Humboldt-Universität zu Berlin

D i s s e r t a t i o n

Functional genome analysis of the plant-growth promoting bacterium *Bacillus amyloliquefaciens* strain FZB42; characterizing its production and regulation of nonribosomal peptide synthetases

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

Mathematisch-Naturwissenschaftlichen Fakultät I

Dipl. Chem. Alexandra Koumoutsi

Dekan: **Prof. Dr. Christian Limberg**

Gutachter: 1. **Professor Dr. Rainer Borriss**

2. **PD Dr. habil. Joachim Vater**

eingereicht: 11.10.2006

Datum der Promotion: **04.12.2006**

List of contents

List of figures

List of tables

Abbreviations

aa	amino acid
aa-tRNA	aminoacyl-transfer RNA
ABC transporter	ATP-binding cassette transporter
ACS medium	sucrose-ammonium citrate medium
Ap	ampicillin
APS	ammoniumpersulfate
ATP	adenosine-5'-triphosphate
BSA	Bovine Serum Albumin
cpm	counts per minute
DIG	digoxigenin
FAS	fatty acid synthase
Fig.	figure
h	hours
IM-HK	intramembrane-sensing histidine kinases
IPTG	isopropyl β-D-thiogalactoside
Km	kanamycin
LB	Luria-Broth
MALDI-TOF MS	matrix-assisted laser desorption/ionization-time of flight mass spectrometry
min	minutes
mRNA	messenger RNA
MS	mass spectrometry
MU	Miller units
NRPS	nonribosomal peptide synthetase
OD	optical density
ORF	open reading frame
P_{bmy}, P_{yczE}	promoter of the *bmy* operon, promoter of *yczE*
PCR	polymerase chain reaction
PKS	polyketide synthase
4'-PP cofactor	4'-phosphopantetheine cofactor
PPi	pyrophosphate
4'-PPTase	4'-phosphopantetheinyl transferase

Rap	response regulator aspartate phosphatase
rpm	rounds per minute
SDS	Sodiumdodecylsulfate
sec	seconds
TCS	two-component regulatory system
T_{den}, T_{anneal}, T_{ext}	denaturating, annealing and extension temperature
TEMED	N,N,N,N-Tetramethylethylendiamine
TRP	tetratricopeptide
X-Gal	5-Bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

Abbreviations for domains present on NRPSs, FASs and PKSs

A	Adenylation
ACP	acyl carrier protein
AT	Acyltransferase
C	Condensation
C-MT	*C*-Methyltransferase
Cy	Heterocyclization
DH	β-hydroxyacyl-ACP dehydratase
E	epimerization
ER	enoyl-ACP reductase
KR	ketoacyl-ACP reductase
KS	keto synthase
N-MT	*N*-Methyltransferase
PCP	peptidyl carrier protein
T	thiolation
TE	thioesterase
TE II	thioesterase II

Summary

Bacillus amyloliquefaciens is a Gram-positive bacterium that is widely distributed in the soil. It colonizes the plant roots and several of its natural isolates, such as the FZB42 strain, are used as bio-fertilizers, since they can promote plant growth and suppress plant pathogenic organisms. The features and mechanisms governing the biocontrol-related function of the strain have not yet been fully characterized. The domesticated strain of *B. subtilis* 168, a model organism for studies on Gram-positive bacteria, is closely related to *B. amyloliquefaciens* FZB42, but does not promote plant growth.

As a first approach to detect gene differentiation between *B. amyloliquefaciens* FZB42 and *B. subtilis* 168, and since only the genome sequence of the latter was known at that point, Suppression Subtractive Hybridization (SSH) was employed. Thereby, several unique genes of *B. amyloliquefaciens* FZB42 could be identified. Among others, it was established that the genome of *B. amyloliquefaciens* FZB42 harbours genes with high similarity to nonribosomal peptide synthetases and polyketide synthases of various *Bacillus* species, yet different from the ones present in the genome of *B. subtilis* 168.

Meanwhile, our laboratory became engaged in a project aiming to define the complete genome sequence of *B. amyloliquefaciens* FZB42, in collaboration with the GenoMik Network in Göttingen. The major part of the work and the co-ordination of the whole process were performed by Xiao-Hua Chen and myself. Shotgun and fosmid library approaches, primer walking and multiplex PCR were used in order to decipher the complete sequence of *B. amyloliquefaciens* FZB42. Sequencing of the whole genome has since been completed and the second round of annotation is currently in process (performed by Xiao-Hua Chen).

Strain FZB42 is the first member of the *B. amyloliquefaciens* species to have its genome sequenced. The genome of strain FZB42 consists of a single circular chromosome of 3916 kb, and thus is smaller than that of *B. subtilis* 168 (4214 kb). It contains 3931 genes, 80% of which show more than 50% amino acid similarity to genes of *B. subtilis* 168. Comparative genome analysis revealed several characteristics of the bacterium that might be associated with its biocontrol activity. Striking is the presence of eight gene clusters that control the non-conventional synthesis of secondary metabolites, some of which with reported antifungal and antibacterial activities.

B. amyloliquefaciens FZB42 possesses the *srf*, *fen*, *pks1* (*bae*), *bac* and *dhb* operons, which are responsible for the synthesis of surfactin, fengycin, bacillaene, bacilysin and bacillibactin, respectively, and are also shared by *B. subtilis* 168. In addition, and as initially detected by the SSH experiments, the genome of *B. amyloliquefaciens* FZB42 contains the

bmy, *dif*, *pks2* gene clusters, which control the synthesis of bacillomycin D, difficidin/oxydifficidin and macrolactin respectively. The functionality of all eight gene-clusters was verified by a series of mass spectrometry analysis (MALDI-TOF MS and HPLC-ESI MS), in collaboration with Xiao-Hua Chen and Dr J. Vater. It is conceivable that the profound genetic capacity of *B. amyloliquefaciens* FZB42 to produce antagonistically acting secondary metabolites enables the strain to cope successfully with competitors within its natural environment and to promote plant growth. Therefore, the biological activity of those compounds was further examined. Bacillomycin D and fengycin were the only antibiotics produced by the strain, which could exhibit a general inhibitory effect on fungal growth, acting in a synergistic manner.

A further issue pursued in this work was to identify the regulatory pathways that govern the expression of bacillomycin D. Global regulators, such as DegU, DegQ and ComA, the alternative sigma factors, σ^B and σ^H, and a novel Rap protein were found to affect the transcriptional activation of the main promoter of the *bmy* operon identified in this work. In particular, DegU was shown to mediate its effects, after binding directly to two distinct A/T-rich sites at the *bmy* promoter region. The other regulatory players were associated with more indirect effects, which were mostly exerted via DegQ, a protein that seems to optimise the activity of DegU, or via DegU itself.

DegU was shown to play an additional role on bacillomycin D production, presumably a post-transcriptional one, apart from activating the main promoter of the *bmy* operon. Therefore, its presence was critical for the production of bacillomycin D. Similarly, YczE, a membrane protein of unknown function, encoded adjacently to *sfp* (a 4'-phosphopantetheinyl transferase that post-translationally modifies nonribosomal peptide synthetases and makes them functionally active), proved to be essential for bacillomycin D production, but dispensable for the production of the rest peptide antibiotics produced by *B. amyloliquefaciens* FZB42. The effect was mediated at a post-transcriptional level (prior to the peptide's export) and was independent of DegU.

To conclude, this work provides information concerning the genetic identity of *B. amyloliquefaciens* FZB42, its lifestyle and its production of secondary metabolites by it. In addition, it defines the complex regulatory network that controls the expression of the most abundant lipopeptide of the organism, bacillomycin D. It is the first time that the gene expression of a member of the iturin-group antibiotics has been monitored.

Zusammenfassung

Bacillus amyloliquefaciens ist ein im Boden weit verbreitetes Gram-positives Bakterium. Es kolonisiert Pflanzenwurzeln und mehrere natürliche Isolate, wie zum Beispiel der Stamm FZB42 werden als Biodünger verwendet, da sie in der Lage sind, Pflanzenwachstum zu fördern und Pflanzenpathogene zu unterdrücken. Die Eigenschaften und Mechanismen, welche diese Biokontrollfunktionen steuern wurden bislang noch nicht vollständig charakterisiert. Der domestizierte Stamm *B. subtilis* 168, ein Modellorganismus für Studien an Gram-positiven Bakterien, ist eng verwand mit *B. amyloliquefaciens* FZB42, fördert jedoch kein Pflanzenwachstum.

Als ein erster Ansatz zur Ermittlung von Gendifferenzen zwischen *B. amyloliquefaciens* FZB42 und *B. subtilis* 168 - wobei zum damaligen Zeitpunkt nur die Genomsequenz letzteren Organismus bekannt war - wurde die "Supression Subtractive Hybridisation" (SSH) angewandt. Hierdurch wurden mehrere einzigartige Gene in *B. amyloliquefaziens* identifiziert. Unter anderem wurde gezeigt, dass das Genom von B. *amyloliquefaziens* FZB42 Gene mit starker Ähnlichkeit zu nichtribosomalen Peptid-Synthetasen und Polyketid-Synthasen verschiedener *Bacillus*-Arten beinhaltet, die sich jedoch von den im *B. subtilis* 168-Genom enthaltenen Genen unterscheiden.

Unterdessen beteiligte sich unser Labor in Kollaboration mit dem GenoMik Network in Göttingen an einem Projekt, dessen Ziel die komplette Sequenzierung des Genoms von *B. amyloliquefaciens* war. Der Hauptanteil der Arbeit, sowie die Koordination des gesamten Projekts wurden von Xiao-Hua Chen und mir selbst durchgeführt. Zur Entschlüsslung der vollständigen Genomsequenz von *B. amyloliquefaciens* wurden Shotgun und Fosmid-Library Ansätze, Primer walking und Multiplex-PCR angewandt. Die Sequenzierung des gesamten Genoms wurde mittlerweile abgeschlossen und derzeitige Arbeiten sind bis zur zweiten Annotationsrunde vorangeschritten (durchgeführt von Xiao-Hua Chen).

Der Stamm FZB42 ist das erste Mitglied der *B. amyloliquefaziens*-Art, dessen Genom sequenziert wurde. Das Genome von Stamm FZB42 besitzt ein einziges kreisförmiges und 3916 kb großes Chromosom, das damit kleiner ist als das Chromosom von *B. subtilis* 168 (4214 kb). Es enthält 3931 Gene, von denen 80% mehr als 50%ige Aminosäuren-Ähnlichkeit mit Genen von *B. subtilis* zeigen. Vergleichende Genomanalysen offenbarten mehrere Charakteristika des Bakteriums, welche mit seiner Biokontrollaktivität assoziiert sein könnten. Auffällig ist die Präsenz von acht Genclustern, die die unkonventionelle Synthese von sekundären Metaboliten kontrollieren, von denen einige bereits beschriebene antifungale und antibakterielle Aktivitäten besitzen.

B. amyloliquefaciens FZB42 besitzt die *srf*, *fen*, *pks1* (*bae*), *bac* und *dhb* Operons, welche für die Synthese von Surfactin, Fengycin, Bacillaene, Bacilysin und Bacillibactin verantwortlich sind und die ebenfalls im Genom von *B. subtilis* 168 enthalten sind. Wie bereits durch die anfänglichen SSH-Experimente gezeigt worden war, beinhaltet das Genom von *B. amyloliquefaciens* FZB42 die *bmy*-, *dif*-, *pks2*-Gencluster, die die Synthese von Bacillomycin D, Difficidin/Oxydifficidin und Macrolactin kontrollieren. Die Funktionalität dieser acht Gencluster wurde in Zusammenarbeit mit Xiao-Hua Chen und Dr. J. Vater durch eine Serie von Massenspektrometrie-Analysen (MALDI-TOF MS and HPLC-ESI MS) nachgewiesen. Es ist vorstellbar, dass die umfangreiche genetische Kapazität, antagonistisch wirkende sekundäre Metabolite zu produzieren, es *B. amyloliquefaciens* FZB42 ermöglicht, erfolgreich gegen Konkurrenten in seiner natürlichen Umgebung vorzugehen und das Pflanzenwachstum zu fördern. Daher wurde die biologische Aktivität dieser Komponenten weiter untersucht. Bacillomycin D und Fengycin waren die einzigen von diesem Stamm produzierten Antibiotika, welche einen generellen inhibitorischen Effekt auf das Wachstum von Pilzen zeigten, wobei sie in synergistischer Weise wirkten.

Ein weiteres in dieser Arbeit verfolgtes Ziel war die Identifizierung der regulatorischen Wege, die die Expression von Bacillomycin D steuern. Es wurde gezeigt, dass globale Regulatoren, wie beispielsweise DegU, DegQ und ComA, die alternativen Sigmafaktoren σ^B und σ^H und ein neuartiges Rap-Protein die transkriptionale Aktivität des in dieser Arbeit identifizierten Hauptpromotors des *bmy*-Operons beeinflussen. Insbesondere wurde gezeigt, dass DegU seine Effekte nach direkter Bindung an zwei unterschiedliche A/T-reiche Regionen im *bmy*-Promotor ausübt. Die anderen Regulatoren wurden mit eher indirekten Effekten assoziiert, welche meist über DegU oder DegQ ausgeübt wurden. Letzteres Protein scheint die Aktivität von DegU auf unbekannte Weise zu optimieren.

Es wurde außerdem gezeigt, dass DegU abgesehen von der Aktivierung des Hauptpromoters des *bmy*-Operons eine zusätzliche, vermutlich post-transkriptionale Rolle bei der Bacillomycin D-Produktion spielt. Daher war die Präsenz von DegH essentiell für die Produktion von Bacillomycin D. Auf ähnliche Weise erwies sich YczE, ein Membranprotein unbekannter Funktion, das neben *sfp* (eine 4´-Phosphopantetheinyl-Transferase, die nichtribosomale Peptide post-translational modifiziert und sie aktiviert) kodiert liegt, als essentiell für die Bacillomycin D-Produktion, jedoch als entbehrlich für die Produktion der restlichen von *B. amyloliquefaciens* FZB42 produzierten Peptid-Antibiotika. Der Effekt wurde auf einem post-transkriptionalen Level ausgeübt (vor dem Peptid-Export) und war unabhängig von DegU.

Abschließend kann gesagt werden, dass diese Arbeit Informationen über die genetische Identität von *B. amyloliquefaciens* FZB42, seine Lebensweise und die Produktion sekundärer Metabolite durch das Bakterium liefert. Außerdem definiert sie das komplexe regulatorische Netzwerk, das die Expression des meistvorhandenen Lipopeptides des Organismus, Bacillomycin D, kontrolliert. Es ist die erste Untersuchung der Genexpression eines Mitglieds Gruppe der Antibiotika von Iturin.

Introduction

Bacillus amyloliquefaciens strain FZB42

Bacilli are aerobic, rod-shaped, Gram-positive bacteria, with low G/C content. They are widely distributed in soil, air and water, and form oval endospores, as a consequence to deprived environmental conditions. Representatives of this genus, comprising some 51 validly described species, are being used in a wide range of industrial processes, mainly due to their ability to produce extracellular enzymes, antibiotics and insecticides, and secrete them in high concentrations [1].

In particular, *Bacillus amyloliquefaciens* and its numerous natural isolated strains serve for the production of a-amylase, an enzyme necessary for liquefaction of starch prior to saccharification for the production of sugar syrups in food industry [2, 3]. The habitat of this species is the soil, especially the rhizosphere where it colonizes plant roots [4, 5]. The commercially available strain of *B. amyloliquefaciens* FZB24 is applied as bio-fertilizer, as it stimulates plant growth and suppresses plant pathogenic organisms. These abilities are also shared by strain FZB42 [6].

The soil is also the natural environment of *Bacillus subtilis*, the best characterized member of the genus. Strain 168 was the first Gram-positive bacterium to be sequenced and has been used as a model organism to study the behavior of microorganisms for more than a century [7]. It is closely related to *Bacillus amyloliquefaciens* strain FZB42, but does not promote plant growth. The features and mechanisms governing the biocontrol-related function of *B. amyloliquefaciens* FZB42, which are obviously not active in the domesticated strain of *B. subtilis* 168, have not been fully characterized yet.

Genome sequencing

The genome contains the complete set of genetic information that organisms require to live and thrive. Therefore, the complete sequencing of the genomic DNA of an organism offers better understanding in respect with the mechanisms the organism adopts to withstand its environment. The function of several sequenced genes can be predicted as a basis of the genetic organisation of the gene's surrounding region, the conserved regions within the gene and the degree of its alignment with other genes of established function. The existence of

databases that compile information about sequenced organisms/genes in combination with powerful bioinformatics tools that perform multi-level gene comparisons / annotations in a very short time (seconds), makes the task of assigning a gene's putative function easier, faster and more successful. The information obtained by comparisons with such databases can serve as a basis for further molecular and biochemical work. Finally comparisons of complete genome sequences are very informative from an evolutionary aspect, as they allow better phylogenetic taxonomy of sequenced organisms and highlight the genetic reorganisation that evolution has imposed within closely related organisms.

The first complete genome sequence of a microbial organism was that of *Haemophilus influenzae* Rd KW20, published in 1995 [8]. *Bacillus subtilis* was the first Gram-positive bacterium to be sequenced as its biochemistry, physiology and genetics had been already thoroughly studied for many years [7]. In the following years, various other *Bacilli* were sequenced [9, 10, 11]. It is interesting to mention that the fully sequenced microbial genomes has rapidly risen from only 30 in year 2000 to more than 300 today. Shortly after the completion of sequencing of the first prokaryotic organism, the complete sequence of major eukaryotic organisms (such as drosophila, mouse and human) was accomplished and this marked a major breakthrough in science of the last century [12, 13, 14]. It is therefore apparent that information concerning complete sequenced genomes accumulates exponentially and combined with the development of more powerful databases (that reflect the advances in bioinformatics) provides better understanding / prediction of the abilities and functions of newly sequenced organisms.

Antibiotic production from Bacilli

Antibiotics are a diverse group of chemical substances produced by both prokaryotes and eukaryotes and are of great importance in medicine due to their ability to disrupt pathogenic microbial metabolism, by various mechanisms. They can be classified according to their structure or their action (Fig. 1). In the medical field, the two most important groups of antibiotics are the β-lactam and tetracyclines. Members of the first group, such as penicillins and cephalosporins, are produced by fungi and are potent inhibitors of cell wall synthesis of bacteria. The tetracyclines consist of a naphthacene ring system that can be substituted at several positions to form new analogs. They are produced by prokaryotes and inhibit almost all Gram-positive and negative bacteria, by interfering with 30S ribosomal subunit function.

In addition, aminoglycoside antibiotics, a separate group of antibiotics, exhibit the same way of function. They contain amino sugars bound to each other by glycosidic linkage, as in the case of streptomycin and kanamycin. Furthermore, macrolide antibiotics are widespread antibiotics in medicine that contain large lactone rings connected to sugar moieties. Erythromycin belongs to this group and inhibits protein synthesis at the 50S subunit of the ribosome (Fig. 1) [15].

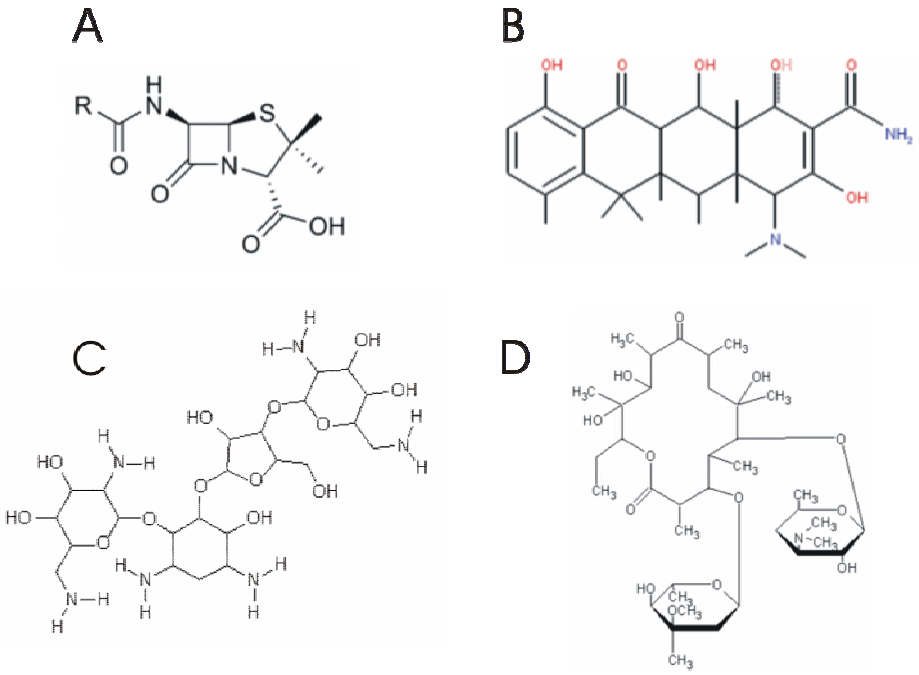

Figure 1: Chemical structural representation of different classes of antibiotics with major importance.

A. Penicillin core, B. Tetracycline core, C. Neomycin; aminoglycoside antibiotic and D. Erythromycin; a macrolide antibiotic

Bacilli are widely known and used microorganisms for production of a wide range of antibiotics, such as polymyxins (*B. polymyxa*) which destroy membrane integrity as well as edeines (*B. brevis*) which inhibit the formation of the initiation complex on the 30S ribosomal subunit [16]. The predominant class though of antibiotics produced by *Bacilli* are peptide antibiotics. These exhibit highly rigid hydrophobic and/or cyclic structures with unusual constituents like D-amino acids, and are generally resistant to hydrolysis by peptidases and proteases [17]. Furthermore, they are insensitive to oxidation, because cysteine residues are either oxidized to disulphides and/or modified to characteristic C-S (thioether) linkages. The peptide synthesis is achieved ribosomally, followed by post-translational modifications, or nonribosomally by multienzyme complexes [18].

Ribosomally synthesized peptide antibiotics

Synthesis

For the production of proteins and peptides three basic enzymes are required: tRNA synthetases, tRNAs and the ribosome [19]. First the aa-tRNA synthetase selects the cognate amino acid and loads it onto the 2'- or 3'- hydroxyl group of the corresponding t-RNA [20, 21]. Subsequently, with help of the elongation factor (EF-Tu), the ribosome selects the correct aa-tRNA during each cycle of polypeptide elongation, according to the mRNA sequence [22]. Therefore a complex comprising aa-tRNA, EF-Tu and GTP enters the acceptor site of the ribosome. The large ribosomal subunit stimulates GTP hydrolysis when there is complementary base-pairing between the mRNA and the cognate aa-tRNA. Peptide bond is formed when the aa-tRNA has been accommodated to the acceptor site, whereupon translocation can occur regenerating the ribosome. Eventually, post-translational modification events lead to the completion of synthesis of these peptide antibiotics [23].

Ribosomally synthesized peptide antibiotics in *Bacilli*; classification and control of gene regulation

Lantibiotics are the major group of ribosomally synthesized antibiotics in *Bacilli*. They contain lanthionine, which is formed post-translationally through dehydration of serine or threonine residues followed by addition of neighbouring cysteine thiol groups, leading to inter-residual thioether bonds [24, 25]. Based on structural properties, two types of lantibiotics are distinguished: type A, with a more linear secondary structure, and type B, with a more globular one [26].

Subtilin and ericin are members of the type A group and are lethal against Gram-positive bacteria by forming voltage-dependent pores into the cytoplasmic membrane [27]. Mersacidin belongs to the type B group and inhibits cell wall biosynthesis by complexing lipid II [28]. Other unusual lantibiotics produced by *Bacilli* are sublacin and subtilosin, which also act against Gram-positive bacteria through yet unknown mechanisms. The organization of these gene clusters is shown in figure 2.

Subtilin biosynthesis is mediated by the prepeptide SpaS [29], which is post-translationally modified by SpaBC [30]. Furthermore, the translocator SpaT exports the lantibiotic. Immunity to the producer strain is conferred by the lipoprotein SpaI and the ABC transporter SpaFEG [31]. In a positive feedback loop, subtilin activates the two component regulatory system SpaRK (response regulator and sensor histidine kinase) and directly

stimulates expression of genes involved in biosynthesis and immunity [32, , 33]. SpaRK expression is also controlled by the sporulation transcription factor SigH [33].

Ericin has the same gene cluster organization as subtilin, but surprisingly two structural genes, *eriA* and *eriS*. However, the production of ericin A and ericin S, that differ in amino acid composition and ring structure, is under the control of the same synthetase (EriBC) [34].

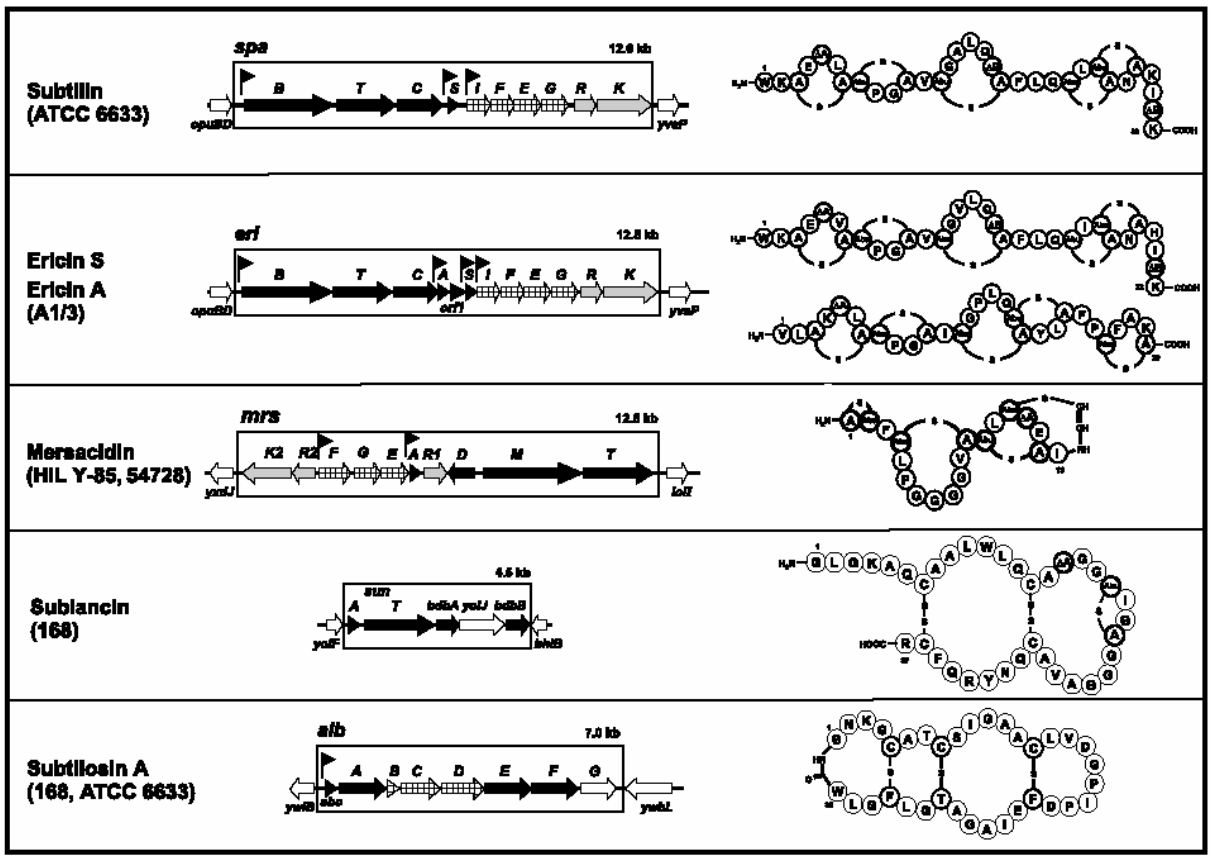

Figure 2: *Bacillus subtilis* lantibiotics, lantibiotic-like peptides and specifying gene clusters.

The organisation of gene clusters (boxed) specifying lantibiotic and lantibiotic-like peptides are presented along with schematic structure representations of the mature peptides. The size of gene clusters is given in kilobases (kb). Black boxes indicate structural genes and genes involved in post-translational modification and transport; grey boxes indicate regulatory genes; filled boxes stand for immunity genes. The figure is reproduced from [35].

mrsA is the structural gene in the mersacidin gene cluster, whereas the genes *mrsM* and *mrsD* are involved in its post-translational modification [36]. Furthermore, *mrsT*, coding for a transporter with an associated protease domain, mediates the transport while the operon *mrsFGE*, an ABC transporter, confers self-protection against the lantibiotic. m*rsR1* is a

response regulator that controls biosynthesis of mersacidin whereas the putative two component regulatory system *mrsR2K2* controls immunity [36, 37].

The structural gene for sublancin biosynthesis is *sunA* and it belongs to the *B. subtilis* temperate bacteriophage SPβ. An ABC transporter (SunT) and two thiol-disulphide oxidoreductases (BdbAB) belong to the same locus [38]. Until now, only BdbB is proven to be involved in the sublancin production, most probably for the formation of the disulphide bonds [39]. The genes conferring immunity are unidentified.

Finally, the gene cluster of subtilosin (*sbo-alb*) encodes AlbA protein, probably involved in post-translational modification of presubtilosin, and AlbBCD proteins, a putative ATP-binding transporter, involved in immunity [40]. The expression of *alb* genes is under the negative control of AbrB [41].

Nonribosomally synthesized peptide antibiotics

Structural diversity is a predominant feature of nonribosomally synthesized peptides, as they are assembled from an exceedingly heterogeneous group of precursors. There are more than 300 members in this group including pseudo, nonproteinogenic, hydroxy, *N*-methylated and D-amino acids [42]. In contrast, ribosomal synthesis of peptides is restricted to 20 amino acids.

Synthesis

In spite of their structural heterogeneity, the peptide antibiotics of this group share a common mode of synthesis, the **multicarrier thiotemplate mechanism** [43]. According to this model, peptide bond formation takes place on multienzymes designated nonribosomal peptide synthetases (NRPS), which are arranged in **modules**. Modules are the units responsible for the incorporation and/or modification of a specific amino acid into the peptide product, and their arrangement and number are usually colinear to the amino acid sequence and the length of the peptide respectively (colinearity rule) [44, 45, 46]. Modules are further divided into **domains**; the enzymatic units involved in a specific step of synthesis, such as substrate activation, covalent binding, elongation *etc* [47].

According to the multicarrier thiotemplate mechanism, the carboxy group of amino acid is activated to the corresponding adenylate by ATP hydrolysis and then it is transferred onto the free thiol-group of an enzyme bound 4'-phosphopantetheinyl cofactor (4'-PP), forming a thioester. At this stage, the substrates can undergo modifications such as epimerization or *N*-

methylation. Peptide assembly is achieved via peptide bond formation steps, by binding of the thioester-activated carboxyl group of the upstream module to the free amino group of the adjacent downstream module. During this N to C stepwise elongation, the intermediates are covalently attached to the multienzyme complex. The termination of the synthesis is induced by the release of the thioester-bound peptide product by hydrolysis, cyclization or transfer to a functional group [19, 44, 48]. As an example figure 3 shows a prototype NRPS assembly line for the cyclic lipoheptapeptide surfactin [49].

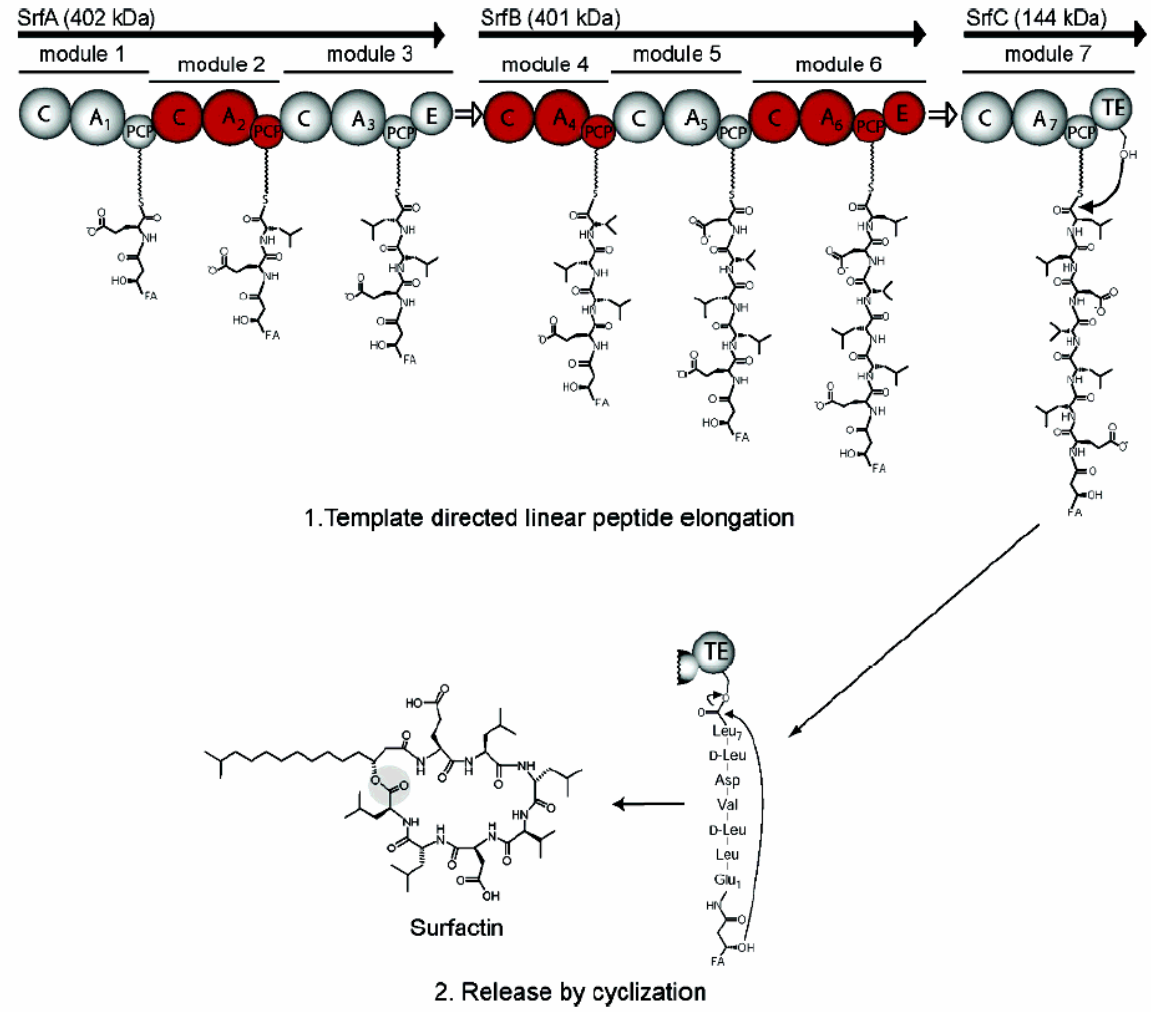

Figure 3: Surfactin assembly line.

The multienzyme complex consists of seven modules which are responsible for the incorporation of seven amino acids. 24 domains catalyse the same number of chemical reactions. The peptide chain is elongated stepwise from N to C end. The last domain is responsible for release and cyclization surfactin. The figure is reproduced from [48].

Domains of nonribosomal peptide synthetases

Domains are not just imaginary sections in the module. They are enzymatically active, as well as structurally and catalytically independent. They can be excised from the peptide chain and still retain their activity [45].

Adenylation domain

Nonribosomal peptide synthesis is initiated by the recognition and activation of the designated substrate. This is the role of the adenylation domain (A), which recognizes and incorporates the suitable amino acid substrate into the peptide. At the expense of Mg^{+2}-ATP and release of PPi, the amino acid is activated as aminoacyl adenylate (Fig. 4A) [50, 51, 52]. There is a specific adenylation domain for each amino acid included in the peptide antibiotic and its location indicates the primary structure of the product [45]. Sequence comparison of the A-domains (ca 550aa) deriving from various genes that code for peptide synthetases revealed 10 residues as the major determinants of substrate specifity; this result was also confirmed by introducing specific point mutations at these sites [45, 53, 54].

Thiolation domain (peptidyl carrier protein domain)

The thiolation domain (T), also known as peptidyl carrier domain (PCP), accepts the activated amino acid. The prerequisite for the functionality of the T-domain is its post-translational modification with the 4'-phosphopantetheine cofactor (4'-PP). Associated 4'-phosphopantetheinyl transferases catalyze the transfer of the 4'-PP moiety from the coenzyme A to a conserved serine residue of the T-domain, converting thus the inactive *apo*-T to its active *holo*-T (see also Fig. 6) [43, 55, 56]. Furthermore, the aminoacyl adenylate from the A-domain forms a thioester with the cysteamine thiol group of 4'-PP cofactor and therefore can be transported to the next module (Fig. 4B) [43, 47, 55, 57, 58]. The thiolation domain has around 100 amino acid residues and is located downstream of the adenylation domain [45]. It represents the transport unit that enables the elongation intermediates to move between the catalytic centers. The combination of adenylation and thiolation domains is referred to as **initiation module**, since both domains are required to activate and covalently tether the first building block in the peptide synthesis.

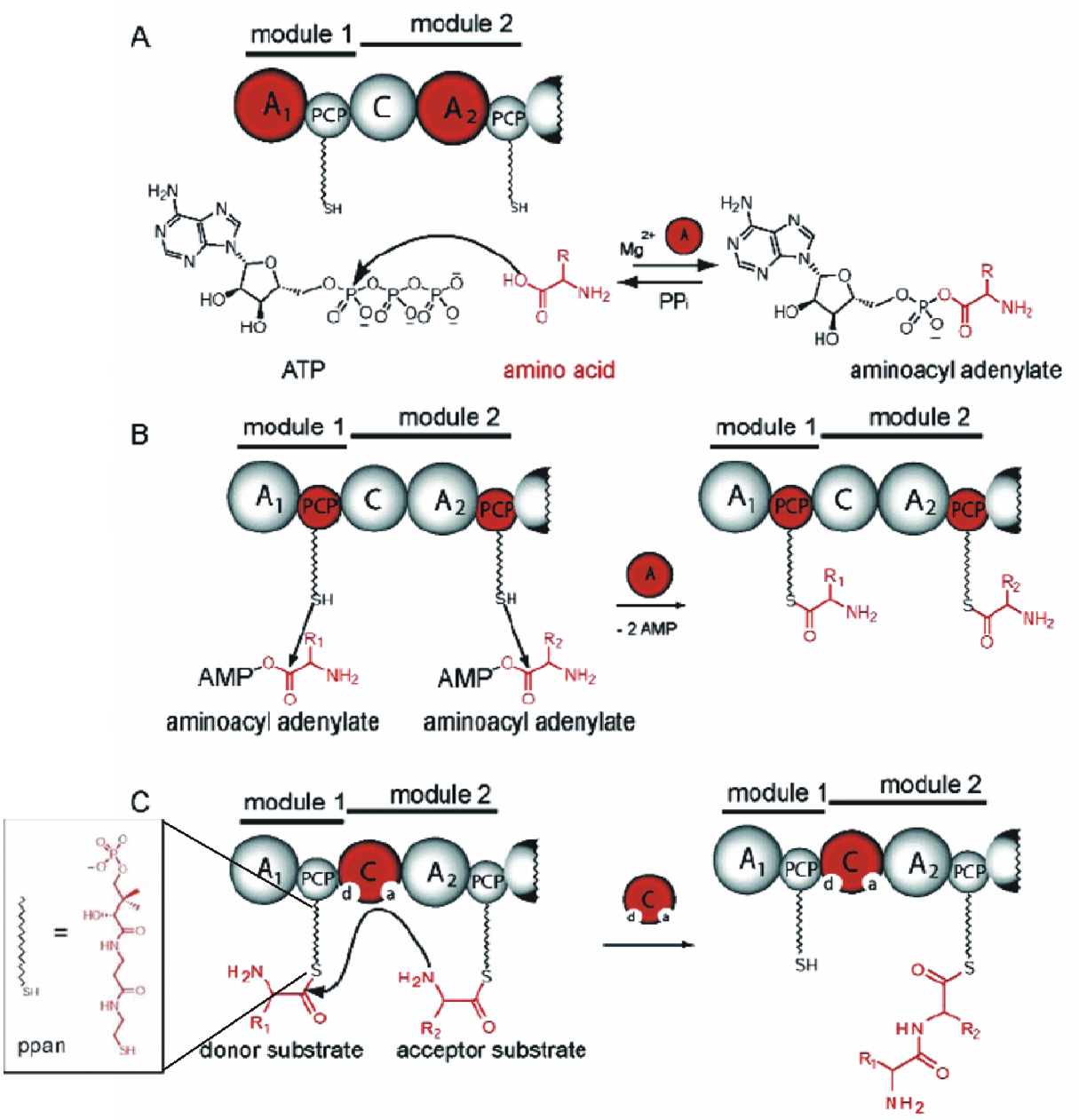

Figure 4: Domain catalyzed reactions.

A) The adenylation domain recognizes and activates the suitable amino acid as aminoacyl adenylate at the expense of ATP. B) Covalent attachment of the activated aminoacyl adenylate onto the free thiol group of the 4'-phosphopantetheine cofactor bound to the peptidyl carrier domain. C) Peptide elongation by the condensation domain, which catalyses the attack of the nucleophilic amine of the acceptor substrate onto the electrophilic thioester of the donor substrate. A_1- A_2, adenylation domains; PCP, thiolation / peptidyl carrier domain; C, condensation domain; d and a, donor and acceptor sites on condensation domain; ppan, 4'-phosphopantetheine cofactor. Domains in action are indicated in red. The figure is reproduced from [48].

Condensation domain

The condensation domain (C), ca. 450 amino acid length, is responsible for the formation of the peptide bond between two activated amino acids on adjacent modules and therefore controls the elongation of the growing peptide chain [19]. It catalyses the attack of the nucleophile aminoacyl-S-4'-PP-T to the electrophile aminoacyl/peptidyl-S-4'-PP-Ts, that lye downstream and upstream of the C-domain respectively (Fig. 4C) [59]. For this scope the C-domain harbors two selective substrate-binding sites: an enantioselective electrophilic donor site and an amino acid selective nucleophilic acceptor site [60]. The amino acid acceptor site is responsible for preventing internal mis-initiation as well as for controlling the timing of substrate epimerization [61], whereas the donor site for incorporating the correct isomer [60, 62].

The C-domain is found between two consecutive initiation modules located on the same synthetase (intramolecular amino acid transfer). In case the initiation modules belong to different synthetases, the C-domain is located at the N-terminus of the one accepting the substrates (intermolecular amino acid transfer). Peptide synthetases involved in lipopeptide biosynthesis contain an additional C-domain preceding the first initiation module, probably involved in the coupling of the fatty acid moiety to the first amino acid of the peptide moiety [63].

Variations on the peptide backbone can be obtained by the replacement of C-domains with the structurally and mechanically related heterocyclization (Cy) domains. Five-membered heterocyclic rings such as oxazoline in vibriobactin or thiazoline in bacitracin are common features of nonribosomal peptides and significant for chelating metals and interaction with proteins, RNA, DNA [48, 64]. The formation of such heterocyclic rings and the subsequent peptide elongation is catalyzed by Cy-domains with the nucleophilic attack of a T-bound cysteine, threonine or serine acceptor substrate onto the thioester of the donor substrate. As observed for C-domains, the free α-amino group of the cysteine, threonine or serine is the nucleophile. Subsequently, the side chain hydroxyl or thiol group carries out a nucleophilic attack onto the α-carbonyl C atom of the donor amino acid, producing a heterocyclic ring. Finally, the product is dehydrated to form oxazoline or thiazoline (Fig. 5A) [65, 66].

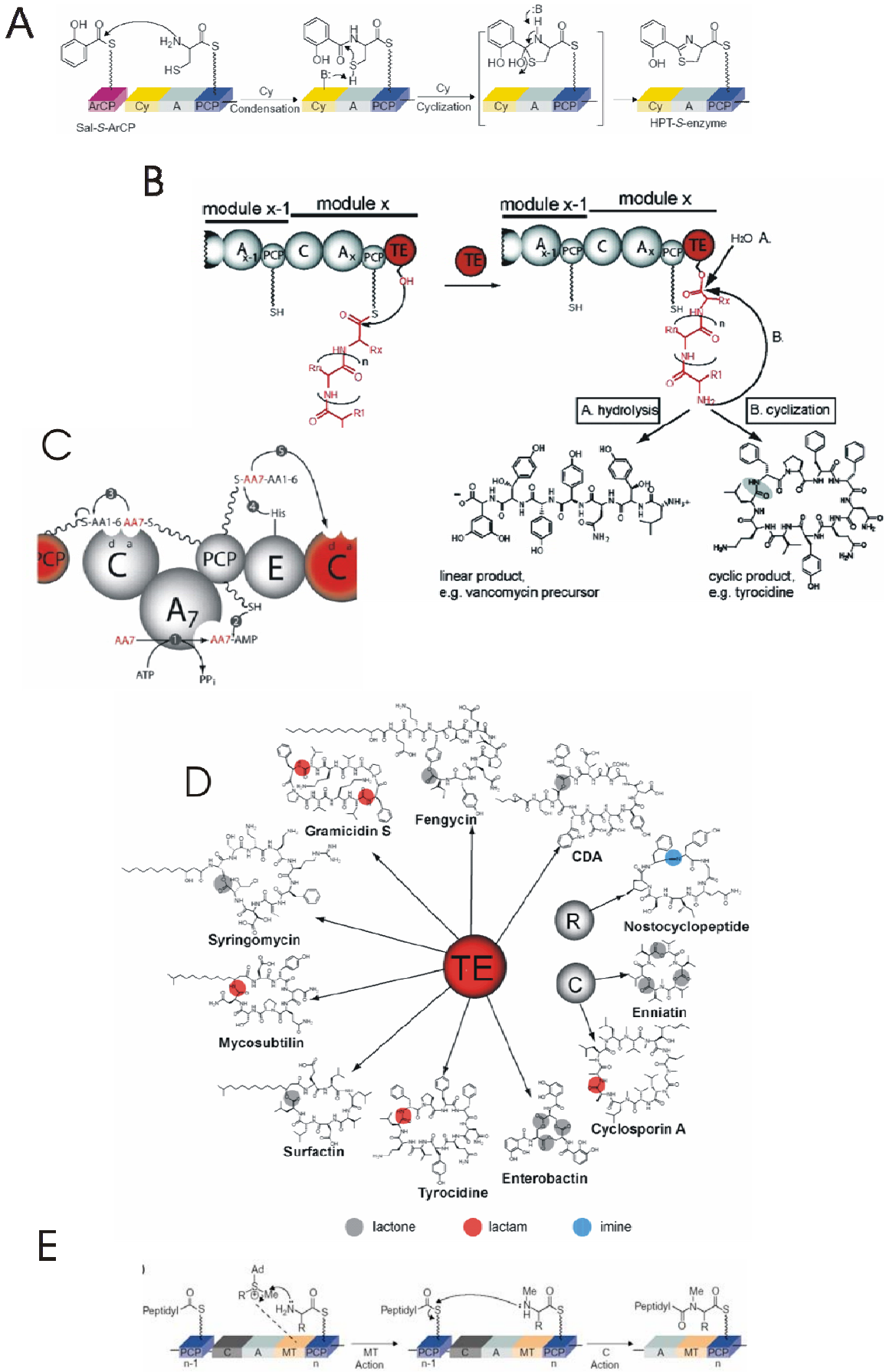

Figure 5: Schematic representation of the catalytic functions of Cy-, TE-, E- and N-MT-domains.

A) Formation of thiazoline heterocyclic rings from cysteine precursors catalyzed by Cy-domain. Three reactions are catalyzed by the Cy-domains: amide bond formation, cyclization and dehydration. Cy, heterocylization domain. Example from yersiniabactin nonribosomal synthetase present in *Yersinia pestis*. ArCP, aryl carrier protein; Sal-S-ArCP, activated salicyl group onto the N-terminal ArCP. The figure is reproduced from [65].

B) Peptide release by the TE-domain. Peptide release is achieved either by external nucleophile water resulting in a linear product (A) or by an internal nucleophile resulting in a cyclic product (B), depending on the NRPS template. The figure is reproduced from [48].

C) Peptide synthesis order in the presence of an E-domain within the elongation modules. 1. Substrate adenylation by A-domain. 2. Transfer of the activated amino acid to the PCP domain. 3. Binding on the upstream C-domain acceptor site and formation of peptide bond. 4. The resulting peptidyl-PCP has lower affinity for the acceptor site and is transferred to the subsequent E-domain. Equilibrium of D/L isomers is produced. 5. Binding of the D-isomer on the donor site of the downstream C-domain. AA, amino acid; AAx$_n$, upstream peptidyl chain; E-epimerization domain. The figure is reproduced from [48]

D) Cyclization strategies. The majority of cyclization reactions within NRPS are catalyzed by TE-domains. A putative C-domain accounts for cyclization of cyclosporine A while a T-C domain controls oligomerization of the trilactone enniatin. A reductase domain (R) is responsible for cyclization of the imine nostocyclopeptide. The figure is reproduced from [48]

E) N-Methylation of nonribosomal peptides by embedded N-MT domains. N-methylation occurs on the aminoacyl thioester monomer prior to amide bond formation with the upstream peptidyl chain. Example from yersiniabactin nonribosomal synthestase in *Yersinis pestis*. The figure is reproduced from [65].

Thioesterase domain

The thioesterase domain (TE), ca 250 amino acid length, is responsible for the release of the peptide from the multienzyme complex. During synthesis, the growing peptide chain is transported between the T-domains of the subsequent modules from the N to the C-terminus of the synthetase until it reaches the final module. This module usually contains the TE-domain, causing product liberation by a two-step process. This involves an acyl-*O*-TE-enzyme intermediate that is attacked by either a peptide-internal nucleophile [67, 68] or water [69], and results in a macrocyclic [70] or a linear product [71] (Fig. 5B).

TE-domains are very diverse since they catalyze various reactions (Fig. 5D) [72]. In the case of tyrocidine (*B. brevis*), head to tail cyclization is achieved by amide bond formation between the N-terminal amine and the C terminus of the peptide, yielding a lactam product

[52], whereas for surfactin and mycosubtilin (*B. subtilis et al*.) lipo branched chain cyclization is accomplished by connection of a *β*-hydroxy and a *β*-amino fatty acid to the C-terminus, yielding a lactone and a lactam respectively [63, 70]. The same situation is observed for the calcium dependent antibiotic (CDA) produced by *Streptomyces coelicolor* A3(2) [73]. For fengycin (*B. subtilis et al*.) and syringomycin (*P. syringae*), amino acid branched chain cyclization occurs by using a tyrosine and a serine from the peptide chain as nucleophiles, discriminating them from other peptide antibiotics that use the *β*-hydroxyl group of the attached fatty acid moiety [74, 75, 76]. Some TE-domains do not permit cyclization of one peptide chain, but force the multienzyme to repeat the synthesis once or twice more. Subsequently, they have the ability to count the assembled synthetase monomers at the end and initiate release by cyclic dimer or trimer formation, when the desired length is achieved [48]. This mechanism, though not yet fully characterized, applies for the synthesis of gramicidin S (*B. brevis*) [67], enterobactin (*E. coli*) [77] and bacillibactin (*B. subtilis)* [78]. As they control such different mechanisms of cyclization, TE-domains show high degree of specialization and therefore share low sequence homology (10%-15%) [72].

Nevertheless, cyclization is not accomplished exclusively by TE-domains. For cyclosporin A (*Tolypocladium niveum*), a putative C-domain is responsible for the final peptide bond [79], whereas for enniatin (*Fusarium script*), a T-C didomain accounts for the oligomerization [80]. In the case of nostocyclopeptide (*Nostoc sp*.), the C-terminal residue of the linear peptide is reduced by a reductase domain (R-domain) to give an aldehyde, that is intramolecularly captured by the α-amino group of the N-terminal amino acid residue to produce a cyclic imine [81].

Epimerization domain

The epimerization domain (E) controls the conversion of amino acids, that belong to the attached growing peptide chain, from L to D-configuration. Usually these domains (ca 450 amino acid length) are located internally in the synthetases upstream of the condensation domain [82]. They represent a class of cofactor independent amino acid epimerases that catalyze the de- and reprotonation of the α-carbon atom of an enzyme bound aminoacyl or peptidyl-S-4'-PP thioester in both directions (L-to-D, D-to-L), resulting in a mixture of both isomers. However, the L-isomer is rejected by the enantioselective donor site of the following C-domain, whereas the D-isomer is used from the same domain for the elongation of the peptide chain [83, 84].

If the E-domain is part of the initiation module, an equilibration between the two isomers takes place as the amino acid is bound as thioester at the thiolation domain, prior to

peptide bond formation [83, 85, 86]. The downstream C-domain is selective for the D-isomer, which is eventually incorporated [60, 62]. However, if an E-domain is embedded in the elongation modules, epimerization occurs at the peptidyl-4'-PP-T stage. The corresponding A-domain recognizes and activates the L-isomer, which is then transferred onto the following T-domain and then onto the upstream C-domain for peptide bond formation. Then, E-domain acts to produce a D/L equilibrium of peptidyl-S-4'-PP thioesters. Furthermore, downstream C-domain catalyses only the transfer of the D-isomer to the next elongation module (Fig. 5C) [61, 84, 87].

Quite rarely, D-amino amino acids are present in the peptides independently from the catalytic function of E-domains. These substrates are first epimerized by racemases, which are not intergrated in the peptide synthetase, and then recognized and incorporated by the corresponding A-domain. This is the case for D-Ala_1 in the cyclosporine synthetase [88].

N- and C-Methyltransferase domains

The *N*-Methyltransferase (N-MT) and *C*-Methyltransferase (C-MT) domains are responsible for the *N*-or *C*-methylation of amino acid residues, thus making the peptide less susceptible to proteolytic breakdown. N-MT, which is usually located between the corresponding A- and T-domains, catalyzes the transfer of *S*-methyl group from *S*-adenosyl methionine (SAM) to the α-amino group of the thiosterified amino acid (Fig. 5E) [89]. This reaction is accomplished prior to peptide bond formation, as determined for the enniatin synthetase [90]. C-MT domains appear rarely in nonribosomal peptide synthetases, but use also SAM as the methyl donor [65].

Posttranslational modification

Nonribosomal peptide synthetases require posttranslational modification to be functionally active. As it has been already mentioned, thiolation domains are unable to serve as transport proteins immediately after translation, resulting in blocking of peptide synthesis. A modification by transfer of the 4'-PP moiety of coenzyme A onto a conserved serine residue of each T-domain, converts the latter from *apo*- to *holo*-form and unblocks the synthesis. The mobile 4'-PP prosthetic group is about 20Å in length and since it is covalently bound as a phosphothioester to the multienzyme [91], it serves as a “flexible arm”, which initially accepts the activated substrates and later on delivers them to the next building-block [43, 55, 57]. The conversion of T-domain is catalyzed by a dedicated 4'-phosphopantetheinyl

transferase (4'-PPTase) in a Mg^{+2}-dependent way, thereby releasing 3', 5'-ADP (Fig. 6) [92, 93]. Sfp and Gsp proteins control this reaction in *B. subtilis* and *B. brevis*, respectively [56, 92, 93].

Recent studies have shown that Sfp accepts as substrates CoA derivatives, such as acetyl-CoA and aminoacyl-CoA [60, 94]. It is therefore likely that PPTases also modify the T-domains of NRPSs with acyl-4'-PP, rendering the enzyme inactive, as misprimed transport units are unable to accept activated amino acids. The activity can be restored by thioesterases II (TE-II) which hydrolyze the acyl-4'-PP, leaving only the 4'-PP bound, and are found in association with the peptide synthetases [95]. TE-IIs contribute as proofreading enzymes, since they preferentially hydrolyze acetyl-Ts versus aminoacyl or peptidyl-Ts [96]. Consequently, the capable of nonribosomal peptide synthesis *holo*-Ts are made either by direct priming of the *apo*-derivatives, catalyzed by PPTases ,or by deblocking misprimed derivatives, catalyzed by TE-IIs.

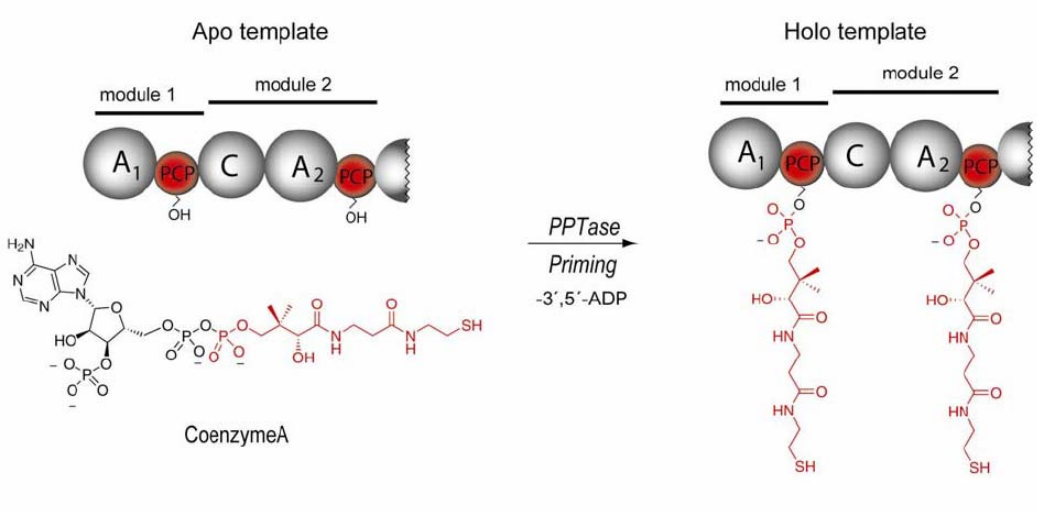

Figure 6: Conversion of thiolation domain from *apo*- to *holo*-form.

The 4'-phosphopantetheine moiety of coenzyme A is covalently attached onto an invariant serine residue of the thiolation domain (PCP) by dedicated phosphopantetheinyl transferases; thus PCP-domains are activated. The figure is reproduced from [48].

Hybrid synthetases

In recent years increasingly more peptide synthetases have been identified that contain domains normally present in fatty acid (FASs) or polyketide (PKSs) synthases. The first determined mixed NRPS-PKS biosynthetic gene cluster was that of rapamycin in *Streptomyces hydroscopius*, that contains a NRPS module for the incorporation of pipecolic

acid into the polyketide [97, 98]. In addition, synthesis of melithiazole and myxothiazole requires six multifunctional enzymes that switch back and forth between NRPS and PKS [99, 100]. Furthermore, hybrid systems of peptide synthetase and fatty acid synthase, such as mycosubtilin and iturin were characterized in various *Bacillus* strains [63, 101]. Most recently, a genomic island (54kb) that consists of three nonribosomal peptide synthetases, three polyketide synthases and two hybrid NRPS/PKS synthases was identified among pathogenic *E. coli* strains of the B2 group. Interestingly, it was shown that *E. coli* strains expressing this gene cluster induce double-strand breaks in eukaryotic cells leading to cell death [102].

Fatty acid synthases (FASs)

Fatty acids are essential for primary and secondary metabolism, because they are used as a form of energy storage, but also as building blocks for cell membranes or for nonribosomally synthesized peptides. The fatty acid synthase (FAS) of bacteria is a multienzyme complex that consists of individual, highly conserved enzymes [103, 104].

The first step in fatty acid production is the synthesis of malonyl-CoA from acetyl-CoA and CO_2, which involves the biotin carboxyl carrier protein and is catalyzed by biotin carboxylase [105, 106]. The manolyl units are subsequently transferred to the 4'-PP of the *holo*-acyl carrier protein (ACP) by action of malonyl-CoA:ACP transacylase [107]. The acylated β-ketoacyl-ACP synthase III is then in the position to initiate chain elongation via condensation with malonyl-ACP and release of CO_2, resulting in an ACP-bound acyl chain that is extended by C_2 [108]. The β-carbon of the intermediate tethered to the ACP is reduced by a ketoacyl-ACP reductase (KR) and then dehydrated by a β-hydroxyacyl-ACP dehydratase (DH) (Fig. 7A). Finally, the enoyl-ACP reductase (ER) catalyzes reduction of the β-carbon to CH_2. This elongated acyl-ACP can participate in subsequent rounds of synthesis that involve additional keto synthases (KSs) with different substrate selectivities [19, 100].

Polyketide synthases (PKSs)

Polyketides are secondary metabolites which are synthesized on modularly organized giant multienzymes (polyketide synthases, PKSs) by decarboxylative Claisen condensations. In general, their biosynthetic pathway shares similarities to nonribosomally synthesized peptides and requires at least three domains [19].

The acyltransferase (AT) domain is responsible for the selection of substrate, which can be malonyl-, methyl-, ethyl- or propylmalonyl-CoA [109]. This appears to be a significant difference to FASs whose substrate selectivity is limited only to malonyl-CoA. Further on, the AT-domain transfers the chosen substrate to the 4'-PP of the corresponding *holo*-ACP, which is analogous to the transport protein of FASs. Like in NRPSs, ACPs are posttranslationally modified by 4'-phosphopantetheinyl transferases [110]. Relocation of the malonyl-derivative occurs to an active cysteine residue of the KS-domain. The substrate of the next module binds to the ACP-domain and is decarboxylated, resulting in the free nucleophile necessary for the subsequent Claisen-condensation with the KS-bound ketide. Therefore, an enzyme-bound β-ketoacyl intermediate is generated. Moreover, the produced intermediates are always transferred on the synthase according to the indicated elongation steps and finally a TE-domain catalyzes the cleavage of the product by macrocyclization. Like in the case of NRPSs, the order of modules determines the sequence of polyketide synthesis (Fig. 7B) [19].

"Optional" domains, such as KR-, DH-, ER- domains, are also observed in PKSs, such as and they operate in a similar manner to those used by FASs [110, 111]. In general, even though fatty acid and polyketide synthases share striking architectural and organizational similarities with the peptide synthetases, they are more closely related to each other.

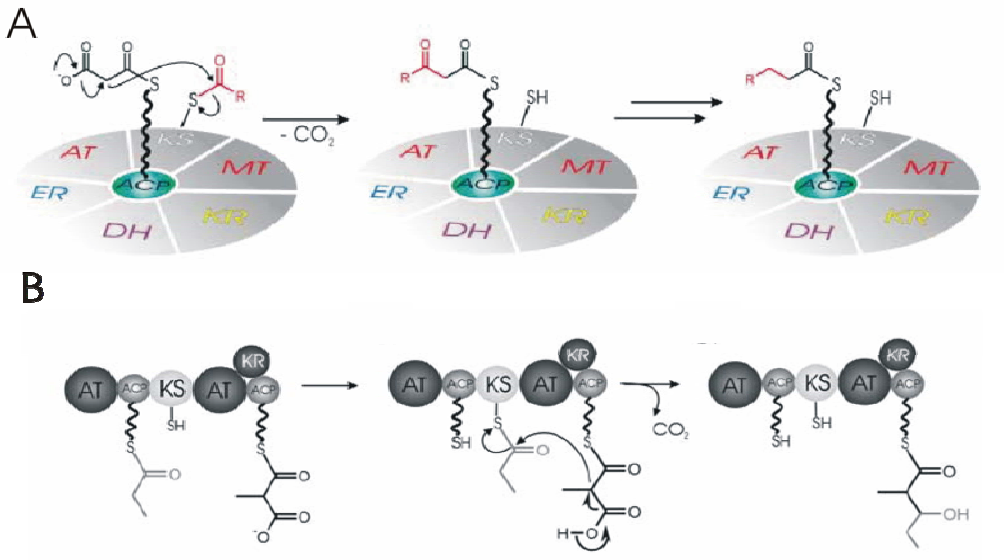

Figure 7: FASs and PKSs; multienzyme complexes with distinct domains.

A. Fatty acid synthases (FASs). A malonyl residue loaded onto the central ACP is condensed with an acyl chain bound to the KS. After condensation with release of CO_2, the β-keto group is first reduced by a KR, dehydrated by a DH and finally reduced to the methyl group by an ER. ACP, acyl carrier protein; KS, keto synthase; KR, ketoacyl-ACP reductase; DH, β-hydroxyacyl-ACP dehydratase; ER, enoyl-ACP reductase.B. A fictitious dimodular polyketide synthase (PKS). The ACP of the first module is loaded with propionyl by the AT domain of the first module, while the second AT domain

loads its ACP with methylmalonyl. The propionyl residue is translocated to an active-site cysteine of the KS-domain, whereas the methylmalonyl is decarboxylated resulting in the nucleophile for the condensation with the KS-bound propionyl. The product of condensation is covalently tethered to the 4'-PP present at the ACP of the second module. KR domain causes reduction of the β-carbonyl group to a hydroxyl one. ACP, acyl carrier protein; AT, acyl transferase; KS, keto synthase; KR, ketoacyl-ACP reductase. The figure is reproduced from [19].

Distribution-organization-function of peptide synthetase operons in *Bacilli*

Nonribosomally synthesized peptide antibiotics are widespread among *Bacilli*. Some of them are characteristically produced by only one member of the genus whereas others are more conserved. Nowadays more information concerning their diversity and distribution has accumulated, partly as a result of the increased number of sequenced genomes. Due to their conserved genetic structure and huge size, these synthetases can be easily recognized. Together with the polyketide synthases, they are the largest operons in the genome. In this section, an attempt will be made to summarize the current knowledge in respect with how the most well studied antibiotics of this group are organized and operate.

Different *Bacillus* strains produce small cyclic peptides with long fatty moiety, the so-called lipopeptides. Based on their structure, they can be generally classified into three different groups: i) the surfactin [112], ii) the fengycin [76, 113, 114] and iii) the iturin group [115].

Surfactin is a heptapetide linked via lactone bond to a β-hydroxy fatty acid composed of 13 to 15 carbon atoms (Fig. 8A) [116, 117]. Its operon comprises four open reading frames (ORFs) codifying the proteins SrfAA, SrfAB, SrfAC, SrfAD (Fig. 9A) [49, 118, 119, 120, 121]. SrfAC protein ends with a TE-domain, responsible for peptide release and cyclization, whereas the following protein SrfAD shows high homology to TE-IIs. Remarkably, disruption of this gene leads to severe reduction but not abolishment of the antibiotic's production [95, 96]. Furthermore, SrfAD acts in a double manner by hydrolyzing 4'-PP bound acetyl groups of misprimed NRPSs, according to the TEII ability [95] as well as by mediating the transfer of the fatty acid substrate to the Glu-module and stimulating β-hydroxyacyl-glutamate formation [122]. In general, the number of amino acids and their configuration agrees totally with the organization of modules and domains on the surfactin synthetase, confirming the colinearity rule mentioned earlier. An example is the presence of two D-

configurated amino acids that correspond exactly to the position of two epimerization domains.

Surfactin is one of the best characterized lipopeptides, since it possesses various beneficial abilities. Firstly, surfactin is able to lower surface and interfacial tension, thanks to its amphiphilic structure. In particular, surfactin produced by *B. subtilis* ATCC 21332 is considered one of the most powerful biosurfactans, since it can lower the surface tension of water from 72 to 28 mN/m at concentrations as low as 24μM [123, 124]. Furthermore, surfactin is responsible for inhibition of fibrin clot formation [124] and for erythrocytes lysis [125]. Other beneficial properties, with potential biotechnological and pharmaceutical applications are. i) antitumor activity [126], ii) activity against enveloped viruses [127], iii) antibiotic function against the protoplast of *B. megaterium* [128] and *Mycoplasma* [129, 130]. Furthermore, the *srf* operon encodes the regulatory gene, *comS* [131], which is involved in the development of genetic competence, an active process aimed at acquiring new genetic material that enables the cell to survive under changing environmental conditions [1]. Surfactin is also essential for swarming motility [132, 133, 134, 135], a flagellum-driven social form of surface locomotion, as well as for formation of biofilms, i.e. surface-associated multicellular communities [136, 137].

Fengycin, synonymous to plipastatin, is a cyclic decapeptide linked to a β-hydroxy fatty acid moiety, with lengths that vary from 14 to 18 carbon atoms (Fig. 8E, 9B) [138, 139, 140]. Fengycin demonstrates strong surface activity, although lower compared to surfactin [141]. Fengycin is active against filamentous fungi [76, 139, 140], and inhibits the enzymes phospholipase A2 [142] and aromatase [143].

Iturin, **mycosubtilin** and **bacillomycin** belong to the same group of lipopeptides. These compounds consist of seven α-amino acids and one β-amino fatty acid, that distinguishes them from the already mentioned groups. The peptide moiety contains a tyrosine in the D-configuration at the second amino acid position as well as two additional D-amino acids at positions three and six (Fig. 8B, 8C, 8D). Gene sequences encoding enzymes for biosynthesis of iturin A and mycosubtilin, but not bacillomycin D, have been reported (Fig. 9C) [63, 101]. Thereby it has been revealed that these lipopeptides are synthesized on hybrid synthases, since domains homologous to fatty acid and polyketide synthases are situated at their N-terminus [63]. These domains are absent from the peptide synthetases of surfactin and fengycin groups, so it appears very likely that these domains are involved in the incorporation of the β-amino fatty acid moiety into the peptides of the iturin group lipopeptides [63].

Moreover, these antibiotics exhibit strong antifungal and hemolytic activities, whereas their antibacterial function is more limited [76, 115].

Figure 8: Schematic structure of various lipopeptides produced by *Bacilli.*

A Surfactin, n = 10-12, B Iturin A, n = 10-13, C Mycosubtilin, n = 10-13, D Bacillomycin D, n = 10-13, E Fengycin, n = 13-17, F Lichenysin, n = 9 -14

The above mentioned lipopeptides are produced by different *Bacilli*, such as *B. subtilis* and *B. cereus*. However, one lipopeptide with similar structure to surfactin is exclusively composed by *B. licheniformis* [144, 145, 146]. It is designated as lichenysin and is a cyclic heptapeptide with a β-hydroxy fatty acid moiety, composed of 12-17 carbon atoms (Fig. 8F, 9D) [146]. It demonstrates antimicrobial properties and reduces the surface tension of water [144, 146]. In particular, lichenysin A can cause a similar reduction in water surface tension as surfactin from *B. subtilis* ATCC 21332, albeit in lower concentration (12µM versus 24µM) [145].

Another nonribosomally synthesized antibiotic compound is **bacitracin** found in *B. licheniformis* [147, 148]. This thiazoline ring-containing dodecapeptide is synthesized by the large multienzyme complex BacABC (Fig. 9E) [149]. Bacitracin is a prominent inhibitor of cell wall biosynthesis and most active against Gram-positive bacteria [147]. However, *B. licheniformis* and several other Gram-positive bacteria are not susceptible to this antibiotic suggesting the existence of specific resistance mechanisms [150]. Its primary mode of action is the formation of a tight ternary complex with the peptidoglycan carrier C_{55}-isoprenyl pyrophosphate (IPP) and a divalent metal cation. This carrier is responsible for the

translocation of cell envelope building blocks from the cytosol to the external side of the cytoplasmic membrane, where they are incorporated to the macromolecular network of the cell envelope (i.e. peptidoglycan, teichoic acids and polysaccharide capsule). Binding of bacitracin to IPP prevents its recycling by dephosphorylation to the monophosphate form that is normally reloaded on the inner face of the membrane [150, 151].

Another member of the *Bacillus* genus, *B. brevis*, produces two cyclic decapeptides, **tyrocidine** and **gramicidin S** (Fig. 5D, 9F) [52, 84, 152]. The first one characteristically contains a nonproteinogenetic residue, the L-ornithine and acts as antibiotic by membrane perturbation [17, 52]. Gramicidin S is synthesized on the enzymes GrsTAB, where only five amino acids are activated and incorporated. However, the peptide is dimerized to the decapeptide prior to its release. Furthermore, gramicidin S exhibits strong antibacterial activities against Gram positive and negative bacteria [153, 154], probably due to an interaction with membrane phospholipids. Thereby, gramicidin S causes a phase separation of negatively charged phospholipids from other lipids leading to a disturbance of the membrane's osmotic barrier [155, 156].

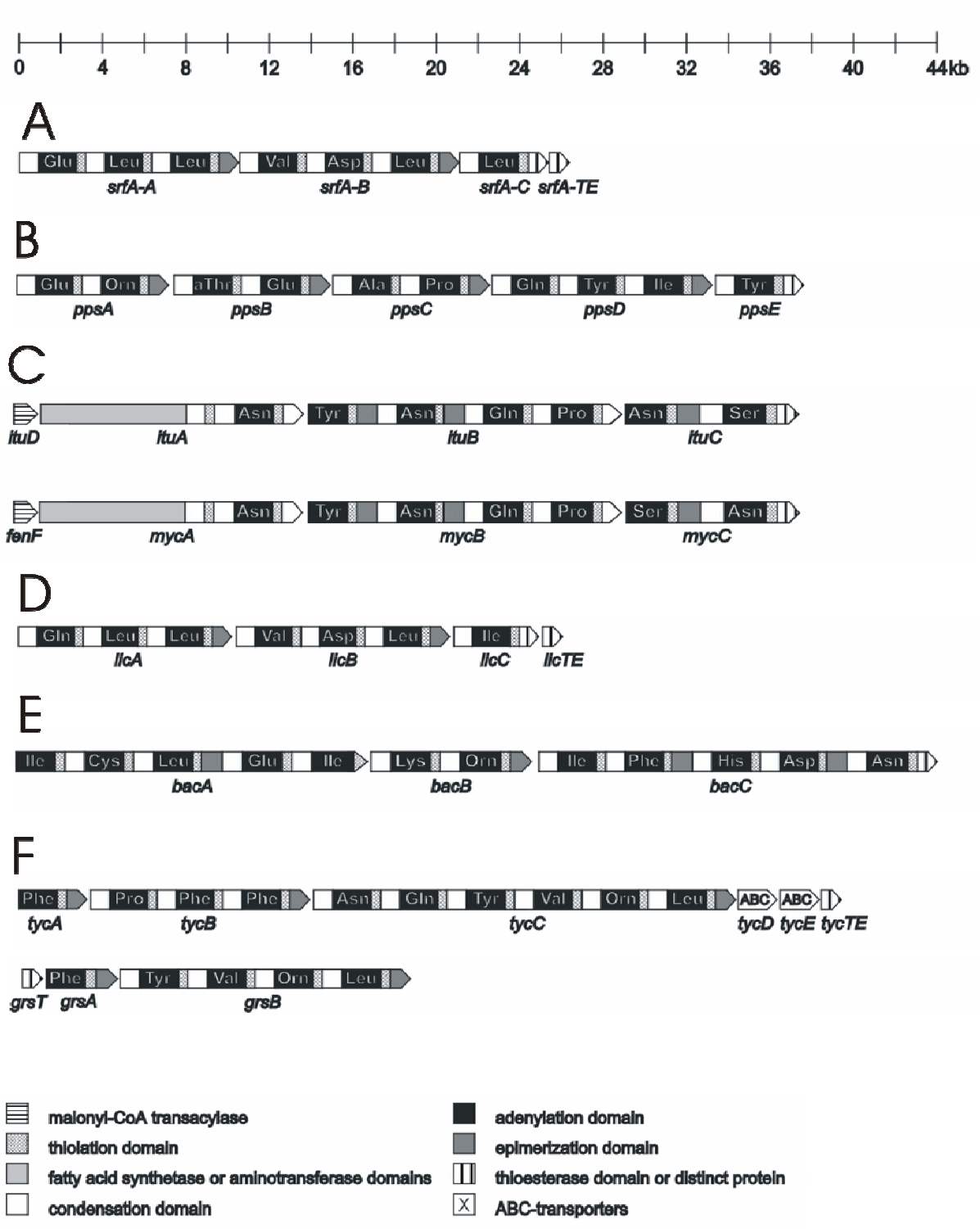

Figure 9: Schematic representation of peptide synthetase operons in *Bacilli*.

The genes comprising each peptide synthetase operon and their sizes are indicated. Organisation within the modules is presented, while the respective activated amino acid are depicted within the adenylation domains. A. Surfactin operon in *B. subtilis* [49]. B. Fengycin operon in *B. subtilis* [140]. C. Iturin A and mycosubtilin operons in *B. subtilis* [63, 101]. D. Lichenysin A operon in *B. licheniformis* [146]. E. bacitracin operon in *B. licheniformis* [149]. F. Tyrocidine and gramicidin S operons in *B. brevis* [52, 152]. The figure is adapted from [157].

Multiple control of expression of peptide synthetase operons in *Bacilli*. Export and immunity mechanisms.

In the last few decades, the pathways that govern the synthesis of antibiotics on large multienzymes have been thoroughly studied. Significant progress has been made on the functional analysis of various domains as well as on the role of their assembly in the peptide synthetases. Moreover, high resolution structures obtained for several enzymatic subunits from different antibiotics led to a better understanding of their architectural organization, substrate specificity and catalytic action [70, 158, 159, 160]. In contrast, our knowledge concerning how the organism regulates expression of these systems or the mechanisms which govern export of the peptides and/or resistance to them is rather limited. An exception is the case of surfactin, for which studying the regulation of gene expression received increased attention due its connection with the development of genetic competence.

The expression of **surfactin** is growth-phase dependent and is induced during transition to stationary phase [161]. Its transcription is driven by a σ^A-dependent promoter [161] and its expression is regulated via a complex network, including the two component regulatory system, ComAP [161, 162]. ComP is the sensor histidine kinase that is autophosphorylated after sensing increase in the concentration of the pheromone ComX [163]. The phosphoryl group is then transferred to the response regulator, ComA and activates it. Phosphorylated ComA can bind upstream of the *srf* operon and induce its expression. Therefore, systems involved in the phosphorylation / DNA-binding ability of ComA (ComXQ, RapC-CSF, RapF) modify indirectly the antibiotic's expression [163, 164, 165, 166, 167]. PerR, a general repressor of the peroxide stress regulon, is shown to positively regulate surfactin in a direct manner, independently of ComA [168]. In contrast CodY, a GTP-activated global regulator, acts as a direct repressor under casamino acids rich conditions [169]. Furthermore, YerP, a protein homologous to the RND (resistance, nodulation and cell division) family of efflux pumps in Gram-negative bacteria, seems to contribute in secretion of surfactin and self-resistance of the producer strain against it [170].

Knowledge on transcriptional regulation of the remaining lipopeptides is rather limited. The promoters of **fengycin** and **iturin** operons have been successfully identified and show similarity to a housekeeping σ^A promoter [101, 113]. Furthermore, deletion of *degQ*, a pleiotropic regulator gene that controls the production of several secreted and degradative enzymes [171], reduces severely the production of these antibiotics, via an unidentified mechanism [172, 173].

All these lipopeptides are post-translationally regulated by *sfp*, a 4'-phosphopantetheinyl transferase which converts T-domains to their active form (see corresponding chapter; [92, 137]. The importance of this gene is demonstrated in strains that contain intact synthetases but dysfunctional *sfp*. *B. subtilis* strain 168 contains intact *srf* and *fen* operons but is unable to produce the antibiotics, due to a frameshift mutation on the *sfp* gene [174]. However, when complemented with a functional 4'-PPTase, the antibiotic production of the strain is restored [172, 175].

Mechanisms that govern regulation of **lichenysin** and **bacitracin** are studied only in a preliminary basis. Lichenysin expression is dependent on the two component regulatory system ComAP [176]. In the case of bacitracin, an ABC transporter (BcrABC) conferring resistance to the producer strain against the antibiotic was determined [177, 178]. It is located about 3 kb downstream of the bacitracin biosynthetic operon *bacABC* and its expression is induced by the dodecylpeptide [150, 179]. Moreover, a two component regulatory system BacRS, situated between the *bac* operon and the *bcrABC* genes, negatively regulates expression of the transporter genes [150].

Transcription of the **tyrocidine** operon is driven by a typical σ^A promoter and its expression is induced at the end of exponential phase of growth. Spo0A, Spo0B and Spo0E, involved in the sporulation process, are required for full activation of the operon, whereas AbrB, a transition-phase regulator, acts as its repressor [180]. Further studies revealed that AbrB inhibits tyrocidine expression directly by binding to the upstream region of *tycA* [181]. Moreover, *tycD* and *tycE*, which are located downstream of the operon, show high similarity to members of the ABC transporter family and thus may confer immunity to the producer strain [52]. However, their role remains to be verified.

Approaches to new antibiotics

Years of research revealed that NRPS and NRPS-PKS hybrids can produce biologically active compounds exhibiting high antimicrobial activity. Their modular architecture allows the possibility to manipulate the enzymatic machinery in order to increase or alter their biological action. In the last decade, successful steps have been made in creating novel improved antibiotics by genetically redesigning natural synthesized compounds.

Genetic engineering has been achieved using different approaches. The first approach was based on exchanging the A-T units of the terminal module of surfactin synthetase that is originally responsible for the incorporation of leucine. Different A-T units have replaced the already existing one and novel surfactins with aliphatic (Val), charged (Orn) and aromatic

(Phe) residues at position 7 were created. However, their hemolytic activity did not differ significantly from that of the wild type product [182]. Nevertheless, swapping of numerous domains indicated for the first time that a rational design of antibiotics is accomplishable [183].

A further strategy for constructing synthetic antibiotics involves entire module swapping as well as insertions or deletions of modules. Conistent to this concept, deletion of the second module of the *srf* operon produced a new hexapeptide surfactin [184]. Alternatively, the manipulation of the A-domain's specificity via point mutagenesis can also result in novel antibiotics. The altered A-domain recognizes and activates a different amino acid, which is then incorporated in the polypeptide chain to yield a new product [185]. Another pathway to novel antibiotic production involves the replacement of TE-domains on the synthetase to force earlier release and cyclization [186]. It has been already shown that the bioactivity of many peptide antibiotics is attributed to small heterocyclic compounds, such as thiazoline and oxazoline, which are composed by heterocyclization domains present on the synthetases. Therefore incorporation of such domains on peptide synthetases could lead to new pharmaceutical substances [187].

Nowadays, there are an increasing number of examples for functional engineered peptide synthetases. Genetic redesign requires well-defined sequence information about the biosynthetic system that will be altered. Although this is often provided, manipulation has been unsuccessful in some cases, due to possible disruptions on some catalytic site(s) [182]. Therefore, information on domain structures as well as on possible protein-protein interaction sites between domains would improve manufacturing of novel antibiotic compounds.

Miscellaneous antibiotics produced by *Bacilli*

Bacilli do not produce only peptide antibiotics, but also several other secondary metabolites such as polyketides. Their biosynthesis occurs on PKSs by step-wise decarboxylative condensations (see chapter 1.3.2.4.2). Difficidin, oxydifficin as well as bacillaene are polyketides produced by various *Bacilli* strains and exhibit antibacterial activity. Posttranslational modification occurs by the 4'-PPTase Sfp [137]. Therefore, strains containing intact PKSs but defective *sfp* gene are deprived of polyketide production.

Furthermore, some new antibiotics have been recently isolated from various *Bacilli*. One of them is bacilysocin, a phospholipid that accumulates within the cells. It possibly derives from phosphatidylglycerol via acyl ester hydrolysis, a reaction controlled by YtpA. Bacilysocin inhibits the growth of various organisms such as *Staphylococcus aureus*,

Saccharomyces cerevisiae and the fungi *Candida pseudotropicalis*, *Cryptococcus neoformans* [188]. Furthermore, *Bacilli* produce low weight phenylpropanol derivative substances named isocoumarins with antibacterial and anti-inflammatory activity. Among them, amicoumacins could be used for treatment of chronic gastritis and peptic ulcer in humans, as they act against *Heliobacter pylori* [189]. Moreover, 3, 3'-neotrehalosadiamine (NTD), an aminosugar antibiotic produced by *B. pumilus* and *B. circulans*, inhibits the growth of *Staphylococcus aureus* and *Klebsiella pneumoniae*. In *B. subtilis* production is achieved only in RNA polymerase mutated strains that show resistance to rifampicin and is driven by the operon *ntdABC*. Expression is induced by NTD itself, via the regulatory protein NtdR [190].

Goal setting

Bacillus amyloliquefaciens and its numerous natural isolates are closely related to the already sequenced "Methuselah of the labs" *Bacillus subtilis* 168, but in parallel show broad biotechnological interest and often unique and remarkable characteristics. Since no representative of the *B. amyloliquefaciens* species had been yet sequenced, the molecular and biochemical work on these strains was hindered and the elucidation of the pathways that contribute to the organism's characteristics remained incomplete. Therefore, our laboratory, in collaboration with the GenoMik Network in Göttingen, set out to map the sequence of the plant growth promoting strain of *B. amyloliquefaciens* FZB42. This strenuous work that started in 2001 comprised a big part of my research the past years.

Nevertheless the primary focus of my work has been the elucidation and characterisation of pathways involved in the beneficial features of *Bacillus amyloliquefaciens* FZB42. For this scope and since the genome sequencing project did not immediately deliver results, alternative methods had to be employed in order to compare the FZB42 strain with its sequenced relative *B. subtilis* 168 and find the unique genomic regions that might be associated with the plant growth promoting abilities of *Bacillus amyloliquefaciens* FZB42.

The finding of such gene candidates (by both genomic and non-genomic approaches) generated new questions that I explored to answer. What are their products and what is the mechanism of action? When and how are they produced? How is their expression regulated? What is the effect of global regulators in their expression? But before all these questions could be answered, a protocol had to be established for the genetic manipulation of the natural isolate strain of *Bacillus amyloliquefaciens* FZB42.

To conclude, my thesis aimed to provide a first insight view of the unique features that enable *Bacillus amyloliquefaciens* FZB42 to promote plant growth. In order to accomplish such a task, a dual genomics and functional genomics approach was adopted. In parallel to this, the elucidation of the organism's genome sequence sets the ground for future work with other isolated strains of the same species and adds important information to the function and evolution of the *Bacilli* genus.

Materials and Methods

Chemicals and materials

All chemicals and materials used in the present study are listed in table 1.

Table 1: Chemicals and materials used in the present study

Manufacturer	Product
Amersham Pharmacia	[γ- ^{32}P]ATP, Plus One Tris-Base, Plus One EDTA, Plus One boric acid, Ready to Go DNA labelled Beads
BD	Difco medium 3
Biorad	Blotting grade blotter non-fat dry milk
Bioron	Taq polymerase
Fermentas	DNA markers, dNTPs, prestained protein ladder, RevertAid M-MuLV reverse transcriptase (200U/µl), restriction endonucleases, RiboLock ribonuclease inhibitor (40U/ µl), T4 DNA ligase, T4 kinase, T4 Polynucleotide kinase
Fluka	$CaCl_2$, EDTA
Macherey-Nagel	Nitrocellulose membrane porablot NCL, Nucleo Spin ® Extract II, Nucleo Spin RNA L, Porablot NY plus, Protino® Ni-1000 kit
Merck	Meat extract
MP Biomedicals	Urea pure
New England Biolabs	MidRange II PFG marker, Vent Polymerase
Promega	BCIP (50 mg/ml), NBT (50 mg/ml), pGEM-T® Vector systems
Qiagen	QIAEX II gel extraction kit, QIAprep Spin mini prep kit, Qiaquick PCR purification kit
Roche	Anti-DIG AP, Ampicillin, blocking reagent, DIG-dUTP, kanamycin
Roth	Agarose, chloramphenicol, citric acid, $CuSO_4$, DEPC, $FeCl_2$, $FeCl_3$, $Fe_2(SO_4)_3$, formaldehyde, L-glutamic acid, glycerol, HEPES, IPTG, KCl, K_2HPO_4, H_2KPO_4, maleic acid, $MgSO_4$, $MnCl_2$, $MnSO_4$, Na-acetate, Na-citrate, Na_2CO_3, NaCl, NaOH, $(NH_4)_2SO_4$, peptone, SDS, Proteinase K, Rotiphorese Gel 40 (19:1), Rotiphorese Gel 40 (29:1), TEMED, Tris, Triton-X 100, Tween 20, XGal, yeast extract, $ZnCl_2$
Santa Cruz Biot.	His-probe H15 sc-803 rabbit polyclonal IgG (200 mg/ml)
Serva	Agar, APS, boric acid, casamino acids, DTT, EGTA, erythromycin, glucose, N-Lauroylsarcosine-sodium, lincomycin/HCl, $MgCl_2$, MOPS, NaN_3, Na_2SO_4, ONPG, L-tryptophan
Sigma	Oligonucleotides, Anti-rabbit IgG AP
USB	Low-melting point agarose, Thermo Sequenase cycle Sequencing kit

Plasmids, bacterial strains and primers

The plasmids, bacterial strains and primers used in this study are listed in tables 2, 3, 4 respectively.

Table 2: Plasmids used in the present study

Plasmid/reference	Description
pDG148/[191]	*E. coli* and *B. subtilis* shuttle vector, IPTG-inducible P_{spac} promoter, Ap^r Km^r $Phleo^r$
pDG268/[192]	pBR322 derivarive with promoterless *lacZ*, integrative vector for recombination into *B. subtilis amyE*, Ap^r Cm^r
pGEM-T/Promega	Cloning vector, Ap^r
pECE73/[193]	$Cm^r \rightarrow Km^r$ exchange vector, Ap^r
pMX39/[194]	*E. coli* and *B. subtilis* shuttle vector based on pBR322 and PDB101, Ap^r Em^r
pQE60/Qiagen	Expression vector, IPTG-inducible promoter, His_6-Taq, Ap^r
pREP4/Qiagen	Repressor plasmid encoding *lacI*, Km^r
pAK1[a]	pGEM-T carrying 1,2 kb fragment of *bmyA*
pAK2	pGEM-T carrying *bmyA*::*Em*^{r}
pAK3	pGEM-T carrying 1,3 kb fragment of *fenA*
pAK4	pGEM-T carrying *fenA*::C*m*^{r}
pAK5	pDG268 carrying a fragment of *bmyD* from -400 to +126 bp (relative to the start codon)
pAK6	pDG268 carrying a fragment of *bmyD* from -183 to +126 bp (relative to the start codon)
pAK7	pDG268 carrying a fragment of *bmyD* from -120 to +126 bp (relative to the start codon)
pAK8	pDG268 carrying a fragment of *bmyD* from -30 to +126 bp (relative to the start codon)
pAK9	*B. amyloliquefaciens* FZB42 integrative vector *amy*::*lacZ*, Ap^r Cm^r, pDG268 derivative
pAK10	pGEM-T carrying a kanamycin cassette, Km^r
pAK12	pAK10 derivative carrying flanking regions of *yvrGyvrH*
pAK15	pAK12 derivative;Km^r is replaced by Cm^r
pAK16	pAK9 carrying a fragment of *bmyD* from -400 to +126 bp (relative to the start codon)
pAK17	pAK9 carrying a fragment of *bmyD* from -183 to +126 bp (relative to the start codon)

Plasmid/reference	Description
pAK18	pAK9 carrying a fragment of *bmyD* from -120 to +126 bp (relative to the start codon)
pAK19	pAK9 carrying a fragment of *bmyD* from -30 to +126 bp (relative to the start codon)
pAK25	Integrative vector carrying Cm^r cassette flanked by *bmyA-His$_6$- Taq* and *BmyB* sequences; pAK15 derivative
pAK27	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *yerPyerO*; pAK15 derivative; used for Δ*yerPyerO*::*Cm*^{r}
pAK29	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *sig01*; pAK15 derivative; used for Δ*sig01*::*Cm*^{r}
pAK33	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *spaR*; pAK15 derivative; used for Δ*spaR*::*Cm*^{r}
pAK35	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *sigW*; pAK15 derivative; used for Δ*sigW*::*Cm*^{r}
pAK39	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *aat*; pAK15 derivative; used for Δ*aat*::*Cm*^{r}
pAK41	Integrative vector carrying Cm^r cassette flanked by sequences of *bmyB-moduleB1* and *moduleB2* ; pAK15 derivative
pAK43	Integrative vector carrying Cm^r cassette flanked by sequences of *bmyB-moduleB2* and *moduleB3* ; pAK15 derivative
pAK45	Integrative vector carrying Cm^r cassette flanked by sequences of *bmyB-moduleB3* and *moduleB4* ; pAK15 derivative
pAK47	Integrative vector carrying Cm^r cassette flanked by sequences of *bmyB-moduleB4* and *bmyC- moduleC1* ; pAK15 derivative
pAK49	Integrative vector carrying Cm^r cassette flanked by sequences of *bmyC-moduleC1* and *moduleC2* ; pAK15 derivative
pAK51	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *codY*; pAK15 derivative; used for Δ*codY*::*Cm*^{r}
pAK54	pQE60 derivative carrying *degU*
pAK58	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *sigD*; pAK15 derivative; used for Δ*sigD*::*Cm*^{r}
pAK60	Integrative vector carrying Cm^r cassette flanked by neighbouring sequences of *sigH*; pAK15 derivative; used for Δ*sigH*::*Cm*^{r}

Plasmid/reference	Description
pAK61	pGEM-T carrying 1,2 kb fragment of *rapX*
pAK63	pGEM-T carrying *rapX*::*Cm*^{r}
pAK64	pDG148 carrying a bp fragment of *degQ*

[a] The reference is omitted in case the plasmids were prepared for this study.

Table 3: Bacterial strains used in the present study

Strain	Genotype	Reference
B. amyloliquefaciens FZB42	Wild type	FZB Berlin
B. amyloliquefaciens FZB45	Wild type	FZB Berlin
B. amyloliquefaciens FZB24	Wild type	FZB Berlin
B. subtilis 168	*trpC2*	Laboratory stock
B. subtilis MO1099	JH642; MLS^r; *amyE*::*erm trpC2 pheA1*	[195]
B. subtilis FZB37	Wild type	FZB Berlin
AK1	FZB42 *bmyA*::*Em*^{r}	This study
AK2	FZB42 *fenA*::*Cm*^{r}	This study
AK3	FZB42 *bmyA*::*Em*^{r} *fenA*::C*m*^{r}	This study
AK4	MO1099 *amyE*::P$_{bmyD400bp}$-*lacZ* (*Cm*^{r})	This study
AK5	MO1099 *amyE*::P$_{bmyD183bp}$-*lacZ* (*Cm*^{r})	This study
AK6	MO1099 *amyE*::P$_{bmyD120bp}$-*lacZ* (*Cm*^{r})	This study
AK7	MO1099 *amyE*::P$_{bmyD30bp}$-*lacZ* (*Cm*^{r})	This study
AK8	FZB42 Δ*RBAM01839/RBAM01840*::*Cm*^{r}	This study
AK9	FZB42 *amyE*::P$_{bmyD400bp}$-*lacZ* (*Cm*^{r})	This study
AK10	FZB42 *amyE*::P$_{bmyD183bp}$-*lacZ* (*Cm*^{r})	This study
AK11	FZB42 *amyE*::P$_{bmyD120bp}$-*lacZ* (*Cm*^{r})	This study
AK12	FZB42 *amyE*::P$_{bmyD30bp}$-*lacZ* (*Cm*^{r})	This study
AK13	FZB42 *amyE*::P$_0$-*lacZ* (*Cm*^{r})	This study
AK14	AK9 *yczE*::*Em*^{r}	This study
AK15	FZB42 *bmyA-His$_6$-Taq*::*Cm*^{r}	This study
AK16	FZB42 *amyE*::P$_{bmyD400bp}$-*lacZ* (*Km*^{r})	This study
AK17	FZB42 *amyE*::P$_{bmyD183bp}$-*lacZ* (*Km*^{r})	This study

Strain	Genotype	Reference
AK18	AK16 Δ*RBAM01839/RBAM01840*::*Cm*^{r}	This study
AK19	AK17 Δ*RBAM01839/RBAM01840*::*Cm*^{r}	This study
AK20	AK16 Δ*yerPyerO*::*Cm*^{r}	This study
AK21	AK17 Δ*yerPyerO*::*Cm*^{r}	This study
AK22	AK16 *comA*::*Em*^{r}	This study
AK23	AK17 *comA*::*Em*^{r}	This study
AK24	FZB42 Δ*yerPyerO*::*Cm*^{r}	This study
AK25	FZB42 Δ*sig01*::*Cm*^{r}	This study
AK26	AK16 *yczE*::*Em*^{r}	This study
AK27	AK17 *yczE*::*Em*^{r}	This study
AK28	AK16 Δ*sig01*::*Cm*^{r}	This study
AK30	AK16 Δ*spaR*::*Cm*^{r}	This study
AK31	FZB42 Δ*spaR*::*Cm*^{r}	This study
AK32	AK16 *degU*::*Em*^{r}	This study
AK33	AK17 *degU*::*Em*^{r}	This study
AK34	AK16 Δ*sigW*::*Cm*^{r}	This study
AK35	AK17 Δ*sigW*::*Cm*^{r}	This study
AK36	FZB42 Δ*sigW*::*Cm*^{r}	This study
AK37	AK17 Δ*spaR*::*Cm*^{r}	This study
AK39	FZB42 *bmyB-moduleB2*::*Cm*^{r}	This study
AK40	FZB42 *bmyB-moduleB3*::*Cm*^{r}	This study
AK41	FZB42 *bmyB-moduleB4*::*Cm*^{r}	This study
AK42	FZB42 *bmyC-moduleC1*::*Cm*^{r}	This study
AK43	FZB42 *bmyC-moduleC2*::*Cm*^{r}	This study
AK44	FZB42 Δ*aat*::*Cm*^{r}	This study
AK45	FZB42 Δ*codY*::*Cm*^{r}	This study
AK46	AK16 Δ*aat*::*Cm*^{r}	This study
AK47	AK17 Δ*aat*::*Cm*^{r}	This study
AK48	FZB42 *sigX*::*Km*^{r}	This study (pECE73→UL1)
AK49	FZB42 *sigX*::*Km*^{r} Δ*sigW*::*Cm*^{r}	This study
AK50	FZB42 Δ*sigH*::*Cm*^{r}	This study
AK51	FZB42 Δ*sigD*::*Cm*^{r}	This study
AK52	AK16 Δ*sigH*::*Cm*^{r}	This study

Strain	Genotype	Reference
AK53	AK17 Δ*sigH*::*Cm*^{r}	This study
AK56	AK16 Δ*sig01*::*Cm*^{r}	This study
AK57	FZB42 *sigB*::*Em*^{r} *rapX*::*Cm*^{r}	This study
AK58	FZB42 *degU*::*Em*^{r} with pAK64 (*Km*^{r}, *Phleo*^{r})	This study
AK59	FZB42 *rapX*::*Cm*^{r}	This study
AK60	AK4 with pAK64 (*Km*^{r}, *Phleo*^{r})	This study
AK61	AK5 with pAK64 (*Km*^{r}, *Phleo*^{r})	This study
CH1	FZB42 *srfAA*::*Em*^{r}	[196]
CH3	FZB42 *sfp*::*Em*^{r}	[197]
CH4	FZB42 *yczE*::*Em*^{r}	X.-H.Chen, unpublished
CH23	FZB42 *comA*::*Em*^{r}	X.-H.Chen, unpublished
CH30	FZB42 *sigV*::*Em*^{r}	X.-H.Chen, unpublished
CH33	FZB42 *sigB*::*Em*^{r}	X.-H.Chen, unpublished
TF1	FZB42 *degU*::*Em*^{r}	T.-F. Huang, unpublished
UL1	FZB42 *sigX*::*Em*^{r}	U. Leppert, diploma work
E. coli DH5α	*supE44* Δ*lacU169* (*Φ80 lacZΔM15*) *hsdR17 recA1 gyrA96 thi-1 relA1*	Laboratory stock
E. coli JM101	*supE thiA (lac-proAB) tra D36, pro AB , lac 9,Z A M15*	Laboratory stock
AK38	*E. coli* DH5α pREP4 pAK54	This study

Table 4: Primers used in this study

Primer name (restriction site)	Sequence (5' to 3' end)	Use
pRB1601 [6]	TAATACATGCAAGTCGAGCGG	Riboprint analysis
pRB1602 [6]	ACGTATTACCGCGGCTGCTGGC	Riboprint analysis
ssh1	TCGAGCGGCCGCCCGGGCAGGT	SSH
ssh2	AGCGTGGTCGCGGCCGAGGT	SSH
bmyAa	AAAGCGGCTCAAGAAGCGAAACCC	pAK2

Primer name (restriction site)	**Sequence (5' to 3' end)**	**Use**
bmyAb	CGATTCAGCTCATCGACCAGGTAGGC	pAK2
fenAa	AAGAGATTCAGTAAGTGGCCCATCCAG	pAK3
fenAb	CGCCCTTTGGGAAGAGGTGC	pAK3
cm1*Kpn*I	TGAGGTACCATGTTTGACAGCTTATCATCGGC	pAK4
cm2*Hind*III	TATGCCAAGCTTTTCTTCAACTAACGGGGCAGG	pAK4, pAK63
bmyA1(*Apa*I)	TTACTGGGCCCAAGACTTTGCAGTTTGGCAGC	pAK25
bmyA2(*Sph*I)-His_6-Taq	TTATCGCATGCTCAGTGGTGGTGGTGGTGGTGAA AGTTCAATTGAATAGAATCAAGCG	pAK25
bmyB1(*Spe*I)	TCATACTAGTGATGAAGAAGACGGCCTAAGCG	pAK25
bmyB2(*Pst*I)	TCATCTGCAGATTCGCCTTCTCATTCAGTTCCC	pAK25
bmyB12b1(*Sph*I)	TCATGCATGCAACAGCTTTTGGAGCAGACGCG	pAK41
bmyB12b2(*Age*I)	AACTACCGGTTCGGAGCTTATGTCACACG	pAK41
bmyB12f1(*Spe*I)	TCATACTAGTAGCGTCTCAACTAGTTGAGACAC ACC	pAK41
bmyB12f2(*Sal*I)	GCAAGTCGACGATTAGTACGCCTTTTGGCC	pAK41
bmyB23b1(*Sph*I)	TCATGCATGCTATCCATTGACGAATTGGATCAGC	pAK43
bmyB23b2(*Age*I)	AACTACCGGTGAAGACAACGTCTGCGGACCC	pAK43
bmyB23f1(*Spe*I)	GAGGACAGCACTAGTGCTGATACG	pAK43
bmyB23f2(*Sal*I)	GCAAGTCGACGTCACAACATGAGTGCAGCTGC	pAK43
bmyB34b1(*Sph*I)	ATTCGCATGCTCAAGCGAAAGAAGAACAGGCGG	pAK45
bmyB34b2(*Age*I)	TTACACCGGTCGTTCGTATCGCCAGTGTGCC	pAK45
bmyB34f1(*Spe*I)	TCTTACTAGTTTCCGGGAGTACGTGCAGG	pAK45
bmyB34f2(*Sal*I)	GCTTTTCTCGTCGACTGCGGC	pAK45
bmyB4C1b1(*Sph*I)	ATTCGCATGCCAGGAGTTGTTCTGGAGCAGC	pAK47
bmyB4C1b2(*Age*I)	CGTTTGATCACCGGTACGTTCCG	pAK47
bmyB4C1f1(*Spe*I)	AAAGGGCGAATCAACTAGTTCG	pAK47
bmyB4C1f2(*Sal*I)	CAAACTTCTCGGCCGTCGACTCGGG	pAK47
bmyC12b1(*Sph*I)	ATCAACAAGATCACAAGCATGCGTCAG	pAK49
bmyC12b2(*Age*I)	AATGCGTCTGCAACCGGTCGACACTTGC	pAK49
bmyC12f1(*Spe*I)	TCTTACTAGTAAATTATGAAGCAAATGGCGGACG	pAK49
bmyC12f2(*Sal*I)	GCAAGTCGACTTCGGAATAATCACTAATTTGCCC	pAK49
bmyD1(*Eco*RI)	CCGGAATTCCAGATCCATCTCTTGCGCC	pAK5, pAK16, EMSA, FT
bmyD2(*Eco*RI)	CCGGAATTCGATTTTCGGTGAAACCCC	pAK6, pAK17, EMSA

Primer name (restriction site)	**Sequence (5' to 3' end)**	**Use**
bmyD3(*Eco*RI)	CCGGAATTCCGAACAATAACTCCTCCG	pAK7, pAK18
bmyD4(*Eco*RI)	CCGGAATTCTCCCCTGTTCAATATGATCGGAGG	pAK8, pAK19
bmyD5(*Bam*HI)	TCGGGATCCCAAGGAGATCGCATCGCTCG	pAK5 to pAK8, pAK16 to pAK19
FM2	TATCGGCCTCAGGAAGATCGCACTC	Fusions control
amyEf1(*Xba*I)	TCGATTCTAGACGTCATCGGTCAAAAACGGG	pAK9
amyEf2(*Xho*I)	TGACTCTCGAGCGGGAACCAATCACTGCCC	pAK9
degQ1(*Hind*III)	ACTCAAGCTTAAAAAAAGGAGTGTGGAAACGG	pAK64
degQ2(*Sph*I)	ACTCGCATGCTGCACAAAAAAAAGACTTGTTTCC	pAK64
spac	GACTATTCGGCACTGAAATTATGGG	pAK64 control
rev1	CCTACAAATTGAGACCCTTGTCCAGG	PE (P_{bmy}), EMSA, FT
rev2	TAAAACATGGGGGTTTCACCG	PE (P_{bmy})
sigHb1(*Apa*I)	ATTCGGGCCCACATGATTGGAGCTTGGCCG	pAK60
sigHb2(*Age*I)	TTACACCGGTAATGACCTGCTCGTCCTCC	pAK60
sigHf1(*Spe*I)	ATTCACTAGTGATAATGCCCTGCAGCGCG	pAK60
sigHf2(*Sal*I)	TTACGTCGACGGCTCAGGGCCTATGAATCC	pAK60
yvrGb1(*Spe*I)	TTTCACTAGTATCACCATTCACAGCACCGC	pAK15
yvrGb2(*Pst*I)	TCTTCTGCAGCTCCTTCGCATCATTTTGGC	pAK15
yvrHf1(*Sph*I)	TTCTGCATGCTTTGAACGATCCGCAGGC	pAK15
yvrHf2(*NcoI*)	TATCTCCATGGCTTATTGCGATGCTGATGCC	pAK15
gatB1(*Apa*I)	TACTGGGCCCTTTGAACTGCGAAATCGCAACGG	pAK27
gatB2(*Sph*I)	TACTGCATGCATCTTGTTGACCATCGGCGGG	pAK27
yerQ1(*Spe*I)	TACTACTAGTAATCCGACTTCAGGACGGGAGC	pAK27
yerQ2(*Pst*I)	TTACTCTGCAGTTCGCCGTCCAGGTTCAGCTGC	pAK27
spaG1(*Age*I)	GTTTGCCACCGGTCGAATCGCTCC	pAK33
spaG2(*Sph*I)	CCGTGCTTTTACGATAGCATGCGGGCCG	pAK33
spaK1(*Pst*I)	GTAAGCCCCCTGCAGTGATGCCCC	pAK33
spaK2(*Spe*I)	GGGGTGTCGGATACTAGTGGGAATAGC	pAK33
sigWb1(*Sph*I)	CGTAACGTCTTCGCCGCATGC	pAK35
sigWb2(*Age*I)	CCTCTGCCCTTCACCGGTCTG	pAK35
sigWf1(*Spe*I)	GGCTCTTAGAAAACTAGTGAGGG	pAK35
sigWf2(*Pst*I)	GTTATCGCTTGGTCCTGCAGCC	pAK35
srfDDb1(*Sph*I)	ATTCGCATGCTATTCCGCATCATTCCGCC	pAK39
srfDDb2(*Age*I)	AGTTACCGGTCTGTTCAGCTCTTTTGCTGC	pAK39

Primer name (restriction site)	Sequence (5' to 3' end)	Use
aatf1(*Spe*I)	ACTTACTAGTGTTGAAGAAGAACACATCGC	pAK39
aatf2(*Sal*I)	TCTTGTCGACTTTCCTGATCCTGTTGTCCG	pAK39
asig01a(*Eco*RI)	TTCGGAATTCGAAGCAGGAGCTGGAAAAGGAGG	pAK29
Asigo1b(*Pst*I)	TTCGCTGCAGGCTTTCGGGTCTATCGGTTTGC	pAK29
guaA1(*Sph*I)	CAAGGCATGCATGAAGCGGACAAGCTGAAAGG	pAK29
guaA2(*Age*I)	CAAGACCGGTCTTCCTTCACCTTATCCACCTCC	pAK29
sigDb1(*Apa*I)	GATTCGGGCCCGCTTTATGAGCCGTGCGG	pAK58
sigDb1(*Age*I)	TTACACCGGTCCGGCTTTAGGATCTTTCC	pAK58
sigDf1(*Spe*I)	ATTCACTAGTACAGATTCATTCAAAGGCGC	pAK58
sigDf2(*Sal*I)	TTACGTCGACCGTTTGCAGCACCCTCTGC	pAK58
codYb1(*Sph*I)	ATTCGCATGCCAGGCAAATTAATCGATATGG	pAK51
codYb2(*Age*I)	TTACACCGGTATAAATAATCCTCCTAGAATTCC	pAK51
codYf1(*Spe*I)	TTCTGAACAACTAGTTCCGTATCG	pAK51
codYf2(*Sal*I)	GCAAGTCGACATTTTCCTCCTGTCAAGACGG	pAK51
degU60a(*Nco*I)	AATCCATGGCTAAAGTAAATATTGTTATTATCG	pAK54
degU60b(*Bgl*II)	AATAGATCTACGCATCTCTACCCAGCCG	pAK54
bmyD6	AGTCTTAAAGAGAGATGATGAAAGCC	n.r.EMSA
rapX1	GATTTGTTCGGCTTGTGCCGTTGAAC	pAK63
rapX2	TACTTGTCAGACTGTGACGGCG	pAK63
cm1(*Hind*III)	TTCTAAGCTTCATGTTTGACAGCTTATCATCG	pAK63
yczeu	CGGCAAAATAAAACGTCCAGCG	PE(P_{yczE})

FT, DNAse I footprinting; PE, primer extension; P_{bmy}, promoter of the *bmy* operon; P_{yczE}, promoter of *yczE*; n.r. EMSA, non-radioactive EMSA; the enzyme recognition site within each primer is underlined.

2.3 Media and supplements

All media used in this work were prepared and sterilized according to [198, 199]. Supplements with different antibiotics and compounds are listed in table 5. For antibiotic production and mass spectrometry measurements, bacteria were grown either in Landy medium [200] or sucrose-ammonium citrate medium (ACS) [201].

●LB (Luria-Broth) medium

1 % w/v	peptone
0,5 % w/v	yeast extract
0,5 % w/v	NaCl

●Landy medium

2% w/v	glucose
0,5 % w/v	Na-glutamate
2 mM	$MgSO_4$
6,7 mM	KCl
0,007 mM	KH_2PO_4
$0{,}015x10^{-3}$ % w/v	$Fe_2(SO_4)_3$
$5x10^{-4}$ % w/v	$MnSO_4$
$0{,}016x10^{-3}$ w/v	$CuSO_4$

●Naehr agar

0,5 % w/v	peptone
0,3 % w/v	meat extract
1,5 % w/v	agar

●ACS medium

10% w/v	sucrose
0,06 M	citric acid
0,028 M	Na_2SO_4
0,5 % w/v	yeast extract
0,032 M	$(NH4)_2HPO_4$
0,01 M	KCl
2 mM	$MgCl_2$ x6H_2O
0,001% w/v	$ZnCl_2$
0,0024% w/v	$FeCl_2$ x 6H_2O
0,018 % w/v	$MnCl_2$ x 4H_2O

Fungi were grown on "potato agar" at room temperature. When fungi and bacteria had to be simultaneously grown on plates, then Waksman agar was used and the microorganisms were let to grow at 27°C.

●Potato agar

4% w/v	Potato puree (without milk)
2% w/v	glucose
2% w/v	agar

●Waksman agar

0,5% w/v	peptone
1% w/v	glucose
0,085 M	NaCl
0,3% w/v	yeast extract
2% w/v	agar

Table 5: Supplements

Supplement	Final concentration
Agar	1,5 % w/v, 0,75 % w/v (soft agar plates)
Amplicillin	100 µg/ml
Chloramphenicol	20 µg/ml (for *E. coli*), 5 µg/ml (for *Bacilli*)
Erythromycin	1 µg/ml (for *Bacilli*)
IPTG	1 mM
Kanamycin	20 µg/ml (for *E. coli*), 5 µg/ml (for *Bacilli*)
Lincomycin	25 µg/ml (for *Bacilli*)
XGal	40 µg/ml

Molecular Biology techniques

Standard molecular biology methods

DNA manipulation, such as digestion with restriction endonucleases and ligation, was performed according to the instructions supplied by the manufacturer. Agarose-gel-electrophoresis, fluorescent visualization of DNA with ethidium bromide, spectrophotometric quantitation of DNA as well as preparation of $CaCl_2$-competent *E. coli* cells followed by transformation of plasmid DNA were carried out with standard procedures described by [198]. Bacterial chromosomal DNA from *Bacilli* was prepared as described by [202]. Polymerase chain reaction (PCR) was done using the GeneAmp PCR system 2700 (Applied Biosciences) according to [203], under the appropriate conditions in each case. Ligation of PCR products to pGEM-T vector was carried out following the instructions of the manufacturer (Promega). Plasmid DNA isolation and recovery of DNA from agarose gels were performed with QIAprep Spin mini prep kit and QIAEX II gel extraction kit, respectively.

Transformation in *Bacillus subtilis*

Competent cells of *Bacillus subtilis* were prepared according to the protocol published by [204]. Cells were grown overnight in 10 ml KM1 buffer at 32°C on a rotary shaker (150 rpm) and the next morning were 1:10 diluted in 50 ml KM1 buffer. Cells were further grown at 37°C under vigorous shaking (175 rpm). Every 30 minutes, samples were collected and the

optical density at 600 nm was determined. At the beginning of stationary phase the culture was diluted 1:10 in 100 ml KM2 buffer and was further incubated at 37°C for 75 minutes (75 rpm). Subsequently the cells were harvested by a 10 minute centrifugation at 5000 rpm (room temperature) and the pellet was resuspended in 2 ml of the supernatant. Aliquots of 0,5 ml competent cells with 10% glycerol were stored at -80°C.

For the transformation, one aliquot was unfrozen by short incubation at 37°C. 1 μg of the desired DNA (chromosomal or linearized/circular plasmid DNA) was added and cells were incubated at 37°C for 30 minutes (50 rpm). Subsequently, 0,5 ml of LB medium, containing inducing concentration (0,1 μg/ml) of the appropriate antibiotic was added to the cells and they were further grown at 37°C for 75 minutes (200 rpm). Aliquots of the culture were plated on selective agar plates.

Buffers

●KM1 buffer

1 x SMM buffer
1mM $MgSO_4$
0,025 M glucose
0,04% w/v casamino acids
0,005% w/v tryptophan

●KM2 buffer

1 x SMM buffer
1 mM $MgSO_4$
0,015M glucose
$5x10^{-4}$ % w/v casamino acids
0,005% w/v tryptophan
1 mM $CaCl_2$

●10 x SMM buffer

0,15 M $(NH_4)_2SO_4$
0,8 M K_2HPO_4
0,45 M H_2KPO_4
0,034 M Na-citrate

Transformation in *Bacillus amyloliquefaciens*

Competent cells of *Bacillus amyloliquefaciens* were obtained by modifying the two-step protocol published by [205]. Cells were grown overnight in LB medium at 28°C (170 rpm). The next day, they were diluted in glucose-casein hydrolysate-potassium phosphate (GCHE) buffer to an OD_{600} of 0,3. The cell culture was then incubated at 37°C under vigorous shaking (200 rpm) until the middle of exponential growth (OD_{600} ~1,4). Dilution with an equal volume of GC medium followed and the cells were further incubated under the same conditions for 1 hour. Further on, the culture was divided in 2 ml Eppendorf tubes and cells were harvested by centrifugation at 6000 rpm for 5 minutes (room temperature). The pellets were resuspended in 200 μl of the supernatant and the desired DNA (1 μg) with 2 ml transformation buffer was added to them. After incubation at 37°C under shaking at 75 rpm for 20 minutes, 1 ml LB medium containing sublethal concentration (0,1 μg/ml) of the appropriate antibiotic was

added. The cells were grown under vigorous shaking for 90 minutes and platted on selective agar plates.

Buffers

●GCHE buffer

1 x	PC buffer
0,1 M	glucose
0,005% w/v	tryptophan
0,04 M	$FeCl_3$ / Na-citrate
0,25% w/v	potassium glutamate
3 mM	$MgSO_4$
0,1% w/v	casein hydrolysate

●GC buffer

1 x	PC buffer
0,1 M	glucose
0,005% w/v	tryptophan
0,04 M	$FeCl_3$ / Na-citrate
3 mM	$MgSO_4$

●10 x PC buffer

0,8 M	K_2HPO_4
0,45 M	H_2KPO_4
0,028 M	Na-citrate

●Transformation buffer

1 x	SMM buffer
1 mM	EGTA
0,025 M	glucose
0,02 M	$MgCl_2$

Suppression Subtractive Hybridization (SSH)

Suppression Subtractive Hybridization (SSH) is applied to two strains of the same species or genus and aims to find major sequence differences between them. SSH identifies unique DNA sequences of target strain (tester) that are absent from the reference strain (driver). The method was performed according to the protocol published by [206] and [207].

In principal, genomic DNA from two strains was digested separately with RsaI, yielding fragments of 100 to 1000 bp. The tester DNA was subdivided into two portions, each of which was ligated with a different adaptor (1 and 2R, see their sequence at the end of this section). The ends of the adaptors are unphosphorylated and thus only one strand of each adaptor attaches to the 5' end of the DNAs. At first ligation of the adaptor to the fragments of the tester strain was performed for 16 hours at room temperature. The mixture was then heated at 72°C for 5 min in order to inactivate the ligase.

Subsequently two hybridizations were performed. In the first one, excess of driver DNA was added to each adaptor-ligated lot separately. After denaturation of the two mixtures at 98°C for 2 minutes, the samples were allowed to anneal at 63°C for 90 minutes thus generating type *a*, *b*, *c*, *d* molecules (Fig. 10). During the second hybridization, the two primary hybridization samples are mixed together without denaturation in order to assure that the remaining single-stranded tester-specific DNAs can form the new type *e* molecules. The *e* molecules are double-stranded tester-specific DNAs with different ends that have resulted from ligation with different adaptors. Fresh denaturated driver is added to the mix to further enrich fraction *e* for tester-specific sequences. The samples were allowed to anneal at 63°C for 16 hours.

The entire population of molecules was then subjected to PCR. First the reaction mix was incubated in the thermal cycler at 72°C for 2 min in order for the adaptors to be extended (their recessed 3' ends were filled in during this step). PCR was performed using primers ssh1/ssh2 (Table 4), that annealed on the adaptors 1 and 2R respectively [T_{den}=94°C (30 sec), T_{anneal}=66°C (30 sec), T_{ext}=72°C (30 sec) for 35 cycles]. During the PCR, molecules *a* and *d* were missing the primer-annealing sites while type *b* molecules formed a panhandle-like structure. As a result, these three types of molecules could not be amplified. Type *c* molecules had only one primer annealing site and were thus amplified linearly. Due to the suppression PCR effect only type *e* molecules that had two different adaptors and contained tester-specific sequences could be exponentially amplified. The substracted DNAs were cloned into pGEM-T vector and sequenced.

The adaptors' sequences are given below.

Adaptor 1

5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT-3'

3'-GGCCCGTCCA-5'

Adaptor 2R

5'-CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT-3'

3'-GCCGGCTCCA-5'

●Hybridization buffer

50 mM HEPES-HCl PH=8.0

0,5 mM $MgCl_2$

0,2 mM EDTA PH=8.0

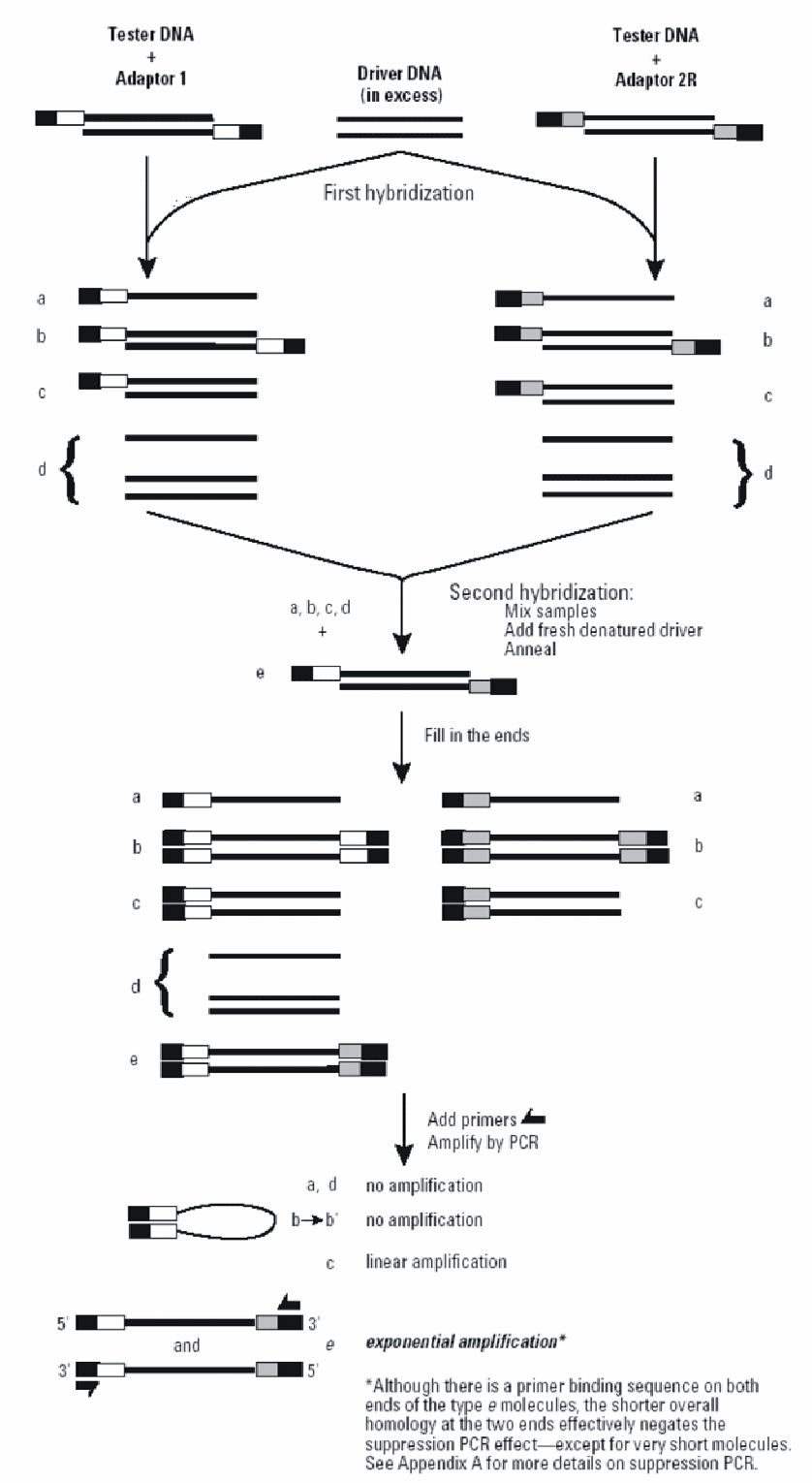

Figure 10: Schematic diagram of Suppression Subtractive Hybridization.

Tester DNA fragments that are ligated with adaptors 1 and 2R separately are further hybridized separately with excess of driver DNA. The samples are mixed together without denaturation and are hybridized in the presence of fresh denaturated driver. After the second hybridization, the PCR mixture is incubated at 72°C for 2 minutes in order for the recessed 3' ends of the adaptors to be filled in. Type *a*, *b* and *d* molecules cannot be amplified, due to lack of primer annealing site (*a* and *d*) and to formation of a panhandle-like structure (*b*).Type *c* molecules have only one primer annealing site

and thus are amplified linearly. Type *e* molecules are amplified exponential only if the sequence is present in the tester strain but absent from the driver strain. Solid lines stand for *Rsa*I digested DNAs. Filled boxes represent the outer identical parts of adaptors 1 and 2R. Clear and shaded boxes indicate the inner parts of adaptors 1 and 2R, respectively and correspond to the sequence of primers ssh1, ssh2. The figure is reproduced from [207].

Pulsed Field Gel Electrophoresis (PFGE)

Pulsed Field Gel Electrophoresis allows size separation of DNA fragments ranging from a few kilobase pairs to 10 megabase pairs. It operates by applying electric fields from different angles, thus making even very large DNA fragments to move through the gel and be efficiently separated. DNA is embedded in agarose in order to prevent shearing during purification.

Mid-exponential-phase *Bacillus* cells were used to prepare DNA for the PFGE [208]. After centrifugation for 10 minutes at 4°C and 4000 rpm, the pellet was resuspended in wash buffer. Plugs of cell suspensions prepared by mixing with 1% low-melting point agarose were first incubated overnight with lysis buffer at 56°C and were subsequently incubated overnight with digestion buffer at 50°C. After incubation with 1 x TE buffer containing 100 μM PMSF for 1 hour at 37°C, the plugs were stored at 4°C in 1 x TE buffer. Digestion was performed overnight using *SfiI* according to the instructions of the manufacturer and was stopped by addition of stop buffer. The plugs were loaded on a 1,2 % agarose gel in TBE buffer and PFGE was performed at 10°C using the Gene navigator electrophoresis unit of Pharmacia Biotech. Direction of the applied electric fields (160V) changed every 2 seconds for the first 2 hours of the run, every 10 seconds for the next 8 hours, every 25 seconds for the next 8,5 hours and every 40 seconds for the last 6 hours of the run. Visualization of the DNA was performed with ethidium bromide.

Buffers

●Wash buffer

10 mM	Tris-HCl PH=7.2
200 mM	NaCl
100 mM	EDTA PH=8.0

●Lysis buffer

10 mM	Tris-HCl PH=7.5
50 mM	NaCl
100 mM	EDTA PH=8.0
0,1 % w/v	lysozyme

●**Digestion buffer**		●**Stop buffer**	
0,5 M	EDTA PH=8.0	10 mM	Tris-HCl PH=8.0
1 % v/v	N-Lauroylsarcosine-sodium	50 mM	EDTA PH=8.0
0,05 % w/v	Proteinase K		

Hybridization analysis of Southern blots

Southern blot is a way of permanently immobilizing DNA (that has been separated by agarose gel electrophoresis) to a solid support. It is designed to locate a particular sequence of DNA within a complex mixture, such as an entire genome. Hybridization and detection occurs by "anealling" with a complementary labelled DNA probe.

Synthesis of DIG-labelled probe

For each southern hybridization, an appropriate probe was labelled with Digoxigenin-11-dUTP (DIG-dUTP), according to the Ready-to-Go kit from Roche. The desired DNA region was amplified by PCR and purified, prior to labelling. 100 ng of the PCR fragment were denaturated by heating at 100°C for 10 minutes and then mixed with 5 µl dCTP (10 mM), 2,5 µl DIG-dUTP (1mM) to a final volume of 50 µl. The mixture was incubated at 37°C for 1,5 hours and was stored at -20°C until use.

Preparation of samples; transfer and fixation on a membrane

1-2 µg of the chromosomal DNA in question were digested overnight with a suitable restriction endonuclease. Samples were initially separated on a 0,8 % agarose gel in 1 x TAE buffer at 70 Volt. The gel was washed twice for 20 minutes, initially with denaturation buffer and subsequently with neutralization buffer. Transfer on a nylon membrane was performed using the Biorad vacuum blotter (model 785). The DNA was fixed permanently on the membrane by cross-linking using UV radiation.

Buffers

●**Denaturation buffer**		●**Neutralization buffer**	
1,5 M	NaCl	1,5 M	NaCl
0,5 M	NaOH	1 M	Tris-HCl PH=8.0

Hybridization and detection

The membrane was initially incubated for 1 hour at 65°C with 40 ml hybridization buffer and was hybridized overnight at 55°C with 5-10 ml hybridization buffer containing 5-25 ng/ml of denaturated DIG-labelled probe. The membrane was washed twice for 15 minutes, first with 2 x SSC/0,1 % SDS at room temperature and then with 0,5 x SSC/0,1 % SDS at 55°C.

Detection was achieved by a colorimetric approach. The membrane was first equilibrated with P1-DIG buffer and was then incubated for 30 minutes with P1-DIG buffer containing 3,75 units of the antibody Anti-Digoxigenin-Alkaline-Phosphatase. Unbound antibody was removed after a fifteen minute washing step. Addition of 10 ml Ap buffer containing 2,25 mg nitroblue tetrazolium salt (NBT) and 1,75 mg 5-bromo-4-chloro-3 – indolyl phosphate (BCIP) to the membrane and incubation in the dark allowed visualization of the hybridized DNA with our labelled probe.

Buffers

●Hybridization buffer

5 x	SSC
1 % w/v	blocking reagent
0,1 % v/	N-Lauroylsarcosine-sodium
0,02 % w/v	SDS

●20 x SSC

3 M	NaCl
0,3 M	Na-citrate

●P1-DIG buffer

0,1 M	Maleic acid
0,15 M	NaCl
1 % w/v	blocking reagent

●Wash buffer

0,1 M	Maleic acid
0,15 M	NaCl
0,3 % v/v	Tween-20

●Ap buffer

0,1 M	Tris-HCl PH=9.5
0,1 M	NaCl
0,05 M	$MgCl_2$

Denaturating Gel Electrophoresis for Sequencing

Samples from primer extension, DNAse I footprinting and sequencing reactions were analysed on denaturating sequencing gels. High concentrations of urea in the gel secured that the DNA was completely denaturated and thus could be better separated.

The gel was let to prerun before loading the samples for 1 hour at 60 Watt in 1x TBE buffer, using the SequinGen Sequencing Cell of Biorad. After loading the samples, DNA separation was allowed for approximately 2 hours more using the same running conditions. The gel was dried at 80°C for 1 hour using the vacuum SlaB Gel Dryer Model SE1160. An IP

screen was put on the top of the dried gel and visualization was achieved using the Molecular Imager FX scanner (Biorad) or the phosphoimager 445SI (Molecular Dynamics).

●7 %	**Denaturating gel**
8 M	Urea
1 x	TBE
7%	Rotiphorese Gel 40 (29:1)
0,08% w/v	APS
0,06 % v/v	TEMED

Radioactive labelling of oligonucleotides

Oligonucleotides can be radio-labelled at their 5'-OH end by the T4 Polynucleotide kinase (T4 PNK) that catalyses the transfer of the γ-phosphate from ^{32}P- ATP.

Therefore, 40 pmol of primer were mixed with 4 µl [γ-^{32}P]ATP (10µCi/ml) and phosphorylation took place by incubation of the mixture with T4-Kinase at 37°C for 30 minutes. The reaction was stopped by heat inactivation at 70°C for 10 minutes.

Radioactive sequencing DNA

Sequencing reactions were carried out using the Thermo Sequenase cycle Sequencing kit (USB) according to the manufacturer's instructions. 300 ng of plasmid DNA containing the desired fragment and 1 pmol of the radioactive primer were included in the reaction. Amplification was performed using a 23 cycle PCR program [T_{den}=94°C (30 sec), T_{anneal}=58°C (sec), T_{ext}=72°C (30 sec)]

RNA preparation

Stationary-phase cells of *Bacillus amyloliquefaciens* were harvested for preparation of total RNA. 20 ml of the culture was mixed with 10 ml "killing" buffer (stops mRNA production) and centrifuged for 10 minutes at 4°C and 12000 rpm. The pellet was washed once more with 1 ml "killing" buffer and was then stored at -80°C.

Isolation of RNA was performed using the Nucleo Spin RNA L (Macherey Nagel). In order to remove possible DNA contaminations, the isolated RNA was additionally extracted with an acidic Phenol:Chloform:Isoamylalcohol (25:24:1) mixture and then chloroform. Ethanol precipitation followed and the pellet was resuspended in 20 µl DEPC-H_2O. The

concentration of total RNA was spectrophotometrically determined, according to [198] whereas its quality was checked on a 1,5% RNA agarose gel under denaturating conditions (1xMEN, 16% formaldehyde). The samples were mixed with 1,6 volume loading buffer and were incubated at 65°C for 5 minutes prior to loading on the gel. The gel was run in 1 x MEN buffer at 60 Volt.

Buffers

●"Killing buffer"		**●10 x**	**MEN**
20 mM	Tris-HCl PH=7.5	200 mM	MOPS
5 mM	$MgCl_2$	50 mM	Na-acetate
20 mM	NaN_3	10 mM	EDTA PH=7.0

Primer extension

Primer extension was used to map the 5' termini of mRNAs. 40 µg of total RNA was mixed with 0,15 µM radioactively (^{32}P) labelled primer at 70°C for 5 minutes. Then 4 µl 5 x reverse transcriptase buffer, 2 µl dNTPS (10 mM each) and 1 µl Ribonuclease inhibitor (40 units) were added to a final volume of 19 µl. After incubation at 37°C for 5 minutes, 1 µl reverse transcriptase (200 units) was added to the mixture and further incubation was allowed for 1 hour at 42°C. The primers used for identifying the transcriptional start(s) of *bmy* operon and *yczE* can be seen in table 4.

Electrophoretic Mobility Shift Assay (EMSA)

Electrophoretic Mobility Shift Assay (EMSA) is a technique used for determining protein-DNA interactions. It is based on the observation that DNA-protein complexes migrate slower through a non-denaturating polyacrylamide gel than free DNA fragments. Therefore, EMSA is a useful tool to determine if a protein binds directly to a DNA fragment or not.

In our case, the desired DNA fragment of the *bmyD* promoter region was amplified by PCR using primers bmyD1 and rev1 (Table 4), one of which was previously labelled at its 5'-end with [γ-^{32}P]ATP. The radio-labelled product (450bp) was purified with the Qiagen PCR purification kit. After dilution of the labelled DNA fragment to attain final activity of 10.000 cpm, the DNA was incubated at 37°C for 20 minutes with increasing concentrations of DegU protein in the 1xbinding buffer. The reaction mixtures were separated on 8% polyacrylamide

gels, under non-denaturating conditions, in 1 x TBE buffer at 60 V. The gels were visualized using the Biorad Molecular Imager FX scanner.

Non radioactive EMSA experiments were performed in a similar manner, but visualization was done fluorescently using ethidium bromide. In particular, two smaller DNA fragments were amplified using primers bmyD1 / bmyD6 and bmyD2 / rev1. The obtained fragments, D1 (217 bp) and D2 (233 bp) respectively, result together in the whole 450 bp fragment used in the radioactive EMSA.

●5 x Binding buffer

100 mM	Tris-HCl PH=8.0
500 mM	KCl
25 mM	$MgCl_2$
2,5 mM	DTT
50%	glycerol
0,25 %	Nonidet P40
0,025 % w/v	poly(dI-dC)
0,025 % w/v	BSA

●8% polyacrylamide gel (non-denaturating)

2 ml Rotiphorese Gel 40 (29:1)
1 ml 10 x TBE
0,285 µl glycerol
0,075 µl 10% APS
7,5 µl TEMED
6,63 ml H_2O

DNase I footprinting

DNase I footprinting is a method of studying protein-DNA interactions and identifying the DNA region to which a protein binds. These experiments were done as described by [209]. A DNA fragment (450 bp) obtained by PCR, using primers bmyD1 and rev1, or plasmid DNA carrying the same fragment (pAK16), were incubated in binding buffer with 0, 0,8 and 1,6 µM DegU protein for 20 min at 37°C. Complexes were then treated with DNase I (0,6 µg/ml) for 20 seconds and the reaction was stopped by addition of 10 µl stop buffer containing 1,75 ng/µl non-specific DNA (salmon sperm) and rapid chilling on ice. Primer extension followed with ^{32}P-labelled primers bmyD1 and rev1, for the template strand and for the non template strand, respectively [T_{den}=94°C (30 sec), T_{anneal}=58°C (30 sec), T_{ext}=72°C (30 sec) for 23 cycles].

●Stop buffer (10 µl)

2 µl	10xTaq buffer
1 pmol	^{32}P-labelled primer
0,2 µl	dNTPS (100mM)

0,5 µl	non-specific DNA (salmon sperm, 35 ng/ µl)
0,2 µl	Taq polymerase
6,6 µl	H_2O

Biological tests

For the antifungal tests, *B. amyloliquefaciens* FZB42 and its derivatives were grown in Landy medium at 37°C for 24 hours. The cultures were centrifuged and 2 µl of the supernant were spotted on Waksman agar together with regularly arranged growing fungi. The plates were incubated at 27°C.

For the antibacterial tests, *B. amyloliquefaciens* FZB42 and its derivatives were grown treated in the same manner. The indicator strain was grown overnight at 37°C under vigorous shaking. 300 µl of the culture was mixed with 3 ml soft agar and poured on LB dishes. Supernatants obtained from the *B. amyloliquefaciens* FZB42 strains, grown in Landy medium for 24 hours, were applied on the plates and were incubated at 37°C.

Biochemical methods

MS analysis

B. amyloliquefaciens FZB42 was grown overnight on agar plates of Landy medium at 37°C. To record mass spectra, cell material was picked from the plate, spotted onto the target and covered with matrix medium, i.e. a saturated solution of α-cyanocinnamic acid in 40% acetonitrile-0,1% trifluoroacetic acid. It was air dried and then analyzed by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS), as described in [210].

Alternatively, culture filtrate extracts were prepared by lyophilisation of the supernatants that resulted from cultures grown for 12, 24 or 48 hours at 37°C in Landy or ACS medium. A small sample of the culture filtrate was extracted with 70% acetonitrile-0,1% trifluoroacetic acid and then mixed with an equal volume of matrix medium. 1 µl aliquots were spotted on the target and were air dried prior to MS measurement [211]. Postsource decay (PSD) mass spectra were obtained with the same samples. Monoisotopic mass numbers were recorded.

Quantification of specific β-galactosidase enzymatic activity

Specific β-galactosidase activity was determined from growing liquid cultures in Difco medium, according to [199]. At different times of the growth curve the optical density of the culture at 600nm was determined and cells were harvested. Their pellets were frozen in order to be further used in the β-galactosidase assay. Pellets were resuspended in 640 μl Z-buffer and mixed with 160 μl lysozyme-buffer. After short vortexing, they were incubated at 37°C for 10 min. Further on, 8 μl of a 10% Triton-X solution was added to the samples, followed by ten-minute incubation on ice. The reaction began by addition of 200 μl ortho-nitrophenyl-β-D-galactopyranoside (ONPG) 4 mg/ml at 30°C and was stopped by addition of 400 μl 1M Na_2CO_3 when their colour changed to yellow. The samples were then centrifuged for 5 min and the supernatant's absorbance was measured at 420 and 550 nm. Specific β-galactosidase activity was calculated in Miller units (MU) [212], according to the formula.

Miller units (MU) = 1000 x (OD_{420} – 1,75 x OD_{550}) / (t x V x OD_{600})

OD_{420}, OD_{550}, OD_{600} = optical density at 420, 550, 600 nm

T = reaction time (min)

V = volume of the sample of bacterial cells used for the reaction (ml)

Buffers

●Z-buffer		●Lysozyme-buffer	●ONPG-buffer
60 mM	Na_2HPO_4 PH=7.0	2,5 mg/ml Lysozyme	4 mg/ml ONPG
40 mM	NaH_2PO_4	50 U/L Benzonase	in Z-buffer
10 mM	KCl	in Z-buffer	
1 mM	$MgSO_4$ x H_2O		

SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis was performed according to [198]. The proteins were separated by 4% upper and 10% lower gels, using the "Mini-Protean II" apparatus of Biorad. Gels were run at 200 Volt in 1 x running buffer and were stained with Coomasie Brilliant Blue.

●4% upper gel	**●10% lower gel**	
0,5 ml Rotiphorese Gel 40 (19:1)	3 ml	Rotiphorese Gel 40 (19:1)
1,25 ml 0,5 M Tris-HCl PH=6.83,28ml	1,5 M	Tris-HCl PH=8.8
0,05 ml 10% SDS	0,13 ml	10% SDS
0,025 ml 10% APS	0,066 ml	10% APS
0,006 ml TEMED	0,005 ml	TEMED
5,52 ml H_2O	3,2 ml	H_2O

●1 x running buffer

25 mM	Tris-HCl PH=8.3
1,92 M	Glycin
1%	SDS

Western Blot

The gels were electroblotted onto nitrocellulose membranes (0,45 µm pore size) using the "Trans-Blot SD Semi-Dry transfer cell" (Biorad). After incubation for 15 min with TBST buffer containing 10% powdered milk, the membranes were incubated overnight at 4°C with TBST buffer containing 1% powdered milk and 20 µl antibody "His-probe H15 sc-803 rabbit polyclonal IgG" (200 mg/ml; Santa Cruz). The membranes were washed 3 times with TBST buffer for 10 minutes and were then incubated with TBST buffer containing 2% powdered milk and 6 µl of the "Anti-rabbit IgG alkaline phoshatase conjugate" antibody (dilution 1:10.000) for 1 hour. Detection was achieved colorimetricly by addition of 10 ml Ap buffer containing 2,25 mg nitroblue tetrazolium salt (NBT) and 1,75 mg 5-bromo-4-chloro-3 – indolyl phosphate (BCIP).

●Transfer buffer

10%	Methanol
25 mm	Tris-HCl PH=8,7
150 mM	Glycin

Overexpression and purification of 6xHis-tagged DegU

A DNA fragment containing the whole *degU* gene of *B. amyloliquefaciens* FZB42 was amplified using primers degU60a / degU60b (table 4), that contain restriction sites for the

endonucleases *Nco*I and *Bgl*II respectively. DegU was cloned in the expression plasmid pQE60 (Qiagen) as a C-terminal His_6-tag fusion under regulation of an isopropyl β-D-thiogalactoside (IPTG) –inducible promoter. The resulting plasmid, pAK54, was used to overexpress DegU-His_6 fusion protein in *E. coli*. For this scope, DH5α strain was simultaneously transformed with pREP4 (a repressor plasmid carrying the *lacI* repressor) and pAK54 resulting in strain AK38.

Strain AK38 was grown overnight at 37°C in LB medium containing 100 μg/ml ampicillin and 20 μg/ml kanamycin. The culture was diluted in 500 ml LB-Ap/Km to an OD_{600} of 0,03 and was further grown at 30°C under vigorous shaking. When the cells grew to an OD_{600} of 1, ethanol was added to a final concentration of 3% to induce chaperone synthesis and minimize formation of inclusion bodies. After 15 minutes, IPTG was added to the cultures at 1 mM final concentration. The cultures were grown for 2,5 hours and were then centrifuged at 6000 rpm for 20 minutes. The pellets were stored at -80°C. The protein was then purified with the Protino® Ni-1000 kit according to the manufacture's instructions (Macherey Nagel) and was subsequently dialysed overnight against storage buffer at 4°C.

●Storage buffer

50 mM	Tris-HCl PH=8.0
0,5 mM	EDTA
100 mM	NaCl
0,5 mM	DTT
50% v/v	glycerol

Complete genome sequencing and annotation strategies

The genome of *B. amyloliquefaciens* FZB42 was sequenced in collaboration with the GenoMik Network in Göttingen using the random shotgun approach. Total genomic DNA was shared randomly or partially digested with *Sau*3AI, and DNA fragments 1 to 3 kb in size were cloned into pTZ19R or pCR2.2 TOPO (Invitrogen) to establish a shotgun library. The inserts of the recombinant plasmids were sequenced from both ends using the MegaBACE DNA Sequencing Systems 1000 and 4000 (Amersham-Biosciences) and ABI Prism 377 sequencers (Applied Biosystems) with dye terminator chemistry. Fosmid library and combinatorial multiplex PCR were performed in order to determine the RNA sequences present in the genome and their location within it.

Approximately 44068 sequences were processed with PHRED, assembled into contigs by using the PHRAP assembling tool [213] and edited with GAP4, which is a part of the STADEN package software [214]. The resulting contigs of *B. amyloliquefaciens* FZB42 were sorted using the genome of *B. subtilis* 168 as scaffold [7]. PCR-based techniques and primer walking on recombinant plasmids were applied in order to close remaining sequence gaps.

Identification of ORFs in all six different frames was performed using the Glimmer2 program. Annotation of the genome was done using the GeneSOAP program provided by Rainer Cramm. The full length of each ORF was determined according to the presence of putative ribosome-binding sites and putative start codons as well as by comparison to the ortholog ORFs in *B. subtilis* 168 [7] and in *B. licheniformis* DSM13 [11]. In addition, using GeneSOAP it was possible to look for conserved protein domains within the ORFs of *B. amyloliquefaciens* FZB42 by comparison to PFAM [215].

Results

Identifying unique DNA regions in the genome of B. amyloliquefaciens strain FZB42

Taxonomic classification of *Bacillus* strains FZB24, FZB37, FZB42, FZB45 and 168

Bacillus strains FZB24, FZB37, FZB42 and FZB45 have been isolated from plant-pathogen-infested soil and their contribution in plant growth promotion as well as in suppression of plant pathogenic organisms has been documented [5]. Initial studies on growth characteristics and carbon source utilization of those strains showed that they are closely related to the *Bacillus subtilis* and *Bacillus amyloliquefaciens* groups [6]. The main criterion for distinguishing between the two *Bacillus* subspecies was the ability shared within the *B. amyloliquefaciens* strains to produce lipase and acid from lactose [216]. Thereby it was concluded that FZB24, FZB42 and FZB45 belong to the *B. amyloliquefaciens* group whereas FZB37 is more related to *B. subtilis* [5]. However, another study classified strain FZB24 as a member of the *Bacillus subtilis* group. In order to further verify these results, ribotyping analysis and macrorestriction profiling by PFGE were performed.

The same amount (2 μg) of genomic DNA from FZB24, FZB37, FZB42, FZB45 and *B. subtilis* 168 was digested overnight at 37°C using the restriction endonuclease *Eco*RI. After transfer and fixation of the samples on a nylon membrane, overnight hybridization at 55°C was performed (see materials and methods). A DNA fragment, part of the 16S *rrnE* gene of *B. subtilis*, was amplified by PCR using primers pRB1601 and pRB1602 [6] and after labelling with DIG-dUTP, it was used as the probe for Southern hybridization. The ribotyping analysis revealed that the patterns obtained for FZB24 and FZB42 were almost identical (Fig. 11). FZB45 displayed a unique riboprint with high similarity to those belonging to FZB24 and FZB42. In contrast, FZB37 and *B. subtilis* 168 provide profiles that are identical to each other but quite distinct from the ones observed for the rest FZB strains. Comparison of these patterns with a database of known riboprints was performed in DSMZ (Deutsche Sammlung von Mikroorganismen and Zellkulturen GmbH, Braunschweig, Germany). Thereby it was confirmed that FZB24, FZB42 and FZB45 belong to the *B. amyloliquefaciens* group, whereas FZB37 belongs to the *B. subtilis* group.

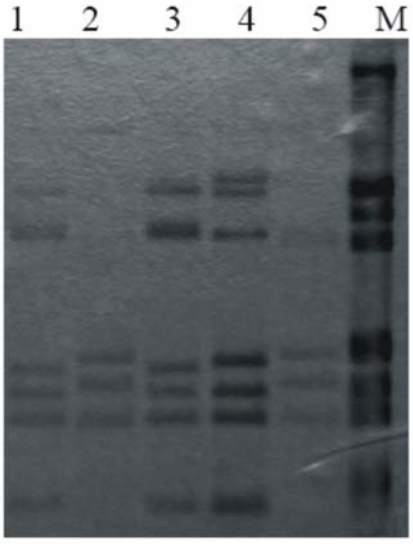

Figure 11: Riboprints of various *B. subtilis* / *B. amyloliquefaciens* strains.

Ribotype patterns obtained after digestion of genomic DNA of *B. amyloliquefaciens* FZB24 (1), FZB37 (2), FZB42 (3), FZB45 (4) and *B. subtilis* 168 (5) with *Eco*RI and hybridization with a DIG-labelled 16S rDNA probe. M, *Eco*RI / *Hind*III digested phage λ DNA; bands from bottom to top 0,9/1,4/1,6/2/3,5/4,2/5,1/21,2 kb

Moreover, *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 were further investigated by macrorestriction profiling. PFGE is a more analytical approach than riboprinting since it allows better separation of larger DNA fragments. The profiles of *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 obtained by PFGE after digestion with the restriction endonuclease *Sfi*I are very distinct from each other, as seen in (Fig. 12). Furthermore, the patterns obtained with the commercially available *B. amyloliquefaciens* strain FZB24, after digestion with the same restriction endonuclease, were identical to the ones resulting from FZB42 as reported by [6]. Similarly, FZB37 has the same macrorestriction pattern as *B. subtilis* 168 [6].

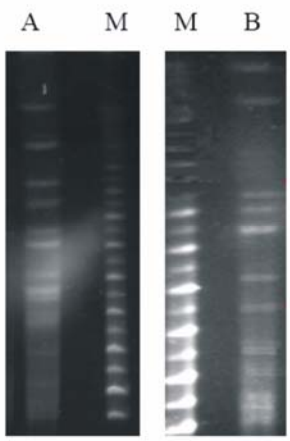

Figure 12: Genomic DNA macrorestriction profiles of *B. subtilis* 168 and *B. amyloliquefaciens* FZB42.

Genomic DNA of *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 was digested with the rare-cutter restriction endonuclease *Sfi*I and then separated by pulsed field gel electrophoresis (PFGE). The obtained macrorestriction profiles of *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 can be seen in panels A and B respectively. M, molecular mass marker (MidRange II PFG marker); bands from bottom to top 24,5/48,5/73/97/121,5/145,5/170/194/218,5/242,5/267 kb.

Suppression Subtractive Hybridization (SSH)

Suppression Subtractive Hybridization (SSH) was used as means of identifying extensive gene differentiation between *B. amyloliquefaciens* FZB42 and *B. subtilis* 168. At the time point that these experiments were performed, only preliminary data existed for the genome sequence of *B. amyloliquefaciens* FZB42 (see chapter 3.2 for the updated data), whereas the complete genome of *B. subtilis* 168 had already been published [7]. Therefore SSH provided us with a rapid but thorough first view of genetic variation between the strains, long before that was possible by direct comparison of both strains' complete sequences. For these experiments, *B. amyloliquefaciens* FZB42 was used as tester strain in order to rapidly detect its unique DNA sequences that are absent from *B. subtilis* 168 (driver strain).

Sixty-six clones were obtained by this approach. After sequencing analysis, the clones were aligned at nucleotide level to the known genome of *B. subtilis* 168. Three of these clones appeared twice in the screen and ten displayed more than 60% nucleotide homology to the driver strain. The SSH application can be thus considered as very successful, since 84% of the gained clones contained sequences with low (less than 60%) nucleotide homology to the driver strain. Thereby for the first time fifty-three DNA segments of various lengths present in *B. amyloliquefaciens* FZB42 but not in *B. subtilis* 168 were identified. Furthermore, their putative function was deduced by basic alignment search tool analysis (BLAST). The results are presented in table 6.

Most interestingly, eight clones showed high similarity to genes of nonribosomal peptide synthetases and polyketide synthases. In particular, clone cAK6 displayed 81% amino acid homology to MycC, involved in mycosubtilin biosynthesis, whereas clone cAK49 was 98% similar to ItuB, involved in iturin A biosynthesis. These findings suggest that the genome of *B. amyloliquefaciens* FZB42 may contain an operon for the nonribosomal biosynthesis of an iturin-like antibiotic. In contrast, such operon is not part of *B. subtilis* 168

genome. In this strain only the peptide sythetase operons encoding for surfactin and fengycin are present. Furthermore, six clones displayed similarity to polyketide synthases. For example, clones cAK24 and cAK48 were respectively 58% and 31% homologous, at amino acid level, to PksR and PksM, proteins that are encoded within the single polyketide synthase operon (*pksX*) in *B. subtilis* 168 (Table 6). Low similarities between the *pks* operon present in the driver strain and the six sequences from the tester strain indicated that *B. amyloliquefaciens* FZB42 might contain operon(s) responsible for polyketide biosynthesis that differ from the one present in *B. subtilis* 168.

Clone cAK30 displayed 99% similarity to MrsG, a protein probably involved in the immunity against the lantibiotic mersacidin [36] (see also 1.3.1.2). *mrsG* is transcribed from the same operon as *mrsF* and *mrsE*, in the mersacidin-producer Bacillus sp. strain HIL Y-85,54728. These genes encode an ABC transporter that could be involved in protection against the antibiotic [36]. The question that arose was whether *B. amyloliquefaciens* FZB42 possessed the whole *mrsFGE* operon or even the entire biosynthetic gene cluster of mersacidin; none of these genes are encoded in *B. subtilis* 168. For this purpose, primer walking was performed on the region neighbouring clone cAK30. Thereby the presence of the *mrsFGE* operon and of *mrsR2K2*, the two-component regulatory system that controls the operon's transcription, was demonstrated; the biosynthetic genes of mersacidin were not found in this genomic region.

A putative *IS3*-like transposase was identified in clone cAK2, indicating that horizontal gene transfer in *B. amyloliquefaciens* FZB42 might be achieved by transposases. Interestingly, such phenomena are not observed in *B. subtilis* 168 which does not contain any transposase in its genome. Moreover, two clones (cAK9 and cAK38) are similar to phage-related proteins. In addition to the phage-related protein, cAK38 harbours a hypothetical protein, conserved in *B. licheniformis* DSM13 and absent from the driver strain. The remaining clones exhibit similarity to proteins with various functions, such as transcriptional regulators (cAK10), thymidylate synthases (cAK11), membrane or cation efflux proteins (cAK17/cAK20) etc. Moreover, several obtained sequences, parts of ORFs or putative proteins, score the best homology to unknown proteins of *B. subtilis* 168. Considering that the nucleotide and amino acid homology is very low, it seems plausible that these proteins are new and non-existing in the already sequenced strains of the *Bacilli* family.

Since SSH is based exclusively on DNA similarity, it was only expected that some of the clones would contain non-coding regions. These regions are of high interest and can have regulatory function since RNA polymerase and various transcriptional regulators might bind

to them. Moreover, these can be regions where non-coding RNAs are located. Even though such clones are considered as positive attempts towards identifying sequence variations between *B. amyloliquefaciens* FZB42 and the driver strain, since they exhibit low nucleotide similarity to *B. subtilis* 168, they are not included in the table 6; 39,6% of the positive clones mostly contained non-coding region.

Table 6: FZB42 strain-specific SSH clones

Clone	**Size (bp)**	**Nucleotide similarity to *B. subtilis* 168**	**Putative function/accession number**	**Identities** (aa level)	**Organism**
cAK1	258	n.s.	Predicted hydrolase (HAD superfamily)/AAU39182	57/85 (67%)	*B. licheniformis* DSM 13
cAK2	211	n.s.	Putative transposase part of *IS* element/YP_080245	37/68 (54%)	*B. licheniformis* ATCC 14580
cAK4	419	171/419 (40%)	Tmk, Thymidylate kinase/AAU39021	55/132 (41%)	*B. licheniformis* DSM 13
cAK5	314	n.s.	Acetyltransferase/ZP_00962455	20/38 (52%)	*Sulfitobacter* sp. NAS-14.1
cAK6	120	n.s.	MycC, Mycosubtilin synthetase C/AAF0879	31/38 (81%)	*B. subtilis* ATCC 6633
cAK7	191	n.s.	LicR, transcriptional regulator/AAU42903	18/31 (58%)	*B. licheniformis* DSM 13
cAK9	437	259/437 (59%)	YobO, similar to phage-related pre-neck appendage protein/CAB13795	111/137 (81%)	*B. subtilis* 168
cAK10	126	n.s	Negative transcriptional regulator/ZP_01185138	28/42 (66%)	*B. weihenstephanensis* KBAB4
cAK11	377	n.s.	Thymidylate synthase/ZP_00063805	59/96 (61%)	*Leuconostoc mesenteroides* ATCC 8293
cAK12	526	109/526 (20%)	YqeD, hypothetical protein/CAB14513	60/89 (67%)	*B. subtilis* 168
cAK13	491	n.s.	Florfenicol-chloramphenicol resistance protein/NP_899167	43/136 (31%)	*Staphylococcus sciuri*
cAK14	392	n.s.	ykpA, hypothetical protein/CAB13316	31/103 (30%)	*B. subtilis* 168
cAK17	402	n.s	Putative membrane protein/ZP_00231274	46/126 (36%)	*Listeria monocytogenes* str. 4b 7858

Clone	**Size (bp)**	**Nucleotide similarity to *B. subtilis* 168**	**Putative function/accession number**	**Identities** (aa level)	**Organism**
cAK18	322	141/322 (43%)	TuaB, Teichuronic acid biosynthesis protein/O32273	87/107 (81%)	*B. subtilis* 168
cAK20	262	n.s.	cation efflux family protein/ZP_00230783	57/87 (65%)	*Listeria monocytogenes* str. 4b H7858
cAK22	310	n.s.	TreA, phospho-alpha-(1,1)-glucosidasephospho-alpha-(1,1)-glucosidase/CAA91015	58/91 (63%)	*B. subtilis* 168
cAK23	383	n.s.	YyaL, hypothetical protein/CAB16119	56/93 (60%)	*B. subtilis* 168
cAK24	380	106/380 (28%)	PksR, polyketide synthase/CAB13606	72/123 (58%)	*B. subtilis* 168
cAK26	339	n.s.	conserved hypothetical cytosolic protein/ZP_01182157	70/72 (97%)	*B. cereus subsp. cytotoxis* NVH 391-98
cAK27	296	n.s.	PksN, polyketide synthase/CAB13604	24/39 (61%)	*B. subtilis* 168
cAK30	378	n.s.	MrsG, putative ABC-transporter integral membrane protein/AB60256	104/105 (99%)	*Bacillus* sp. HIL-Y85/54728
cAK36	351	160/351 (45%)	YbdN, hypothetical protein/NP_388086	78/113 (69%)	*B. subtilis* 168
cAK38	496	n.s.	Phage-like element PBSX protein xkdC/P39782-hypothetical protein BLi01445/AAU40348	32/55 (58%)-24/46 (52%)	*B. subtilis* 168-*B. licheniformis* ATCC 14580
cAK39	353	n.s.	hypothetical protein, putative carbohydrate esterase/AAU41672	108/119 (90%)	*B. licheniformis* ATCC 14580
cAK47	423	174/423 (41%)	YwbD, hypothetical protein/NP_391715	77/129 (59%)	*B. subtilis* 168
cAK48	261	n.s.	PksM, polyketide synthase/P40872	27/87 (31%)	*B. subtilis* 168
cAK49	390	n.s.	ItuB, iturin A synthetase B/BAB69699	127/129 (98%)	*B. subtilis* RB14
cAK53	344	n.s.	PksL, polyketide synthase/Q05470	21/54 (38%)	*B. subtilis* 168
cAK54	346	n.s.	PksN, polyketide synthase/BG12652	27/102 (26%)	*B. subtilis* 168
cAK56	493	n.s	YwmC, hypothetical protein / AB03680	62/104 (59%)	*B. subtilis* 168

Clone	Size (bp)	Nucleotide similarity to *B. subtilis* 168	Putative function/accession number	Identities(aa level)	Organism
cAK58	443	n.s.	Hypothetical cytosolic protein/ZP_00740431	66/131 (50%)	*B. thuringiensis serovar israelensis* ATCC 35646
cAK59	497	183/497 (37%)	PksE, polyketide synthesis/ P_389593	108/155 (69%)	*B. subtilis* 168

Clones obtained by SSH containing sequences of *B. amyloliquefaciens* FZB42 that exhibit less than 60% nucleotide similarity to *B. subtilis* 168. The sequence's size and its exact overall nucleotide similarity to the driver strain are indicated; less than 20% similarity is considered non significant (n.s.). The putative functions of the DNA segments are presented, as derived by BLASTX alignment. Similarities on amino acid level are indicated for the aligned part of the sequences. Clones that mostly comprise of non-coding regions are not included in the table.

When later the sequencing task of *B. amyloliquefaciens* FZB42 advanced, the identification of the regions flanking the sequences obtained by SSH was possible. Moreover, in many cases sequencing of the strain was directed by the clones obtained by SSH. Our focus was particularly drawn on those regions that contained genes coding for nonribosomal peptide synthetases and polyketide synthases and further attempts to obtain detailed sequence information of these regions were conducted (mainly by primer walking).

Sequence analysis of these regions revealed the presence of three distinct *pks* gene clusters (*pks1*-72442 bp, *pks2*-54350 bp, *pks3*-69548 bp) (Fig. 13). Clone cAK24 belongs to *pks1* gene cluster whereas clones cAK27, cAK53, cAK54 belong to *pks2* gene cluster. *pks3* polyketide synthase includes sequences obtained by clones cAK48 and cAK59. From these three gene clusters only the *pks1* system from *B. amyloliquefaciens* FZB42 is similar to the *pksX* operon present in *B. subtilis* 168; a strain unable to synthesize polyketides due to a mutation in the *sfp* gene [174]. The polyketide synthases *pks2* and *pks3* are novel gene clusters. Various types of mass spectrometry, mutant construction and biological tests were used for verifying the functionality of these gene clusters. These experiments were in majority performed by Xiao-Hua Chen and can been seen in detail in [197].

Finally, clones cAK6 and cAK49 were found to be part of a ~37 kb operon that showed homology to the iturin A operon of *B. subtilis* RB14. Further experiments were performed for the characterization of this gene cluster (see 3.3 and 3.4).

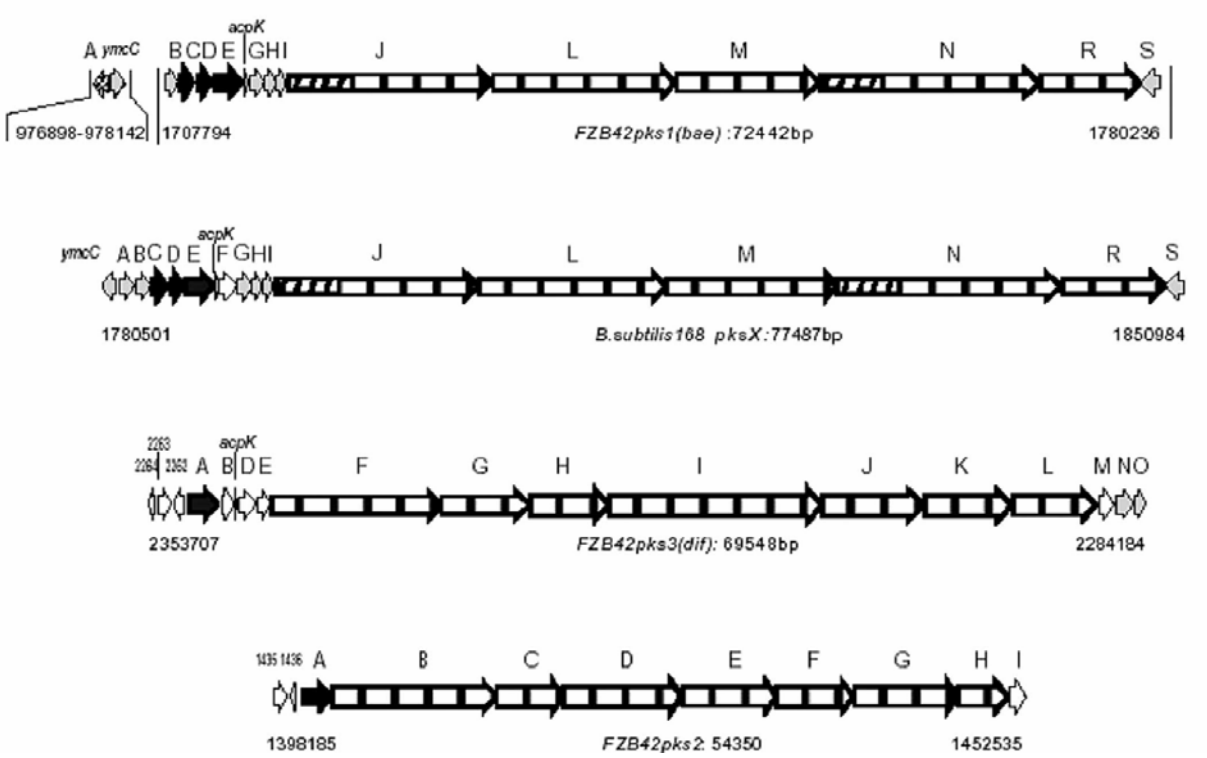

Figure 13: Organization of the gene clusters involved in polyketide biosynthesis in *B. amyloliquefaciens* FZB42 (*pks1*, *pks2*, *pks3*) and *B. subtilis* 168 (*pksX*).

The size and location of the three polyketide gene clusters in the genome of *B.amyloliquefaciens* FZB42 are shown. Filled and bold arrows indicate discrete AT domains and modular PKS respectively. NRPS portions occurring in hybrid NRPS-PKS enzymes are shaded. Gene clusters *pks1*, *pks2* and *pks3* are responsible for the biosynthesis of bacillaene, macrolactin and difficidin/oxydifficidin respectively. AT, acyltransferase [197];(K.Schneider and Xiao-Hua Chen, unpublished results).

Sequence analysis of B. amyloliquefaciens FZB42 genome

Sequencing of the *B. amyloliquefaciens* FZB42 genome was performed as a joint collaboration between our laboratory and the GenoMik Network in Göttingen. The major part of the work and the co-ordination of the whole process were done by Xiao-Hua Chen and myself. Shotgun and fosmid library approaches, primer walking and multiplex PCR were used in order to obtain the complete genetic information encoded in the chromosome of *B. amyloliquefaciens* FZB42 (for more details see Materials and Methods, chapter 2.6). Sequencing of the whole genome of the strain has been completed whereas the second round of annotation using the GeneSOAP program is currently in process (performed by Xiao-Hua Chen).

The genome of *B. amyloliquefaciens* FZB42 is a single chromosome consisting of 3916 kb. The G+C content is about 46% and it contains 11 rRNA clusters. Even though the genome annotation is not yet completed, preliminary data revealed the presence of 3931 genes. BLAST comparison with SUBTILIST (a database containing all annotated genes of *B. subtilis* 168) showed that around 80% of the genes (3125) encoded by *B. amyloliquefaciens* FZB42 are more than 50% homologous at amino acid level to genes of *B. subtilis* 168. However, more than 200 of them are located in regions different than in the *B. subtilis* 168 genome, possibly due to rearrangement events that occurred during evolution of the two genomes. Moreover, co-linear regions exhibiting high similarity to the *B. subtilis* genome, which are then interrupted by regions of variable length containing genes unique for FZB42 were also detected. The unique genes of *B. amyloliquefaciens* FZB42 were found distributed in at least 14 DNA islands and islets around the whole genome. In contrast to *B. subtilis* 168, horizontal gene transfer is achieved not only by phages but also by different types of *IS* elements which are present in different copy numbers within the FZB42 genome (Table 7). The circular map of the chromosome of *B. amyloliquefaciens* FZB42 demonstrating some of its basic characteristics as well as an illustrated comparison to the *B. subtilis* 168 genome are presented in figure 14.

Table 7: Transposases present in *B.amyloliquefaciens* FZB42 genome

Transposase type	**Position**	**Size (aa)**	**Identities (aa level)**	**Organism/accession number**
IS231-related transposase	1999	96	65/80 (81%)	*B. weihenstephanensis* KBAB4 /ZP_01184801
IS3Bli1	2452	405	208/405 (51%)	*B. licheniformis* ATCC 14580/AF459921
transposase	3456	199	87/147 (59%)	*Staphylococcus epidermidis* ATCC 12228/AAO03698
IS3Bli1	3640	405	208/405 (51%)	*B. licheniformis* ATCC 14580/AF459921

IS3Bli1	3764	405	208/405 (51%)	*B. licheniformis* ATCC 14580/AF459921

The positions of the transposases on the genome of *B. amyloliquefaciens* FZB42 are indicated in kb. Similarities are indicated on amino acid level for the aligned part of the sequences, as derived by BLASTX alignment.

Striking is the presence of eight gene clusters encoding for secondary metabolites. In addition to *srf*, *fen* and *pks1* (*bae*) operons that are responsible for the synthesis of surfactin, fengycin and bacillaene and are also present in the *B. subtilis* genome [7, 197], *B. amyloliquefaciens* FZB42 contains several additional gene clusters coding for peptide/polyketide antibiotics. *bmy*, *pks2* and *pks3* operons are involved in bacillomycin D, macrolactin, difficidin / oxydifficidin polyketide synthesis and further information about them can be found in other sections (see 3.3 and 3.1) [196, 197]; K.Schneider and Xiao-Hua Chen, unpublished results). Moreover, *B. amyloliquefaciens* FZB42 genome contains the *bac* operon responsible for the biosynthesis of the dipeptide bacilysin [217]. This antibiotic consists of an L-alanine at the N-terminus and a unusual amino acid, L-anticapsin, at the C-terminus and displays antibacterial activity [218]. The unusual epoxy-modified amino acid anticapsin is probably generated through the action of a prephenate dehydratase and an aminotransferase encoded by *bacA* and *ywfG* respectively, as a branching off from the prephenate of the aromatic amino acid pathway [219]. Additionally, the *dhb* operon is present in the genome of *B. amyloliquefaciens* FZB42 (see also section 3.5). The *dhbACEBF* operon is involved in the synthesis of 2,3-dihydroxybenzoate (DHB) as well as its modification and esterification to the iron siderophore bacillibactin [78] that enables microorganisms to efficiently scavenge iron [220, 221]. DhbE is a stand-alone adenylation domain that activates DHB in an ATP-dependent reaction. The activated DHB is subsequently transferred to the free thiol group of the co-factor phosphopantetheine of the bifunctional isochorismate lyase/aryl carrier protein DhbB. The third synthetase, DhbF, is a dimodular nonribosomal peptide synthetase that specifically adenylates threonine (and to a lesser extent glycine) as well as covalently loads both amino acids onto the corresponding peptidyl carrier domains [78].

The eight gene clusters encoding peptide/polyketide antibiotics and a siderophore represent about 8% of the total genome and control synthesis of bioactive compounds by processes based on nonconventional translation. Interestingly, three of these gene clusters

(*bmy*, *fen*, *pks1*) are localized at the replication terminus, indicating that this region is probably more susceptible to horizontal gene transfer.

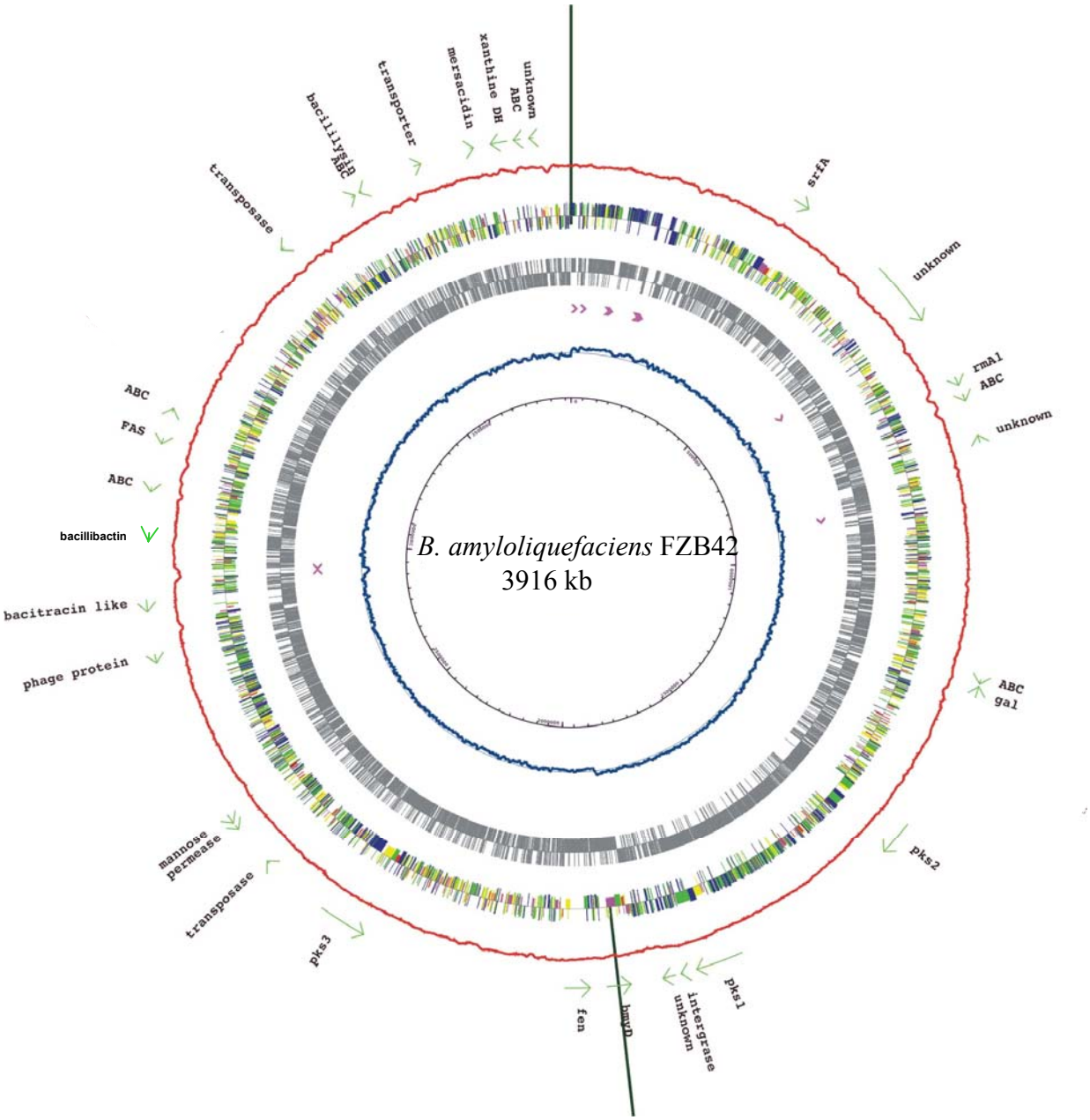

Figure 14: Whole genome map of *B. amyloliquefaciens* FZB42 (kindly provided by Xiao-Hua Chen).

The scale on the inner circle shows coordinates, in bp. The blue circle is the GC-skew, which is correlated with the replication start point. Total 11 sets of rRNA are shown by the pink arrows. The grey circle represents all 3931 genes detected in FZB42. The colored circle displays the distribution of the homologous genes with *B. subtilis* 168, detected by BLASTX and BLASTP comparison. Around 80% of the total genes bear similarity of more than 50%. The color code indicates identities greater than 90%, 80%, 70%, 60%, 50%. The green arrows within the outer circle indicate genes which are unique in the FZB42 and might contribute to the plant growth promotion or involved in horizontal

gene transfer. The orange GC-content circle shows consistency between horizontal gene transfer and low GC-content.

Lipopeptides produced by B. amyloliquefaciens strain FZB42

Organization of nonribosomal peptide synthetases on the FZB42 chromosome

The SSH experiments revealed the presence of nonribosomal peptide synthetases and polyketide synthases in the genome of *B. amyloliquefaciens* FZB42. With the later acquirement of the first assembly of the organism's genome sequence, it became clear that *B. amyloliquefaciens* FZB42 encodes operons *srf*, *fen* and *bmy* which are responsible for the synthesis of three lipopeptides: surfactin, fengycin and bacillomycin D. This was the first report revealing the coding sequence of bacillomycin D and evidence for its functionality was provided by MALDI-TOF MS analysis (see 3.3.2). *B. amyloliquefaciens* FZB42 also encodes three polyketide synthases [197].

The cluster of *bmy* is a FZB42-specific DNA island comprising of 4 genes (*bmyD*, *bmyA*, *bmyB* and *bmyC*) and is close to the *fen* operon on the chromosome. Regions flanking the large gene cluster are characterized by DNA rearrangements joining the antibiotic DNA islands with sequences originally present in different regions of the *B. subtilis* 168 chromosome (Fig. 15A). In particular, right from the 37.2 kb *bmy* gene cluster two rearranged clusters are situated: *yxjCDEF* and *bioIBDFAW*, that are present in *B. subtilis* 168 at positions 3999 to 4002 kb and 3088 to 3094 kb, respectively. On the left site, regions located in *B. subtilis* 168 at positions 1910 to 1943 kb (*yndG*, *bglC*, *ynfJ* and *xynD*) were detected. Interestingly, the *bmy* operon is inserted at the same position as the iturin A gene cluster in *B. subtilis* RB14. This "coincidence" and the high homology between bacillomycin D and iturin A made us initially assume (before the MS results; see 3.3.2) that FZB42 could encode the *itu* operon.

The *fen* five-gene cluster (*fenA-E*) present in *B. amyloliquefaciens* FZB42 is related to the *pps* operon in *B. subtilis* 168 and is situated at the same locus as this, about 25 kb distant from the *bmy* operon (Fig. 15A). Amino acid similarity between the operons of the two strains is between 60% (*fenB*) and 65% (*fenC*). The nonribosomal peptide synthetase directing biosynthesis of fengycin is also present in *B. subtilis* strains F29-3 [222] and A1/3 [140]. However, in *B. subtilis* ATCC 6633, mycosubtilin, another iturin-like lipopeptide, is found at the same genetic locus as fengycin [63]. There seems to be an exchange of these two operons

between different strains, implying high degree of genetic flexibility in this region. Moreover, this suggests that additional NRPS operons might be integrated in this area either as insertions or as substitutions of already existing NRPS operons.

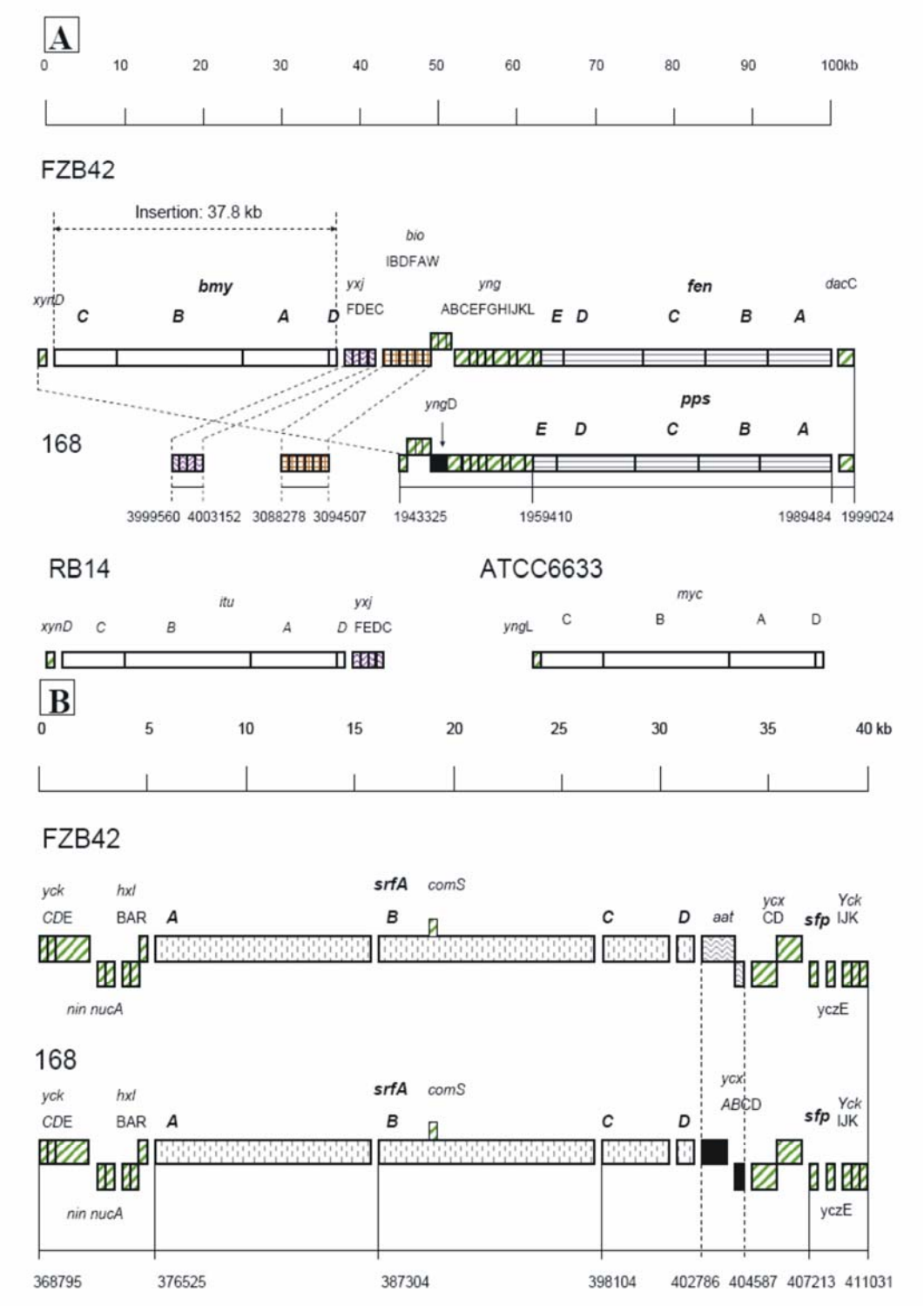

Figure 15: Organisation of the bacillomycin D, fengycin and surfactin operons in *B. amyloliquefaciens* FZB42.

In panel A is presented the chromosomal organization of bacillomycin D and fengycin operons in *B. amyloliquefaciens* FZB42 as well as of operons producing highly homologous antibiotics in other

Bacilli strains. In panel B is compared the organisation of the surfactin operon in *B. amyloliquefaciens* FZB42 with that of *B. subtilis* 168. The intersecting dotted lines indicate events of insertion or rearrangement in FZB42 compared to the respective *B. subtilis* 168 genome region, whereas full lines demonstrate conservation of gene order between the two strains. Black-filled boxes indicate genes present in *B. subtilis* 168 but absent from the respective genome region of *B. amyloliquefaciens* FZB42. The organisation of the homologous gene clusters in *B. subtilis* 168 (fengycin operon-*pps*, surfactin-*srf*), *B. subtilis* RB14 (iturin A operon-*itu*) and *B. subtilis* ATCC6633 (mycosubtilin operon-*myc*) are presented according to [7, 49, 63, 101].

The 26.5 kb *srf* operon present in *B. amyloliquefaciens* FZB42 genome is organized in a similar manner as in *B. subtilis* 168 (see introduction). The corresponding genes of these two strains exhibited similarity between 72% (*srfAA*) and 83% (*srfAC*) on amino acid level. The genes present at the left flanking region of the operon are *hxlBAR*, like in the case of *B. subtilis* 168. However, on the right flank of *srfAD*, the *B. subtilis* 168 *ycxAB* are substituted by two ORFs with unknown function (Fig. 15B). Moreover, the *comS* gene, encoding a competence signal molecule, is embedded in the *srfAB* sequence, as already detected for various *Bacillus* strains and displays 63% homology to its orthologue in *B. subtilis* 168. The 4'-phosphopantetheinyl transferase Sfp is located 4kb downstream and exhibits 70% amino acid homology to the one encoded by strain 168.

Functional analysis of lipopeptide production in *B. amyloliquefaciens* FZB42

MS identification of the lipopeptide products of *B. amyloliquefaciens* FZB42

In order to check the functionality of the lipopeptide-encoding gene clusters, culture filtrate extracts and whole cells were investigated by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS). The spectra obtained by both methods were found identical and revealed the presence of three different lipopeptides, as three groups of mass peaks were detected (Fig. 16). A summary of their mass numbers is presented in table 8.

Surfactins and fengycins have been identified by comparing their mass data with those previously obtained by MS analysis of numerous *B. subtilis* strains [211]. Moreover, *B. amyloliquefaciens* FZB42 produces surfactins and fengycins with fatty acid side chains of 13 to 15 and 15 to 17 carbon atoms, respectively.

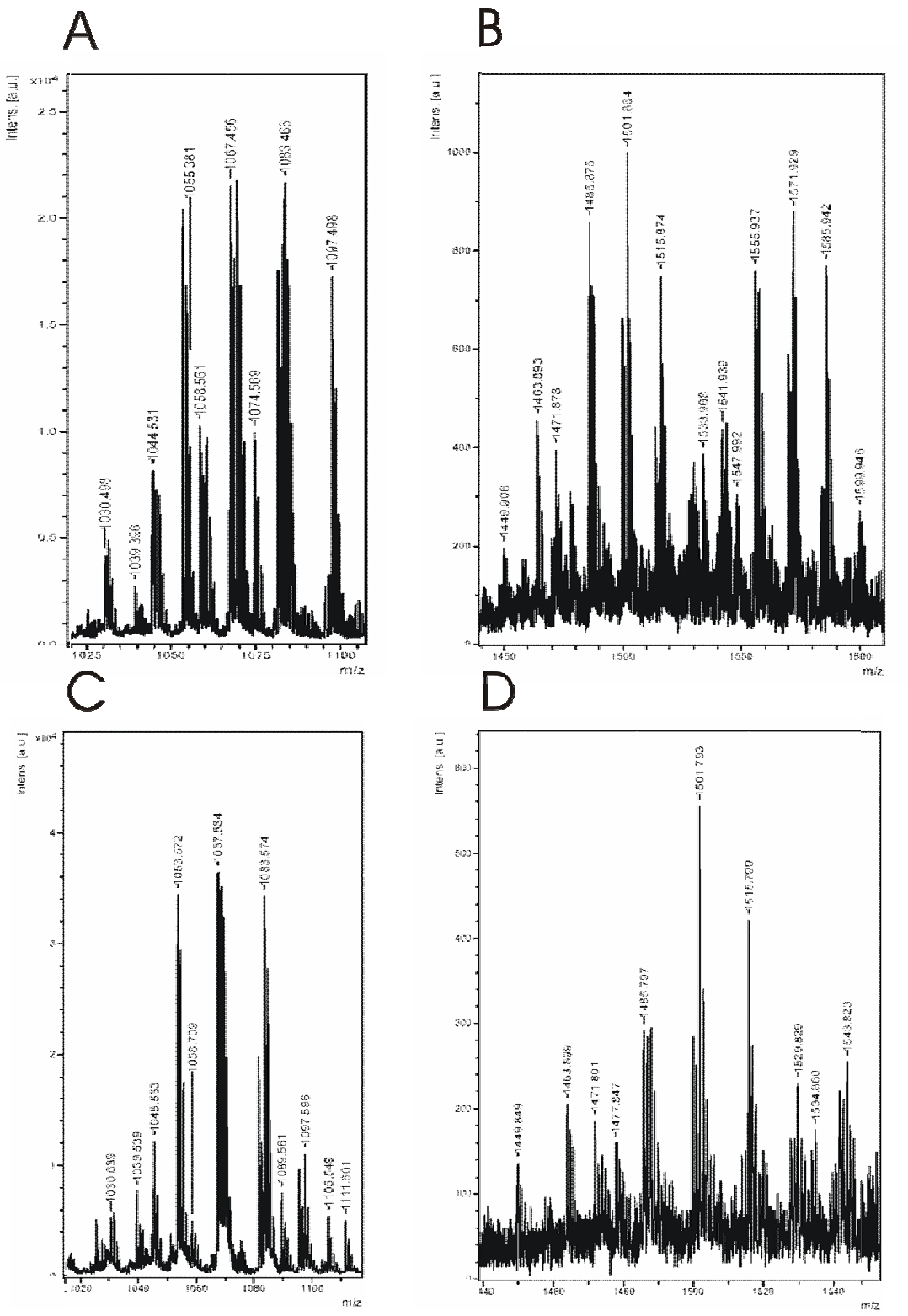

Figure 16: MALDI-TOF MS analysis of lipopeptides produced by *B. amyloliquefaciens* FZB42 (performed in collaboration with Dr. J. Vater).

Detection of surfactin (A), bacillomycin D (A) and fengycin (B) mass peaks in culture filtrate extracts prepared from *B. amyloliquefaciens* FZB42 grown for 24 hours in Landy medium. Spectra of intact whole cells grown on Landy medium agar plates; detection of surfactin (C), bacillomycin D (C) and fengycin (D) mass peaks. See table 8 for peak identification.

Table 8: Lipopeptide products of *B. amyloliquefaciens* FZB42 detected by MALDI-TOF mass spectrometry

Observed peaks m/z	**Assignment**
Surfactin	
1030.8*/1046.8	C13-surfactin $[M + Na, K]^+$
1044.8*/1060.8	C14-surfactin $[M + Na, K]^+$
1058.8*/1074.8*	C15-surfactin $[M + Na, K]^+$
Bacillomycin D	
1031.7/1053.7*/1069.7	C14-bacillomycin D $[M + H, Na, K]^+$
1045.7*/1067.7*/1083.7*	C15-bacillomycin D $[M + H, Na, K]^+$
1059.7/1081.7/1097.7*	C16-bacillomycin D $[M + H, Na, K]^+$
1095.7/1111.7*	C17-bacillomycin D $[M + Na, K]^+$
Fengycin	
1449.9*/1471.9*/1487.9	Ala-6-C15-fengycin $[M + H, Na, K]^+$
1463.9*/1485.9*/1501.9*	Ala-6-C16-fengycin $[M + H, Na, K]^+$
1477.9*/1499.9/1515.9*	Ala-6-C17-fengycin $[M + H, Na, K]^+$
1491.8/1513.9/1529.9*	Val-6-C16-fengycin $[M + H, Na, K]^+$
1505.8/1527.8/1543.8*	Val-6-C17-fengycin $[M + H, Na, K]^+$

Peaks indicated in figure 16 are marked by an asterisk.

Furthermore, postsource decay (PSD) MALDI-TOF MS revealed that the third produced lipopeptide is bacillomycin D. In the mass spectra obtained for whole cells and surface extracts the mass peaks of sodium and potassium adducts dominate, whereas the protonated species appear with minor intensities. However, the protonated species are preferred for PSD MALDI-TOF MS mediated sequence analysis because they decompose into fragments more readily than the alkali adducts. For example, the lipopeptide with a mass number of *m/z* 1031.5 produced by *B. amyloliquefaciens* FZB42 was identified as the protonated form of a bacillomycin D isoform with a fatty acid chain of 14 carbon atoms. Its sequence was determined from a series of b_{n1}-, Y_n‘‘$(-H_2O)$-, and proline-directed b_{n2}- fragment ions (Fig. 17). The peptide ring of this bacillomycin D was cleaved both at the peptide bond between its amino fatty acid residue and threonine at position 7 as well as at the N terminus of proline-4. In the first case a series of b_{n1}- and Y_n‘‘$(-H_2O)$-fragment ions were detected. In addition, b_{n2}-ions of highly intensity were observed. Based on all these data, this lipopeptide was identified

as the protonated form of a C14- bacillomycin D. The obtained sequence was corroborated by b_{n1}- ions of dipeptide fragments at *m/z* values of 171.4, 212.3, 226.8 and 268.4, indicating nearest-neighbour relationships in the peptide ring of this lipopeptide for ES(-H_2O), NP, PE and NY respectively.

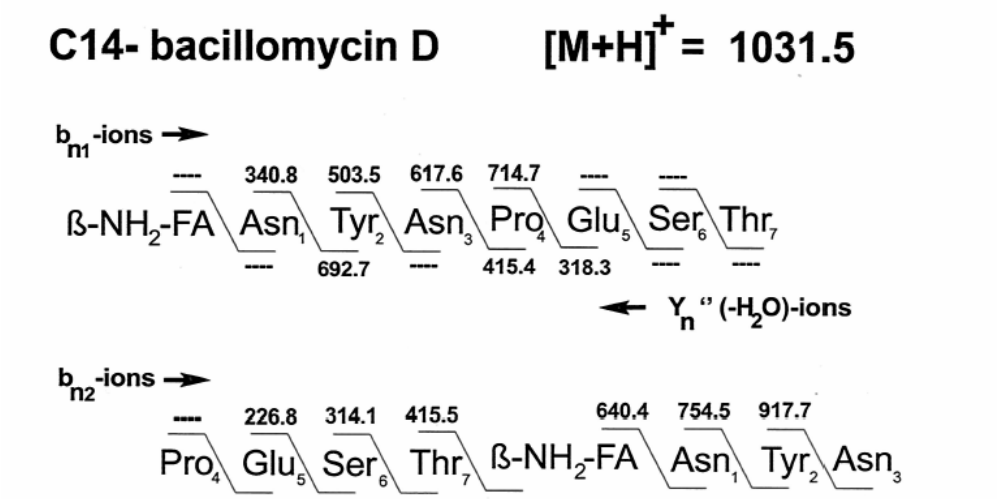

Figure 17: In situ structural analysis of the lipopeptide product of *B. amyloliquefaciens* FZB42 with mass number m/z 1031.5 by PSD-MALDI-TOF-MS (performed in collaboration with Dr. J. Vater).

The structure was derived from a series of f b_{n1}-, Y_n''(-H_2O)-, and proline-directed b_{n2}- fragment ions. FA, fatty acid

Production of lipopeptides along the growth curve

The production of lipopeptides during growth in liquid cultures was monitored by MALDI-TOF MS. This type of spectrometry is not suitable for determining their exact concentrations, mainly because of inhomogeneities in the analytical distribution of the crystalline matrix and different ionization efficiencies of the investigated compounds. However, the relative quantities of the three antibiotics at different points of growth can be estimated by calculating the ratios of the intensity values of the peaks corresponding to the various antibiotics.

For this reason, culture filtrate extracts from *B. amyloliquefaciens* FZB42 grown in ACS medium at 37°C under vigorous shaking for 10, 20, 40 and 60 hours were subjected to MALDI-TOF MS analysis. Thereby it was shown that surfactins and bacillomycins were present at similar intensities but reached their zenith in different stages of growth. Maximum levels of surfactin appeared in samples obtained after 10 to 40 hours of growth, whereas after 60 hours production dropped. On the other hand, bacillomycin D accumulated after 40 to 60

hours of growth. The time course of fengycin resembled that of surfactin, but its intensity was clearly lower compared to the other lipopeptides (Table 9). The same pattern of lipopeptide production was obtained also when *B. amyloliquefaciens* FZB42 was grown in Landy medium and samples were drawn at 12, 24 and 48 hours. In this case, surfactin peaked already at 12 hours, whereas bacillomycin D at 24 hours. The intensities of fengycin peaks maintained rather low all along growth.

Table 9: Time-dependent production of lipopeptides by *B. amyloliquefaciens* FZB42 grown in ACS medium

Lipopeptide	***m/z***	**Species**	**Intensity at:**			
			10 hours	**20 hours**	**40 hours**	**60 hours**
Surfactin	1044.8	C14 $[M + Na]^+$	15500	10300	10100	2500
	1058.8	C15 $[M + Na]^+$	16500	11700	10000	2400
Bacillomycin	1053.7	C14 $[M + Na]^+$	3800	6600	11700	12300
D	1067.7	C15 $[M + Na]^+$	2500	2500	7600	7000
Fengycin	1485.9	C16 $[M + Na]^+$	3500	2950	3630	470
	1449.9	C17 $[M + Na]^+$	1800	1810	2510	170

Lipopeptide production was monitored by MALDI-TOF mass spectrometry. Culture filtrate extracts were prepared after 10, 20, 40 and 60 hours of growth in ACS medium. Only the main peaks (*m/z*) of each lipopeptide were selected for analysis.

Lipopeptide deficient mutants

In order to confirm that *bmy* and *fen* operons are directing bacillomycin D and fengycin biosynthesis, disruption mutants at the *bmyA* and *fenA* genes were created. *B. amyloliquefaciens* FZB42 mutant strains were generated via double-crossover recombination [199] according to a modified protocol that has been originally developed for *B. subtilis* (see materials and methods).

In detail, a 1,2 kb fragment of *bmyA* was amplified by PCR using primers bmyAa and bmyAb (Table 4) and cloned into pGEM-T. After digestion with the restriction endonuclease *Ava*I, an erythromycin cassette was inserted inside the bmyA-fragment, resulting in plasmid pAK2 which was transformed in *B. amyloliquefaciens* FZB42. Disruption of *bmyA* in the resistant colonies was demonstrated by PCR with primers bmyAa and bmyAb and by

Southern hybridization. Disruption of *fenA* was achieved in a similar manner by insertion of a chloramphenicol cassette in a fragment obtained by PCR with primers fenAa and fenAb and digested with *Hind*III and *Kpn*I. Correct integration of the antibiotic cassette was verified by PCR and Southern hybridization. Furthermore, a double mutant of both *bmyA* and *fenA* genes was created.

Analysis of the mutant strains by MALDI-TOF MS, verified that strains Δ*bmyA*::Em^r and Δ*fenA*::Cm^r failed to produce bacillomycin D and fengycin respectively, since the corresponding groups of mass peaks were absent (Fig. 18). Moreover, strain AK3 (Δ*bmyA*::Em^r Δ*fenA*::Cm^r) was deficient in the production of both lipopeptides. Disruption of *srfAA* was performed by Xiao-Hua Chen and resulted in the strain's inability to produce surfactin (data not shown). Consequently, the NRPS gene clusters that were identified on the chromosome are responsible for the biosynthesis of the respective lipopeptides in *B. amyloliquefaciens* FZB42.

Biological activity of wild type and mutant strains

B. amyloliquefaciens FZB42 stimulates plant growth and suppresses plant pathogenic organisms [5, 6], via mechanisms that have not been yet fully characterized. In order to check if the nonribosomal synthesized peptides produced by this bacterium contribute to its biocontrol capacity, the wild type and mutant strains deficient in biosynthesis of lipopeptides were assayed for their biological activities.

For this purpose, growth of various phytopathogenic fungi in the presence of *B. amyloliquefaciens* FZB42 was investigated. Strain FZB42 was shown to inhibit the growth of *Fusarium oxysporum*, *Gaeumannomyces graminis*, *Rhizoctonia solani*, *Alternaria alternate* and *Pythium aphanidermatum*. Moreover, strain AK1 (Δ*bmyA*::Em^r) suppressed growth of all fungi at a smaller extent, suggesting that bacillomycin D contributes to the antifungal activity of *B. amyloliquefaciens* FZB42. Interestingly, even though growth inhibition in the presence of strain AK2 (Δ*fenA*::Cm^r) was comparable to the one caused by the wild type, no inhibition was observed in the double mutant strain AK3 (Δ*bmyA*::Em^r Δ*fenA*::Cm^r) (Fig. 19A). The fact that fungi could grow uninfluenced by the presence of strain AK3 indicates a synergistic action of bacillomycin D and fengycin, against the target microorganism. The surfactin deficient mutant strain Δ*srfAA*::Em^r, provided by Xiao-Hua Chen, still retained its antifungal properties.

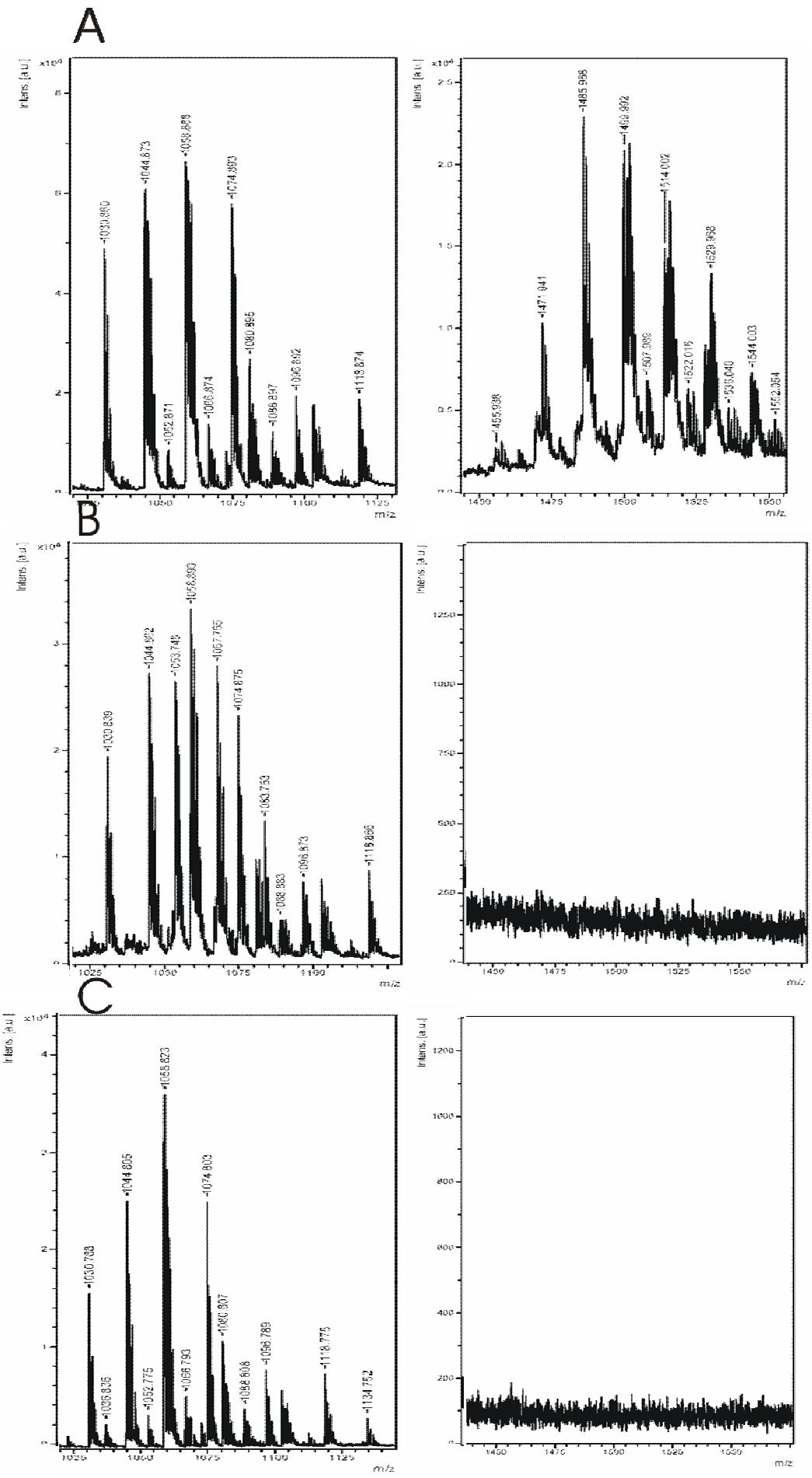
A
B
C
Intens. [a.u.]
m/z

Figure 18: MALDI-TOF MS analysis of mutant strains in nonribosomal peptide synthetases (performed in collaboration with Dr. J. Vater).

Spectra of culture filtrate extracts prepared from strains grown for 24 hours in Landy medium. A) Bacillomycin D is not produced by mutant AK1 (Δ*bmyA*::Em^r), whereas surfactin and fengycin are. B) Strain AK2 (Δ*fenA*::Cm^r) is deficient in fengycin production while (C) strain AK3 (Δ*bmyA*::Em^r Δ*fenA*::Cm^r) is deficient in both bacillomycin D and fengycin production.

In parallel, it was observed that *B. amyloliquefaciens* FZB42 suppresses the growth of *B. megaterium*. This inhibitory activity was also shared by each one of the strains deficient in lipopeptide biosynthesis (Fig. 19B). Consequently, the antibacterial properties of strain FZB42 are driven by some antibiotic(s) produced by the strain other than nonribosomal peptides [197, 223].

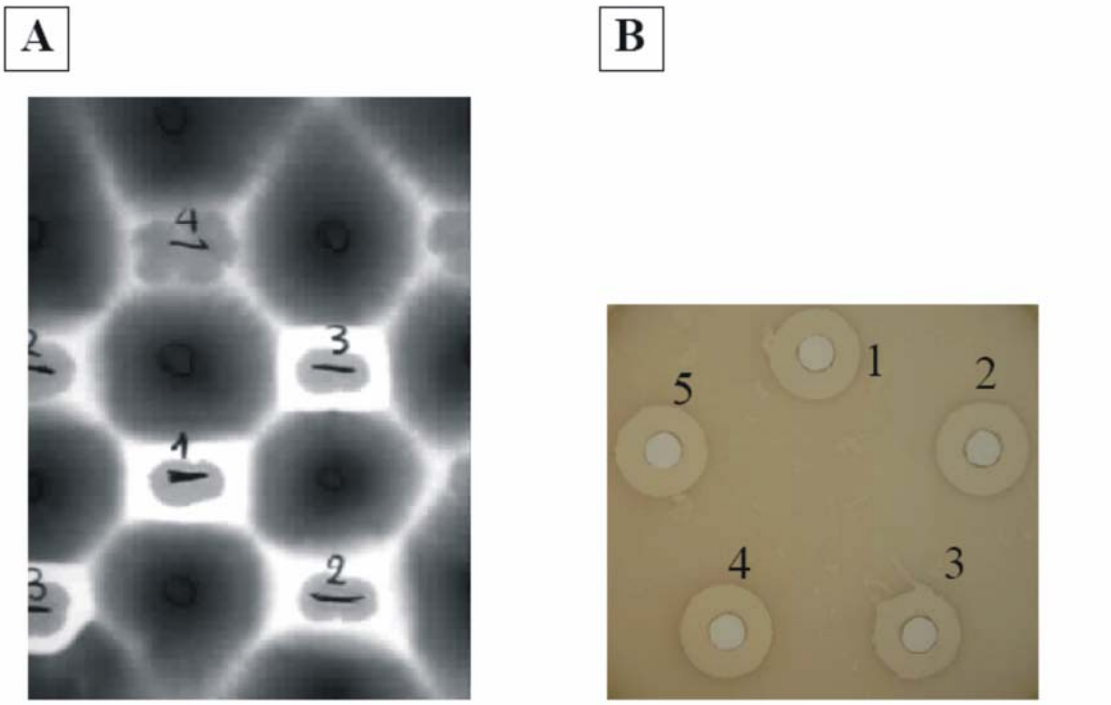

Figure 19: Biological activity of *B. amyloliquefaciens* FZB42 and lipopeptide deficient mutant strains.

A) Growth of *Fusarium oxysporum* f.sp.*cucumerinum* DSMZ 62313 in the presence of *B. amyloliquefaciens* FZB42 and lipopeptide deficient mutants. B) Growth of *B. amyloliquefaciens* FZB42 and lipopeptide deficient mutant strains on *B. megaterium* lawn. (1) wild type strain, (2) AK1 (Δ*bmyA*::Em^r), (3) AK2 (Δ*fenA*::Cm^r), (4) AK3 (Δ*bmyA*::Em^r Δ*fenA*::Cm^r) and (5) CH1 (Δ*srfAA*::Em^r).

Analysis of functional domains in *bmy* operon

Sequence analysis of the *bmy* operon revealed the presence of a cluster of four ORFs designated *bmyD*, *bmyA*, *bmyB* and *bmyC*, respectively (Fig. 20). *bmyD* encodes a putative malonyl coenzyme A transacylase and displays strong similarity to FabD, that is involved in fatty acid synthesis. Moreover, the protein is 98% and 80% identical to ItuD and FenF, respectively (ItuD and FenF are the first proteins participating in iturin and mycosubtilin biosynthesis[63, 101]. The ORFs encoding BmyA (3,982 aa), BmyB (5,633 aa) and BmyC (2,619 aa) show high similarity to members of the nonribosomal peptide synthetase family and display the ordered assembly of conserved condensation, adenylation and thiolation domains characteristic for such multienzymes (see introduction). As shown in figure 20, seven amino acid activating modules can be distinguished: one in BmyA (A1), four in BmyB (B1, B2, B3 and B4) and two in BmyC (C1, C2). Modules B1, B2 and C1 contain an epimerization domain, indicating that the activated amino acids are converted into D-configuration. The number of modules corresponds to the number of incorporated amino acids while the location of epimerization domains within the peptide synthetase agrees with the position of D-configurated amino acids in the peptide moiety of bacillomycin D. The last domain of this multienzyme system is a thioesterase domain, which is presumably required for release and cyclization of the synthesized lipopeptide. The organisation of this nonribosomal peptide synthetase is similar to that already described for the closely related lipopeptides iturin A and mycosubtilin (see figure 9C) [63, 101]. Analogously bacillomycin D is also synthesized according to the multicarrier thiotemplate mechanism (see 1.3.2.1).

Furthermore, the adenylation domains of bacillomycin D were compared to those of iturin A and mycosubtilin (Table 10). In the first case, more than 97% amino acid homology was observed within the first three modules of these synthetases, whereas in the second case the homology was less pronounced (>70%). However, homologies were lower within the adenylation domains responsible for activation of amino acids 4 to 7. This correlates well with the sequence variability between these antibiotics at amino acid positions 4 to 7, as shown in figure 8. The highest homology in this region was obtained for bmy_C1 and myc_C1 (81.4%), which both activate the amino acid serine in the sixth module. Furthermore, comparison of the 10 selectivity-conferring amino acid residues of adenylation domains (see introduction) revealed that Pro, Glu, Ser and Thr are activated by the last four modules, i.e. B3, B4, C1 and C2 respectively.

Table 10: Homologies and selectivity-conferring code of amino acid-specific adenylation domains (A-domains) of the bacillomycin D operon compared to the respective A domains extracted from the iturin A and mycosubtilin gene clusters

			Position of selectivity conferring amino acids[1]									
A-domain	**amino acid[2]**	**Identity[3]**	**235**	**236**	**239**	**278**	**299**	**301**	**322**	**330**	**331**	**517**
bmy_A1_Asn	**Asn (1)**	**100%**	**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
itu_A1_Asn	Asn (1)	98.6 %	**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
myc_A1_Asn	Asn (1)	80.5 %	**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
BacC, TycC[4]	Asn		**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
bmy_B1_Tyr	**Tyr (2)**	**100%**	**D**	**A**	**L**	**S**	**V**	**G**	**E**	**V**	**V**	**K**
itu_B1_Tyr	Tyr (2)	99.5 %	**D**	**A**	**L**	**S**	**V**	**G**	**E**	**V**	**V**	**K**
myc_B1_Tyr	Tyr (2)	85.2 %	**D**	**A**	**L**	**S**	**V**	**G**	**E**	**V**	**V**	**K**
TycB, TycC[1]	Tyr		**D**	**A**	**L**	V	T	**G**	A	**V**	**V**	**K**
bmy_B2_Asn	**Asn (3)**	**100%**	**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
itu_B2_Asn	Asn (3)	97.7 %	**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
myc_B2_Asn	Asn (3)	80.1 %	**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
BacC, TycC[4]	Asn		**D**	**L**	**T**	**K**	**I**	**G**	**E**	**V**	**G**	**K**
bmy_B3_Pro	**Pro (4)**	**100%**	**D**	**V**	**Q**	**F**	**I**	**A**	**H**	**V**	**V**	**K**
myc_B4_Pro	Pro (5)	44.8 %	**D**	**V**	**Q**	**F**	**I**	**A**	**H**	**V**	**V**	**K**
itu_B4_Pro	Pro (5)	42.7 %	**D**	**V**	**Q**	**F**	**I**	**A**	**H**	**V**	**V**	**K**
Pps4[4]	Pro		**D**	**V**	**Q**	**F**	**I**	**A**	**H**	**V**	**V**	**K**
bmy_B4_Glu	**Glu (5)**	**100%**	**D**	**A**	**K**	**D**	**L**	**G**	**V**	**V**	**D**	**K**
myc_B3_Gln	Gln (4)	59.8 %	**D**	**A**	Q	**D**	**L**	**G**	**V**	**V**	**D**	**K**
itu_B3_Gln	Gln(4)	58.2 %	**D**	**A**	Q	**D**	**L**	**G**	**V**	**V**	**D**	**K**
SrfAA[4]	Glu		**D**	**A**	**K**	**D**	**L**	**G**	**V**	**V**	**D**	**K**
bmy_C1_Ser	**Ser (6)**	**100%**	**D**	**V**	**W**	**H**	**F**	**S**	**L**	**I**	**D**	**K**
Myc_C1_Ser	Ser (6)	81.4%	**D**	**V**	**W**	**H**	**F**	**S**	**L**	**I**	**D**	**K**
Itu_A_C2	Ser (7)	72.4%	**D**	**V**	**W**	**H**	**F**	**S**	**L**	**I**	**D**	**K**
EntF, CdaI[4]	Ser		**D**	**V**	**W**	**H**	**F**	**S**	**L**	**I**	**D**	**K**
itu_C1_Asn	Asn (6)	43.8%	**D**	L	T	K	I	G	E	V	G	**K**
bmyC2_Thr	**Thr (7)**	**100%**	**D**	**F**	**W**	**N**	**I**	**G**	**M**	**V**	**H**	**K**
FenD, Pps2, PvD[4]	Thr		**D**	**F**	**W**	**N**	**I**	**G**	**M**	**V**	**H**	**K**
A_C2	Ser (7)	50.2%	**D**	V	**W**	H	F	S	L	I	D	**K**
mycC2_Asn	Asn(7)	47.6%	**D**	L	T	K	I	**G**	E	V	G	**K**

[1]As determined by [53]. Domains and conserved residues lining the substrate-binding pockets of adenylation domains of assigned functions are indicated in boldface.[2]The positions of the activated amino acid within the respective lipopeptides are given in parentheses.[3]This stands for the overall homology of the whole adenylation domain, about 440 amino acids, compared to the respective domain of the *bmy* operon. [4]Domains and residues lining the substrate-binding pocket as described by [54]

Interestingly, BmyA displays a remarkable complexity, similar to MycA and ItuA. The first amino acid module present in these three nonribosomal peptide synthetases is preceded by several domains with homology to proteins involved in the synthesis of fatty acids and polyketides. Four different domains could be distinguished (Fig. 20). In BmyA, the first domain (AL) shows high similarity to long-chain fatty acid CoA-ligases as well as 98% and 85% homology with the corresponding domains of ItuA and MycA, respectively. Furthermore, two domains similar to acyl carrier proteins (ACP) were recognized as well as one similar to β-ketoacyl synthetases (KS). Finally, one domain homologous to glutamate-1-semialdehyde aminotransferase (AMT) was detected. These domains presumambly play a role in the incorporation of the β-amino fatty acid into the peptide moiety [63]. The condensation domain lying directly upstream of the first adenylation domain in BmyA, responsible for the activation and incorporation of Asn, probably catalyzes the transfer of the β-amino fatty acid to the first amino acid.

According to the colinearity rule, arrangement of modules within a peptide synthetase determines the order of incorporation of specific amino acids in the peptide moiety. As the multicarrier thiotemplate mechanism proposes, elongation of the peptide occurs stepwise from the N to C end. However, in the case of bacillomycin D there was no experimental evidence proving the consequent activation and incorporation of the seven amino acids in the peptide chain. Therefore in order to verify our assumptions about the biosynthetic pathway of bacillomycin D six mutants were created by disrupting one by one the last six modules of the nonribosomal peptide synthetase. Thereby the multienzyme system was silenced at different points after the incorporation of a new amino acid resulting in intermediate products which reflect the stepwise elongation of the peptide. By identifying the intermediate elongation variants that were produced as the peptide moiety groew, it would be possible to monitor biosynthesis of bacillomycin D.

For this purpose, a chloramphenicol cassette was integrated via double-crossover recombination at the beginning of each adenylation domain (bmy_B1, bmy_B2, bmy_B3, bmy_B4, bmy_C1 and bmy_C2) resulting in the bacillomycin D deficient mutant strains AK15, AK39, AK40, AK41, AK42 and AK43 respectively (Fig. 20, Table 3). The mutant strains were grown in Landy medium at 37°C for 24 hours under vigorous shaking. Culture filtrates and sonificated cell extracts from them were prepared and were subsequently analysed by MALDI-TOF MS. The peaks of the expected products, according to the multicarrier thiotemplate mechanism, were not detected in the spectra obtained from the culture filtrates. This probably means that the intermediates remained attached to the

multienzyme system and were not secreted from the cells; only the cyclic compound could be exported from the cell. Having this concept in our mind, we sought for the expected peaks in the spectra obtained from cell extracts. Even in these spectra, those peaks had very low intensity and in most cases were hardly distinguished from the background. This result indicates that the elongation variants were tightly attached to the complex and could be only partially detached from it by sonification (see also Discussion).

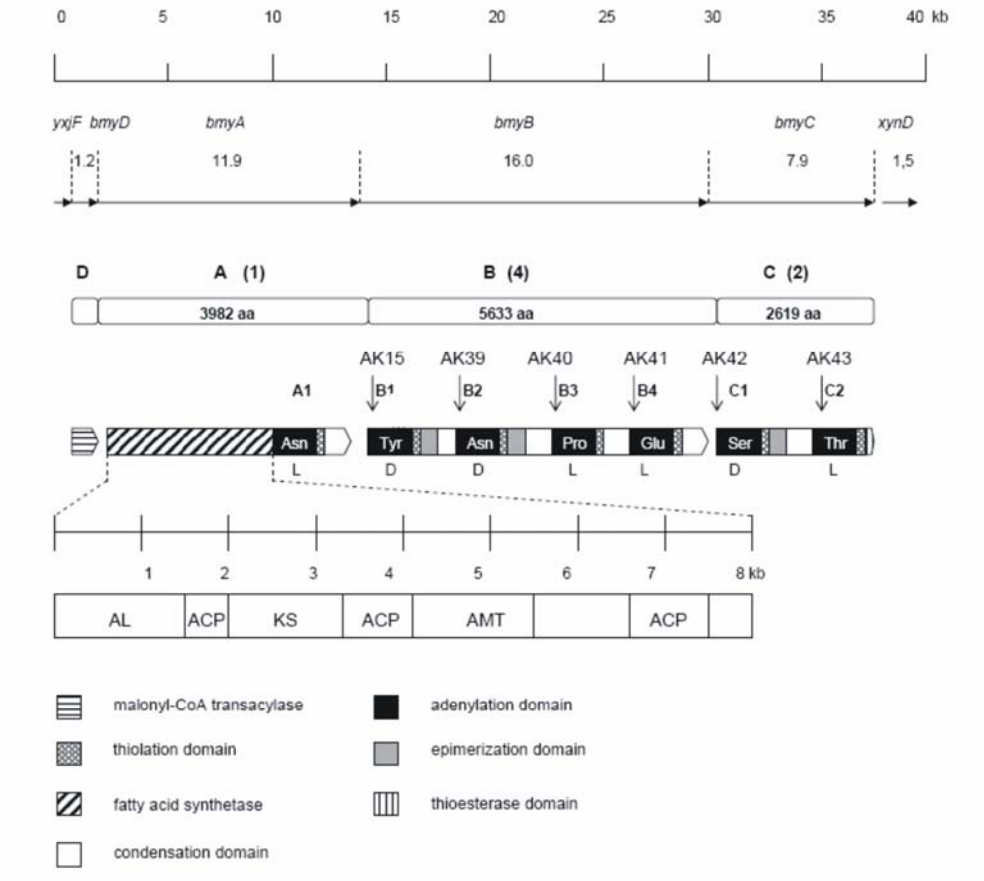

Figure 20: Schematic representation of the bacillomycin D operon in *B. amyloliquefaciens* FZB42.

The operon comprises of four ORFs *bmyD*, *bmyA*, *bmyB* and *bmyC*; their sizes are given in kb. The number of modules within each protein is indicated in parentheses. A more schematic overview of the four proteins depicts the exact location of the seven modules (A1, B1 *etc*). In parallel, the domains organisation within the modules is demonstrated. The activated amino acids are depicted within the adenylation domains while their configuration is presented under the respective domains. The arrows indicate the position where the chloramphenicol cassette was introduced within the *bmy* operon in order to construct strains that produce only intermediate products of bacillomycin D; the names of the obtained strains are also noted. AL, acyl coenzyme A ligase domain; ACP, acyl carrier protein domain; KS, β-ketoacyl synthetase domain; AMT, aminotransferase domain.

Regulation of bacillomycin D production

5'-deletion analysis of the *bmy* promoter region

Determination of *bmy* expression in *B. subtilis* MO1099

In order to monitor the transcriptional regulation of bacillomycin D of *B. amyloliquefaciens* FZB42, four reporter fusions of the postulated *bmy* promoter region (the upstream region of the first gene of the *bmy* operon, *bmyD*) to *lacZ* were generated. A series of nested fragments with a common downstream end (by the 42nd codon of BmyD) and variable upstream ends (400, 183, 120 and 30 bps upstream of the translational start) were amplified by PCR, using primers bmyD1 to bmyD5 (Fig. 21; see also Table 4). The obtained products carried suitable restriction sites (*Eco*RI and *Bam*HI; embedded on the primers) in order to be cloned into pDG268, a plasmid extensively used for constructing transcriptional reporter fusions in *B. subtilis* that can be later integrated at the *amyE* locus of its chromosome [192]. The new pDG268 derivatives (pAK5 to pAK8; see also Table 2) were subsequently used for integrating our series of 5'-deletion *bmy* promoter variants into the chromosome of *B. subtilis* MO1099 as single-copies, via double-crossover recombination. The correct chromosomal integration of the transcriptional fusions was verified by Southern hybridization. The new strains contained decreasing lengths of the *bmy* promoter region fused to *lacZ* and were named AK4, AK5, AK6, and AK7 (Fig. 21; see also Table 3).

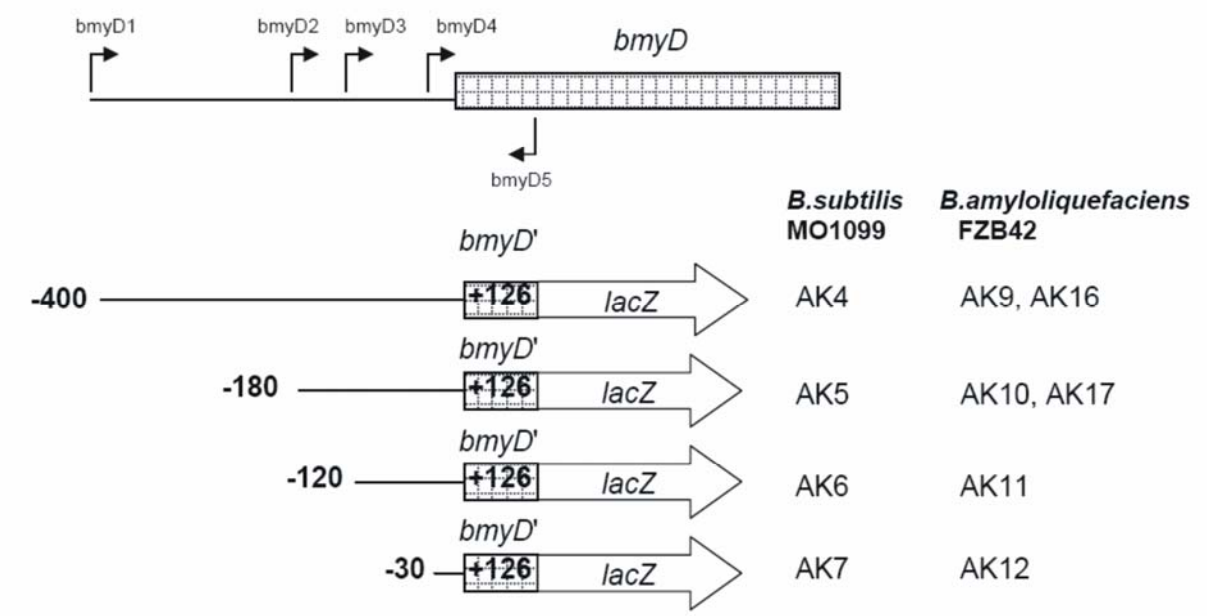

Figure 21: Schematic representation of the 5'-deletion analysis conducted for the *bmy* promoter region.

The dark arrows indicate the primers used for generating *bmyD*::*lacZ* reporter fusions (bmyD1-D5). The 5' and 3' end termini of *bmyD* promoter regions are indicated by their nucleotide position relative to the translational start. The derivative strains of *B. subtilis* MO1099 and *B. amyloliquefaciens* FZB42 that carry the respective fusions are presented on the right side of the scheme (Strains AK16-17 are isogenic to AK9-10 but have a different antibiotic cassette, see Table 3).

The expression of the *bmyD*::*lacZ* reporter fusions was determined throughout the growth cycle. All four strains were silent during logarithmic phase, indicating that the promoter(s) of bacillomycin D is not active at this time point (Fig. 22A), in total agreement with the late production of the lipopeptide observed by the MALDI-TOF MS analysis (Table 9). Upon entry into stationary phase, the transcriptional activity of AK4, AK5 and AK6 increased and reached its maximum after 3 hours. On the contrary AK7 remained silent and did not show any β-galactosidase activity all along the growth curve, in consistence with its white colour on LB agar plates containing 40μg/ml X-Gal (all three other strains appeared blue on plates). This result was partially expected, as the DNA promoter region contained in the smallest fusion (AK7) reached only up to 19bp upstream of the potential Shine-Dalgarno site, and therefore could not include an entire promoter site. Furthermore, it is apparent that the smallest DNA fragment containing an intact promoter is that encoded in AK6 (reaching up to 120bp upstream of the *bmyD* coding region). The fact that the fusions of AK6 and AK5 showed no difference in their expression pattern throughout the whole growth cycle suggested that no additional trans-activating factor binds to the region between –183 and –120bp (in respect with the *bmyD* translational start). On the other hand AK4 exhibited slightly but reproducibly higher activity than the other two strains (AK6 and AK5) during stationary phase (about 25%; Fig. 22A). This means that the *bmyD* promoter region between –400bp and –183bp (in respect with the *bmyD* translational start), harboured only by AK4 (see Fig. 21), possibly carries additional cis-activating elements. The question whether these elements code for a transcriptional activator's binding-site or for an additional promoter was addressed later.

Determination of *bmy* expression in *B. amyloliquefaciens* FZB42

To monitor the expression pattern of bacillomycin D in its natural environment, plasmids pAK5 to pAK8 (derivatives of pDG268, carrying different 5'-end deletions of the *bmy* promoter region; see previous paragraph) were attempted to be integrated at the chromosome of *B. amyloliquefaciens* FZB42. Unfortunately, pDG268 carries parts of the *amyE* gene of *B.*

subtilis (used for the double-crossover recombination of the fusion to the chromosome) that show relatively low homology (less than 80%) to their corresponding regions of *B. amyloliquefaciens* FZB42. Therefore, our initial attempts to obtain single-copy *bmyD*::*lacZ* fusions as part of the chromosome of *B. amyloliquefaciens* FZB42 were unsuccessful. In order to overcome this problem, a new plasmid, pAK9, was constructed by replacing the amylase sequences from pDG268 with the respective sequences of *B. amyloliquefaciens* FZB42. All our *bmyD*::*lacZ* transcriptional fusions were further cloned to the new vector and then successfully integrated at the *amyE* locus of *B. amyloliquefaciens* FZB42. Therefore the four new strains, AK9, AK10, AK11 and AK12, carry decreasing sizes of the *bmy* promoter region, in complete analogy to strains AK4-AK7 (note also that strains AK16-17 are isogenic to AK9-10, but have a different antibiotic cassette; see also Table 3. Exchange of the chloramphenicol cassette to a kanamycin one was performed using the marker exchange plasmid pECE73; [193].

The β-galactosidase activity of the four new strains was also examined throughout the growth cycle. The overall expression pattern of the various fusions was quite similar to that observed in the background of *B. subtilis* MO1099 (compare Fig. 22.A and B), and therefore most of the conclusions drawn in the previous section are also valid here. In other words *bmy* was only expressed during stationary phase, the entire core promoter was encoded within the first 120bp upstream of the *bmyD* start codon, and the region directly upstream of that (between –183 and –120) did not play any role in *bmy* expression. However, the strain carrying the longest upstream promoter region, AK9, exhibited 4-5-fold higher β-galactosidase activity in middle stationary phase (3-4 h after entering stationary phase when the β-galactosidase levels have reached their plateau; Fig. 22B) than the strains carrying the shorter fusions (AK10 and AK11). This difference in the expression levels between AK9 and AK10/AK11 is considerably higher than that observed for the corresponding strains of *B. subtilis* MO1099 (Fig. 22A), underlining thus the importance of this DNA upstream region in the full transcriptional activation of the *bmy* operon in its natural environment.

Two straight forward explanations can be provided for the different influence of the far upstream DNA region (between bps –400 and –183) on *bmy* expression in the two *Bacilli* strains: the cis-acting element situated at this region is optimally bound i) by a regulator (transcriptional factor or sigma factor) only present in *B. amyloliquefaciens* FZB42 (in *B. subtilis* a regulator of the same family only weakly recognises these sequences and offers basal levels of activation) or ii) by a regulator that is significantly less expressed in *B. subtilis* MO1099. No matter of the nature of this regulator, transcriptional analysis of the *bmy* operon

proceeded further in its natural environment (*B. amyloliquefaciens* FZB42), despite the practical disadvantages that such a decision had (*B. amyloliquefaciens* FZB42 is more difficult to genetically manipulate and any regulatory mutant to be tested has to be *de novo* constructed), so that important regulatory elements of bacillomycin D expression were not to be missed or underestimated.

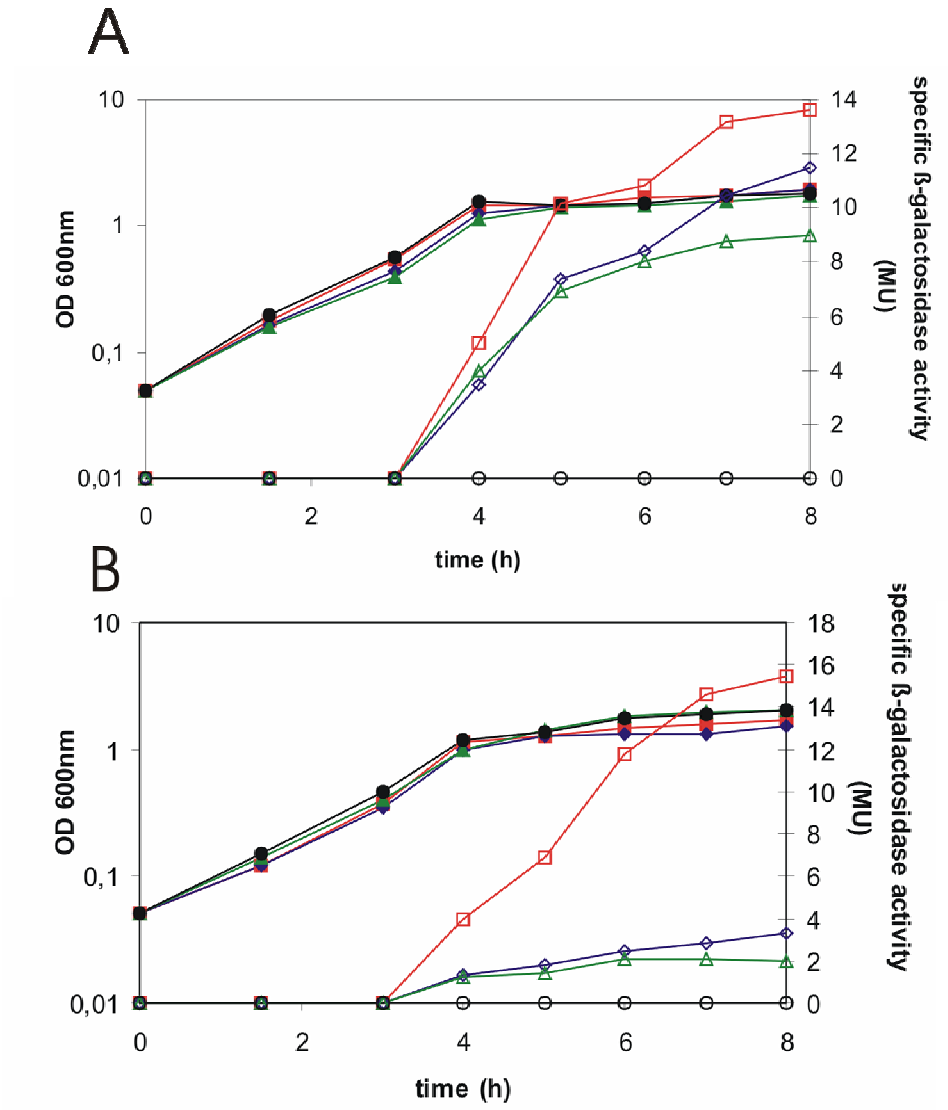

Figure 22: Expression of *bmyD*::*lacZ* fusions carrying different 5′-deletions of the region upstream of *bmyD*.

The expression of a series of transcriptional fusions of the *bmy* operon's promoter region to *lacZ* (see also Fig. 21) was monitored both in *B. subtilis* MO1099 (panel A) and in *B. amyloliquefaciens* FZB42 (panel B). Strains harbouring single copies of the 5′-deletion *bmy* promoter variants were grown in Difco medium at 37°C and optical densities (closed symbols) and specific β-galactosidase activities (in Miller Units; open symbols) were determined along the growth curve. The expression patterns shown here represent the average of more than three independent experiments. Squares, strains

AK4/AK9, carrying the longest promoter region of *bmy* (-400, +126 relative to the translational start of *bmyD*); diamonds, strains AK5/AK10, carrying the *bmy* promoter region between -183 and +126; triangles, strains AK6/AK11, harbouring the *bmy* promoter region between -120 and +126; circles, strains AK7/AK12, containing the shortest *bmy* promoter region between -30 and +126.

DegQ is partially responsible for the differences in *bmy* expression in *B. amyloliquefaciens* FZB42 and *B. subtilis* MO1099

It has been demonstrated that the horizontal transfer of functional gene clusters coding for peptide antibiotics and found in natural *Bacilli* isolates, in the chromosome of the domesticated *B. subtilis* 168 requires additional steps for the conversion of the latter into an antibiotic producer strain [172, 173]. Firstly the introduction of a functional *sfp* is absolutely necessary for peptide antibiotic production by *B. subtilis* 168, which carries a frame-shift mutation in this gene (see also introduction). Secondly it has been exhibited that increased expression of the pleiotropic regulator DegQ in *B. subtilis* 168 [224] enhances the antibiotic production [172, 173]. Interestingly most of the natural *Bacilli* isolates that express peptide antibiotics show significantly elevated *degQ* expression compared to that of *B. subtilis* 168, due to the fact that the *degQ* promoter has a more σ^A consensus-like -10 hexamer in those strains (**TACACT** instead of **CACACT**) [225]. However whether DegQ directly influences the transcriptional regulation of the antibiotic operons or it controls the expression of a post-transcriptional regulator involved in the antibiotic synthesis (for example Sfp) has not been clarified.

B. amyloliquefaciens FZB42 has the promoter version of *degQ* that yields higher DegQ cellular levels (data not shown). In contrast to that, *B. subtilis* MO1099, a derivative of the strain 168, carries the defected *degQ* promoter version. To test whether DegQ is responsible for the differential *bmy* expression patterns in its host strain and its *B. subtilis* counterpart, I constructed a plasmid that carried *degQ* under the IPTG-inducible P_{spac} promoter, pAK64 (Table 2), using the replicated vector in *B. subtilis*, pDG148 [191]. pAK64 was subsequently transformed into AK4 and AK5 (the strains harbouring the two longer *bmy* promoter regions fused to *lacZ*; see also Fig. 21) and the expression of the reporter fusions was monitored along the growth curve for the following cases: i) AK4 and AK5 without the plasmid (control), ii) AK4 and AK5 with the plasmid uninduced and iii) AK4 and AK5 with the plasmid induced at OD_{600}~0.7. As seen in Fig. 23 the presence of uninduced pAK64 hardly changed *bmy* expression from the two strains. Upon induction of pAK64 though, a significant 2.5-fold increase could be observed in the activity of AK4 (Fig. 23), a strain that carries the whole

upstream promoter region of *bmyD*. In other words, the *B. subtilis* strains carrying the two longer *bmy* promoter fusions, AK4 and AK5, exhibit a more pronounced difference in their activity when DegQ is expressed in higher levels from a plasmid. The magnitude of the difference does not match that observed for the corresponding *B. amyloliquefaciens* FZB42 strains (AK9 and AK10; Fig 22B), but with higher amounts of DegQ present in the cell, the pattern of *bmy* expression in *B. subtilis* approximates more the pattern observed in *B. amyloliquefaciens* FZB42. Since DegQ levels in *B. amyloliquefaciens* FZB42 are probably considerably higher than those in *B. subtilis* MO1099 (see above), DegQ accounts, at least partially, for the different *bmy* expression in the two host strains. Moreover, this is the first direct evidence that DegQ affects the transcriptional regulation of peptide antibiotics and not a subsequent step of their production. In the case of bacillomycin D the exertion of this effect seems to be dependent on a far upstream region of the promoter. Having in mind that DegQ is not a DNA-binding protein, the effect on bacillomycin D should be mediated in an indirect manner, possibly through modulating the activity of other transcriptional regulator(s) (see also Discussion).

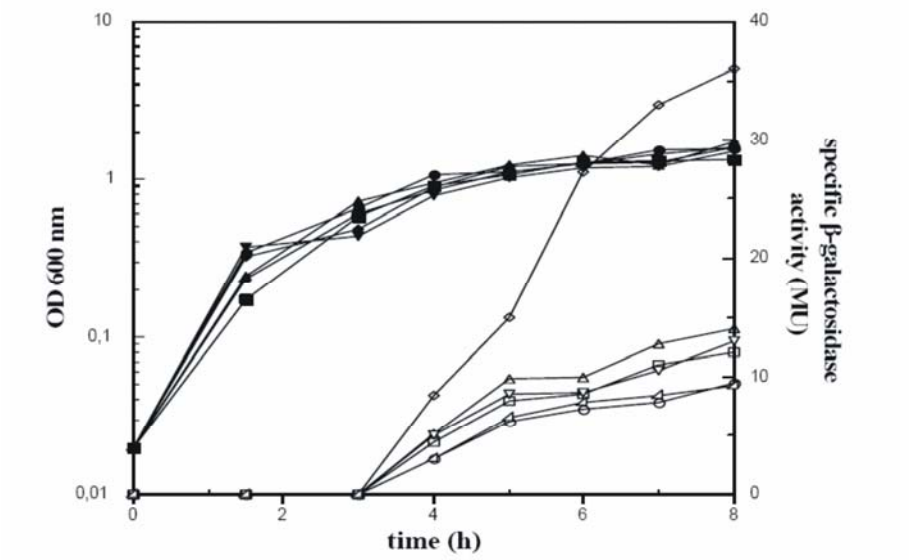

Figure 23: The effect of DegQ on the expression pattern of *bmyD*::*lacZ* fusions in *B. subtilis* MO1099.

Strains AK4 and AK5, harbouring the longest promoter region of *bmy* fused to *lacZ*, were transformed with pAK64 (a replicated plasmid carrying DegQ under an IPTG-inducible promoter). After 3h of growth, pAK64 was induced in two of the cultures by addition of 1mM IPTG. Cells were grown in Difco medium (supplemented with 5 μg/ml kanamycin where applicable) at 37°C and optical densities (closed symbols) and specific β-galactosidase activities (in Miller Units; open symbols) were determined along the growth curve. Squares, strain AK4, carrying the longest promoter region of *bmy* (-400, +126 relative to the translational start of *bmyD*), without plasmid; cirles, strain AK5, carrying

the *bmy* promoter region between -183 and +126, without plasmid; diamonds, strain AK60 (AK4+pAK64) with plasmid induced; triangles facing down, strain AK60 with plasmid uninduced; triangles facing up, strain AK61 (AK5+pAK64) with plasmid induced; rectangular triangle, strain AK61 with plasmid uninduced.

Identifying the transcriptional start site of the *bmy* operon

The transcriptional start site of the *bmy* operon was determined by primer extension. Total RNA of *B. amyloliquefaciens* FZB42 was isolated from cultures growing in Difco medium, 2,5-3 hours after their entry into stationary phase, when expression of *bmyD* promoter peaked, as shown by the *lacZ* reporter fusions (Fig. 22).

Primer extension using the radiolabelled primer rev1 (binding within the coding region of *bmyD*; see also Table 4) revealed two overlapping transcriptional start sites for *bmyD*: an adenine (A) and a thymine (T) nucleotide located 58 and 57 bp upstream of the gene's initiation codon respectively (see Figs 24 and 25). Sequences resembling the consensus of the -10 and -35 elements of the housekeeping sigma factor, σ^A, were found upstream of the mapped transcriptional start site (Fig. 24). In particular, the -35 (TATACA) and -10 (TAGGAT) hexamers identified upstream of *bmyD* carry 4/6 matches to the corresponding consensus sequences of σ^A [226]. The spacer between them is 18 bp long and within it also lays an extended -10 region, directly upstream of the –10 hexamer, i.e. C**ATG**c (the bold faced nucleotides match the consensus, TRTGn; [227, 228]. Therefore the promoter of *bmyD* (P_{bmy}) seems to be recognised and utilised by the vegetative sigma factor, despite the fact that *bmy* shows stationary-phase induced expression. Consistently the entire core promoter is encompassed within the first 75 bp upstream of the gene's translational start, as predicted by the deletion analysis of the reporter fusions (see Figs. 22 and 25).

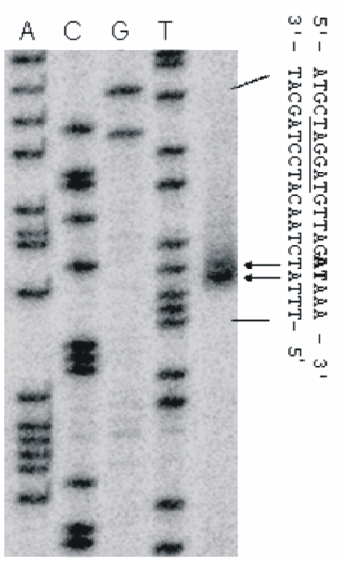

Figure 24: Mapping of the transcriptional start of the *bmy* operon by primer extension analysis.

Total RNA was isolated from *B. amyloliquefaciens* FZB42 cells grown in Difco medium, 2.5-3 hours after their entry into stationary phase. Primer extension was performed using the 5'-end ^{32}P-labelled primer rev1. The first four lanes result from dideoxynucleotide sequencing reactions using the same primer. The positions of the transcripts corresponding to the *bmyD* transcriptional starts are indicated by arrows. The respective DNA sequences are shown on the right of the picture. Transcriptional starts are highlighted in bold face whereas the -10 hexamer is underlined.

Extensive attempts to identify further upstream-situated promoters that would account for the differences in gene expression between strains AK9 and AK10 failed (data not shown). Consequently, a transcriptional regulator, binding at the region between -400 and -183 bp upstream of the gene's start codon, is partially responsible for the stimulation of the *bmy* operon's expression during stationary phase.

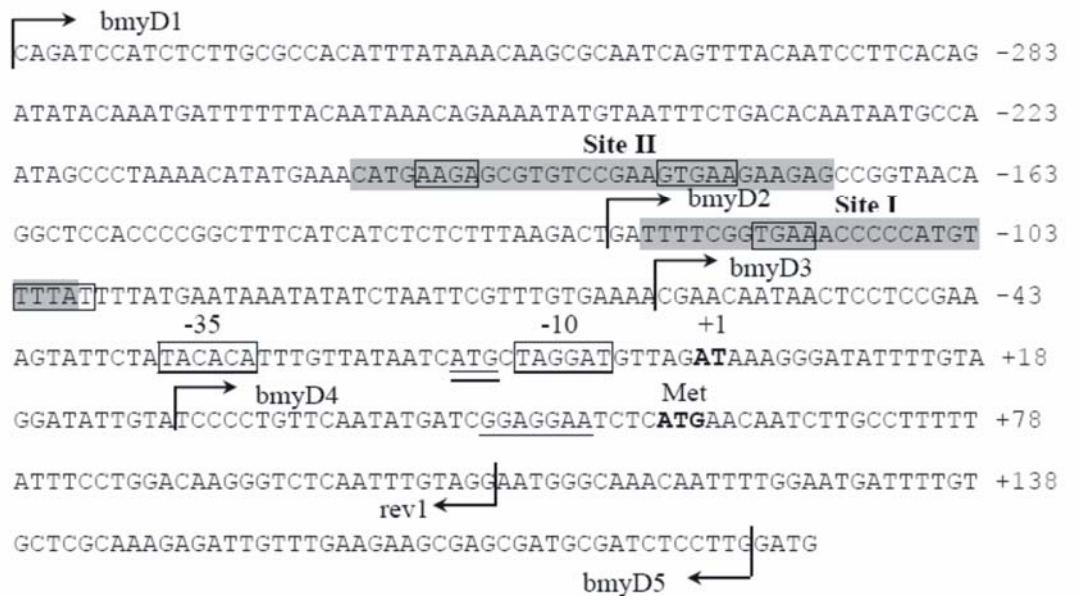

Figure 25: Nucleotide sequence of the *bmyD* promoter region.

The positions of the two adjacent transcriptional starts (bold face, +1), the translational start (bold face, Met), the putative ribosome binding site (underlined), the -35 and -10 hexamers (boxes) and the extended -10 region (double underlined) are indicated. The annealing site (5'-end) of the oligonucleotides used for primer extension (rev1) and for construction of reporter fusions (bmyD1-D5) are shown with arrows. Site I and II (shaded) represent the two DegU binding sites (see 3.4.4.3). The degenerate forms of the motif AGAA-N_{11}-TTCAG, which was proposed by Dartois *et al* (1998) as the recognition site for DegU, are indicated by boxes within the shaded sites.

Global regulators control the production of bacillomycin D

Effect of global regulators on the activity of *bmyD*::*lacZ* reporter fusions

In order to obtain further information concerning the transcriptional regulation of bacillomycin D, a series of mutations were introduced in transcriptional regulatory proteins and sigma factors of the bacterium. Some of the genes mutated are also found in the genome of *B. subtilis*, whereas others are novel members of the genome of *B. amyloliquefaciens* FZB42. The several mutations were then crossed over into strains AK9 and AK10 (AK11 was excluded from these experiments, since it displays the same pattern of gene expression as AK10). Thereby, it was possible to monitor the behaviour of the mutant strains by calculating their β-galactosidase activities along the growth curve and comparing them to that of the wild type strain.

Using this approach, several players involved in the regulation of the *bmy* operon were identified. Two two-component response regulator proteins were found to be essential for full

activation of the operon: DegU, known to control the expression of degradative enzymes [229, 230] and to be involved in the initiation of competence [231] and ComA, a regulator of late-competence genes [232] and surfactin production [112]. Moreover, the sporulation sigma factor σ^H [233] regulates expression of bacillomycin D.

The effects of DegU, ComA and σ^H on the activity of reporter fusions are analytically presented in Figure 26. Inactivation of the genes coding for each of the above proteins resulted in severely impaired promoter activity during stationary phase, especially in strains harbouring the whole promoter region (AK9 derivatives). *bmy* expression was 3-4-fold lower in the *degU*, *comA* and *sigH* mutant derivatives of AK9 (AK32, AK22 and AK52 respectively) compared to the parental strain AK9 (Fig. 26), and similar to that of wild type strain AK10, carrying the shorter promoter region (its 5'-end is deleted 183 bp upstream of the *bmyD* translational start). Furthermore, in the absence of DegU, ComA or σ^H, the activity of AK10 derivatives (AK33, AK23 and AK53 respectively) was also reduced in comparison to that of the parental strain AK10, albeit to a much smaller extent (Fig. 26).

Therefore DegU, ComA and σ^H are all required for the full activation of the *bmyD* promoter and their effects seem to be mostly exerted through the DNA region located between -400 and -183 bps upstream of the gene's translational start. Despite the severity of the reported effects, the expression of *bmy* is not completely silenced by any of those mutations (on the contrary to the promoter silencing observed with strain AK12; (Fig. 22B), insinuating that the σ^A-dependent promoter retains both stationary phase induction and basic levels of *bmy* expression without any of these regulators (see also Discussion).

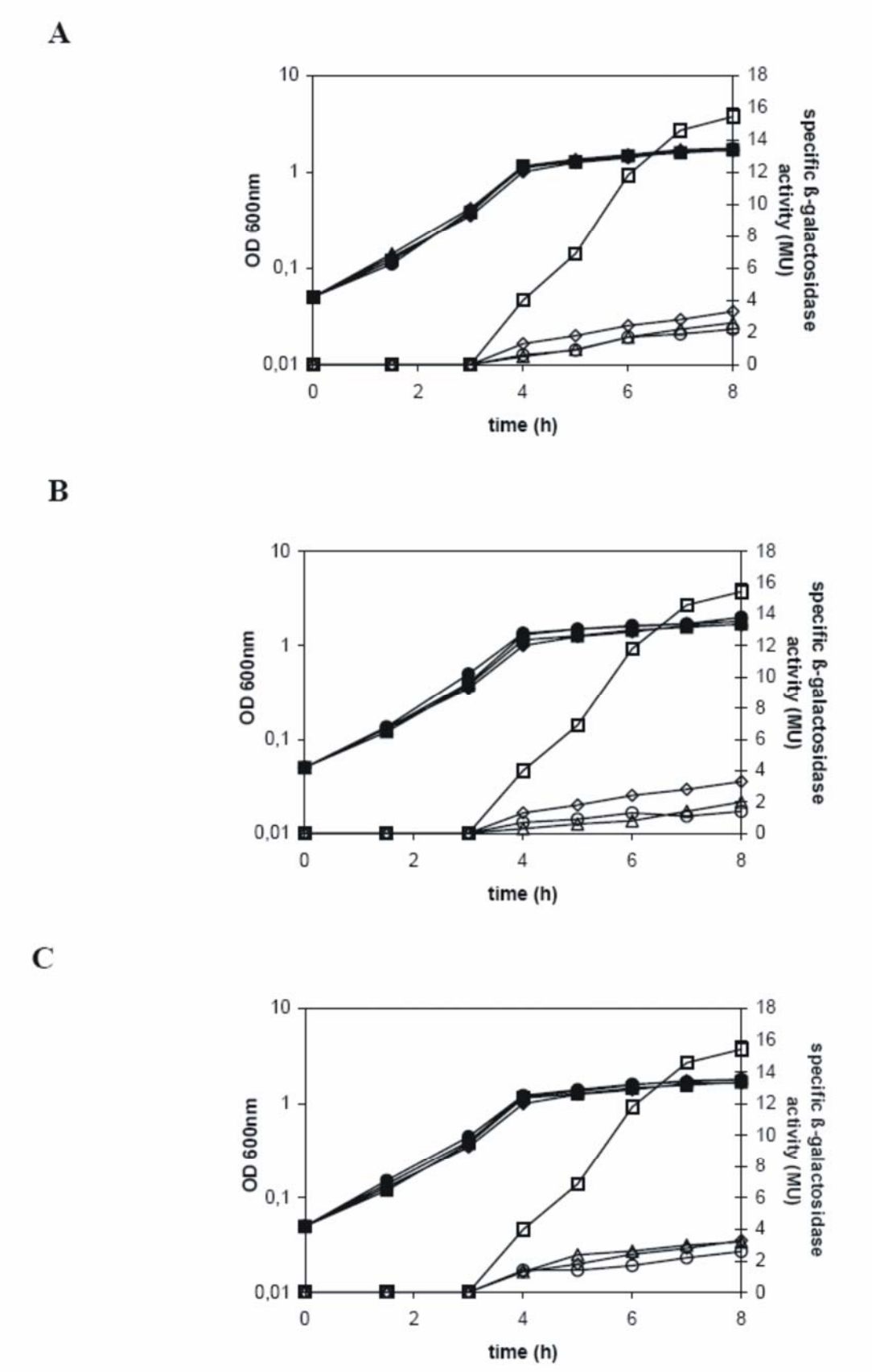

Figure 26: Effects of ComA, DegU and σ^H on the expression of the various *bmyD*::*lacZ* fusions.

Strains AK9 and AK10, carrying the two *bmyD*::*lacZ* fusions with the longest upstream promoter region and the highest activity (see also Fig. 21, 22), and their *degU* (A), *comA* (B) and *sigH* (C) mutant derivatives were grown in Difco medium at 37°C and optical densities (closed symbols) and specific β-galactosidase activities (in Miller Units; open symbols) were determined along the growth curve. In each panel the expression of the fusions in the wild-type background is provided for direct comparison to that of the mutants. Squares, AK9; diamonds, AK10; triangles, mutant derivatives of AK9 (AK32 in A, AK22 in B, AK52 in C); circles, mutant derivatives of AK10 (AK33 in A, AK23 in B, AK53 in C).

In addition, the effect of σ^H cannot be associated with the presence of a second upstream promoter, since primer extension experiments ruled out such as a scenario (see above). Consistently, no similarity with the promoter consensus sequence of σ^H (AGGANNT-15-17bp-GAAT; [234] could be found in the entire *bmy* promoter region. Thus, the effect of σ^H exerted on P_{bmy} is indirect.

Deletion mutants of several other transcriptional regulators and sigma factors did not significantly influence the expression of the *bmy* operon, as shown by monitoring the activities of the respective reporter strains along the growth curve (data not shown). The genes/operons tested were: an uncharacterised two-component system *RBAM01839-RBAM01840* in the close proximity of *bmy* operon, present only in FZB42 (the homology to other known two-component systems of *B. subtilis* is very low; higher scores with *yvrGH*; [235]; *yerPO,* YerP belongs to a resistance-nodulation-cell division (RND) family proteins and is involved in surfactin immunity/production [170] whereas YerO is a putative transcriptional regulator encoded adjacent to it; *spaR*, encoding a response regulator that activates subtilin production and immunity [33], present in *B. amyloliquefaciens* FZB42 but not in *B. subtilis* MO1099; *sigW*, encoding an extracytoplasmic sigma factor involved in antibiotic resistance [236]; and *aat*, encoding a putative transcriptional regulator direct downstream of *srfAD*, that is only present in *B. amyloliquefaciens* FZB42 (see also Fig. 15).

Effects of *degU*, *comA*, *sigB* and *sigH* mutations on transcriptional initiation by the identified promoter of *bmy* operon (P_{bmy})

Reporter fusions showed that DegU, ComA and σ^H positively regulate transcription of the *bmy* operon in *B. amyloliquefaciens* FZB42. To obtain further proof, the effects of these mutations on the activity of P_{bmy} were examined by primer extension. In addition the role of

σ^B, the general stress sigma factor in *Bacilli* [237, 238], was also investigated. Total RNA was extracted from wild type or mutant cells that had reached middle stationary phase, and primer extension was performed as described earlier (see 3.4.2 and materials and methods).

A unique transcriptional start site was identified for the *bmy* operon in both the wild type and the four mutant strains, and as expected coincided with the one reported in the previous section (Fig. 24). However the intensity of the obtained transcripts varied enormously at the different genetic backgrounds; the wild type strain produced a strong and clearly distinguishable transcript whereas *degU*, *comA*, *sigH* and *sigB* mutant strains gave only weak but reproducible signals (Fig. 27), a sign of decreased P_{bmy} promoter activity. This finding is in perfect agreement with our previous results (Fig. 26), identifying DegU, ComA and σ^H as positive regulators of P_{bmy} in stationary phase. It also verifies our previous statement that cells lacking those regulators are still able to show basal expression of the σ^A-dependent P_{bmy}. Moreover, σ^B serves as a new positive transcriptional regulator of the *bmy* operon; its effects were of similar magnitude to those of the other three regulators (Fig. 27), and most probably exerted through an indirect mechanism since there are no sequences in the *bmyD* promoter region that apparently resemble the promoter consensus sequence of σ^B (GTTT-15-17bp-GGGWAW, where W stands for A/T; [239] see also section 3.4.5).

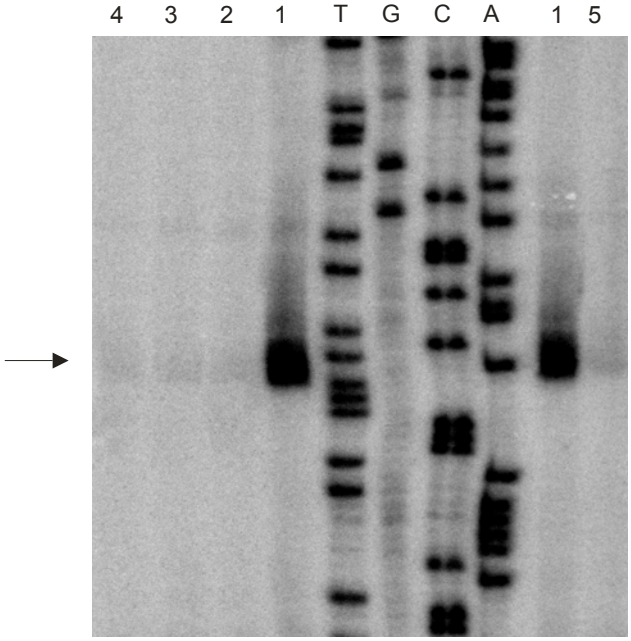

Figure 27: Effects of *degU*, *comA*, *sigB* and *sigH* mutations on the activity of the *bmy* operon promoter (P_{bmy}).

Cells were grown in Difco medium and 2.5-3 hours after entry into stationary phase, total RNA was extracted from *B. amyloliquefaciens* FZB42 and its mutant *degU* (TF1), *comA* (CH23), *sigB* (CH33) and *sigH* (AK50) derivative strains. Primer extension was performed using the same amount of total

RNA (40 μg) and the 5'-end ^{32}P radiolabelled primer rev1. Sequencing ladders were generated with the same primer. Intensities of the obtained transcripts indicate the effects exerted by the respective mutations on the activity of the P_{bmy} promoter. 1, *B. amyloliquefaciens* FZB42; 2, TF1; 3, CH23; 4, AK50; 5, CH33; +1 and the arrow both indicate the transcriptional start.

Several more transcriptional regulators and sigma factors were deleted or disrupted and their role on *bmy* operon's transcription was further assessed by primer extension analysis: *sig01*, an extracytoplasmic sigma factor identified for the first time in FZB42; *sigD*, the *Bacillus* sigma factor involved in chemotaxis, flagella synthesis, motility [240]; three more extracytoplasmic sigma factors, encoded by *sigX*, *sigV*, *sigW* and different combinations of them [236, 241, 242]; and *codY*, encoding a global transcriptional regulator in *B. subtilis* [243] and a direct repressor of surfactin expression [169]. All these mutant strains produced an equally strong transcript signal as that of the wild type strain, indicating that they do not play a role in transcription of the *bmy* operon (data not shown).

MALDI-TOF MS analysis of *B. amyloliquefaciens* FZB42 strains deficient of global regulators that are involved in transcription of the *bmy* operon; DegU has a post-transcriptional effect on bacillomycin D production

In order to decipher how the reported effects of DegU, ComA, σ^B and σ^H on *bmy* transcription reflect to the end-production of bacillomycin D, we decided to monitor the antibiotic's synthesis in the different mutant strains by mass spectrometric analysis. MALDI-TOF MS was performed using culture filtrates of the *degU*, *comA*, *sigH* and *sigB* mutant derivatives of *B. amyloliquefaciens* FZB42 (grown with aeration for 24h in Landy medium), and the obtained spectra were compared to that of the wild type, in order to evaluate the relative production levels of bacillomycin D in the different genetic backgrounds. Even though this type of analysis does not provide accurate quantitative results, relative production of bacillomycin D can be roughly estimated by comparing the intensity values of the peaks that correspond to bacillomycin D with the intensity values of the surrounding peaks (belonging to other peptide products) in the wild type and mutant strains.

It has been previously shown that mutations in *degU*, *comA*, *sigH* and *sigB* severely impair but do not silence transcription of the *bmy* operon. Consistently, bacillomycin D production was defected but not entirely blocked in the *comA*, *sigH* and *sigB* mutant strains, since the intensities of the peaks reflecting the presence of bacillomycin D were considerably weaker in the spectra of the mutant strains compared to the spectrum of the wild type strain

(Fig. 28). However, no bacillomycin D could be detected from a culture filtrate stemming from the *degU* strain (Fig. 28.E and F), suggesting that the role of DegU in bacillomycin D production goes beyond than the mere activation of the P_{bmy} promoter.

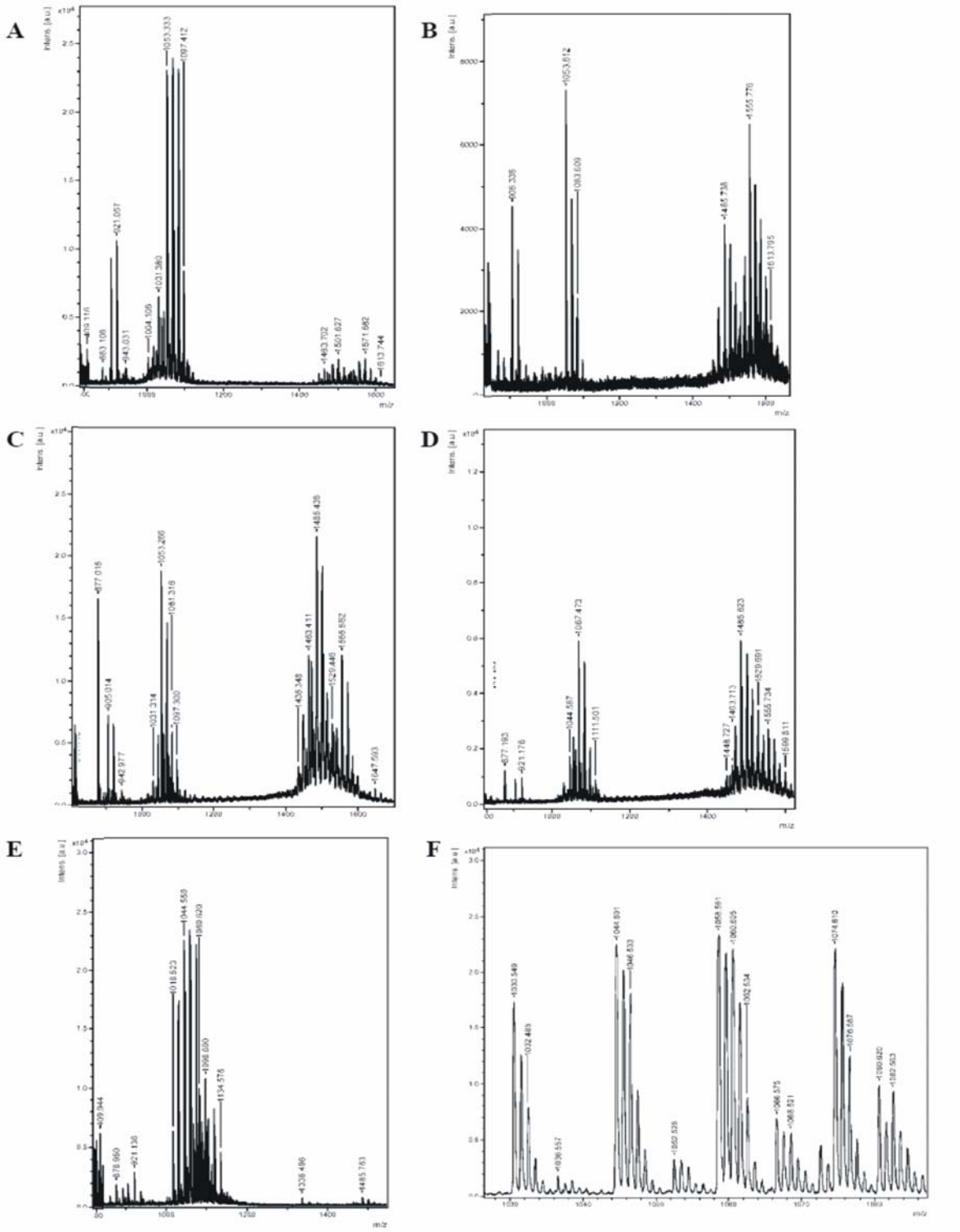

Figure 28: MALDI-TOF MS analysis of *comA*, *sigB*, *sigH* and *degU* mutant strains. The absence of DegU deprives the cell of bacillomycin D production (performed in collaboration with Dr. J. Vater and Xiao-Hua Chen).

Spectra of culture filtrate extracts prepared from *B. amyloliquefaciens* FZB42 and its mutant derivatives, after 24 hours growth in Landy medium. A. In the wild type strain, bacillomycin D (*m/z* 1053.3, 1097.4) is the antibiotic produced in the highest amounts under these conditions (see also 3.3), judged from the intensities of its peaks, whereas the siderophore bacillibactin (*m/z* 833, 905 and 921; see also 3.5) and surfactin (*m/z* 1044) follow in production scale. The peaks corresponding to fengycin production (*m/z* 1463.7, 1501.6) are the ones with the lowest intensity. For more details concerning the exact number of peaks that define the presence of an antibiotic in the external milieu of *B. amyloliquefaciens* FZB42 see also Table 8. B-D. The production pattern of the four compounds changes in *comA* (B), *sigB* (C) and *sigH* (D) mutants. In all three cases the levels of bacillomycin D are relatively low in comparison with the other the antibiotics and in most cases hardly exceed them. In addition the intensities of the peaks corresponding to fengycin seem increased in all three mutants, rendering thus fengycin as one of the major antibiotics produced by *B. amyloliquefaciens* FZB42 in these genetic backgrounds (see also 3.5 and Table 11). As expected, the *comA* mutant strain (B) is deficient of surfactin production [161, 162]. E-F Bacillomycin D is absent in the *degU* mutant strain. A zoomed in version of the panel E spectrum is presented in panel F, where only the peaks attributed to surfactin are apparent while those belonging to bacillomycin D (*m/z* 1053.7, 1067.7, 1069.7) are absent.

Since in the absence of all four regulatory proteins, i.e. DegU, ComA, σ^B and σ^H, P_{bmy} retained its basal expression (Figs 26 and 27), but only in the absence of DegU the production of bacillomycin D was completely abolished, DegU seems to have an additional "control-point" on *bmy* regulation, at a post-transcriptional level. To test whether the post-transcriptional effect of DegU was associated with the export of bacillomycin D to the extracellular milieu, sonificated cell extracts of the *degU* strain (grown with aeration for 24h in Landy medium) were also analysed by MALDI-TOF MS. No bacillomycin D was apparent in the cell extracts too (data not shown), suggesting that DegU influences both transcription of the *bmy* operon, and also a step post to transcription of *bmy* but prior to its export to the surrounding environment.

DegU directly binds to the bacillomycin D promoter

DegU was shown to be essential for bacillomycin D production, both at transcriptional and post-transcriptional level. In order to fully decipher its role on the activation of P_{bmy}, the response regulator was purified as a C-terminal His_6-Taq fusion protein [231], and was further used in EMSA and DNAse I experiments. Overexpression and purification of 6xHis-tagged

DegU have been already described in detail in material and methods. Analysis by SDS-PAGE (Fig. 29) and Western blot revealed that the purification of native 6xHis-tagged DegU (35 kDa) protein was successful.

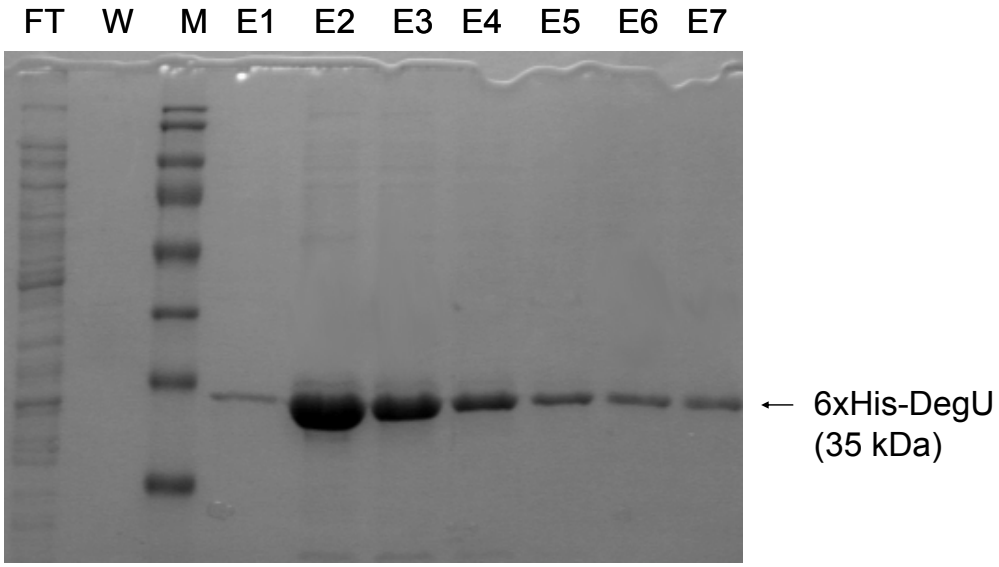

Figure 29: Overexpression and purification of the 6xHis-tagged DegU

Overexpression and purification of the 6xHis-tagged DegU protein was performed as described in detail in chapter 2.5.5. Samples collected from different steps of the purification process were analyzed by SDS-PAGE electrophoresis. 6xHis-tagged DegU was the only purified protein. FT, sample from flow through; W, sample from a wash step; E1 to E7, samples from elution steps 1 to 7; M, Prestained Protein Ladder (Fermentas), bands from bottom to top 25, 35, 45, 55, 70, 100, 130, 170 kDa.

EMSA shows that DegU is a direct activator of the *bmy* promoter

Our genetic analyses identified a positive role of the response regulator DegU in *bmy* transcription. To examine whether the *bmy* promoter is a direct target for DegU, gel retardation mobility shift assays (EMSA) of a 450 bp DNA fragment, harbouring regions between -342 and +108 relative to the transcriptional start (or -400 to +50 relative to the translational start) were performed using increasing concentrations of the response regulator. As seen in Figure 30A, the *bmyD* promoter indeed contains specific binding sites for DegU. The DNA fragment was already shifted with 0,2 and 0,4 µM unphosphorylated DegU, in an analogous manner to the DegU binding at the *comK* promoter [231]. Mobility of the 450 bp DNA fragment changed dramatically upon incubation with higher amounts of the protein (0,8 and 1,6 µM), indicating the presence of more than one DegU binding-sites in the promoter

region of *bmy*. The observed shifts are DegU-specific since no migration of the DNA fragment was observed upon incubation with the same amounts of BSA (Fig. 30A).

Since our reporter fusions' results indicated that DegU influences both expression of the shorter and the longer versions of the *bmyD*::*lacZ* fusions (present in strains AK9 and AK10; Fig. 22), we sought to narrow down the regions that DegU binds. In order to accomplish this, gel retardation mobility shift assays were also performed with two smaller DNA fragments. Fragment D1 encompassed the region between -342 to –126 bp relative to the transcriptional start (or -400 to –184 bp relative to the translational start of *bmyD*), present only in AK9, while fragment D2 encompassed the DNA region between -125 to +108 bp relative to the transcriptional start (or -183 to +50 bp relative to the translational start of *bmyD*), present in both strains, AK9 and AK10. EMSA experiments with fragments D1, D2 and increasing concentrations of the response regulator revealed that unphosphorylated DegU directly bound to both DNA regions, with similar affinities (Fig. 30B). It is also apparent that higher amounts of DegU (0,4 or 0,8 μM) lead to a supershift of the two DNA fragments (similar to that observed in the EMSA experiments with the DNA fragment encompassing the whole promoter region; Fig. 30A). This indicated that the initial DegU binding to the *bmy* promoter region might trigger the co-operative binding of more DegU molecules to the promoter and/or change the promoter's architecture (see also Discussion).

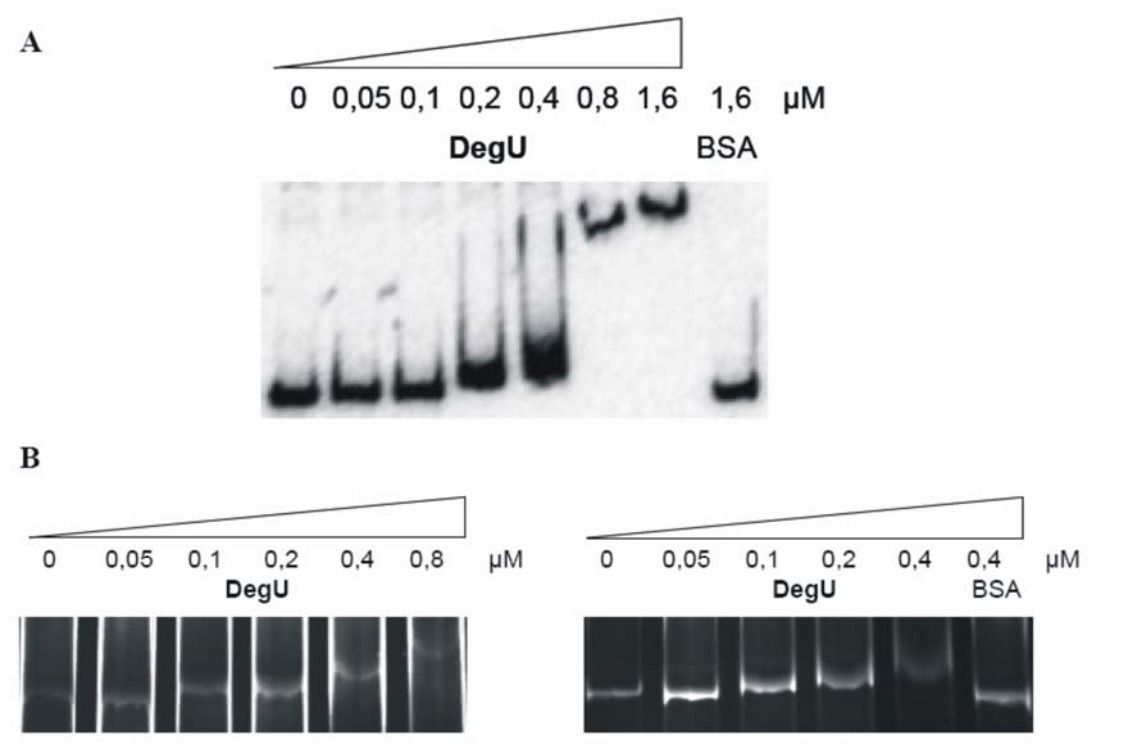

Figure 30: Gel retardation mobility shift assays (EMSA) of the *bmyD* promoter region

A) Gel retardation mobility shift assays of a ^{32}P-labeled *bmyD* promoter fragment using increasing concentrations of DegU, as indicated on the top of the gel. The DNA fragment used harbours regions

between -342 and +108 bp, relative to the transcriptional start. The incubation of the *bmy* promoter region with the BSA protein (same amount as that of the highest DegU concentration used) did not cause migration of the DNA fragment.

B) Gel retardation mobility shift assays using two smaller DNA fragments of the *bmyD* promoter region, D1 and D2, ranging from -342 to -126 bp and -125 to +108 bp respectively (relative to the transcriptional start). Gel mobility shifts were performed using 30ng of DNA fragments D1 (left gel) and D2 (right gel) and increasing concentrations of DegU, indicated on the top of the gels. A control using BSA protein was also performed. Visualization was obtained by ethidium bromide staining.

Mapping the location of the DNA-binding sites of DegU on the *bmy* promoter region

DNase I footprinting analysis was performed in order to determine the exact location of the DegU binding-sites in the *bmy* promoter region. DNase I treatment followed incubation of either a linear DNA fragment encompassing the whole *bmyD* promoter region or the same promoter fragment cloned in a plasmid, and thus supercoiled, with unphosphorylated DegU (0, 0,8, 1,6 μM DegU). Primer extension with two different primers enabled the visualization of both the template and non-template strand (only the data obtained using the supercoiled DNA can be seen in Fig. 31; similar data were obtained with the linear DNA fragment).

The footprinting patterns obtained with the coding and the non-coding strand lead to similar conclusions. In detail, the presence of DegU lead to the protection of a region spanning from -123 to -106 (relative to the transcriptional start) on the top strand, followed by an extended region of hypersensitive sites ranging from -103 to -85 (Fig. 31). Moreover, directly downstream of a series of hypersensitive sites at around -210 of the top strand, is situated a relative protected region from DegU between -201 and -172. Consistently the bottom strand revealed a strongly protected area from -116 to -99 and a relatively more weakly protected region between -198 and -172 upon DegU addition (Fig. 31). Strong hypersensitive sites could be observed in the region between -98 and –66, and at -201 and -203 of the same strand (Fig. 31).

In conclusion, unphosphorylated DegU binds two distinct sites at the promoter of *bmy*, i.e. Site I (-123 to -99) and Site II (-201 to -172) (see also Fig. 26), inducing bends and local changes in the DNA architecture adjacently to these sites (seen as hypersensitive sites). The two DNA binding-sites bear the A/T-rich signature of DegU recognition-sites (see also Discussion; [244, 245], and their location is also consistent with the data obtained by the gel

retardation assays. It is worth mentioning that phosphorylated DegU (after incubation with acetyl phosphate) produced identical DNase I footprints (data not shown).

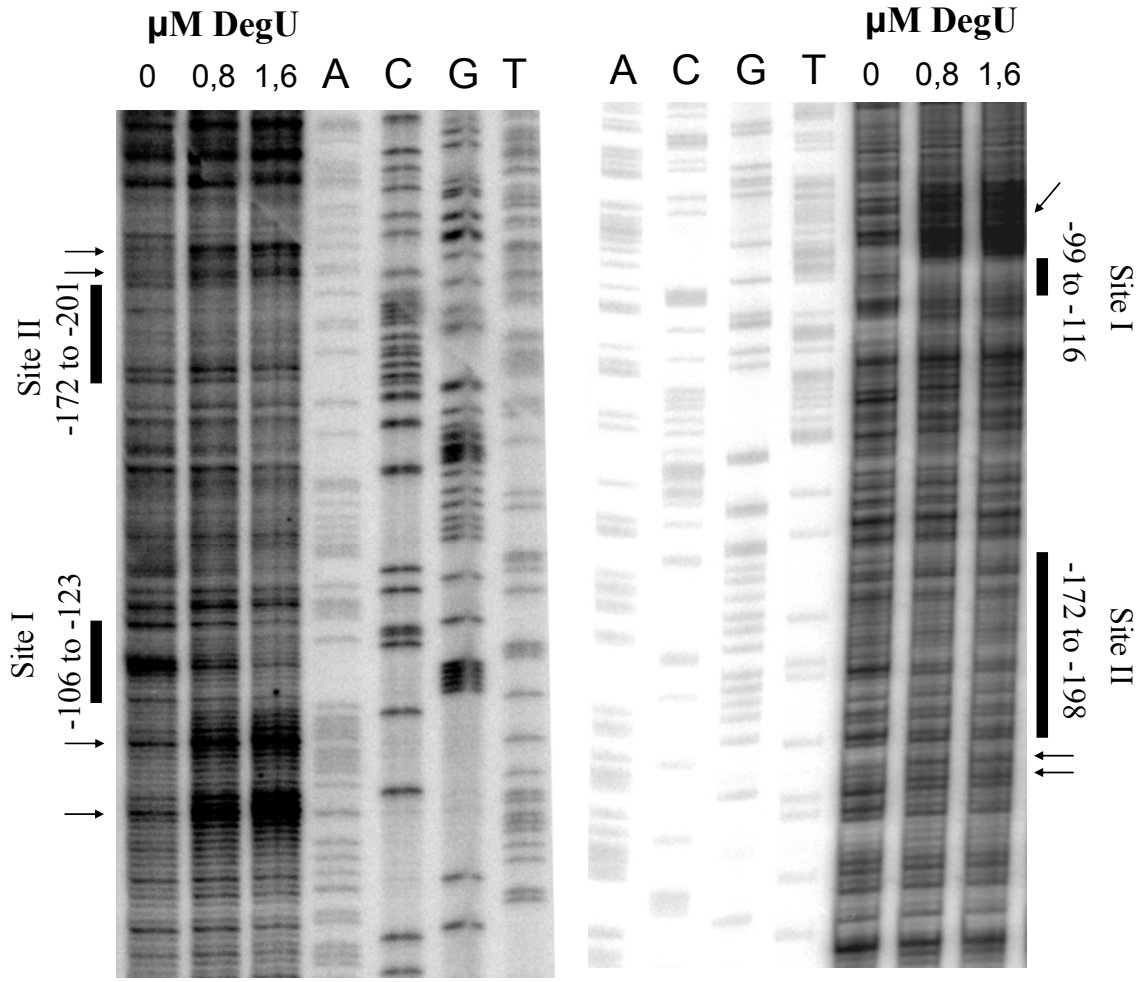

Figure 31: DNase I footprinting analysis of DegU at the *bmy* promoter region

The DNase I digestion patterns were obtained using primer extension in order to be able to monitor both linear and supercoiled DNA (here shown the results with the latter). On the left panel can be seen the footprint pattern of the coding strand in the presence of increasing concentrations of DegU, whereas on the right panel is presented the footprint pattern of the non-coding strand. The protected and hypersensitive regions are marked with bars and arrows respectively. The sequence reactions of the appropriate DNA strand were used as size markers.

The effect of DegU on *bmy* transcription is epistatic to that of DegQ

It has been previously observed that DegQ exerts a positive effect on bacillomycin D expression (see Fig. 23 and corresponding text), similarly to its role in the production of other peptide antibiotics such as iturin A and plipastatin [172, 173]. This effect is most probably

mediated in an indirect manner (no DNA-binding ability is predicted for DegQ), requiring the existence of a far upstream region in the case of the P_{bmy} promoter region (between –400 and –183 bps in respect with the translational start of the first gene of the operon, *bmyD*). Since DegU has a DNA binding-site in this region (see Figs. 25 and 31), a plausible scenario is that the effect of DegQ is mediated through DegU. In addition DegU is known to stimulate the expression of *degQ* [171], and an alternative scenario would be that the transcriptional effects of DegU on *bmy* expression are indirect and are due to decreased DegQ levels. To test these hypothesises, the plasmid carrying an IPTG inducible copy of *degQ*, pAK64, was transformed in the *degU* mutant derivative of *B. amyloliquefaciens* FZB42. The strain was grown until middle-late exponential phase (OD_{600}~0.7), pAK64 was induced with IPTG, and cells were harvested after 4h for total RNA preparation. The supernatants were lyophilized and further analyzed by MALDI-TOF mass spectrometry (see Fig. 32B). Primer extension analysis revealed that the P_{bmy} activity was equally low in both the *degU* mutant and the *degU* mutant with an overexpressed DegQ, in contrast to the strong activity that P_{bmy} showed in the wild type background (Fig. 32A). In addition, bacillomycin D was not detected in the spectrum of the *degU* mutant strain carrying an induced pAK64 (Fig. 32B), similarly to the spectrum obtained by the *degU* mutant strain (Fig. 28F). This means that the effect of DegU on *bmy* expression is epistatic to that of DegQ, and that the latter needs the former in order to exert its role.

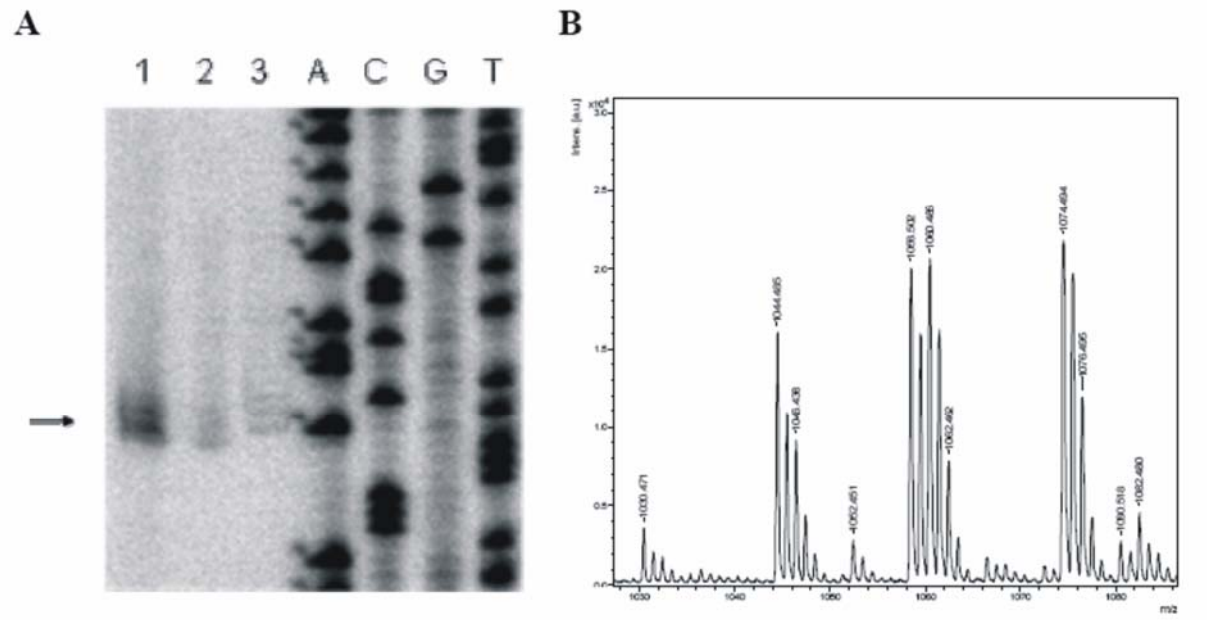

Figure 32: Increased DegQ cellular levels cannot restore bacillomycin D production in a *degU*⁻ background

Strain TF1 (*degU*⁻) containing pAK64, which carries the *degQ* gene under an IPTG inducible promoter (strain AK58), was grown in Difco medium, supplemented with 5 μg/ml kanamycin, until

mid-exponential phase. The plasmid was then induced with 1mM IPTG, the strain was further grown for 4 hours and the cells were harvested for total RNA extraction and MALDI-TOF MS analysis.
A. Primer extension analysis was performed as described previously (see Material and Methods and Fig. 27). Intensities of the obtained transcripts indicate the effects exerted on P_{bmy} promoter activity by the respective mutations. 1, *B. amyloliquefaciens* FZB42; 2, TF1; 3, AK58 (TF1 with pAK64). Induction of DegQ expression cannot alleviate the defect that the absence of DegU imposes on P_{bmy} activity.
B. MALDI-TOF MS analysis of the culture filtrate extracts of strain AK58. Bacillomycin D is not synthesized (the obtained peaks are only due to surfactin production, see Table 8). The analysis has been done in collaboration with Dr. J. Vater.

The fact that ComA is also known to positively regulate the expression of *degQ* and to have an even more pronounced effect than DegU [171], motivated us to clarify whether the effect of ComA on *bmy* expression is imposed through DegQ. However extensive efforts to transform a wild type *B. amyloliquefaciens* FZB42 with pAK64 (that subsequently would be modified to a *comA*$^-$ mutant) were unsuccessful. This result deprived us of testing the truth of this hypothesis, but generated interesting implications considering the cellular role of DegQ (see Discussion).

σ^B mediates its control on P_{bmy} by indirectly controlling the repression of a novel member of the Rap protein family

The effects of σ^B and σ^H on P_{bmy} are most probably indirect for reasons explicitly stated above (see chapter 3.4.3.2). Since DegU and ComA are also involved in *bmy* transcription (the former also acts directly; see above), we reasoned that several Rap proteins should be also indirectly involved in the expression of bacillomycin D. Rap proteins extensively regulate the activity of response regulators in *B. subtilis*, and five of them (out of eleven present in *B. subtilis*) have been shown to directly inhibit the DNA-binding ability of either ComA or DegU, i.e. RapC [246], RapF [167], RapK [247], RapG [248] and RapH [249]. Most Rap proteins (including all five members named before) come as pair with a small-sized Phr protein, which after its synthesis is excreted from the cell and processed to a signal pentapeptide [250]. The pentapeptide is re-imported to the cell by an oligopeptide permease [251] and inhibits the activity of its cognate Rap protein [166]. The *rap-phr* gene pairs usually constitute operons, but the expression of *phr* is often controlled by an extra σ^H- dependent

promoter ([234]; see also discussion). Recently RghR (formerly YvaN) was reported to repress the expression of *rapG* and *rapH* and thereby to have a positive effect in the expression of downstream targets of ComA and DegU [249]. Interestingly *yvaN* had been previously identified as a member of the σ^B regulon [252]. These data provided possible links-suggestions to the indirect effects of σ^B and σ^H on P_{bmy} that were further addressed.

B. amyloliquefaciens FZB42 possesses 9 putative Rap proteins (carrying a TRP-tetratricopeptide domain), among which only 6 are highly homologous to their *B. subtilis* counterparts (see also Discussion). Both *rap* targets of RghR (*rapG* and *rapH*) are not conserved in *B. amyloliquefaciens* FZB42, but tandem RghR binding-sites [249] are predicted in front of one of the novel *rap* members of the strain, *rapX* (data not shown). *rapX* is located at the same chromosomal locus as *rapG* in *B. subtilis*, but is encoded in the opposite strand, carries no obvious *phr* counterpart (in contrast to *rapG*) and shares only basic homology to *rapG* (27% in protein level; a percentage similar to that exhibited to other Rap proteins of *B. subtilis*; see also Discussion).

Having in mind that σ^B exerts an indirect positive effect on P_{bmy}, we hypothesized that this effect can be mediated through a pathway involving RghR and RapX. To test this hypothesis a *rapX* $^-$ and a *rapX* $^-$*sigB*$^-$ double mutant strain were constructed and assayed together with the wild type and *sigB*$^-$ strains for P_{bmy} activity by primer extension (Fig. 33). While the absence of σ^B (theoretically causes a decrease in RghR levels, and thereby an increase in those of RapX) significantly reduced P_{bmy} activity, the effect was alleviated when *rapX* was also mutated. This derepression observed verified that σ^B and RapX work on the same pathway (with RghR -conserved in *B. amyloliquefaciens* FZB42- as intermediate) to influence the activity of P_{bmy}. Nevertheless, the P_{bmy} activity was only slightly higher in the *rapX* $^-$ background, compared to that of the wild type (Fig. 33), indicating that RapX is not substantially repressing DegU or ComA in stationary phase, where σ^B activates the expression of RghR (see also Discussion and Fig. 39)

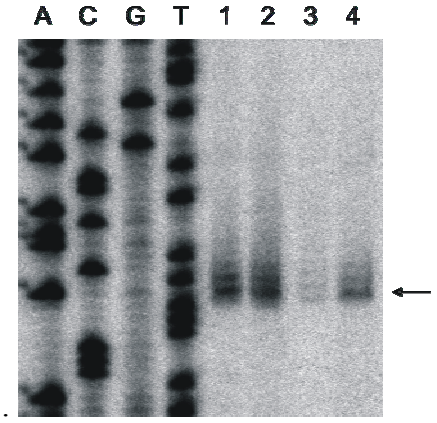

Figure 33: σ^B activates expression of P_{bmy} due to the repression it exerts on a novel Rap protein found in *B. amyloliquefaciens* FZB42, RapX

Cells were grown in Difco medium and 2.5-3 hours after entry into stationary phase, total RNA was extracted from *B. amyloliquefaciens* FZB42 and its mutant derivatives. Primer extension analysis was performed as described previously in detail (see Materials and Methods and Fig. 27). 1, *B. amyloliquefaciens* FZB42; 2, AK59 (*rapX*::Cm^r); 3, CH33 (*sigB*::Em^r); 4, AK57 (*rapX*::Cm^r *sigB*::Em^r). The absence of RapX increases P_{bmy} activity compared to that of the wild type, whereas introducing a *rapX* mutation on the *sigB* mutant strain leads to a de-repression of the P_{bmy} activity. This indicates that the positive effect of σ^B on *bmy* expression is mediated through the repression of RapX, which negatively regulates the promoter activity of P_{bmy} (see also text).

From the Raps known to inhibit the function of DegU or ComA in *B. subtilis*, only RapC and RapF are also encoded in *B. amyloliquefaciens* FZB42. Their cognate Phrs (PhrC and PhrF) are both controlled by σ^H in *B. subtilis* [234] and the σ^H-dependent promoter sites are conserved in *B. amyloliquefaciens* FZB42 (data not shown). This insinuates that the indirect effect of σ^H in *bmy* expression can be mediated through PhrC and PhrF (see also Discussion and Fig. 39).

Post-transcriptional effects in bacillomycin D production

Sfp and YczE control bacillomycin D production in a post-transcriptional manner

Nonribosomal peptide synthetases require posttranslational modification to be functionally active, as already mentioned in the introduction. Sfp, the 4'-phosphopantetheinyl transferase, catalyses conversion of the T-domain from its *apo*- to its *holo*-form and unblocks synthesis (see 1.3.2.3). As expected, MALDI-TOF MS analysis revealed that disruption of *sfp* gene in *B. amyloliquefaciens* FZB42 (strain CH3) resulted in deficiency in both lipopeptide and polyketide synthesis (Fig. 34B) [197].

Furthermore, *yczE* encodes a predicted membrane protein of unknown function and is located directly downstream of *sfp*. Disruption of the *yczE* gene in *B. amyloliquefaciens* FZB42 shut down the production of bacillomycin D (and polyketides; [197], whereas synthesis of fengycin and surfactin remained unimpaired, as demonstrated by MALDI-TOF MS analysis of culture filtrate extracts from the corresponding strains (Fig. 34C). Since YczE is similar to a membrane protein, its role could be associated with the export of bacillomycin

D. If this were the case, the antibiotic would be synthesized but would be entrapped inside the cell. To test this hypothesis, sonificated cell extracts of the *yczE* mutant strain were analyzed by MALDI-TOF mass spectrometry. Bacillomycin D was not detected inside the cells (data not shown) indicating that YczE is not involved in the antibiotic's export but rather in its synthesis in a yet unidentified manner.

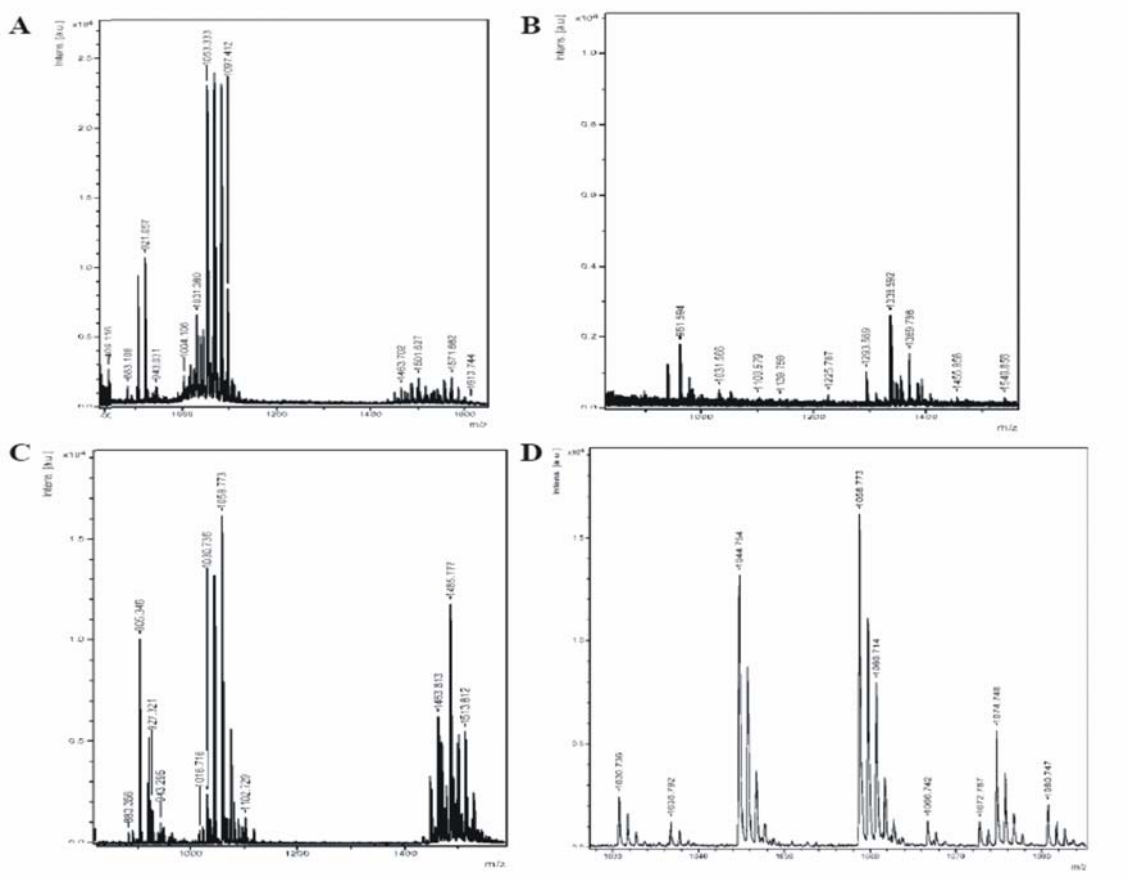

Figure 34: MALDI-TOF MS analysis of *sfp* and *yczE* mutant strains (performed in collaboration with Dr. J. Vater and Xiao-Hua Chen)

Spectra of culture filtrate extracts prepared from strains grown for 24 hours in Landy medium. **A**) *B. amyloliquefaciens* FZB42 produces the lipopeptides bacillomycin D (*m/z* 1053.7, 1097.4), surfactin (*m/z* 1030) and fengycin (*m/z* 1463, 1501) (see also Table 8 and Fig. 16). **B**. Introduction of an *sfp* mutation in *B. amyloliquefaciens* FZB42 causes deficiency in lipopeptide synthesis. **C**. A *yczE* mutant strain (CH4) produces surfactin (*m/z* 1030.7, 1058.7) and fengycin (*m/z* 1463.8, 1485.7) but no bacillomycin D. A zoomed in version of this spectrum is presented in panel **D**, where one can clearly see the peaks deriving from surfactin production, but not those expected for bacillomycin D production (*m/z* 1053.7, 1067.7, 1069.7).

To test whether the effect of YczE on bacillomycin D synthesis was exerted at the transcriptional level, the *yczE*::*Em*^{r} was introduced in strains AK9 and AK10 that carry

different sized *bmyD*::*lacZ* fusions (see also Fig. 22). The expression of the fusions was not influenced by the presence of the *yczE* mutation (Fig. 35). Therefore YczE does not have an impact on transcription of *bmy* operon but exerts its effect post-transcriptionally. The same conclusion was reached after checking the P_{bmy} activity in the *yczE* background by primer extension analysis (data not shown).

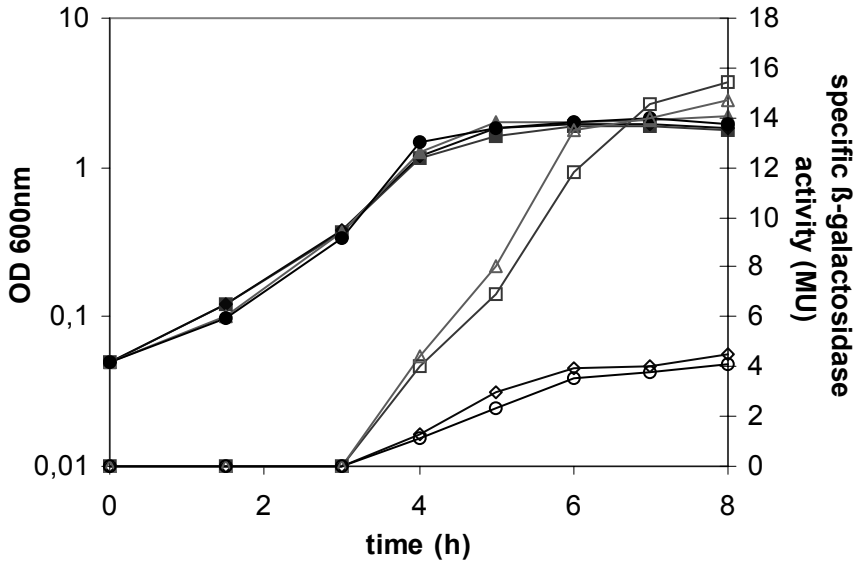

Figure 35: YczE does not influence the expression of the *bmy* operon

Strains AK9 and AK10, harbouring *lacZ* fusions of the *bmyD* upstream region between -400 and +126 bp, and between -183 and +126 bp, respectively (see also Fig. 21), and their *yczE* mutant derivatives (AK26, AK27), were grown in Difco medium at 37°C and optical densities (closed symbols) and specific β-galactosidase activities (in Miller Units; open symbols) were determined along the growth curve. Promoter activity was not altered by the presence of a *yczE* mutation. Squares, parental strain AK9; diamonds, parental strain AK10; triangles, AK26, *yczE* mutant derivative of AK9; cycles, AK27, *yczE* mutant derivative of AK10.

The post-transcriptional effect of DegU on *bmy* production is not mediated through YczE

It has already been demonstrated that both DegU and YczE influence bacillomycin D synthesis in a post-transcriptional manner. Since DegU is a response regulator controlling various post-exponential phase responses, it was postulated that it could be involved in transcriptional activation of YczE. Thereby the post-transcriptional effect of DegU on bacillomycin D production could be also mediated through YczE. In order to test this scenario, the transcriptional site of *yczE* was mapped by primer extension analysis and the

promoter activity was measured in different genetic backgrounds; i.e. wild type strain (*B. amyloliquefaciens* FZB42), *degU* mutant (TF1) and *comA* (CH23) mutant strains. The latter strain was used as a control, since ComA was predicted not to affect *yczE* transcription, (remember that the disruption of *comA* does not abolish production of bacillomycin D; Fig. 28).

The primer extension analysis conducted in the wild type strain revealed three possible overlapping transcriptional starts for *yczE*, but no apparent promoter recognition sequences could be traced in front of them (Fig. 36). On the other hand a weaker transcriptional start is located directly upstream of these three signals (one and half turn of the DNA helix) and just downstream of a perfect extended -10 promoter recognition sequence for the housekeeping sigma factor σ^A (Fig. 36). We propose that this is the real transcriptional start and the other signals just derive from 5'-end processing of this main mRNA. Moreover, no significant difference in the intensity of the transcripts was detected in the *degU* and *comA* mutant strains (Fig. 36). These findings indicate that the post-transcriptional effects of DegU and YczE on bacillomycin D biosynthesis are mediated through different pathways (see also Discussion).

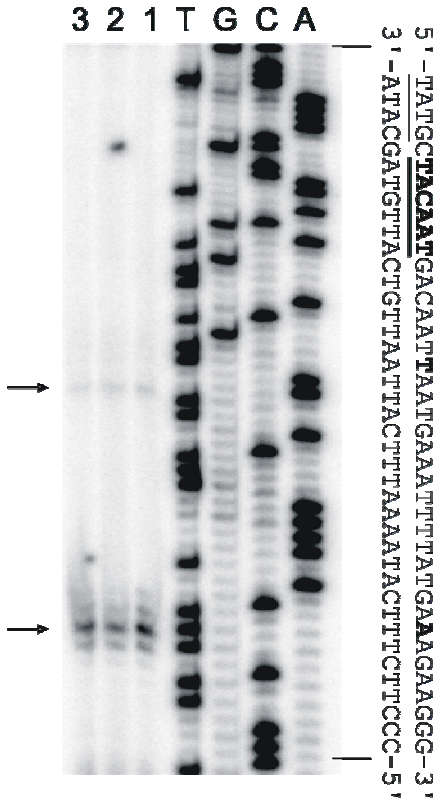

Figure 36: Mapping of the transcriptional start of *yczE* by primer extension analysis. DegU and ComA do not influence transcriptional initiation from the identified *yczE* promoter (P_{yczE})

Cells were grown and total RNA was prepared from them as described in detail in Fig. 24 and in Materials and Methods. Primer extension was performed using the 5'-end ^{32}P-labelled primer yczeu

(see also Table 4). The first four lanes result from dideoxynucleotide sequencing reactions using the same primer. The positions corresponding to the putative *yczE* transcriptional starts are indicated by arrows. The respective DNA sequences are shown in the outside lanes. Transcriptional starts are highlighted in bold face, the putative -10 hexamer is underlined and highlighted in bold face, whereas the putative extended -10 element is only underlined. Primer extension analysis performed in the wild type strain and its mutant derivatives showed no change in the promoter activity. 1, wild type *B. amyloliquefaciens* FZB42; 2, TF1 (*degU* $^{-}$); 3, CH23 (*comA*$^{-}$). Therefore DegU and YczE eliminate bacillomycin D production in *B. amyloliquefaciens* FZB42 using different pathways.

Global regulators affect the production of surfactin, fengycin and bacillibactin

The MALDI-TOF MS analysis enabled us to monitor the effects that transcriptional regulatory proteins and sigma factors exert on the production of other non-ribosomally synthesized compounds, apart from bacillomycin D. In the spectra obtained by culture filtrate extracts of *B. amyloliquefaciens* FZB42, fengycin, surfactin, bacillomycin D and the siderophore bacillibactin are visible with distinct peaks (see Table 8 and later on in this paragraph). Although, as already mentioned before, this type of mass spectrometry does not provide accurate quantitative results in respect with the concentrations of the four compounds, a rough estimation of their relative abundance can be achieved by comparing the intensity values of the peaks belonging to each compound with those belonging to the other three in the different genetic backgrounds. Thereby, after 24 hours of growth at 37^{o}C in Landy medium, bacillomycin D is the prominent lipopeptide synthesized by *B. amyloliquefaciens* FZB42, followed by bacillibactin and surfactin. Fengycin is the antibiotic with the lowest intensity values for its peaks and thus, the lowest abundance (see Figs. 16 and 28A).

Although the presence of the *dhb* operon in the genome of *B. amyloliquefaciens* FZB42 genome was known (see 3.2), it was uncertain whether the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester, **bacillibactin** [78], was functional. MALDI-TOF mass spectrometry of culture filtrate extracts prepared from *B. amyloliquefaciens* FZB42, grown for 24 hours at 37°C in Landy medium, verified the production of the compound. Peaks with mass numbers of 883, 905 and 921 *m/z* (Figs. 28A and 37A) correspond to the protonated $[M+H]^{+}$ form and the $[M+Na, K]^{+}$ alkali adducts of the compound, respectively [78].

The peaks attributed to bacillibactin were prominent in the spectrum of the wild type strain (Fig. 37A), showing higher intensity values than those of surfactin (and fengycin; see

Fig. 28A) but lower than the intensities of the bacillomycin D peaks. However, the production pattern of the three compounds changed significantly in the *degU* (TF1) and *sigW* (AK36) mutants (Fig. 37B and C). In the first case (Fig. 37B), surfactin peaks are significantly stronger than those of bacillibactin, in contrast to the wild type pattern (Fig. 37A), suggesting that the siderophore production is impaired in a *degU* $^{-}$ genetic background. In consistence with our results, microarray analysis combined with northern blot analysis and reporter fusions activity assays have previously shown that DegU is a positive regulator of the *dhb* operon in *B. subtilis* [253]. Moreover, the production of bacillibactin appeared also impaired in the *sigW* mutant strain (Fig 37C). There the intensity values of the peaks representing bacillibactin repeatedly dropped to the same levels (or lower) of their surfactin counterparts, whereas their difference to the bacillomycin D peaks increased compared to the wild-type strain (compare Figs. 37A and C). These findings suggest that σ^{w} might positively regulate the production and/or export of bacillibactin. Interestingly, stress conditions that provoke iron limitation for *Bacillus* (e.g. high salinity; [221]), trigger both DegU and σ^{w} mediated responses in order to aid the cell cope with the new conditions [254]. Iron limitation also triggers the production of bacillibactin [255], which is then secreted from the cell to work as an iron scavenger and then be re-imported into the cell, where its hydrolysis leads to release of cytosolic iron [256]. Therefore, a link between the three components, bacillibactin, DegU and σ^{w} seems plausible (see also Discussion).

On the other hand, **fengycin** levels appeared significantly elevated in many cases where bacillomycin D production was defected. The *comA* (CH23), *sigH* (AK50) and *sigB* (CH33) mutant strains exhibited this behaviour (Fig. 28.B-D and Table 11). However, both the *degU* (TF1) and *bmyD* (AK1) mutant did not show any significant increase in fengycin synthesis, even though they are completely unable to produce bacillomycin D (Figs. 18 and 28D). These two results implied that although the same regulatory pathways may differentially act in fengycin and bacillomycin D production, it is not the *per se* production of bacillomycin D that represses fengycin expression (see also Discussion).

Finally, most of the introduced mutations on transcriptional regulatory proteins and sigma factors had no apparent effect on the production of **surfactin**, as observed by MALDI-TOF MS analysis, since the ratios of the intensity values of surfactin peaks to that of other peptide antibiotics were the same for the mutant strains and the wild type. The *comA* mutant strain (CH23) was deficient in surfactin production (Fig 28B), as already described for *B. subtilis* [161, 162]. On the contrary, surfactin production was moderately pronounced in the

degU mutant strain, in accordance to the reported repressing effect of DegU on *srf* expression [253, 257].

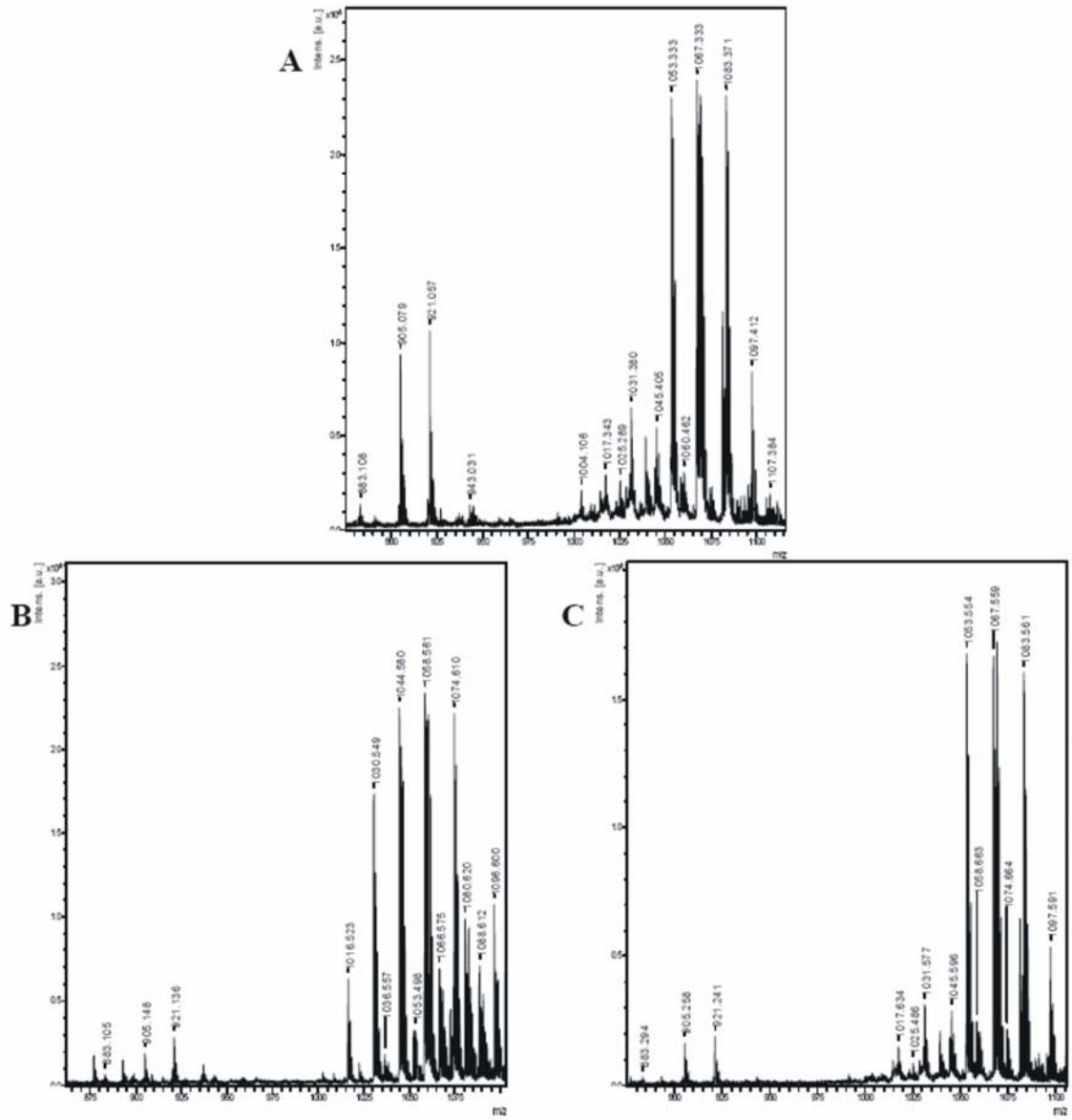

Figure 37: DegU and σ^W influence bacillibactin production (performed in collaboration with Dr. J. Vater).

Spectra of culture filtrate extracts prepared from strains grown for 24 hours in Landy medium. **A**. In the spectrum of *B. amyloliquefaciens* FZB42, the levels of bacillibactin (*m/z* 883, 905, 921) are lower than those of bacilllomycin D (*m/z* 1053, 1067, 1083) and clearly higher than those of surfactin (*m/z* 1060), as judged by the intensity of their peaks. The ratios of the intensity values between the three compounds' peaks change in the *degU* (**B**) and *sigW* (**C**) mutant strains at the cost of bacillibactin. As noticed before, bacillomycin D is not produced (and its peaks are absent) in a *degU* mutant strain (**B**), whereas the peaks belonging to surfactin (*m/z* 1044, 1058, 1074) have higher intensity than those of

bacillibactin. In the *sigW* mutant (**C**) the levels of bacillibactin are significantly lower than in the wild type strain (**A**) and are equivalent to those of surfactin (*m/z* 1058, 1074).

Table 11: MALDI-TOF MS analysis reveals increased production of fengycin in *comA*, *sigB* and *sigH* mutant strains of *B. amyloliquefaciens* FZB42

Lipopeptide	***m/z***	**Species**	**Intensity in:**			
			FZB42	**CH23**	**CH33**	**AK50**
Surfactin	1044.8	C14 $[M + Na]^+$	6000	-	6000	1900
	1058.8	C15 $[M + Na]^+$	3000	-	5000	1500
Bacillomycin	1053.7	C14 $[M + Na]^+$	23000	7200	19000	2600
D	1067.7	C15 $[M + Na]^+$	24000	4500	8000	6000
Fengycin	1485.9	C16 $[M + Na]^+$	1600	4000	22000	6000
	1449.9	C17 $[M + Na]^+$	1200	3200	18000	3400
Bacillibactin	905	$[M+Na]^+$	9500	4500	7500	1000
	921	$[M+K]^+$	11000	3500	6500	1000

Here are presented the intensities of the main peaks corresponding to the four compounds produced by *B. amyloliquefaciens* FZB42 and its *comA* (CH23), *sigB* (CH33) and *sigH* (AK50) mutant derivatives. The entire spectra can be seen in detail in Fig. 28. Note that the intensity values of the peaks are not comparable between the different MALDI-TOF spectra (this method does not offer quantitative results), but the relative production pattern of the four compounds in the different genetic backgrounds can be directly compared.

Discussion

Functional genomic analysis of B. amyloliquefaciens strain FZB42 reveals features of the bacterium that might be associated with its biocontrol activity

The first complete genome sequence of a Gram-positive bacterium was published in 1997 and belonged to the best characterized member of the group, i.e. *B. subtilis* strain 168. During the following years, complete genomes of various representatives of the *Bacilli* family were sequenced, such as that of the alkaliphilic bacterium *B. halodurans* (strain C-125, [258]), and those of the pathogenic bacteria *B. cereus* (strains ATCC 14579, [9]; ATCC 10987, [259] and E33L, accession number: NC_006274.1) and *B. anthracis* (strains Ames, [10] and Sterne, accession number: NC_005945). Meanwhile, the genome sequences of *B. clausii* KSM-K16 (accession number: NC_006582.1), *B. thuringiensis serovar konkukian* str.97-27 (NC_005957.1) and the industrially used bacterium *B. licheniformis* strain ATCC 14580 (note that the strain DSM13 is identical to ATCC 14580, [11, 260] have been completely defined.

However, until recently no strain of the *B. amyloliquefaciens* species had been completely sequenced. We have determined the complete genomic sequence of the plant growth promoting strain FZB42 of *B. amyloliquefaciens*, which is already commercially available as bio-fertilizer. Furthermore, we have tried to identify pathways that contribute to the organism's beneficial role.

Having the complete sequence of the organism in hand, functional and comparative genomic studies can be initiated to fully comprehend the lifestyle of *B. amyloliquefaciens* FZB42. These studies may ultimately lead to the development of new strategies to improve the properties of other plant growth promoting *Bacilli* or, even more, to construction of novel bio-fertilizers with more predictable action. Moreover, such comparative genomic analyses can provide a better understanding of evolution and phylogenetic relatedness among the *Bacillus* species.

Although the second round of annotation of the genome of *B. amyloliquefaciens* FZB42 has not yet been completed, most of the microorganism's genes have been assigned with a function by the first round and, therefore, can be compared to those of other sequenced *Bacilli*.

General features of the *B. amyloliquefaciens* FZB42 genome and comparison with genomes of other members of the *Bacillus* family

Recent taxonomic studies have shown that *B. amyloliquefaciens* is closely related to *B. subtilis* and *B. licheniformis*, based on comparisons of their 16S rDNA and 16S-23S internal transcribed spacer (ITS) nucleotide sequences [261]. However, the genome size of *B. amyloliquefaciens* FZB42 is 3,9 Mb and thus significantly smaller than the genomes of all other sequenced *Bacilli*, including that of its close relatives *B. subtilis* (4,214 Mb; [7] and *B. licheniformis* (4,222 Mb; [11, 260]; Table 12). Furthermore, preliminary data revealed the presence of 3931 genes in the *B. amyloliquefaciens* FZB42 genome, whereas 4112 genes and 4286 genes are present in the genomes of *B. subtilis* 168 and *B. licheniformis* ATCC 14580 [7, 11], respectively. Almost 80% of the *B. amyloliquefaciens* FZB42 genes show more than 50% homology at amino acid level to genes of *B. subtilis* 168.

Table 12: Features of the *B. amyloliquefaciens* FZB42 genome and comparison with genomes of other *Bacillus* species

Feature	*B. amyloliquefaciens* FZB42	*B. subtilis* 168	*B. licheniformis* ATCC 14580	*B. halodurans* C-125	*B. anthracis* Ames	*B. cereus* ATCC 14579
Chromosome size (kb)	3916	4214	4222	4202	5227	5427
G+C content (%)	46	43,5	46,2	43,7	35,4	35,4
Number of genes	3931	4112	4286	4006	5508	5642
rRNA operons	11	10	7	8	11	12
Transposase genes	5	0	10	93	18	10

B. subtilis 168 [7], *B. licheniformis* ATCC 14580 [11, 260], *B. halodurans* C-125 [258], *B. anthracis* Ames [10] and *B. cereus* ATCC 14579 [9]

Horizontal gene transfer

The unique genes of *B. amyloliquefaciens* FZB42 were found distributed in at least 14 DNA islands and islets around the whole genome. Interestingly, most of the DNA islands are situated in the same genetic locus where prophage-like elements are found in the *B. subtilis* 168 genome, indicating that those regions are susceptible to genetic rearrangements. This is the case for the first two DNA islands present in *B. amyloliquefaciens* FZB42. Both of them are inserted in the position of prophage-like elements 1 and 2 from *B. subtilis* 168, where they have replaced the majority of genes. The genome of *B. amyloliquefaciens* FZB42 includes only part of the prophage-like elements found in *B. subtilis* 168, but in addition it contains three transposases in different copy numbers (see Table 7). Interestingly, an insertion sequence element showing 51% homology to the *IS3Bli1* element of *B. licheniformis* ATCC 14580, appears in three copies in the genome of *B. amyloliquefaciens* FZB42. In contrast, *B. licheniformis* ATCC 14580 has nine copies of the *IS3Bli1* element and one more putative transposase that is closely related to a transposase previously identified in the *Thermoanaerobacter tengcogensis* genome (Table 12, [260]). *B. subtilis* 168 does not contain any transposases, and it is assumed that horizontal gene transfer is mainly achieved by bacteriophages [7]. Obviously, this is not the case for *B. amyloliquefaciens* FZB42 and the other sequenced *Bacillus* species. An extreme case is that of *B. halodurans*, which contains 93 transposase genes of IS elements (Table 12, [258]). On the other hand, the bacteriophages or the bacteriophage-like elements (SPβ, PBSX and the skin element) present in the genome of *B. subtilis* 168 are not found in the genomes of any other sequenced *Bacilli*, and that is also the case for *B. amyloliquefaciens* FZB42.

Signal transduction proteins

One of the most important and divergent features of microorganisms is their ability to receive and respond to different environmental signals. This ability to sense fluctuations in their environment defines the capacity of the microorganism to adapt and proliferate in its natural habitat. Signal transduction systems are responsible for such processes. **Two-component regulatory systems** (TCS) comprise a large family of signal transducing proteins that accomplish the task of monitoring, processing, and responding to a plethora of divergent environmental stimuli [262]. They usually consist of a membrane-bound sensor kinase, which sensing an environmental stimulus autophosphorylates at a specific histidine residue. Subsequently, this phosphate is transferred to a specific aspartate residue in the second

component of the system, the cytoplasmatic response regulator protein [263, 264]. Phosphoryl transfer to the Asp residue in the N-terminal receiver (REC) domain of the response regulator affects the properties of its C-terminal [265, 266, 267]. The activated response regulator initiates adaptive changes in behaviour, structure, or physiology of the cell, with most response regulators acting as transcriptional repressors or activators [263, 268].

TCS are widespread among prokaryotes. Thirty kinase-regulator pairs, that reside in an operon, have been found in *B. subtilis* 168 [7, 269]. Several of those two-component systems have been thoroughly characterized. BceRS, LiaRS, YxdJK and YvcPQ, located next to ABC transporter or transmembrane proteins, are the only TCS that control the cell envelope stress response in *B. subtilis* 168 [270, 271]. Only the biological function of BceRS and its neighbouring ABC transporter, BceAB, is known and that is to control the bacterium's resistance against the cell wall antibiotic bacitracin [270]. BceRS, senses the presence of the toxic compound and activates the expression of the ABC transporter BceAB, which consequently facilitates the removal of the antibiotic [270, 272].

The sensor histidine kinases BceS, LiaS, YxdK and YvcQ of *B. subtilis* 168 belong to the recently-introduced subfamily of **intramembrane-sensing histidine kinases** (IM-HK) [273]. These proteins have striking similarities in their overall domain organisation: they are relatively small (less than 400aa) and their N-terminal sensing domain consists of two deduced transmembrane helices with a spacing of less than 25 amino acids. Therefore the N-terminal domain is almost entirely buried in the cytoplasmic membrane, indicating that no extracellular stimulus is detected [270]. Moreover, the cytoplasmic transmitter domain harbors only the standard features characteristic for all histidine-kinases (HisKA, HATPase_c for kinase activity and in some cases the dimerization domain HAMP), but it lacks any additional domains that would allow signal detection within the cytoplasm [270]. A very recent screen, [273], searching for this group of histidine kinases in completely sequenced microbial genomes, revealed 147 intramembrane-sensing histidine kinases (out of 5000 sensor kinases) with the majority of them found in the *Firmicutes* phylum (110). One striking feature of all studied IM-HKs is their common physiological role: they all seem to sense cell envelope stress and regulate genes important for the cell membrane organisation and integrity, detoxification and virulence [273]. Furthermore, most of those IM-HKs are located, together with their partner response regulator, adjacent to genes encoding ABC transporters or conserved transmembrane proteins [273].

The genome sequence of *B. amyloliquefaciens* FZB42 harbours thirty-one gene pairs encoding classical TCS. Twenty-one of them are orthologues to respective systems in *B.*

subtilis 168, whereas ten are novel TCS that exhibit high similarity to respective systems of other bacteria from the *Firmicutes* phylum (Table 13). *B. amyloliquefaciens* FZB42 lacks nine two-component regulatory systems encoded by *B. subtilis* 168 (YbdKJ, YcbLM, YccHG, YesMN, YfiJK, YkoGH, YvcPQ, YvfUT, YxjML). Among them, only YvcPQ is of known function, and that is associated with the sensing of cell envelope stress, as already mentioned above.

The analysis of the thirty-one sensor kinases present in the genome of *B. amyloliquefaciens* FZB42, using the simple modular architecture research tool (SMART, http://smart.embl-heidelberg.de; [274]), accompanied by searches for genomic context conservation and sequence homology, revealed that the bacterium encodes five potential IM-HK. Three of them (BceS, YxdK and LiaS) are direct orthologues to corresponding proteins encoded in *B. subtilis* 168 (70%, 97% and 75% similarity on amino acid level respectively) and two (RBAM00197, RBAM03294) are novel proteins. *bceS*, *yxdK*, *liaS* and their cognate response regulators are localized in the same genomic context in *B. amyloliquefaciens* FZB42 as in *B. subtilis* 168, i.e. next to genes encoding ABC transporters (*ytsCD*, *yxdLM*) or a transmembrane protein (*yvqF*). The high degree of conservation of those TCS between *B. amyloliquefaciens* FZB42, *B. subtilis* 168 and *B. licheniformis* ATCC 14580, implies that their role in *B. amyloliquefaciens* FZB42 has also to do with sensing the cell envelope stress as it is the case for the other two organisms [270, 271, 275].

The histidine kinases RBAM00197 and RBAM03294 display all the structural characteristics of IM-HK. RBAM00197 (340 aa) and its cognate response regulator RBAM00196 are located next to genes that display similarity to ABC transporters (RBAM00198/00199) (Table 13). In addition, the RBAM00196/00197-RBAM00198/00199 system displays similarity to the BceRS-BceAB, YxdJK- YxdLM systems of *B. subtilis* 168, strengthening our prediction that it also plays a role in cell envelope stress response. Interestingly, these sequences comprise part of a 22 kb size DNA island inserted in the genome of *B. amyloliquefaciens* FZB42 at the position where the prophage-like element 1 is located in the *B. subtilis* 168 genome.

Table 13: Novel two-component regulatory systems in *B. amyloliquefaciens* FZB42

RBAM name/position of HK-RR	Protein name/accession number/ identities of closest homologue	RBAM name of neighbou-ring ABC transpo-rters	Protein name/accession number/identities of closest homologue to ABC transporters	Organism
00197 **(IM-HK)**/210-00196/211	BceS/O35044/26%-BceR/O34951/42%	00198-00199	BceA/O34697/46%-BceB/O34741/23%	*B. subtilis* 168
03294 **(IM-HK)**/332-003295/3321	AAP08928/39%-AAP08927/56%	03295-03296	AAP08927/85%	*B. cereus* ATCC14579/ *B. subtilis* 168
00207/221-00208/222	NP_348870/28%-NP_346816/56%	00209-00210-00211	NP_346812/66%-NP_346813/37%-NP_346814/32%	*Clostridium acetobutylicum* ATCC 824
00546/559-00545/558	EAO53032/64%-EAO53033/86%	-	-	*B. thuringiensis serovar israelensis* ATCC 35646
01839/1866-01840/1867	EAT23926/46%-EAT23927/55%	-	-	*Clostridium phytofermentans* ISDg
02015/2088-02014/2087	AAU24296/78%-AAU24295/75%	-	-	*B. licheniformis* ATCC 14580
03132/3166-03131/ 3165	NP_978181/50%-NP_978180/76%	03133-03134-03135	NP_978182/74%-NP_978183/48%-NP_978184/48%	*B. cereus* ATCC 10987
03180/3211-03181/3212	SpaK/AAB91593/54%-SpaR/AAB91594 80%	03182-03183-03184	SpaG/AAB91595/49%-SpaE/AB91596/50%-SpaF/AAB91597/74%	*B. subtilis* ATCC 6633

03606/3610-03605/3610	YP_085861/51%-YP_085862/58%	03607-03608-03609	YP_085858/39%-YP_085859/59%-YP_085860/86%	*B. cereus* E33L
03780/3771-03781/3772	MrsK2/CAB60253/98%-MrsR2/CAB60254/99%	03782-03783-03784	MrsF/CAB60255/99%-MrsG/CAB60256/98%-MrsE/CAB60257/99%	*Bacillus sp.* HIL-Y85

RBAM name is the protein name in *B. amyloliquefaciens* FZB42. The name of the proteins (omitting the RBAM prefix) and their positions, indicated in kb, are given in the first column. The two-component sensor histidine kinase protein (HK) is written first, followed by the two-component response regulator (RR). Similarities to the closest homologue are derived by BLASTX alignment and are indicated on amino acid level for the overall protein length. The closest homologue's protein name is given if it is known. Minus indicates the absence of a neighbouring ABC transporter to the two-component system. IM-HK; intramembrane histidine kinase

On the other hand, the two-component system RBAM03294/03295 comprises a two-gene insertion in a region that is conserved between the genomes of *B. amyloliquefaciens* FZB42 and *B. subtilis* 168. It is located next to *bmrA*, a multidrug ABC transporter that is functionally active in *B. subtilis* and is constitutively expressed throughout growth [276]. Therefore, it would be intriguing to check if the inserted two-component system has a functional link with the multidrug ABC transporter, and whether it alters the regulation of *bmrA* in *B. amyloliquefaciens* FZB42. Furthermore, a more general function of the IM-HK RBAM03294 in the cell envelope stress response should not be excluded.

The cell envelope is the first and major line of defence against threats from the environment. It gives the cell its shape, counteracts the high inner osmotic pressure and provides an important sensory interface and molecular sieve between a bacterial cell and its surroundings, mediating both information flow and controlled transport of solutes. Therefore, monitoring the cell envelope integrity and adequately changing its composition is critical for survival. *B. amyloliquefaciens* FZB42 possesses five candidate two-component systems involved in the cell envelope stress response, two of which are novel members among the sequenced bacteria, as mentioned above. The closely related bacterium *B. licheniformis* ATCC 14580 possesses only three TCS for the same scope, BceRS, LiaRS and YxdJK, shared by *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 [275]. *B. subtilis* 168 has an

additional TCS, YvcPQ, that is not found in the two other organisms. All these data indicate substantial overlap, but also a degree of differentiation between the three closely related bacteria in respect with their response to envelope stress. Different environmental cues trigger presumably distinct responses in the three bacteria, which allow them to adopt different strategies to survive in their natural habitat, the soil.

Apart from RBAM00196/00197 and RBAM03294/03295, *B. amyloliquefaciens* FZB42 has eight more novel two-component regulatory systems that show similarity to systems present in other *Bacilli* or *Clostridia* (Table 13). Five of them are located adjacent to novel putative ABC transporters. For example, the TCS RBAM03780/03781 is highly homologous to the MrsK2R2 proteins of *Bacillus* sp.HIL-Y85 (92% and 99% similarity on amino acid level respectively), which control the immunity against the lantibiotic mersacidin, produced by the strain [36, 37]. In parallel, the adjacent putative ABC transporter proteins RBAM03782-03784 are almost identical to the MrsFGE transport system (99%, 87% and 84% similarity on amino acid level respectively), which confers self-protection against mersacidin to the producer bacterium [36, 37]. Thereby, it can be assumed that *B. amyloliquefaciens* FZB42 is also immune to mersacidin. Taking into consideration that the combination of two-component and ABC transporter systems is characteristic of detoxification units, which can selectively sense a harmful for the cell compound and export it into the extracellular space [277, 278], we can postulate that the respective systems in *B. amyloliquefaciens* FZB42 control immunity against various antibiotics, produced either by *B. amyloliquefaciens* FZB42 itself or by other competing microorganisms. Thereby, these detoxification units provide FZB42 with defensive mechanisms in order to survive in a highly competitive environment such as the soil. This resistant capacity of *B. amyloliquefaciens* FZB42 makes the bacterium more competent of surviving in the plant roots, where it exerts its biocontrol activity.

A key issue for the proper functioning of a signal transduction system is its ability to balance the input signalling with the output response. This is thought to occur through regulation of the overall phosphorylation state of the system and/or through regulation of the activity of the output domain of the response regulator. The **Rap** (response regulator aspartate phosphatase) **phosphatases** are a conserved family of regulatory proteins that negatively influence many response regulators [234]. *B. subtilis* 168 encodes 11 Rap proteins, eight of which constitute operons with downstream *phr* genes [249]. However, the expression of *phr* genes is usually controlled by an additional σ^H-dependent promoter [234]. Pre-Phr is synthesized as a small protein with a putative signal peptide, which is cleaved and secreted as

a pentapeptide to the external milieu [250]. The Phr pentapeptide is imported again into the cells by an oligopeptide permease [251], and inhibits the activity of its cognate Rap protein [166]. The Rap proteins inhibit the action of the target response regulators either by dephosphorylating them [279] or by binding to the DNA-binding domain of the response regulator [167, 246].

B. amyloliquefaciens FZB42 has six Rap proteins that are also present in *B. subtilis* 168; RapA, RapC, RapF, RapB, RapD and RapJ (also the cognate Phr of the first three are conserved). Due to high similarity of the Rap proteins and their target response regulators in *B. amyloliquefaciens* FZB42 and *B. subtilis* 168, it can be assumed that the function of the Rap proteins is conserved in the two bacteria. Therefore, RapA, RapB probably have a negative influence on the initiation of sporulation, by dephosphorylating Spo0F [279], while RapC and RapF probably inhibit binding of ComA to its target genes [167, 246]. It is noteworthy that *B. amyloliquefaciens* FZB42 lacks orthologues of RapG and RapH that negatively regulate DegU in *B. subtilis* 168.

In addition, *B. amyloliquefaciens* FZB42 possesses three novel putative Rap proteins (Table 14). The novel Rap proteins contain tetratricopeptide (TRP) domains, similarly to other members of the Rap family [280]. The TRP domains are thought to be directly implicated in protein-protein interactions [281]. It is considered that TRP domains play an important role in the interaction between the Rap protein and its cognate Phr [250] and it is speculated that Rap proteins, whose inhibitory function is not associated with dephosphorylation of their target response regulators, bind to the target regulator through their TRP domains [167, 248]. Moreover, no cognate *phr* genes were identified downstream of the three novel *rap* genes of *B. amyloliquefaciens* FZB42. Interestingly, RBAM03282 is situated at the same genetic locus where RapG is in *B. subtilis* 168. However, RBAM03282 ows very low homology (maximum 27% on amino acid level) to all Rap proteins of *B. subtilis* 168 but relatively high homology to a putative Rap protein of *Bacillus licheniformis* ATCC 14580 (YP_080123), which has not been studied so far (48% on amino acid level). We have shown that RBAM03282 is involved in the regulation of the *bmyD* operon and, therefore, designated it with a new name, RapX (see chapter 3.4 and later on in Discussion). The functions of the two other novel Rap proteins remain to be identified.

Table 14: Novel Rap (response regulator aspartate phosphatase) proteins in the genome of *B. amyloliquefaciens* FZB42

Protein name	Position	Size (aa)	Protein name/accession number of closest homologue	Identities (aa level)	Organism
RBAM00430	462	358	RapH/P40771	152/353 (43%)	*B. subtilis* 168
RBAM02010	2082	378	RapA/Q00828*	189/379 (49%)	*B. subtilis* 168
RBAM03832 (RapX)	3830	371	Putative response regulator aspartate phosphatase/YP_080123	173/358 (48%)	*B. licheniformis* ATCC 14580

The position of the proteins is given in kb and the closest homologue is presented, as derived by BLASTX alignment. Similarities on amino acid level are indicated for the aligned part of the sequences. * The orthologous RapA protein of *B. subtilis* 168 is present in another position on the genome of *B. amyloliquefaciens* FZB42

Sigma factors

The model Gram-positive bacterium *B. subtilis* 168 has 17 σ factors [7], seven of which deal with extracytoplasmic functions and therefore are designated as ECF σ factors (σ^M, σ^V, σ^W, σ^X, σ^Y, σ^Z, σ^{ylaC}; [282]). On the other hand, *B. amyloliquefaciens* FZB42 possesses 16 σ factors with six of them being ECF σ factors. The two organisms retain conserved all their non-ECF σ factors. Five of the ECF σ factors are common between *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 (σ^M, σ^V, σ^W, σ^X, σ^{ylaC}), whereas the latter lacks $\sigma^Y$and σ^Z, but has in addition a novel putative ECF σ factor, Sig01.

ECF σ factors are typically regulated by a co-transcribed membrane-bound anti-sigma factor that keeps the sigma factor inactive, bound in the cell membrane [283]. Sig01 is not an exception of this rule, since a putative anti-sigma factor is located downstream of its coding region. σ^{01} and anti-σ^{01} of *B. amyloliquefaciens* FZB42 display low similarity on amino acid level to a novel ECF σ factor (Bli04171) and its cognate anti-sigma factor (Bli04170) (40% and 27%, respectively) found in *B.licheniformis* strain ATCC 14580 [11]. Recently it was shown that Bli04171 ECF σ factor (designated σ^{ecfH} hereafter) is part of the regulatory

network that controls the cell envelope stress response in *B. licheniformis* ATCC 14580, since its expression was induced seven- and five-fold after vancomycin and bacitracin treatment, respectively [275]. These results indicate that σ^{01} could be also involved in the cell envelope stress response of *B. amyloliquefaciens* FZB42. Until now, the only knowledge we have about the function of σ^{01} is that it does not control the expression of bacillomycin D (see chapter 3.4) and most probably the expression of all lipopeptides and polyketides produced by the strain (data not shown). Therefore, it would be intriguing to find out whether and how this novel ECF σ factor contributes to cell envelope stress response.

It is noteworthy that a core of five ECF sigma factors are conserved in *B. amyloliquefaciens* FZB42, *B. licheniformis* ATCC 14580 and *B. subtilis* 168 (σ^{M}, σ^{V}, σ^{W}, σ^{X}, σ^{ylaC}). *B. amyloliquefaciens* FZB42 has one additional ECF σ factor (Sig01), *B. subtilis* 168 two (σ^{Y} and σ^{Z}) [282] and *B. licheniformis* ATCC 14580 has three (σ^{Y}, σ^{efcG} and σ^{ecfH}) [275]. These findings indicate, once more again, regulatory divergence, but also a partial overlap between the three *Bacilli* in respect with their response to envelope stress. Interestingly, *B. halodurans* strain C-125 has 20 σ factors with only half of them conserved in *B. subtilis* 168 [258]. Eleven σ factors belong to the ECF family, but only one (σ^{W}) is homologous to the ECF σ factors of *B. subtilis* 168, indicating that its unique ECF σ factors regulate special mechanisms that allow the bacterium to live in an alkaline environment [258].

Competence genes

Genetic or natural competence is a physiological differentiation state in which bacteria are able to take up exogenous DNA from the medium. The molecular processes involved in the competence development in the model gram-positive bacterium *B. subtilis* have been studied extensively over the last decades. The establishment of competence requires at least 25 different genes, acting together in a finely intertwined cascade of signal transduction pathways and regulatory circuits, reviewed in [284]. *B. amyloliquefaciens* FZB42 is a natural competent strain (deviating from the transformation protocol published for *B. subtilis* 168 [205])was developed in this study, see Materials and Methods) and its genome contains orthologs of all genes involved in the development of competence in *B. subtilis* 168. In contrast, *B. licheniformis* ATCC 14580 is not naturally transformable due to the lack of a *comS* homologue and to a transposon insertion into the *comP* gene [260].

Despite that the majority of competence genes in *B. amyloliquefaciens* FZB42 are highly homologous to their counterparts of *B. subtilis* 168, the genes that control the competence quorum-sensing system of *B. amyloliquefaciens* FZB42 (*comQ*, *comX*, *comP*) exhibit low similarity to the respective genes of *B. subtilis* 168 (36%, 31% and 55%,

respectively). Such low sequence similarity of the competence quorum-sensing system has been already observed among various *Bacillus* isolates [285]. The genetic polymorphism extends through c*omQ*, c*omX* and the 5' two-thirds of *comP* [285], as it is the case in *B. amyloliquefaciens* FZB42. Furthermore, it was exhibited that this genetic variability is correlated with specificity in the quorum-sensing response, so that each pheromone is sensed only by its cognate receptor [286]. The quorum-sensing locus may have been introduced by horizontal transmission into a common ancestor of *Bacillus* strains and thereafter subjected to strong positive selection, which resulted into a dramatic sequence polymorphism and pheromone specificity [287].

In addition, *B. amyloliquefaciens* FZB42 was found to be competent in an earlier stage of growth than its closely related *B. subtilis* 168; the former showed increased transformation rates during mid to late exponential phase (see Materials and Methods), whereas the latter is known to become competent upon entry into stationary phase [204]. It is tempting to speculate that *B. amyloliquefaciens* FZB42 exhibits a distinct temporal regulation of its competence gene circuit from its sibling *B. subtilis* 168, apart from maintaining its specific pheromone (ComX)-modificator (ComQ) pair for initiating the competence process. Identifying the differentially regulated competence genes between the two organisms would be a future challenge, since it will permit the genetic manipulation of two organisms in order to modify/improve their DNA uptake, both in terms of yield and of chronological occurrence.

Secondary metabolites

B. amyloliquefaciens FZB42 encodes eight gene clusters which are responsible for the nonribosomal synthesis of secondary metabolites. These operons comprise 8% of the bacterium's genome and encode for peptide/polyketide antibiotics and a siderophore. We have verified the functionality of all eight gene clusters and we believe that the secondary metabolites produced enable *B. amyloliquefaciens* FZB42 to dominate over competing organisms within its natural environment and/or serve as signals that trigger cellular responses to the receiving organisms in the surrounding [288, 289].

In detail, *B. amyloliquefaciens* FZB42 is able to produce three distinct lipopeptide antibiotics: surfactin, fengycin and bacillomycin D. All three lipopeptides are synthesized nonribosomally according to the multicarrier thiotemplate mechanism (see introduction; review by [48]. **Surfactin** is encoded by the *srf* operon, which is also found in the genome of *B. subtilis* 168, a strain unable to produce lipopeptides or polyketides due to frameshift mutation on the *sfp* gene [174]. The chromosomal locus, as well as the organisation of the

genes and modules within the *srf*, operon are identical among the two bacteria; only the downstream-flanking genes of the *srf* operon vary. One of these genes is *aat*, a putative transcriptional regulator. No dramatic change in the production of surfactin or of other lipopeptides was observed when *aat* was deleted in strain FZB42 (data not shown). Moreover, the *aat* deletion had no effect on the transcriptional regulation of bacillomycin D (see chapter 3.4). However, we have no data about the effects caused by the *aat* deletion on the transcriptional regulation of surfactin, fengycin and the polyketides. Therefore, the putative function of *aat* in the regulation of lipopeptides and polyketides should be more closely examined.

Surfactin is known to provide antibacterial activity to the producer strain, since it can penetrate bacterial membranes and disturb their function [290]. In addition, it is essential for the swarming motility of the microorganism [132, 133, 134, 135], as well as for the formation of biofilms [136, 137]. Thereby, surfactin controls colonization of surfaces and can aid in acquisition of nutrients though its surface-wetting and detergent properties [291]. Recently, it was shown that surfactin is required for the development of aerial structures in the biofilms produced by *B. subtilis*, which resemble the fruiting-body formation by myxobacteria [288, 292]. Moreover, it was shown that the surfactin produced by *B. subtilis* acts antagonistically against *Streptomyces coelicolor* by inhibiting its development of aerial hyphae and spores [288]. Interestingly, surfactin did not inhibit the vegetative growth of *Streptomyces coelicolor*, as a typical antibiotic would do, but prevented a specific developmental process of *Streptomyces coelicolor* [288]. Therefore, surfactin protects *B. amyloliquefaciens* FZB42 against bacteria [197] and enables it to form biofilms, equipping thus the bacterium with powerful antagonistic advantages during surface colonization.

The *bmy* and *fen* operons are responsible for the biosynthesis of **bacillomycin D** and **fengycin** in *B. amyloliquefaciens* FZB42, respectively. These gene clusters are located at the same chromosomal locus with a distance of about 25 kb between them. Interestingly, the gene clusters directing the biosynthesis of bacillomycin L in *B. subtilis* A1/3 and iturin A (a lipopeptide with similar structure as bacillomycin D) in *B. subtilis* RB14 are situated at the same position as the *bmy* operon in *B. amyloliquefaciens* FZB42. In addition, the *pps* operon in *B. subtilis* 168, which is assigned to fengycin biosynthesis (despite of the strain's inability to produce it), as well as the *fen* operon in the producer *B. subtilis* strains F29-3 [222] and A1/3 [140], are located at the same genetic locus as the *fen* operon in *B. amyloliquefaciens* FZB42. On the other hand, the genome of *B. subtilis* ATCC 6633 contains the *myc* operon (directing the biosynthesis of mycosubtilin, an iturin-like lipopeptide) at the same position

that the *fen* operon occupies in strains F29-3 and A1/3 [63]. These findings indicate high degree of genetic flexibility in this region and suggest that additional nonribosomal peptide synthetases (NRPS) can be integrated in it either as an insertion or as a substitution of already existing NRPS operons.

Synthesis of bacillomycin D occurs according to the multicarrier thiotemplate mechanism. We have tried to verify the biosynthetic pathway of bacillomycin D by disrupting one by one the last six modules (in specific the respective adenylation domains) of the nonribosomal peptide synthetase and then by trying to identify the intermediate elongation variants (see Fig. 20 and chapter 3.3). However, the expected products could not be detected by MALDI-TOF MS analysis of neither culture filtrate extracts nor sonificated cell extracts. This indicates that only the full length lipopeptide is exerted from the cell, whereas the intermediate products are covalently attached to the multienzyme system, from which they can not be completely detached, even after sonification. A possible way to achieve detachment of the products from the enzymes would be reaction with a suitable thiol-compound, such as cysteine or cysteamine. Reaction with such a compound, under the appropriate conditions, could lead in the transfer of the thioester bound product onto the free thiol-group, rendering thus possible the identification of the obtained intermediate variants of bacillomycin D by MALDI-TOF MS. We are currently pursuing this issue further, in collaboration with Dr. J. Vater.

Bacillomycin D and fengycin inhibit the growth of various phytopathogenic fungi. Abolishment of each antibiotic led to decreased inhibition of the fungal growth, compared to the wild type strain; the effect of fengycin was smaller than that of bacillomycin D. Deletion of both antibiotics deprived *B. amyloliquefaciens* FZB42 of its antifungal abilities (see chapter 3.3). Thereby, we have demonstrated a synergistic action of both lipopeptide antibiotics against the target microorganism, a phenomenon previously described for secondary metabolites produced by actinomycetes. The synergistic activity of the antibiotics had been interpreted as an evolved adaptation mechanism of the producer organism in order to compete with other microorganisms and maintain its sessile lifestyle [293]. In the case of *B. amyloliquefaciens* FZB42, the level of fengycin production is considerably lower than that of bacillomycin D and thus the observed synergistic effect of the antifungal compounds was unexpected.

Interestingly, several of the mutant derivatives of *B. amyloliquefaciens* FZB42 have opposing effects on the production of bacillomycin D and fengycin. In particular, we have shown that the *comA*, *sigH* and *sigB* mutations reduce *bmy* expression by several-fold (Fig.

28A-D), whereas preliminary data obtained by MALDI-TOF MS analysis of the respective mutant strains show enhanced production of fengycin (see chapter 3.5). On the other hand, the *bmyD* (AK1) and *degU* (TF1) mutant strains, that completely lack bacillomycin D, did not display an elevated production of fengycin. These results suggest that the same regulatory pathways (and not itself the production of bacillomycin D) may opposingly direct the regulation of both antifungal compounds. The bacterium can, thereby, enhance the expression of fengycin in conditions where the expression of bacillomycin D is low. In this way, any single fengycin or bacillomycin D mutant retains a considerable inhibitory effect on fungal growth compared to the double mutant.

It is noteworthy that bacillomycin D and fengycin, in contrast to surfactin, have no effect on biofilm formation (data not shown; [137]). Recently, iturin A (that belongs to the same family of peptide antibiotics like bacillomycin D) was shown to inhibit sporulation of *Streptomyces scabies*, but not its growth [289]. This suggests that bacillomycin D and fengycin might have additional roles as secondary messengers.

The genome of *B. amyloliquefaciens* FZB42 contains three giant modular polyketide gene clusters (for details see [197]. The *bae* operon is responsible for the biosynthesis of **bacillaene**, a conjugated hexaene with a linear structure [294], whose chemical structure is still unknown. The *dif* gene cluster is devoted to the synthesis of **difficidin** and **oxydifficidin**, which are highly unsaturated 22-member macrolides with a rare phosphate group [295]. The third polyketide gene cluster is designated *pks2* and is involved in the synthesis of **macrolactin** (K.Schneider and Xiao-Hua Chen, unpublished results). Notably, this is the first time that the complete gene clusters involved in the biosynthesis of bacillaene and difficidin/oxydifficidin are defined. Modular organisation of the three *pks* clusters in *B. amyloliquefaciens* FZB42 revealed an unusual *trans*-AT architecture, which indicates that all PKS modules lack an AT domain and are complemented by ATs encoded on isolated genes [197]. This unusual *trans*-AT architecture was recently described for a polyketide synthase-peptide synthetase gene cluster of an uncultured bacterial symbiont of *Paederus* beetles [296].

B. subtilis 168 possesses only one large polyketide gene cluster, designated *pksX*. However, this strain is unable to produce the respective polyketide, due to a mutation in the *sfp* (4'-phosphopantetheinyl transferase) gene [174]. Therefore, until recently it was not known which polyketide is synthesized by the *pksX* cluster. We have demonstrated that *B. subtilis* OKB105, a *sfp*$^+$ derivative of *B. subtilis* 168, is able to produce bacillaene indicating that *pksX* directs synthesis of this polyketide [197].

Bacillaene and difficidin/oxydifficidin exhibit strong antibacterial activities, whereas macrolactin inhibits the growth of *B. megaterium* and *E. carotovora* only weakly [197]. Interestingly, bioautographs of the wild type strain and the *sfp* (CH3) mutant derivative of *B. amyloliquefaciens* FZB42 (that is deficient in lipopeptide and polyketide synthesis) on *B. megaterium* lawn, revealed the production of an antibacterial compound with unknown structure [197].

In addition, the genome of *B. amyloliquefaciens* FZB42 contains the *bac* operon that controls the synthesis of the dipeptide **bacilysin** [217]. Organisation and localization of the bac operon in the genomes of *B. amyloliquefaciens* FZB42 and *B. subtilis* 168 are identical. Recently it was shown in *B. subtilis* 168 that genes *bacDE* are involved in amino acid ligation and bacilysin immunity, respectively [297].

Bacilysin is active against a wide range of bacteria [218]. It was suggested that its antibacterial spectrum overlaps with that of the polyketide compounds bacillaene and difficidin [137]. However, bacilysin does not account for the remaining antibacterial compound detected in the *sfp* mutant derivative of *B. amyloliquefaciens* FZB42, as observed in bioautographs on *B. megaterium* lawn [298]. This is the first evidence for an additional ribosomally produced, antibacterial compound of *B. amyloliquefaciens* FZB42.

The last operon involved in the nonribosomal synthesis of a compound in *B. amyloliquefaciens* FZB42 is that of *dhb*. The *dhbACEBF* operon is involved in the synthesis of 2,3-dihydroxybenzoate (DHB) as well as its modification and esterification to the iron siderophore **bacillibactin** [78] that enables microorganisms to efficiently scavenge iron [220, 221]. Bacillibactin was detected in the culture filtrate extracts of *B. amyloliquefaciens* FZB42, verifying the functionality of the *dhb* operon (Fig. 37). Both the organisation and the localization of the operon are conserved between *B. amyloliquefaciens* FZB42 and *B. subtilis* 168.

Iron is an essential trace element for all bacteria [299]. In many aerobic, neutral or alkaline environments, Fe^{+2} is present in only suboptimal concentrations due to its low solubility. Microorganisms have therefore developed elaborate systems for scavenging iron from environmental sources. These systems frequently involve the synthesis of high-affinity chelators, their excertion into the environment, and the recapturing of the iron-loaded chelator via affinity transport systems [300, 301]. Similarly, iron limitation triggers the production of bacillibactin [255] in *Bacilli*, which is then secreted from the cell to act as an iron scavenger and then is re-imported into the cell, where its hydrolysis leads to release of cytosolic iron [256]. In a highly competitive environment, such as the plant rhizosphere, the microorganisms

that can make use of the environmental iron are more likely to survive. Therefore, it is possible that *B. amyloliquefaciens* FZB42 enhances plant growth by depriving soil pathogenic microorganims of iron, like already proposed for other plant growth promoting rhizobacteria (PGPR) [302].

In conclusion, the genome of *B. amyloliquefaciens* FZB42 contains eight operons that direct nonribosomal synthesis of three lipopeptides, three polyketides, one dipeptide and a siderophore. These compounds exhibit strong antifungal and antibacterial activities and enable the bacterium to survive in its natural environment. As *B. amyloliquefaciens* FZB42 colonizes the plant roots, it inhibits growth of phytopathogenic bacteria or fungi either by depriving them of the essential iron (through the action of bacillibactin) or by directly inhibiting their growth and/or certain of their developmental processes (through the actions of lipopeptides and polyketides). We must note that antibiotic activity is possibly not the only function of lipopeptides and polyketides produced by *B. amyloliquefaciens* FZB42. Surfactin is involved in intercellular signalling [288] and may be other secondary metabolites play also a role in interspecies communication and thereby affect the developmental pathway of a bacterium without influencing its vegetative growth. Until now, only preliminary studies have been performed with cocultivated bacteria, a situation that resembles more the natural settings.

Interestingly, *B. amyloliquefaciens* FZB42 does not produce most of the ribosomally synthesized peptide antibiotics that *B. subtilis* 168 does. The genome of *B. amyloliquefaciens* FZB42 does not contain the gene clusters of bacteriocins subtilosin (*sbo-alb*) and the SPβ proghage-encoded sublancin (see chapter 1.3.1). Moreover, the bacterium does not produce the antibiotic-like killing factor Skf (sporulation killing factor) or the toxic protein SdpC (sporulation delay protein) [303]. Notably, SdpC is present only in *B. subtilis* strains and orthologues of it have not been identified in other bacteria including all *Bacillus* species sequenced to date [223].

Recently, it was reported that the *sfp* derivative of *B. amyloliquefaciens* FZB42 inhibits the growth of a *sigW* deficient strain of *B. subtilis* as strongly as the wild-type strain FZB42 [223]. *B. amyloliquefaciens* FZB42 exhibited one of the strongest inhibitory effects on a *sigW* mutant of *B. subtilis*, among several members of the *Bacilli* family tested. This indicates that *B. amyloliquefaciens* FZB42 encodes ribosomally synthesized peptide(s) or toxic protein(s) with antibacterial function, as observed in experiments on *B. megaterium* lawn performed in our lab. The *ydbST* and *fosB* (*yndN*) genes, present also in the genome of *B. amyloliquefaciens* FZB42, contribute to resistance against these antimicrobial compound(s), albeit to a smaller

extent than σ^W [223]. Other members of the σ^W regulon could be also involved in promoting resistance against the ribosomally synthesized antibacterial compounds of *B. amyloliquefaciens* FZB42.

A complex network controls the expression of bacillomycin D in B. amyloliquefaciens FZB42

Bacillomycin D is a nonribosomally-synthesised heptapeptide with a β-amino fatty acid moiety that belongs to the same structural family of peptide antibiotics as iturin and mycosubtilin. Several studies have successfully elucidated the physicochemical and biological properties of several peptides that belong to this group [115, 304, 305]. Furthermore the mechanism of the compounds' synthesis has been documented and the multienzyme complexes responsible for the biosynthesis of mycosubtilin and iturin A have been identified and partially characterized [63, 101]. In contrast, neither the regulatory pathways that control the expression of the iturin-like lipopeptides, nor the mechanisms that govern their export into the surrounding milieu of the cell, have been studied until now.

In this study we have shown that the expression of bacillomycin D is driven by a stationary-phase induced σ^A promoter, P_{bmy}, in *B. amyloliquefaciens* FZB42 (Figs. 22B and 24). An identically organised promoter has been reported to control the expression of iturin A in *B. subtilis* RB14 [101], though the reported transcriptional start differs from the one we identify here (it is situated 1bp downstream). In addition, we have identified three global regulators, DegU, DegQ, ComA and two sigma factors σ^B and σ^H that positively influence the transcriptional activation of P_{bmy} in *B. amyloliquefaciens* FZB42, and a novel Rap protein that exerts a negative effect on P_{bmy}. Interestingly, P_{bmy} retains basal levels of activity even in the absence of the above-mentioned activators. Taking into consideration the strong similarity of the upstream regions between the promoters of *itu* and *bmy*, it would be not surprising if the same global regulators control the expression of iturin A.

The role of DegU on *bmy* expression and bacillomycin D production

DegU is a two-component system response regulator of the LuxR-FixJ family, whose members have a helix-turn-helix (HTH) structure at their C-terminus [306]. It is known to control many cellular processes, including exoprotease production, competence development,

motility and to trigger post-exponential-phase responses under growth limiting conditions [245, 307]. Recently, two genome-wide transcriptional profiling studies have been published for the role of DegU in *B. subtilis* [253, 257]. Although none of them directly compared the gene expression in the wild-type strain versus that of the isogenic *degU* mutant, an extensive regulon was identified for DegU. In addition, DegU has been associated with response of *B. subtilis* to high salinity [205, 254].

To our knowledge, this is the first time that DegU is demonstrated to play a *central* role in the regulation of a nonribosomally synthesized antibiotic. A series of in-vivo and in-vitro data demonstrated that DegU directly activates the expression of P_{bmy} (see also Figs. 26, 27 and 30). In addition, the results from the EMSA and the DNase I footprinting experiments (see also Figs. 30 and 31) coincided and pointed out that DegU retains two distinct DNA binding-sites at the *bmy* promoter. The first site, Site I, is located relative near to the transcriptional start, between bps -123 and -99 (relative to the transcriptional start), whereas the second one, Site II, is situated further upstream between -201 and -172 (see Figs 25 and 31). Binding of DegU to the latter upstream site is absolutely essential for the optimal activation of the promoter (see Figs. 22B and 26A). The existence of a third DegU binding-site that is located more upstream than –230 bps should not be excluded, since our in-vitro footprint data do not provide conclusive evidence for this region.

This is the second study to date, which has directly monitored the binding of DegU to a promoter by footprinting analysis. The protection that DegU offers to the DNA at its two binding-sites is quite weak, similarly to that exhibited in the previous study by Hamoen *et al.* (2000). On the contrary, strong hypersensitive sites can be observed adjacently to the two DNA binding-sites, implying that the binding of DegU to its sites rearranges the local DNA architecture, probably by inducing strong DNA-bending, constraint or even unwinding, which makes the DNA more accessible to DNase I attack. This correlates well to the role of DegU in the activation of the *comK* promoter [231]. Based on a series of data, Hamoen *et al.* proposed that DegU alters the shape of the ~ 4 DNA helixes that separate the tandem ComK boxes (possibly by unwinding and/or bending the DNA), and, thereby, facilitates the binding of ComK to them; ComK can then stimulate the transcription of its own promoter.

There are several reasons why the binding of DegU to the DNA only weakly protects the latter against DNase I attack. First, both in our experiments and the Hamoen *et al.* (2000) study, unphosphorylated DegU was used for the footprinting analysis. Although in many studies response regulators are used in their unphosphorylated form in order to demonstrate DNA-binding, the use of the phosphorylated response regulator can often result in more

distinct/extended regions being protected against DNAse I cleavage [162, 308]. I also performed the footprinitng analysis with phosphorylated DegU (after incubation with "cold" acetyl phosphate) and obtained very similar results to those of the unphosphorylated DegU. Even though incubation of a response regulator with acetyl phosphate should result in its phosphorylation, no direct proof can be provided whether transphosphorylation actually took place, without using radioactive acetyl phosphate. Nevertheless, experiments with unphosphorylated response regulator can provide important information on its DNA-binding ability as seen before in many cases, such as that of UhpA, ComA and Spo0A [162, 308].

Another reason for the weak protection patterns of DegU is the nature of its binding-sites. A/T-rich DNA regions, such as the DNA-binding-sites of DegU, are more curved and therefore less accessible to DNase I, even when the DNA is naked without any protein bound to it. Thus, the A/T-rich DNA-binding-sites appear protected even in the absence of their binding partner. Hydroxyl radical footprinting has given more clear results in such cases, and should be considered as an alternative method in future studies.

Despite the fact that in-vitro assays monitoring the binding of DegU to DNA promoter regions are limited [231, 244, 248], two possible motifs have been suggested as putative DegU recognition-sites [244, 245]. Shimane *et al* (2004) based on in-vivo data from the *aprE* and *comK* promoters proposed that DegU recognises an A/T-rich motif (either a tandem repeat of a 5-nucloetide sequence TAAAT or an inverted repeat of ATTTA-N7-TAAAT), whereas Dartois *et al.* (1998) based on in-vivo studies in the *wapA* promoter and an alignment of DegU-regulated promoters, proposed AGAA-N_{11}-TTCAG as the recognition site for DegU. Although none of these studies provides conclusive evidence and they are contradicting to each other, degenerate forms of the latter motif could be identified in the DegU protected regions at the *bmy* promoter region (both sites I and II; see also Fig. 25), whereas the A/T-rich motifs proposed by Shimane *et al* (2004) were part of the hypersensitive sites that were generated at the *bmy* promoter region upon addition of DegU in the DNase I footprints. In any case, further experimental evidence, involving extensive site-directed mutagenesis, will be required to identify the consensus sequence recognised by DegU in P_{bmy} and/or other promoters.

All previous studies which have carefully assessed the binding of DegU (always unphosphorylated DegU used) to different promoter regions (*comK* and *aprE*; [231, 244, 248]) have shown a picture similar to the one exhibited in this study (Fig. 30). Increasing amounts of DegU cause a gradual shift of the DNA fragment. In most cases, DNA binding-proteins that recognise defined motifs at the DNA and bind tightly to it produce distinct shifts

that their number reflects how many binding-sites are present at this DNA fragment. If there is not enough protein in the assay to fully occupy the DNA binding-site(s) then the bound (shifted) and the unbound DNA are in equilibrium. The pattern of the band-shift assays produced by DegU at the *bmy* promoter raise interesting mechanistic scenarios in respect with how DegU binds to the promoter and activates transcription. It seems plausible that initial DegU binding serves as an anchor to further recruit DegU molecules to the promoter. However till now, little is known about the multimerisation state of DegU when it binds to its target sites, or what the helix-turn-helix of each DegU molecule recognises as DNA binding-motif.

As mentioned above, the in-vivo data (see also Figs. 22 and 26) of this study pinpoint an upstream regulatory region as absolutely essential for the maximal activation of the P_{bmy} promoter by DegU and the rest of the global regulators identified here to be involved in the *bmy* expression (see also below; most of them are shown or proposed to mediate their effects indirectly, via DegU). Nevertheless, the in-vitro data (Figs. 30 and 31) suggest that DegU retains at least two DNA binding-sites. The first of them (site I) is centred in a region that is not shown to be able to activate *per se* the P_{bmy} promoter, i.e site I is included in AK10, which does not show a significant difference in its activity from AK11, which lacks the DegU recognition site I (Figs. 21, 22 and 25). On the contrary site II of DegU is located within the upstream DNA region that is necessary for the promoter activation (Figs. 21, 22 and 25). We propose that the binding of DegU to Site I triggers a sharp DNA bend directly downstream of it and thus enables the DegU bound to Site II to activate the promoter (Fig. 38). This is a rather common transcriptional activation mechanism.

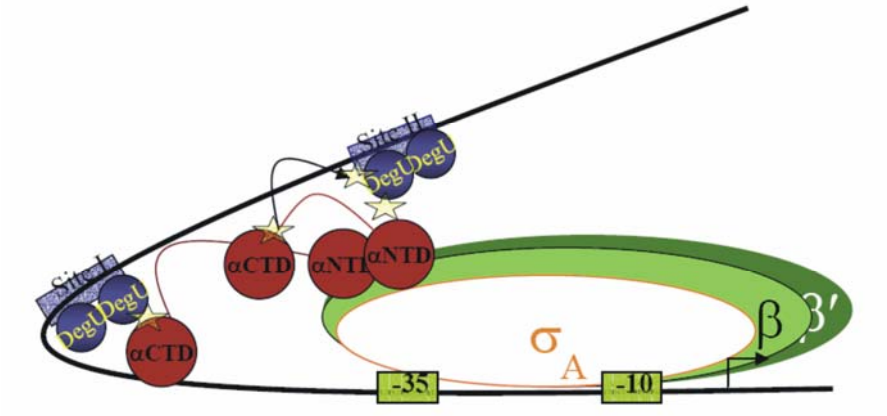

Figure 38: Proposed mechanism of action of DegU on the P_{bmy} promoter.

This figure illustrates how DegU might activate the function of the vegetative RNA polymerase (RNAP) on the P_{bmy} promoter. DegU is shown to bind to its identified DNA sites in dimers for presentational reasons (direct information for this is missing). Site I and site II are located between -

123 and -99 bps, and between -201 and -177 bps (relative to the transcriptional start), respectively. The DNA U-turn, shown directly downstream of site I, reflects to the strong hypersensitivity that this site exhibits in the DNase I footrpinting experiments. The DNA bending that possibly occurs directly upstream of site II (see DNase I footprints) is not shown in this picture for presentational reasons. The two C-terminal domains of the α subunit (designated as αCTD) are tethered with a flexible linker to the RNAP-bound N-terminal domains of the α subunit (designated as αNTD). Stars indicate possible interactions between RNAP and DegU. The binding of DegU to its site I alters the relative location of the DegU, bound to the site II, towards RNAP and renders the two in position to interact with each other.

It is generally accepted that DegU has two modes of action: *phosphorylated* DegU directly activates degradative enzyme production and represses motility, whereas *unphosphorylated* DegU directly stimulates competence [307] through binding to the *comK* promoter [231]. The belief that only *unphosphorylated* DegU is required for competence was supported by the observation that hyperphosphorylation of DegU (*degU32*(Hy) shows a 7-fold increase in the stability of the phosphorylated form of DegU) or inactivation of the *degSU* operon decreased competence, whereas inactivation of *degS* alone left competence unaffected. Moreover, a DegU mutant with an impaired phosphorylation site had no effect in competence [307]. However there are alternative explanations why the hyperactive form of DegU, or the complete absence of DegU, hinder competence, whereas the modest activity of the unphosphorylated DegU is enough to activate competence. DegU has opposing effects to different members of the DNA uptake gene-cascade. On one hand, it co-activates with ComK the *comK* promoter [231], but on the other hand it represses the *srf* operon [257] and therefore, also inhibits the expression of *comS*. Reduced ComS levels result into an enhanced MecA/ClpCP-mediated degradation of ComK [309]. Thus, it may well be that the final output of DegU on ComK is only positive, when the levels of DegU and/or its DNA binding affinity are relatively low (remember that the unphosphorylated form of response regulators has usually weaker binding affinity to its DNA targets; see above). On the contrary, when the cellular amounts or activity of DegU increase then the negative effect on *comS* expression prevails. Further evidence for such a scenario can be deduced by the genome-wide transcriptional profiling by Ogura *et al* (2001). In this study, the DegU regulon was identified by comparing a *degS* mutant strain with its isogenic strain (*degS* mutant too), having though DegU overexpressed from a plasmid. *comK* was not part of the induced genes, whereas all the phosphorylated DegU-dependent genes were. This insinuated that the phosphorylation state of

DegU alone does not dictate the targets of DegU. It is probably the DegU amounts and relative activity (which can be modulated by the phosphorylation state of the protein) that do so.

In our case, DegU activates the expression of the *bmy* operon during stationary phase growth, and therefore it seems plausible that the phosphorylated form of DegU is more suitable for optimal promoter binding and activation. However more direct evidence would be required for verifying this suggestion (see also above).

Finally, DegU seems to have a pronounced role in the synthesis of bacillomycin D since in its absence, *B. amyloliquefaciens* FZB42 is defected solely in the production of this peptide antibiotic (Fig. 28). This detrimental effect on bacillomycin D production cannot be only due to the reported effect of DegU on the P_{bmy} promoter, since several other regulators exert effects of similar extent on the promoter activity, but do not completely inhibit the synthesis of the antibiotic (Figs. 26-28). Two scenarios can explain this situation. First, it is possible that DegU also controls the activity of a second, yet unidentified, internal promoter in the *bmy* operon. In this case, the mutation of *degU* would be deleterious for bacillomycin D biosynthesis, since more than one promoter responsible for the expression of the *bmy* operon would be strongly hindered. However, till now, there are no reports about internal promoters regulating the expression of gene clusters encoding nonribosomal peptide synthetases. A second more plausible scenario would be that DegU is involved in the post-transcriptional regulation of bacillomycin D. DegU would then have to control the expression of a protein involved in the synthesis of bacillomycin D, but not in its export (no bacillomycin D was detected in sonificated cell extracts of the *degU*⁻ mutant). The possibility that this protein is Sfp should be ruled out, since DegU exerts a specific effect only on bacillomycin D. In contrast, production of surfactin, fengycin and the three polyketides was not impaired in the *degU* mutant stain. Moreover, DegU is not involved in the transcriptional regulation of *yczE*, which also controls the production of bacillomycin D in a post-transcriptional manner (see later and Fig. 36). Therefore, the putative post-transcriptional effect of DegU on the synthesis of bacillomycin D should be mediated through pathways independent of Sfp and YczE, and prior to the antibiotic's export out of the cell.

The role of DegQ on *bmy* expression

DegQ is a small pleiotropic regulatory protein, which consists of 46 amino acids and controls the expression of degradative enzymes, intracellular proteases and several other secreted enzymes (levansucrase, β-glucanase, xylanase, subtilisin and α-amylase) [171, 224]. Lately it

was also shown to stimulate the expression of several peptide antibiotics [172, 173]. DegQ shares no homology to typical transcriptional regulators, i.e. DNA-binding proteins. It may be located adjacently to the competence genes in the chromosome of different *Bacilli* organisms, but its function has been associated with that of DegU, with which they exhibit a significant target overlap [171]. In the absence of DegU, DegQ ceases to control the expression of *sacB* (encoding a levansucrase), implying that the effects of DegQ on *sacB* expression are indirect and mediated through DegU [171]. Our results show also that the effects of DegU are epistatic to those of DegQ on bacillomycin D production, since DegQ overexpression cannot complement for the loss of DegU in terms of bacillomycin D synthesis (see also Fig. 32). Thus, it seems plausible that DegQ regulates the transcription of its target genes only in an indirect manner, via DegU. DegQ possibly modulates the activity of DegU, via a yet unidentified mechanism. It is worth mentioning that DegQ shows homology to a region of the eukaryotic A-kinase anchor proteins (Dransfield *et al.*, 1997), and therefore a plausible role of it would be that it anchors DegS and facilitates the transphosphorylation to DegU.

Earlier studies had shown that *Bacilli* harbour two different versions of the σ^A-dependent promoter that is responsible for the transcription of *degQ*. *B. subtilis* 168 (and its derivative MO1099 used here) carry the degenerated promoter version, whereas *B. amyloliquefaciens* FZB42 possesses the optimised promoter version with a more consensus-like -10 hexamer, designated as *degQ36*(Hy) (for more details see corresponding text in results). Consistently, strains that carry the *degQ36*(Hy) show more prominent production of the enzymes DegQ regulates [225]. I have shown here that supplying the defected on *degQ* expression, *B. subtilis* MO1099, with ectopically produced DegQ, results into a 3-fold increase in the activity of the bacillomycin D promoter (see also Fig. 23). Moreover, this increase could be observed only when both DegU recognition sites were intact in the promoter region, verifying that DegQ exerts its role on the promoter activity via the action of DegU. In addition, this has been the first time that *degQ* was demonstrated to have an effect on the *transcriptional* regulation of a nonribosomally synthesized antibiotic. However, this effect was not as pronounced as the effect of DegQ on the overall production levels of iturin A or plipastatin, where an increase of 8- to 10-fold was observed [172, 173]. This insinuates that DegQ has an additional post-transcriptional role on lipopeptide synthesis. Consistently, DegU seems to exert a post-transcriptional effect on *bmy* expression (see above), and therefore, the two proteins may act again as a "pair" in the post-transcriptional control of the bacillomycin D synthesis.

The role of ComA on *bmy* expression

A further player that positively influenced the expression of bacillomycin D was the two-component system response regulator **ComA** (see also Figs 26 and 27). ComA is known to be involved in the regulation of several central developmental processes in the cell. Phosphorylated ComA activates the promoter of the *srf* operon [162], which encodes the enzyme complex that catalyzes the synthesis of the surfactin and also the competence regulation factor ComS, that lies within and out-of-frame in the *srfAB* gene. Consequently, ComS destabilizes the ternary ComK/MecA/ClpC complex with which ComK is degraded [309], releasing, thereby, the competence transcription factor that acts as a key regulator element in the development of competence [310]. Thus, ComA triggers the expression of surfactin and that of late competence genes. In addition, ComA controls the expression of *rapA* [311], a phosphatase which negatively regulates the initiation of sporulation by dephosphorylating Spo0F [279]. *rapC* and *rapF*, are also activated by ComA, creating thus a negative feedback loop, since both Rap proteins inhibit the function of ComA [167, 311]. Finally, ComA has a crucial role in the activation of *degQ*, along with DegU, which shows a more subtle effect on this process [171].

Here, we have shown that ComA exhibits similar effects in the activity of P_{bmy} as DegU (Figs. 26 and 27). The effects of ComA were mostly dependent on the presence of an upstream DNA region (-342 to -126 bp, relative to the transcriptional start), again similarly to those of DegU, raising the possibility that the two proteins mediate their effects on P_{bmy} through the same pathway. Since DegU is shown to directly bind to a DNA-site within this region, and ComA controls the expression of DegQ [171], which presumably serves as an auxiliary factor to DegU (see above), it would be plausible that the effects of ComA on the expression of the *bmy* operon are indirect and mediated through the DegQ-DegU system (Fig. 39). It is noteworthy that two ComA-boxes are located upstream of the *degQ* gene in *B. amyloliquefaciens* FZB42, similarly to the situation in *B. subtilis* 168 (data not shown). In addition, ComA activates both *degQ* promoter versions [171], and therefore it would be expected to promote the expression of DegQ in *B. amyloliquefaciens* FZB42, too. Unfortunately, our attempts to verify the proposed indirect role of ComA, by constructing a *comA* deficient strain of *B. amyloliquefaciens* FZB42, with *degQ* being expressed from an IPTG-inducible promoter, were unsuccessful (see corresponding section in the Results).

However, the recognition sequences of ComA have already been identified and consist of a palindromic segments, termed as ComA-box, i.e. TTGCGG-N_4-CCGCAA [162, 312]. The centres of dyad symmetry of the ComA-boxes are separated by about 45 bp. A screen for

the above-mentioned motif did not reveal any putative ComA-binding sites at the *bmyD* upstream region. This supports our suggestion that ComA only indirectly controls the transcriptional regulation of the *bmy* operon. Nevertheless, further experimental proof has to be provided for this statement. Either band-shift assays or assaying the role of ComA (in the presence of *degQ* being expressed from an IPTG-inducible promoter) on *bmy* expression in *B. subtilis* MO1099 would tackle the problem.

It is worth mentioning that the effect of ComA on the final production of bacillomycin D was not as devastating as that of DegU (Fig. 28), indicating that the transcriptional control of DegU and ComA on *bmy* expression might be exerted through the same pathway, but this is not the case for the post-transcriptional effects on bacillomycin D production (Fig. 39).

The role of σ^B and σ^H on *bmy* expression

Two sigma factors were shown to positively regulate bacillomycin D transcription: $\boldsymbol{\sigma^H}$, the sporulation sigma factor [233] and regulator of late-growth activities [162], and $\boldsymbol{\sigma^B}$, the general stress sigma factor in *Bacilli* [237, 238]. Both of them stimulate the activity of the σ^A-dependent P_{bmy} promoter (Figs. 26 and 27). Their effects on *bmy* expression are of the same magnitude to those of DegU and ComA (Figs. 26 and 27), and are most probably exerted in an indirect manner, since there are no sequences in the *bmyD* promoter region that resemble the promoter consensus sequences of σ^H (AGGANNT-15-17bp-GAAT; [234]) and σ^B (GTTT-15-17bp-GGGWAW, where W stands for A/T; [239]).

Bacillomycin D production was not silenced in the absence of σ^H and σ^B, similarly to the *comA* deletion. This indicates that either these sigma factors act principally on ComA or that they just moderately modulate the activity of DegU and do not completely abolish it. Based on our results and on former studies, we propose that the effects of σ^H and σ^B are mediated through various Rap proteins that control the activities of ComA and DegU.

RapC/RapF/RapK and RapG/RapH have been shown to inhibit ComA and DegU, respectively, from binding to their target sequences, in *B. subtilis* 168 [167, 246, 247, 248, 249]. These Rap proteins directly bind to the C-terminally located DNA-binding domain of the two response regulators and, thereby, hinder their transcriptional regulatory function [279]. The activity of the above-mentioned five Rap proteins is inhibited by specific, adjacently encoded, Phr pentapeptides (see also section 4.1.3). Interestingly, *rap* and *phr* genes are co-transcribed by a σ^A-dependent promoter [312], while the *phr* genes are additionally controlled by a σ^H-dependent promoter [234].

B. amyloliquefaciens FZB42 also encodes *rapC*, *rapF*, and their cognate *phr* genes, but lacks orthologues of *rapK*, *rapG*, *rapH* and/or their cognate *phr* genes (see also section 4.1.3). In addition, the bacterium possesses three novel Rap proteins (see also Table 14), which do not have a cognate Phr partner. Based on studies performed in *B. subtilis* 168 [167, 246], and our results demonstrating that ComA positively regulates expression of bacillomycin D, it is very likely that the effect of σ^H on the transcriptional regulation of the *bmy* operon is mediated through RapC and RapF. Deletion of σ^H decreases expression of PhrC and PhrF, and, thereby, RapC and RapF can more efficiently inhibit ComA from activating the expression of bacillomycin D (Fig. 39).

In *B. subtilis* 168, the σ^B–controlled RghR [252] was recently shown to specifically repress *rapG* and *rapH* by directly binding to their promoter regions [249]. RghR has no effect on other Rap proteins of *B. subtilis* 168 [249]. Although *B. amyloliquefaciens* FZB42 lacks *rapG* and *rapH* orthologues, RghR binding-sites were found upstream of one of its novel *rap* members, *rapX* (see also section 3.4.5). We have shown that a deletion of *rapX* results in enhancement of the P_{bmy} promoter activity (Fig. 33), which indicates the participation of RapX in the antibiotic's complex regulatory circuit (Fig. 39). However, due to its low homology to any of the Rap proteins of *B. subtilis* 168, it remains unclear whether RapX inhibits ComA or DegU or both of them. If the target of RapX is DegU, then the presence of increased amounts of RapX (in a *sigB* mutant) do not completely silence the activity of DegU since the *sigB* deficient strain can still produce bacillomycin D (see also Fig. 28). RapX could either dephosphorylate its target response regulator(s) or bind to its DNA-binding site and inhibit its function. Even though there is no direct evidence, we postulate that ComA and DegU are inhibited by Rap proteins via the same mechanisms in *B. subtilis* 168 and in *B. amyloliquefaciens* FZB42, i.e. the Rap protein binds to the DNA-binding site of the response regulator and blocks its action.

Furthermore, a double *sigB rapX* mutation clearly derepressed the expression of bacillomycin D, which was defected in the *sigB* single mutant (Fig. 33). This indicates that the effect of σ^B is mediated through RapX. We presume that the intermediate link is RghR (Fig. 39), since *rghR* (and its promoter region) is highly conserved between *B. subtilis* 168 and *B. amyloliquefaciens* FZB42 and the promoter region of *rapX* carries optimal DNA binding-sites for RghR. Further experimental evidence will be required for our assumption to be verified. In addition, it seems plausible that RghR might repress further Rap proteins (that inhibit the function of DegU or ComA), since the derepression effect on *bmy* expression observed after introducing a *rapX* mutation on the *sigB* mutant strain was not complete. A

good candidate would be *RBAM00430*, which shows 43% similarity on amino acid level to RapH of *B.subtilis* 168, but preliminary searches for RghR binding-sites on its promoter region revealed only relatively degenerate motifs in comparison to the published consensus sequence of the RghR DNA binding-site [249].

Post-transcriptional control of bacillomycin D expression

Sfp and **YczE** were both shown to post-transcriptionally regulate the expression of bacillomycin D. The essentiality of Sfp on nonribosomal synthesis is already known and thereby, the strain's deficiency to produce lipopeptides and polyketides in a *sfp*$^-$ strain was expected. Surprisingly, the deletion of the adjacently located gene, *yczE*, encoding for a predicted membrane protein, specifically abolished the production of bacillomycin D (Fig. 34.C), even though the activity of the P_{bmy} promoter was not impaired (Fig. 35). YczE is not involved in the export of the lipopeptide into the external milieu, similarly to DegU, and it exerts its effects through a separate pathway than that of DegU (Fig. 36). Both DegU (see also above) and YczE exert distinct control over the expression and the synthesis of bacillomycin D than Sfp, and therefore their mechanism of action remains an issue for further research.

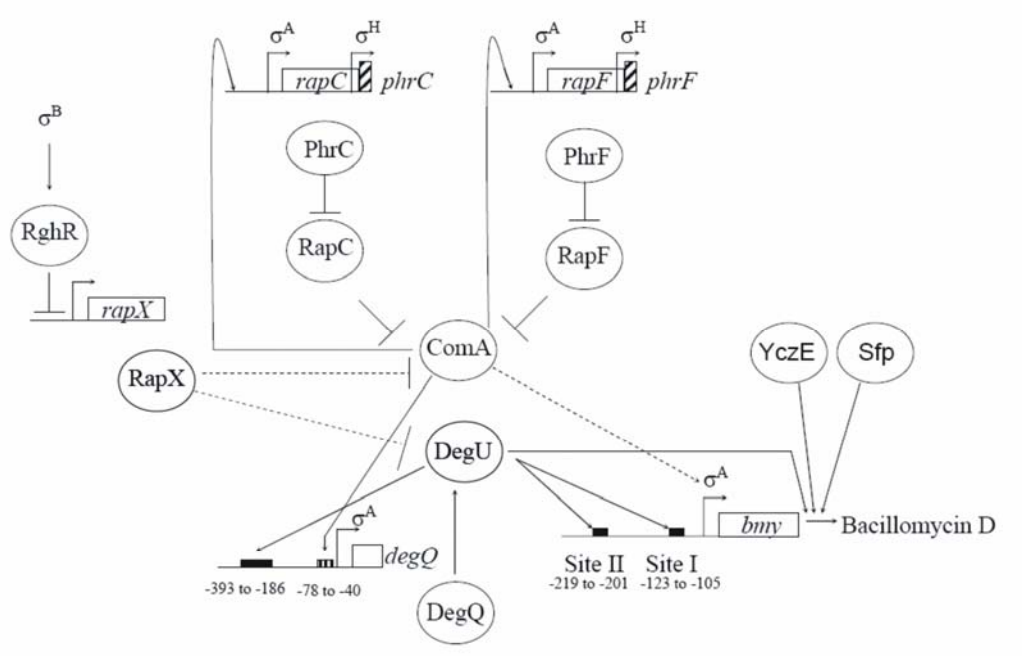

Figure 39: A complex regulatory network governs bacillomycin D production in *Bacillus amyloliquefaciens* strain FZB42

Boxes and cycles indicate ORFs and proteins respectively. Arrows and T-bars indicate activation and repression respectively. Interactions that have not been proven are represented by the dotted lines. σ^A and σ^H represent the promoters of the corresponding genes. Sites I and II are binding sites of DegU at the upstream region of *bmyD*.

Finally, the genes governing the export of bacillomycin D or securing the organism's immunity against the lipopeptide have not been identified yet. A novel TCS located upstream of the *bmy* operon (RBAM01839/RBAM01840) was investigated for its role on export of bacillomycin D (or other peptide antibiotics) and/or on self-resistance to *B. amyloliquefaciens* FZB42 against the antibacterial compounds produced by the strain. Deletion of this TCS did not impair the export of lipopeptides/polyketides, nor did the mutant strain show growth disadvantages when mixed with equal amounts of wild-type cells and let grow for several generations (data not shown). This indicates that the TCS RBAM01839/RBAM01840 is not involved in the release of lipopeptides/polyketides to the external milieu or in the self-resistance mechanisms.

Selbständigkeitserklärung

Hiermit versichere ich, die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Berlin, 11.10. 2006

Teile dieser Arbeit sind in folgenden Veröffentlichungen erhalten:

1. **Alexandra Koumoutsi**, Xiao-Hua Chen, Anke Henne, Heiko Liesegang, Gabriele Hietzeroth, Peter Franke, Joachim Vater and Rainer Borriss (2004) Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42, ***J.Bacteriol***, 186(4), 1084-96
2. Xiao-Hua Chen, Joachim Vater, Jörn Piel, Peter Franke, Romy Scholz, Katrin Schneider, **Alexandra Koumoutsi**, Gabriele Hietzeroth, Nicolas Grammel, Axel W. Strittmatter, Gerhard Gottschalk, Roderich Sussmuth and Rainer Borriss (2006) Structural and functional characterization of three Polyketide Sythase Gene Clusters in *Bacillus amyloliquefaciens* FZB42 ***J.Bacteriol***, 188, 4024-4036
3. **Alexandra Koumoutsi**, Xiao-Hua Chen, Joachim Vater and Rainer Borriss
   A complex regulatory network governs bacillomycin D expression in *Bacillus amyloliquefaciens* strain FZB42, in Vorbereitung.
4. Xiao Hua-Chen, **Alexandra Koumoutsi**, Romy Scholz, Axel Strittmatter, Heiko Liesegang, Gerhard Gottschalk, Oleg Reva, Helmut Junge, Rainer Borriss
   Complete genome sequence of the plant growth promoting *Bacillus amyloliquefaciens* strain FZB42, in Vorbereitung.
5. Xiao-Hua Chen, **Alexandra Koumoutsi**, Joachim Vater and Rainer Borriss
   More than anticipated - production of novel antibiotics by *Bacillus amyloliquefaciens* FZB42, Review inVorbereitung.
6. Patent title: Sequenz für die Bacillomycin D Synthese in *Bacillus amyloliquefaciens* FZB42
   Erfinder: Prof. Dr. Rainer Borriss, **Alexandra Koumoutsi**, Dr. Joachim Vater, Dr. Helmut Junge
   Anmeldetag: 14.12.2004
   Amtliches Aktenzeichen : DE 10326394.2 Int. Cl. C12N 15/31
7. Patent title: Sequences for the synthesis of antibacterial polyketides
   Erfinder: Prof. Dr. Rainer Borriss, Xiao-Hua Chen, **Alexandra Koumoutsi**, Dr. Joachim Vater
   Internationales Veröffentlichungsdatum: 08.12.2005
   Internationales Veröffentlichungsnummer: WO 2005/116061 A2

Lebenslauf

Name	Alexandra Koumoutsi
Geburtsdatum	12.10.1978
Geburtsort	Larisa/Griechenland
Staatsangehörigkeit	Griechisch
Familienstand	Ledig

Schulischer Werdegang

1984-990	3^{e} Grundschule Volos / Griechenland
1990–1993	3^{e} Gymnasium Volos / Griechenland
1990–1996	3^{e} Lyzeum Volos / Griechenland

Studium

10.1996–09.2001 Studium in der Fachrichtung Chemie an der Aristoteles Universität Thessaloniki/Griechenland, Diplomabschluss, Benotung: 7.3/10

02-09.2000 Diplomarbeit im Biochemischen Labor, Abteilung Biochemie und Biotechnologie, Aristoteles Universität Thessaloniki, unter die Betreuung von Prof. Dr. P. Arzoglou

Thema:“Studies on pancreatic lipase activity”

Benotung: 9/10

Wissenschaftlichen Aktivitäten

02.2001-06.2001 Freie Mitarbeiterin im Labor von Genetik und Biotechnologie, am Institut für Biologie, Universität Athenes, unter die Betreuung von Prof. Dr. M. Typas

Thema:“Mitochondrial gene polymorphisms in isolates of the phytogenic fungi of the genus *Verticillium*”

11.2001-10.2006 Promotion am Institut für Biologie, AG Bakteriengenetik, Prof. Dr.R. Borriss

Acknowledgements

I would like to thank Prof. Dr. Rainer Borriss for giving me the opportunity to work in his group and to spend five very creative years in his laboratory. I greatly appreciate his continuous advice and interest in my work, but also in my future career.

I would also like to thank Dr. Joachim Vater for his numerous mass spectrometric analyses and for stimulating scientific discussions. Apart from that, I am indebted to him for accepting to be a referee for this thesis, but also for always be willing to help me out with my future career.

I would like to thank Dr. Martin Meixner for his support with the SSH experiments.

Of course, I am grateful to all my colleagues from the "Bakteriengenetik" group for the pleasant atmosphere in the laboratory, numerous discussions and their continuous support. The time with them in the laboratory, but also outside of it, will always stay in my memory. I would especially like to thank Xiao-Hua Chen for the perfect collaboration in the sequencing project as well as for sharing with me several of her mutant strains. My special thanks to Christiana Müller and Sybille Striegl for their help in several experiments, but also for the continuous solving of several everyday problems in the lab; without them things could have only been more difficult for the completion of this thesis.

I want also to thank Christina Pesavento for helping me out with writing the german version of the summary in my thesis.

I would like to thank my parents, my sister and especially my dear Nassos.

Literature

[1] Harwood, Colin (1989): *Bacillus* (vol. 2), Biotechnology Handbooks Plenum Press.

[2] Ingle, M. B. and Erickson, R. J. (1978): Bacterial alpha-amylases, Adv Appl Microbiol (vol. 24), pp. 257-78.

[3] Fogarty, W.M. (1983): Microbial amylases, Fogarty, W.M., Microbial enzymes and Biotechnology pp. pp 1-92, Applied Science, London.

[4] M. Kilian, U.Steiner, B.Krebs, H.Junge, G.Schmiedeknecht, R.Hain (2000): FZB24 *Bacillus subtilis*-mode of action of a microbial agent enhancing plant vitality, 1/00, Pflazenschutz-Nachrichten Bayer. 72-93

[5] Krebs, B., Höding, B., Kübart, S. M., Workie, A., Junge, H., Schmiedeknecht, G., Grosch, R., Bochow, H. & Hevesi, M. (1998): Use of *Bacillus subtilis* as biocontrol agent. 1. Activities and characterization of *Bacillus subtilis* strains., J Plant Dis Prot (vol. 105), pp. 181-197.

[6] Idriss, E. E.; Makarewicz, O.; Farouk, A.; Rosner, K.; Greiner, R.; Bochow, H.; Richter, T. and Borriss, R. (2002): Extracellular phytase activity of *Bacillus amyloliquefaciens* FZB45 contributes to its plant-growth-promoting effect, Microbiology (vol. 148), No. Pt 7, pp. 2097-109.

[7] Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A. M.; Alloni, G.; Azevedo, V.; Bertero, M. G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Borriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S. C.; Bron, S.; Brouillet, S.; Bruschi, C. V.; Caldwell, B.; Capuano, V.; Carter, N. M.; Choi, S. K.; Codani, J. J.; Connerton, I. F.; Danchin, A. and et al. (1997): The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*, Nature (vol. 390), No. 6657, pp. 249-56.

[8] Fleischmann, R. D.; Adams, M. D.; White, O.; Clayton, R. A.; Kirkness, E. F.; Kerlavage, A. R.; Bult, C. J.; Tomb, J. F.; Dougherty, B. A.; Merrick, J. M. and et al. (1995): Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, Science (vol. 269), No. 5223, pp. 496-512.

[9] Ivanova, N.; Sorokin, A.; Anderson, I.; Galleron, N.; Candelon, B.; Kapatral, V.; Bhattacharyya, A.; Reznik, G.; Mikhailova, N.; Lapidus, A.; Chu, L.; Mazur, M.; Goltsman, E.; Larsen, N.; D'Souza, M.; Walunas, T.; Grechkin, Y.; Pusch, G.; Haselkorn, R.; Fonstein, M.; Ehrlich, S. D.; Overbeek, R. and Kyrpides, N. (2003): Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*, Nature (vol. 423), No. 6935, pp. 87-91.

[10] Read, T. D.; Peterson, S. N.; Tourasse, N.; Baillie, L. W.; Paulsen, I. T.; Nelson, K. E.; Tettelin, H.; Fouts, D. E.; Eisen, J. A.; Gill, S. R.; Holtzapple, E. K.; Okstad, O. A.; Helgason, E.; Rilstone, J.; Wu, M.; Kolonay, J. F.; Beanan, M. J.; Dodson, R. J.; Brinkac, L. M.; Gwinn, M.; DeBoy, R. T.; Madpu, R.; Daugherty, S. C.; Durkin, A. S.; Haft, D. H.; Nelson, W. C.; Peterson, J. D.; Pop, M.; Khouri, H. M.; Radune, D.; Benton, J. L.; Mahamoud, Y.; Jiang, L.; Hance, I. R.; Weidman, J. F.; Berry, K. J.; Plaut, R. D.; Wolf, A. M.; Watkins, K. L.; Nierman, W. C.; Hazen, A.; Cline, R.; Redmond, C.; Thwaite, J. E.; White, O.; Salzberg, S. L.; Thomason, B.; Friedlander, A. M.; Koehler, T. M.; Hanna, P. C.; Kolsto, A. B. and Fraser, C. M. (2003): The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria, Nature (vol. 423), No. 6935, pp. 81-6.

[11] Veith, B.; Herzberg, C.; Steckel, S.; Feesche, J.; Maurer, K. H.; Ehrenreich, P.; Baumer, S.; Henne, A.; Liesegang, H.; Merkl, R.; Ehrenreich, A. and Gottschalk, G. (2004): The complete genome sequence of *Bacillus licheniformis* DSM13, an

organism with great industrial potential, J Mol Microbiol Biotechnol (vol. 7), No. 4, pp. 204-11.

[12] Adams, M. D.; Celniker, S. E.; Holt, R. A.; Evans, C. A.; Gocayne, J. D.; Amanatides, P. G.; Scherer, S. E.; Li, P. W.; Hoskins, R. A.; Galle, R. F.; George, R. A.; Lewis, S. E.; Richards, S.; Ashburner, M.; Henderson, S. N.; Sutton, G. G.; Wortman, J. R.; Yandell, M. D.; Zhang, Q.; Chen, L. X.; Brandon, R. C.; Rogers, Y. H.; Blazej, R. G.; Champe, M.; Pfeiffer, B. D.; Wan, K. H.; Doyle, C.; Baxter, E. G.; Helt, G.; Nelson, C. R.; Gabor, G. L.; Abril, J. F.; Agbayani, A.; An, H. J.; Andrews-Pfannkoch, C.; Baldwin, D.; Ballew, R. M.; Basu, A.; Baxendale, J.; Bayraktaroglu, L.; Beasley, E. M.; Beeson, K. Y.; Benos, P. V.; Berman, B. P.; Bhandari, D.; Bolshakov, S.; Borkova, D.; Botchan, M. R.; Bouck, J.; Brokstein, P.; Brottier, P.; Burtis, K. C.; Busam, D. A.; Butler, H.; Cadieu, E.; Center, A.; Chandra, I.; Cherry, J. M.; Cawley, S.; Dahlke, C.; Davenport, L. B.; Davies, P.; de Pablos, B.; Delcher, A.; Deng, Z.; Mays, A. D.; Dew, I.; Dietz, S. M.; Dodson, K.; Doup, L. E.; Downes, M.; Dugan-Rocha, S.; Dunkov, B. C.; Dunn, P.; Durbin, K. J.; Evangelista, C. C.; Ferraz, C.; Ferriera, S.; Fleischmann, W.; Fosler, C.; Gabrielian, A. E.; Garg, N. S.; Gelbart, W. M.; Glasser, K.; Glodek, A.; Gong, F.; Gorrell, J. H.; Gu, Z.; Guan, P.; Harris, M.; Harris, N. L.; Harvey, D.; Heiman, T. J.; Hernandez, J. R.; Houck, J.; Hostin, D.; Houston, K. A.; Howland, T. J.; Wei, M. H.; Ibegwam, C.; Jalali, M.; Kalush, F.; Karpen, G. H.; Ke, Z.; Kennison, J. A.; Ketchum, K. A.; Kimmel, B. E.; Kodira, C. D.; Kraft, C.; Kravitz, S.; Kulp, D.; Lai, Z.; Lasko, P.; Lei, Y.; Levitsky, A. A.; Li, J.; Li, Z.; Liang, Y.; Lin, X.; Liu, X.; Mattei, B.; McIntosh, T. C.; McLeod, M. P.; McPherson, D.; Merkulov, G.; Milshina, N. V.; Mobarry, C.; Morris, J.; Moshrefi, A.; Mount, S. M.; Moy, M.; Murphy, B.; Murphy, L.; Muzny, D. M.; Nelson, D. L.; Nelson, D. R.; Nelson, K. A.; Nixon, K.; Nusskern, D. R.; Pacleb, J. M.; Palazzolo, M.; Pittman, G. S.; Pan, S.; Pollard, J.; Puri, V.; Reese, M. G.; Reinert, K.; Remington, K.; Saunders, R. D.; Scheeler, F.; Shen, H.; Shue, B. C.; Siden-Kiamos, I.; Simpson, M.; Skupski, M. P.; Smith, T.; Spier, E.; Spradling, A. C.; Stapleton, M.; Strong, R.; Sun, E.; Svirskas, R.; Tector, C.; Turner, R.; Venter, E.; Wang, A. H.; Wang, X.; Wang, Z. Y.; Wassarman, D. A.; Weinstock, G. M.; Weissenbach, J.; Williams, S. M.; WoodageT; Worley, K. C.; Wu, D.; Yang, S.; Yao, Q. A.; Ye, J.; Yeh, R. F.; Zaveri, J. S.; Zhan, M.; Zhang, G.; Zhao, Q.; Zheng, L.; Zheng, X. H.; Zhong, F. N.; Zhong, W.; Zhou, X.; Zhu, S.; Zhu, X.; Smith, H. O.; Gibbs, R. A.; Myers, E. W.; Rubin, G. M. and Venter, J. C. (2000): The genome sequence of *Drosophila melanogaster*, Science (vol. 287), No. 5461, pp. 2185-95.

[13] Mural, R. J.; Adams, M. D.; Myers, E. W.; Smith, H. O.; Miklos, G. L.; Wides, R.; Halpern, A.; Li, P. W.; Sutton, G. G.; Nadeau, J.; Salzberg, S. L.; Holt, R. A.; Kodira, C. D.; Lu, F.; Chen, L.; Deng, Z.; Evangelista, C. C.; Gan, W.; Heiman, T. J.; Li, J.; Li, Z.; Merkulov, G. V.; Milshina, N. V.; Naik, A. K.; Qi, R.; Shue, B. C.; Wang, A.; Wang, J.; Wang, X.; Yan, X.; Ye, J.; Yooseph, S.; Zhao, Q.; Zheng, L.; Zhu, S. C.; Biddick, K.; Bolanos, R.; Delcher, A. L.; Dew, I. M.; Fasulo, D.; Flanigan, M. J.; Huson, D. H.; Kravitz, S. A.; Miller, J. R.; Mobarry, C. M.; Reinert, K.; Remington, K. A.; Zhang, Q.; Zheng, X. H.; Nusskern, D. R.; Lai, Z.; Lei, Y.; Zhong, W.; Yao, A.; Guan, P.; Ji, R. R.; Gu, Z.; Wang, Z. Y.; Zhong, F.; Xiao, C.; Chiang, C. C.; Yandell, M.; Wortman, J. R.; Amanatides, P. G.; Hladun, S. L.; Pratts, E. C.; Johnson, J. E.; Dodson, K. L.; Woodford, K. J.; Evans, C. A.; Gropman, B.; Rusch, D. B.; Venter, E.; Wang, M.; Smith, T. J.; Houck, J. T.; Tompkins, D. E.; Haynes, C.; Jacob, D.; Chin, S. H.; Allen, D. R.; Dahlke, C. E.; Sanders, R.; Li, K.; Liu, X.; Levitsky, A. A.; Majoros, W. H.; Chen, Q.; Xia, A. C.; Lopez, J. R.; Donnelly, M. T.; Newman, M. H.; Glodek, A.; Kraft, C. L.; Nodell, M.; Ali, F.; An, H. J.; Baldwin-Pitts, D.; Beeson, K. Y.; Cai, S.; Carnes, M.; Carver, A.; Caulk, P. M.; Center, A.; Chen, Y. H.; Cheng,

M. L.; Coyne, M. D.; Crowder, M.; Danaher, S.; Davenport, L. B.; Desilets, R.; Dietz, S. M.; Doup, L.; Dullaghan, P.; Ferriera, S.; Fosler, C. R.; Gire, H. C.; Gluecksmann, A.; Gocayne, J. D.; Gray, J.; Hart, B.; Haynes, J.; Hoover, J.; Howland, T.; Ibegwam, C.; Jalali, M.; Johns, D.; Kline, L.; Ma, D. S.; MacCawley, S.; Magoon, A.; Mann, F.; May, D.; McIntosh, T. C.; Mehta, S.; Moy, L.; Moy, M. C.; Murphy, B. J.; Murphy, S. D.; Nelson, K. A.; Nuri, Z.; Parker, K. A.; Prudhomme, A. C.; Puri, V. N.; Qureshi, H.; Raley, J. C.; Reardon, M. S.; Regier, M. A.; Rogers, Y. H.; Romblad, D. L.; Schutz, J.; Scott, J. L.; Scott, R.; Sitter, C. D.; Smallwood, M.; Sprague, A. C.; Stewart, E.; Strong, R. V.; Suh, E.; Sylvester, K.; Thomas, R.; Tint, N. N.; Tsonis, C.; Wang, G.; Wang, G.; Williams, M. S.; Williams, S. M.; Windsor, S. M.; Wolfe, K.; Wu, M. M.; Zaveri, J.; Chaturvedi, K.; Gabrielian, A. E.; Ke, Z.; Sun, J.; Subramanian, G.; Venter, J. C.; Pfannkoch, C. M.; Barnstead, M. and Stephenson, L. D. (2002): A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome, Science (vol. 296), No. 5573, pp. 1661-71.

[14] Dunham, I.; Shimizu, N.; Roe, B. A.; Chissoe, S.; Hunt, A. R.; Collins, J. E.; Bruskiewich, R.; Beare, D. M.; Clamp, M.; Smink, L. J.; Ainscough, R.; Almeida, J. P.; Babbage, A.; Bagguley, C.; Bailey, J.; Barlow, K.; Bates, K. N.; Beasley, O.; Bird, C. P.; Blakey, S.; Bridgeman, A. M.; Buck, D.; Burgess, J.; Burrill, W. D.; O'Brien, K. P. and et al. (1999): The DNA sequence of human chromosome 22, Nature (vol. 402), No. 6761, pp. 489-95.

[15] Madigan M.T., J.M. Martinko, J. Parker (2000): Microbial growth control, P.F.Corey, Brock biology of microorganisms, ninth. ed., Prentice Hall, New Jersey.

[16] Priest, F.G. (1989): Products and Applications, Harwood, C. R., *Bacillus* (vol. 2) pp. 303-305, Plenum Press, New York.

[17] Katz, E. and Demain, A. L. (1977): The peptide antibiotics of *Bacillus*: chemistry, biogenesis, and possible functions, Bacteriol Rev (vol. 41), No. 2, pp. 449-74.

[18] Hancock, R. E. and Chapple, D. S. (1999): Peptide antibiotics, Antimicrob Agents Chemother (vol. 43), No. 6, pp. 1317-23.

[19] Finking, R. and Marahiel, M. A. (2004): Biosynthesis of nonribosomal peptides, Annu Rev Microbiol (vol. 58), pp. 453-88.

[20] Sankaranarayanan, R.; Dock-Bregeon, A. C.; Romby, P.; Caillet, J.; Springer, M.; Rees, B.; Ehresmann, C.; Ehresmann, B. and Moras, D. (1999): The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site, Cell (vol. 97), No. 3, pp. 371-81.

[21] Eriani, G.; Delarue, M.; Poch, O.; Gangloff, J. and Moras, D. (1990): Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs, Nature (vol. 347), No. 6289, pp. 203-6.

[22] Ogle, J. M.; Carter, A. P. and Ramakrishnan, V. (2003): Insights into the decoding mechanism from recent ribosome structures, Trends Biochem Sci (vol. 28), No. 5, pp. 259-66.

[23] McAuliffe, O.; Ross, R. P. and Hill, C. (2001): Lantibiotics: structure, biosynthesis and mode of action, FEMS Microbiol Rev (vol. 25), No. 3, pp. 285-308.

[24] Sahl, H. G. and Bierbaum, G. (1998): Lantibiotics: biosynthesis and biological activities of uniquely modified peptides from gram-positive bacteria, Annu Rev Microbiol (vol. 52), pp. 41-79.

[25] Sahl, H. G.; Jack, R. W. and Bierbaum, G. (1995): Biosynthesis and biological activities of lantibiotics with unique post-translational modifications, Eur J Biochem (vol. 230), No. 3, pp. 827-53.

[26] Jung, G. (1991): Lantobiotics: a survey, Nisin and novel lantibiotics pp. 1-34, Escom, Leiden.

[27] Breukink, E. and de Kruijff, B. (1999): The lantibiotic nisin, a special case or not?, Biochim Biophys Acta (vol. 1462), No. 1-2, pp. 223-34.
[28] Brotz, H.; Bierbaum, G.; Reynolds, P. E. and Sahl, H. G. (1997): The lantibiotic mersacidin inhibits peptidoglycan biosynthesis at the level of transglycosylation, Eur J Biochem (vol. 246), No. 1, pp. 193-9.
[29] Banerjee, S. and Hansen, J. N. (1988): Structure and expression of a gene encoding the precursor of subtilin, a small protein antibiotic, J Biol Chem (vol. 263), No. 19, pp. 9508-14.
[30] Klein, C.; Kaletta, C.; Schnell, N. and Entian, K. D. (1992): Analysis of genes involved in biosynthesis of the lantibiotic subtilin, Appl Environ Microbiol (vol. 58), No. 1, pp. 132-42.
[31] Klein, C. and Entian, K. D. (1994): Genes involved in self-protection against the lantibiotic subtilin produced by *Bacillus subtilis* ATCC 6633, Appl Environ Microbiol (vol. 60), No. 8, pp. 2793-801.
[32] Klein, C.; Kaletta, C. and Entian, K. D. (1993): Biosynthesis of the lantibiotic subtilin is regulated by a histidine kinase/response regulator system, Appl Environ Microbiol (vol. 59), No. 1, pp. 296-303.
[33] Stein, T.; Borchert, S.; Kiesau, P.; Heinzmann, S.; Kloss, S.; Klein, C.; Helfrich, M. and Entian, K. D. (2002): Dual control of subtilin biosynthesis and immunity in *Bacillus subtilis*, Mol Microbiol (vol. 44), No. 2, pp. 403-16.
[34] Stein, T.; Borchert, S.; Conrad, B.; Feesche, J.; Hofemeister, B.; Hofemeister, J. and Entian, K. D. (2002): Two different lantibiotic-like peptides originate from the ericin gene cluster of *Bacillus subtilis* A1/3, J Bacteriol (vol. 184), No. 6, pp. 1703-11.
[35] Stein, T. (2005): *Bacillus subtilis* antibiotics: structures, syntheses and specific functions, Mol Microbiol (vol. 56), No. 4, pp. 845-57.
[36] Altena, K.; Guder, A.; Cramer, C. and Bierbaum, G. (2000): Biosynthesis of the lantibiotic mersacidin: organization of a type B lantibiotic gene cluster, Appl Environ Microbiol (vol. 66), No. 6, pp. 2565-71.
[37] Guder, A.; Schmitter, T.; Wiedemann, I.; Sahl, H. G. and 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 (vol. 68), No. 1, pp. 106-13.
[38] Paik, S. H.; Chakicherla, A. and Hansen, J. N. (1998): Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168, J Biol Chem (vol. 273), No. 36, pp. 23134-42.
[39] Dorenbos, R.; Stein, T.; Kabel, J.; Bruand, C.; Bolhuis, A.; Bron, S.; Quax, W. J. and Van Dijl, J. M. (2002): Thiol-disulfide oxidoreductases are essential for the production of the lantibiotic sublancin 168, J Biol Chem (vol. 277), No. 19, pp. 16682-8.
[40] Zheng, G.; Hehn, R. and Zuber, P. (2000): Mutational analysis of the *sbo-alb* locus of *Bacillus subtilis*: identification of genes required for subtilosin production and immunity, J Bacteriol (vol. 182), No. 11, pp. 3266-73.
[41] Zheng, G.; Yan, L. Z.; Vederas, J. C. and Zuber, P. (1999): Genes of the *sbo-alb* locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilosin, J Bacteriol (vol. 181), No. 23, pp. 7346-55.
[42] Kleinkauf, H. and von Dohren, H. (1990): Nonribosomal biosynthesis of peptide antibiotics, Eur J Biochem (vol. 192), No. 1, pp. 1-15.
[43] Stein, T.; Vater, J.; Kruft, V.; Otto, A.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; McDowell, R. and Morris, H. R. (1996): The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates, J Biol Chem (vol. 271), No. 26, pp. 15428-35.

[44] von Döhren, H.; Keller, U.; Vater, J. and Zocher, R. (1997): Multifunctional Peptide Synthetases, Chem Rev (vol. 97), No. 7, pp. 2675-2706.
[45] Marahiel, M. A. (1997): Protein templates for the biosynthesis of peptide antibiotics, Chem Biol (vol. 4), No. 8, pp. 561-7.
[46] Mootz, H. D.; Schwarzer, D. and Marahiel, M. A. (2002): Ways of assembling complex natural products on modular nonribosomal peptide synthetases, Chembiochem (vol. 3), No. 6, pp. 490-504.
[47] Stachelhaus, T. and Marahiel, M. A. (1995): Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis, FEMS Microbiol Lett (vol. 125), No. 1, pp. 3-14.
[48] Sieber, S. A. and Marahiel, M. A. (2005): Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics, Chem Rev (vol. 105), No. 2, pp. 715-38.
[49] Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G. and van Sinderen, D. (1993): Sequence and analysis of the genetic locus responsible for surfactin synthesis in Bacillus subtilis, Mol Microbiol (vol. 8), No. 5, pp. 821-31.
[50] Dieckmann, R.; Lee, Y. O.; van Liempt, H.; von Dohren, H. and Kleinkauf, H. (1995): Expression of an active adenylate-forming domain of peptide synthetases corresponding to acyl-CoA-synthetases, FEBS Lett (vol. 357), No. 2, pp. 212-6.
[51] May, J. J.; Kessler, N.; Marahiel, M. A. and Stubbs, M. T. (2002): Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases, Proc Natl Acad Sci U S A (vol. 99), No. 19, pp. 12120-5.
[52] Mootz, H. D. and Marahiel, M. A. (1997): The tyrocidine biosynthesis operon of Bacillus brevis: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains, J Bacteriol (vol. 179), No. 21, pp. 6843-50.
[53] Stachelhaus, T.; Mootz, H. D. and Marahiel, M. A. (1999): The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, Chem Biol (vol. 6), No. 8, pp. 493-505.
[54] Challis, G. L.; Ravel, J. and Townsend, C. A. (2000): Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, Chem Biol (vol. 7), No. 3, pp. 211-24.
[55] Stein, T.; Vater, J.; Kruft, V.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; Mc Dowell, R. and Morris, H. R. (1994): Detection of 4'-phosphopantetheine at the thioester binding site for L-valine of gramicidinS synthetase 2, FEBS Lett (vol. 340), No. 1-2, pp. 39-44.
[56] Gocht, M. and Marahiel, M. A. (1994): Analysis of core sequences in the D-Phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis, J Bacteriol (vol. 176), No. 9, pp. 2654-62.
[57] Schlumbohm, W.; Stein, T.; Ullrich, C.; Vater, J.; Krause, M.; Marahiel, M. A.; Kruft, V. and Wittmann-Liebold, B. (1991): An active serine is involved in covalent substrate amino acid binding at each reaction center of gramicidin S synthetase, J Biol Chem (vol. 266), No. 34, pp. 23135-41.
[58] Stachelhaus, T.; Huser, A. and Marahiel, M. A. (1996): Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases, Chem Biol (vol. 3), No. 11, pp. 913-21.
[59] Stachelhaus, T.; Mootz, H. D.; Bergendahl, V. and Marahiel, M. A. (1998): Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain, J Biol Chem (vol. 273), No. 35, pp. 22773-81.
[60] Belshaw, P. J.; Walsh, C. T. and Stachelhaus, T. (1999): Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis, Science (vol. 284), No. 5413, pp. 486-9.

[61] Linne, U. and Marahiel, M. A. (2000): Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization, Biochemistry (vol. 39), No. 34, pp. 10439-47.
[62] Ehmann, D. E.; Trauger, J. W.; Stachelhaus, T. and Walsh, C. T. (2000): Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases, Chem Biol (vol. 7), No. 10, pp. 765-72.
[63] Duitman, E. H.; Hamoen, L. W.; Rembold, M.; Venema, G.; Seitz, H.; Saenger, W.; Bernhard, F.; Reinhardt, R.; Schmidt, M.; Ullrich, C.; Stein, T.; Leenders, F. and Vater, J. (1999): The mycosubtilin synthetase of Bacillus subtilis ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase, Proc Natl Acad Sci U S A (vol. 96), No. 23, pp. 13294-9.
[64] Roy, R. S.; Gehring, A. M.; Milne, J. C.; Belshaw, P. J. and Walsh, C. T. (1999): Thiazole and oxazole peptides: biosynthesis and molecular machinery, Nat Prod Rep (vol. 16), No. 2, pp. 249-63.
[65] Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; Marshall, C. G.; Miller, D. A. and Patel, H. M. (2001): Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines, Curr Opin Chem Biol (vol. 5), No. 5, pp. 525-34.
[66] Marshall, C. G.; Hillson, N. J. and Walsh, C. T. (2002): Catalytic mapping of the vibriobactin biosynthetic enzyme VibF, Biochemistry (vol. 41), No. 1, pp. 244-50.
[67] Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A. and Walsh, C. T. (2001): Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases, Biochemistry (vol. 40), No. 24, pp. 7099-108.
[68] Trauger, J. W.; Kohli, R. M. and Walsh, C. T. (2001): Cyclization of backbone-substituted peptides catalyzed by the thioesterase domain from the tyrocidine nonribosomal peptide synthetase, Biochemistry (vol. 40), No. 24, pp. 7092-8.
[69] Miller, D. A.; Luo, L.; Hillson, N.; Keating, T. A. and Walsh, C. T. (2002): Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*, Chem Biol (vol. 9), No. 3, pp. 333-44.
[70] Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T. and Sieber, S. A. (2002): Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase, Biochemistry (vol. 41), No. 45, pp. 13350-9.
[71] Patel, H. M. and Walsh, C. T. (2001): In vitro reconstitution of the *Pseudomonas aeruginosa* nonribosomal peptide synthesis of pyochelin: characterization of backbone tailoring thiazoline reductase and N-methyltransferase activities, Biochemistry (vol. 40), No. 30, pp. 9023-31.
[72] Sieber, S. A. and Marahiel, M. A. (2003): Learning from nature's drug factories: nonribosomal synthesis of macrocyclic peptides, J Bacteriol (vol. 185), No. 24, pp. 7036-43.
[73] Grunewald, J.; Sieber, S. A. and Marahiel, M. A. (2004): Chemo- and regioselective peptide cyclization triggered by the N-terminal fatty acid chain length: the recombinant cyclase of the calcium-dependent antibiotic from *Streptomyces coelicolor*, Biochemistry (vol. 43), No. 10, pp. 2915-25.
[74] Scholz-Schroeder, B. K.; Soule, J. D.; Lu, S. E.; Grgurina, I. and Gross, D. C. (2001): A physical map of the syringomycin and syringopeptin gene clusters localized to an approximately 145-kb DNA region of *Pseudomonas syringae* pv. syringae strain B301D, Mol Plant Microbe Interact (vol. 14), No. 12, pp. 1426-35.

[75] Segre, A.; Bachmann, R. C.; Ballio, A.; Bossa, F.; Grgurina, I.; Iacobellis, N. S.; Marino, G.; Pucci, P.; Simmaco, M. and Takemoto, J. Y. (1989): The structure of syringomycins A1, E and G, FEBS Lett (vol. 255), No. 1, pp. 27-31.

[76] Vanittanakom, N.; Loeffler, W.; Koch, U. and Jung, G. (1986): Fengycin--a novel antifungal lipopeptide antibiotic produced by *Bacillus subtilis* F-29-3, J Antibiot (Tokyo) (vol. 39), No. 7, pp. 888-901.

[77] Shaw-Reid, C. A.; Kelleher, N. L.; Losey, H. C.; Gehring, A. M.; Berg, C. and Walsh, C. T. (1999): Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization, Chem Biol (vol. 6), No. 6, pp. 385-400.

[78] May, J. J.; Wendrich, T. M. and Marahiel, M. A. (2001): The *dhb* operon of *Bacillus subtilis* encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin, J Biol Chem (vol. 276), No. 10, pp. 7209-17.

[79] Weber, G.; Schorgendorfer, K.; Schneider-Scherzer, E. and Leitner, E. (1994): The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame, Curr Genet (vol. 26), No. 2, pp. 120-5.

[80] Haese, A.; Schubert, M.; Herrmann, M. and Zocher, R. (1993): Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing N-methyldepsipeptide formation in *Fusarium scirpi*, Mol Microbiol (vol. 7), No. 6, pp. 905-14.

[81] Becker, J. E.; Moore, R. E. and Moore, B. S. (2004): Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization, Gene (vol. 325), pp. 35-42.

[82] Pfeifer, E.; Pavela-Vrancic, M.; von Dohren, H. and Kleinkauf, H. (1995): Characterization of tyrocidine synthetase 1 (TY1): requirement of posttranslational modification for peptide biosynthesis, Biochemistry (vol. 34), No. 22, pp. 7450-9.

[83] Stachelhaus, T. and Walsh, C. T. (2000): Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase, Biochemistry (vol. 39), No. 19, pp. 5775-87.

[84] Clugston, S. L.; Sieber, S. A.; Marahiel, M. A. and Walsh, C. T. (2003): Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst, Biochemistry (vol. 42), No. 41, pp. 12095-104.

[85] Luo, L. and Walsh, C. T. (2001): Kinetic analysis of three activated phenylalanyl intermediates generated by the initiation module PheATE of gramicidin S synthetase, Biochemistry (vol. 40), No. 18, pp. 5329-37.

[86] Luo, L.; Burkart, M. D.; Stachelhaus, T. and Walsh, C. T. (2001): Substrate recognition and selection by the initiation module PheATE of gramicidin S synthetase, J Am Chem Soc (vol. 123), No. 45, pp. 11208-18.

[87] Luo, L.; Kohli, R. M.; Onishi, M.; Linne, U.; Marahiel, M. A. and Walsh, C. T. (2002): Timing of epimerization and condensation reactions in nonribosomal peptide assembly lines: kinetic analysis of phenylalanine activating elongation modules of tyrocidine synthetase B, Biochemistry (vol. 41), No. 29, pp. 9184-96.

[88] Konz, D. and Marahiel, M. A. (1999): How do peptide synthetases generate structural diversity?, Chem Biol (vol. 6), No. 2, pp. R39-48.

[89] Gehring, A. M.; DeMoll, E.; Fetherston, J. D.; Mori, I.; Mayhew, G. F.; Blattner, F. R.; Walsh, C. T. and Perry, R. D. (1998): Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*, Chem Biol (vol. 5), No. 10, pp. 573-86.

[90] Billich, A.; Zocher, R.; Kleinkauf, H.; Braun, D. G.; Lavanchy, D. and Hochkeppel, H. K. (1987): Monoclonal antibodies to the multienzyme enniatin synthetase. Production and use in structural studies, Biol Chem Hoppe Seyler (vol. 368), No. 5, pp. 521-9.
[91] Mofid, M. R.; Finking, R.; Essen, L. O. and Marahiel, M. A. (2004): Structure-based mutational analysis of the 4'-phosphopantetheinyl transferases Sfp from *Bacillus subtilis*: carrier protein recognition and reaction mechanism, Biochemistry (vol. 43), No. 14, pp. 4128-36.
[92] Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C. and Walsh, C. T. (1996): A new enzyme superfamily - the phosphopantetheinyl transferases, Chem Biol (vol. 3), No. 11, pp. 923-36.
[93] Walsh, C. T.; Gehring, A. M.; Weinreb, P. H.; Quadri, L. E. and Flugel, R. S. (1997): Post-translational modification of polyketide and nonribosomal peptide synthases, Curr Opin Chem Biol (vol. 1), No. 3, pp. 309-15.
[94] Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P. and Walsh, C. T. (1998): Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases, Biochemistry (vol. 37), No. 6, pp. 1585-95.
[95] Schwarzer, D.; Mootz, H. D.; Linne, U. and Marahiel, M. A. (2002): Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases, Proc Natl Acad Sci U S A (vol. 99), No. 22, pp. 14083-8.
[96] Schneider, A. and Marahiel, M. A. (1998): Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*, Arch Microbiol (vol. 169), No. 5, pp. 404-10.
[97] Schwecke, T.; Aparicio, J. F.; Molnar, I.; Konig, A.; Khaw, L. E.; Haydock, S. F.; Oliynyk, M.; Caffrey, P.; Cortes, J.; Lester, J. B. and et al. (1995): The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin, Proc Natl Acad Sci U S A (vol. 92), No. 17, pp. 7839-43.
[98] Doekel, S. and Marahiel, M. A. (2001): Biosynthesis of natural products on modular peptide synthetases, Metab Eng (vol. 3), No. 1, pp. 64-77.
[99] Weinig, S.; Hecht, H. J.; Mahmud, T. and Muller, R. (2003): Melithiazol biosynthesis: further insights into myxobacterial PKS/NRPS systems and evidence for a new subclass of methyl transferases, Chem Biol (vol. 10), No. 10, pp. 939-52.
[100] Garwin, J. L.; Klages, A. L. and Cronan, J. E., Jr. (1980): Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*, J Biol Chem (vol. 255), No. 24, pp. 11949-56.
[101] Tsuge, K.; Akiyama, T. and Shoda, M. (2001): Cloning, sequencing, and characterization of the iturin A operon, J Bacteriol (vol. 183), No. 21, pp. 6265-73.
[102] Nougayrede, J. P.; Homburg, S.; Taieb, F.; Boury, M.; Brzuszkiewicz, E.; Gottschalk, G.; Buchrieser, C.; Hacker, J.; Dobrindt, U. and Oswald, E. (2006): *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells, Science (vol. 313), No. 5788, pp. 848-51.
[103] Black, P. N. and DiRusso, C. C. (1994): Molecular and biochemical analyses of fatty acid transport, metabolism, and gene regulation in *Escherichia coli*, Biochim Biophys Acta (vol. 1210), No. 2, pp. 123-45.
[104] Marrakchi, H.; Zhang, Y. M. and Rock, C. O. (2002): Mechanistic diversity and regulation of Type II fatty acid synthesis, Biochem Soc Trans (vol. 30), No. Pt 6, pp. 1050-5.
[105] Marini, P.; Li, S. J.; Gardiol, D.; Cronan, J. E., Jr. and de Mendoza, D. (1995): The genes encoding the biotin carboxyl carrier protein and biotin carboxylase subunits of

Bacillus subtilis acetyl coenzyme A carboxylase, the first enzyme of fatty acid synthesis, J Bacteriol (vol. 177), No. 23, pp. 7003-6.
[106] Guchhait, R. B.; Polakis, S. E.; Dimroth, P.; Stoll, E.; Moss, J. and Lane, M. D. (1974): Acetyl coenzyme A carboxylase system of *Escherichia coli*. Purification and properties of the biotin carboxylase, carboxyltransferase, and carboxyl carrier protein components, J Biol Chem (vol. 249), No. 20, pp. 6633-45.
[107] Harder, M. E.; Ladenson, R. C.; Schimmel, S. D. and Silbert, D. F. (1974): Mutants of *Escherichia coli* with temperature-sensitive malonyl coenzyme A-acyl carrier protein transacylase, J Biol Chem (vol. 249), No. 23, pp. 7468-75.
[108] D'Agnolo, G.; Rosenfeld, I. S.; Awaya, J.; Omura, S. and Vagelos, P. R. (1973): Inhibition of fatty acid synthesis by the antibiotic cerulenin. Specific inactivation of beta-ketoacyl-acyl carrier protein synthetase, Biochim Biophys Acta (vol. 326), No. 2, pp. 155-6.
[109] Katz, L. (1997): Manipulation of Modular Polyketide Synthases, Chem Rev (vol. 97), No. 7, pp. 2557-2576.
[110] Cane, D. E. (1997): Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From Collie to *Coli*, Chem Rev (vol. 97), No. 7, pp. 2463-2464.
[111] Hopwood, D. A. (1997): Genetic Contributions to Understanding Polyketide Synthases, Chem Rev (vol. 97), No. 7, pp. 2465-2498.
[112] Peypoux, F.; Bonmatin, J. M. and Wallach, J. (1999): Recent trends in the biochemistry of surfactin, Appl Microbiol Biotechnol (vol. 51), No. 5, pp. 553-63.
[113] Lin, T. P.; Chen, C. L.; Chang, L. K.; Tschen, J. S. and Liu, S. T. (1999): Functional and transcriptional analyses of a fengycin synthetase gene, *fenC*, from *Bacillus subtilis*, J Bacteriol (vol. 181), No. 16, pp. 5060-7.
[114] Umezawa, H.; Aoyagi, T.; Nishikiori, T.; Okuyama, A.; Yamagishi, Y.; Hamada, M. and Takeuchi, T. (1986): Plipastatins: new inhibitors of phospholipase A2, produced by *Bacillus cereus* BMG302-fF67. I. Taxonomy, production, isolation and preliminary characterization, J Antibiot (Tokyo) (vol. 39), No. 6, pp. 737-44.
[115] Maget-Dana, R. and Peypoux, F. (1994): Iturins, a special class of pore-forming lipopeptides: biological and physicochemical properties, Toxicology (vol. 87), No. 1-3, pp. 151-74.
[116] Hosono, K. and Suzuki, H. (1983): Acylpeptides, the inhibitors of cyclic adenosine 3',5'-monophosphate phosphodiesterase. II. Amino acid sequence and location of lactone linkage, J Antibiot (Tokyo) (vol. 36), No. 6, pp. 674-8.
[117] Hosono, K. and Suzuki, H. (1983): Acylpeptides, the inhibitors of cyclic adenosine 3',5'-monophosphate phosphodiesterase. I. Purification, physicochemical properties and structures of fatty acid residues, J Antibiot (Tokyo) (vol. 36), No. 6, pp. 667-73.
[118] Menkhaus, M.; Ullrich, C.; Kluge, B.; Vater, J.; Vollenbroich, D. and Kamp, R. M. (1993): Structural and functional organization of the surfactin synthetase multienzyme system, J Biol Chem (vol. 268), No. 11, pp. 7678-84.
[119] Fuma, S.; Fujishima, Y.; Corbell, N.; D'Souza, C.; Nakano, M. M.; Zuber, P. and Yamane, K. (1993): Nucleotide sequence of 5' portion of srfA that contains the region required for competence establishment in Bacillus subtilus, Nucleic Acids Res (vol. 21), No. 1, pp. 93-7.
[120] Galli, G.; Rodriguez, F.; Cosmina, P.; Pratesi, C.; Nogarotto, R.; de Ferra, F. and Grandi, G. (1994): Characterization of the surfactin synthetase multi-enzyme complex, Biochim Biophys Acta (vol. 1205), No. 1, pp. 19-28.
[121] Nakano, M. M.; Marahiel, M. A. and Zuber, P. (1988): Identification of a genetic locus required for biosynthesis of the lipopeptide antibiotic surfactin in Bacillus subtilis, J Bacteriol (vol. 170), No. 12, pp. 5662-8.

[122] Steller, S.; Sokoll, A.; Wilde, C.; Bernhard, F.; Franke, P. and Vater, J. (2004): Initiation of surfactin biosynthesis and the role of the SrfD-thioesterase protein, Biochemistry (vol. 43), No. 35, pp. 11331-43.
[123] Desai, J. D. and Banat, I. M. (1997): Microbial production of surfactants and their commercial potential, Microbiol Mol Biol Rev (vol. 61), No. 1, pp. 47-64.
[124] Arima, K.; Kakinuma, A. and Tamura, G. (1968): Surfactin, a crystalline peptidelipid surfactant produced by *Bacillus subtilis*: isolation, characterization and its inhibition of fibrin clot formation, Biochem Biophys Res Commun (vol. 31), No. 3, pp. 488-94.
[125] Bernheimer, A. W. and Avigad, L. S. (1970): Nature and properties of a cytolytic agent produced by Bacillus subtilis, J Gen Microbiol (vol. 61), No. 3, pp. 361-9.
[126] Kameda, Y.; Oira, S.; Matsui, K.; Kanatomo, S. and Hase, T. (1974): Antitumor activity of *bacillus natto*. V. Isolation and characterization of surfactin in the culture medium of *Bacillus natto* KMD 2311, Chem Pharm Bull (Tokyo) (vol. 22), No. 4, pp. 938-44.
[127] Vollenbroich, D.; Ozel, M.; Vater, J.; Kamp, R. M. and Pauli, G. (1997): Mechanism of inactivation of enveloped viruses by the biosurfactant surfactin from *Bacillus subtilis*, Biologicals (vol. 25), No. 3, pp. 289-97.
[128] Tsukagoshi, N.; Tamura, G. and Arima, K. (1970): A novel protoplast-bursting factor (surfactin) obtained from *Bacillus subtilis* IAM 1213. I. The effects of surfactin on *bacillus megaterium* KM, Biochim Biophys Acta (vol. 196), No. 2, pp. 204-10.
[129] Vollenbroich, D.; Pauli, G.; Ozel, M. and Vater, J. (1997): Antimycoplasma properties and application in cell culture of surfactin, a lipopeptide antibiotic from *Bacillus subtilis*, Appl Environ Microbiol (vol. 63), No. 1, pp. 44-9.
[130] Beven, L. and Wroblewski, H. (1997): Effect of natural amphipathic peptides on viability, membrane potential, cell shape and motility of mollicutes, Res Microbiol (vol. 148), No. 2, pp. 163-75.
[131] Hamoen, L. W.; Eshuis, H.; Jongbloed, J.; Venema, G. and van Sinderen, D. (1995): A small gene, designated *comS*, located within the coding region of the fourth amino acid-activation domain of *srfA*, is required for competence development in *Bacillus subtilis*, Mol Microbiol (vol. 15), No. 1, pp. 55-63.
[132] Kearns, D. B. and Losick, R. (2003): Swarming motility in undomesticated Bacillus subtilis, Mol Microbiol (vol. 49), No. 3, pp. 581-90.
[133] Julkowska, D.; Obuchowski, M.; Holland, I. B. and Seror, S. J. (2005): Comparative analysis of the development of swarming communities of Bacillus subtilis 168 and a natural wild type: critical effects of surfactin and the composition of the medium, J Bacteriol (vol. 187), No. 1, pp. 65-76.
[134] Connelly, M. B.; Young, G. M. and Sloma, A. (2004): Extracellular proteolytic activity plays a central role in swarming motility in *Bacillus subtilis*, J Bacteriol (vol. 186), No. 13, pp. 4159-67.
[135] Kinsinger, R. F.; Shirk, M. C. and Fall, R. (2003): Rapid surface motility in *Bacillus subtilis* is dependent on extracellular surfactin and potassium ion, J Bacteriol (vol. 185), No. 18, pp. 5627-31.
[136] Branda, S. S.; Gonzalez-Pastor, J. E.; Ben-Yehuda, S.; Losick, R. and Kolter, R. (2001): Fruiting body formation by *Bacillus subtilis*, Proc Natl Acad Sci U S A (vol. 98), No. 20, pp. 11621-6.
[137] Hofemeister, J.; Conrad, B.; Adler, B.; Hofemeister, B.; Feesche, J.; Kucheryava, N.; Steinborn, G.; Franke, P.; Grammel, N.; Zwintscher, A.; Leenders, F.; Hitzeroth, G. and 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 (vol. 272), No. 4, pp. 363-78.

[138] Nishikiori, T.; Naganawa, H.; Muraoka, Y.; Aoyagi, T. and Umezawa, H. (1986): Plipastatins: new inhibitors of phospholipase A2, produced by *Bacillus cereus* BMG302-fF67. II. Structure of fatty acid residue and amino acid sequence, J Antibiot (Tokyo) (vol. 39), No. 6, pp. 745-54.
[139] Ongena, M.; Jacques, P.; Toure, Y.; Destain, J.; Jabrane, A. and Thonart, P. (2005): Involvement of fengycin-type lipopeptides in the multifaceted biocontrol potential of *Bacillus subtilis*, Appl Microbiol Biotechnol (vol. 69), No. 1, pp. 29-38.
[140] Steller, S.; Vollenbroich, D.; Leenders, F.; Stein, T.; Conrad, B.; Hofemeister, J.; Jacques, P.; Thonart, P. and Vater, J. (1999): Structural and functional organization of the fengycin synthetase multienzyme system from *Bacillus subtilis* b213 and A1/3, Chem Biol (vol. 6), No. 1, pp. 31-41.
[141] Deleu, M.; Paquot, M. and Nylander, T. (2005): Fengycin interaction with lipid monolayers at the air-aqueous interface-implications for the effect of fengycin on biological membranes, J Colloid Interface Sci (vol. 283), No. 2, pp. 358-65.
[142] Nishikiori, T.; Naganawa, H.; Muraoka, Y.; Aoyagi, T. and Umezawa, H. (1986): Plipastatins: new inhibitors of phospholipase A2, produced by *Bacillus cereus* BMG302-fF67. III. Structural elucidation of plipastatins, J Antibiot (Tokyo) (vol. 39), No. 6, pp. 755-61.
[143] Kimura, K.; Nakayama, S.; Nakamura, J.; Takada, T.; Yoshihama, M.; Esumi, Y.; Itoh, Y. and Uramoto, M. (1997): SNA-60-367, new peptide enzyme inhibitors against aromatase, J Antibiot (Tokyo) (vol. 50), No. 6, pp. 529-31.
[144] Jenny, K.; Kappeli, O. and Fiechter, A. (1991): Biosurfactants from *Bacillus licheniformis*: structural analysis and characterization, Appl Microbiol Biotechnol (vol. 36), No. 1, pp. 5-13.
[145] Yakimov, M. M.; Timmis, K. N.; Wray, V. and Fredrickson, H. L. (1995): Characterization of a new lipopeptide surfactant produced by thermotolerant and halotolerant subsurface *Bacillus licheniformis* BAS50, Appl Environ Microbiol (vol. 61), No. 5, pp. 1706-13.
[146] Konz, D.; Doekel, S. and Marahiel, M. A. (1999): Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin, J Bacteriol (vol. 181), No. 1, pp. 133-40.
[147] Azevedo, E. C.; Rios, E. M.; Fukushima, K. and Campos-Takaki, G. M. (1993): Bacitracin production by a new strain of *Bacillus subtilis*. Extraction, purification, and characterization, Appl Biochem Biotechnol (vol. 42), No. 1, pp. 1-7.
[148] Anker, H. S.; Johnson, B. A.; Goldberg, J. and Meleney, F. L. (1948): Bacitracin: Methods of Production, Concentration, and Partial Purification, with a Summary of the Chemical Properties of Crude Bacitracin, J Bacteriol (vol. 55), No. 2, pp. 249-55.
[149] Konz, D.; Klens, A.; Schorgendorfer, K. and Marahiel, M. A. (1997): The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases, Chem Biol (vol. 4), No. 12, pp. 927-37.
[150] Neumuller, A. M.; Konz, D. and Marahiel, M. A. (2001): The two-component regulatory system BacRS is associated with bacitracin 'self-resistance' of *Bacillus licheniformis* ATCC 10716, Eur J Biochem (vol. 268), No. 11, pp. 3180-9.
[151] Storm, D. R. and Strominger, J. L. (1973): Complex formation between bacitracin peptides and isoprenyl pyrophosphates. The specificity of lipid-peptide interactions, J Biol Chem (vol. 248), No. 11, pp. 3940-5.
[152] Kratzschmar, J.; Krause, M. and Marahiel, M. A. (1989): Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases, J Bacteriol (vol. 171), No. 10, pp. 5422-9.

[153] Kondejewski, L. H.; Farmer, S. W.; Wishart, D. S.; Kay, C. M.; Hancock, R. E. and Hodges, R. S. (1996): Modulation of structure and antibacterial and hemolytic activity by ring size in cyclic gramicidin S analogs, J Biol Chem (vol. 271), No. 41, pp. 25261-8.
[154] Kondejewski, L. H.; Farmer, S. W.; Wishart, D. S.; Hancock, R. E. and Hodges, R. S. (1996): Gramicidin S is active against both gram-positive and gram-negative bacteria, Int J Pept Protein Res (vol. 47), No. 6, pp. 460-6.
[155] Kaprel'yants, A.S., V.V. Nikiforou, A.I. Miroshnikov, L.G. Snezhkova, V.A. Eremin, D.N. Ostrovskij (1977): Bacterial membranes and action of gramicidin S, Biokhimiya (vol. 42), pp. 329-337.
[156] Izumiya, N., T. Kato, H. Aoyaki, M. Waki, M.Kondo (1979), Synthetic aspects of biologically active cyclic peptides; gramicidin S and tyrocidines pp. 49-108, Kondansha International Ltd, Tokyo.
[157] Duitman, E. (2003): Nonribosomal peptide synthesis in *Bacillus subtilis*, Groningen University, Groningen.
[158] Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. and Stubbs, M. T. (2002): Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE, Structure (vol. 10), No. 3, pp. 301-10.
[159] Conti, E.; Stachelhaus, T.; Marahiel, M. A. and Brick, P. (1997): Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S, Embo J (vol. 16), No. 14, pp. 4174-83.
[160] Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M. and Escalante-Semerena, J. C. (2003): The 1.75 A crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A, Biochemistry (vol. 42), No. 10, pp. 2866-73.
[161] Nakano, M. M.; Xia, L. A. and Zuber, P. (1991): Transcription initiation region of the *srfA* operon, which is controlled by the *comP-comA* signal transduction system in *Bacillus subtilis*, J Bacteriol (vol. 173), No. 17, pp. 5487-93.
[162] Roggiani, M. and Dubnau, D. (1993): ComA, a phosphorylated response regulator protein of *Bacillus subtilis*, binds to the promoter region of *srfA*, J Bacteriol (vol. 175), No. 10, pp. 3182-7.
[163] Magnuson, R.; Solomon, J. and Grossman, A. D. (1994): Biochemical and genetic characterization of a competence pheromone from *B. subtilis*, Cell (vol. 77), No. 2, pp. 207-16.
[164] Solomon, J. M.; Lazazzera, B. A. and Grossman, A. D. (1996): Purification and characterization of an extracellular peptide factor that affects two different developmental pathways in *Bacillus subtilis*, Genes Dev (vol. 10), No. 16, pp. 2014-24.
[165] Grossman, A. D. (1995): Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*, Annu Rev Genet (vol. 29), pp. 477-508.
[166] Perego, M. (1997): A peptide export-import control circuit modulating bacterial development regulates protein phosphatases of the phosphorelay, Proc Natl Acad Sci U S A (vol. 94), No. 16, pp. 8612-7.
[167] Bongiorni, C.; Ishikawa, S.; Stephenson, S.; Ogasawara, N. and Perego, M. (2005): Synergistic regulation of competence development in *Bacillus subtilis* by two Rap-Phr systems, J Bacteriol (vol. 187), No. 13, pp. 4353-61.
[168] Hayashi, K.; Ohsawa, T.; Kobayashi, K.; Ogasawara, N. and Ogura, M. (2005): The H_2O_2 stress-responsive regulator PerR positively regulates *srfA* expression in Bacillus subtilis, J Bacteriol (vol. 187), No. 19, pp. 6659-67.

[169] Serror, P. and Sonenshein, A. L. (1996): CodY is required for nutritional repression of *Bacillus subtilis* genetic competence, J Bacteriol (vol. 178), No. 20, pp. 5910-5.
[170] Tsuge, K.; Ohata, Y. and Shoda, M. (2001): Gene *yerP,* involved in surfactin self-resistance in *Bacillus subtilis*, Antimicrob Agents Chemother (vol. 45), No. 12, pp. 3566-73.
[171] Msadek, T.; Kunst, F.; Klier, A. and Rapoport, G. (1991): DegS-DegU and ComP-ComA modulator-effector pairs control expression of the *Bacillus subtilis* pleiotropic regulatory gene *degQ*, J Bacteriol (vol. 173), No. 7, pp. 2366-77.
[172] Tsuge, K.; Ano, T.; Hirai, M.; Nakamura, Y. and 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 (vol. 43), No. 9, pp. 2183-92.
[173] Tsuge, K.; Inoue, S.; Ano, T.; Itaya, M. and Shoda, M. (2005): Horizontal transfer of iturin A operon, itu, to *Bacillus subtilis* 168 and conversion into an iturin A producer, Antimicrob Agents Chemother (vol. 49), No. 11, pp. 4641-8.
[174] Mootz, H. D.; Finking, R. and Marahiel, M. A. (2001): 4'-phosphopantetheine transfer in primary and secondary metabolism of *Bacillus subtilis*, J Biol Chem (vol. 276), No. 40, pp. 37289-98.
[175] Nakano, M. M.; Corbell, N.; Besson, J. and Zuber, P. (1992): Isolation and characterization of *sfp*: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*, Mol Gen Genet (vol. 232), No. 2, pp. 313-21.
[176] Yakimov, M. M. and Golyshin, P. N. (1997): ComA-dependent transcriptional activation of lichenysin A synthetase promoter in *Bacillus subtilis* cells, Biotechnol Prog (vol. 13), No. 6, pp. 757-61.
[177] Podlesek, Z.; Comino, A.; Herzog-Velikonja, B.; Zgur-Bertok, D.; Komel, R. and Grabnar, M. (1995): Bacillus licheniformis bacitracin-resistance ABC transporter: relationship to mammalian multidrug resistance, Mol Microbiol (vol. 16), No. 5, pp. 969-76.
[178] Podlesek, Z.; Comino, A.; Herzog-Velikonja, B. and Grabnar, M. (2000): The role of the bacitracin ABC transporter in bacitracin resistance and collateral detergent sensitivity, FEMS Microbiol Lett (vol. 188), No. 1, pp. 103-6.
[179] Bernard, R.; Joseph, P.; Guiseppi, A.; Chippaux, M. and Denizot, F. (2003): YtsCD and YwoA, two independent systems that confer bacitracin resistance to *Bacillus subtilis*, FEMS Microbiol Lett (vol. 228), No. 1, pp. 93-7.
[180] Marahiel, M. A.; Zuber, P.; Czekay, G. and Losick, R. (1987): Identification of the promoter for a peptide antibiotic biosynthesis gene from *Bacillus brevis* and its regulation in *Bacillus subtilis*, J Bacteriol (vol. 169), No. 5, pp. 2215-22.
[181] Robertson, J. B.; Gocht, M.; Marahiel, M. A. and Zuber, P. (1989): AbrB, a regulator of gene expression in *Bacillus*, interacts with the transcription initiation regions of a sporulation gene and an antibiotic biosynthesis gene, Proc Natl Acad Sci U S A (vol. 86), No. 21, pp. 8457-61.
[182] Stachelhaus, T.; Schneider, A. and Marahiel, M. A. (1995): Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains, Science (vol. 269), No. 5220, pp. 69-72.
[183] Stachelhaus, T.; Schneider, A. and Marahiel, M. A. (1996): Engineered biosynthesis of peptide antibiotics, Biochem Pharmacol (vol. 52), No. 2, pp. 177-86.
[184] Mootz, H. D.; Kessler, N.; Linne, U.; Eppelmann, K.; Schwarzer, D. and Marahiel, M. A. (2002): Decreasing the ring size of a cyclic nonribosomal peptide antibiotic by in-frame module deletion in the biosynthetic genes, J Am Chem Soc (vol. 124), No. 37, pp. 10980-1.

[185] Eppelmann, K.; Stachelhaus, T. and Marahiel, M. A. (2002): Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics, Biochemistry (vol. 41), No. 30, pp. 9718-26.
[186] de Ferra, F.; Rodriguez, F.; Tortora, O.; Tosi, C. and Grandi, G. (1997): Engineering of peptide synthetases. Key role of the thioesterase-like domain for efficient production of recombinant peptides, J Biol Chem (vol. 272), No. 40, pp. 25304-9.
[187] Duerfahrt, T.; Eppelmann, K.; Muller, R. and Marahiel, M. A. (2004): Rational design of a bimodular model system for the investigation of heterocyclization in nonribosomal peptide biosynthesis, Chem Biol (vol. 11), No. 2, pp. 261-71.
[188] Tamehiro, N.; Okamoto-Hosoya, Y.; Okamoto, S.; Ubukata, M.; Hamada, M.; Naganawa, H. and Ochi, K. (2002): Bacilysocin, a novel phospholipid antibiotic produced by *Bacillus subtilis* 168, Antimicrob Agents Chemother (vol. 46), No. 2, pp. 315-20.
[189] Pinchuk, I. V.; Bressollier, P.; Verneuil, B.; Fenet, B.; Sorokulova, I. B.; Megraud, F. and Urdaci, M. C. (2001): In vitro anti-*Helicobacter pylori* activity of the probiotic strain *Bacillus subtilis* 3 is due to secretion of antibiotics, Antimicrob Agents Chemother (vol. 45), No. 11, pp. 3156-61.
[190] Inaoka, T.; Takahashi, K.; Yada, H.; Yoshida, M. and Ochi, K. (2004): RNA polymerase mutation activates the production of a dormant antibiotic 3,3'-neotrehalosadiamine via an autoinduction mechanism in *Bacillus subtilis*, J Biol Chem (vol. 279), No. 5, pp. 3885-92.
[191] Stragier, P.; Bonamy, C. and Karmazyn-Campelli, C. (1988): Processing of a sporulation sigma factor in *Bacillus subtilis*: how morphological structure could control gene expression, Cell (vol. 52), No. 5, pp. 697-704.
[192] Antoniewski, C.; Savelli, B. and Stragier, P. (1990): The *spoIIJ* gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes, J Bacteriol (vol. 172), No. 1, pp. 86-93.
[193] Steinmetz, M. and Richter, R. (1994): Plasmids designed to alter the antibiotic resistance expressed by insertion mutations in *Bacillus subtilis*, through in vivo recombination, Gene (vol. 142), No. 1, pp. 79-83.
[194] Ceglowski, P. and Alonso, J. C. (1994): Gene organization of the *Streptococcus pyogenes* plasmid pDB101: sequence analysis of the orf eta-copS region, Gene (vol. 145), No. 1, pp. 33-9.
[195] Gonzy-Treboul, G.; Karmazyn-Campelli, C. and Stragier, P. (1992): Developmental regulation of transcription of the *Bacillus subtilis ftsAZ* operon, J Mol Biol (vol. 224), No. 4, pp. 967-79.
[196] Koumoutsi, A.; Chen, X. H.; Henne, A.; Liesegang, H.; Hitzeroth, G.; Franke, P.; Vater, J. and Borriss, R. (2004): Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42, J Bacteriol (vol. 186), No. 4, pp. 1084-96.
[197] Chen, X. H.; Vater, J.; Piel, J.; Franke, P.; Scholz, R.; Schneider, K.; Koumoutsi, A.; Hitzeroth, G.; Grammel, N.; Strittmatter, A. W.; Gottschalk, G.; Sussmuth, R. D. and Borriss, R. (2006): Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42, J Bacteriol (vol. 188), No. 11, pp. 4024-36.
[198] Sambrook J.; Fritsch, E.F. and Maniatis, T. (1989): Molecular Cloning: a laboratory manual, Cold Spring Harbor Laboratory, NY.
[199] Cutting, S.M. and Vander Horn, P.B. (1990): Molecular Biological methods for *Bacillus* Harwood C., Cutting S.M., Ed, Wiley, Chichester.

[200] Landy M., Warren, G.H., Roseman S.B., Colio L.G. (1948): Bacillomycin, an antibiotic from *Bacillus subtilis* active against pathogenic fungi, Proc. Soc. Exp. Biol. Med (vol. 67), pp. 539-541.
[201] Ebata, M.; Miyazaki, K. and Takahashi, Y. (1969): Studies on subsporin. I. Isolation and characterization of subsporins A, B and C, J Antibiot (Tokyo) (vol. 22), No. 10, pp. 467-72.
[202] Cutting, S.M. and Van der Horn, P.B. (1990): Genetic Analysis, Harwood C.R., Cutting S.M, Molecular biological methods for *Bacillus* pp. 27-74, Wiley Interscience, Chichister, United Kingdom.
[203] Dieffenbach, C.W. and Dveksler, G.S. (1995): PCR primer, a laboratory manual Cold Spring Harbor Laboratory Press.
[204] Spizizen, J. (1958): Transformation of Biochemically Deficient Strains of *Bacillus Subtilis* by Deoxyribonucleate, Proc Natl Acad Sci U S A (vol. 44), No. 10, pp. 1072-8.
[205] Kunst, F. and Rapoport, G. (1995): Salt stress is an environmental signal affecting degradative enzyme synthesis in *Bacillus subtilis*, J Bacteriol (vol. 177), No. 9, pp. 2403-7.
[206] Diatchenko, L.; Lau, Y. F.; Campbell, A. P.; Chenchik, A.; Moqadam, F.; Huang, B.; Lukyanov, S.; Lukyanov, K.; Gurskaya, N.; Sverdlov, E. D. and Siebert, P. D. (1996): Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries, Proc Natl Acad Sci U S A (vol. 93), No. 12, pp. 6025-30.
[207] Clontech Laboratories, Inc. (2006): PCR-Select^{TM} Bacterial Genome Subtraction Kit User Manual
[208] Birren, B. and Lai, E. (1993): Pulsed Field Gel Electrophoresis. a practical Guide, Academic Press, San Diego.
[209] Typas, A. and Hengge, R. (2006): Role of the spacer between the -35 and -10 regions in σ^S promoter selectivity in *Escherichia coli*, Mol Microbiol (vol. 59), No. 3, pp. 1037-51.
[210] Leenders F., Stein T.H., Kablitz B., Franke P., Vater J. (1999): Rapid typing of *Bacillus subtilis* strains by their secondary metabolites using matrix-assisted laser desorption/ionisation mass spectrometry of intact cells, Rapid Commun. Mass spectrom. (vol. 13), pp. 943-949.
[211] Vater, J.; Kablitz, B.; Wilde, C.; Franke, P.; Mehta, N. and Cameotra, S. S. (2002): Matrix-assisted laser desorption ionization--time of flight mass spectrometry of lipopeptide biosurfactants in whole cells and culture filtrates of Bacillus subtilis C-1 isolated from petroleum sludge, Appl Environ Microbiol (vol. 68), No. 12, pp. 6210-9.
[212] Miller, J.H. (1972): Experiments in molecular genetics, Cold Spring Harbor Laboratory Press, N.Y.
[213] Ewing, B.; Hillier, L.; Wendl, M. C. and Green, P. (1998): Base-calling of automated sequencer traces using phred. I. Accuracy assessment, Genome Res (vol. 8), No. 3, pp. 175-85.
[214] Staden, R.; Beal, K. F. and Bonfield, J. K. (2000): The Staden package, 1998, Methods Mol Biol (vol. 132), pp. 115-30.
[215] Finn, R. D.; Mistry, J.; Schuster-Bockler, B.; Griffiths-Jones, S.; Hollich, V.; Lassmann, T.; Moxon, S.; Marshall, M.; Khanna, A.; Durbin, R.; Eddy, S. R.; Sonnhammer, E. L. and Bateman, A. (2006): Pfam: clans, web tools and services, Nucleic Acids Res (vol. 34), No. Database issue, pp. D247-51.
[216] Priest, F.G., Goodfellow, M., Shute L.A., Berkeley R.C.W. (1987): *Bacillus amyloliquefaciens* sp. nov., nom. rev., Int J Syst Bacteriol (vol. 37), pp. 69-71.

[217] Inaoka, T.; Takahashi, K.; Ohnishi-Kameyama, M.; Yoshida, M. and 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 (vol. 278), No. 4, pp. 2169-76.
[218] Kenig, M. and Abraham, E. P. (1976): Antimicrobial activities and antagonists of bacilysin and anticapsin, J Gen Microbiol (vol. 94), No. 1, pp. 37-45.
[219] Hilton, M. D.; Alaeddinoglu, N. G. and Demain, A. L. (1988): Synthesis of bacilysin by *Bacillus subtilis* branches from prephenate of the aromatic amino acid pathway, J Bacteriol (vol. 170), No. 1, pp. 482-4.
[220] Ollinger, J.; Song, K. B.; Antelmann, H.; Hecker, M. and Helmann, J. D. (2006): Role of the Fur regulon in iron transport in *Bacillus subtilis*, J Bacteriol (vol. 188), No. 10, pp. 3664-73.
[221] Hoffmann, T.; Schutz, A.; Brosius, M.; Volker, A.; Volker, U. and Bremer, E. (2002): High-salinity-induced iron limitation in *Bacillus subtilis*, J Bacteriol (vol. 184), No. 3, pp. 718-27.
[222] Chen, C. L.; Chang, L. K.; Chang, Y. S.; Liu, S. T. and Tschen, J. S. (1995): Transposon mutagenesis and cloning of the genes encoding the enzymes of fengycin biosynthesis in *Bacillus subtilis*, Mol Gen Genet (vol. 248), No. 2, pp. 121-5.
[223] Butcher, B. G. and Helmann, J. D. (2006): Identification of *Bacillus subtilis* sigma-dependent genes that provide intrinsic resistance to antimicrobial compounds produced by *Bacilli*, Mol Microbiol (vol. 60), No. 3, pp. 765-82.
[224] Amory, A.; Kunst, F.; Aubert, E.; Klier, A. and Rapoport, G. (1987): Characterization of the sacQ genes from *Bacillus licheniformis* and *Bacillus subtilis*, J Bacteriol (vol. 169), No. 1, pp. 324-33.
[225] Yang, M.; Ferrari, E.; Chen, E. and Henner, D. J. (1986): Identification of the pleiotropic *sacQ* gene of *Bacillus subtilis*, J Bacteriol (vol. 166), No. 1, pp. 113-9.
[226] deHaseth, P. L.; Zupancic, M. L. and Record, M. T., Jr. (1998): RNA polymerase-promoter interactions: the comings and goings of RNA polymerase, J Bacteriol (vol. 180), No. 12, pp. 3019-25.
[227] Voskuil, M. I. and Chambliss, G. H. (1998): The -16 region of *Bacillus subtilis* and other gram-positive bacterial promoters, Nucleic Acids Res (vol. 26), No. 15, pp. 3584-90.
[228] Barne, K. A.; Bown, J. A.; Busby, S. J. and Minchin, S. D. (1997): Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters, Embo J (vol. 16), No. 13, pp. 4034-40.
[229] Kunst, F.; Pascal, M.; Lepesant-Kejzlarova, J.; Lepesant, J. A.; Billault, A. and Dedonder, R. (1974): Pleiotropic mutations affecting sporulation conditions and the syntheses of extracellular enzymes in *Bacillus subtilis* 168, Biochimie (vol. 56), No. 11-12, pp. 1481-9.
[230] Msadek, T.; Kunst, F.; Henner, D.; Klier, A.; Rapoport, G. and Dedonder, R. (1990): Signal transduction pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: expression of the regulatory genes and analysis of mutations in *degS* and *degU*, J Bacteriol (vol. 172), No. 2, pp. 824-34.
[231] Hamoen, L. W.; Van Werkhoven, A. F.; Venema, G. and Dubnau, D. (2000): The pleiotropic response regulator DegU functions as a priming protein in competence development in *Bacillus subtilis*, Proc Natl Acad Sci U S A (vol. 97), No. 16, pp. 9246-51.
[232] Guillen, N.; Weinrauch, Y. and Dubnau, D. A. (1989): Cloning and characterization of the regulatory *Bacillus subtilis* competence genes *comA* and *comB*, J Bacteriol (vol. 171), No. 10, pp. 5354-61.

[233] Gaidenko, T. A. and Price, C. W. (1998): General stress transcription factor sigma(B) and sporulation transcription factor sigma(H) each contribute to survival of *Bacillus subtilis* under extreme growth conditions, J Bacteriol (vol. 180), No. 14, pp. 3730-3.
[234] McQuade, R. S.; Comella, N. and Grossman, A. D. (2001): Control of a family of phosphatase regulatory genes (*phr*) by the alternate sigma factor σ^H of *Bacillus subtilis*, J Bacteriol (vol. 183), No. 16, pp. 4905-9.
[235] Serizawa, M.; Kodama, K.; Yamamoto, H.; Kobayashi, K.; Ogasawara, N. and Sekiguchi, J. (2005): Functional analysis of the YvrGHb two-component system of *Bacillus subtilis*: identification of the regulated genes by DNA microarray and northern blot analyses, Biosci Biotechnol Biochem (vol. 69), No. 11, pp. 2155-69.
[236] Huang, X.; Fredrick, K. L. and Helmann, J. D. (1998): Promoter recognition by *Bacillus subtilis* sigmaW: autoregulation and partial overlap with the sigmaX regulon, J Bacteriol (vol. 180), No. 15, pp. 3765-70.
[237] Hecker, M. and Volker, U. (1998): Non-specific, general and multiple stress resistance of growth-restricted *Bacillus subtilis* cells by the expression of the sigmaB regulon, Mol Microbiol (vol. 29), No. 5, pp. 1129-36.
[238] Hecker, M. and Volker, U. (2001): General stress response of *Bacillus subtilis* and other bacteria, Adv Microb Physiol (vol. 44), pp. 35-91.
[239] Petersohn, A.; Bernhardt, J.; Gerth, U.; Hoper, D.; Koburger, T.; Volker, U. and Hecker, M. (1999): Identification of sigma(B)-dependent genes in *Bacillus subtilis* using a promoter consensus-directed search and oligonucleotide hybridization, J Bacteriol (vol. 181), No. 18, pp. 5718-24.
[240] Marquez, L. M.; Helmann, J. D.; Ferrari, E.; Parker, H. M.; Ordal, G. W. and Chamberlin, M. J. (1990): Studies of sigma D-dependent functions in *Bacillus subtilis*, J Bacteriol (vol. 172), No. 6, pp. 3435-43.
[241] Huang, X.; Decatur, A.; Sorokin, A. and Helmann, J. D. (1997): The *Bacillus subtilis* sigma(X) protein is an extracytoplasmic function sigma factor contributing to survival at high temperature, J Bacteriol (vol. 179), No. 9, pp. 2915-21.
[242] Sorokin, A.; Bolotin, A.; Purnelle, B.; Hilbert, H.; Lauber, J.; Dusterhoft, A. and Ehrlich, S. D. (1997): Sequence of the *Bacillus subtilis* genome region in the vicinity of the lev operon reveals two new extracytoplasmic function RNA polymerase sigma factors SigV and SigZ, Microbiology (vol. 143 (Pt 9)), No. Pt 9, pp. 2939-43.
[243] Sonenshein, A. L. (2005): CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria, Curr Opin Microbiol (vol. 8), No. 2, pp. 203-7.
[244] Shimane, K. and Ogura, M. (2004): Mutational analysis of the helix-turn-helix region of *Bacillus subtilis* response regulator DegU, and identification of cis-acting sequences for DegU in the aprE and comK promoters, J Biochem (Tokyo) (vol. 136), No. 3, pp. 387-97.
[245] Dartois, V.; Debarbouille, M.; Kunst, F. and Rapoport, G. (1998): Characterization of a novel member of the DegS-DegU regulon affected by salt stress in *Bacillus subtilis*, J Bacteriol (vol. 180), No. 7, pp. 1855-61.
[246] Core, L. and Perego, M. (2003): TPR-mediated interaction of RapC with ComA inhibits response regulator-DNA binding for competence development in *Bacillus subtilis*, Mol Microbiol (vol. 49), No. 6, pp. 1509-22.
[247] Auchtung, J. M.; Lee, C. A. and Grossman, A. D. (2006): Modulation of the ComA-dependent quorum response in *Bacillus subtilis* by multiple Rap proteins and Phr peptides, J Bacteriol (vol. 188), No. 14, pp. 5273-85.
[248] Ogura, M.; Shimane, K.; Asai, K.; Ogasawara, N. and Tanaka, T. (2003): Binding of response regulator DegU to the *aprE* promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*, Mol Microbiol (vol. 49), No. 6, pp. 1685-97.

[249] Hayashi, K.; Kensuke, T.; Kobayashi, K.; Ogasawara, N. and Ogura, M. (2006): *Bacillus subtilis* RghR (YvaN) represses *rapG* and *rapH*, which encode inhibitors of expression of the *srfA* operon, Mol Microbiol (vol. 59), No. 6, pp. 1714-29.
[250] Pottathil, M. and Lazazzera, B. A. (2003): The extracellular Phr peptide-Rap phosphatase signaling circuit of *Bacillus subtilis*, Front Biosci (vol. 8), pp. d32-45.
[251] Perego, M.; Higgins, C. F.; Pearce, S. R.; Gallagher, M. P. and Hoch, J. A. (1991): The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation, Mol Microbiol (vol. 5), No. 1, pp. 173-85.
[252] Price, C. W.; Fawcett, P.; Ceremonie, H.; Su, N.; Murphy, C. K. and Youngman, P. (2001): Genome-wide analysis of the general stress response in *Bacillus subtilis*, Mol Microbiol (vol. 41), No. 4, pp. 757-74.
[253] Ogura, M.; Yamaguchi, H.; Yoshida, Ki; Fujita, Y. and 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 (vol. 29), No. 18, pp. 3804-13.
[254] Steil, L.; Hoffmann, T.; Budde, I.; Volker, U. and Bremer, E. (2003): Genome-wide transcriptional profiling analysis of adaptation of *Bacillus subtilis* to high salinity, J Bacteriol (vol. 185), No. 21, pp. 6358-70.
[255] Bsat, N. and Helmann, J. D. (1999): Interaction of *Bacillus subtilis* Fur (ferric uptake repressor) with the dhb operator in vitro and in vivo, J Bacteriol (vol. 181), No. 14, pp. 4299-307.
[256] Miethke, M.; Klotz, O.; Linne, U.; May, J. J.; Beckering, C. L. and Marahiel, M. A. (2006): Ferri-bacillibactin uptake and hydrolysis in *Bacillus subtilis*, Mol Microbiol (vol. 3), p. 3.
[257] Mader, U.; Antelmann, H.; Buder, T.; Dahl, M. K.; Hecker, M. and Homuth, G. (2002): *Bacillus subtilis* functional genomics: genome-wide analysis of the DegS-DegU regulon by transcriptomics and proteomics, Mol Genet Genomics (vol. 268), No. 4, pp. 455-67.
[258] Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S. and Horikoshi, K. (2000): Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*, Nucleic Acids Res (vol. 28), No. 21, pp. 4317-31.
[259] Rasko, D. A.; Ravel, J.; Okstad, O. A.; Helgason, E.; Cer, R. Z.; Jiang, L.; Shores, K. A.; Fouts, D. E.; Tourasse, N. J.; Angiuoli, S. V.; Kolonay, J.; Nelson, W. C.; Kolsto, A. B.; Fraser, C. M. and Read, T. D. (2004): The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1, Nucleic Acids Res (vol. 32), No. 3, pp. 977-88.
[260] Rey, M. W.; Ramaiya, P.; Nelson, B. A.; Brody-Karpin, S. D.; Zaretsky, E. J.; Tang, M.; Lopez de Leon, A.; Xiang, H.; Gusti, V.; Clausen, I. G.; Olsen, P. B.; Rasmussen, M. D.; Andersen, J. T.; Jorgensen, P. L.; Larsen, T. S.; Sorokin, A.; Bolotin, A.; Lapidus, A.; Galleron, N.; Ehrlich, S. D. and Berka, R. M. (2004): Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species, Genome Biol (vol. 5), No. 10, p. R77.
[261] Xu, D. and Cote, J. C. (2003): Phylogenetic relationships between *Bacillus* species and related genera inferred from comparison of 3' end 16S rDNA and 5' end 16S-23S ITS nucleotide sequences, Int J Syst Evol Microbiol (vol. 53), No. Pt 3, pp. 695-704.
[262] Hoch, JA. and Silhavy, TJ. (1995): Two-Component Signal Transduction, American Society of Microbiology, Washington D.C.
[263] Parkinson, J. S. (1993): Signal transduction schemes of bacteria, Cell (vol. 73), No. 5, pp. 857-71.

[264] Stock, JB., Surette MG, Levit M., Park P. (1995): Two-component signal transduction systems: structure-function relationships and mechanisms of catalysis Hoch, JA. Silhavy TJ., Two-Component Signal Transduction pp. 25-51, American Society for Microbiology, Washington DC.
[265] Stock, A. M.; Robinson, V. L. and Goudreau, P. N. (2000): Two-component signal transduction, Annu Rev Biochem (vol. 69), pp. 183-215.
[266] West, A. H. and Stock, A. M. (2001): Histidine kinases and response regulator proteins in two-component signaling systems, Trends Biochem Sci (vol. 26), No. 6, pp. 369-76.
[267] Galperin, M. Y. (2006): Structural classification of bacterial response regulators: diversity of output domains and domain combinations, J Bacteriol (vol. 188), No. 12, pp. 4169-82.
[268] Stock, J. B.; Stock, A. M. and Mottonen, J. M. (1990): Signal transduction in bacteria, Nature (vol. 344), No. 6265, pp. 395-400.
[269] Fabret, C.; Feher, V. A. and Hoch, J. A. (1999): Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world, J Bacteriol (vol. 181), No. 7, pp. 1975-83.
[270] Mascher, T.; Margulis, N. G.; Wang, T.; Ye, R. W. and Helmann, J. D. (2003): Cell wall stress responses in *Bacillus subtilis*: the regulatory network of the bacitracin stimulon, Mol Microbiol (vol. 50), No. 5, pp. 1591-604.
[271] Pietiainen, M.; Gardemeister, M.; Mecklin, M.; Leskela, S.; Sarvas, M. and Kontinen, V. P. (2005): Cationic antimicrobial peptides elicit a complex stress response in *Bacillus subtilis* that involves ECF-type sigma factors and two-component signal transduction systems, Microbiology (vol. 151), No. Pt 5, pp. 1577-92.
[272] Ohki, R.; Giyanto; Tateno, K.; Masuyama, W.; Moriya, S.; Kobayashi, K. and Ogasawara, N. (2003): The BceRS two-component regulatory system induces expression of the bacitracin transporter, BceAB, in *Bacillus subtilis*, Mol Microbiol (vol. 49), No. 4, pp. 1135-44.
[273] Mascher, T. (2006): Intramembrane-sensing histidine kinases: a new family of cell envelope stress sensors in *Firmicutes* bacteria, FEMS Microbiol Lett.
[274] Schultz, J.; Milpetz, F.; Bork, P. and Ponting, C. P. (1998): SMART, a simple modular architecture research tool: identification of signaling domains, Proc Natl Acad Sci U S A (vol. 95), No. 11, pp. 5857-64.
[275] Wecke, T.; Veith, B.; Ehrenreich, A. and Mascher, T. (2006): Cell Envelope Stress Response in *Bacillus licheniformis*: Integrating Comparative Genomics, Transcriptional Profiling and Regulon Mining to Decipher a Complex Regulatory Network, J Bacteriol (vol. 25), p. 25.
[276] Steinfels, E.; Orelle, C.; Fantino, J. R.; Dalmas, O.; Rigaud, J. L.; Denizot, F.; Di Pietro, A. and Jault, J. M. (2004): Characterization of YvcC (BmrA), a multidrug ABC transporter constitutively expressed in *Bacillus subtilis*, Biochemistry (vol. 43), No. 23, pp. 7491-502.
[277] Joseph, P.; Fichant, G.; Quentin, Y. and Denizot, F. (2002): Regulatory relationship of two-component and ABC transport systems and clustering of their genes in the *Bacillus*/*Clostridium* group, suggest a functional link between them, J Mol Microbiol Biotechnol (vol. 4), No. 5, pp. 503-13.
[278] Stein, T.; Heinzmann, S.; Kiesau, P.; Himmel, B. and Entian, K. D. (2003): The spa-box for transcriptional activation of subtilin biosynthesis and immunity in *Bacillus subtilis*, Mol Microbiol (vol. 47), No. 6, pp. 1627-36.
[279] Perego, M.; Glaser, P. and Hoch, J. A. (1996): Aspartyl-phosphate phosphatases deactivate the response regulator components of the sporulation signal transduction system in *Bacillus subtilis*, Mol Microbiol (vol. 19), No. 6, pp. 1151-7.

[280] Perego, M. and Brannigan, J. A. (2001): Pentapeptide regulation of aspartyl-phosphate phosphatases, Peptides (vol. 22), No. 10, pp. 1541-7.
[281] Blatch, G. L. and Lassle, M. (1999): The tetratricopeptide repeat: a structural motif mediating protein-protein interactions, Bioessays (vol. 21), No. 11, pp. 932-9.
[282] Turner, M. S. and Helmann, J. D. (2000): Mutations in multidrug efflux homologs, sugar isomerases, and antimicrobial biosynthesis genes differentially elevate activity of the sigma(X) and sigma(W) factors in *Bacillus subtilis*, J Bacteriol (vol. 182), No. 18, pp. 5202-10.
[283] Helmann, J. D. (2002): The extracytoplasmic function (ECF) sigma factors, Adv Microb Physiol (vol. 46), pp. 47-110.
[284] Albano, M.; Smits, W. K.; Ho, L. T.; Kraigher, B.; Mandic-Mulec, I.; Kuipers, O. P. and Dubnau, D. (2005): The Rok protein of *Bacillus subtilis* represses genes for cell surface and extracellular functions, J Bacteriol (vol. 187), No. 6, pp. 2010-9.
[285] Tortosa, P.; Logsdon, L.; Kraigher, B.; Itoh, Y.; Mandic-Mulec, I. and Dubnau, D. (2001): Specificity and genetic polymorphism of the *Bacillus* competence quorum-sensing system, J Bacteriol (vol. 183), No. 2, pp. 451-60.
[286] Ansaldi, M.; Marolt, D.; Stebe, T.; Mandic-Mulec, I. and Dubnau, D. (2002): Specific activation of the *Bacillus* quorum-sensing systems by isoprenylated pheromone variants, Mol Microbiol (vol. 44), No. 6, pp. 1561-73.
[287] Ansaldi, M. and Dubnau, D. (2004): Diversifying selection at the *Bacillus* quorum-sensing locus and determinants of modification specificity during synthesis of the ComX pheromone, J Bacteriol (vol. 186), No. 1, pp. 15-21.
[288] Straight, P. D.; Willey, J. M. and Kolter, R. (2006): Interactions between *Streptomyces coelicolor* and *Bacillus subtilis*: Role of surfactants in raising aerial structures, J Bacteriol (vol. 188), No. 13, pp. 4918-25.
[289] Han, J. S.; Cheng, J. H.; Yoon, T. M.; Song, J.; Rajkarnikar, A.; Kim, W. G.; Yoo, I. D.; Yang, Y. Y. and Suh, J. W. (2005): Biological control agent of common scab disease by antagonistic strain *Bacillus sp. sunhua*, J Appl Microbiol (vol. 99), No. 1, pp. 213-21.
[290] Ron, E. Z. and Rosenberg, E. (2001): Natural roles of biosurfactants, Environ Microbiol (vol. 3), No. 4, pp. 229-36.
[291] Neu, T. R. (1996): Significance of bacterial surface-active compounds in interaction of bacteria with interfaces, Microbiol Rev (vol. 60), No. 1, pp. 151-66.
[292] Elliot, M. A. and Talbot, N. J. (2004): Building filaments in the air: aerial morphogenesis in bacteria and fungi, Curr Opin Microbiol (vol. 7), No. 6, pp. 594-601.
[293] Challis, G. L. and Hopwood, D. A. (2003): Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces species*, Proc Natl Acad Sci U S A (vol. 100 Suppl 2), pp. 14555-61.
[294] Patel, P. S.; Huang, S.; Fisher, S.; Pirnik, D.; Aklonis, C.; Dean, L.; Meyers, E.; Fernandes, P. and Mayerl, F. (1995): Bacillaene, a novel inhibitor of procaryotic protein synthesis produced by *Bacillus subtilis*: production, taxonomy, isolation, physico-chemical characterization and biological activity, J Antibiot (Tokyo) (vol. 48), No. 9, pp. 997-1003.
[295] Wilson, K. E.; Flor, J. E.; Schwartz, R. E.; Joshua, H.; Smith, J. L.; Pelak, B. A.; Liesch, J. M. and Hensens, O. D. (1987): Difficidin and oxydifficidin: novel broad spectrum antibacterial antibiotics produced by *Bacillus subtilis*. II. Isolation and physico-chemical characterization, J Antibiot (Tokyo) (vol. 40), No. 12, pp. 1682-91.
[296] Piel, J. (2002): A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus beetles*, Proc Natl Acad Sci U S A (vol. 99), No. 22, pp. 14002-7.

[297] Steinborn, G.; Hajirezaei, M. R. and Hofemeister, J. (2005): bac genes for recombinant bacilysin and anticapsin production in *Bacillus* host strains, Arch Microbiol (vol. 183), No. 2, pp. 71-9.
[298] Scholz, Romy (2005): Diploma Arbeit, Berlin
[299] Andrews, S. C.; Robinson, A. K. and Rodriguez-Quinones, F. (2003): Bacterial iron homeostasis, FEMS Microbiol Rev (vol. 27), No. 2-3, pp. 215-37.
[300] Escolar, L.; Perez-Martin, J. and de Lorenzo, V. (1999): Opening the iron box: transcriptional metalloregulation by the Fur protein, J Bacteriol (vol. 181), No. 20, pp. 6223-9.
[301] Hantke, K. (2001): Bacterial zinc transporters and regulators, Biometals (vol. 14), No. 3-4, pp. 239-49.
[302] Kloepper, J.W., Leong J., Teintze M., Schroth M.N. (1980): Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria., Nature (vol. 286), pp. 885-886.
[303] Gonzalez-Pastor, J. E.; Hobbs, E. C. and Losick, R. (2003): Cannibalism by sporulating bacteria, Science (vol. 301), No. 5632, pp. 510-3.
[304] Besson, F.; Peypoux, F.; Michel, G. and Delcambe, L. (1978): Mode of action of iturin A, an antibiotic isolated from *Bacillus subtilis*, on *Micrococcus luteus*, Biochem Biophys Res Commun (vol. 81), No. 2, pp. 297-304.
[305] Peypoux, F.; Besson, F.; Michel, G. and Delcambe, L. (1981): Structure of bacillomycin D, a new antibiotic of the iturin group, Eur J Biochem (vol. 118), No. 2, pp. 323-7.
[306] Pao, G. M.; Tam, R.; Lipschitz, L. S. and Saier, M. H., Jr. (1994): Response regulators: structure, function and evolution, Res Microbiol (vol. 145), No. 5-6, pp. 356-62.
[307] Dahl, M. K.; Msadek, T.; Kunst, F. and Rapoport, G. (1992): The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of genetic competence in *Bacillus subtilis*, J Biol Chem (vol. 267), No. 20, pp. 14509-14.
[308] Dahl, J. L.; Wei, B. Y. and Kadner, R. J. (1997): Protein phosphorylation affects binding of the *Escherichia coli* transcription activator UhpA to the *uhpT* promoter, J Biol Chem (vol. 272), No. 3, pp. 1910-9.
[309] Turgay, K.; Hamoen, L. W.; Venema, G. and Dubnau, D. (1997): Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of *Bacillus subtilis*, Genes Dev (vol. 11), No. 1, pp. 119-28.
[310] Dubnau, D.; Hahn, J.; Roggiani, M.; Piazza, F. and Weinrauch, Y. (1994): Two-component regulators and genetic competence in *Bacillus subtilis*, Res Microbiol (vol. 145), No. 5-6, pp. 403-11.
[311] Mueller, J. P.; Bukusoglu, G. and Sonenshein, A. L. (1992): Transcriptional regulation of *Bacillus subtilis* glucose starvation-inducible genes: control of *gsiA* by the ComP-ComA signal transduction system, J Bacteriol (vol. 174), No. 13, pp. 4361-73.
[312] Lazazzera, B. A.; Kurtser, I. G.; McQuade, R. S. and Grossman, A. D. (1999): An autoregulatory circuit affecting peptide signaling in *Bacillus subtilis*, J Bacteriol (vol. 181), No. 17, pp. 5193-200.