene segments gained via extensive horizontal gene transfer between bacteria (31). Therefore, the architecture of the modules is highly variable and may include repeated domain sets, modules split between proteins and apparent superfluous domain that contradict the colinearity rule and make it difficult to predict function based on PKS structure (29, 31). In consequence, the majority of the efforts in relation with *trans*-AT PKSs have been mainly focused on the prediction of their molecular architectures and synthetic mechanisms. The coordinate regulation for the expression of genes that encode the synthetic proteins is poorly understood.

Bacillaene, a prototype for *trans*-AT PKSs

Bacillaene, a protein synthesis inhibitor, was the first product of a *trans*-AT PKSs to be discovered, in 1993 (33). The whole genome sequencing of *B. subtilis* strain 168 established the presence of a PKS but the strain was not able to synthesize its product, leading to the conclusion that the *pks* genes was an evolutionary relic (31). The isolation and structure characterization of bacillaene was reported two years later but lacked two amino acid building blocks that were predicted from the *pks* genes, making it difficult to predict the relationship between bacillaene and the *pks* genes (34). The

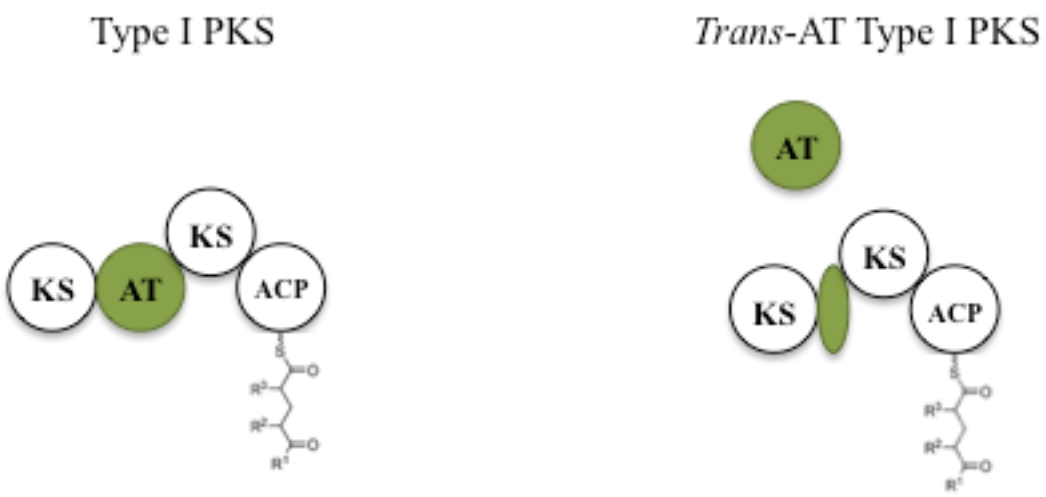


Figure 1- 2. Organization of *trans*-AT type I PKSs.
(A) A prototype of type I PKS that contains the AT domain. (B) A *trans*-AT type I PKS lacks the integral AT domain within each module. It may contain a short segment of remnant AT residues (green). Abbreviations are: AT, acyltransferase; ACP, acyl carrier protein; KS, ketosynthase; KR, ketoreductase (21, 35).

connection between the *B. subtilis pks* genes and the product, bacillaene, was found more than a decade later with biological competition assays between *B. subtilis* str. NCIB 3610 and *Streptomyces coelicolor* and bacillaene structure was confirmed by NMR (36, 37). Another independent line of research developed a comparative analysis between *Bacillus amyloliquefaciens* strain FZB42, *B. subtilis* A1/3 and *B. subtilis* ATCC 39320 and found that *B. amyloliquefaciens* also produce bacillaene (38)(39). Since its discovery, the Pks synthetase that produces bacillaene has been a prototype for *trans*-AT PKSs (28).

The structure of the Pks synthetase and the synthesis of bacillaene have been broadly characterized. Bacillaene is a linear molecule and has a high degree of conjugation with several *cis* double bonds (Figure 1-3A). Two main isomers are synthesized by the Pks enzymatic complex, bacillaene ($C_{34}H_{49}N_2O_6$) and dihydrobacillaene ($C_{34}H_{51}N_2O_6$), with masses of 581 D and 583 D, respectively. The structure of bacillaene allowed a more accurate annotation of the Pks proteins and the *pks* genes (Fig. 1-3B-C). The hybrid PKS/NRPS enzymatic complex contains 13 PKS and 3 NRPS modules (PksJ-PksR), three free-standing AT domains (PksCDE) and a protein subcluster responsible for a β-methylation in the middle of the hexaene moiety (PksCFGHI and AcpK) (36). PksS is a cytochrome P450 homolog that oxidizes dihydrobacillaene to bacillaene and PksB is a predicted zinc-dependent hydrolase (40). Several other unusual biochemical reactions of *trans*-AT PKSs are also present in the Pks synthetase (31). The first module of PksJ initiates with a α-hydroxy-isocaproic acid (α-HIC) as starter unit, instead of the typical acetyl-CoA or malonyl-CoA. There are two

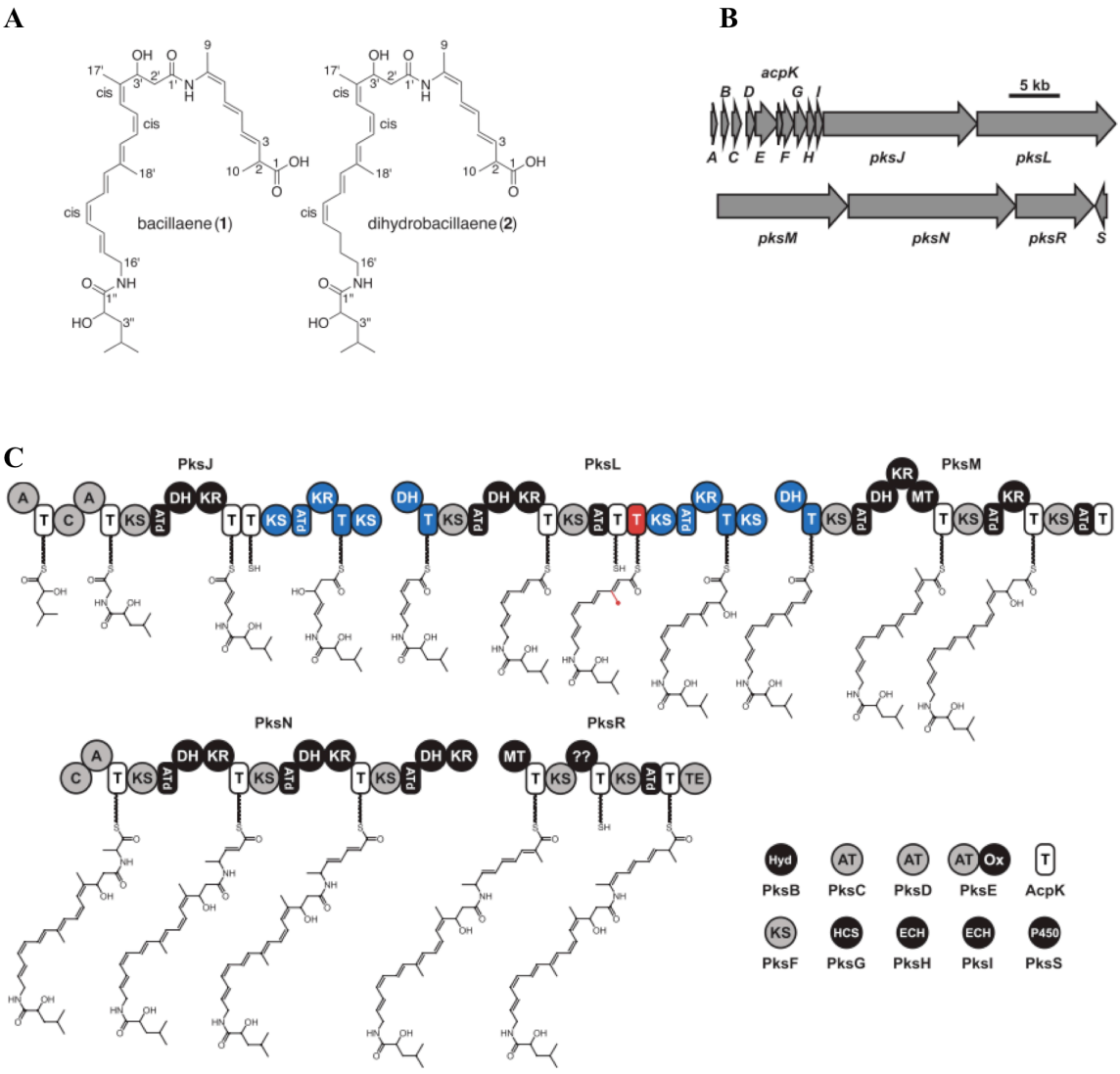


Figure 1- 3. Biosynthesis of bacillaene in *B. subtilis*.
(A) Structure of bacillaene and dihydrobacillaene. The C14'-C15' bond is olefinic in bacillaene and saturated in the precursor dihydrobacillaene. (B) The *pks* gene cluster in *B. subtilis* encode the Pks enzymatic complex that produces bacillaene. (C) General proposed model for the synthesis of bacillaene. The two split modules are depicted in blue. In red, the timing of the β- branching by the PksCFGHI/AcpK subcluster. Abbreviations: KS, ketosynthase; AT, acyltransferase; T, thiolation; DH, dehydratase; KR, ketoreductase; MT, methyltransferase; A, adenylation; C, condensation; ATd, AT-docking; Hyd, Zn-dependent hydrolase; Ox, flavin mononucleotide-dependent oxidase; HCS, HMG-CoA synthase; ECH, enoyl-CoA hydratase/isomerase; TE, thioesterase. Adapted from Butcher et al (2008) (36).

split modules at PksJ-PksL and PksL-PksM that appear to divide reduction and dehydration into discrete steps (36). The timing of the β-branching occurs during elongation (as opposed to after), and there is formation of Z double bonds by type A dehydrating bimodules (41). The pathway of bacillaene biosynthesis has been elucidated to near completion, but little emphasis has been put on the regulation of the *pks* gene cluster nor the function of the molecule.

Relevance of bacillaene in the ecology of *B. subtilis*

Antibiotics are not only bacterial weapons for fighting competitors but also molecules with multiple functions that may contribute to the stability of entire microbial communities (42). Antibiotics were considered for decades to be the microbial deadly weapons to kill competitors. The reason for this misconception is that research on these natural compounds has been mainly focused on their role as drugs and not much attention has been put on their function in natural environments (43). However, it is plausible to think that sub-lethal concentrations of antibiotics are the rule, not the exception in nature (44). Indeed, several studies have shown that sub-lethal concentrations of antibiotics are intercellular signals to promote the fitness of the bacteria that produce them (6, 42, 45). Bacterial species other than the producers may also identify antibiotics as signals to activate biological functions that benefit the entire community (42, 46). However, antibiotics may also be synthesized to compete for space and nutrients, although with no lethal effects for the competitors. Non-lethal effects of antibiotics, such as developmental arrest or nutrient piracy, would relax the pressure on

the competitor to survive and develop resistance (47). Numerous functions have been found for antibiotics but a complete understanding of their role in nature still remains to be elucidated. These molecules may even have multiple roles in nature and understanding these functions is crucial to address the increasing clinical problems of antibiotic resistance (14).

Bacillaene is not only a broad-spectrum antibacterial compound but also a molecule with multiple functions, as apparent from competitions assays between *B. subtilis* and *Streptomyces spp* (37, 48, 49). Bacillaene was first reported as a bacteriostatic antibiotic (34). The instability of bacillaene makes it unsuitable as a drug but competition assays have revealed the importance of this molecule in the competitive fitness of *B. subtilis* (Fig. 1-4). For example, the growth inhibition effect of bacillaene is clearly observed in competition assays pairing *B. subtilis* with *Streptomyces avermitilis* (Fig. 1-4A). *B. subtilis* is able to compete with *S. avermitilis* when bacillaene is present; however in its absence (due to a deletion of the *pks* genes) *S. avermitilis* inhibits the growth of *B. subtilis*. The interactions of *B. subtilis* with several other species of *Streptomyces* reveal intriguing phenotypes other than growth inhibition that suggest multiple functions of this molecule.

Competition assays between *B. subtilis* and *Streptomyces coelicolor* suggest that bacillaene interferes with the secondary metabolism of *S. coelicolor*. In particular, production of undecylprodigiosin (a.k.a. RED) and actinorhodin (a.k.a. BLUE) by *S. coelicolor* seem to be affected in the presence of bacillaene (Figure 1-3B) (37). RED and BLUE are two antibiotics whose gene clusters get activated during the transition

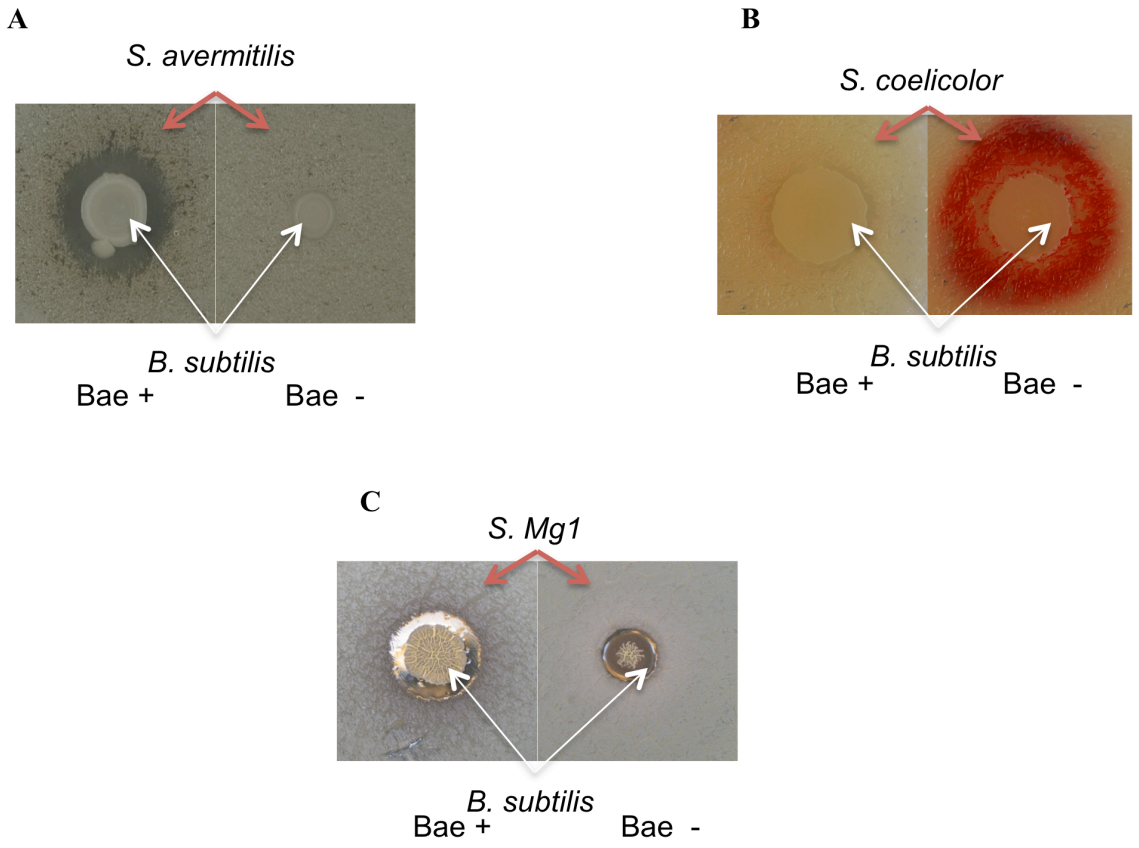
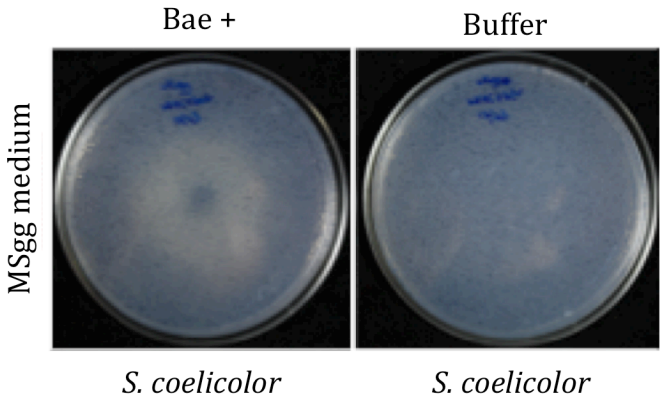


Figure 1- 4. Absence of bacillaene changes competition outcomes of *B. subtilis* and *Streptomyces spp*.
A lawn of 10^5 spores/ml of *Streptomyces spp* was plated in G7 medium and 2.5 µL of a *B. subtilis* overnight –wild type or Δ*pks*- was spotted in the middle of the plate. Plates were incubated at 30˚C and tracked over time. (A) Bacillaene is an antibacterial compound against *S. avermitilis*. (B) Bacillaene modulates the synthesis of prodiginines, RED and BLUE, in *S. coelicolor*. (C) Bacillaene contributes to the defense mechanisms of *B. subtilis* in competition with *S.* sp Mg1.

between mycelial growth and sporulation (12, 50). During interactions of wild-type *B. subtilis* and *S. coelicolor* no changes in the synthesis of antibiotics by the latter are observed. However, in the absence of bacillaene (*B. subtilis* Δ*pks*), the premature synthesis of RED and BLUE in the adjacent *S. coelicolor* becomes apparent (Fig. 1-4B) (37). To determine whether bacillaene was affecting the expression of RED and BLUE gene clusters, we purified bacillaene and added it to MSgg plates that contained lawns of *S. coelicolor* (10^5 spores/ml). After 24 hours of growth at 30˚C, the plates that contained bacillaene had a weak decrease in pigment production (Fig. 1-5A). However, qRT-PCR analysis of gene targets from the RED and BLUE gene clusters showed no significant differences in the mRNA abundance between bacillaene treated and not treated control (Fig. 1-5B). Therefore, bacillaene modulates synthesis of RED and BLUE by altering steps in RED and BLUE production that are subsequent to gene expression. In addition, three unknown metabolites of *m/z* 407, 641 and 812 are also repressed by bacillaene, as detected by IMS (48). These metabolites are known as bacillaene-repressed molecules (BRM) and the mechanism by which they are repressed remains to be elucidated.

Cocultures of *B. subtilis* with S*treptomyces* sp. Mg1 highlight the detrimental effects of the absence of bacillaene in this competitive interaction (Fig. 1-4C). Bacillaene provides a defense mechanism to *B. subtilis* in competition against *S.* sp Mg1 (49). *Streptomyces* sp. Mg1 causes cellular lysis and disrupts the colony extracellular matrix of *B. subtilis*. Strains of *B. subtilis* are hypersensitive to the lytic activity when

A



B

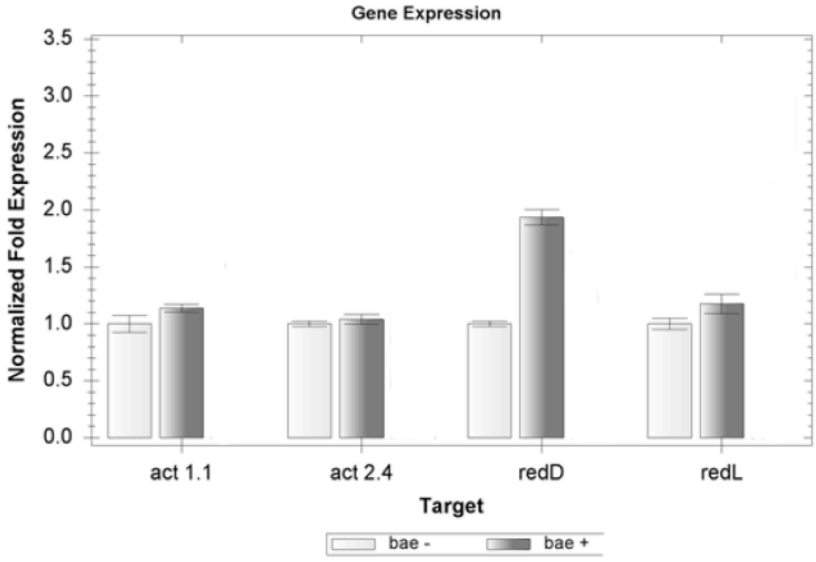


Figure 1- 5. Modulatory effect of bacillaene on production of prodiginines by *S. coelicolor*.

(A) Bacillaene decreases prodiginine production by S. coelicolor. MSgg plates that contained lawns of S. coelicolor (10^5 spores/mL) were incubated at 30˚C after addition of bacillaene and tracked over time. Plates that contained bacillaene but not sodium phosphate buffer (negative control) decreased the production of bacillaene. Susceptibility to bacillaene of *S*. sp Mg1 Δs.c. 37 was used as positive control. (B) RNA abundance of targets act 1.1, act 2.4 (actinorhodin gene cluster), redD and redL (undecylprodigiosin gene cluster) were measured by qRT-PCR. No significant differences were found between bacillaene treated and not treated samples. C_q values were determined for each target and normalized using C_q for *rpoB*. Fold-expression values reported are relative to the untreated sample for each data set.

bacillaene synthesis is disrupted by deletion of the *pks* operon (Unpublished work by Reed Stubbendieck, Straight lab).

Bacillaene also seems to arrest the development of aerial hyphae of *S.* sp. Mg1, phenotype known as bald (Fig. 1-6). It has been previously reported that *B. subtilis* uses surfactin to inhibit aerial growth in several *Streptomyces spp.* (51). *Streptomyces* sp. Mg1 is not sensitive to surfactin due to the production of surfactin hydrolase, an enzyme that degrades the molecule (52). Consequently, no bald phenotype is observed when surfactin is applied to lawns of *Streptomyces* sp. Mg1. However, the interaction between *B. subtilis* and *S.* sp. Mg1 still causes a bald phenotype in the latter, indicating that *B. subtilis* uses alternative mechanisms to inhibit aerial growth in *S.* sp. Mg1 (51). I have concluded that bacillaene arrests aerial growth of *S.* sp. Mg1 by competition assays between *S.* sp. Mg1 and Δ*pks* strain (Fig. 1-6A). Lawns of *S.* sp. Mg1 (10^7 spores/mL) in G7 plates with 2.5 μL spots of a Δ*pks B. subtilis* overnight lost the ability to cause the bald phenotype. In addition, direct application of bacillaene to a plate culture is sufficient to arrest aerial growth development in *S.* sp. Mg1 (Fig. 1-6B). Taken together, these competition assays indicate the multiple effects of bacillaene on competitor streptomycetes. The mechanism(s) of bacillaene action remain to be elucidated.

Regulation of bacillaene production

In general, synthesis of secondary metabolites by bacteria is regulated at the transcriptional level (50, 53–56). Cells recognize physiological and environmental

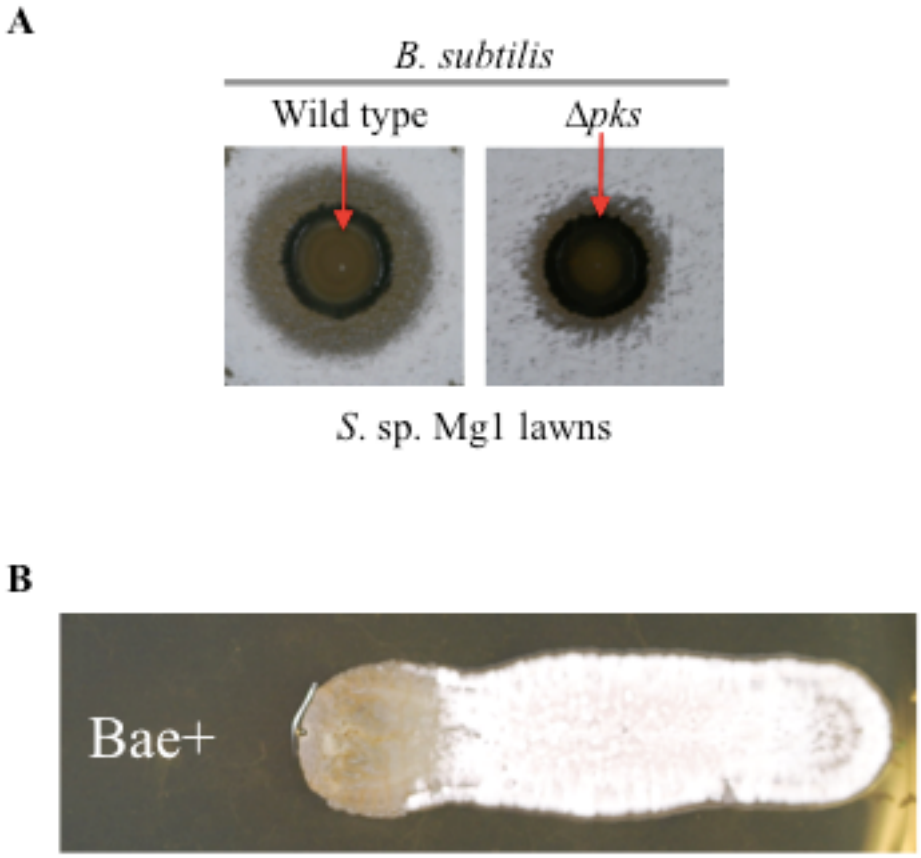


Figure 1- 6. Bacillaene arrests aerial growth in *Streptomyces* sp Mg1.
(A) co-cultures of *S.* sp. Mg1 and *B. subtilis* wild type and Δ*pks* in G7 medium show that bacillaene is necessary to cause aerial growth arrestment in *S.* sp. Mg1. (B) Addition of purified bacillaene is sufficient to cause bald phenotype in *S.* sp. Mg1.

signals to activate transcription of genes for secondary metabolites via pathway-specific or pleitropic regulators. A pathway-specific regulator is dedicated to control the expression of the genes encoding the production of the secondary metabolite. Often the regulatory gene is encoded within the gene cluster. A pleiotropic regulator controls multiple metabolic functions, one of which may be regulation of the secondary metabolite biosynthetic gene cluster. Pleiotropic regulators may exert control on the pathway-specific regulator, if present, or directly control the biosynthetic genes. Insights on the coordination of those regulatory processes will contribute to a better understanding of the regulation of secondary metabolites.

The regulation and synthesis of bacillaene implies a substantial amount of cellular resources (31). The *pks* gene cluster, which encodes the enzymatic complex that produces bacillaene, is one of the earliest AT-less PKS gene clusters to be reported (33). The *pks* genes span almost 80 kb of *B. subtilis* genome (~2%) and therefore they compose the biggest gene cluster in *B. subtilis* (37). Expression of this massive region of DNA has to be precisely controlled such as synthesis of bacillaene does not become a burden for the cells. The *pks* gene cluster encodes a Pks enzymatic assembly line that exceeds 100 nm, being larger than that of the ribosome (31). A single Pks assembly line would comprise a core synthetase of nearly 2.5 MD but multiple assembly lines assemble into a single organelle-like complex of about 10-100 MD (37). Moreover, transcription of the *pks* genes has to be coordinated with other processes that involve a high demand of energy consumption, such as synthesis of flagella for motility and formation of spores (57). As an archetype of *trans*-AT PKS encoding gene cluster, the

pks gene cluster contains three ATs (PksCDE) that iteratively load extender units into the ACP domains (32, 35, 36). The regulatory mechanisms to activate the *pks* genes have not been elucidated.

Dissertation overview

The purpose of this dissertation is to investigate the regulatory mechanisms used by *B. subtilis* to control the expression of the *pks* genes and the production of bacillaene. The regulatory proteins involved in the transcriptional activation of the *pks* genes will be discussed in chapter II. The regulatory features present in the *pks* gene cluster, including a novel *cis*-RNA element that allows the read-through of transcriptional terminators, will be discussed in chapter III. The function of the first protein encoded by the *pks* genes, PksA, will be described in chapter IV. Finally, the conclusions and future directions of each line of investigation are discussed in chapter V. This dissertation contributes to the current knowledge of the ecology of *B. subtilis* and the activation of biosynthetic pathways of natural products.

In chapter II, I report an intricate network of transcriptional regulators that coordinate the expression of the *pks* genes. I provide evidence that PksA does not participate in the regulation of these genes, as it was previously assumed. Instead, transcription of the *pks* genes requires the coordination of multiple regulatory proteins that are also involved in developmental transitions and nutrient stress response. In particular, CodY, a stationary-phase global regulator that senses nutrient scarcity, induces the highest expression of the *pks* genes at the entry into stationary phase. A

master regulator for development, Spo0A, is also responsible for full induction of the *pks* genes. The intervention of multiple regulators leads to a transient repression of the *pks* genes during motility and their activation during stationary-phase functions, such as biofilm and sporulation. With this work I provide a better understanding of cell specialization for the biosynthesis of antibiotics. I hypothesize how this knowledge can contribute to the fitness of *B. subtilis* population during bacterial competition.

Chapter III presents intriguing data about the dynamic regulatory elements along the *pks* genes that allow differential gene expression during vegetative growth and sporulation. Promoters upstream of *pksB*, *pksC* and *pksD* are active during vegetative growth. The *pksC* promoter seems to control the expression of almost 77 kb, form *pksC* to *pksR*. Surprisingly, a promoter at *pksG* becomes very active during sporulation. This activation of gene expression is not only differential respect to the genes upstream of *pksG* but also compartmentalized inside the forespores. The data also suggest that the intergenic region between *pksC* and *pksD* is a novel *cis*-RNA element that promote read-through of termination sites at the *pks* genes. To my knowledge, no regulatory RNAs had been reported to control production of natural products. I speculate about the biological functions of the peculiar characteristics of the *pks* operon. With this work, I provide new insights on antibiotic regulation in many organisms that encode biosynthetic enzymatic complexes in gene operons with similar organization.

In chapter IV, I extend my investigation about PksA, the previously presumed regulator of the pks genes. In chapter II, I found that PksA is not involved in the control of bacillaene production. In chapter IV, I present evidence that PksA controls the

expression of its own gene and *ymcC*, a gene with unknown function that sits adjacent and divergently oriented to *pksA*. The control exerted by PksA on *ymcC* is strong and no detectable levels of YmcC are observed at any time during vegetative growth. However, YmcC is observed during sporulation, suggesting a potential role during spore maturation or germination. Possible roles during these stages of development and potential experiments to solve these questions are also proposed.

Last, chapter V contains the general conclusions of this dissertation and discusses the future directions of this research. I propose connections between the different transcriptional regulators of the *pks* genes, based in previous and current work. I include ideas related with temporal control of the regulatory features of the *pks* genes to optimize the production of bacillaene. I also speculate about the roles of the differential activation of *pks* promoters and the *cis*-regulatory RNA during cell sporulation.

CHAPTER II

BACTERIAL COMPETITION REVEALS DIFFERENTIAL REGULATION OF THE *pks* GENES BY *Bacillus subtilis*[*]

Summary

Bacillus subtilis is adaptable to many environments in part due to its ability to produce a broad range of bioactive compounds. One such compound, bacillaene, is a linear polyketide/non-ribosomal peptide produced by *B. subtilis*. The *pks* genes encode the enzymatic megacomplex that synthesizes bacillaene. The majority of the *pks* genes appear to be organized as a giant operon (> 74kb from *pksC-pksR*). In previous work, a deletion of the *pks* operon in *B. subtilis* was found to induce prodiginine production by *Streptomyces coelicolor*. Here, colonies of wild type *B. subtilis* formed a spreading population that induced prodiginine production from *Streptomyces lividans,* suggesting differential regulation of *pks* genes and thus bacillaene. While the parent colony showed widespread induction of *pks* expression among cells in the population, we found that the spreading cells uniformly and transiently repressed the expression of the *pks* genes. To identify regulators that control *pks* genes, we first determined the pattern of *pks* gene expression in liquid culture. Next we identified mutations in regulatory genes that disrupted the wildtype pattern of *pks* gene expression. We found expression of the *pks*

---

[*] This chapter, written by Carol Vargas-Bautista, was adapted from the original publication in Journal of Bacteriology (149): **Vargas-Bautista C, Rahlwes K, Straight P**. 2014. Bacterial Competition Reveals Differential Regulation of the *pks* Genes by *Bacillus subtilis*. J. Bacteriol. **196**:717–28 (Copyright © American Society for Microbiology. Reprinted with Permission.)

genes requires the master regulator of development, Spo0A, through its repression of AbrB, and the stationary-phase regulator, CodY. Deletions of *degU*, *comA*, and *scoC* had moderate effects, disrupting the timing and level of *pks* gene expression. The observed patterns of expression suggest that complex regulation of bacillaene and other antibiotics optimizes competitive fitness for *B. subtilis.*

Introduction

Bacillus subtilis is a globally dispersed bacterial species that is competitive in diverse environments and produces numerous bioactive compounds. *B. subtilis* dedicates 4-5% of its genome to produce secondary metabolites (19). In particular, three massive gene clusters encode enzyme complexes for dedicated synthesis of their cognate products. Two of the gene clusters encode the non-ribosomal peptide synthetases (NRPS) for surfactin (*srfAA-srfAD;* 27 kb) and plipastatin (*ppsA-ppsE;* 37 kb), respectively. Surfactin is a multifunctional lipopeptide that provides surfactant and signaling activities required for motility and biofilm development (58–60). Plipastatin is a lipopeptide with antifungal properties (61, 62). A third gene cluster (*pksA-pksS;* 78 kb) encodes machinery for the production of bacillaene, a hybrid non-ribosomal peptide/polyketide (NRP/PK) produced by *B. subtilis* (36, 37).

The diverse functions of bacillaene are apparent from competition studies pairing *B. subtilis* with species of *Streptomyces* (37, 48, 49) (See chapter I). The importance of bacillaene for competitive fitness of *B. subtilis* raises the question of how the organism regulates *pks* gene expression and bacillaene biosynthesis. The *pks* gene cluster has been

annotated as sixteen genes, five of them encoding the multimodular synthetase (*pksJ, L, M, N, and R)*, and another ten (*pksB-pksI* and *pksS*) genes encoding individual enzymes that function *in trans* to the assembly line (Fig. 2-1A). The first fifteen genes, *pksA-pksR*, are oriented in the forward direction (positive strand) and the last gene, *pksS*, in the reverse direction (negative strand). In many cases, modular type I PKS, NRPS and hybrid PKS-NRPS gene clusters include associated regulators that coordinate the expression of the synthesis genes (24, 63). The *pksA* gene sits adjacent to the gene cluster and encodes a putative TetR-family regulatory protein (http://genolist.pasteur.fr/SubtiList/). PksA is predicted to function as a pathway-specific regulator of the *pks* genes, but the regulatory function has not been experimentally confirmed (38–40). In addition to pathway-specific regulation, secondary metabolic pathways are commonly controlled by global regulatory functions that respond to changes in nutrient conditions or environmental cues to activate different physiological responses (17, 55). Differentially regulated functions in *B. subtilis* include genetic competence, motility, biofilm formation, and sporulation, in addition to production of antibiotics and degradative enzymes (64). Regulation of developmental processes has been studied in detail for *B. subtilis*, and in many instances the regulatory functions are known to influence secondary metabolism (19, 57, 64). Studies of surfactin, bacilysin, and other metabolites highlight the integration of secondary metabolism with different physiological states (58, 60, 65, 66)

In the present study, we identified a competitive interaction with *S. lividans* that suggested differential regulation of bacillaene production between morphologically

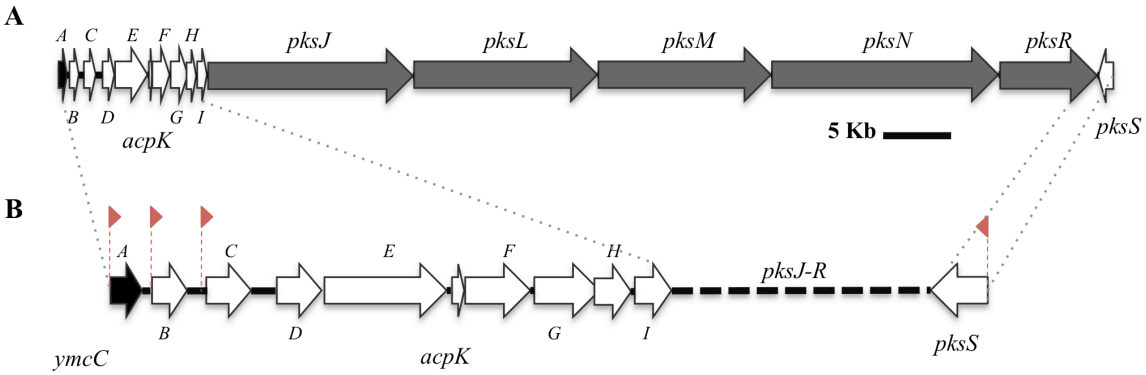


Figure 2- 1. The *pks* gene cluster in *B. subtilis*
(A) Sixteen genes from *pksA* to *pksS* (78.6 kb) comprise the *pks* gene cluster as annotated in the *B. subtilis* 168 genome. Dark gray arrows represent the genes encoding the multimodular PKS enzymes that synthesize bacillaene. White arrows represent genes encoding functions required in trans to the multimodular enzymes. The black arrow represents *pksA*, which encodes a predicted TetR-family transcriptional regulator. Arrows are drawn to scale. (B) Expansion of the genes *pksA* – *pksI* and *pksS* highlights the intergenic regions (not to scale). Potential transcriptional control regions are indicated with red flags (67).

different subpopulations of *B. subtilis*. We investigated the regulation of *pks* gene transcription to determine whether bacillaene production is segregated in different *B. subtilis* subpopulations. Initially using liquid cultures, we show that the 5'UTR of *pksC* is active in promoting expression of the apparent *pks* operon, which extends nearly 75kb from the *pksC* to *pksR* genes (http://subtiwiki.uni-goettingen.de/apps/expression/) (67). Also, we show that the gene annotated as *pksA* does not encode a pathway regulator for bacillaene. Using transcriptional reporters fused to the *pksC* promoter element, we identified multiple global regulators that influence expression of the *pks* genes. We show that Spo0A is required to activate *pks* gene expression through repression of the transition state regulator, AbrB (19, 68). Expression of *pks* genes is also dependent on CodY, which regulates metabolism in response to nutrient status and was recently shown to bind to multiple sites in the *pks* operon (69, 70). DegU, ComA and ScoC are also required for full induction of *pks* gene expression. Using transcriptional reporters, we show that the expression of *pks* genes is homogeneously and transiently repressed in cells that spread toward *S. lividans* in a competitive interaction. Our data indicate that *B. subtilis* uses multiple regulatory functions to exert dynamic control of bacillaene production, which may benefit the overall competitive fitness of the colony.

Material and methods

Bacterial strains, primers, media and growth conditions

Table 2-1 contains a list of strains used in this study. The undomesticated strain *Bacillus subtilis* NCIB 3610 was used for all the experiments in this work. Unless

otherwise stated, all *B. subtilis* strains were cultured at 37°C in CH medium (1% casein hydrolysate, 0.47% L-glutamate, 0.16% L-asparagine, 0.12% L-alanine, 1 mM KH_2PO_4, 25 mM NH_4Cl, 0.22 mg/ml Na_2SO_4, 0.2 mg/ml NH_4NO_3, 1 μg/ml $FeCl_3$.6H2O, 25 mg/liter $CaCl_2.2H_2O$, 50 mg/ liter $MgSO_4$, 15 mg/liter $MnSO_4.H_2O$, 20 μg/ml L-tryptophan, pH 7.0), which is commonly used for consistent timing of developmental transitions and optimal for live cell microscopy (71). To generate a uniform population of cells in early exponential growth phase, overnight cultures of *B. subtilis* were diluted to an OD_{600} = 0.085, cultured to approximately OD_{600} = 0.2, and re-diluted to OD_{600} = 0.085. This cycle was repeated three times before initiation of the experiments. Genetic manipulations of *B. subtilis* were initially made using the PY79 strain, and then transduced via bacteriophage SPP1 into *B. subtilis* NCIB 3610 as previously described (72). All manipulations were confirmed by genomic extraction, amplification of genetic targets and sequencing. *Escherichia coli* XL1 blue was used for plasmid manipulations and storage.

Table 2- 1. Strains used in this study

Strain	Relevant Genotype	Source
PSK0531	*Streptomyces lividans* wild-type strain TK24	Laboratory collection
PDS0066	*B. subtilis* NCIB3610 Wild-type	Laboratory collection
PKS0212	*B. subtilis* NCIB3610 *pksR:yfp* spec	(37)
PDS0184	*B. subtilis* NCIB3610 *pksA::Kn*	This study
PDS0480	*B. subtilis* NCIB3610 *pksA::Kn lacA::pksA (mls)*	This study
PDS0183	*B. subtilis* NCIB3610 Δ*pksA amyE::Phyperspac:pksA:lacI (cat)*	This study
PDS0032	*B. subtilis* NCIB3610 *amyE:: P_{pksB}-yfp (cat)*	This study
PDS0036	*B. subtilis* NCIB3610 *amyE:: P_{pksC}-yfp (cat)*	This study
PDS0035	*B. subtilis* NCIB3610 *amyE:: P_{pksS}-yfp (cat)*	This study
PDS0189	*B. subtilis* NCIB3610 *amyE:: P_{pksB}-lacZ (cat)*	This study
PDS0227	*B. subtilis* NCIB3610 *amyE:: P_{pksC}-lacZ (cat)*	This study
PDS0201	*B. subtilis* NCIB3610 *amyE:: P_{pksS}-lacZ (cat)*	This study
PDS0430	*B. subtilis* NCIB3610 *amyE::P_{pksC}-yfp (cat) lacA::P_{hag}-cfp (mls)*	This study
PDS0432	*B. subtilis* NCIB3610 *amyE::P_{pksC}-yfp (cat) lacA::P_{tapA}-cfp (mls)*	This study
PDS0431	*B. subtilis* NCIB3610 *amyE::P_{pksC}-yfp (cat) lacA::P_{sspB}-cfp (mls)*	This study
PDS0327	*B. subtilis* NCIB3610 Δ*spo0A::mls*	This study
PDS0382	*B. subtilis* NCIB3610 Δ*degU::tet*	Kolter lab
PDS0247	*B. subtilis* NCIB3610 Δ*abrB::tet*	Kolter lab
PDS0262	*B. subtilis* NCIB3610 Δ*abh::km*	Kolter lab
PDS0512	*B. subtilis* NCIB3610 Δ*comA::cat*	This study
PDS0368	*B. subtilis* NCIB3610 Δ*scoC::km*	This study
PDS0525	*B. subtilis* NCIB3610 Δ*codY::mls*	This study
PDS0337	*B. subtilis* NCIB3610 Δ*sigD::mls*	Kearns Lab
PDS0311	*B. subtilis* NCIB3610 Δ*spo0A::mls* Δ*abrB::tet amyE:: PpksC:lacZ (cat)*	This study

Antibiotics used in this study were chloramphenicol (5 μg/ml), spectinomycin (100 μg/ml), tetracycline (10 μg/mL), kanamycin (10 μg/mL) and mls (1 μg/mL of erythromycin, 25 μg/mL of lincomycin).

Co-culture assays

G7 plates (1.5% Bacto agar, 1% Bacto Malt Extract, 0.4% yeast extract and 0.4% D-glucose, buffered with 100 mM morpholinepropanesulfonic acid [MOPS] and 5 mM potassium phosphate) were used to co-culture *B. subtilis* and *Streptomyces lividans*. X-gal (300 μg/ml) was added to the plates when needed. Briefly, 2 μl of *S. lividans* spores (10^7 spores/ml) were spotted on solid media and incubated at 30°C for 12 hours. Following initial incubation of the *S. lividans*, 1.5 μl aliquots of a *B. subtilis* overnight culture were spotted in a cross-wise pattern to the *S. lividans*, and plates were returned to incubation at 30°C. Taking as time zero when *B. subtilis* was spotted, the co-culture was observed over time and images were captured at the indicated time points.

Extraction and quantification of bacillaene

Time course experiments were done in triplicate with cells growing at 30°C, in 500 ml of CH medium, under constant agitation (250 rpm) and complete darkness. To extract bacillaene, 15 ml of the culture supernatants were mixed 1:1 with dichloromethane. Bacillaene was recovered by evaporation of the organic phase followed by resuspension in methanol. The methanol was then evaporated and the samples resuspended in a buffer of 65% 20 mM sodium phosphate/ 35% Acetonitrile

and H_2O. HPLC analysis was performed with a C18 reverse-phase column (Phenomenex). Samples were eluted with a gradient of 35% - 40% acetonitrile and 65% - 60% of 20mM sodium phosphate. Bacillaene was detected by UV absorption using a wavelength of 361 nm as previously reported (9). The amount of bacillaene in each sample was determined by integrating the area under the relevant peaks on the elution chromatograph. We confirmed the specificity of bacillaene peaks in the HPLC chromatographs by comparison to a sample from a *B. subtilis* Δ*pks* strain. LC-MS analysis determined that the relevant peaks were all different isoforms of bacillaene (not shown). Quantitative data was normalized to the sample cell density (OD_{600}) in order to compare synthesis of the molecule over time between strains.

Quantitative RT-PCR (qRT-PCR)

Cell samples were stabilized using RNAprotect Bacteria Reagent (Qiagen) and RNA isolation was performed using RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. Subsequently, RNA samples were treated with Turbo DNA-free kit (Applied Biosystems) to remove DNA traces and total RNA was quantified. Thermo Scientific DyNAmo Flash SYBR green qPCR kit was used with target specific primers listed in Table 2-2 and 200 µg of total RNA as template to synthesize cDNA. After the RT step, quantitative PCR was done in a CFX96 Touch Real Time PCR thermocycler (BioRad). The protocol was as follows: denaturation at 95.0°C for 15 min; 39 cycles of denaturation at 94.0°C for 10 s, annealing at 58.0°C for 25 s, extension at 72.0°C for 30 s; and final melt curve from 60.0°C to 95.0° for 6 min. We determined

that *gyrB* transcript abundance per cell did not significantly change from OD_{600} = 0.2 to 6.8 (not shown). Consequently, we used *gyrB* as the reference gene. The samples were run in triplicate for each target gene and negative controls where included for each sample as reaction mixtures with total RNA after DNase treatment (no RT performed). Primer efficiency and C_q values were calculated using the software LinReg (73). Gene study analysis for comparison between independent experiments was performed based on the primer efficiency calculated by the software LinReg and the analysis of the CFX Manager software (Bio-Rad).

Western blots

Cell growth conditions are the same as described for extraction and quantification of bacillaene (see above). Fifteen milliliters of cell pellets from each time point were lysed by incubation in 500 µl of lysis buffer (50 mM Tris 7.5, 200 mM NaCl, 0.5 mM EDTA, 5 mM $MgCl_2$, 1mg/ml lysozyme, 1 mM 4-(2-Aminoethyl)benzenesulfonylfluoride hydrochloride (AEBSF), 1 mM DTT) at 37ºC for 15 min. After treatment, protein concentration was measured by Bradford Assay (Bio-Rad Protein Assay) and lysates were diluted to 1 mg/ml of total protein. Addition of 2X loading buffer in a 1:1 ratio and heating at 100ºC for 5 min was done before loading 30 µl of the samples in 8% acrylamide gel for SDS-PAGE. Proteins were transferred onto an immobilon PVDF membrane (Sigma). Rabbit anti-GFP (1:1000) and goat anti-rabbit HRP (1:5000) (Invitrogen) served as primary and secondary antibody, respectively. The

blotting was visualized using Pierce ECL western blotting substrate (ThermoScientific) according to manufacturer instructions.

Fluorescence microscopy

Samples from shaken liquid cultures in CH medium were taken for fluorescence imaging, centrifuged at 8,000 rpm and washed once with PBS. Cells were resuspended in 20 µM 1-(4-trimethylammoniumphenyl)-6-phenyl-1,3,5-hexatriene p-toluenesulfonate (TMA) (Molecular Probes) and fluorescence images were captured using a Nikon Ti-E inverted microscope equipped with a CFI Plan Apo Lambda DM 100X objective, TI-DH Diascopic Illuminator, and a CoolSNAP HQ2 Monochrome Camera. Exposure time was 2,000 ms for YFP, 200 ms for CFP and 1,000 ms for TMA. The NIS-elements AR software was used to capture and process the images identically for comparative analysis. Samples from solid media were scraped, dissolved in PBS, passed repeated times through a 25g 1½ needle to disrupt aggregated cells and centrifuged at 8,000 rpm. All the subsequent steps were the same as described for samples from liquid medium.

Construction of pksA mutants

Deletion of *pksA* was performed by long-flanking homology PCR, using the primers *pksA* KO_P1, *pksA* KO_P2, *pksA* KO_P3 and *pksA* KO_P4 (Table 2-2) to amplify the region flanking *pksA* and the intervening kanamycin cassette (74). To overexpress *pksA,* primer pair *pksA*-90_FHIII/*pksA*-90_RSalI were use to amplify *pksA*. The amplified regions were cut with the restriction enzymes HindIII and SalI (NEB) and

Table 2- 2. Primers used in this study

Name	Sequence (5'-3')
pksA KO_P1	gatggccgcgataaaagtaa
pksA KO_P2	cctatcacctcaaatggttcgctgcgttgcttctgcaatttgtt
pksA KO_P3	cgagcgcctacgaggaatttgtatcggcgtggaagatacacgtgag
pksA KO_P4	aacaccttctatgtaatcattttcg
pksA compl-F	atgcatgctagcatctcgagaacccaaaacgcaatttcac
pksA compl-R	aacgtcccggggagctcatgaattccaagaatcgcttttcgcac
pksA-90_FHIII	taaagcttaatccattcccctcttttc
pksA-90_RSalI	ttaagtcgaccaacaagaatcgcttt
pA-F(EcoRI)	ttagaattcataagcgatcgatatacc
pA-R(HindIII)	taggaagcttagctttattgtaacaagaaa
pC-F(EcoRI)	ttagaattcccattcgataaaggat
pC-R(HindIII)	tatgaagcttgattagtagatgtgtttcac
rtPCR_pksC_1	aaagccgcatctcttttga
rtPCR_pksC_2	gcatgaaggaactcctcgaa
qPCR-pksE1	tacgtgagctggatgcaaag
qPCR-pksE2	atgcttcgggttttgttcag
qPCR-pksR-F	acagcgtaacggaattttgg
qPCR-pksR-R	ttgattgcccttccttatcg
gyrB qPCR_F	gggcaactcagaagcacggacg
gyrB qPCR_R	gccattcttgctcttgccgcc

ligated with T4 ligase (NEB) into pPST001 (*amyE::Phyperspac lacI cat amp*). The plasmid was recovered by transformation into *E. coli* XL1Blue, transformed into *B. subtilis* PY79 and transduced into PDS0184 (Table 2-1), as previously described (72). For *pksA* complementation, *pksA* compl-F/*pksA* compl-R primers (Table 2-2) were used to clone the *pksA* gene with 203 bases of upstream sequence into pDR183 (*lacA:: mls amp*) by enzymatic assembly, as previously described (75).

Transcriptional fusions of pks promoters

Primer pairs pC-F(EcoRI)/pC-R(HindIII) (Table 2-2), were used to amplify 300-500 bps upstream of *pksA*, *pksB*, *pksC* and *pksS*, respectively. The amplified regions were cut with the restriction enzymes EcoRI and HindIII (NEB) and ligated with T4 ligase (NEB) into pCW001 (*amyE::yfp cat amp*) and pDG1661 (*amyE::lacZ cat amp*). The transformations were performed using *E. coli* XL1Blue for recovery of the plasmids. Subsequent transformation into PY79 and transduction into NCIB 3610 *B. subtilis* strains were performed as previously described (72). Recovered clones were grown in CH medium for subsequent analysis by fluorescence microscopy and β-galactosidase assays.

β-galactosidase assays

Samples were taken over time and cell density measured at OD_{600}. β-galactosidase assays were done as previously described by Miller (1972) (76). Briefly, one ml samples were lysed with Z buffer (60 mM $Na_2HPO_4 \cdot 7H_2O$, 40 mM

$NaH_2PO_4 \cdot H_2O$, 10 mM KCl, 1mM $MgSO_4 \cdot 7H_2O$) that contained 0.27% BME and lysozyme (200 µg/ml) at 30°C for 20 min. Then, serial dilutions of the samples were done to find an optimal range for colorimetric detection with ONPG (400 µg) at OD_{420} and OD_{550}. The values are reported in Miller Units (MU).

Results

Co-culture of B. subtilis with Streptomyces lividans suggests that bacillaene synthesis is inactive within spreading populations of B. subtilis

In a previous study, we found that a bacillaene-deficient *B. subtilis* strain (Δ*pksB-R,* from this point forward Δ*pks*) induces the production of red-pigmented prodiginines (RED) by *S. coelicolor* (37, 48). Based on the observed pattern of induction, we associate RED with the absence of bacillaene in our co-culture assays. In the present study, we plated colonies of *S. lividans*, which also encodes the RED genes, adjacent to wild-type *B. subtilis* colonies (50). Over the course of 4 days, the *B. subtilis* colonies spread on the plates toward the *S. lividans* colonies (Fig. 2-2). We observed that the RED pigment was induced where the spreading *B. subtilis* population contacts the colonies of *S. lividans*. The observed RED induction is similar to prior observations

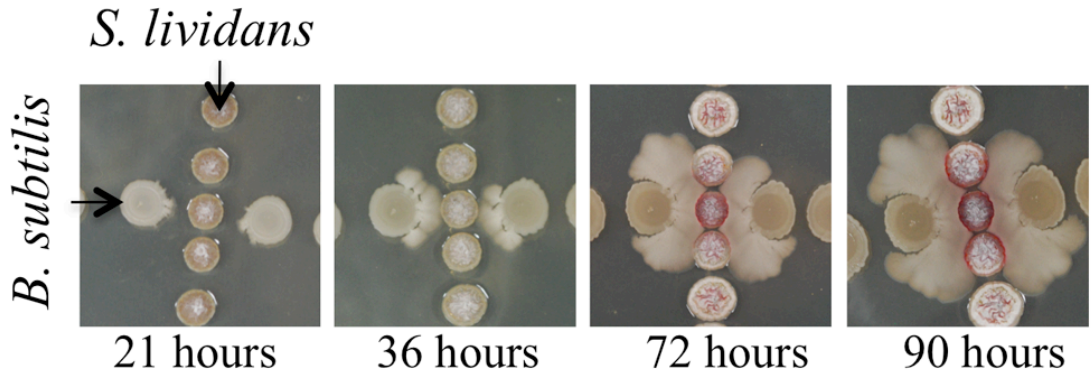


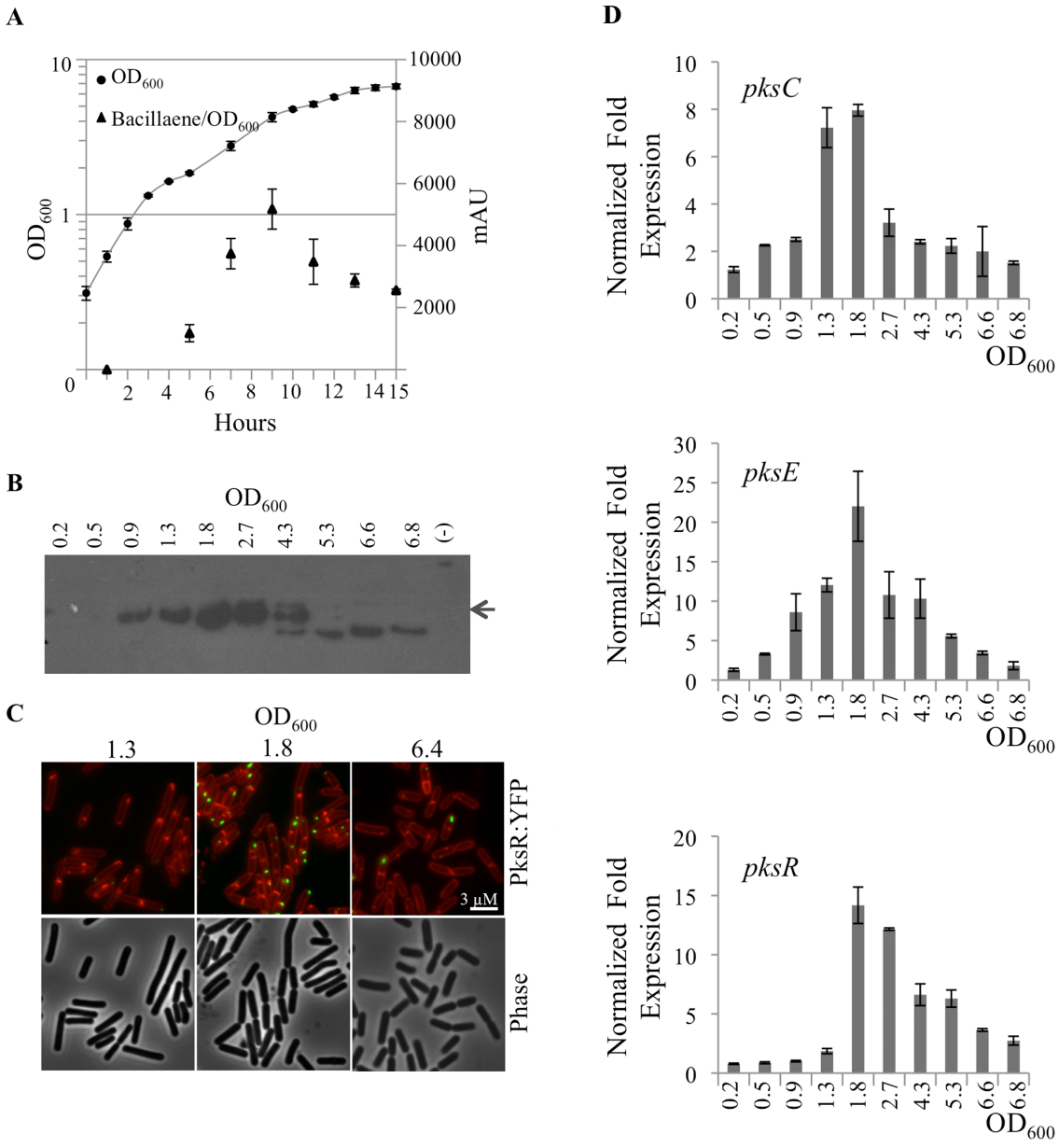
Figure 2- 2. Induction of RED pigment by *S. lividans* is associated with absence of bacillaene.
B. subtilis spotted cross-wise with *S. lividans* inoculated 12 hours prior from a spore suspension. Time zero corresponded to inoculation of *B. subtilis*. Initially, both species formed round colonies. After 21 hours, the *B. subtilis* colonies begun to migrate toward *S. lividans*. Upon contact with *B. subtilis* (36-72 hours) *S. lividans* induced prodiginines (RED pigment), which are enhanced with extended incubation (90 hours). No RED pigment is detected in the absence of colony contact in the time frame studied. The images shown represent the results of multiple experiments done in duplicate.

using bacillaene-deficient Δ*pks* strains cocultured with *S. coelicolor*. The presence of RED suggested the possibility that the spreading cells do not produce bacillaene and raised the question whether differential expression of the *pks* genes occurred in different subpopulations of *B. subtilis*.

Bacillaene production peaks at the onset of stationary phase in liquid culture

To understand the regulatory functions that control bacillaene production, we first used classical growth in liquid culture to follow the pattern of bacillaene synthesis and to identify the relevant regulatory elements. In previous work, fluorescence microscopy of cells expressing individual Pks proteins fused to yellow fluorescent protein (YFP) or cyan fluorescent protein (CFP) revealed that the bacillaene megacomplex synthetase accumulates within *B. subtilis* cells as cultures approach high cell density (37). This pattern of megacomplex assembly suggests that regulation of bacillaene synthesis is coordinated with cellular growth. To build a comprehensive view of bacillaene synthesis, we sought to determine whether *pks* gene expression and bacillaene secretion follow a similar pattern as observed for megacomplex assembly. Thus, we monitored bacillaene synthesis, megacomplex formation, and *pks* gene expression in samples taken from a liquid culture of the strain PKS0212, which expressed YFP fused to the C-terminal end of the PksR protein (Fig. 2-3) (37). We chose to use PksR as a representative of the assembly-line enzymes required for bacillaene synthesis because the *pksR* gene resides at the 3' end of the nearly 75 kb *pks* operon, as described for the *pks* gene cluster in the SubtiExpress database

Figure 2- 3. Bacillaene production during liquid culture of *B. subtilis* NCIB 3610.
Strain PSK0212 (PksR-YFP) was cultured in CH medium (30°C), and sampled over 15 hours. All quantitative data shown are average values with standard deviations from triplicate experiments. (A) Growth curve of PSK0212 and HPLC quantitation of bacillaene. Equal culture volumes were sampled for OD_{600} measurements (circles). Bacillaene extracted from cell-free supernatants was quantitated by HPLC (triangles) (mAU (λ=361nm)/OD_{600}). Peak bacillaene accumulation per OD_{600} was detected at OD_{600} = 4.2. (B) Western blot (α-GFP) of PksR-YFP from *B. subtilis* cell lysates. A single PksR-YFP band (indicated with an arrow) was detected at low cell densities, and increased in intensity to a maximal level observed between OD_{600} = 1.8 - 2.7. The signal intensity for PksR-YFP decreased at higher cell density and lower molecular mass forms appeared, suggesting degradation of PksR. (C) Upper panels- Fluorescence images of PksR-YFP (green) in cells stained with TMA (red) to visualize membranes. Lower panels- Phase contrast images of cells. PksR-YFP signal intensity and number changed with cell density. Maximal intensity was observed at the end of log phase (OD_{600} = 1.8) and was diminished at high cell density (OD_{600} = 6.4). Scale bar, 3 μM. (D) qRT-PCR of representative *pks* genes. C_q values were determined for *pksC*, *pksE* and *pksR* and normalized using C_q for *gyrB*. Fold-expression values reported are relative to the wild-type lowest cell density sample (OD_{600} = 0.2) for each data set. The maximal fold expression for each transcript occurred at the OD_{600} = 1.8 time point.


A
OD600
Bacillaene/OD600
OD600
mAU
Hours
B
OD600
0.2 0.5 0.9 1.3 1.8 2.7 4.3 5.3 6.6 6.8 (-)
C
OD600
1.3
1.8
6.4
PksR:YFP
Phase
3 µM
D
pksC
pksE
pksR
Normalized Fold Expression
OD600

(http://subtiwiki.uni-goettingen.de/apps/expression/) (67). As the final product transcribed from the *pks* operon, we postulated that the accumulation of PksR protein approximates the amount of completely assembled enzymatic complexes within the cell. *B. subtilis* PSK0212 cultures growing in CH medium at 30°C were sampled at multiple times over 15 hours and monitored using three approaches. First, we used HPLC to quantitate bacillaene in the culture supernatant (Fig. 2-3A). Second, we used the PksR-YFP chimera to monitor the protein accumulation by western blot (Fig. 2-3B) and the formation of megacomplex by fluorescence microscopy (Fig. 2-3C) (37). Third, we measured the abundance of three transcripts that span the length of the operon, *pksC*, *pksE,* and *pksR* by quantitative RT-PCR (qRT-PCR) to determine the pattern of *pks* gene expression (Fig. 2-3D).

The production of bacillaene by *B. subtilis* followed a pattern typical of many antibiotics produced during the transition from exponential growth to stationary phase (19, 22, 77). Bacillaene was not detected by HPLC in cultures of low cell density (< 0.5 OD_{600}). However, the amount of bacillaene per unit OD_{600} in the culture broth increased over time until the onset of stationary phase (Fig. 2-3A). The transition from exponential to stationary phase in CH medium occurred above OD_{600} ~ 1.5 under the culture conditions used. Above this cell density, the increase in detectable bacillaene per unit OD_{600} was pronounced, reaching a peak accumulation at OD_{600} = 4.2. Upon further incubation, the amount of bacillaene/unit OD_{600} in culture supernatants declined, suggesting that active synthesis is diminished as cells progress into stationary phase.

We hypothesized that the rate of bacillaene synthesis would change with cell density if the megacomplex enzymes underwent assembly and subsequent turnover or inactivation during the course of growth. In a prior study, PksR-YFP was found to increase with cell density up to an OD_{600} = 1.7 (37). Here, we extended the cultures to OD_{600} = 6.8 in order to track the protein during stationary phase. We examined the levels of PksR-YFP protein in cells taken from the culture at the same time points as the samples taken for HPLC (Fig. 2-3A). Equivalent amounts of protein from whole cell lysates were probed with an anti-GFP antibody to detect the presence of PksR-YFP. At low culture density, PksR-YFP was below the level of detection, consistent with previous results (37). However, a band corresponding to 311 kD, the expected molecular mass of the PksR-YFP fusion, was readily detected at OD_{600} above 0.9, and reached peak intensity between OD_{600} 1.8 and 2.7, corresponding to the stationary phase transition. A second high molecular mass band became visible from OD_{600} 1.8. This may represent a modified form of PksR-YFP. Upon further incubation, PksR appears to be processed or degraded, as seen by the diminished signal of higher and lower molecular mass bands on the western blot (Fig. 2-3B).

The diminished PksR-YFP signal is consistent with the enzyme being turned over during stationary-phase, which would account for reduced bacillaene production in culture. We predicted that the fluorescent signal from PksR-YFP in assembled megacomplexes would decline in cultures of stationary phase cells in accord with degradation of PksR-YFP. Using the same culture conditions as above, we examined cells expressing PksR-YFP by fluorescence microscopy to follow the assembly of

megacomplexes and their subsequent disruption. As seen in Figure 2-3C, megacomplexes became visible as fluorescent foci within cells grown to intermediate cell density (OD_{600} = 1.8). At higher cell density (OD_{600} = 6.6), fewer PksR-YFP positive cells were observed and the overall intensity of the signal per cell is reduced. We counted cells with detectable, punctate YFP signal at each sample point and found that cells positive for megacomplexes first increased and then reduced to less than 50% of the population at high OD_{600} (Table 2-3). However, an intense fluorescent signal persists for a percentage of the cells at high cell density. Whether these cells actively produce bacillaene is unknown. Comparison of bacillaene production in Figure 2-3A with the fluorescence signal in figure 2-3C reveals a consistent pattern of megacomplex assembly and bacillaene synthesis that peaks during the transition to stationary phase and decreases upon continued incubation.

Antibiotic biosynthesis is commonly regulated by transcriptional activation of the biosynthetic gene clusters during transition from exponential to stationary growth phase. We next sought to determine if the *pks* genes are expressed in a pattern similar to the pattern of bacillaene production. We selected three open reading frames within the apparent *pks* operon for targeted expression analysis. Two of the genes, *pksC* and *pksE,* are positioned near the 5' end of the *pks* operon (Fig. 2-1A). The third gene we analyzed, *pksR*, is the final ORF before the predicted transcriptional terminator and encodes a multimodular

Table 2- 3. Formation of Pks megacomplex as seen by fluorescence microscopy[a]

OD_{600}	Total Cell Count[b]	PksR-YFP Count	% PksR-YFP
0.66	185	17	9
1.06	357	33	9
1.39	283	49	17
1.44	277	57	19
1.66	326	173	49
1.84	459	385	81
3.17	594	415	68
4.07	183	119	63
5.26	600	368	61
6.46	367	157	43

[a]Fresh cells were washed once with PBS, resuspended in 20μM TMA, and observed in phase contrast, TMA and yfp, as described in Materials and Methods.

[b]Number of cells per microscopic field.

PKS enzyme. To compare their patterns of expression at the beginning and end of the apparent operon, equivalent amounts of total RNA were used to measure relative amounts of transcripts for *pksC, pksE* and *pksR* using qRT-PCR. Here we found that all of the *pks* genes followed the same expression pattern (Fig. 2-3D). The transcripts were at the lowest level during exponential growth and peaked near the transition to stationary phase (1.8 OD_{600}). Consistent with the bacillaene and PksR-YFP results, *pks* transcripts diminished as cells progressed through stationary phase. This pattern of *pks* gene expression and bacillaene synthesis suggests that the production of bacillaene is tied to the levels of *pks* transcript in the cells.

The TetR-family protein PksA *is not involved in bacillaene regulation*

Many loci that encode assembly line enzyme complexes also encode transcription factors that control the expression of the biosynthetic genes (24, 54, 78). The *pksA* gene, located adjacent to the *pks* gene cluster, encodes a putative TetR-family regulatory protein that is predicted to function as the associated regulator of the *pks* genes (38–40). To determine whether PksA regulates *pks* gene expression, we replaced the endogenous *pksA* gene with a kanamycin resistance gene and examined the effect on *pks* gene transcripts. We measured the level of *pksC*, *pksE*, and *pksR* transcripts in the wild type and Δ*pksA* strains during the induction phase of *pks* gene expression (OD_{600} 0.2 to 1.8), by qRT-PCR (Fig. 2-4). The transcripts of all three genes increased several-fold for both wild type and Δ*pksA* strains as cultures exited log phase (Fig 2-4, OD_{600} = 1.8). We complemented the Δ*pksA* mutation with insertion of the *pksA* gene at

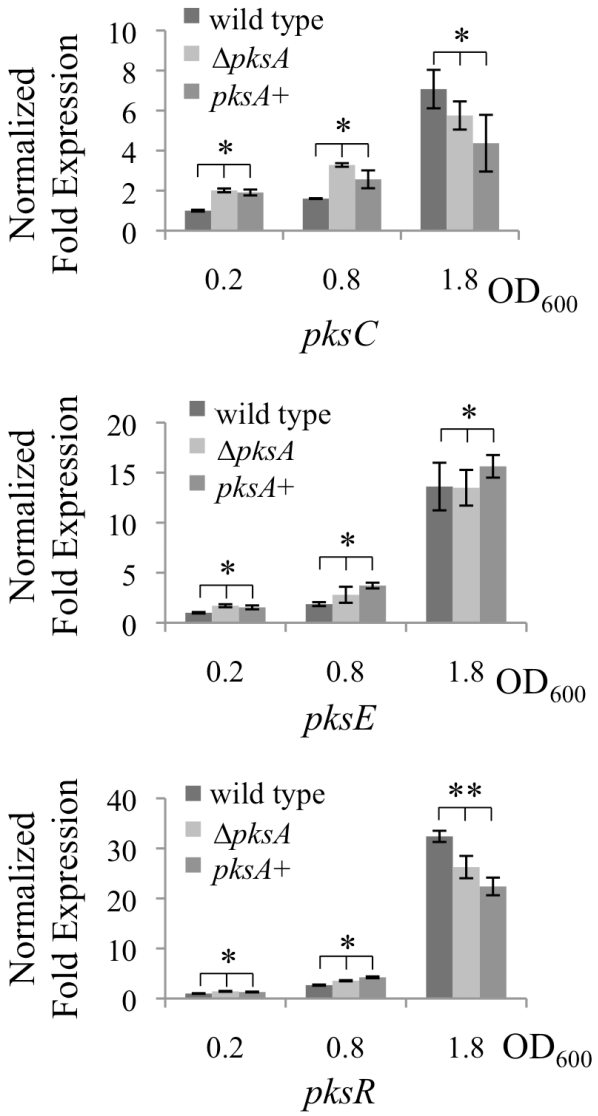


Figure 2- 4. PksA function is unrelated to regulation of bacillaene synthesis.
qRT-PCR data are presented as described for Figure 2-3, panel D. The *pksC*, *pksE*, and *pksR* transcripts measured by qRT-PCR were induced in wild type, Δ*pksA* and the Δ*pksA* genetically complemented mutant (*pksA*+) strains. Comparison of low (0.2), mid (0.8) and high (1.8) OD_{600} showed induction during growth. Two-factor analysis of variance showed no significant effect of the Δ*pksA* and *pksA*+ genetic background on the *pks* genes tested (*p-value > 0.05. **p-value >0.01).

the *amyE* locus. The complemented strain showed a similar pattern of gene expression as the Δ*pksA* and wild-type strains (Fig. 2-4). In addition, the Δ*pksA* deletion had no discernable effect on bacillaene production as determined by HPLC (Appendix I. Fig. 1A). The absence of a phenotype for Δ*pksA* does not preclude the function of PksA as a repressor of *pks* gene expression. To determine if overexpression of *pksA* would repress *pks* gene expression, we introduced an IPTG-inducible copy of *pksA* into the Δ*pksA* strain and quantitated *pksC* and *pksR* transcripts. No significant effect on either *pks* transcript was detected in the *pksA* overexpression condition, despite a 35-fold elevation in abundance of *pksA* transcript (Appendix I. Fig. 1B). Thus, neither deletion nor overexpression of *pksA* significantly perturbed the induction of the *pks* genes during growth of *B. subtilis*, leading us to conclude that the target of PksA regulation is not the *pks* operon.

The promoter PpksC controls expression of the pks gene cluster

Collectively, these data indicate that the regulation of the *pks* operon is coupled to cellular growth by an undetermined mechanism. To identify regulatory functions that activate bacillaene production, we generated a reporter for transcriptional activation of the *pks* operon. Based on a previous report of *B. subtilis* global gene expression, three putative upstream regulatory sequences are active within the *pks* gene cluster: P_{pksB}, P_{pksC}, and P_{pksS} (Fig. 2-1B) (http://subtiwiki.uni-goettingen.de/apps/expression/) (67). We isolated the 5'UTR of *pksC*, and fused it to the *yfp* and *lacZ* genes for fluorescence and β-galactosidase assays, respectively. Both fluorescence microscopy and β-galactosidase

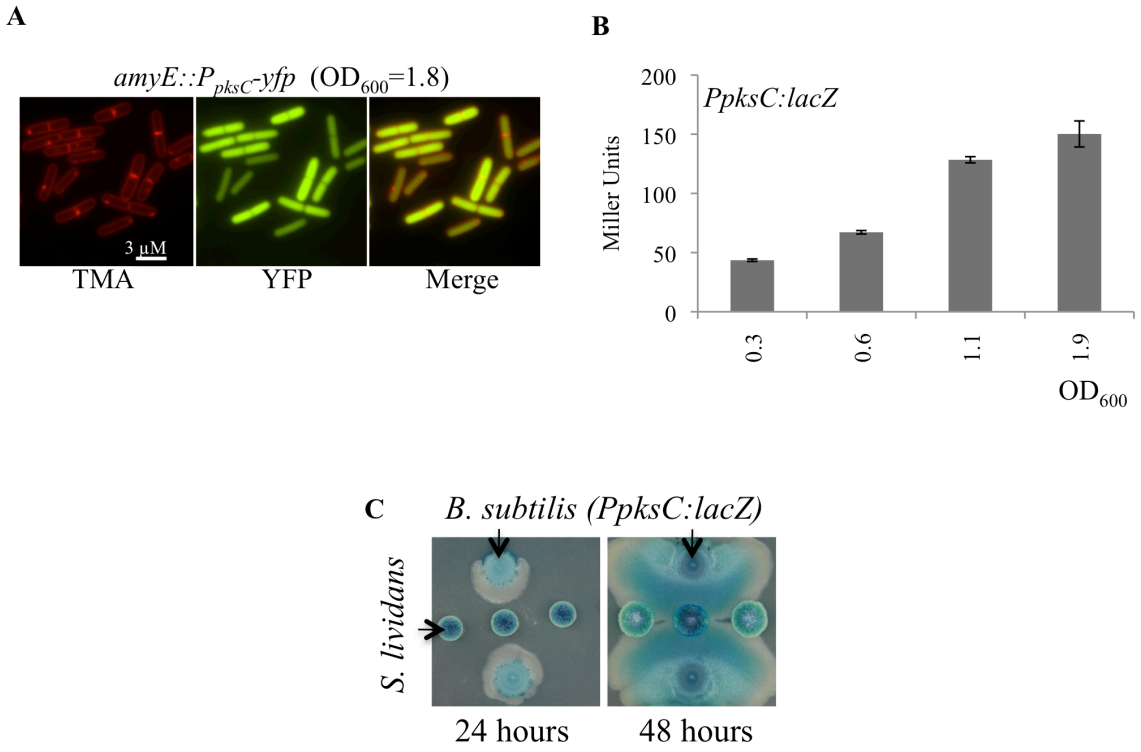


Figure 2- 5. Activity of promoters of the *pks* gene cluster.
(A) Fluorescence of transcriptional reporter for *pksC* promoter fused to *yfp*. Cells growing in liquid CH medium at 37°C were taken at indicated culture densities to observe activation of the promoters. Images represent several microscopic fields from samples of two independent experiments. TMA-stained membranes - red. Promoter-yfp fusions - green. Scale bar = 3 µm. (B) β-galactosidase assay of the P_{pksC}-*lacZ* strain. The pattern of β-galactosidase activity indicates the *pksC* promoter is activated during the transition to stationary phase. Miller units are average from triplicate experiments with reported standard deviation. (C) Co-culture of *S. lividans* and *B. subtilis* (P_{pksC}-*lacZ*). G7 plates (300 µg/ml of X-gal) were inoculated with *S. lividans* and *B. subtilis* as described in Fig. 2-2. P_{pksC}-*lacZ* activity was differentially localized between spreading and static cells in the colony. *S. lividans* endogenous β-galactosidase activity results in blue colonies. Images represent three independent experiments, each time performed in duplicate.

assays revealed that the *pksC* promoter is highly active and induced at the same cell density as *pksC* transcripts, as described above (Fig. 2-5A and 2-5B). Thus the P*pksC* reporter fusions provide a tool to determine patterns of *pks* gene expression in cultures of *B. subtilis*.

Differential activation of PpksC in colonies and motile subpopulations

We predicted that if bacillaene synthesis were inactivated in the spreading populations, as we hypothesized based on the induction of RED synthesis by *S. lividans* (Fig. 2-2), then P_{pksC}-*lacZ* activity would be differentially localized between the parent colony and the spreading subpopulation. *B. subtilis* carrying the P_{pksC}-*lacZ* reporter was challenged with *S. lividans* on plates containing X-gal. As previously described, the *B. subtilis* cells spread toward *S. lividans* on the agar plate, and P_{pksC}-*lacZ* was differentially activated within the colonies (Fig 2-5C). Endogenous β-galactosidase activity of *S. lividans* produced blue streptomycete colonies and obscured the visibility of RED pigment in these assays (79). However, β-galactosidase activity from *B. subtilis* was co-incident with the primary colony. As seen in Figure 2-5C, at early time points (24 hours) the β-galactosidase activity was absent within the spreading population of cells. These results supported the observation that the spreading cells are inactive for bacillaene production. Upon further incubation (48 hours), P_{pksC}-*lacZ* activity increased on the interior of the spreading population, but remained repressed at the leading edge where contact with *S. lividans* is initiated. We conclude from these data that the *pks* operon is

likely to be activated by regulatory pathways that at least transiently differentiate highly motile from static populations.

Multiple regulatory networks control pks gene expression

B. subtilis uses a complex network of regulatory proteins to control antibiotic production, developmental transitions, and specification of cell fates within a population (19, 64, 80). To identify regulatory functions that control *pks* gene expression, we surveyed induction of *pks* gene expression in several strains carrying gene deletions for regulators that control antibiotic production, developmental transitions, or nutrient stress response. We used liquid cultures of *B. subtilis* for standardized comparison of *pks* expression levels between mutant strains. Several strains were compared to wild type for levels of *pks* gene expression (Fig. 2-6). A moderate reduction was observed with Δ*degU* and Δ*comA* strains, which showed disruption of induced *pks* gene expression as cells transition from exponential to stationary phase (OD_{600} = 1.8) (Fig 2-6A). In contrast, moderate elevation of *pksC* expression was found for the intermediate sample (OD_{600} = 0.9) in the Δ*scoC* mutant strain, which subsequently failed to reach wild-type levels of transcript at high cell density. Reproducibly, the Δ*spo0A* and Δ*codY* strains had the lowest detectable level of *pksC* expression compared to wild type, suggesting that *pks* gene expression requires dual activation through CodY and Spo0A.

Spo0A represses transcription of AbrB, which controls multiple antibiotic biosynthesis pathways and other transition state processes in *B. subtilis* (68, 81–83).

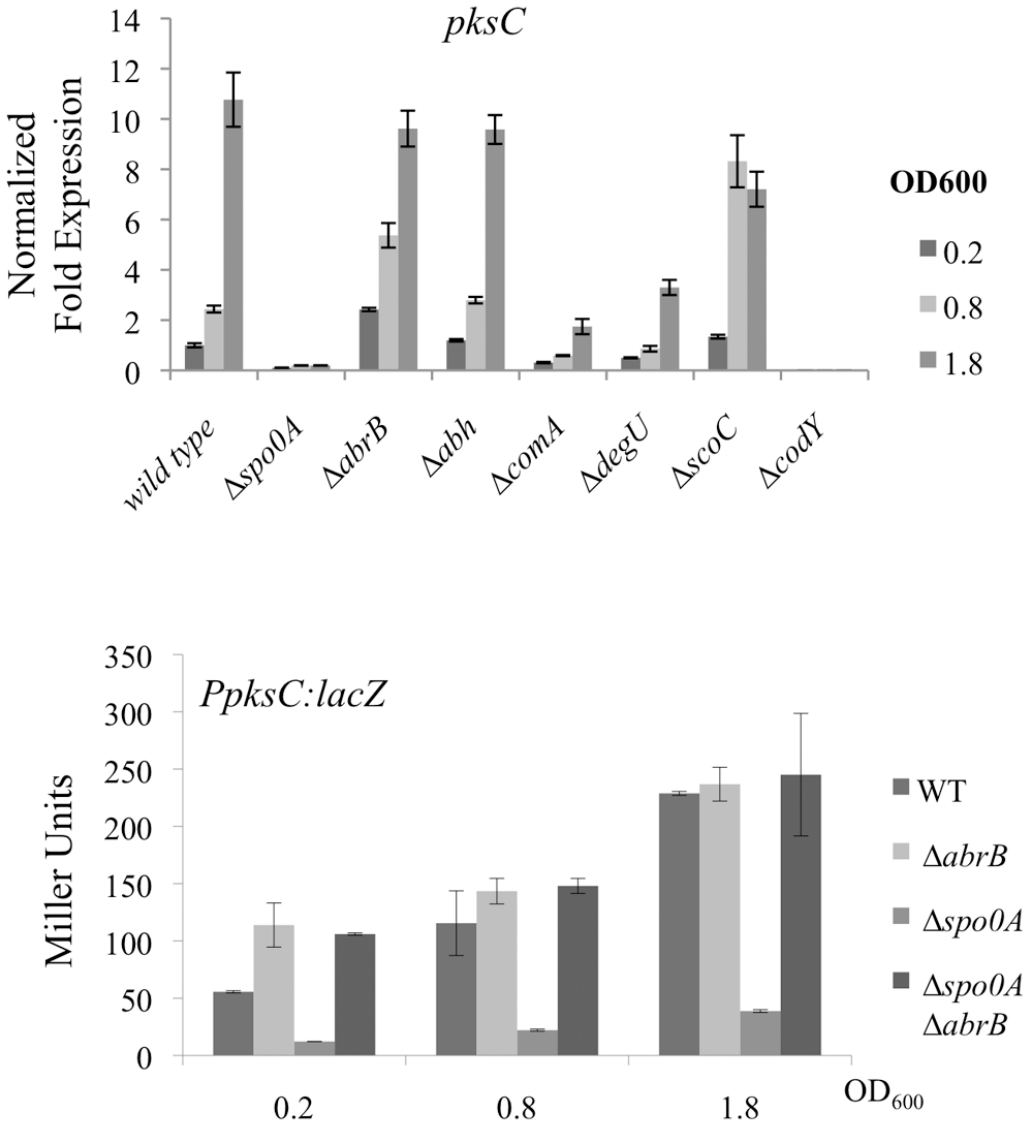


Figure 2- 6. Regulatory pathways for *pks* gene expression.
(A) Quantitative RT-PCR of *pksC* in liquid cultures of strains: Δ*spo0A,* Δ*abrB*, Δ*abh,* Δ*comA,* Δ*degU,* Δ*scoC* and Δ*codY*. Results of qRT-PCR reported as described for Figure 2-3, panel D. Induction of *pksC* expression is reduced in Δ*spo0A*, Δ*comA*, Δ*degU* and Δ*codY* strains. The Δ*abrB* maintains *pksC* induction. Average values and standard deviation from triplicate independent experiments are reported. (B) β-galactosidase assay of P_{pksC}-*lacZ* activity in the single mutants, Δ*spo0A*, Δ*abrB*, and the double mutant Δ*spo0A* Δ*abrB*. Cells were cultured in liquid CH medium at 37°C and cellular equivalents compared from samples taken at the indicated OD_{600}. Deletion of *abrB* restored P_{pksC} activation to the Δ*spo0A* strain. Miller units are averaged from triplicate experiments with standard deviation reported.

However, a Δ*abrB* strain showed a *pks* gene expression pattern similar to wild type (Fig. 2-6A). Because the Δ*spo0A* strain disrupted *pks* gene expression, we tested whether Δ*spo0A*-dependent block to *pks* expression requires AbrB by determining the level of *pks* gene expression in a Δ*spo0A* Δ*abrB* double mutant strain (Fig. 2-6B). To do this, we used the P_{pksC}-*lacZ* strain in order to accommodate existing markers for strain construction. Strains with deletions of the *spo0A* and *abrB* genes, individually and in combination, were used to quantitate *pksC* promoter activity by β-galactosidase assay. The deletion of *abrB* in a Δ*spo0A* background restored promoter activity of *pksC* at all time points. We conclude that *pks* expression is activated by Spo0A through repression of AbrB, a pattern shared by several *B. subtilis* gene clusters encoding antibiotics (19, 68).

Heterogeneous PpksC activity in liquid cultures of B. subtilis

Spo0A and CodY are stationary-phase regulators with functions that intersect with DegU, ComA, and ScoC-dependent processes, including transitions between motile populations, antibiotic production, extracellular matrix production, and sporulation (66, 70, 84–87). We generated P_{pksC}-*yfp* reporter strains that also encode fusions of *cfp* to reporters for motility, extracellular matrix, and sporulation to determine whether the *pks* genes are coordinately controlled by these pleiotropic regulators during the switch between static and spreading populations. We used a fusion of the promoter for the *hag* gene (P_{hag}-*cfp*), which encodes the principal flagellar protein, to indicate δ^D-dependent motile cells (80, 88). A P_{tapA}-*cfp* fusion was used to indicate biofilm matrix-producing

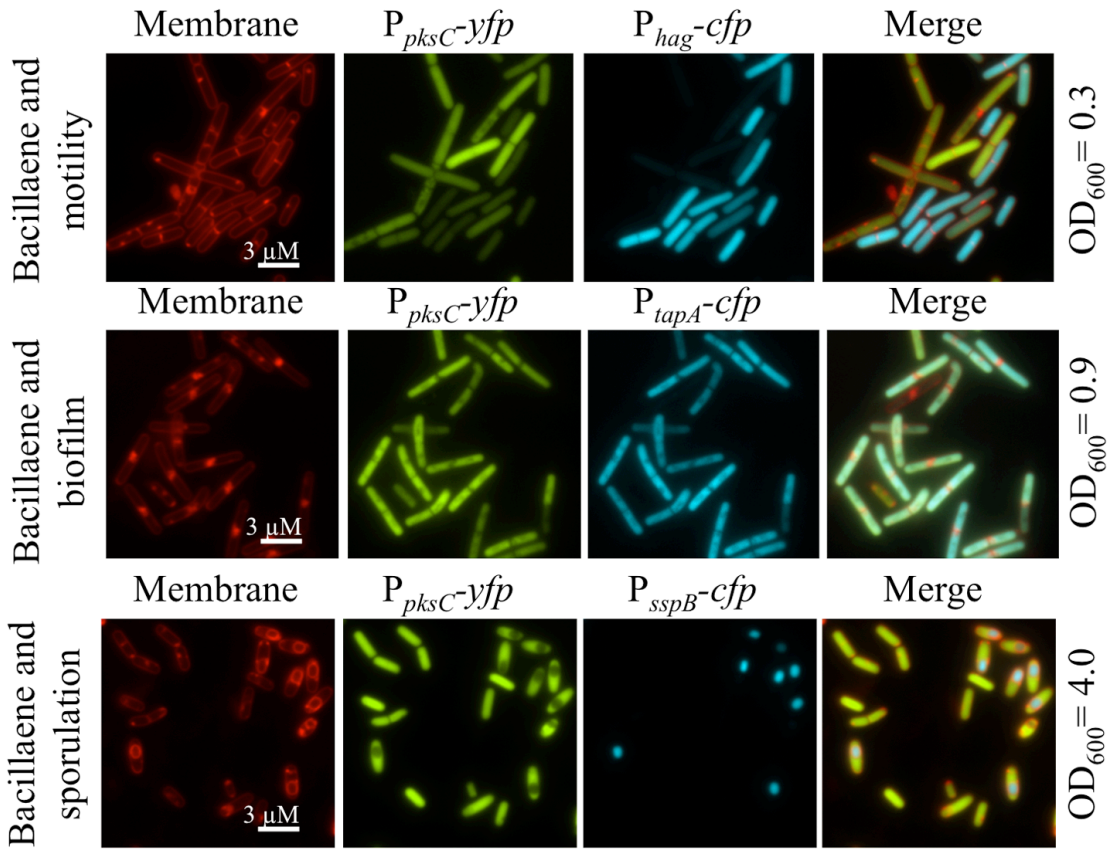


Figure 2- 7. Activation of the *pksC* promoter coincides with activation of biofilm and spore formation.

Fluorescence imaging of *pks* activation (*amyE*::P_{pksC}-*yfp*) with reporters for motility (*lacA*::P_{hag}-*cfp*), biofilm matrix production (*lacA*::P_{tapA}-*cfp*) and sporulation (*lacA*::P_{sspB}-*cfp*). Cells were cultured in liquid CH medium at 37°C and monitored by fluorescence microscopy over time. Images shown were taken at indicated culture densities to observe activation of the relevant pathway reporters. Flagellum dependent motile cells showed low signal intensity for P_{pksC}-*yfp*. Cells active for matrix production (P_{tapA}-*cfp*) and spore formation (P_{sspB}-*cfp*) activated P_{pksC}-*yfp*. TMA- stained membranes - red. Promoter-cfp fusions - blue. Promoter-yfp fusions - green. Scale bar = 3 µm.

subpopulations. The *tapA* gene encodes a component of the biofilm extracellular matrix and is dependent on Spo0A repression of AbrB for activation (68, 80, 89). In addition to P_{hag}-*cfp* and P_{tapA}-*cfp*, we used P_{sspB}-*cfp* to monitor sporulating cells, which are indicative of highly phosphorylated Spo0A and CodY derepression under conditions of nutrient depletion (57, 80, 90, 91). Fluorescence microscopy was used to examine promoter activities at cell densities associated with induction of the respective pathway-specific reporters. As seen in Figure 2-7, the cells within each field show heterogeneous intensity of P_{pksC}-*yfp* fluorescence, suggesting differential *pks* expression in distinct subpopulations of cells in a liquid culture (57). The observed pattern of P_{pksC}-*lacZ* activation suggested that *pks* gene expression is repressed in motile cells expressing P_{hag}-*cfp*. Conversely, cells expressing P_{tapA}-*cfp* also showed elevated levels of P_{pksC}-*yfp*. Thus, the bacillaene operon appears to be induced in matrix-producing populations and not in motile subpopulations when *B. subtilis* is grown in liquid culture. The observed pattern is consistent with a pattern of Spo0A-dependent activation and with the P_{pksC}-*lacZ* expression we have observed on agar plates with *S. lividans*. Upon starvation at high cell density, Spo0A is highly phosphorylated and induces sporulation (92). We examined the level of P_{pksC}-*yfp* fluorescence in strains expressing P_{sspB}-*cfp* as a marker of sporulation. The P_{pksC}-*yfp* signal was detectable in the majority of cells at high cell density (OD_{600} = 4.0). In a percentage of the cells, nascent spores were visible by fluorescence of both TMA-stained membranes and the P_{sspB}-*cfp* reporter (Fig. 2-7). Within the visibly sporulating population, P_{pksC}-*yfp* reporter expression was restricted to the mother cells and not in developing spores. These observations are consistent with activation of *pks*

gene expression during the transition from exponential growth to stationary phase, processes controlled by the master regulatory protein, Spo0A, and the nutrient status regulator, CodY (91, 92).

PpksC is homogeneously repressed upon spreading of B. subtilis colonies

Based on the observed patterns of P_{pksC}-*yfp* activation in liquid culture, we asked whether *pks* gene expression on solid surfaces was heterogenous and exclusive to biofilm matrix-producing cells but not motile cells. We cultured the P_{hag}-*cfp*, P_{pksC}-*yfp* and P_{tapA}-*cfp*, P_{pksC}-*yfp* strains on agar media with *S. lividans* (Fig. 2-8). We observed in these strains that P_{pksC}-*yfp* activation was uniformly low in the spreading population and high in the parent colony, as was also observed with the P_{pksC}-lacZ reporter. The patterns of P_{hag}-*cfp* and P_{tapA}-*cfp* expression, on the other hand, were heterogenous within these populations. This expression pattern indicated that activation of *pks* gene expression is not co-regulated with matrix production per se, which we inferred from its coincidence with P_{tapA}-*cfp* in liquid cultures. Nor is P_{pksC}-*yfp* strictly repressed in the P_{hag}-*cfp* marked motile cells. Instead, the level of *pks* gene expression is largely determined by differentiation of the spreading population from a static colony. In an effort to define the type of motility observed in these assays, we determined that the spreading population is dependent upon both surfactin (Δ*srfAA*) and δ^D (Δ*sigD*) for motility (Appendix I. Fig. S2A). Thus, we think the cells are using swarming motility, which requires δ^D for expression of flagellar genes, to propel themselves toward the *S. lividans* (72). Colony

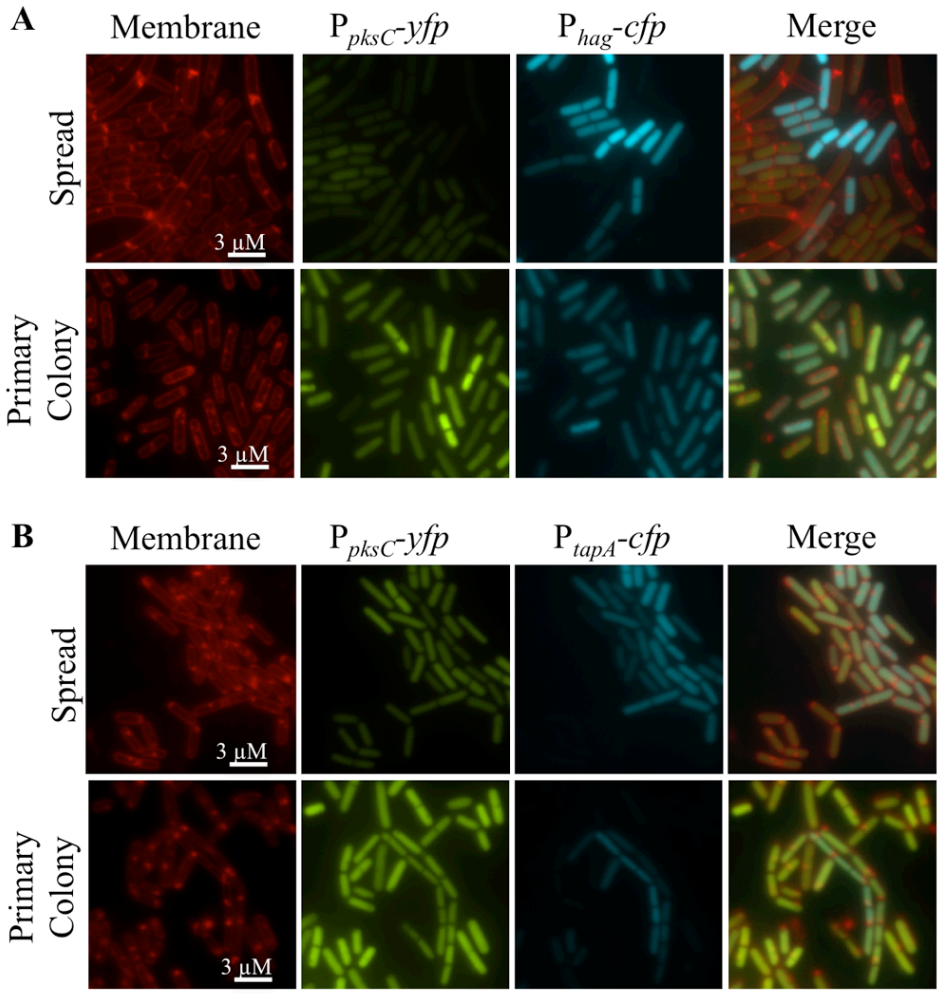


Figure 2- 8. Differential *pks* genes expression of spreading and static cells of *B. subtilis* in competition with *Streptomyces lividans*.
B. subtilis and *S. lividans* co-cultures were prepared as described in Fig. 2-2. Following 36 hours of incubation, cells from the leading edge and primary colony of *B. subtilis* were scraped from the agar and prepared for fluorescence microscopy. (A) A reporter strain for *pks* activation (*amyE*::P$_{pksC}$-*yfp*) and flagellar expression (*lacA*::P$_{hag}$-*cfp*). Low levels of P$_{pksC}$-*yfp* activity were detected in the spreading population, while P$_{pksC}$-*yfp* activity was elevated in *B. subtilis* cells from the primary colony. The spreading populations had a subpopulation of cells with high levels of P*hag-cfp* activity that were negative for P*pksC-yfp* activity. (B) A reporter strain for *pks* activation (*amyE*::P$_{pksC}$-*yfp*) and extracellular matrix (*lacA*::P$_{tapA}$-*cfp*). Lower levels of P$_{pksC}$-*yfp* activity were also detected in the spreading population, compared to the primary colony. Expression of the *pks* genes occurred broadly into the population and not exclusively into the subpopulation of producers of extracellular matrix. TMA- stained membranes - red. Promoter-*cfp* fusions - blue. Promoter-*yfp* fusions - green. Scale bar = 3 µm.

spreading also occurs by spontaneous mutation or targeted disruption of competence and DNA metabolism genes (93). We found that the spreading cells in our assays are not formed of spontaneous mutants, suggesting that the presence of *S. lividans* either results in swarming motility or directional growth by *B. subtilis* (Appendix I. Fig. 2B). In either case, the expression of the *pks* genes is minimal upon emergence of the swarming population. The pattern of transient repression during motility and activation in the parent colony is consistent with complex control of *pks* gene expression by regulators that converge on switching between motility and stationary-phase functions, including the energy and extracellular responsiveness of CodY and Spo0A.

Discussion

Bacillaene is an important determinant of outcomes during interactions between *B. subtilis* and competitor species. Bacillaene is essential for survival in competition with the predatory-like species *Streptomyces* sp. Mg1 (49). Also, the presence or absence of bacillaene influences how a competitor responds to *B. subtilis*, as illustrated by the induction of prodiginines (RED) by *S. coelicolor* and as found in the present study by *S. lividans* (37). The present study directly addressed *pks* gene regulation and the control of bacillaene production by *B. subtilis*. We took a multistep approach to identify regulatory functions that control bacillaene production and considered several existing transcriptomic studies that suggest modes of bacillaene regulation (67, 69, 83, 94, 95). We first determined that *B. subtilis* in liquid culture induces transcription of the >74kb *pks* operon as the cells exit exponential growth and transition to stationary phase. As the

cultures progressed into stationary phase, the production of bacillaene was diminished. The level of *pks* transcripts decreased during stationary phase culture and the Pks megacomplexes were degraded as observed by fluorescence of PksR-YFP. Thus a similar pattern in liquid culture of induction and subsequent reduction is apparent for *pks* transcript levels, PksR abundance, and the presence of secreted bacillaene. This pattern suggests transcriptional regulation is a primary determinant for bacillaene production, as opposed for instance, to activation or deactivation of enzymatic assembly lines.

The regulation of *pks* gene transcription was previously assigned to the PksA protein, annotated as a TetR-family regulator of *pks* genes. Our results indicate that PksA does not regulate *pks* gene expression, at least under the experimental conditions we tested. Our preliminary data suggests that the function of PksA may be directed toward an adjacent, divergently transcribed gene, *ymcC* (not shown). Related organisms such as *Bacillus amyloliquefaciens* FZB42 also produce bacillaene, encoding the enzymatic complex in the *bae* gene cluster (38). In contrast to *B. subtilis*, the orthologous *pksA* gene of *B. amyloliquefaciens* FZB42 is located in a region of the chromosome separate from the *bae* biosynthetic gene cluster, suggesting the protein is not a pathway-specific regulator (39). Thus these data support a model for *pks* gene regulation and bacillaene production that relies on global regulatory circuits and not a pathway-specific regulator such as PksA.

Evidence for differential activation of *pks* gene expression emerged from the interaction *of B. subtilis* with *S. lividans*, which suggested bacillaene is repressed in spreading subpopulations. To understand regulatory processes that control bacillaene

production, we focused our attention on regulation of the *pks* operon that extends from *pksC* to *pksR*. We identified the 5'UTR of *pksC* as the control element for induction of the *pks* operon, which is consistent with results from a genome-wide study of *B. subtilis* transcription (67). Using P_{pksC}-*yfp* and P_{pksC}-*lacZ* reporters, we confirmed that *pks* gene expression was transiently repressed in populations of cells that migrate across agar toward *S. lividans*. We have shown that bacillaene production is principally under the control of the Spo0A and CodY stationary phase regulators. However, full induction of *pks* gene expression is also dependent on DegU, ComA, and ScoC, suggesting that *B. subtilis* uses multiple mechanisms to integrate bacillaene synthesis with other cellular functions (64, 96–98). The observed patterns of *pks* gene expression suggest that *B. subtilis* activates and represses bacillaene production in response to nutrient conditions and developmental transitions. Similar observations for regulation of several *B. subtilis* antibiotics in liquid culture have been described (19, 60, 65, 66, 68, 72, 99–101)

Our results support a model wherein *B. subtilis* inactivates bacillaene production during a motile phase induced by growth in the presence of *S. lividans*. The parent colony actively expresses the *pks* genes. The swarming population initially repressed *pks* expression, which over time becomes active within the motile populations. The described pattern of synthesis is consistent with the observation that Spo0A, which controls a switch between motile and biofilm matrix-producing cells, are required for full induction of the *pks* genes (80). Additionally, the loss of *pks* gene expression in the Δ*codY* strain revealed that the *pks* operon is one of a few targets dependent upon CodY for expression (102). CodY binds directly to GTP and branched-chain amino acids

(BCAAs) as a mechanism for sensing nutrient-rich conditions (91, 103). When bound to these signals, CodY represses functions that include motility (e.g. *hag* and *fla-che*) and antibiotic production (e.g. surfactin and bacilysin) through increased affinity for their regulatory DNA elements (87, 104). The Δ*codY* phenotype for *pks* gene expression suggests that the *pks* operon is tuned to changes in nutrient availability, and is repressed when cells divert resources to motility. We speculate that maintaining low *pks* gene expression in motile populations may be important for energy resource allocation in *B. subtilis*, because synthesis of megacomplexes and bacillaene are likely to require considerable energy input.

Our results suggest a hypothesis that dual regulation of *pks* expression may ensure bacillaene production in high cell-density populations, such as biofilms (Spo0A-dependent), and under conditions of nutrient availability that permit *pks* genes expression (CodY). In addition, secondary control mechanisms may fine tune expression in response to external and internal conditions or signals through DegU, ComA, and ScoC. For example, DegU is induced by the presence of glucose through catabolite regulator function (105). Thus, a DegU-dependent timing mechanism may exist for initial repression by dephospho-DegU, followed by activation through DegU phosphorylation and inhibition of motility (106). ComA and ScoC also function in both motility control and secondary metabolism. ComA regulates surfactin production, which is required for swarming motility, and DegQ, which enhances DegU phosphorylation (95, 107). ScoC, which directly controls bacilysin production, also controls the transition between motile and biofilm-forming populations through its regulation of *flgM* and *sinI*,

respectively (66, 98, 108). Thus, ScoC may coordinate the timing of induction for *pks* gene expression, which we found is activated early in a Δ*scoC* mutant strain. Other, unidentified regulatory functions may contribute to *pks* control as well. How the relative timing of interactions between regulatory pathways integrates downstream functions is complex and incompletely understood.

Exploring the divergent functions in static and motile populations of competing bacteria provides an experimental system to better understand pathway integration. Based on our results, synthesis of bacillaene enters a growing list of processes that are divided among subpopulations of clonal *B. subtilis* cells (109). Regulatory mechanisms for antibiotic biosynthesis have been generally studied by culturing bacteria in liquid and following patterns of synthesis correlated with cell density. As we have shown, the production of bacillaene in liquid culture fits a pattern typical of many antibiotics, i.e. induction upon transition out of logarithmic growth. Historically, this timing defines antibiotic production within the idiophase (3). When grown on solid surfaces, patterns of differentiation suggest that many bacteria have sophisticated mechanisms for determining the timing of pathway activation (80). Antibiotic biosynthesis is no exception. Developmental regulatory processes also control antibiotic biosynthesis (19). For example, many *Streptomyces* species couple antibiotic biosynthesis with developmental pathways of aerial growth and sporulation through complex regulatory circuits (17). In some cases, antibiotic biosynthesis is activated by pathway specific regulatory genes that may be directed by developmental regulators (27). When no

pathway specific regulatory proteins are present, identifying the specific determinants of activation requires understanding the developmental control networks of the organism.

Studies of antibiotic regulation under competition suggest that coordinated control of multiple antibiotics with developmental transitions serves to optimize the competitive fitness of *B. subtilis* by ensuring efficient resource allocation (110). The convergence of bacterial developmental regulation with control of antibiotic synthesis may highlight new approaches to activate or optimize production of molecules of interest (111). For example, strategic genetic manipulations of the producing species could be used to restrict the organism to a high antibiotic output state, which may also be an effective approach to activate cryptic secondary metabolic pathways (112). Future examination of developmental controls for antibiotic biosynthesis will likely inform key principles of bacterial competition as well as new strategies to induce antibiotic production from microorganisms.

CHAPTER III

REGULATORY FEATURES OF THE *pks* GENE CLUSTER

Summary

Gene operons are common in bacteria to coordinate the expression of multiple genes that participate in a similar metabolic pathway or other concerted processes. In addition to regulatory proteins that initiate the transcription, numerous regulatory elements are usually present along the operons to elongate transcription of the distal genes. The *pks* genes, from *pksB* to *pksR*, span nearly 76 kb and form the largest operon in the *B. subtilis* genome. Several promoters and transcription termination sites along the *pks* genes have been predicted but not tested. In this chapter, we present evidence about the activity of the putative promoters upstream of *pksB*, *pksC* and *pksS* (P_{pksB}, P_{pksC} and P_{pksS}). Our data suggest that P_{pksC} is the most active promoter and likely controls the expression from *pksC* to *pksR* (~75 kb) during vegetative growth. A promoter candidate at the intergenic region between *pksC* and *pksD* did not show any activity. Instead, our results indicate that the *pksC* to *pksD* intergenic region is a novel *cis*-RNA element that promotes read-through of terminations sites in the *pks* genes. This a novel mechanism of genetic control not previously reported for a natural product encoding gene cluster. Our results serve as a model for antibiotic regulation in many organisms with similar long operons. We also report the activation of an internal promoter at *pksG* (P_{pksG}) during sporulation that suggests differential *pks* gene expression at this developmental stage.

Introduction

Gene expression requires a series of concerted regulatory activities at different levels of transcription. Regulatory proteins that either block or modulate RNA polymerase activity are only the first step of the control of gene transcription (113)(Chapter II). In addition, the RNA polymerase must recognize regulatory elements in the chromosome, called promoters, to initiate transcription (114). Elongation of transcription is accomplished by either the recruitment of elongation factors or antitermination systems that promote read-through at transcription termination sites (115). Finally, cells stop transcription by two different mechanisms, Rho-dependent termination and Rho-independent termination. In a Rho-dependent termination mechanism, a protein called Rho destabilizes the complex formed by the RNA polymerase and the DNA template. In a Rho-independent termination, also known as intrinsic termination, a G-C-rich hairpin loop precedes a poly-U stretch in the nascent RNA transcript and forms a mechanical stress that pulls the poly-U stretch out of the RNA polymerase active site (116, 117). Intrinsic terminator hairpins (Rho-independent terminators) are the predominant mechanism of transcription termination in *B. subtilis* and other Firmicutes (67, 118). Additional genetic features such as operators and enhancers may also contribute to regulation of gene expression.

Bacterial operons are organized to coordinate simultaneous expression of several genes through multiple overlapping genetic controls (119, 120). Operons are composed of two or more open reading frames that are transcribed as a single mRNA, spanning from a few up to several hundreds of kilobases (114). Due to the length of many of these

operons, numerous regulatory strategies are necessary beyond transcription initiation to extend the expression of every gene of the operon (121). Transcription of the operon genes usually initiates at a promoter upstream of the proximal gene of the operon. Then, other regulatory sequences along the operon may have auxiliary roles in transcribing distal genes. For example, internal promoters are usually present along the operon to contribute to increase levels of particular distant genes (122). *Cis*- acting regulatory RNAs may also serve as regulatory elements to induce transcription of entire operons (120).

It is estimated that 4.1% of *B. subtilis* genes are controlled, at least in part, by RNA-mediated genetic control strategies (121). *Cis*-non-coding small RNAs located at the 5' end or within intergenic regions of long operons may adopt structural configurations that affect transcriptional and post-transcriptional regulation (123). Most often, these *cis*- acting regulatory RNAs may affect access to ribosome binding sites, mRNA stability or transcription termination signals (124). *Cis*- RNA elements may control intrinsic transcription termination by promoting the read-through of termination sites, also known as processive antitermination ((120) and references within). One example of these antitermination systems is the EAR element located at the intergenic region of the *eps* operon in *B. subtilis* (120). The *eps* operon encodes the synthesis of exopolysaccharides required for biofilm formation and it spans nearly 16 kb. The EAR element is located in the intergenic region between *epsB* and *epsC* and provides read-through of distally located termination sites to ensure the synthesis of the entire *eps* transcript (120).

The *pks* genes, from *pksC* to *pksR*, span nearly 75 kb and form the largest operon in the *B. subtilis* genome. Predictive bioinformatic analysis of the *pks* gene cluster reveals several candidate regulatory elements of the operon (Fig. 3-1). The Subtiexpress database reports upshifts in transcription at *pksB*, *pksC* and *pksS*, suggesting the presence of promoter elements in these intergenic regions (http://subtiwiki.uni-goettingen.de/apps/expression/) (Chapter II). Additional promoters are predicted by Subtilist at the intergenic region of *pksC* and *pksD*, and in the overlapping region between *pksF* and *pksG* open reading frames (http://genolist.pasteur.fr/SubtiList/). Subtilist also predicts a transcription termination site upstream of *pksG*. The regulatory element organization described in Subtiexpress suggests that the *pks* genes are organized in a single operom from *pksC* and *pksR*. In contrast, Subtilist predicts one operon from *pksC* to *pksF,* and a second operon from *pksG* to *pksR*. The presence of these putative *pks* regulatory elements suggests that the control of *pks* gene expression is complex. None of the predicted regulatory elements have been investigated prior to this work.

In chapter II, we investigated the regulatory proteins that coordinate the initiation of transcription of the *pks* genes at the entry into stationary phase. The question of which regulatory elements determine the length of the *pks* transcript remains to be elucidated. In this chapter, we investigate the regulatory features present in the *pks* gene cluster and their role in the expression of the *pks* genes. In agreement with the predictions, we have found that the regions upstream *pksB, pksC* and *pksS* are active promoters during vegetative growth that contribute to the independent expression of *pksB* and *pksS*, respectively. Instead, *pksC* promoter appears to control the expression of the genes from

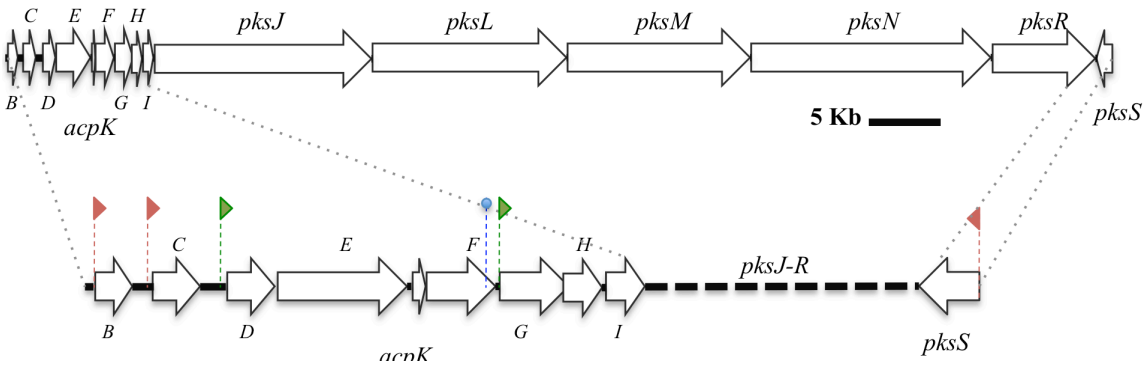


Figure 3- 1. Predicted promoters and terminators of the *pks* gene cluster in *B. subtilis*. The *pks* genes from *pksB-pksR* are about 76 kb long, as annotated in the *B. subtilis* str. 168 genome. *pksA* has been removed from the illustration as our data indicates that PksA has no role in the synthesis of bacillaene. Above, arrows are drawn to scale. Below, there is a magnification of the genes where additional features are predicted (Not to scale). Red flags indicate predicted promoters according to Subtiexpress database (http://subtiwiki.uni-goettingen.de/apps/expression/). Green flags indicate additional promoters predicted by Subtilist (http://genolist.pasteur.fr/SubtiList/). The feature in blue is predicted by Subtilist as a transcriptional terminator upstream of the end of pksF.

pksC to *pksR*. We also found that *pksG* promoter is highly active in sporulating but not vegetative cells, suggesting differential *pks* gene expression during sporulation. Also, our data indicates that the predicted termination site upstream of *pksG* does not stop transcription elongation, at least under the conditions tested. However, abundance of the specific transcript between *pksF* and *pksG* was low and multiple non-specific bands were present, suggesting that the predicted termination site upstream of *pksG* produce some level of transcription interference. We also report what appears to be a novel *cis*-regulatory RNA located at the intergenic region between *pksC* and *pksD* that promotes processive antitermination of the *pks* operon.

Material and methods

Bacterial strains, primers, media and growth conditions

All the experiments of this study were performed with the undomesticated strain *Bacillus subtilis* NCIB 3610. Table 3-1 contains a complete list of the strains used in this work. Transcriptional reporters P_{pksB}-*yfp*, P_{pksD}-*yfp*, P_{pksG}-*yfp*, and P_{pksS}-*yfp* were constructed as described in chapter II for P_{pksC}-*yfp,* using the primer pairs pB-F(EcoRI)/pB-R(HindIII), pD-F(EcoRI)/pD-R(HindIII), pG-F(EcoRI)/pG-R(HindIII) and pS-F(EcoRI)/pS-R(HindIII), respectively (Table 3-2). CH medium was used for vegetative growth and 37°C was the default temperature, unless otherwise stated.

Fluorescence microscopy

Samples from shaken liquid cultures were taken for fluorescence imaging,

Table 3- 1. *B. subtilis* strains used in this study

Strain	Relevant Genotype	Source
PDS0066	*B. subtilis* 3610 wild type	Laboratory collection
PDS0067	*B. subtilis* 3610 Δ*pksB-R*	Laboratory collection
DS7817	*B. subtilis* 3610 wild type	Winkler Lab
GEE5174	*B. subtilis 3610 ΔpksC-D IR*	Winkler Lab
GEE5175	*B. subtilis 3610 pksC-D IR:: EAR*	Winkler Lab
PDS0525	*B. subtilis 3610 ΔcodY*	Laboratory collection
PSK0209	*B. subtilis* 3610 *pksE-cfp*	Laboratory collection
PDS0032	*B. subtilis NCIB3610 amyE:: P_{pksB}-yfp (cat)*	Laboratory collection
PDS0036	*B. subtilis NCIB3610 amyE:: P_{pksC}-yfp (cat)*	Laboratory collection
PDS0204	*B. subtilis NCIB3610 amyE:: P_{pksD}-yfp (cat)*	Laboratory collection
PDS0207	*B. subtilis NCIB3610 amyE:: P_{pksG}-yfp (cat)*	Laboratory collection
PDS0035	*B. subtilis NCIB3610 amyE:: PpksS-yfp (cat)*	Laboratory collection

Table 3- 2. Primers used in this study

Name	Sequence (5'-3')
pB-F(EcoRI)	ctagaattcctgagagactttacgc
pB-R(HindIII)	attcaagcttatcatgtaaagttcttaaac
PpksD_RI_F	atgaattcaccacatccagagcacca
PpksD_HIII_R	gaataagcttcccatgtgatagtatt
PpksG_RI_F	atgaattcatcaaaaatgcgctgagc
PpksG_HIII_R	gaataagcttccggattcagagcaag
pS-F(EcoRI)	aatgaattcgcgctaatagggtaaataga
pS-R(HindIII)	tataaagcttgctatacgcagtacgaatc
qPCR-pksR-F	acagcgtaacggaattttgg
qPCR-pksR-R	ttgattgcccttccttatcg
gyrB qPCR_F	gggcaactcagaagcacggacg
gyrB qPCR_R	gccattcttgctcttgccgcc
eRT-pksJ-R	agtcctagtcgcagaccacgg tcctgatgcatcccccggtcc
eRT-pksI-F	cggcaatgcagggccatgga
eRT-pksH-F	tcagtcgtgcgaaagccaccg
eRT-pksG-F	ggtctcaaccgcaactgccgt
eRT-pksF-F	cttcctcgctgctgtccggc

centrifuged at 8,000 rpm and washed once with PBS. Cells were resuspended in 20 μM 1-(4-trimethylammoniumphenyl)-6-phenyl-1,3,5-hexatriene p-toluenesulfonate (TMA) (Molecular Probes) and fluorescence images were captured using a Nikon Ti-E inverted microscope equipped with a CFI Plan Apo Lambda DM 100X objective, TI-DH Diascopic Illuminator, and a CoolSNAP HQ2 Monochrome Camera. Exposure time was 1,000 ms for TMA and 2000 ms for YFP. The NIS-elements AR software was used to capture and process the images for comparative analysis. Pseudocoloring within a range of 600-4600 LUTs was applied to YFP (green) images for P_{pksB}-*yfp*, P_{pksD}-*yfp*, P_{pksG}-*yfp*, and P_{pksS}-*yfp*. However, due to the intensity of the P_{pksC}-*yfp* signal, the range applied for this promoter was from 600-16000 LUTs.

Induction of sporulation

Cells growing in CH medium were taken at early exponential phase (OD_{600} = 0.5-0.8), centrifuged and resuspended in warm sporulation medium as described by Sterlini and Mandelstam (125). The initiation of sporulation was designated as S0, the time of resuspension. Samples were harvested at hourly intervals for eight hours (S0-S7).

Germination and outgrowth efficiency

B. subtilis strains were induced to sporulate in DSM plates for 48 h at 37°C and then collected and treated with lysozyme to eliminate vegetative cells, as previously

described (71). The purity of the spore collection was 98% as evaluated by phase-contrast microscopy. Spores were heat-activated at 80˚C for 20 min, put on ice for 20 min and incubated at 37˚C in LB broth (initial OD_{600} = 0.3). Optical density (OD_{600}) was measured every five minutes during germination and every ten minutes during outgrowth. Each experiment was done in triplicates.

Quantification of bacillaene

Production of bacillaene was determined as described in Chapter II. Briefly, three early exponential growth cycles were done to generate a uniform population before the beginning of the experiment. Then, cell cultures were grown in CH medium, at 30°C, constant agitation (250 rpm) and complete darkness. Samples of 50 ml were taken at OD_{600} of approximately 2.8 and OD_{600} of approximately 4.5. Supernatants of the samples were used to extract and quantify bacillaene. Multiple peaks in the HPLC chromatograms represent isoforms of bacillaene that come at different elution times. As indicated in Chapter II, LC-MS analysis has confirmed that the relevant peaks are all different isoforms of bacillaene. The amount of bacillaene in each sample is determined by integrating the area under the relevant peaks on the elution chromatograph. This experiment was done in duplicates and repeated two independent times

Exact RT-PCR

Cells growing in CH medium were taken at mid exponential growth and stabilized using RNAprotect Bacteria Reagent (Qiagen). RNA isolation was performed

using RNeasy Mini Kit, according to the manufacturer's instructions. In this modification of RT-PCR, DNase treatment is not necessary due to the use of a hybrid primer (pksJ-R) to synthesize cDNA (126). Total RNA (100 ng) was used to perform cDNA synthesis with the pkJ-R hybrid- primer (Table 3-2) and using the ProtoScript AMV LongAmp Taq RT-PCR Kit (NEB). Negative controls were included (No reverse transcriptase) to examine the DNA contamination in the samples Downstream PCR amplification was performed with the cDNA product (1/10 of PCR reaction volume) and specific primers that annealed several hundreds bp upstream of *pksJ* (eRT-pksI-F, eRT-pksH-F, eRT-pksG-F and eRT-pksF-F), as indicated in Table 3-2. The PCR protocol was as follows: denaturation at 95.0°C for 30 s; 30 cycles of denaturation at 94.0°C for 15 s, annealing at 57.0°C for 30 s, extension at 72.0°C for 50 s/kb; and a final hold at 4.0°. PCR amplifications were observed in 0.8% agarose gels.

Quantitative RT-PCR (qRT-PCR)

RNA isolation, DNase treatment, cDNA synthesis and quantification of transcript abundance were performed as described in Chapter II, using the primer pairs listed in Table 3-2 (qPCR-pksR-F/qPCR-pksR-R, gyrB qPCR_F/gyrB qPCR_R). Specifically, to determine the *pksH* transcript abundance, wild type cells were grown in CH medium and induced to sporulate by resuspension in starvation medium at 37˚C. Samples were taken every hour for 8 hour, from the time of resuspension (S0 – S7) and processed for *pksH* mRNA quantification. To determine the functionality of *pksF* termination site, cells of the strains PSK0209 (*B. subtilis pksE-cfp*) and PDS0066 (wild type) were grown in LB

broth and allowed to reach exponential growth. Then cultures were induced for four hours with 1.0 mM IPTG. A control with no IPTG was also used during this experiment. The experiments were done in triplicates, two independent times.

Results

Promoters of the pks operon

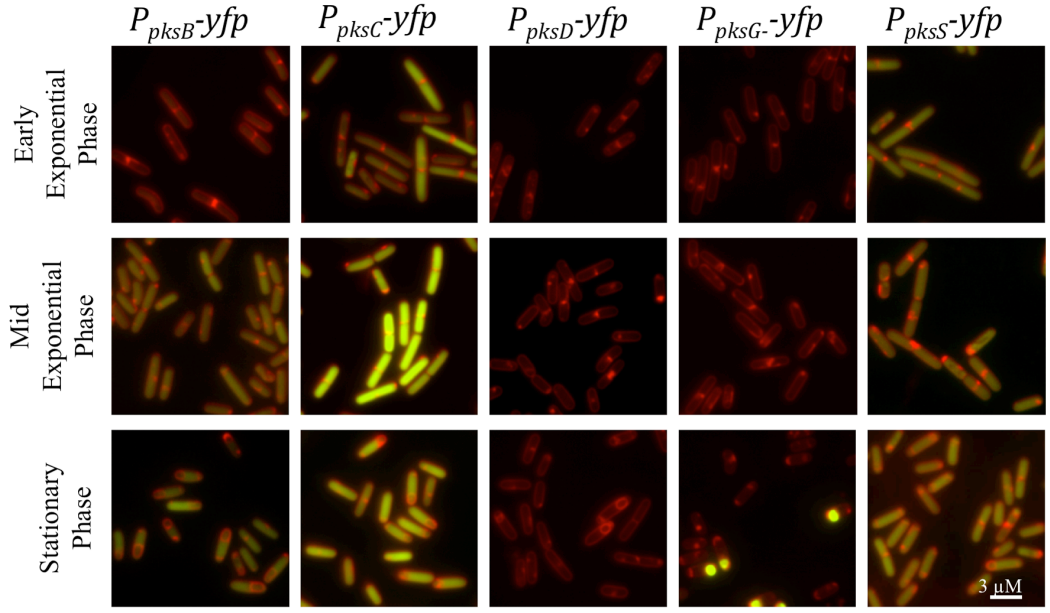
The SubtiExpress database and our own data suggested that the region upstream of *pksC* act as an active promoter in the *pks* gene expression (http://subtiwiki.uni-goettingen.de/apps/expression/) (See Chapter II). However, other transcriptional initiation regions have been predicted to exist in the *pks* gene cluster. As mentioned, SubtiExpress database shows three expression upshifts or promoters within the *pks* genes, namely, P_{pksB}, P_{pksC}, and P_{pksS} (Fig. 3-1) (http://subtiwiki.uni-goettingen.de/apps/expression/) (67). In addition, Subtilist database also predicts one region with -35 and -10 signals upstream of *pksD*, and one region with a -35 signal located 5 bp upstream of *pksG* and a -10 signal located 13 bp downstream of the initiation codon of the *pksG* open reading frame (http://genolist.pasteur.fr/SubtiList/). Therefore, in addition to the *pksC* promoter (P_{pksC}) studied in chapter II, we isolated the 5'UTRs of *pksB* (P_{pksB}), *pksD* (P_{pksD}), *pksG* (P_{pksG}) and *pksS* (P_{pksS}) and fused them to *yfp* for fluorescence microscopy. To determine the activity of these predicted promoters at different stages of growth, we obtained samples at mid (OD_{600} ~ 0.8), late exponential growth (OD_{600} ~ 1.8) and stationary phase (OD_{600} ~3.0), and observed them under the

fluorescence microscope (Fig. 3-2A). Fluorescence signal from P_{pksD} was not detected at any time point studied during this experiment. Low fluorescence signal was detected from P_{pksB} and P_{pksS} at mid exponential growth but the signal increased as the cells reached stationary phase, suggesting P_{pksB} and P_{pksS} are involved in activation of the *pks* genes. Consistent with our previous results, P_{pksC} had the strongest fluorescence signal at all observed time points, indicating that P_{pksC} is the most active promoter in the *pks* operon. To our surprise, P_{pksG} fluorescence signal was not observed at any time during vegetative growth but a strong signal was observed inside the forespores at stationary phase. The strong signal of P_{pksG} suggests that high expression of genes downstream of this promoter is occurring during spore formation. In summary, our data suggest that P_{pksB}, P_{pksC} and P_{pksS} are active promoters of the *pks* operon and reach a peak of activity at the entry into stationary phase. P_{pksG}, is active only during sporulation and its activity seems to be exclusive to the forespore but not the mother cells.

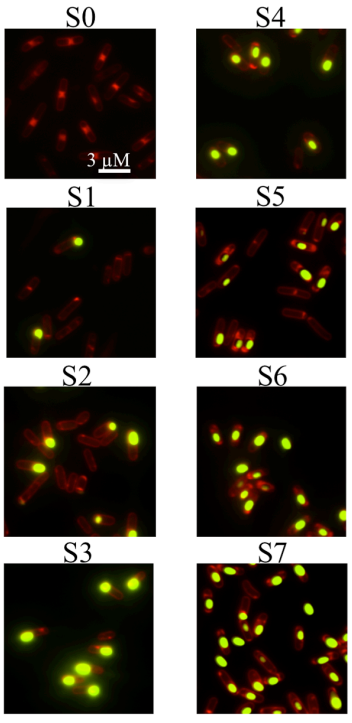
The intriguing results of the *pksG* promoter prompted us to determine the sporulation stage at which P_{pksG} becomes active. We induced spore formation by resuspending in starvation medium a culture of the P_{pksG}-*yfp* strain in mid-exponential growth. We tracked the stages of spore formation by fluorescence microscopy (Fig. 3-2B). As expected, asymmetric division and formation of polar septa occurred at S1.

Figure 3-2. Activity of promoters of the *pks* gene cluster.
Fluorescence signal was analyzed as described in methods and Fig. 2-2. Images represent several microscopic fields from samples of three independent experiments. TMA -stained membranes - red. Promoter-yfp fusions - green. Scale bar = 3 μm. (A) Fluorescence of transcriptional reporters for promoters upstream of *pksB* (P_{pksB}), *pksC* (P_{pksC}), *pksD* (P_{pksD}), *pksG* (P_{pksG}) and *pksS* (P_{pksS}) were fused to *yfp*. Cells growing in liquid CH medium at 37°C were taken at indicated time points to observe activation of the promoters. $P_{pksB,}$ P_{pksC} *and* P_{pksS} showed fluorescence signal at all times. P_{pksG} fluorescence signal was observed inside forespores at late stationary phase. P_{pksD} fluorescence signal was not observed at any time. (B) Fluorescence of P_{pksG} during spore formation. Spore formation was induced by resuspension of exponential growing cells in Resuspension Medium. Samples for fluorescence microscopy were taken hourly from S0 – S7. P_{pksG} is active in the majority of forespore-forming cells at S2. (C) Abundance of *pksH* mRNA as measured by qRT-PCR. Wild type cells were induced to sporulate as described in Material and Methods. Fold-expression values reported are relative to S0. The maximal fold expression occurred at S2 and then expression rapidly decreased. The experiment was done in triplicates, at two independent times.

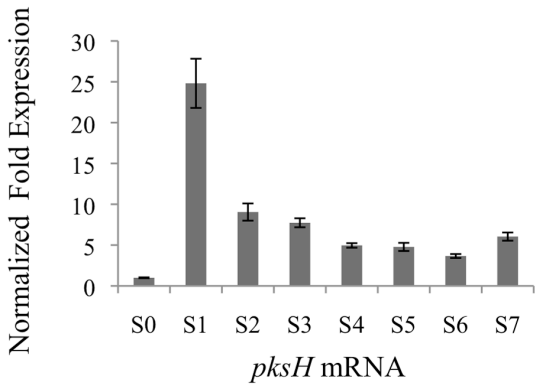
A


P_{pksB}-yfp
P_{pksC}-yfp
P_{pksD}-yfp
P_{pksG}-yfp
P_{pksS}-yfp
Early Exponential Phase
Mid Exponential Phase
Stationary Phase
3 µM

B


S0
S4
3 µM
S1
S5
S2
S6
S3
S7

C


Normalized Fold Expression
30
25
20
15
10
5
0
S0
S1
S2
S3
S4
S5
S6
S7
pksH mRNA

P_{pksG} became highly active soon after formation of polar septum, and P_{pksG}-*yfp* signal remained strong at all stages of sporulation. Given that at late sporulation stages the forespores are already entering dormancy, we suspected that the strong *yfp* signal at the latest stages of sporulation could be caused by the absence of degradation of *yfp* rather than activity of P_{pksG}. Using qRT-PCR, we detected that the abundance of the *pksH* transcript reached a 25-fold increase at S1 but rapidlly decreases over the subsequent phases of sporulation (S2-S7) (Fig. 3-2C). Taken together, these results suggest that compartmentalization of *pksG* downstream gene expression occurs at early stages of spore formation.

The P_{pksG} activation during sporulation is suggestive of a function for bacillaene during either sporulation or germination. In an attempt to determine whether bacillaene is essential for *B. subtilis* spore formation, germination or outgrowth, we investigated the efficiency of sporulation and germination of a Δ*pks* mutant (Fig. 3-3). First, we induced sporulation of the strains by resuspension in starvation medium, as described above. Formation of polar septum and development of spores was similar in the wild-type and the Δ*pks* strains, as observed by phase-contrast microscopy (Fig. 3-3). Heat-resistant spores were also counted on LB agar plates after overnight incubation in starvation medium. No differences in the number of spores that initiated vegetative growth was found (data not shown). Second, to determine the efficiency of germination and outgrowth, spores suspension of wild-type and Δ*pks* strains were incubated in LB broth at 37˚C and OD_{600} was measured over time (Fig. 3-4). No significant differences were

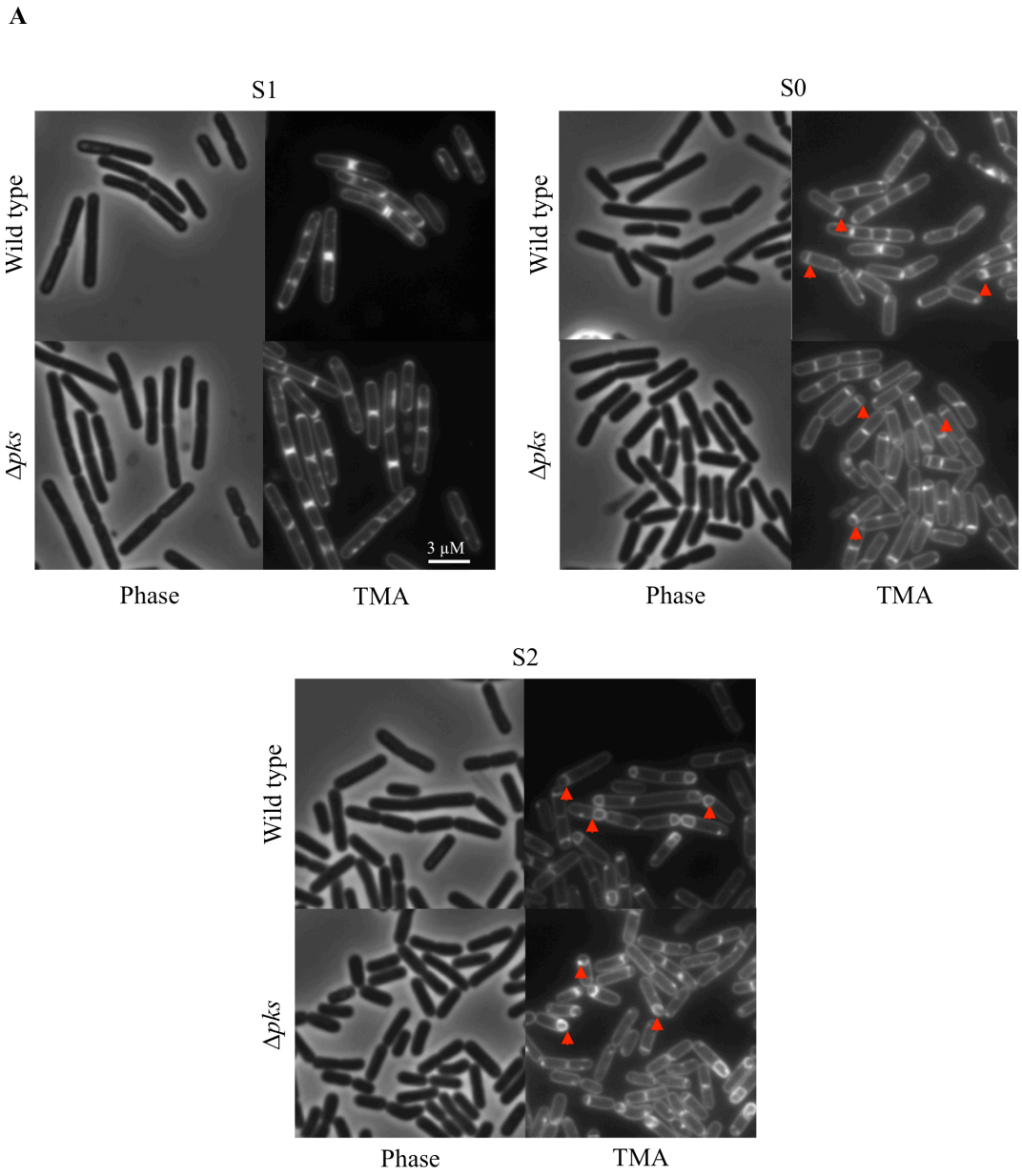


Figure 3- 2. *B. subtilis* sporulation in a Δ*pks* mutant.
Fluorescence microscopy of wild-type and Δ*pks* strains during sporulation. Induction of spores was performed as indicated in Figure 3-2B and methods. Formation of polar septum is observed at the expected stages for both, wild type and Δ*pks* strains (red arrow heads).

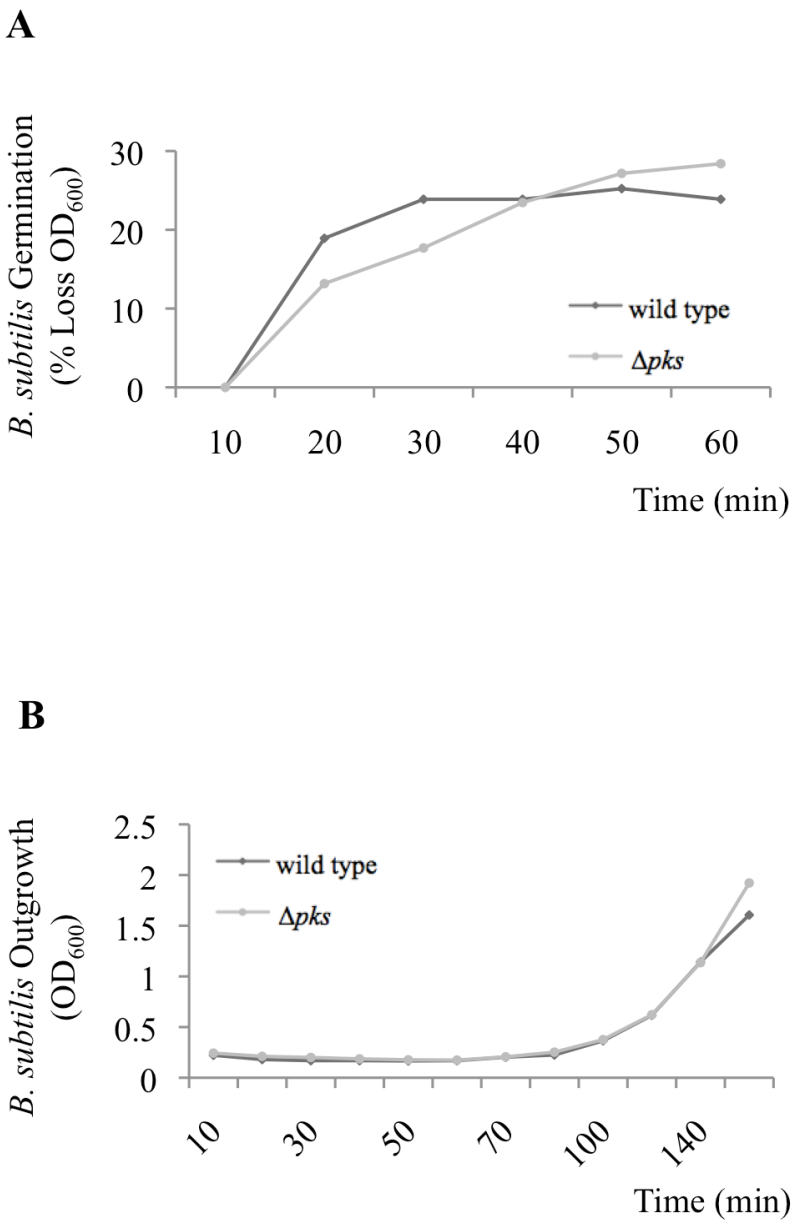


Figure 3- 3. *B. subtilis* germination and outgrowth in a Δ*pks* mutant.
(A) *B. subtilis* germination. The y-axis represents the outgrowth as % of OD_{600} loss given that the optical density decreases during the outgrowth. No significant differences are found in wild type and Δ*pks* outgrowth (B) *B. subtilis* outgrowth in LB medium. Wild type and Δ*pks* spores were transferred to LB medium and incubated at 37˚C. OD_{600} was measured every five minutes during outgrowth and 10 minutes during germination. Changes in OD_{600} were similar in both, wild type and Δ*pks* strain.

found in *B. subtilis* germination and outgrowth between wild type and the Δ*pks* mutant under the conditions tested. Therefore, the role of bacillaene and the differential promoter activation of the *pks* genes downstream of *pksG* remain to be elucidated.

Transcription termination site at pksF

Transcription termination sites are sequence elements necessary to mark the end of single coding sequences and operons. According to Subtilist, the *pks* genes span almost 75 kb from *pksC* to *pksR*, in the positive strand, with a single intrinsic termination site before the end of the *pksF* locus and a bidirectional terminator between *pksR* and divergently transcribed *pksS*. (http://genolist.pasteur.fr/SubtiList/) (Fig. 3-1 and Table 3-3). To visualize whether a transcript truncation occurs at the end of *pksF*, we conducted an Exact RT-PCR analysis from *pksF* to *pksJ*. We predicted that if transcription termination occurs at *pksF* we would detect transcripts from *pksG* to *pksJ* but no transcript would be observed from *pksF-pksJ*. As shown in Figure 3-5A, single amplification products ranging from *pksG* to *pksJ* were observed at the expected sizes. However, amplification from *pksF* to *pksJ* at the expected size (4.4 kb) was not observed, and multiple non-specific bands were visible. In the absence of reverse transcriptase (negative controls) no PCR products were observed (data not shown). These results suggest the presence of a terminator site at *pksF* that prevents the elongation of the transcript from *pksF* to *pksJ*. To discard the possibility that the RT reaction didn't make sufficiently long cDNA, we use a second reverse primer that annealed at *pksG*. We conducted Exact RT-PCR with several forward primers to walk

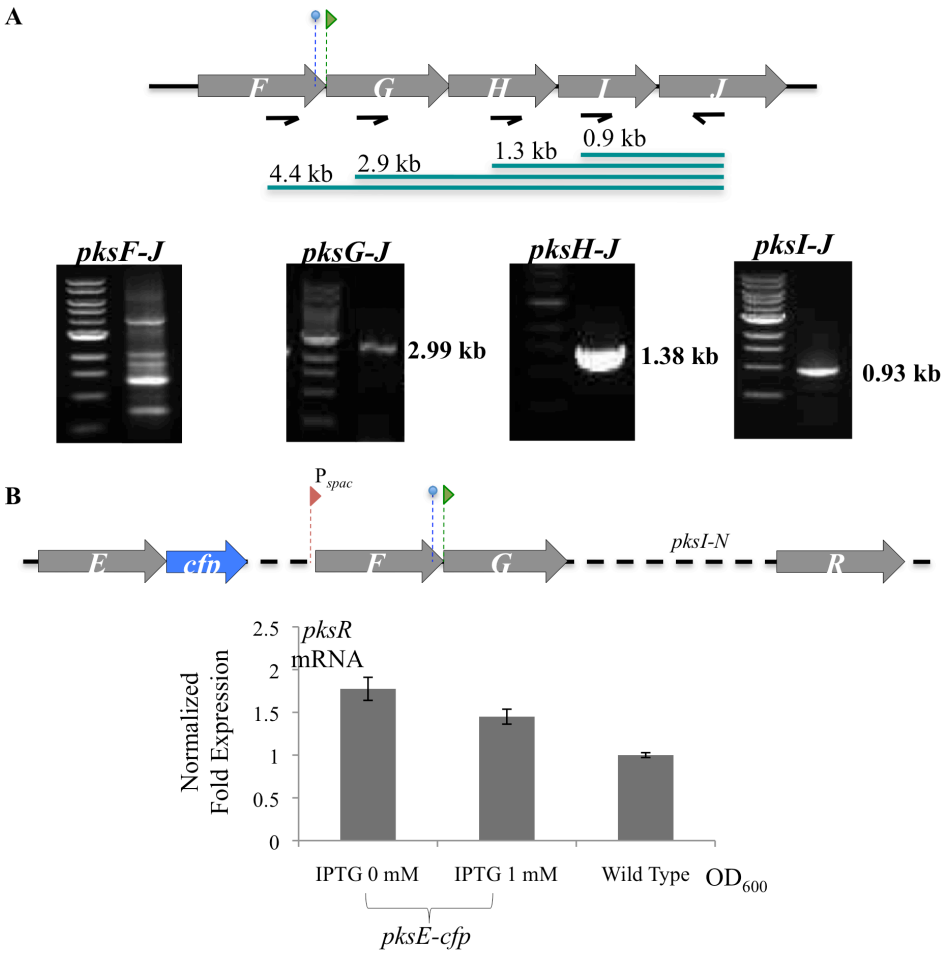


Figure 3- 4. Putative transcription terminator at *pksF*.
(A) Exact RT-PCR was performed to determine RNA transcript length by "walking" along the *pks* gene cluster. Above is depicted the location of the primers forward and reverse, the expected length of the transcripts (green), and the putative transcription terminator (blue) and native promoter at *pksG* (red flag). A *B. subtilis* wild-type strain was cultured in CH medium and total RNA extracted to obtain cDNA with a specific primer annealing at *pksJ*, as described in methods. PCR was used to amplify the cDNA and samples were run in 0.8% agarose. A single band of the expected size was found when the *pks* genes were "walked" from *pksG-J*. Multiple non-specific bands appear when the *pks* genes were "walked" from *pksF-pksJ*. (B) qRT-PCR of *pksR* transcript. Cartoon depicts the location of a *pksE-cfp* fusion that contains a P_{spac} promoter dependent of IPTG, and the putative transcription terminator and native promoter at *pksG*. A strain containing a *pksE-cfp* fusion was cultured in CH medium and IPTG-induced after reaching exponential growth. cDNA was obtained as described in methods and quantification of *pksR* transcript was obtained by qRT-PCR. Data is reported as described in Fig. 2-3. No changes were observed in the abundance of *pksR* transcript under induction of IPTG.

Table 3- 3. Location of *pks* genes and other predicted features in *B. subtilis* str .168 genome[a]

Gene	Coordinates	Direction	-35 signal	-10 signal	RBS	Terminator
pksB	1781983-1782657	+	1781928-1781934	1781955- 1781958	1781970- 1781976	
pksC	1783033..1783896	+	1782895-1782900	1782915-1782919	1783013-1783018	
pksD	1784403..1785290	+	1784255-1784260	1784272-1784275	1784388-1784394	
pksE	1785290..1787590	+				
acpK	1787654..1787899	+				
pksF	1787880..1789127	+			1787880-1787874	1789046-1789074
pksG	1789131..1790390	+	1789118-1789125	1789144-1789148	1789339-1789344	
pksH	1790381..1791157	+			1790371-1790376	
pksI	1791200..1791946	+			1791189-1791194	
pksJ	1791994..1807128	+			1791982-1791986	
pksL	1807115..1820728	+			1807422-1807427	
pksM	1820747..1833532	+			1820732-1820739	
pksN	1833603..1850066	+			1833584-1833590	
pksR	1850084..1857712	+			1850068-1850073	
pksS	1857857..1858984	-	1859099-1859104	1859079-1859083	1858986-1858991	1857727-1857750

[a]Information taken from Subtilis database (http://genolist.pasteur.fr/SubtiList/)

the chromosome from *pksC* to *pksG*. We found amplification of the expected sizes along with multiple non-specific bands in all cases (data not shown). These results suggest that transcription read-through may take place between *pksF* and *pksG*. However, the presence of multiple bands in the transcripts from *pksC* to *pksG* may be an indication of low template abundance, thus the *pksF* termination site may interfere to some degree with transcription elongation through *pksF* to *pksG*.

To further confirm these results, we hypothesized that if a terminator site were present at *pksF*, an insertion of a strong promoter upstream of this termination site would not increase the expression of *pksF* downstream genes. Therefore we used the strain PSK0209 that contains a *pksE-cfp* translational fusion, and P_{spac}, a strong IPTG-inducible promoter, located upstream of the *pksF* locus (Figure 3-1). We measured the abundance of *pksR* transcript by qRT-PCR (Figure 3-4B). Abundance of *pksR* transcript was not significantly different when the cells were induced with 1.0 mM IPTG, compared to non-IPTG induction or wild type. This experiment suggests the presence of a transcription termination site between *pksE* and *pksR*. It could be the predicted *pksF* termination site. However, preliminary data from Winkler lab also suggest the presence of termination sites at *pksJ* (Grace Ellis. Personal communication. Unpublished data). More experiments need to be performed to further confirm the termination sites of the *pks* operon.

Processive antitermination mediated by the pksC-D intergenic region

C*is*-acting regulatory elements have been found in bacteria for transcriptional and post-transcriptional gene regulation (121). In particular, *cis*-acting regulatory RNAs have been found within intergenic regions of gene operons that affect expression of the downstream genes (120). Given that P_{pksD} was not active at any condition tested we hypothesized that the intergenic region between *pksC* and *pksD* (*pksC-D* IR) might have an alternative role rather than acting as a promoter. In collaboration with Wade Winkler's lab (U. Maryland), we predicted that the *pksC-D* IR might function as a *cis*-acting regulatory RNA by promoting transcription antitermination. Therefore, we hypothesized that the deletion of *pksC-D* IR would affect transcription and consequently production of bacillaene would be reduced. To test this hypothesis we measured production of bacillaene in a strain that contained a deletion of *pksC-D* IR (Δ*pksC-D* IR). We used wild-type strain as positive control and Δ*pks* and Δ*codY* mutants served as negative controls (see chapter II). Samples were taken from cells growing in CH medium at OD_{600} ~ 2.8 and OD_{600} ~ 4.5, bacillaene was extracted from supernatants and measured by HPLC (Figure 3-6). Production of bacillaene significantly decreased in the Δ*pksC-D* mutant when compared to the wild-type strain at both cell culture densities sampled (between OD_{600} = 2.8 and OD_{600} = 4.5). Next, we substituted the *pksC-D* IR with the EAR element, a *cis*-RNA regulatory element responsible for processive antitermination of the *eps* operon expression in *B. subtilis* (120). The EAR insertion was able to complement the *pksC-D* IR deletion and restored production of bacillaene (Fig. 3-6). Similar levels of bacillaene were detected in the EAR-complemented mutant and

wild type strains at OD_{600} ~ 2.8 but not at OD_{600} ~ 4.5. These results suggest that *pksC-D* IR is necessary for production of bacillaene. The fact that EAR element was able to restore bacillaene production suggest that *pksC-D* IR might have a similar function as EAR, namely, a *cis*-RNA regulatory element necessary for processive antitermination of the *pks* operon. We have begun to call this region the BAR element (bacillaene-associated RNA) and further experiments are currently performed to confirm this hypothesis.

Discussion

Long gene operons are common in bacteria to optimize the transcription of multiple coding sequences that participate in a similar metabolic pathway (122). Gene clusters that encode natural products are some of the longest operons observed in nature. In particular, a set of 14 genes of the *pks* gene cluster (*pksB* - *pksR*) spans almost 76 kb of *B. subtilis* genome in the same direction at the positive strand, making the *pks* genes the longest operon in *B. subtilis*. In this chapter, we investigated the regulatory features that activate the co-expression of the *pks* genes. We found that P_{pksB}, P_{pksC} and P_{pksS} simultaneously coordinate the expression of the *pks* gene cluster at its highest expression during vegetative growth (OD_{600} ~ 1.8). P_{pksG} appears to function as an alternative promoter that differentially activates *pksG* downstream genes in the forespore but not the mother cell. Some of our data also suggest a weak transcription terminator site upstream of the *pksG* locus that might be coordinated with activation of P_{pksG}. Finally we report

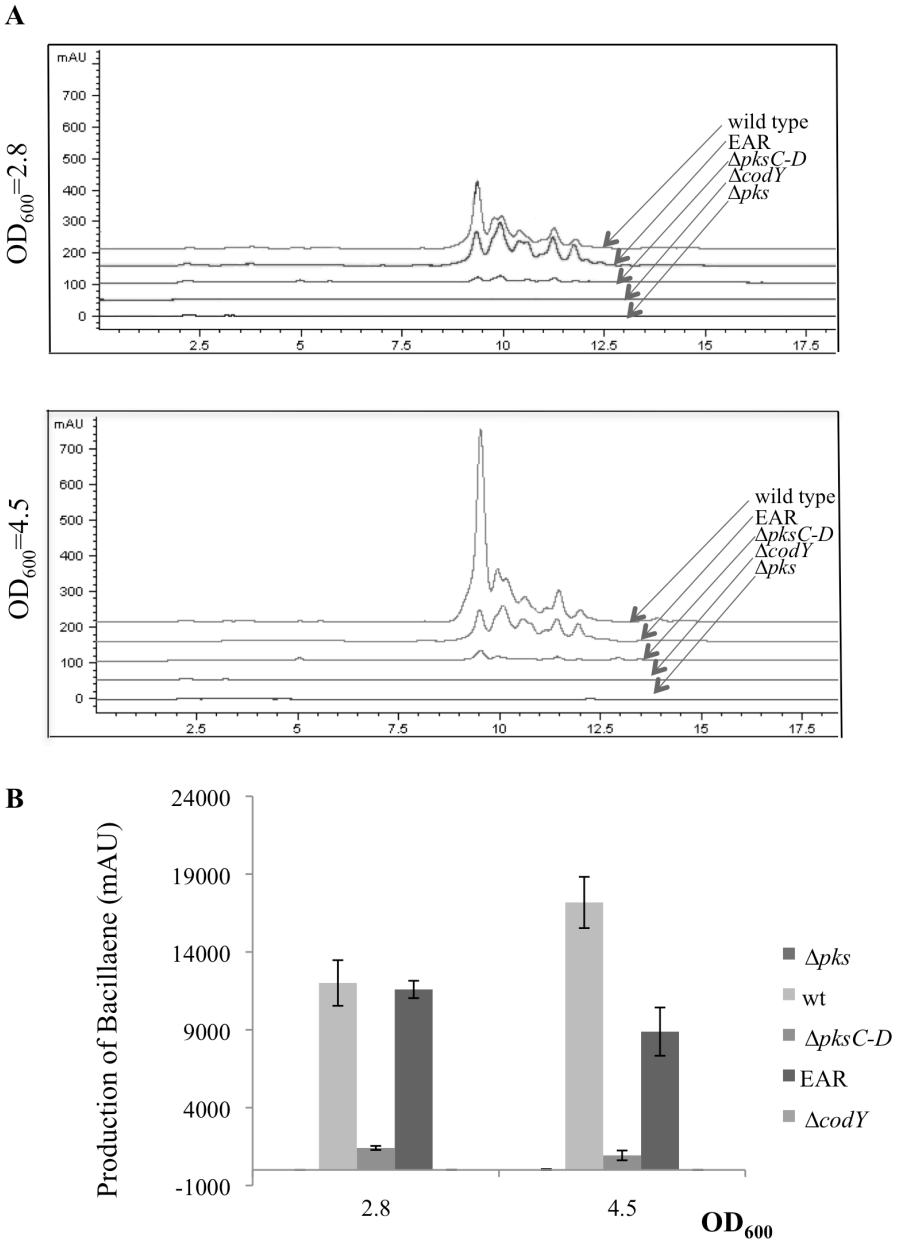


Figure 3- 5. Bacillaene production of *B. subtilis* NCIB 3610 in CH medium.
Wild type strain and the mutants Δ*pks*, Δ*codY*, Δ*pksC-D* intergenic region (Δ*pksC-D*) and *pksC-D::EAR* (EAR) were cultured at 30°C. Samples were taken at OD_{600} ~ 2.8 and OD_{600} ~ 4.5, bacillaene was extracted from supernatants and quantified by HPLC. Multiple peaks represent different isoforms of bacillaene as previously confirmed by LC-MS (See chapter II). Bacillaene production statistically decreases in the Δ*pksC-D* but it is restored in the EAR mutant. All quantitative data shown are average values with standard deviations from triplicates of two independent experiments.

the presence of what appears to be a *cis*-acting regulatory RNA at the intergenic region between *pksC* and *pksD*. The *pksC-D* intergenic region appears to be involved in processive anti-termination of the transcription from *pksC* to *pksR*. To our knowledge, this is the first report of a *cis*-acting regulatory RNA that contributes to the elongation of a natural product encoding- gene cluster.

Transcription of operons is usually initiated with the recognition of a promoter located upstream of the first gene of the operon. This first promoter is usually highly active and other downstream promoters in the operon modestly contribute to elevate the levels of the transcript along the operon (114). In contrast, the proximal promoter of the *pks* operon, P_{pksB}, showed moderate activity, likely contributing to the independent expression of *pksB*. The second promoter of the *pks* genes, P*pksC*, showed the highest activity suggesting that the *pks* operon initiates at *pksC*. In addition, with exact RT-PCR, we were able to observe the effect of the putative transcription termination site located upstream of *pksG*. Although this termination site interferes with transcription elongation at this location, it does not completely interrupt the formation of the transcript. Therefore, our data suggest that P*pksC* is the proximal promoter of an operon that likely extends almost 75 kb, from *pksC* to *pksR*. The weak effect of the transcription termination site at the end of *pksF* may serve to decrease the transcript abundance of downstream genes, such as the ones that encode for the complex synthetase (*pksJ* - *pksR*). This would be consistent with other studies that have shown that the ratio of the protein amount of PksE over PksR is 10 to 100- fold (37). It has been proposed that this

ratio is related to the function of the proteins in the synthesis of bacillaene, i.e., fewer numbers of PksJ to PksR are required to form the core synthetase, but multiple copies from PksB to PksG are needed to function as accessory trans-acting enzymes (37).

To our surprise, we observe high activity of the internal promoter P_{pksG} only inside forespores at early stages of sporulation (S1 – S2). Activity of P_{pksG} may be coupled with the transcription termination site at *pksF* to exclusively promote the expression of *pksG* downstream genes during sporulation, but this hypothesis has yet to be tested. If that is the case, the differential expression of genes downstream of the *pksG* promoter within the spores may compartmentalize downstream *pksG* gene products inside the spore, either as RNA or proteins. We attempted to determine if the *pks* gene products compartmentalization was essential for sporulation or germination but no differences were found between wild type and the Δ*pks* mutant under any condition tested. This data become more intriguing when taken together with some preliminary data that suggest random localization of the Pks megacomplex during vegetative growth but specific localization in the membrane near by the forespore, during early stages of sporulation (S2 - S3) (Appendix II). Further research is necessary to elucidate the function of the *pks* genes during or later development of spore formation.

The apparent processive antitermination role of the BAR element is a novel mechanism of genetic control not previously reported for a natural product encoding gene cluster. It is possible that BAR functions to promote the readthrough of the terminator site at the end of *pksF*. However, other intrinsic termination hairpins have been predicted, in particular, an apparent array of terminators that are located at *pksJ*

(Grace Ellis, Winkler Lab. Personal communication). Whether the *cis*-regulatory element at *pksC-D* intergenic region is promoting read-hrough of the terminator at *pksF* or the predicted array of terminators at *pksJ* is currently under investigation. Understanding the regulatory controls for bacillaene will serve as a model for antibiotic regulation in many organisms with similar long operons.

CHAPTER IV

REGULATORY FUNCTION OF PksA

Summary

The TetR family of transcriptional regulators (TFRs) is widely distributed in bacteria and regulates a variety of cellular functions. PksA is a putative TFR presumed to control the production of bacillaene by modulating the expression of the *pks* gene cluster in *Bacillus subtilis*. In chapter II, we showed that PksA is not involved in the regulation of the *pks* genes. In this chapter, we present evidence that suggest that PksA is a negative regulator of the expression of *ymcC* gene, located upstream and divergently oriented to *pksA*. By qRT-PCR, we show that the expression of *ymcC* is controlled by PksA and increases up to 1000-fold in the *pksA* mutant. We used a YmcC-GFP protein fusion to investigate the cellular localization of YmcC. We do not detect YmcC at any time during vegetative growth but we observed YmcC localization at the membrane during sporulation. A potential function of YmcC during spore maturation or germination is proposed.

Introduction

The ability of bacteria to adapt to the changing conditions of the environment mainly depends on their rapid and adaptive changes in expression of their genes. Natural environments are dynamic, and changes in temperature, nutrient availability and exposure to toxic compounds may occur abruptly (78). Bacterial cells have evolved

mechanisms by which they sense those changes and alter gene expression to appropriately respond to them. Specific gene products are either up regulated or down regulated to respond to specific environmental changes. Adaptive gene expression requires precise regulation in order to overcome difficult conditions in unstable environments (127). A common mechanism used by bacteria to respond to their surroundings is through transcriptional regulators that respond to environmental stimuli and modulate gene transcription and physiological activity accordingly (113).

The TetR family of regulators (TFRs) is a common class of transcriptional regulators of cellular activities (78, 128, 129). The first transcriptional regulator reported in this family was TetR, the repressor of the gene *tetA* that encodes for tetracycline resistance. TetR protein acts as a repressor for both *tetR* and *tetA* genes and, in addition, has a role in the integration, stabilization, or function of the TetA protein in the membrane (130). The N-terminal domain of TetR is a helix-turn-helix DNA-binding domain that binds to an operator in the *tetA* promoter. TetR also has a C-terminal ligand-binding domain that binds to tetracycline in order to inactivate TetR repression of its target (129). The majority of characterized TFRs function as repressors, though there are some TFRs that have been reported as activators, and some others that can act as both repressors and activators (130–132). While TetR is still the paradigm for the family, it is well known that members of this family are involved in a wider range of regulatory activities including central metabolism, pathogenicity and development (129). Still, it is estimated that about 25% of TFRs function as regulators of genes encoding efflux pumps

that confer antibiotic resistance, making this role the most common one among the TFRs (128).

Two characteristics of the TetR family are conserved so that its family members are easily recognized. First, TFRs are almost exclusively α helical and all contain a well-conserved N-terminal helix-turn-helix DNA binding domain that binds to their target gene(s). The C-terminal domain of TFRs is usually highly variable and in the majority of the cases interacts specifically with small-molecule ligands, allowing this family to respond to a diverse range of stimuli (78). Second, the orientation and proximity of TFRs with respect to their target genes is very predictive of a regulatory relationship. TFRs are classified in three major groups according to their proximity and orientation to their target genes (128). Type I TFRs show a divergent orientation to at least one of the adjacent target genes. Type II TFRs are usually in the same orientation of one or more adjacent target genes, which are co-transcribed. Type III TFRs show neither of these relationships and regulate genes located in another region of the genome (129). More than 50% of TFRs belong to class I, so a TFR that shows divergent orientation to an adjacent gene with less than 200 bp of intergenic region is a good candidate for regulating the adjacent gene(s) (129).

PksA was annotated as the TFR of the *pks* genes in *Bacillus subtilis* (33, 38–40). As mentioned in previous chapters, the *pks* genes encode for proteins that synthesize a small molecule with antibiotic properties called bacillaene. *pksA* is upstream of a gene cluster composed of fifteen genes (*pksB-pksS*) that is nearly 76 kb long (Fig. 2-1)(chapter II and references within). Given the *pksA* gene sequence and its orientation

with respect to the *pks* genes, it was thought that *pksA* is a class II TFR that controls the expression of the *pks* genes (38–40, 129). However, our analysis of the *pks* gene regulation led us to conclude that PksA is not involved in such regulation (Chapter II). Therefore, we hypothesize that PksA is either a class I TFR and is regulating the expression of the divergently oriented gene, *ymcC*, or belongs to the class III TFRs and is regulating genes located elsewhere in the *B. subtilis* genome.

The purpose of this chapter is to investigate the function of PksA. Our data confirm that *pksA* encodes a TFR that regulates the expression of the gene adjacent and divergently oriented to *pksA*, *ymcC*. PksA functions as a negative regulator of *ymcC*. PksA also represses its own expression. The expression of *pksA* is constitutive and not *ymcC* transcript or YmcC protein are detected during vegetative growth. In addition, deletion or overexpression of YmcC caused no evident phenotype during vegetative growth either. However, YmcC is detected at late stages of sporulation suggesting a potential role of YmcC during spore maturation or possibly germination.

Material and methods

Bacterial strains, primers, media and growth conditions

The undomesticated strain *Bacillus subtilis* NCIB 3610 was used for all the experiments in this work. Table 4-1 contains a list of the strains used in this study. The construction of the transcriptional reporter P_{pksA}-*yfp* was made using the same methods as described in chapter II. The YmcC-GFP protein fusion was generated by cloning *ymcC* into pLM028, which was then integrated into the bacterial genome at the *ymcC*

Table 4- 1. *Bacillus subtilis* strains used in this study

Strain	Relevant Genotype	Source
PDS0066	*B. subtilis* NCIB3610 wild type	Laboratory collection
PDS0030	*B. subtilis* NCIB3610 *amyE:: P_{pksA}-yfp (cat)*	Laboratory collection
PDS0101	*B. subtilis* NCIB3610 *amyE:: P_{pksA}-yfp (cat)* Δ*pksA*	Laboratory collection
PDS0184	*B. subtilis* NCIB3610 Δ*pksA::kan*	Laboratory collection
PDS0479	*B. subtilis* NCIB3610 Δ*pksA::kan lacA::pksA+(mls)*	Laboratory collection
PDS0150	*B. subtilis* NCIB3610 *amyE::PpksA lacI (cat)*	Laboratory collection
PDS0389	*B. subtilis* NCIB3610 *ymzD-gfp (spc)*	Laboratory collection
PDS0392	*B. subtilis* NCIB3610 *ymzD-gfp (spc)* Δ*pksA::kan*	Laboratory collection
PDS0390	*B. subtilis* NCIB3610 *ymcC:gfp (spc)*	Laboratory collection
PDS0391	*B. subtilis* NCIB3610 *ymcC-gfp (spc)* Δ*pksA::kan*	Laboratory collection
PDS0171	*B. subtilis* NCIB3610 Δ*ymcC::kan*	Laboratory collection
PDS0255	*B. subtilis* NCIB3610 Δ*ymcC::kan amyE::ymcC+:lacZ (cat)*	Laboratory collection

Table 4- 2. Primers used in this study

Name	Sequence (5'-3')
ymcC-gfp-F	accaggaattcactggtgtggacctt
ymcC-gfp-R	gaactctcgagactcctgagattcgc
ymzD-gfp-F	Aggacgaattcttacttttatcgcgg
ymzD-gfp-R	gaactctcgagagctgcttctctatt
Nth_ymcC_F	tgttttggggcttgctgtacgc
Nth_ymcC_R	gaatgccgctcaacgcctcg
qPCR-ymcC-F	gaaattgcgttttgggttgt
qPCR-ymcC-R	ccagttgccgctaacaagat
rtPCR_pksA_1	gcaaaggaagcaggattgtc
rtPCR_pksA_2	aatccctcactgcaatgtcc

locus. Unless otherwise stated, vegetative growth was performed in CH medium at 37°C. Induction of sporulation was performed by centrifuging cells that were growing at exponential phase (OD_{600} between 0.5 – 0.8) in CH medium, and resuspending in sporulation medium as described by Sterlini and Mandelstam (125). The initiation of sporulation was designated as S0, the time of resuspension. Samples were harvested at hourly intervals for eight hours (S0-S7). Table 4-2 contains a list of primers used in this study.

Northern blots

Cells from overnight cultures were inoculated into CH medium at an initial OD_{600} = 0.08, grown at 37 °C with vigorous shaking, and harvested at OD_{600} = 1.5. Total RNA was prepared using an RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. The resulting RNA sample was run in 1% agarose after denaturation with glyoxal and dimethyl sulfoxide, transferred on a nylon membrane (Ambion), and then hybridized with an RNA probe for *ymcC*, as previously described (133). NEBlot Phototope kit (NEB) was used for biotinylation-probe labeling and hybridization, following the manufacturer instructions.

Quantitative RT-PCR

Cells from overnight cultures were inoculated into CH medium at initial OD_{600} = 0.08, grown at 37 °C with vigorous shaking, and harvested either at early (OD_{600} = 0.2), mid (OD_{600} = 0.8) and late (OD_{600} = 1.8) exponential growth, or after four hours of

induction under different concentrations of IPTG. Detailed description of the method is in chapter II. Briefly, RNAprotect Bacteria Reagent (Qiagen) and RNeasy Mini Kit (Qiagene) were used to isolate total RNA, according to the manufacturer's instructions. Subsequently, DNase treatment was performed to the strains using Turbo DNA-free kit (Applied Biosystems). Two Steps qRT-PCR was performed with the Thermo Scientific DyNAmo Flash SYBR green qRTPCR kit. Total RNA of 200 μg was used as template to synthesize cDNA. After the RT step, quantitative PCR was done with 2μL of cDNA, using specific target primers (Table 4-2) in a CFX96 Touch Real Time PCR thermocycler (BioRad). Either *gyrB* or *rpoB* was used as the reference gene as indicated in each experiment. The samples were run in triplicate for each target gene and negative controls were included for each sample where the reaction contained DNase-treated total RNA in place of cDNA. Primer efficiency and C_q values were calculated using the software LinReg (73).

Fluorescence microscopy

Samples from shaken liquid cultures in CH medium were taken for fluorescence imaging, centrifuged at 8,000 rpm and washed once with PBS. To stain the membrane, cells were resuspended in 20 μM 1-(4-trimethylammoniumphenyl)-6-phenyl-1,3,5-hexatriene p-toluenesulfonate TMA (Molecular Probes) or N-(Texas Red sulfonyl)-1,2-dihexadecanoyl-snglycero-3-phosphoethanolamine, triethylammonium salt (TX-RED) (Molecular Probes) as indicated in each experiment. Fluorescence images were captured using a Nikon Ti-E inverted microscope equipped with a CFI Plan Apo Lambda DM

100X objective, TI-DH Diascopic Illuminator, and a CoolSNAP HQ2 Monochrome Camera. Exposure time was 1,000 ms for TMA and varied for YFP and GFP from experiment to experiment as indicated in each figure. The NIS-elements AR software was used to capture and process the images identically for comparative analysis.

Results

Bioinformatic analysis

In our initial purpose of looking for regulators of the *pks* gene cluster we targeted *pksA*, the first gene of the *pks* gene cluster that had been annotated as the regulator of the *pks* gene cluster (Chapter II and references within). PksA has been predicted to belong to the TetR family of regulators (TFRs) (129). Functional domain analysis using the modular architecture research tool ScanProsite (http://prosite.expasy.org/scanprosite/) identified a region between the residues 8 to 68 of PksA as a helix-turn-helix DNA-binding domain (Figure 4-1A). This N-terminal DNA-binding domain is typical of TFR family members and represented by conserved motifs or profiles in the public databases (129). The C- terminal domain of TFRs usually interacts with at least one ligand, altering the regulators ability to bind DNA. In the case of PksA there is not an apparent similarity of its C-terminal domains with other TFRs proteins (134). Therefore, prediction of ligand candidates has not been proposed. Overall bioinformatics analysis based on primary structure predicts that PksA is another member of the family of tetR regulators.

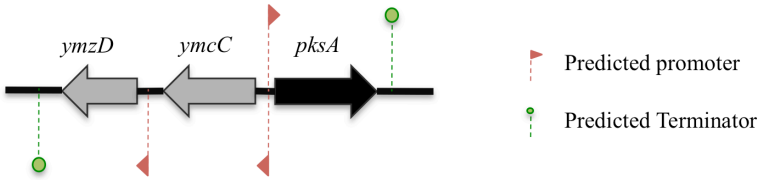
A

PksA (205 aa)

H-T-H motif

MPKQIDHEKRRKQIAEATWRVILERGMEGASARNIAKEAGLSLGALRHYFSTQDELLAFA
MKLVQEKVTDRIKDIAVRDLLPKEKVLQILLEMVPTNEETIREMEVWFAFTAYARHKKDM
FDASHDGIFSGMRNLIAYLDESDLLKQNADKDIEAERLYALVDGLALHAMLDPVRVNKDR
IKRVIMQHVESICVEDTRETQKRHP

B



C

Figure 4- 1. PksA and YmcC organization and conservation among *Bacillus spp*.
(A) PksA aminoacid sequence. In yellow is the predicted tetR-type domain with HTH motif highlighted in red (http://prosite.expasy.org). (B) *pksA*, *ymcC* and *ymzD* loci in *B. subtilis*. Divergent organization of *ymcC* and *ymzD* respect to *pksA* is shown. Predicted promoters and terminators are indicated. A shared promoter for the (+) and (-) strand is present between *ymcC* and *pksA*. No terminator region is predicted between *ymcC* and *ymzD*. (C) PksA is evolutionary conserved among *Bacillus spp*. Multiple-amino-acid-sequence alignment of PksA from *B. subtilis* and the homologous proteins from *B. amyloliquefaciens* FZB42, B. cereus ATCC 10987, B. cereus ATCC14579, B. licheniformis ATCC 14580, B. pumilus SAFR-032 obtained using Clustal Omega (http://www.ebi.ac.uk/Tools/services/web_clustalo/toolform.ebi). Mid blue indicates percentage of agreement >80%, light blue >60%, light grey >40% and white ≤ 40%.

Table 4- 3. Conservation of *pksA-ymcC* pair among *Bacillus spp*[a]

Organism	Gene accesion number[b]	Protein similarity	Protein identity	Predicted product
B. subtilis 168	BSU17080	---	---	Transcriptional regulator, TetR family
B. amyloliquefaciens	RBAM_009910	85%	66%	Transcriptional regulator, TetR family
B. cereus ATCC 10987	BCE_0378	61%	39%	Transcriptional regulator, TetR family
B. cereus ATCC 14579	BC_2410	60%	30%	Transcriptional regulator, TetR family
B. licheniformis	BL02122	63%	41%	Putative uncharacterized protein
B. pumilus	BPUM_3015	64%	40%	Transcriptional regulator

[a] Based on http://genodb.pasteur.fr/cgi-bin/WebObjects/GenoList.woa/wa/goToTaxoRank?level=Bacillus%20subtilis%20168

[b] Genolist accession number

Given that PksA was the presumptive regulator of the *pks* genes, we first hypothesized that PksA was a type II TFR, a regulator with the same orientation of the genes that it regulates. The evidence shown in Chapter II indicated that PksA does not regulate bacillaene production. As shown in Figure 4-1B, *pksA* is adjacent and divergently oriented to *ymcC* gene, a predicted integral inner membrane protein. In addition, the intergenic region between the *pksA* and *ymcC* genes is 120 bp, suggesting that PksA regulates *ymcC* expression (129). *pksA* is also evolutionary conserved and associated with *ymcC* in several species of *Bacillus* (Fig. 4-1C and Table 4-3). PksA (RBAM_009910) from *B. amyloliquefaciens* is 85% similar to PksA in *B. subtilis*. *B. cereus, B. lichenimoformis* and *B. pumilus* also have corresponding proteins homologous to *B. subtilis* PksA, with about 60% similarity. Genes homologous to *ymcC* are adjacent to *pksA* homologous genes in *B. amyloliquefaciens, B. lichenimoformis* and *B. pumilus.* In *B. cereus* ATCC 10987, BCE_0379 is the product adjacent to BCE_0378, the *B.subtilis pksA* homologous gene. Despite having low identity to *ymcC*, BCE_0379 shows amino acid sequence similarity to integral membrane proteins, as *ymcC* does. Therefore, we predicted that *ymcC* is the target of PksA, i.e. PksA is regulating the expression of *ymcC*.

PksA *controls the expression of ymcC*

To determine whether PksA is a type I TFR that regulates the expression of *ymcC*, we directly analyzed the levels of *ymcC* transcripts in a strain that contained a deletion of *pksA* and compared it with wild type. A 1kb transcript was detected by

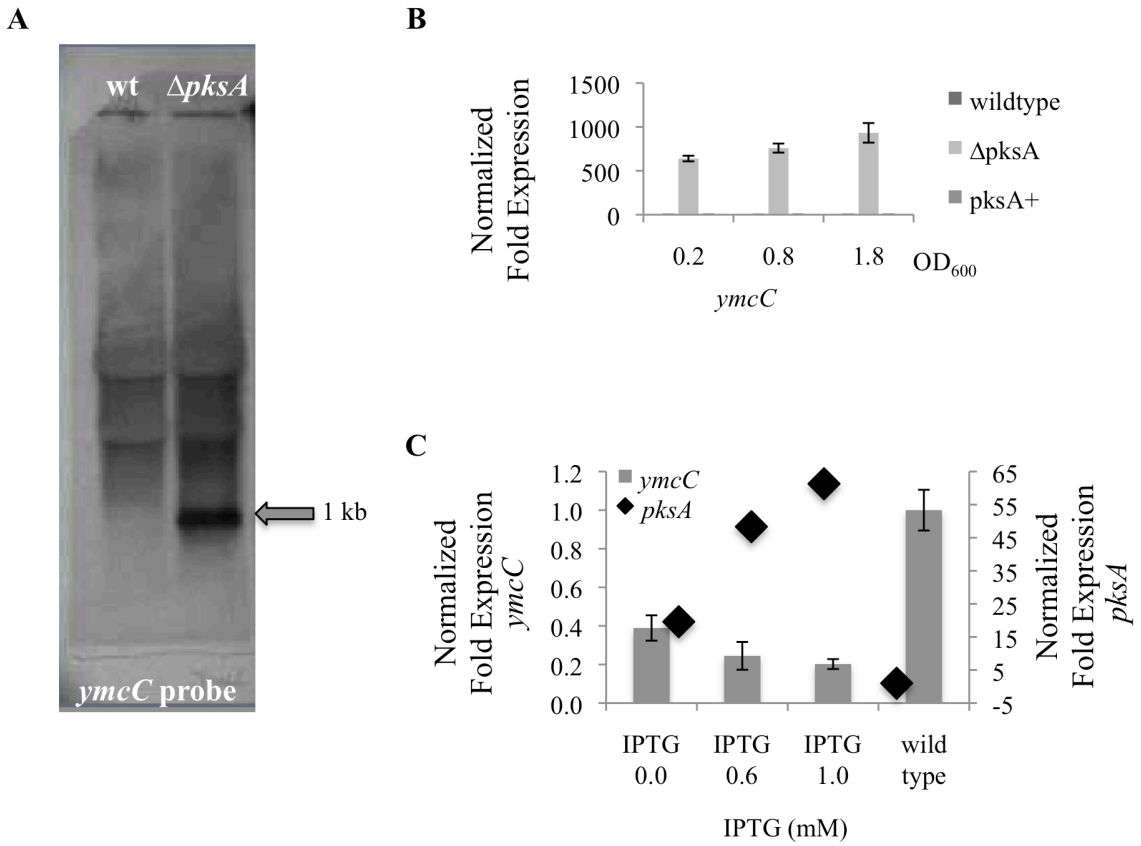


Figure 4- 2. *ymcC* is controlled by PksA.
(A) Northern blot analysis shows that *ymcC* is detectable only in the absence of PksA. *B. subtilis* wild type strain and a Δ*pksA* mutant were grown in CH medium at 37°C, and total RNA extraction at exponential growth was performed. The arrow indicates the band that is detected only in the Δ*pksA* mutant. (B) qRT-PCR indicates that PksA controls *ymcC* expression. Data are presented as described for Figure 2-3, panel D. The *ymcC* transcript was measured by qRT-PCR in wild type, Δ*pksA* and the Δ*pksA* genetically complemented mutant (*pksA*+) strains. The deletion of *pksA* led to an induction of ~1000-fold expression of *ymcC*. Comparison between strains at low (0.2), mid (0.8) and high (1.8) OD_{600} indicates that PksA does not allow expression of *ymcC* during exponential growth. (C). qRT-PCR overexpression of PksA shows a corresponding reduction in *ymcC* expression. Y-axis denotes Normalized Fold Expression for *ymcC* (left) and *pksA* (right)

Northern blotting in the *pksA* mutant but not in wild type (Figure 4-2A). However, the expected size of *ymcC* transcript was 558 bp. Another gene encoding a putative inner membrane protein, *ymzD*, is located downstream and with the same orientation as *ymcC* (Figure 4-1B). *ymzD* is 456 bp long and no transcription terminator sites are detected in the intergenic region between *ymcC* and *ymzD* (http://subtiwiki.uni-goettingen.de/apps/expression/expression.php?search=ymcC). Therefore it is likely that *ymcC* and *ymzD* are being transcribed as a one single operon and negatively regulated by PksA.

Our Northern blotting suggested that the regulation exerted by PksA on the *ymcC-ymzD* pair is very strong since no detectable levels were found in wild type but a strong band appeared in the *pksA* mutant. To determine the strength of this regulation, we performed qRT-PCR analysis of the *ymcC* target in the *pksA* mutant and compared it to wild type (Figure 4-2B). Deletion of *pksA* produced an increase of ~1000-fold in the expression of *ymcC* mRNA transcripts. Complementation of the *pksA* locus ($pksA^+$) restored the repression of *ymcC* transcript to wild type levels. To determine whether the abundance of PksA is correlated with the repression of *ymcC*, we used a strain that contained an IPTG-inducible copy of *pksA* in the Δ*pksA* strain and quantitated *ymcC* transcripts at different concentrations of IPTG. As expected, the increasing concentration of IPTG elevated the levels of *pksA* up to 65-fold above the native levels of the wild type strain (Figure 4-2C). In turn, the increasing levels of PksA gradually decreased the *ymcC* mRNA abundance. These results show that the amount of PksA is related to the repression of the *ymcC* in *B. subtilis* cells.

pksA product negatively controls the expression of its own gene

Many characterized TFRs are known or believed to be autoregulatory to ensure the transcriptional repressor level within optimal concentration limits (78, 129). Given the relationship between the amounts of PksA and the level of repression of *ymcC*, we hypothesized that expression of PksA is also precisely regulated to ensure appropriate concentration of the repressor. To corroborate this hypothesis we measured the abundance of *pksA* transcript over time in a wild type strain (Figure 4-3A). qRT-PCR results showed virtually unchanged levels of *pksA* transcript over time. The abundance of *pksA* transcript had a slight induction of ~1.6- fold from early (OD_{600} = 0.2) to mid (OD_{600} = 0.8) exponential growth and ~3.1- fold from early to late (OD_{600} = 1.5) exponential growth. These results suggest that PksA concentration is substantially uniform during exponential growth.

To assess whether PksA exerts regulatory activity on its own gene, we isolated the 5'UTR of *pksA* and fused it to the *yfp* gene to create a fluorescence transcriptional reporter. We introduced this reporter, P_{pksA}-*yfp*, in a wild type and a Δ*pksA* background (Figure 4-3B). Considerable higher levels of *pksA* promoter activity were detected by fluorescence microscopy in the Δ*pksA* background compared to wild type. Based on these results we concluded that PksA represses the expression of its own gene.

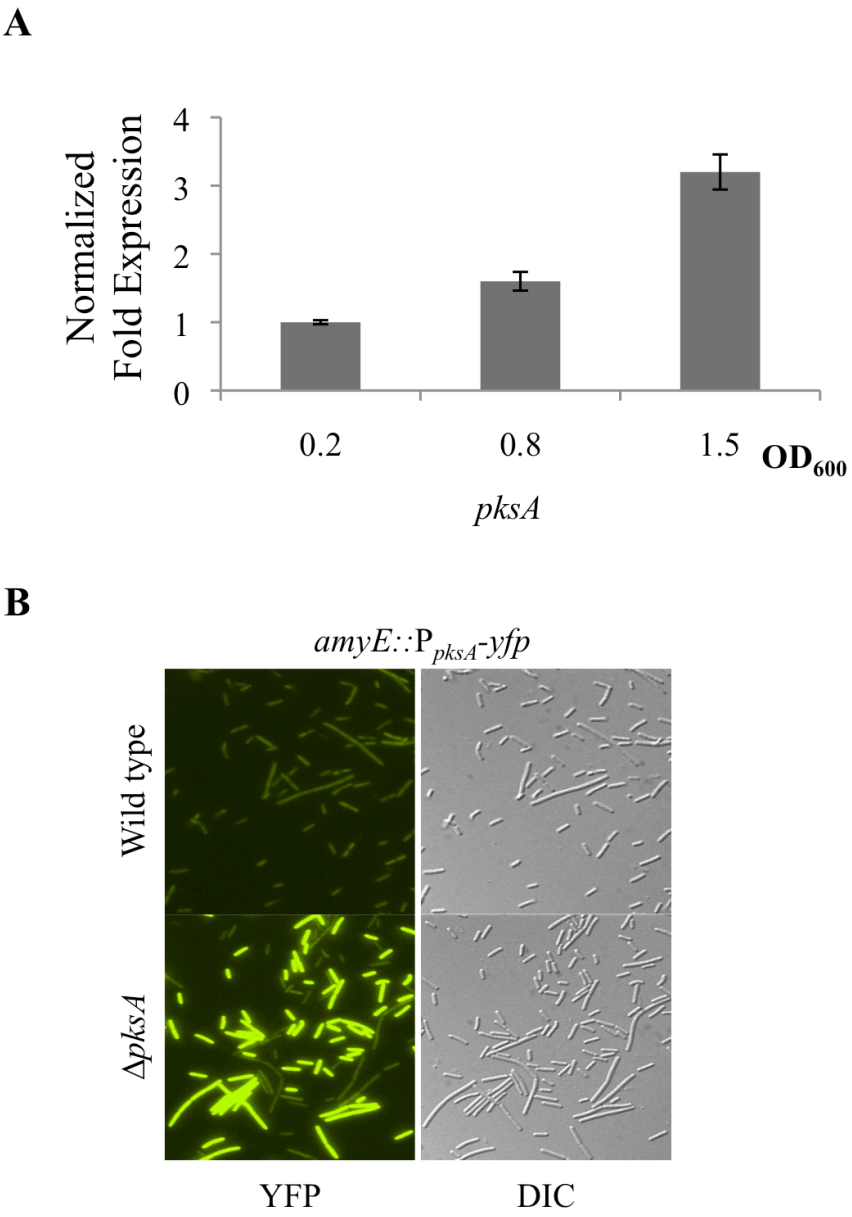


Figure 4- 3. PksA regulates its own expression
(A) qRT-PCR indicates that PksA expression does not dramatically fluctuate over time. Data are presented as described for Figure 2-3, panel D. The *pksA* transcript was measured in a wild type strain at early (OD_{600} = 0.2), mid (OD_{600} = 0.8) and late (OD_{600} = 1.5) exponential growth. Abundance of *pksA* transcript changed ~3-fold from early to late exponential growth. (B) Fluorescence images indicated that activity of *pksA* promoter is higher in the *pksA* mutant than in the wild type strain suggesting a repressive activity of PksA on its own product.YFP (green) and DIC images are shown. Scale bar, 3 μM.

Localization of YmcC *is at the membrane*

ymcC encodes a putative integral inner membrane protein (http://genodb.pasteur.fr). Transmembrane spanning domain (TMD) analysis using the research tool TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) predicts five TMDs, in addition to the periplasmic and cytoplasmic loops (Figure 4-4A). To determine whether YmcC localizes in the membrane, we constructed a *ymcC-gfp* gene fusion that was integrated at the *ymcC* locus. We analyzed the distribution pattern of YmcC-GFP at various points of vegetative growth by fluorescence microscopy (Figure 4-4B). No YmcC-GFP could be detected when cells were harvested during exponential or stationary phase (Fig. 4-4B and data not shown). The same results were seen when GFP was fused to YmzD, which is coexpressed with ymcC and is also predicted to be an integral membrane protein (Fig. 4-4B and data not shown). These results indicate that YmcC and YmzD may not have any role in vegetative growth, at least under the conditions we tested.

We also determined whether the YmcC-GFP and ymzD-GFP fusion proteins could be detected during sporulation. We cultured the strains in resuspension medium and several time points were analyzed by fluorescence microscopy (Figure 4-4C). YmcC-GFP signal was low but observable two hours after the initiation of sporulation (S2) and the signal increased slightly over time. YmcC distribution seemed to be equally distributed in the mother cell and the forespore but with preferential localization in the membrane as it followed the same pattern of signal produced by TMA. These results seemed to be consistent with the prediction that YmcC is a membrane protein but better

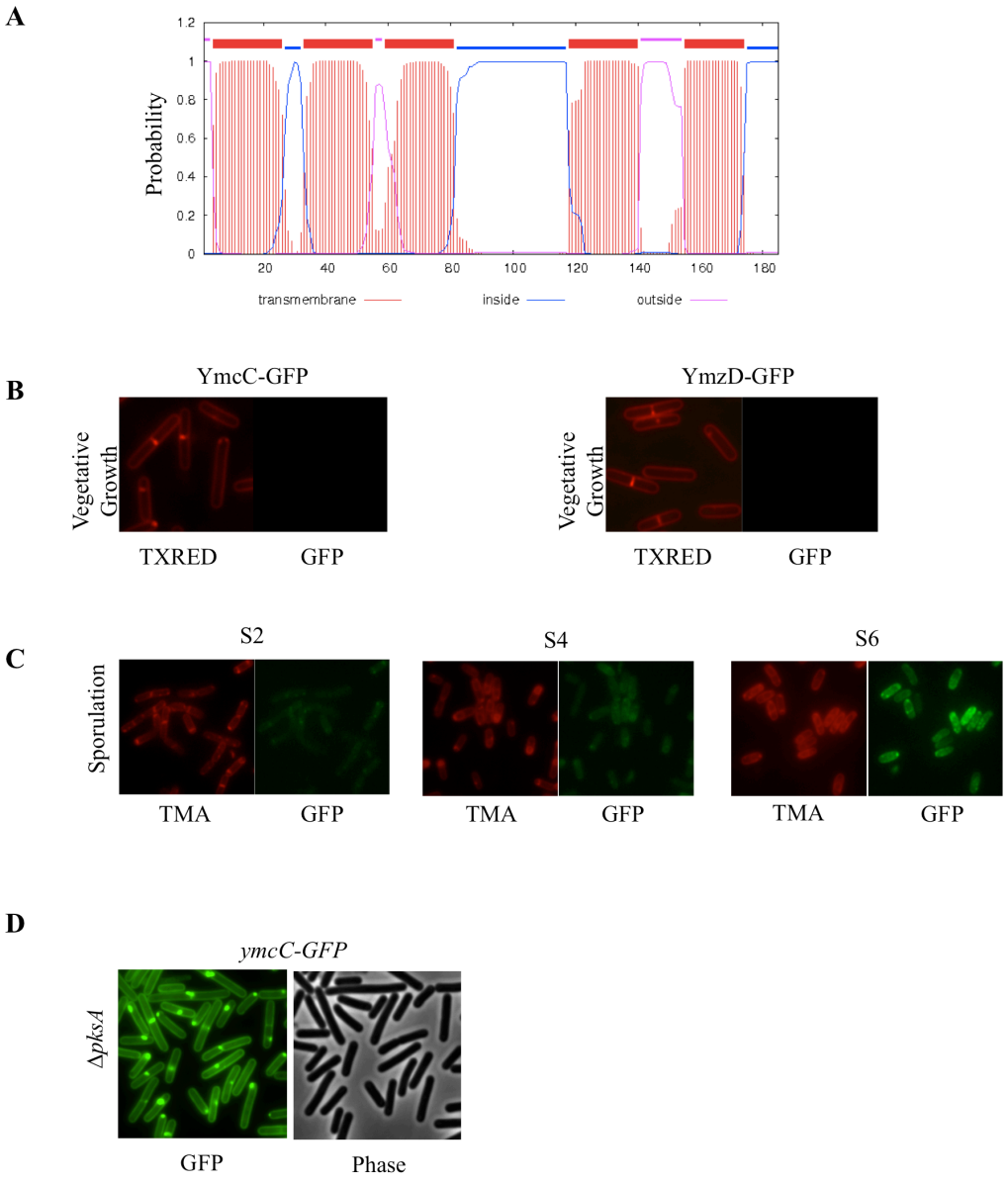


Figure 4- 4. Cellular localization of YmcC.
(A) Predicted membrane-spanning domains of YmcC (http://www.cbs.dtu.dk/services/TMHMM-2.0/). YmcC contains five predicted transmembrane helices (red), three regions located inside the cell (blue) and two regions that contact the exterior (magenta) (B) Fluorescence images show no evident localization of YmcC in vegetative and sporulation growth. YmcC-GFP signal intensity in a wild type background is weak at any time point during exponential growth and spore formation. YmcC-YFP (green), cells stained with TMA (red) or TEXAS-RED (red) as indicated, and phase contrast images are shown. Scale bar, 3 µM. (C) YmcC-GFP signal is observed during sporulation. (D) YmcC-GFP signal in a Δ*pksA* mutant shows localization of YmcC by the cellular membrane.

fluorescence signal would be necessary to confirm these observations. No detectable levels of YmzD were observed at any time during sporulation so YmzD has no apparent role during sporulation either (Data not shown).

To determine whether we could increase the GFP signal of YmcC and YmzD, we introduced the translational fusions YmcC-YFP and YmzD-GFP in two separate Δ*pksA* mutant cell lines. Cells were observed by fluorescence microscopy during vegetative growth and spore formation as described above. Fluorescence signal was prominent for YmcC, even during vegetative growth (S0) (Figure 4-4D). Fluorescence was detected along the membrane, at division septa and sometimes in discrete foci at the mid and pole cells. This pattern of distribution is similar to other proteins that localize in the membrane (135–137).

We attempted to identify the function of YmcC in *B. subtilis*. Given the proximity of *pksA* and *ymcC* loci to the *pks* operon, we first hypothesized that YmcC might function as a bacillaene efflux pump. Therefore, we co-cultured *Streptomyces* sp. Mg1 and a *B. subtilis* strain that contained a *ymcC* deletion to detect changes in competition due to the production of bacillaene. No differences were observed between *B. subtilis* wild type and Δ*ymcC* regarding the production of bacillaene (data not shown). We also used microbial phenotype arrays (http://www.biolog.com) to identify any potential phenotype defect related with the deletion of *ymcC*. We used a wild-type *B. subtilis* strain to compare with the isogenic mutants Δ*ymcC*, Δ*pksA* and a complementation of *ymcC* (*ymcC*$^{+}$) (Table 4-4). Cellular profiles for hundreds of conditions for carbon, nitrogen, phosphorous and sulfur utilization, nutrient stimulation,

Table 4- 4. Biolog results summary

Condition	**Δ*ymcC***	***ymcC*$^{+}$**	**Δ*pksA***
Carbon Utilization	-[a]	-	-
Nitrogen Utilization	-	-	-
Phosphorous and Sulphur	-	-	-
Nutrient Stimulation	-	-	-
Osmotic and pH	-	-	-
Chemical Sensitivity	Increased resistance to oxytetracycline, folate antagonists, aminoglycosides, respiration inhibitors	Increased resistance for folate antagonists, plumbagin, aminoglycosides, tetracyclines, chloramphenicols, and respiration inhibitors	Increased resistance to folate antagonists, aminoglycosides,

[a] No phenotype differences were found between wild type and the mutant

changes in osmotic and pH resistance, and chemical sensitivity (PM1-40) were tested. Wild-type and the isogenic mutant strains of *B. subtilis* showed no growth differences in the media and nutrient supplements tested. Chemical sensitivity decreased similarly in all the mutants. We think the changes in the chemical sensitivity phenotype were due to the presence of the selection markers that were used to identify the mutants rather than an effect of the mutants on the genes under study. Therefore, the phenotype arrays did not show any apparent differences that suggested a potential role for YmcC and further studies are necessary to determine the function of YmcC.

Conclusions

We found that the TFR gene *pksA* encodes a novel negative regulator of *ymcC*. The *pksA* gene is divergently oriented to *ymcC* and the intergenic region that separates them contains the sequence necessary to repress transcription initiation from the promoters of *pksA* and *ymcC*. Our data reveals a near 1000-fold repression of *ymcC* by PksA as measured by qPCR. PksA also represses the expression of its own product as evidenced by the fluorescence signal differences of the P_{pksA} reporter in the presence and absence of the *pksA* gene. Other TFR systems, including the *tetR*-*tetA* model pair, are organized similarly to the *pksA*-*ymcC* pair and the region that represses expression of the gene pair is an apparent palindromic operator sequence that is bound by a dimeric form of the regulator (128). In the case of the *tetR*-*tetA* pair, tetracycline activates *tetA* expression by binding TetR and lowering its affinity for DNA (138). Our data suggest

that the *pksA-ymcC* pair is organized similarly to the *tetR-teA* pair but further studies are necessary to confirm the presence of the palindromic operator sequence.

Levels of *ymcC* expression are slightly increased during sporulation according to the SubtiExpress database. To assess whether we YmcC could be visualized under native conditions during sporulation, we used the YmcC-GFP protein fusion in the wild-type background. Fluorescence was detected as early as two hours after beginning of sporulation (S2) and the fluorescence signal slightly increased over time. The pattern of distribution of YmcC was similar to the pattern observed in the Δ*pksA* mutant, confirming the native localization of YmcC is at the membrane and occurs during spore formation. These results are suggestive of a potential role of YmcC during spore maturation or possibly germination.

Given that the production of YmcC mainly occurs at the latest stages of sporulation, it is likely that YmcC functions during spore maturation or possibly germination. For instance, one possible function based on the localization of YmcC at the membrane would be as a specific receptor in response to an appropriate nutrient during germination. In *B. subtilis*, the operons *gerA* and *gerB* have been discovered to encode germinant receptors. The *gerA* operon encodes three proteins, GerAA, GerAB and GerAC (139). These three proteins are integral membrane proteins and GerAA has a predicted membrane-bound domain that would span the membrane at least five times, a similar pattern as *ymcC* (140) (Fig. 4A). Protein receptors are thought to be germinant specific but no detailed function of any of the germinant receptor protein has been described. It is known that spore-formers often contain several of these receptor operons

and respond to different types of germinant via multiple receptors. It is plausible that YmcC is one of those germinant receptors but further studies are required to validate this hypothesis.

Although Northern blotting experiments suggest that *ymcC* and *ymzD* are being transcribed as one single operon, we were unable to observe YmzD during sporulation, as we expected given the results about YmcC localization. Therefore, more experiments are needed to confirm whether ymzD is part of the transcript controlled by PksA. If coexpression of YmcC and YmzD exist, YmzD may not be a membrane protein as predicted but an RNA element that coexpresses with ymcC to increase the level of regulation of this system.

In the regulatory scheme of divergently oriented TFR systems, synthesis of the repressor and its regulated protein(s) is simultaneously derepressed in the presence of a ligand that binds to the C- terminal domain of the regulator and lowers its affinity to DNA (138). This on/off switch mechanism may be used by PksA to respond to a specific signal but until now the ligand has not been identified. While TFRs can be easily identified based on conserved sequence motifs, identifying their cognate ligands is more challenging. A recent study has reported the use of a phylogenomics approach to organize a TFR database in order to predict cognate ligands for TRFs of unknown function (134). Unfortunately, PksA clusters with other unknown TFRs and the only ligand predicted to bind this TFR cluster is bacillaene due to the previous misconception that *pksA* is part of the *pks* gene cluster (See chapter II and references within). More studies are required to determine if one or multiple ligands are able to bind PksA to alter

its structure and affinity to DNA. If a ligand is part of the regulation of the *pksA-ymcC* pair, we speculate that the ligand must be present at latest stages of sporulation to release the repression exerted by PksA. Alternatively, a ligand that binds PksA may be present during every other *B. subtilis* stage except late sporulation, releasing PksA from the DNA and acting in this way as a catabolite that increases DNA-affinity of PksA (catabolite repressor).

CHAPTER V

CONCLUSIONS AND FUTURE DIRECTIONS

Antiobiotic resistance greatly challenges the medical community's ability to fight infectious diseases. After World War II, the stunning success of antibiotics in treating infections led to the identification and synthesis of hundreds of different antibiotics (141). In part due to increased antibiotic use, the human life expectancy has risen from about 40 years to 77 years in the past century (141). However it did not take long to recognize that bacteria were able to evade antibiotic effects and undermine their efficiency. Resistance mechanisms have emerged and spread globally, resulting in the death and disability of thousands of people a year. According to the World Health Organization, antibiotic resistance is currently a serious threat to global public health (11). In addition, many pharmaceutical firms are eliminating their natural product research due to technical limitations in identifying new compounds with desirable activity and financial considerations for cost of discovery and development (141). Therefore, new approaches need to be investigated to counteract antibiotic resistance. One area of investigation could be the understanding of antibiotic resistance from an ecological perspective, identifying how antibiotics function and how bacteria respond to them in nature (43, 142). Additionally, predictions based on bacterial genome capacity suggest that many antibiotics remain to be discovered since the gene clusters that encode their synthesis do not become active under laboratory conditions. Therefore,

understanding the mechanisms of antibiotic regulation will contribute to activate innumerable natural product-encoding gene clusters that remain silent (112).

Type I *trans*-AT PKSs are a promising source of antibiotics but the transcriptional activation of their gene cluster has been poorly explored (31). Due to their modular architecture (see Chapter I), most of the studies in *trans*-AT PKSs have focused on understanding their biosynthetic mechanisms, genome mining, and the prediction of products from their orphan clusters (29, 112, 143). This knowledge has been useful for the development of combinatorial biosynthesis, an approach to modify natural polyketides in order to improve their functionality (143). However, the ecological role and regulation of *trans*-AT PKSs has not been investigated in depth, despite the fact that this knowledge may also help to predict novel compounds and counteract the emergence bacterial resistance. Bacillaene has been a prototype of *trans*-AT PKS systems but the regulation of the *pks* gene cluster has not been studied (28, 36, 39–41, 144).

In this dissertation, I highlighted some of the intricate mechanisms used by *B. subtilis* to regulate the activation of the *pks* genes and produce bacillaene. I have found that multiple regulatory proteins coordinately modulate the expression of the *pks* genes at different stages of *B. subtilis* development, and discuss this phenomenon from the perspective of possible ecological roles during competition (Chapter II). I also proposed a more complex regulation of the *pks* genes that requires the orchestration of several regulatory elements. One of these regulatory elements is a *cis*-RNA element (BAR element) that contributes to the transcription elongation of almost 75 kb, a phenomenon

not previously reported for regulation of secondary metabolites (Chapter III). Finally, I described the role of PksA in regulation of *ymcC* gene expression, as opposed to the previously predicted function as the regulator of the *pks* genes (Chapter IV). In this chapter, I will discuss the conclusions and future directions that may arise from this research.

Bacillaene is an essential compound for *B. subtilis* survival during competition

The absence of bacillaene dramatically changes the interaction outcomes between *B. subtilis* and *Streptomyces spp*. Bacillaene was first reported as an inhibitor of protein synthesis in bacteria (34). This antibiotic effect seems to be of broad spectrum, causing growth inhibition in species such as *Escherichia coli, Staphylococcus aeureus, Pseudomonas aeruginosa* and some species of *Streptomyces* (34, 37). However, bacillaene is an unstable molecule when purified, so it is suspected that the concentration of the molecule is low in natural environments (41). In this context, bacillaene may have distinct functions as evidenced by the competitive assays between *B. subtilis* and *Streptomyces spp* (Chapter I). In addition to growth inhibitory activity, previous reports have shown that bacillaene interferes with the production of prodiginines by *S. coelicolor,* and in this dissertation I also show the same effect of bacillaene on *S. lividans* (37, 145)(Chapter II). Another outcome is observed in the absence of bacillaene during *B. subtilis* and *Streptomyces* sp. Mg1. During this interaction, bacillaene not only offers some type of protection to *B. subtilis* from the lytic effects of *S.* sp. Mg1, but also

prevents the competitor from developing aerial hyphae. Therefore, *B. subtilis* seem to gain multiple benefits from bacillaene and its different effects on competitors.

What functions might bacillaene have during competition? This is a long-standing question that still remains open. In one case, disruption of bacillaene production leads to increased prodiginines productions by *S. coelicolor* (37). One hypothesis to explain this observation is that when bacillaene production is disrupted, *B. subtilis* increases the production of other secondary metabolites. One of these other molecules could induce *S. coelicolor* to make prodiginines (48). On the other hand, the presence of bacillaene delays prodiginines production, as observed in *S. coelicolor* lawns that have been exposed to pure bacillaene (Chapter I). However, neither production nor absence of prodiginines by *S. coelicolor* appears to harm *B. subtilis*. Therefore, the benefits of bacillaene as modulator of prodiginines do not seem obvious. I speculate that bacillaene must affect other secondary metabolites in *S. coelicolor* to promote *B. subtilis* competitive advantage. Indeed, this speculation has some precedent established by IMS. In the absence of bacillaene, three unknown metabolites are induced in *S. coelicolor*, in addition to prodiginines (48). Whether these unknown metabolites or other undetected metabolites harm *B. subtilis* needs to be determined.

On the other hand, bacillaene also has a defensive function as revealed by the interaction of *B. subtilis* and *S.* sp. Mg1 (Chapter I). The protection offered by bacillaene resides in its capacity to delay the lytic effect on *B. subtilis* caused by *S.* sp. Mg1. In contrast, *B. subtilis* Δ*pks* strains, which do not produce bacillaene, are rapidly killed and lysed by S. sp. Mg1 (Unpublished work by R. Stubbendieck, Straight lab). Moreover, I

have also shown that bacillaene is able to inhibit *S.* sp. Mg1 aerial growth, a developmental process required to raise spores (Chapter I). Therefore, bacillaene may function by simply delaying spore development, which is usually coordinated with antibiotic production in many bacteria (3). Therefore, it is plausible that by inhibiting *S.* sp. Mg1 sporulation, bacillaene is also repressing other unknown secondary metabolites that are triggering the lysis of *B. subtilis*.

The competition phenotypes between *B. subtilis* and *S.* sp. Mg1 discussed in this dissertation may provide a better understanding of the mechanisms used by bacillaene to counteract *B. subtilis* competitors. Additional experiments at transcriptional and post-transcriptional effects of bacillaene on *S.* sp. Mg1 may provide insights on the mechanism of action(s) of bacillaene on this species of *Streptomyces.*

Bacterial competition enlightens our understanding of the differential activation of the *pks* genes

The apparent effects of bacillaene during competition prompted us to study the mechanisms of regulation used by *B. subtilis* to activate the *pks* gene cluster. The results of this dissertation support a competition model wherein *B. subtilis* inactivates bacillaene production during a motile phase in the presence of *S. lividans* while the static colony actively expresses the *pks* genes (Chapter II). The swarming population initially repressed *pks* expression, which over time is activated within the motile populations. The cell-type specific expression of the *pks* cluster raises questions about the ecological costs and benefits of producing bacillaene. As described before, production of bacillaene

requires a considerable expenditure of energy resources (Chapter I). The flagellar machinery is generally used for displacement in order to find new nutritional sources and its assemblage also requires significant resources (146). We speculate that maintaining low *pks* gene expression in motile populations may be important for energy resource allocation in *B. subtilis*, because synthesis of bacillaene is likely to require considerable energy input.

Cell specialization is a common mechanism within a *B. subtilis* population to promote division of labor among its members and express developmental features with a minimal cost of energy (57, 64). Some examples of this differentiation are spore-forming cells, producers of extracellular matrix, genetically competent, and motile cells (64). To ensure proper differentiation, each individual cell regulates an intricate transcriptional network of proteins that control hundreds of genes to activate or repress specific fates. The competitive interactions between *B. subtilis* and *S. lividans* allowed us to visualize the formation of motile and static subpopulations of *B. subtilis* (Chapter II). During the interactions, motile subpopulations delayed activation of the expression of the *pks* genes. Therefore, we hypothesized that *B. subtilis* cells used the transcriptional networks for cell development to activate the production of bacillaene.

What model fits the transcriptional network that controls expression of the *pks* genes? I have shown that activation of bacillaene production occurs due to a highly coordinated network of master regulators for cell development and nutrient conditions (Chapter II) (Fig. 5-1). The development master regulator, Spo0A, and the nutrition response regulator, CodY, appear to be the major players in activation of the *pks* genes.

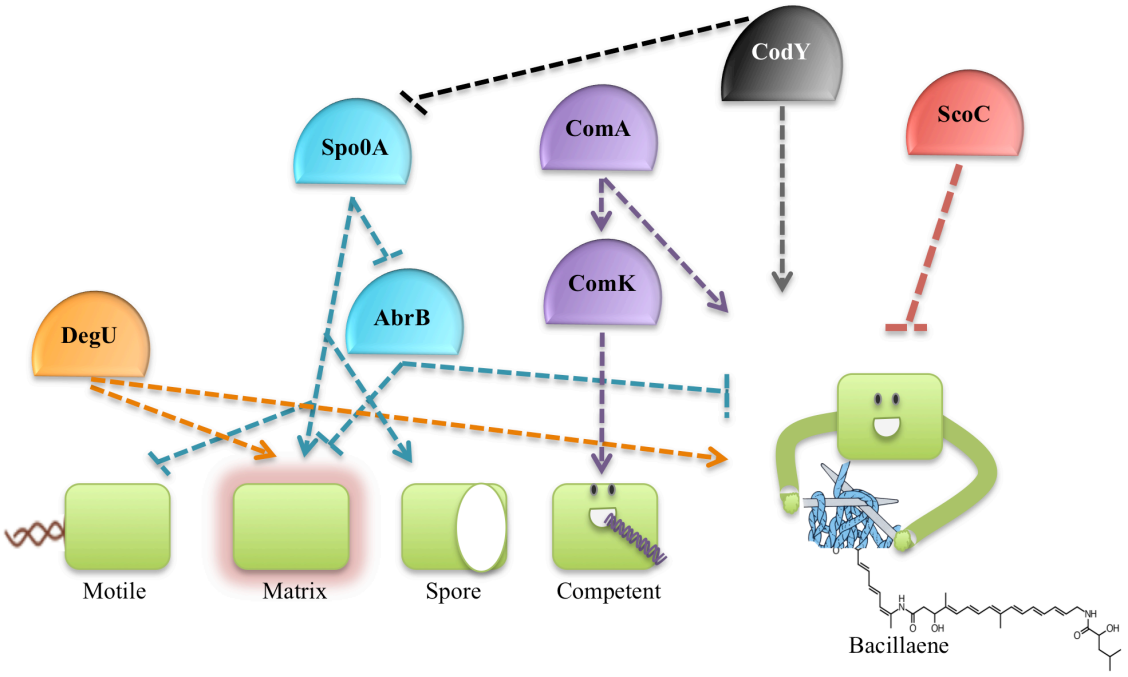


Figure 5- 1. Transcriptional network of the *pks* genes.
Multiple regulatory proteins coordinate the expression of the *pks* genes at the entry into stationary phase. These regulators also activate cell development in *B. subtilis*, leading to biofilm, competent and sporulating cells to produce bacillaene. In particular, the nutrient response regulator, CodY, and the development master regulator, Spo0A are the major regulators of the *pks* genes. Other master regulators such as DegU, ComA and ScoC are also required for full induction.

However, ComA, DegU and ScoC are also required for full induction, highlighting the complex transcriptional network for *pks* regulation.

DNA binding assays are necessary to confirm the master regulators that directly bind to the *pks* genes. There is some precedence that suggest that CodY directly binds the *pks* genes (69). A genome-wide identification of *B. subtilis* CodY binding sites determined that some of the strongest CodY binding sites are located within the *pks* genes (69). Specifically, 20 bp intergenic regions between *pksB* and *pksC*, *pksE* and *acpK*, and inside the coding sequence at *pksJ* and *pksN* were found to bind CodY. My data identified the promoter element between *pksB* and *pksC* that might bind CodY, but internal promoters at *pksE-acpK* and *pksJ-pksN* cannot be discarded and will need further confirmation. Alternatively, these regions may have auxiliary roles in facilitating the expression of the distal genes.

Similarly, a genome-wide binding profile of AbrB found multiple sites throughout the *pks* gene cluster, suggesting that the effect of Spo0A is through AbrB (83, 147). The data we present also suggest that activation of the *pks* genes by Spo0A occurs through the repression of the AbrB repressor, confirming the results of the previous report (Chapter II). Curiously, in the genome-wide binding profile study all the AbrB binding sites in the *pks* genes are within coding sequences and not in intergenic regions (83). AbrB was found to bind regions inside the coding sequence of *pksC*, *pksD*, *pksJ*, *pksL*, *pksN* and *pksR* (83). It is possible that AbrB acts as repressor of the pks genes by altering the topology of the DNA instead of overlapping promoter or activation sites. However, the same study also shows that the expression of the genes downstream

of *pksJ*, *pksL*, *pksN* and *pksR* are not significantly affected by the binding of AbrB (83). It has been previously proposed that AbrB may work as a nucleoid-associated protein given that in many other genes the AbrB binding does not affect transcription (83). Therefore it is possible that the binding of AbrB at distal *pks* genes is related to other biological functions such as replication and higher order of DNA organization, instead of production of bacillaene.

Promoters of the *pks* genes are dynamically activated at different developmental stages

In addition to the complex regulatory network of proteins that controls the expression of the *pks* genes, this dissertation also highlights the organization of several regulatory elements within the *pks* gene cluster. Four promoters along the *pks* genes control the activation of the *pks* genes, namely, P_{pksB}, P_{pksC}, P_{pksG} and P_{pksS}. P_{pksC} is the most active promoter in vegetatively growing cells. P_{pksB}, P_{pksC}, and P_{pksS} activate the *pks* gene cluster, at the entry into stationary phase. Remarkably, P_{pksG} seems to differ in its spatiotemporal activation with respect to the others. P_{pksG} is activated exclusively at late stationary phase and only inside spore-forming cells, but not mother cells.

Why does activation of P_{pksG} delays up to the entry into sporulation? The *pks* genes are dispensable for sporulation and germination, at least under the conditions I tested (Chapter III). However, two types of evidence, transcriptional and cytological, suggest a role of Pks enzymes associated with spores. First, P_{pksG} has a binding motif for SigF, a transcriptional sigma factor that induces differential expression of genes inside

the forespores. This SigF binding motif strongly induces the compartmentalized activation of P_{pksG} inside the forespore but not the mother cells, as observed by fluorescence microscopy (Chapter III). These results suggest that expression of genes downstream of P*pksG* is activated in the spores. Second, the Pks enzymatic complex that synthesizes bacillaene randomly localizes in the cells during vegetative growth. However, during initial stages of sporulation the majority of the Pks enzymatic complexes localize near the forespores (Appendix II). Taken together, these results suggest that production of bacillaene may have an additional, unknown function related with sporulation. Secondary metabolites that increase fitness during sporulation or germination are not unprecedented. Antibiotic production is known to be coordinated with sporulation for many streptomycetes (3). In some cases, antibiotics protect the dormant or germinating spore from consumption by other competitors or delay germination of spores until conditions become favorable ((3) and references within). In *B. subtilis* several secondary metabolites have also been shown to function as signaling molecules for biofilm formation and sporulation (64). Given the protein inhibitory properties of bacillaene, it is plausible to speculate that bacillaene may contribute to maintaining the dormancy of the spores and be secreted from them upon activation of germination. Therefore, further investigation on these subjects may unveil new functions for bacillaene.

To clarify the function of bacillaene, I propose to determine whether bacillaene is able to inhibit germination. Although the Δ*pks* mutant did not show a dramatic change in germination compared to wild type, it is possible that the addition of bacillaene delays

the germination and outgrowth of the cells. Bacillaene would need to be purified and dilution series may establish the minimal concentration of bacillaene to delay germination.

A novel *cis*-regulatory RNA to elongate the transcription of a long operon

The intergenic region between *pksC* and *pksD* was initially predicted to be a promoter of the *pks* genes but transcriptional reporters with YFP showed no activity, negating this hypothesis. However, the deletion of this region led to a significant reduction in the production of bacillaene (Chapter IV). Further analysis performed by Winkler lab (U. Maryland) have shown that the function of this intergenic region is to readthrough transcription terminators located at the *pksJ* locus (Grace Ellis, personal communication). Taken together, these results suggest a role of the *pksC* to *pksD* intergenic region as a *cis*-regulatory RNA that mediates processive antitermination (Fig. 5-2). We called this region 'BAR' (bacillaene-associated RNA). To our knowledge, this is the first report of an RNA regulatory element associated to the production of an antibiotic and the longest transcript controlled by a *cis*-RNA element.

Several questions are currently undergoing further investigation in collaboration with Winkler Lab. Is the BAR element specific for processive antitermination of the *pks* genes or promote read-through termination sites in general? Is BAR permanently active or there are specific conditions and mechanisms that activate an "on" state? What is the structural conformation adopted by BAR and how does conformation cause the read-through terminators?

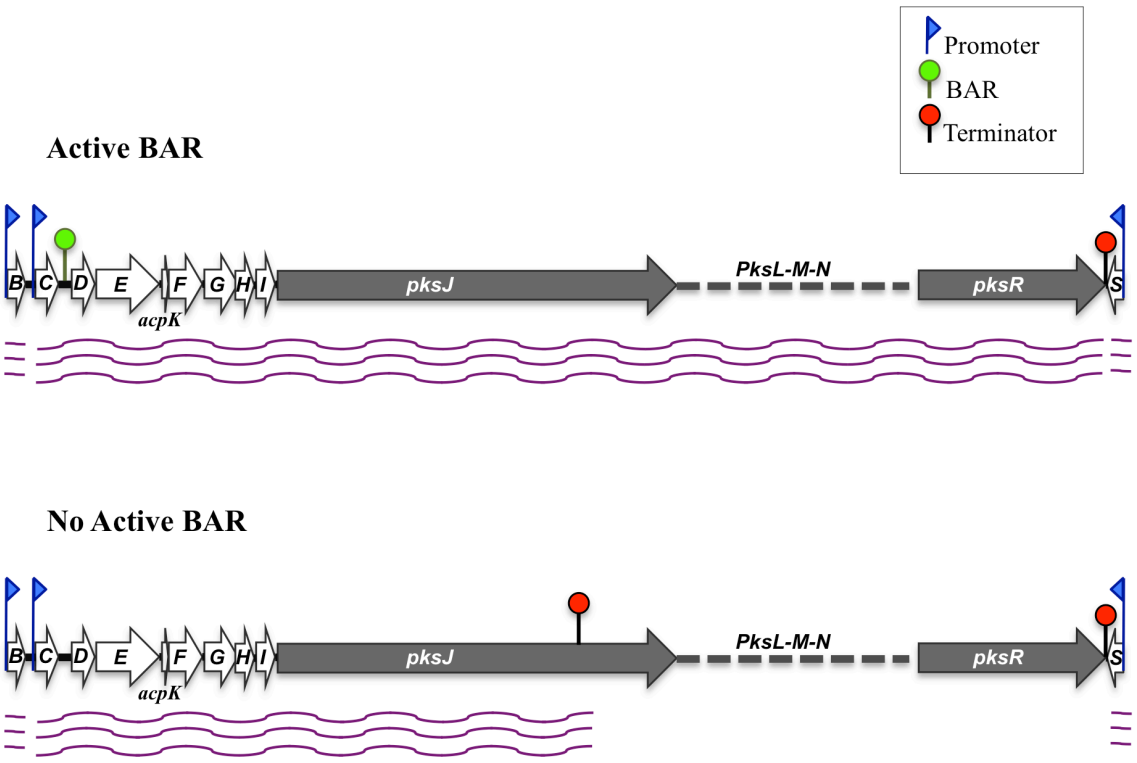


Figure 5- 2. Model for the processive antitermination promoted by the BAR element. The expression of the *pksC-pksR* transcript (~76 kb) begins with the activation of the *pksC* promoter. Transcription elongation initiates and produces the RNA regulatory element at *pksC* and *pksD* intergenic region (BAR element). The BAR element is able to associate with the RNA polymerase and modifies it to become resistant to downstream terminator and pause sites. In the absence of the BAR element, terminators present at the *pksJ* locus truncate the transcription.

PksA is not associated to a regulatory function over the *pks* genes

The regulation of *pks* gene transcription was previously assigned to the PksA protein, annotated as a TetR-family regulator of *pks* genes. I have shown that PksA does not regulate *pks* gene expression. Related organisms, such as *Bacillus amyloliquefaciens* FZB42 also produce bacillaene, encoding the enzymatic complex in the *bae* gene cluster (38). In contrast to *B. subtilis*, the orthologous *pksA* gene of *B. amyloliquefaciens* FZB42 is located almost 1kb apart from the *bae* biosynthetic gene cluster, suggesting that PksA is not a pathway-specific regulator (39). Thus these data support a model for *pks* gene regulation that relies on the global regulatory circuits already described and no pathway-specific regulator is involved in its regulation.

PksA functions as a TetR-like repressor over *ymcC*

I have also shown that the regulatory function of PksA is directed toward an adjacent, divergently transcribed gene, *ymcC.* The function of YmcC is unknown but is predicted to be a membrane protein. The data I present indicates that PksA represses expression of the *ymcC* gene during vegetative growth. In addition, localization of YmcC is observed during spore formation (Chapter IV). I speculate that the protein is involved in some aspect of spore maturation, germination or outgrowth. Further experiments are necessary to expand this hypothesis.

Conclusions

In summary, this dissertation contributes to a better understanding of the regulation of the *pks* genes in *B. subtilis*. The activation of the *pks* genes is more complex than originally hypothesized. Multiple regulatory proteins converge to activate the expression of the *pks* genes at the entry into stationary phase. These regulatory proteins are master regulators involved in developmental transitions and nutrient stress response. Therefore, the activation of the *pks* genes occur in specialized cells that also undergo transitions to competence, biofilm formation or sporulation. In addition, several regulatory elements located at the *pks* genes ensure the transcription of the entire gene cluster (~78 kb). Among these regulatory elements, we found a novel *cis*-RNA element that promotes readthrough of termination sites and guarantees the expression of the transcript from *pksC* to *pksR* (~76 kb). To our knowledge, this mechanism has not been previously reported for a natural product- encoding gene cluster. Last, the expression of the *pks* genes appears to be more complex during sporulation. Activation of a *pksG* promoter inside the forespores suggests differential expression of the *pks* genes during sporulation. Further experiments are necessary to confirm this hypothesis and understand the benefits of this compartmentalization.

The general model is depicted in Fig. 5-3. During early exponential phase of growth, the majority of *B. subtilis* cells are motile and repress the *pks* genes, presumably by the direct binding of AbrB (64, 83)(Fig. 5-3A). At the entry into stationary phase, the *pks* genes become fully activated (Fig. 5-3B). As nutrients become scarce, *B. subtilis*

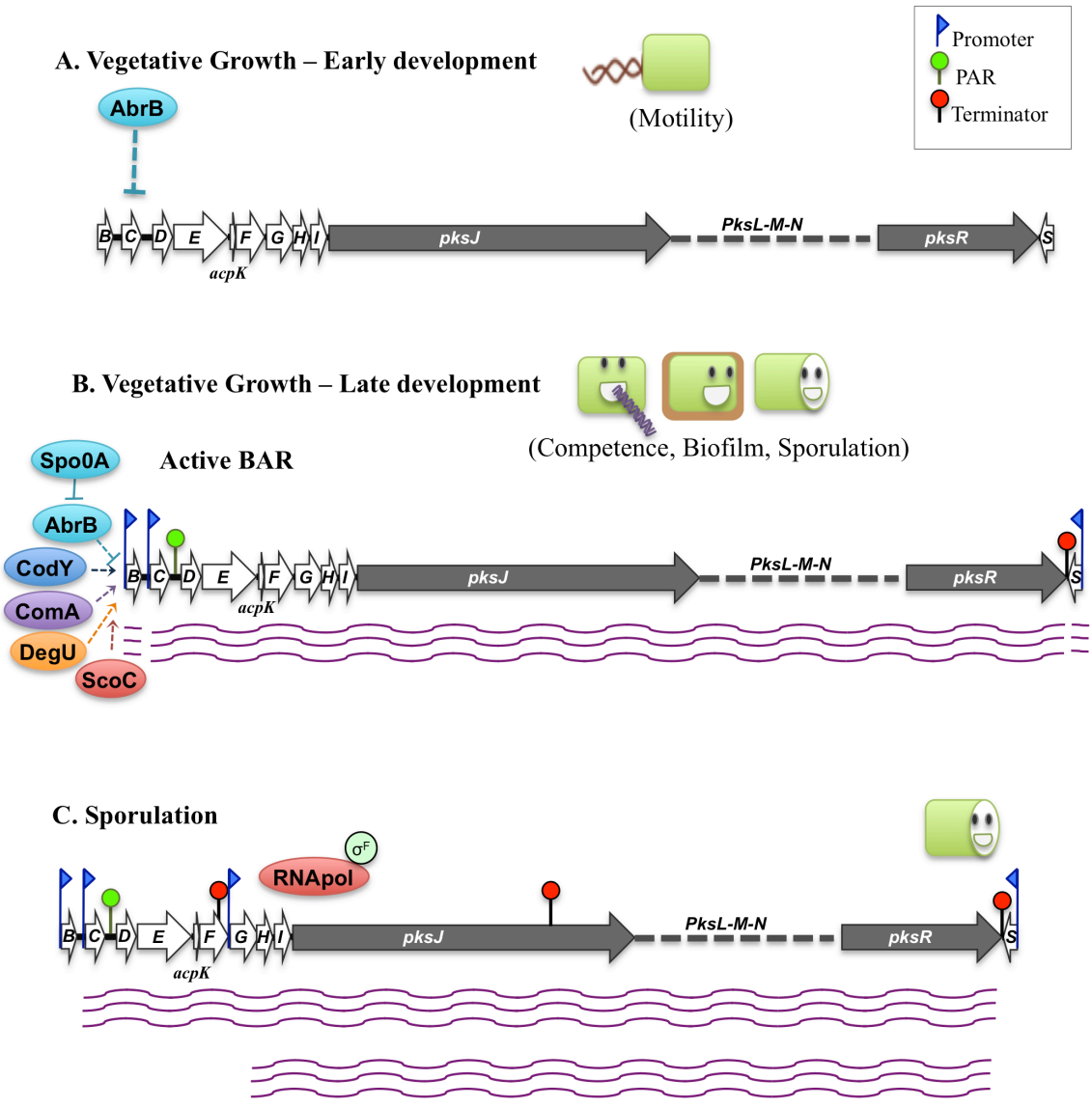


Figure 5- 3. General model for the regulation of the *pks* genes.
(A) *B. subtilis* cells growing at early-exponential phase are mainly motile and block the expression of the *pks* genes through the AbrB repressor. (B) At the entry into stationary phase, Spo0A is activated and blocks the expression of AbrB, releasing the *pks* genes for activation. CodY is also necessary for *pks* gene induction. Other transcriptional regulators play a secondary role in the expression of the *pks* genes. Promoters at *pksB*, *pksC* and *pksS* become active to express the entire gene cluster. The BAR element is transcribed and associates with the RNA polymerase to promoter processive antitermination. (C) During sporulation, a promoter at *pksG* that contains a SigmaF motif is recognized by the RNA polymerase and differentially activated inside the forespores but not the mother cells.

cells differentiate and specialize into developmental pathways, including competence, biofilm and sporulation (148). During these developmental stages, several master regulators become active (57, 64). The major players in the activation of the *pks* genes are the Spo0A and CodY stationary phase regulators. Secondary control mechanisms through DegU, ComA, and ScoC also contribute to *pks* gene expression. In addition, several regulatory elements located at the *pks* genes ensure the expression of the entire gene cluster. Active promoters are present upstream of *pksB*, *pksC* and *pksS*. A BAR element located at the intergenic region of *pksC* and *pksD* promotes read-through of transcription terminators and ensures the transcription from *pksC* to *pksR*. Finally, *B. subtilis* appear to differentially regulate the *pks* genes during sporulation (Fig. 5-3C). The exclusive activation of the *pksG* promoter inside the forespores suggests that *pksG* downstream genes are expressed and their products are compartmentalized inside the forespores.

This dissertation provides insights on the regulation of a *trans*-AT PKS and provides a model to investigate similar gene clusters present in other organisms.

REFERENCES

1. **Jucker E**. 2001. Progress in Drug Research. Molecules **6**:1060.
2. **Kossel A**. 1891. Über die chemische Zusammensetzung der Zelle (On the chemical composition of the cell). Arch Physiol **181**.
3. **Demain L, Fang A**. 2000. The natural functions of secondary metabolites. Adv. Biochem. Eng. Biotechnol. **69**:1–39.
4. **Monds R, O'Toole G**. 2008. Metabolites as intercellular signals for regulation of community-level traits, p. 105–130. *In* Winans, S, Bassler, B (eds.), Chemical Communication among Bacteria. ASM Press, Washington D.C.
5. **Surette MG, Davies J**. 2008. A new look at secondary metabolites, p. 307–322. *In* Winans, SC, Bassler, BL (eds.), Chemical Communication among Bacteria. ASM Press, Washington D.C.
6. **Price-Whelan A, Dietrich LE, Newman DK**. 2006. Rethinking "secondary" metabolism: physiological roles for phenazine antibiotics. Nat. Chem. Biol. **2**:71–78.
7. **Chain B**. 1991. Penicillin and beyond. Nature **353**:492–494.
8. **Comroe JJ**. 1978. Pay dirt: the story of streptomycin. Part I: from Waksman to Waksman. Am. Rev. Respir. Dis. **117**:773–781.
9. **Neu H**. 1992. The crisis in antibiotic resistance. Science. **257**: 1064-1073.
10. **Center for Disease Control US**. 2013. Antibiotic resistance threats in the United States 2013. Access Date: Aug 2014. http://www.cdc.gov/drugresistance/threat-report-2013/index.html

11. **World Health Organization**. 2014. Antimicrobial resistance: global report on surveillance. Access Date: Aug 2014. http://www.who.int/drugresistance/documents/surveillancereport/en/

12. **Wezel G van, McKenzie N, Nodwell J**. 2009. Applying the genetics of secondary metabolism in model actinomycetes to the discovery of new antibiotics. Methods Enzymol. **458**:117–141.

13. **Hede K**. 2014. An infectious arms race. Nature **509**:4–5.

14. **Mlot C**. 2009. Antibiotics in Nature : Beyond Biological Warfare. Science. **324**:1637–1639.

15. **Craney A, Ahmed S, Nodwell J**. 2013. Towards a new science of secondary metabolism. J. Antibiot. (Tokyo). **66**:387–400.

16. **Seyedsayamdost MR, Traxler MF, Clardy J, Kolter R**. 2012. Old meets new: using interspecies interactions to detect secondary metabolite production in actinomycetes. Methods Enzymol. **519**:89-109.

17. **Van Wezel GP, McDowall KJ**. 2011. The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. Nat. Prod. Rep. **28**:1311–33.

18. **Berdy J**. 2005. Bioactive Microbial Metabolites. J. Antibiot. (Tokyo). **58**:1–26.

19. **Stein T**. 2005. *Bacillus subtilis* antibiotics: structures, syntheses and specific functions. Mol. Microbiol. **56**:845–57.

20. **Guthals A, Watrous JD, Dorrestein PC, Bandeira N**. 2012. The spectral networks paradigm in high throughput mass spectrometry. Mol. Biosyst. **8**:2535–44.

21. **Weissman KJ, Müller R**. 2008. Protein-protein interactions in multienzyme megasynthetases. Chembiochem **9**:826–48.

22. **Hofemeister J, Conrad B, Adler B, Hofemeister B, Feesche J, Kucheryava N, Steinborn G, Franke P, Grammel N, Zwintscher A, Leenders F, Hitzeroth G, Vater J**. 2004. Genetic analysis of the biosynthesis of non-ribosomal peptide- and polyketide-like antibiotics, iron uptake and biofilm formation by *Bacillus subtilis* A1/3. Mol. Genet. genomics. **272**:363–378.

23. **Chen X, Koumoutsi A, Scholz R**. 2007. Comparative analysis of the complete genome sequence of the plant growth–promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat Biotechnol. **25**: 1007-1014

24. **Fischbach M a, Walsh CT**. 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem. Rev. **106**:3468–96.

25. **Jenke-Kodama H, Sandmann A, Müller R, Dittmann E**. 2005. Evolutionary implications of bacterial polyketide synthases. Mol. Biol. Evol **22**:2027–2039.

26. **Gross F, Luniak N, Perlova O, Gaitatzis N, Jenke-Kodama H, Gerth K, Gottschalk D, Dittmann E, Müller R**. 2006. Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel

secondary metabolites by heterologous expression in pseudomonads. Arch. Microbiol. **185**:28–38.

27. **Rawlings BJ**. 2001. Type I polyketide biosynthesis in bacteria (Part B) (1995 to mid-2000). Nat. Prod. Rep. **18**:231–281.

28. **Moldenhauer J, Götz DCG, Albert CR, Bischof SK, Schneider K, Süssmuth RD, Engeser M, Gross H, Bringmann G, Piel J**. 2010. The final steps of bacillaene biosynthesis in *Bacillus amyloliquefaciens* FZB42: direct evidence for beta, gamma dehydration by a trans-acyltransferase polyketide synthase. Angew. Chem. Int. Ed. Engl. **49**:1465–7.

29. **Nguyen T, Ishida K, Jenke-Kodama H, Dittmann E, Gurgui C, Hochmuth T, Taudien S, Platzer M, Hertweck C, Piel J**. 2008. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. Nat. Biotechnol. **26**:225–33.

30. **Chen X, Koumoutsi A, Scholz R, Borriss R**. 2009. More than anticipated – production of antibiotics and other secondary metabolites by *Bacillus amyloliquefaciens* FZB42. J. Mol. Microbiol. Biotechnol. **16**:14–24.

31. **Piel J**. 2010. Biosynthesis of polyketides by trans-AT polyketide synthases. Nat. Prod. Rep. **27**:996–1047.

32. **Cheng Y-Q, Tang G-L, Shen B**. 2003. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. Proc. Natl. Acad. Sci. U. S. A. **100**:3149–54.

33. **Scotti C, Piatti M, Cuzzoni A, Perani P, Tognoni A, Grandi G, Galizzi A, Albertini AM**. 1993. A Bacillus subtilis large ORF coding for a polypeptide highly similar to polyketide synthases. Gene. **130**:65–71.

34. **Patel PS, Huangn S, Fisher S, Pirnik D, Aklonis C, Dean L, Meyers E, Fernandes P, Mayerlm F**. 1995. Bacillaene, a novel inhibitor of procaryotic protein synthesis produced by *Bacillus subtilis*: Production , taxonomy, isolation, physico-chemical Activity. J. Antibiot. (Tokyo). **48**:997–1003.

35. **Cheng Y-Q, Coughlin JM, Lim S-K, Shen B**. 2009. Type I polyketide synthases that require discrete acyltransferases. Methods Enzymol. **459**:165-186.

36. **Butcher R a, Schroeder FC, Fischbach M a, Straight PD, Kolter R, Walsh CT, Clardy J**. 2007. The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. **104**:1506–9.

37. **Straight PD, Fischbach M a, Walsh CT, Rudner DZ, Kolter R**. 2007. A singular enzymatic megacomplex from *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. **104**:305–10.

38. **Chen X-H, Vater J, Piel J, Franke P, Scholz R, Schneider K, Koumoutsi A, Hitzeroth G, Grammel N, Strittmatter AW, Gottschalk G, Süssmuth RD, Borriss R**. 2006. Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB42. J. Bacteriol. **188**:4024–36.

39. **Chen XH, Koumoutsi a, Scholz R, Schneider K, Vater J, Süssmuth R, Piel J, Borriss R**. 2009. Genome analysis of *Bacillus amyloliquefaciens* FZB42 reveals its potential for biocontrol of plant pathogens. J. Biotechnol. **140**:27–37.

40. **Reddick JJ, Antolak SA, Raner GM**. 2007. PksS from *Bacillus subtilis* is a cytochrome P450 involved in bacillaene metabolism. Biochem. Biophys. Res. Commun. **358**:363–7.

41. **Moldenhauer J, Chen X-H, Borriss R, Piel J**. 2007. Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family. Angew. Chem. Int. Ed. Engl. **46**:8195–7.

42. **Linares JF, Gustafsson I, Baquero F, Martinez JL**. 2006. Antibiotics as intermicrobial signaling agents instead of weapons. Proc. Natl. Acad. Sci. U. S. A. **103**:19484–9.

43. **Martínez JL**. 2008. Antibiotics and antibiotic resistance genes in natural environments. Science **321**:365–7.

44. **Graff JR, Forschner-Dancause SR, Menden-Deuer S, Long R a, Rowley DC**. 2013. *Vibrio cholerae* exploits sub-lethal concentrations of a competitor-produced antibiotic to avoid toxic interactions. Front. Microbiol. **4**:1-11.

45. **Dietrich LEP, Teal TK, Price-Whelan A, Newman DK**. 2008. Redox-active antibiotics control gene expression and community behavior in divergent bacteria. Science **321**:1203–6.

46. **Goh E-B, Yim G, Tsui W, McClure J, Surette MG, Davies J**. 2002. Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. Proc. Natl. Acad. Sci. U. S. A. **99**:17025–30.

47. **Bobadilla Fazzini R a, Skindersoe ME, Bielecki P, Puchałka J, Givskov M, Martins sos Santos V a P**. 2013. Protoanemonin: a natural quorum sensing inhibitor that selectively activates iron starvation response. Environ. Microbiol. **15**:111–20.

48. **Yang Y-L, Xu Y, Straight P, Dorrestein PC**. 2009. Translating metabolic exchange with imaging mass spectrometry. Nat. Chem. Biol. **5**:885–7.

49. **Barger S, Hoefler B, Cubillos-Ruiz A, Russell W, Russell D, Straight P**. 2012. Imaging secondary metabolism of *Streptomyces* sp. Mg1 during cellular lysis and colony degradation of competing *Bacillus subtilis*. Antonie Van Leeuwenhoek **102**:435–45.

50. **Williamson NR, Fineran PC, Leeper FJ, Salmond GPC**. 2006. The biosynthesis and regulation of bacterial prodiginines. Nat. Rev. Microbiol. **4**:887–99.

51. **Straight PD, Willey JM, Kolter R**. 2006. Interactions between *Streptomyces coelicolor* and *Bacillus subtilis*: Role of surfactants in raising aerial structures. J. Bacteriol. **188**:4918–25.

52. **Hoefler B, Gorzelnik K**. 2012. Enzymatic resistance to the lipopeptide surfactin as identified through imaging mass spectrometry of bacterial competition. Proc. Natl. Acad. Sci. U. S. A. **109**:13082–13087.

53. **Rachid S, Gerth K, Kochems I, Müller R**. 2007. Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium *Sorangium cellulosum* So ce56. Mol. Microbiol. **63**:1783–96.

54. **Ou X, Zhang B, Zhang L, Zhao G, Ding X**. 2009. Characterization of *rrdA*, a TetR family protein gene involved in the regulation of secondary metabolism in *Streptomyces coelicolor*. Appl. Environ. Microbiol. **75**:2158–65.

55. **Bibb MJ**. 2005. Regulation of secondary metabolism in streptomycetes. Curr. Opin. Microbiol. **8**:208–15.

56. **Jones AC, Gerwick L, Gonzalez D, Dorrestein PC, Gerwick WH**. 2009. Transcriptional analysis of the jamaicamide gene cluster from the marine cyanobacterium *Lyngbya majuscula* and identification of possible regulatory proteins. BMC Microbiol. **9**:247.

57. **Lopez D, Vlamakis H, Kolter R**. 2009. Generation of multiple cell types in *Bacillus subtilis*. FEMS Microbiol. Rev. **33**:152–63.

58. **Kearns D, Losick R**. 2003. Swarming motility in undomesticated *Bacillus subtilis*. Mol. Microbiol. **49**:581–590.

59. **Lopez D, Fischbach MA, Chu F, Losick R, Kolter R**. 2009. Structurally diverse natural products that cause potassium leakage trigger multicellularity in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. **106**:280–285.

60. **López D, Vlamakis H, Losick R, Kolter R**. 2009. Paracrine signaling in a bacterium. Genes Dev. **23**:1631–8.

61. **Steller S, Vollenbroich D, Leenders F, Stein T, Conrad B, Hofemeister J, Jacques P, Thonart P VJ**. 1999. Structural and functional organization of the fengycin synthetase multienzyme system from *Bacillus subtilis* b213 and A1/3. Chem. Biol. **6**:31–41.

62. **Jacques P**. 2011. Surfactin and Other Lipopeptides from *Bacillus spp.*, p. 57–91. *In* Soberón-Chávez, G (ed.), Microbiology Monographs Series. Vol.20. Springer Berlin Heidelberg.

63. **Cundliffe E, Demain AL**. 2010. Avoidance of suicide in antibiotic-producing microbes. J. Ind. Microbiol. Biotechnol. **37**:643–72.

64. **López D, Kolter R**. 2010. Extracellular signals that define distinct and coexisting cell fates in *Bacillus subtilis*. FEMS Microbiol. Rev. **34**:134–49.

65. **Karatas A, Cetin S, Ozcengiz G**. 2003. The effects of insertional mutations in *comQ, comP, srfA, spo0H, spo0A* and *abrB* genes on bacilysin biosynthesis in *Bacillus subtilis*. Biochim. Biophys. Acta **1626**:51–56.

66. **Inaoka T, Wang G, Ochi K**. 2009. ScoC Regulates Bacilysin Production at the Transcription Level in *Bacillus subtilis*. J. Bacteriol. **191**:7367–7371.

67. **Nicolas P, Mäder U, Dervyn E, Rochat T, Leduc A, Pigeonneau N, Bidnenko E, Marchadier E, Hoebeke M, Aymerich S, Becher D, Bisicchia P, Botella E, Delumeau O, Doherty G, Denham EL, Fogg MJ, Fromion V, Goelzer A, Hansen A, Härtig E, Harwood CR, Homuth G, Jarmer H, Jules M, Klipp E, Le Chat L, Lecointe F, Lewis P, Liebermeister W, March A, Mars R a T, Nannapaneni P, Noone D, Pohl S, Rinn B, Rügheimer F, Sappa PK, Samson**

F, Schaffer M, Schwikowski B, Steil L, Stülke J, Wiegert T, Devine KM, Wilkinson AJ, van Dijl JM, Hecker M, Völker U, Bessières P, Noirot P. 2012. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. Science. **335**:1103–6.

68. **Strauch MA, Bobay BG, Cavanagh J, Yao F, Wilson A, Le Breton Y**. 2007. Abh and AbrB control of *Bacillus subtilis* antimicrobial gene expression. J. Bacteriol. **189**:7720–32.

69. **Belitsky B, Sonenshein A**. 2013. Genome-wide identification of *Bacillus subtilis* CodY-binding sites at single-nucleotide resolution. Proc. Natl. Acad. Sci. U.S.A. **110**:7026–7031.

70. **Sonenshein AL**. 2005. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr. Opin. Microbiol. **8**:203–7.

71. **Harwood CR, Cutting SM**. 1990. Molecular biological methods for *Bacillus*. Wiley, New York.

72. **Kearns DB, Chu F, Rudner R, Losick R**. 2004. Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol. Microbiol. **52**:357–69.

73. **Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB, Moorman AFM**. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucl. Acids Res. **37**:1–12.

74. **Wach A**. 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. Yeast **12**:259–265.

75. **Gibson D, Young L, Chuang R, Venter J, Hutchison III C, Smith H**. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods **6**:343–345.

76. **Miller JH**. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

77. **Guder A, Schmitter T, Wiedemann I, Sahl H-G, Bierbaum G**. 2002. Role of the single regulator MrsR1 and the two-component system MrsR2/K2 in the regulation of mersacidin production and immunity. Appl. Environ. Microbiol. **68**:106–113.

78. **Ramos JL, Martínez-bueno M, Molina-Henares A, Terán W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R**. 2005. The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. **69**:326–356.

79. **Kieser T, Bibb MJ, Buttner MJ, Chater KF**. 2000. Practical *Streptomyces* genetics. John Innes Foundation, Colney, Norwich NR4 7UH, England.

80. **Vlamakis H, Aguilar C, Losick R, Kolter R**. 2008. Control of cell fate by the formation of an architecturally complex bacterial community. Genes Dev. **22**:945–953.

81. **Shafikhani SH, Leighton T**. 2004. AbrB and Spo0E control the proper timing of sporulation in *Bacillus subtilis*. Curr. Microbiol. **48**:262–9.

82. **Hamon MA, Stanley NR, Britton RA, Grossman AD, Lazazzera BA**. 2004. Identification of AbrB-regulated genes involved in biofilm formation by *Bacillus subtilis*. Mol. Microbiol. **52**:847–860.

83. **Chumsakul O, Takahashi H, Oshima T, Hishimoto T, Kanaya S, Ogasawara N, Ishikawa S**. 2011. Genome-wide binding profiles of the *Bacillus subtilis* transition state regulator AbrB and its homolog Abh reveals their interactive role in transcriptional regulation. Nucleic Acids Res. **39**:414–428.

84. **Verhamme DT, Murray EJ, Stanley-Wall NR**. 2009. DegU and Spo0A jointly control transcription of two loci required for complex colony development by *Bacillus subtilis*. J. Bacteriol. **191**:100–8.

85. **Ogura M, Yamaguchi H, Yoshida Ki, Fujita Y, Tanaka T**. 2001. DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B. subtilis* two-component regulatory systems. Nucleic Acids Res. **29**:3804–13.

86. **Gupta M, Dixit M, Rao K**. 2013. Spo0A positively regulates *epr* expression by negating the repressive effect of co-repressors, SinR and ScoC, in *Bacillus subtilis*. J. Biosci. **38**:291–9.

87. **Inaoka T, Takahashi K, Ohnishi-Kameyama M, Yoshida M, Ochi K**. 2003. Guanine nucleotides guanosine 5’-diphosphate 3'-diphosphate and GTP co-operatively regulate the production of an antibiotic bacilysin in *Bacillus subtilis*. J. Biol. Chem. **278**:2169–76.

88. **Mirel DB, Chamberlin MJ**. 1989. The Bacillus subtilis flagellin gene (*hag*) is transcribed by the sigma 28 form of RNA polymerase. J. Bacteriol. **171**:3095–101.

89. **Branda SS, Chu F, Kearns DB, Losick R, Kolter R**. 2006. A major protein component of the *Bacillus subtilis* biofilm matrix. Mol. Microbiol. **59**:1229–1238.

90. **Connors MJ, Mason JM, Setlow P**. 1986. Cloning and nucleotide sequencing of genes for three small, acid- soluble proteins from *Bacillus subtilis* spores. J. Bacteriol. **166**:417–425.

91. **Ratnayake-lecamwasam M, Serror P, Wong K, Sonenshein AL**. 2001. *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. Genes Dev. **15**:1093–1103.

92. **Fujita M, González-pastor JE, Losick R**. 2005. High- and Low-Threshold Genes in the Spo0A Regulon of *Bacillus subtilis*. J. Bacteriol. **187**:1357–1368.

93. **Zafra O, Lamprecht-Grandío M, de Figueras CG, González-Pastor JE**. 2012. Extracellular DNA release by undomesticated *Bacillus subtilis* is regulated by early competence. PLoS One **7**:1–15.

94. **Fadda A, Fierro AC, Lemmens K, Monsieurs P, Engelen K, Marchal K**. 2009. Inferring the transcriptional network of *Bacillus subtilis*. Mol. Biosyst. **5**:1840–52.

95. **Ogura M, Yamaguchi H, Yoshida Ki, Fujita Y, Tanaka T**. 2001. DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of B.subtilis two-component regulatory systems. Nucleic Acids Res. **29**:3804–13.

96. **Verhamme D, Kiley T, Stanley-Wall N**. 2007. DegU co-ordinates multicellular behaviour exhibited by *Bacillus subtilis*. Mol. Microbiol. **65**:564–568.

97. **Kunst F, Msadek T, Bignon J, Rapoport G**. 1994. The DegS/DegU and ComP/ComA two-component systems are part of a network controlling degradative enzyme synthesis and competence in *Bacillus subtilis*. Res. Microbiol. **145**:393–402.

98. **Shafikhani SH, Mandic-mulec I, Strauch MA, Smith I, Leighton T**. 2002. Postexponential regulation of *sin* operon expression in *Bacillus subtilis*. J. Bacteriol. **184**:564–571.

99. **Tsuge K, Ano T, Hirai M, Nakamura Y, Shoda M**. 1999. The genes *degQ*, *pps*, and lpa-8 (*sfp*) are responsible for conversion of *Bacillus subtilis* 168 to plipastatin Production. Antimicrob. Agents Chemother. **43**:2183–2192.

100. **Tsuge K, Matsui K, Itaya M**. 2007. Production of the non-ribosomal peptide plipastatin in *Bacillus subtilis* regulated by three relevant gene blocks assembled in a single movable DNA segment. J. Biotechnol. **129**:592–603.

101. **Luo Y, Helmann JD**. 2009. Extracytoplasmic function sigma factors with overlapping promoter specificity regulate sublancin production in *Bacillus subtilis*. J. Bacteriol. **191**:4951–8.

102. **Molle V, Nakaura Y, Shivers RP, Yamaguchi H, Losick R, Fujita Y, Sonenshein AL**. 2003. Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. J. Bacteriol. **185**:1911–1922.

103. **Shivers RP, Sonenshein AL**. 2004. Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. Mol. Microbiol. **53**:599–611.

104. **Bergara F, Ibarra C, Iwamasa J, Patarroyo JC, Ma LM**. 2003. CodY is a nutritional repressor of flagellar gene expression in *Bacillus subtilis*. J. Bacteriol. **185**:3118–3126.

105. **Ishii H, Tanaka T, Ogura M**. 2013. The *Bacillus subtilis* response regulator gene *degU* is positively regulated by CcpA and by catabolite-repressed synthesis of ClpC. J. Bacteriol. **195**:193–201.

106. **Kobayashi K**. 2007. Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. Mol. Microbiol. **66**:395–409.

107. **Msadek T, Kunst F, Klier A, Rapoport G**. 1991. DegS-DegU and ComP-ComA modulator-effector pairs control expression of the *Bacillus subtilis* pleiotropic regulatory gene *degQ*. J. Bacteriol. **173**:2366–2377.

108. **Kodgire P, Rao K**. 2009. A dual mode of regulation of *flgM* by ScoC in *Bacillus subtilis*. Can. J. Microbioliogy **55**:983–989.

109. **Aguilar C, Vlamakis H, Losick R, Kolter R**. 2007. Thinking about *Bacillus subtilis* as a multicellular organism. Curr Opin Microbiol. **10**:638–643.

110. **Goel A, Wortel MT, Molenaar D, Teusink B**. 2012. Metabolic shifts: a fitness perspective for microbial cell factories. Biotechnol. Lett. **34**:2147–60.

111. **Breitling R, Achcar F, Takano E**. 2013. Modeling challenges in the synthetic biology of secondary metabolism. ACS Synth. Biol. **2**:373–8.

112. **Chiang Y-M, Chang S-L, Oakley BR, Wang CCC**. 2011. Recent advances in awakening silent biosynthetic gene clusters and linking orphan clusters to natural products in microorganisms. Curr. Opin. Chem. Biol. **15**:137–43.

113. **Tropel D, Meer J Van Der**. 2004. Bacterial transcriptional regulators for degradation pathways of aromatic compounds. Microbiol. Mol. Biol. Rev. **68**:474–500.

114. **Spira B, Aguena M, de Castro Oliveira JV, Yagil E**. 2010. Alternative promoters in the *pst* operon of *Escherichia coli*. Mol. Genet. Genomics. **284**: 489-498

115. **Roberts JW, Shankar S, Filter JJ**. 2008. RNA polymerase elongation factors. Annu. Rev. Microbiol. **62**:211–33.

116. **Yarnell WS, Roberts JW**. 1999. Mechanism of intrinsic transcription termination and antitermination. Science. **284**:611–615.

117. **Richardson JP**. 2002. Rho-dependent termination and ATPases in transcript termination. Biochim. Biophys. Acta **1577**:251–260.

118. **De Hoon MJL, Makita Y, Nakai K, Miyano S**. 2005. Prediction of transcriptional terminators in *Bacillus subtilis* and related species. PLoS Comput. Biol. **1**:e25.

119. **Yanofsky C**. 2004. The different roles of tryptophan transfer RNA in regulating *trp* operon expression in *E. coli* versus *B. subtilis*. Trends Genet. **20**:367–74.

120. **Irnov I, Winkler WC**. 2010. A regulatory RNA required for antitermination of biofilm and capsular polysaccharide operons in Bacillales. Mol. Microbiol. **76**:559–75.

121. **Winkler WC**. 2005. Metabolic monitoring by bacterial mRNAs. Arch. Microbiol. **183**:151–9.

122. **Laing E, Mersinias V, Smith CP, Hubbard SJ**. 2006. Analysis of gene expression in operons of *Streptomyces coelicolor*. Genome Biol. **7**:R46.

123. **Irnov I, Sharma CM, Vogel J, Winkler WC**. 2010. Identification of regulatory RNAs in *Bacillus subtilis*. Nucleic Acids Res. **38**:6637–51.

124. **Waters LS, Storz G**. 2009. Regulatory RNAs in bacteria. Cell **136**:615–28.

125. **Sterlini JM, Mandelstam J**. 1969. Commitment to sporulation in *Bacillus subtilis* and its relationship to development of actinomycin resistance. Biochem. J. **113**:29–37.

126. **Aguena M, Spira B**. 2003. RT-PCR of long prokaryotic operon transcripts without DNase treatment. J. Microbiol. Methods **55**:419–423.

127. **Cases I, de Lorenzo V, Ouzounis C**. 2003. Transcription regulation and environmental adaptation in bacteria. Trends Microbiol. **11**:247–248.

128. **Ahn SK, Cuthbertson L, Nodwell JR**. 2012. Genome context as a predictive tool for identifying regulatory targets of the TetR family transcriptional regulators. PLoS One **7**:e50562.

129. **Cuthbertson L, Nodwell JR**. 2013. The TetR family of regulators. Microbiol. Mol. Biol. Rev. **77**:440–75.

130. **Beck CF, Mutzel R, Barbe J, Muller W**. 1982. A Multifunctional Gene (*tetR*) Controls Tn10-Encoded Tetracycline Resistance. J. Bacteriol. **150**:633–642.

131. **Pompeani AJ, Irgon JJ, Berger MF, Bulyk ML, Wingreen NS, Bassler BL**. 2008. The *Vibrio harveyi* master quorum-sensing regulator, LuxR, a TetR-type protein is both an activator and a repressor: DNA recognition and binding specificity at target promoters. Mol. Microbiol. **70**:76–88.

132. **Christen S, Srinivas A, Bähler P, Zeller A, Pridmore D, Bieniossek C, Baumann U, Erni B**. 2006. Regulation of the Dha operon of *Lactococcus lactis*: a deviation from the rule followed by the TetR family of transcription regulators. J. Biol. Chem. **281**:23129–37.

133. **Brown T, Mackey K, Du T**. 2004. Analysis of RNA by northern and slot blot hybridization, p. 4.9.1–4.9.19. *In* Current Protocols in Molecular Biology. John Wiley & Sons, Inc., New York.

134. **Cuthbertson L, Ahn SK, Nodwell JR**. 2013. Deglycosylation as a mechanism of inducible antibiotic resistance revealed using a global relational tree for one-component regulators. Chem. Biol. **20**:232–40.

135. **Adams D**. 2009. Bacterial cell division: assembly, maintenance and disassembly of the Z ring. Nat. Rev. Microbiol.

136. **Stricker J, Maddox P, Salmon ED, Erickson HP**. 2002. Rapid assembly dynamics of the *Escherichia coli* FtsZ-ring demonstrated by fluorescence recovery after photobleaching. Proc. Natl. Acad. Sci. U. S. A. **99**:3171–5.

137. **Yepes A, Schneider J, Mielich B**. 2012. The biofilm formation defect of a Bacillus subtilis flotillin-defective mutant involves the protease FtsH. Mol. Microbiol. **86**:457–71.

138. **Orth P, Saenger W, Hinrichs W**. 1999. Tetracycline-chelated Mg2+ ion initiates helix unwinding in Tet repressor induction. Biochemistry **38**:191–8.

139. **Moir A, Lafferty E, Smith D**. 1979. Genetic analysis of spore germination mutants of Bacillus subtilis 168: the correlation of phenotype with map location. J. Gen. Microbiol. **111**:165–180.

140. **Moir a**. 2006. How do spores germinate? J. Appl. Microbiol. **101**:526–30.

141. **Li JW-H, Vederas JC**. 2009. Drug discovery and natural products: end of an era or an endless frontier? Science **325**:161–5.

142. **Lee HH, Collins JJ**. 2011. Microbial environments confound antibiotic efficacy. Nat. Chem. Biol. **8**:6–9.

143. **Till M, Race PR**. 2014. Progress challenges and opportunities for the re-engineering of trans-AT polyketide synthases. Biotechnol. Lett. **36**:877–88.

144. **Cuskin F, Solovyova AS, Lewis RJ, Race PR**. 2011. Crystallization and preliminary X-ray analysis of the bacillaene synthase trans-acting acyltransferase PksC. Act. Cryst. **67**:464–6.

145. **Straight PD, Kolter R**. 2009. Interspecies chemical communication in bacterial development. Annu. Rev. Microbiol. **63**:99–118.

146. **Guttenplan SB, Shaw S, Kearns DB**. 2013. The cell biology of peritrichous flagella in *Bacillus subtilis*. Mol. Microbiol. **87**:211–29.

147. **Strauch M, Webb V, Spiegelman G, Hoch J**. 1990. The SpoOA protein of *Bacillus subtilis* is a repressor of the *abrB* gene. Proc. Natl. Acad. Sci. U. S. A. **87**:1801–5.

148. **Zuber P**. 2009. Management of oxidative stress in *Bacillus*. Annu. Rev. Microbiol. **63**:575–97.

149. **Vargas-Bautista C, Rahlwes K, Straight P**. 2014. Bacterial Competition Reveals Differential Regulation of the *pks* Genes by *Bacillus subtilis*. J. Bacteriol. **196**:717–28.

APPENDIX I

PRODUCTION OF BACILLAENE IN Δ*pksA* MUTANT AND CHARACTERIZATION OF SPREADING POPULATIONS OF *B. subtilis*

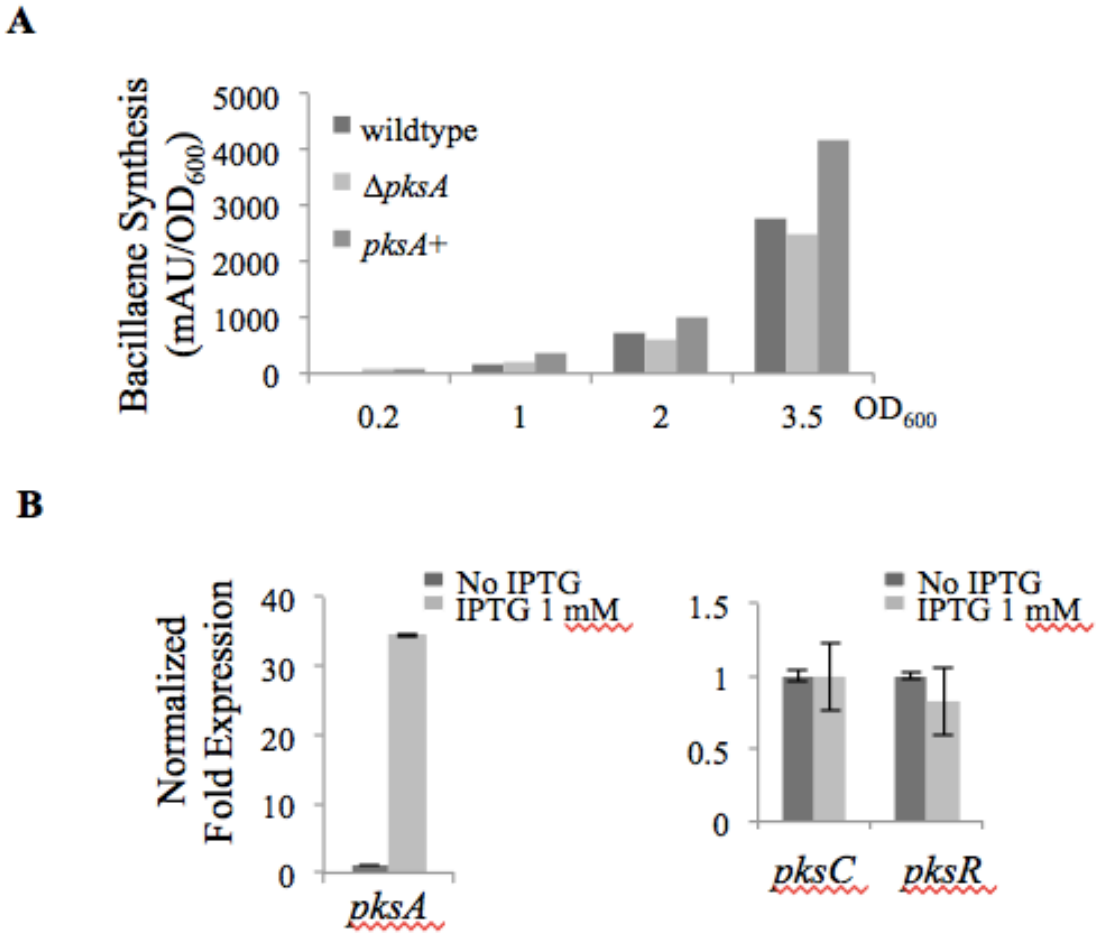


Figure AI-1. (A) Production of bacillaene over time in wild type, Δ*pksA* and *pksA*+. HPLC was performed as described in Materials and Methods. (B) qRT-PCR of *pksA, pksC* and *pksR* in a strain with PksA under the control of P_{hyper}. Similar conditions as described above were used to determine the mRNA abundance of *pksC* and *pksR* after overexpression of *pksA* with 1.0 mM IPTG. 35-fold overexpression of PksA was observed when culture cells where induced with IPTG. No fold changes in mRNA abundance of *pksC* and *pksR* were found.

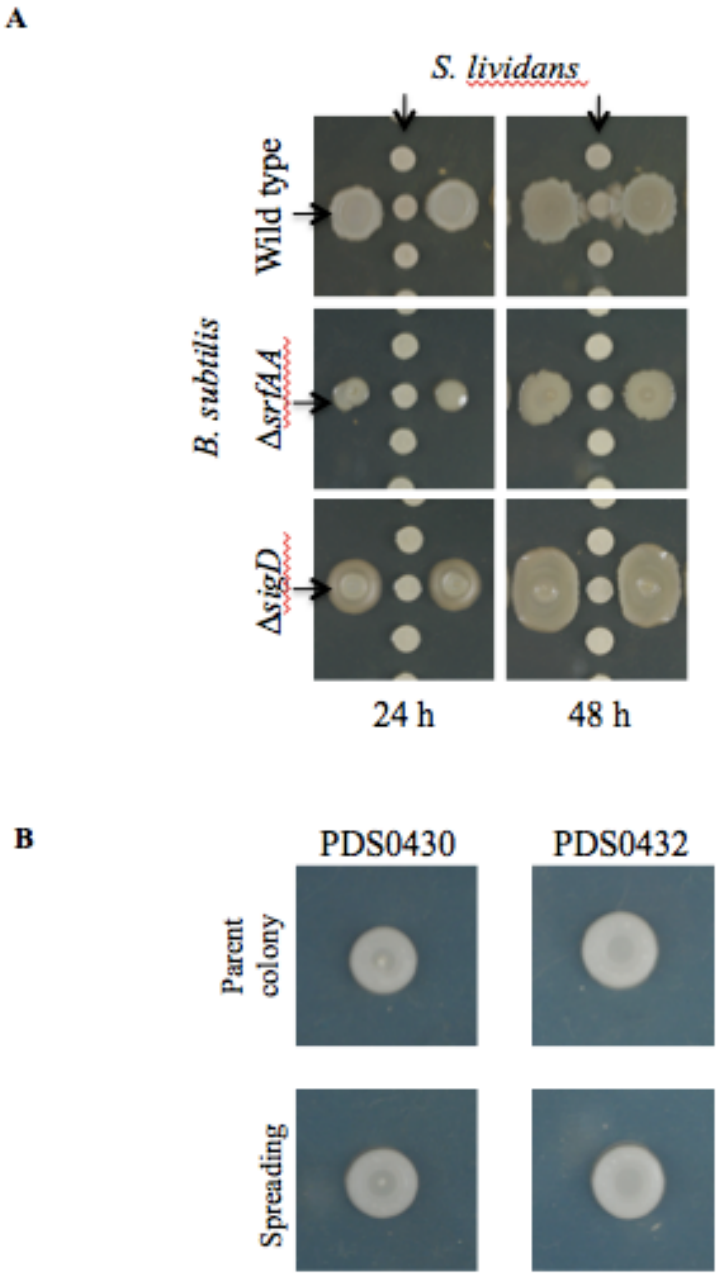


Figure AI-2. Spreading subpopulation resembles swarming motility (A) Co-culture of *S. lividans* and *B. subtilis* wild type, Δ*srfAA* and Δ*sigD* mutants. Conditions are similar to those described in Figure 2. *B. subtilis* mutants that lack *srfAA* or *sigD* genes were unable to spread towards *S. lividans* (B) Cells of PDS0430 (P_{hag}-YFP, P_{pksC}-YFP) and PDS0432 (P_{tapA}-YFP, P_{pksC}-YFP) were picked using a toothpick from the spreading subpopulations and plated on fresh agar medium in parallel with an inoculum from the original parent colony. No spreading occurred for either inoculation source and all colonies maintained the parent morphology in the absence of *S. lividans*.

APPENDIX II

LOCALIZATION OF THE Pks ENZYMATIC COMPLEX DURING SPORULATION

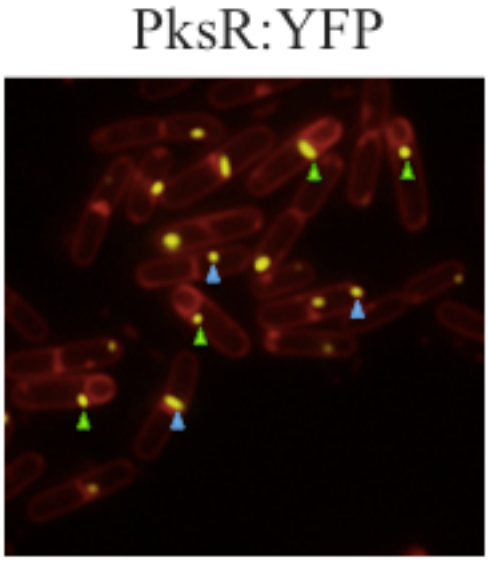


Figure AII-1. Localization of the Pks enzymatic complex at stationary phase (at OD600 = 3.17). A B. subtilis strain that contains a translational fusion of YFP to the C-terminal end of the PksR protein (PSK0212) was used to visualize the localization of the Pks enzymatic complex. Cells were grown in CH medium at 37˚C and samples were taken over time to visualize the Pks enzymatic complex by fluorescence microscopy, as described in Chapter II and Table A2-1. Pks enzymatic complex randomly locates in the vegetative cells (blue head arrows) but seems to have preferential location by the polar septum membrane during spore formation (green head arrows).

OD_{600}	% PksR-YFP	% Forespores	% Forespore with PksR	% pksR at Forespore
0.66	9	0	0	0
1.06	9	0	0	0
1.39	17	0	0	0
1.44	19	0	0	0
1.66	49	0	0	0
1.84	81	0	0	0
3.17	68	3	100	100
4.07	63	10	89	89
5.26	61	3	93	73
6.46	43	4	81	63
8.21	66	3	88	50

Table AII-1. Pks enzymatic complex preferentially locates in the membrane between the mother cell and the forespore during stationary phase. Cells of the strain PSK0212 were grown in CH medium and observed over time under fluorescence microscopy, as described in Fig. A2-1. During vegetative growth, the Pks enzymatic complex was observed to randomly localize in the cells. Formation spores begun at OD_{600} = 3.17 and preferential location of the Pks enzymatic complex was observed at the polar septum membrane (green). The percentage of forespores that harbor the Pks enzymatic complex decreased as the culture goes into late stationary phase.

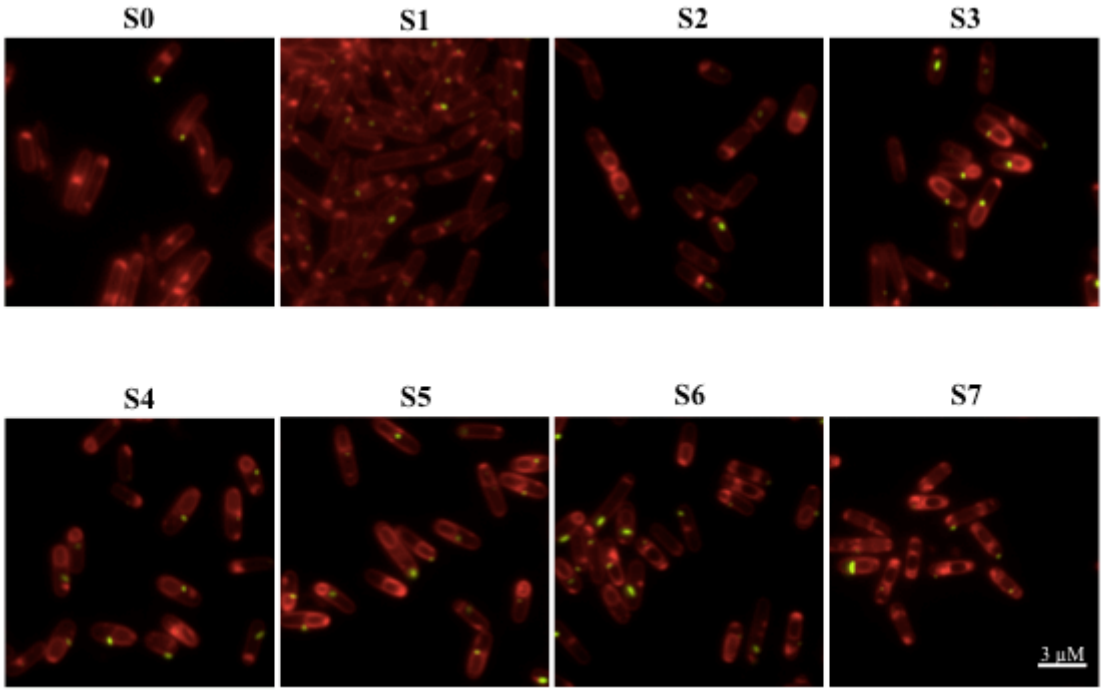


Figure AII-2. Localization of Pks enzymatic complex during induction of sporulation. Cells of PKS0212 were induced to sporulate and samples were taken hourly to visualize under fluorescence microscopy, using the same protocol as in Chapter III and figure 3-2B. Fluorescence images of PksR-YFP (green) in cells stained with TMA-DPH (red) to visualize membranes. Before initiation of sporulation (S0), few vegetative cells had the Pks enzymatic complex and it randomly localized. As spore formation developed, more cells contained the megacomplex and localization by the polar septum became more apparent (S1-S4). Later stages of sporulation (S5- S7) seemed to lack preferential localization of the Pks megacomplex. . Scale bar, 3 µM

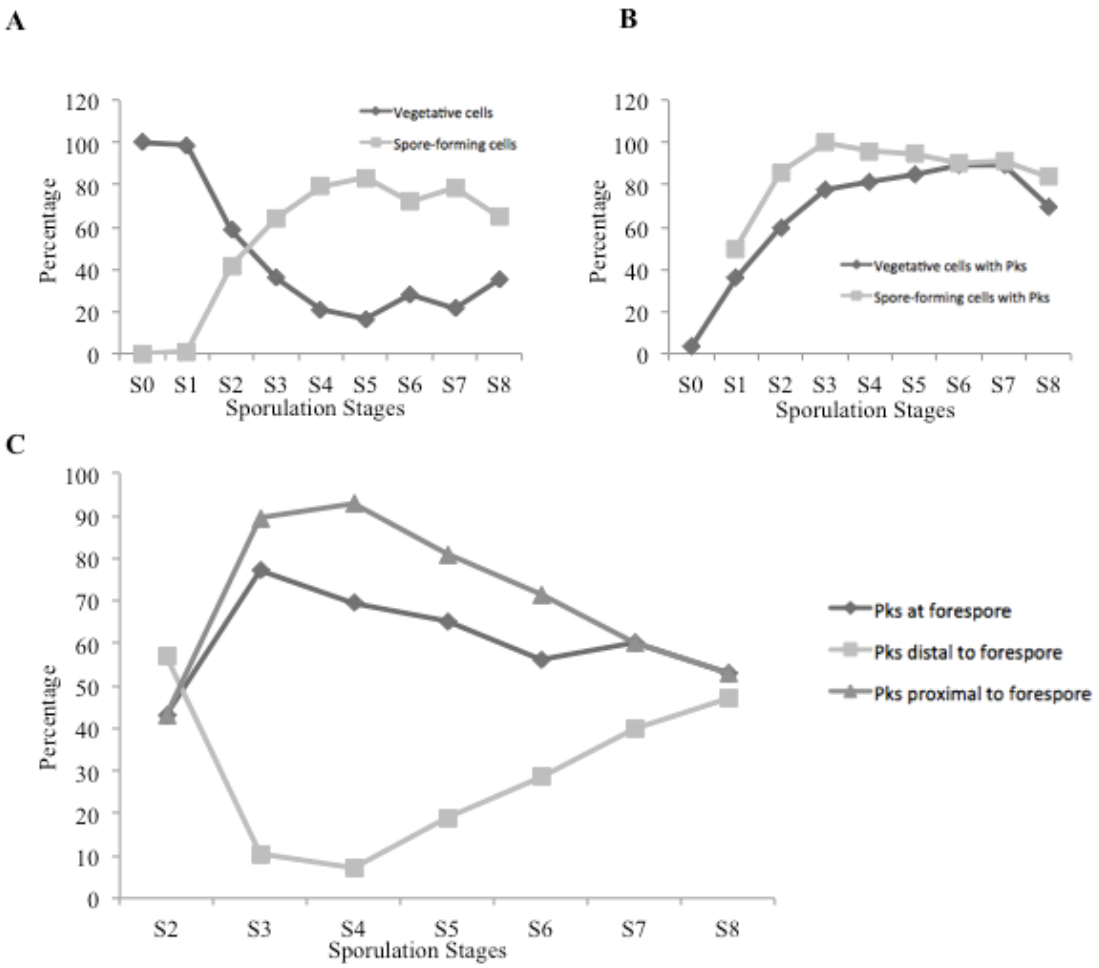


Figure AII-2. Analysis of the localization of the Pks enzymatic complex during induction of sporulation. Cells observed under fluorescence microscopy in Fig. A2-2 were counted and formation of spores and localization of Pks enzymatic complexes were calculated and plotted. (A) Percentage of spore formation over time. (B) Percentage of Pks enzymatic complex in vegetative cells and spore-forming cells. (C) Location of Pks enzymatic complexes in vegetative cells and spore-forming cells. Taken together, the results suggest that the Pks enzymatic complex has a preferential localization at the polar septum membrane during S3 and S4 stages of sporulation