Chapter 83

Comparative Analysis of the Complete Genome Sequence of the Plant Growth-Promoting Bacterium Bacillus amyloliquefaciens FZB42

RAINER BORRIS
ABITEP GmbH, Germany

83.1 INTRODUCTION

Bacteria that are associated with plant roots and exert beneficial effects on plant development are referred to as plant growth-promoting rhizobacteria (PGPR) (Kloepper et al., 1980; see Chapter 53). They competitively colonize plant roots and can simultaneously act as biofertilizers and as antagonists of plant pathogens, including bacteria, fungi, viruses, and nematodes (Biocontrol; see Chapter 54). The aerobic, endospore-forming rhizobacteria belonging to Bacillus amyloliquefaciens subsp. plantarum are known as being especially suitable for commercial use in enhancing yield of crop plants (“biofertilizer” function) and to suppress microbial plant pathogens (“biocontrol” function) (Borriss et al., 2011a; Borriss, 2011). The type strain FZB42T (=BGSC 10A6, DSM23117) was isolated from plant-pathogen-infested soil of a sugar beet field in Brandenburg, Germany (Krebs et al., 1998). FZB42 has been shown to act as an efficient PGPR (Bochow et al., 2001; Grosch et al., 1999; Idriss et al., 2002, Idris et al., 2007) and biocontrol bacterium (Koumoutsi et al., 2004; Chen et al., 2006; Chen et al., 2009a; Schmiedeknecht et al., 2001). Analysis of the whole genome of the plant root-associated bacterium FZB42 revealed an impressive capability to produce a vast array of secondary metabolites aimed to suppress competitive bacteria and fungi within the plant rhizosphere (Chen et al., 2007). Notably, it ruled out that more than 8.5% of the whole genome capacity was devoted to synthesis of antimicrobial secondary compounds. Recently, genome sequences from nonplant-associated B. amyloliquefaciens subsp. amyloliquefaciens strains (Geng et al., 2011; Yang et al., 2011; Zhang et al., 2011), including type strain DSM7T (Rueckert et al., 2011), have become available. In addition, we have completed the sequencing and annotation of the genomes of two further PGPR strains belonging to B. amyloliquefaciens subsp. plantarum, isolated from Chinese soil (Blom et al., 2012; HE617159; HE774679). We discuss here the novel genomic data, available for the two B. amyloliquefaciens subspecies, with their implications for our recent knowledge of this important group of rhizobacteria. Exploiting genomics for further research to identify genes important for plant–bacteria interactions, transposon mutagenesis, and a first proteomics approach are other topics of this chapter.

83.2 COMPARATIVE GENOME ANALYSIS OF ROOT-ASSOCIATED AND NONROOT-ASSOCIATED Bacillus amyloliquefaciens

The whole genome sequence of the type strain of plant-associated B. amyloliquefaciens subsp. plantarum,
FZB42\textsuperscript{T}, has been determined in 2007, as the first representative of Gram-positive, plant growth-promoting bacteria (Chen et al., 2007). Its 3918-kb genome, containing an estimated 3695 protein-coding sequences (CDS), lacks extended phage insertions, which occur ubiquitously in the related \textit{B. subtilis} 168 genome (Kunst et al., 1997; Barbe et al., 2009). The \textit{B. amyloliquefaciens} genome reveals a huge potential to produce secondary metabolites, including the polyketides bacillaene and difficidin. More than 8.5\% of the genome is devoted to synthesizing antibiotics and siderophores by pathways not involving ribosomes. A first comparison of its genomic sequence with that of the \textit{B. amyloliquefaciens}-type strain DSM\textsuperscript{7T} revealed significant differences in the genomic sequences of both strains (Rueckert et al., 2011). The strains have in common 3345 CDS residing in their core genomes (Borriss et al., 2011b), while 547 and 344 CDS were found to be unique in FZB42\textsuperscript{T} and DSM\textsuperscript{7T}, respectively. The core genome shared by both strains exhibited 97.89\% identity on the amino acid level. Notably, the gene clusters encoding nonribosomal synthesis of antibacterial polyketides difficidin and macrolactin are absent in DSM\textsuperscript{7T}. For comparison, \textit{B. subtilis} 168\textsuperscript{T} has a similar number of CDS in common with \textit{B. amyloliquefaciens} strains DSM\textsuperscript{7T} and FZB42\textsuperscript{T} (3222 and 3182 CDS, respectively) (Borriss et al., 2011a). Recently, except DSM\textsuperscript{7T}, the genomes of three other representatives of the \textit{B. amyloliquefaciens} subsp. \textit{plantarum} strains have become available (Geng et al., 2011; Yang et al., 2011; Zhang et al., 2011), enabling a comparative genome analysis of plant root-associated and free-living soil \textit{B. amyloliquefaciens} strains (Table 83.1).

<table>
<thead>
<tr>
<th>Strain</th>
<th>Accession (gb)</th>
<th>Size (bp)</th>
<th>GC content (%)</th>
<th>Protein tRNAs</th>
<th>rRNAs</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{B. amyloliquefaciens} subsp. \textit{amyloliquefaciens}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM\textsuperscript{7T}</td>
<td>FN976644</td>
<td>3,980,199</td>
<td>46.08</td>
<td>3,924</td>
<td>94</td>
<td>30</td>
</tr>
<tr>
<td>TA208</td>
<td>CM002627</td>
<td>3,975,511</td>
<td>45.83</td>
<td>4,089</td>
<td>70</td>
<td>30</td>
</tr>
<tr>
<td>XH7</td>
<td>CM002827</td>
<td>3,939,203</td>
<td>45.82</td>
<td>4,190</td>
<td>75</td>
<td>30</td>
</tr>
<tr>
<td>LL\textsuperscript{3}</td>
<td>CM002534</td>
<td>3,995,227</td>
<td>45.71 %</td>
<td>4,219</td>
<td>72</td>
<td>30</td>
</tr>
<tr>
<td>\textit{B. amyloliquefaciens} subsp. \textit{plantarum}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FZB42\textsuperscript{T}</td>
<td>CP000560</td>
<td>3,918,589</td>
<td>46.48</td>
<td>3,693</td>
<td>87</td>
<td>30</td>
</tr>
<tr>
<td>CAU 946</td>
<td>HL617159</td>
<td>4,019,861</td>
<td>46.51</td>
<td>3,823</td>
<td>95</td>
<td>30</td>
</tr>
<tr>
<td>YAU \textsuperscript{Y2}</td>
<td>HE774679.1</td>
<td>4,242,774</td>
<td>45.85</td>
<td>3,991</td>
<td>91</td>
<td>30</td>
</tr>
<tr>
<td>IT-45</td>
<td>CM001433</td>
<td>3,925,087</td>
<td>46.6</td>
<td>3,898</td>
<td>86</td>
<td>30</td>
</tr>
<tr>
<td>NAU \textsuperscript{B}</td>
<td>Not submitted</td>
<td>4,196,170</td>
<td>45.99</td>
<td>4,250</td>
<td>92</td>
<td>30</td>
</tr>
</tbody>
</table>
83.3 Genes Involved in Plant–Bacteria Interactions

Figure 83.1 Synteny plots generated by EDGAR comparing the core genomes of FZB42\textsuperscript{T} with 
Bacillus amyloliquefaciens plantarum CAU B946 (a), B. amyloliquefaciens amyloliquefaciens DSM7\textsuperscript{T} (b), Bacillus subtilis subtilis 168 (c), and Bacillus licheniformis DSM13\textsuperscript{T} (d). Despite increasing phylogenetic distances, synteny of the core genomes still remained, and no significant rearrangements were observed.

B. amyloliquefaciens-type strains DSM7\textsuperscript{T} and FZB42\textsuperscript{T} (Rueckert et al., 2011). These GI\textsuperscript{s} are defined by local deviations in the tetranucleotide usage patterns from the signature of the whole genome (Reva and Tuemmler, 2005). In some of the FZB42\textsuperscript{T} islands, additional features, such as adjacent rRNAs, remnants of phages, transposase-like sequences, and direct repeats, are indicative of horizontal gene transfer (Chen et al., 2007). The approximately 150 genes only occurring in the plant-associated subspecies “plantarum” were found mainly clustered in the GIs. Generally, such genes are termed singletons, meaning that no orthologous genes can be identified in any other strain of the comparison set (Borriss et al., 2011b).

A comparative analysis between the core genome sequences of B. amyloliquefaciens plantarum and B. amyloliquefaciens strains revealed that the plant-associated strains represent a distinct group and are discernable by their potential to synthesize a huge spectrum of different secondary metabolites. The phylogram based on computing of the Bacillus core genomes underlines this finding (Fig. 83.2). The genomic data separating plant-associated and nonplant-associated strains of B. amyloliquefaciens on the subspecies level are supported by the results of in silico DNA–DNA hybridization (DDH) (Borriss et al., 2011b) and MALDI-TOF MS analysis of cellular components, as well as the spectra of secondary metabolites present in culture fluids (Borriss et al., 2011a).

83.3 GENES INVOLVED IN PLANT–BACTERIA INTERACTIONS

The ability of B. amyloliquefaciens subsp. plantarum to colonize surfaces of plant roots is a prerequisite for phytostimulation. Rhizosphere competence is linked to the capability to form sessile, multicellular communities (biofilms; see Chapter 66). In liquid culture without
Phylogenetic tree reconstructed using the core genomes of several representatives of the *B. subtilis* species complex: *B. amyloliquefaciens* subsp. *plantarum*, *B. amyloliquefaciens* subsp. *amyloliquefaciens*, *B. subtilis* subsp. *subtilis*, and *B. licheniformis*. Bar, 0.1 substitutions per amino acids within the coding regions of the core genomes.

Figure 83.2

The shaking, FZB42T forms robust pellicles at the liquid–air interface, whereas domesticated *B. subtilis* strains usually form thin, fragile pellicles (Chu et al., 2006). Using a green fluorescent protein (GFP)-tagged derivative of FZB42, we recently studied the fate of bacterial root colonization in three different plant species. The studies were performed in a gnotobiotic system, in which FZB42 was applied to seedlings, axenically grown in soft agar containing Murashige–Skoog (MS) medium. They ruled out that the bacterium behaves distinctly in colonizing surfaces of plant roots of different species. In contrast to maize, FZB42 colonizes preferentially the root tips when colonizing *Arabidopsis* roots, for example (Fan et al., 2011; 2012). Root hairs of lettuce plantlets are also a preferred site of colonization (Fig. 83.3).

Figure 83.3

Confocal laser scanning microscopy of GFP-tagged FZB42 colonizing root hairs of lettuce. Source: The image was taken seven days after inoculation (with courtesy of Dr. Kristin Dietel).

An important precondition for root colonization is motility. FZB42T displays a robust swarming phenotype. The proteins encoded by swrA and swrB (vyjD), and the cyclic lipopeptide surfactin present in *B. amyloliquefaciens* plantarium are thought to be essential for swarming motility. These proteins permit colonization of surfaces and nutrient acquisition through their surface wetting and detergent properties (Kearns et al., 2005). The genome of FZB42T contains the complete set of genes implicated in biofilm- and fruiting-body formation in *B. subtilis*, including the 13-gene exopolysaccharide operon epsA-O, apparently required for producing an exopolysaccharide that holds chains of cells together in bundles (Kearns et al., 2005). The unique genes RBAM007750, RBAM007760, and RBAM007770 encode proteins with a collagen-related GXT structural motif and are probably involved in surface adhesion or biofilm formation (Chen et al., 2007). Notably, triple helix repeat-containing collagen proteins were not detected in other representatives of the *B. subtilis* species complex, except for *B. atrophaeus* and *B. pumilus*. However, in the plant-associated *B. amyloliquefaciens* YAU Y2 and NAU B3, genes homologous to RBAM007760 and RBAM007770, were present, while at the same genomic region a gene with putative cell wall anchor function was located in *B. amyloliquefaciens* CAU B946 and DSM7T. CAU B946 possesses a mucin-17-like protein bearing a repetitive PVX motif, unique in all *Bacillus* genomes presently known.

Several mechanisms are involved in plant-beneficiary activities of PGPR; they include synthesizing phytohormones (Idris et al., 2007; see Chapter 27) and volatile organic compounds (Ryu et al., 2003; see Chapter 63) and producing available nutrients for plants (van Loon, 2007) and producing available nutrients for plants (van Loon, 2007; Yang et al., 2008; see Chapter 54).
Analysis by high performance liquid chromatography (HPLC) and gas-chromatography-mass spectrometry (GC-MS) demonstrated the presence of indole-3-acetic acid (IAA) in culture filtrates of FZB42\(^\text{TM}\), \((\Delta_{\text{trp}}\text{BA and } \Delta_{\text{xpy}}\text{ED})\) and two other strains bearing knock-out mutations in genes probably involved in IAA synthesis, SsyE (putative IAA transacetylase) and AyzhX (putative nitri-
lase), produce smaller amounts of IAA than wild type. Three of these mutant strains are less efficient in promot-
ing plant growth (Idris et al., 2007).

A blend of volatile compounds (VOCs) is released by several PGPR \textit{Bacillus} strains, including FZB42\(^\text{TM}\) (Borr-
iss, 2011). The volatiles 3-hydroxy-2-butanoic acid and 2,3 butanediol trigger enhanced plant growth. To syn-
thesise 2,3-butanediol, pyruvate is firstly converted into acetolactate by acetolactate synthase (AlsS) under condi-
tions of low pH and oxygen starvation. The next step of this alternative pathway of pyruvate catabolism, conver-
sion of acetolactate to acetoin, is catalyzed by acetolactate decarboxylase (AlsD). The final step, from acetoin to 2,
3-butanol, is catalyzed by the \textit{bdh} \textit{A} gene product, ace-
toin reductase/2,3-butanediol dehydrogenase (Nicholson 2008). The FZB42\(^\text{TM}\) genome contains all the three genes encoding this pathway. FZB42\(^\text{TM}\) mutant strains, incapable of producing volatiles due to knock-out mutations intro-
duced into the \textit{alsS} and \textit{alsD} genes, are unable to support growth of \textit{Arabidopsis} seedlings (Borris, 2011).

As other environmental members of the \textit{B. subtilis} species complex, \textit{B. amyloliquefaciens} \textit{plantarum} secrete different hydrolas es, enabling them to use external cel-
lulosic and hemicellulosic substrates present in plant cell walls. Microbe-associated hydrolytic enzymes digesting plant cell wall structures, resulting in free oligosacchari-
ides, have been shown to act as elicitors of plant defense (Ebel and Scheel, 1997). A wide range of extracellu-
lar depolymerases, including \textit{A. glican}-, \textit{B. glican}-, and \textit{A. amylo-
lactase}- encoding enzymes are possibly involved in ben-
eficial plant–bacteria interactions, making low molecular weight hydrolysis products of macromolecules for plant uptake and nutrition available. Two of these enzymes, endo-1,4-ß-glucanase (“cellulase,” EC 3.2.1.4) and endo-
1,4-ß-xylanase (“xylanase,” EC 3.2.1.8) are encoded in \textit{B. subtilis} 168 by the \textit{amyA} and \textit{xynB} genes, respectively (Woll et al., 1995). The same genes were also present in the genomes of \textit{B. amyloliquefaciens} \textit{plantarum} but not in the \textit{B. amyloliquefaciens} \textit{amylo-
lactase} strains. Similarly, the genes \textit{syLa} involved in xylose degradation (EC 5.3.1.5), \textit{synH}, encoding 1,4-ß-
xylosidase (EC 3.2.1.71) and \textit{hplC}, encoding an endo-1,4-ß-glucanase (EC 3.2.1.4) are present in \textit{B. subtilis} and \textit{B. amyloliquefaciens} \textit{plantarum} but missing in the \textit{B. amyloliquefaciens} \textit{amy-
lactase} genomes. Unlike DSM7\(^\text{TM}\), the genomes of the \textit{B. amyloliquefaciens} \textit{plantarum} strains do not contain the \textit{amyA} gene encoding liquefying amylase but instead an \textit{amyE}-like sequence encoding a saccharifying amylase common in \textit{B. subtilis} (Chen et al., 2007). About 262 genes encoding proteins with putative secretion signals recognized by signal peptidases type I (150 gene products), type II (98 gene products), and the Tat system (14 gene products) have been found (Chen et al., 2007).

### 83.3.1 Proteomics: Response to Plant Root Exudate

The analysis of the effects of root exudates on expression of extracellular and cytosolic proteins by two-dimensional (2-D) gel electrophoresis revealed 17 and 21 proteins, respectively, as being significantly altered (Kierul, 2012). Most of the differentially expressed proteins were involved in nutrient acquisition, uptake, and utilization. \textit{AlsS} involved in synthesis of volatiles (see above), was 10-fold up-regulated in the presence of root exudate during transitional growth stage. Chitin-binding protein \textit{RBAM} 1754 and malate dehydrogenase (MdH) are greatly enhanced during stationary growth phase in the presence of root exudate. A cysteine protein involved in the conversion of proline to glutamate is induced in \textit{B. amyloliquefaciens}. Phytase is a prominent member of the FZB42\(^\text{TM}\) secretome and its concentration increases in the presence of plant root exudates. By contrast, a promoter mutation prevents phytase expression in the domesticated \textit{B. subtilis} (Makarewicz et al., 2006). The bacterial phytate-degrading activity could generate phytate phosph-
orous to nourish plants during phosphate starvation and thus promote their growth (Idris et al., 2002).

Root exudates also stimulated up-regulation of enzymes probably involved in response to oxidative stress generated in plant roots by thiol peroxidase or enzymes catalyzing compounds secreted by plant roots. Enzymes of this type include bacillopeptidase \textit{F}, \textit{3-glutamyl transpeptidase}, and phosphotransacetylase (Chen et al., 2007).

Our proteome analysis suggested that the hook-
associated flagellar proteins \textit{HAPI} and \textit{HAP2}, and the \textit{HAG} flagellin are differently affected by exudates secreted by maize roots. Although the level of the \textit{fliK} gene product \textit{HAPI} was reduced in the presence of root exudates, expression of \textit{fisD} and \textit{hag} gene products was up-regulated. Flagellin proteins are thought to elicit a general host plant basal defense against potential pathogens (Abramovitch et al., 2006).

### 83.3.2 Identification of Functional Genes by Transposon Mutagenesis

More genes, involved in plant growth promotion, were identified by transposon mutagenesis performed with the mariner transposon \textit{HIMAR} 1 introduced into FZB42\(^\text{TM}\).
wild-type cells (Budiharjo, 2011). Using a miniaturized test system based on L. minor plantlets (Idris et al., 2007), three genes were shown to be efficient in plant growth promotion: nfrA (ywH), abrB, and RBAM_017410.

The nfrA mutant exhibited a reduction in plant growth-promoting activity either in the Lemma system (26%) or Arabidopsis thaliana (40%).

The wild-type gene restored the growth-promoting activity in the mutant strain. The nfrA gene encodes an oxidoreductase that is induced under heat shock and oxidative stress. Disruption of the abrB gene by the TnYLB-1 transposon reduced the growth-promoting activity in Lemma (27%) and A. thaliana (17%). AbfB is a DNA-binding global regulator protein expressed during transition state and under conditions of limited growth. It is involved in the regulation of various cellular functions, such as antibiotic production, competence development, and expression of degradative enzymes, and it is likely that AbfB controls a pathway involved in plant–bacteria interactions. The mutant impaired in the unique RBAM_017410 gene showed a reduction of plant growth-promoting activity in both Lemma (42%) and Arabidopsis (33%). The deduced RBAM_017410 gene product is a small peptide with unknown function, similar to the small subunits of ribonucleoside-diphosphate reductase (NDP reductase). The enzyme catalyzes reduction of ribonucleotides to deoxyribonucleotides. Cells carrying that mutation were found to be not only altered in their morphology but also unable to colonize plant roots (Budiharjo, 2011).

83.4 THIRTEEN GENE CLUSTERS ARE INVOLVED IN BIOCONTROL

Representatives of B. amyloliquefaciens subsp. plantarum possess at least 13 gene clusters involved in nonribosomal and ribosomal synthesis of secondary metabolites with antimicrobial function. Two of them are encoding for nonribosomal peptide synthetases (NRPS, see 3.1) with hitherto unknown products. The cyclic lipopeptides, such as surfactin, and antifungal iturin-like compounds, including bacillomycin D and fengycin, are encoded by large gene clusters (Koumoutsi et al., 2004). Nonribosomal synthesis of the Fe-siderophore bacillibactin is also common in other members of the B. subtilis species complex. Three giant clusters encode polyketide synthases (PKSs) involved in nonribosomal synthesis of the antibacterial polyketides, bacillaeine, difficidin, and macroactin. Synthesis of difficidin and macroactin is unique for B. amyloliquefaciens plantarum (Borriossi et al., 2011b). Sfp-independent nonribosomal synthesis of bacilysin is common in B. amyloliquefaciens plantarum and other representatives of the B. subtilis species complex (Chen et al., 2009a). During the last two years or so, gene clusters were identified, mainly in FZB42T, that are involved in ribosomal synthesis of several bacteriocins, modified peptides with specific action against Gram-positive bacteria. Examples are mersacidin (Hartner et al., 20011), plantazolicin (Scheil et al., 2011), and amylocyclizin (Scholz, 2011). Figure 83.4 demonstrates that most of the 13 gene clusters mentioned above occur in the genomes of B. amyloliquefaciens plantarum, but with some notable exceptions. Gene clusters involved in plantazolicin and mersacidin synthesis and self-immunity are located within GIs acquired by horizontal gene transfer. The same is true for the orpfan nfr gene clusters detected in FZB42T and YAU 10601.

83.4.1 Nonribosomal Synthesis of Antimicrobial Lipopeptides and Polyketides

Polyketides and nonribosomal peptides comprise two families of natural products biosynthesized in a similar manner by multimodular enzymes acting in assembly line arrays. The monomeric building blocks are organic acids or amino acids, respectively (Walsh, 2004).

FZB42T genome analysis revealed the presence of numerous gene clusters involved in synthesis of cyclic lipopeptides (Koumoutsi et al., 2004) and polyketides (Chen et al., 2006; Schneider et al., 2007) with distinct antimicrobial action. Besides the five-gene clusters known in B. subtilis to mediate the nonribosomal synthesis of secondary metabolites, four additional giant clusters, bmyD, mln, dfn, and nrs, have been identified in FZB42T (Chen et al., 2009b). While the function of the two orphan nrs gene clusters found in FZB42T and NAU-Y2 remains still undetermined, the functions of the bmyD (bacillomycin D), mln (macroactin), and dfn (difficidin) genes are known (Koumoutsi et al., 2004; Chen et al., 2006; Schneider et al., 2007). As B. subtilis 168T, the genomes of B. amyloliquefaciens DSM7T and the other representatives of B. amyloliquefaciens harbor a significantly lower number of gene clusters involved in nonribosomal synthesis of secondary metabolites than strain FZB42T. Synthesis of lipopeptides and polyketides depends on Sfp, an enzyme that transfers 4′-phosphopantetheine from coenzyme A to the carrier proteins of nascent peptide or polyketide chains. An exception is the antibacterial dipetide bacilysin (Chen et al., 2009a).

Nonribosomal synthesis of lipopeptides is governed by very large protein templates called peptide synthetases (NRPS) that exhibit a modular organization, allowing polymerization of monomers in an assembly-line-like mechanism. The core of each module of a peptide synthetase is the adenylation (A) domain that catalyzes
83.4 Thirteen Gene Clusters are Involved in Biocontrol

Figure 83.4 Maps of the circular genomes of four *B. amyloliquefaciens* strains. Only gene clusters for surfactin, bacillibacin, bacillomycin D or iturin A, bacillibactin siderophore, and bacilysin synthesis are present in all strains. The giant gene clusters involved in nonribosomal synthesis of secondary metabolites are located in close vicinity of the termination site *ter* in which the bidirectional replication fork meets.

substrate recognition and activation of amino acids to acyl adenylates. The adenylation domain that specifies the amino acids used for peptide synthesis is flanked by two other essential components, the condensation (C) and the thiolation (T) domain, also named peptidyl carrier protein (PCP). Further domains can chemically modify the incorporated residues (e.g., epimerization and N-methylation). The PCP is equipped with cofactor P-pant (4’-phosphopantetheine) that serves as a swinging arm to mediate the peptide elongation reaction between adjacent modular units. It can be positioned at three sites necessary for amine acid loading and peptide elongation. A 4’-phosphopantetheine-transferase (PPTase) converts the apo-form of the PCP into its holo-form by loading the P-pant cofactor to an active serine. *B. subtilis* Sfp is a prototype of PPTases with wide substrate tolerance. The formation of a new peptide bond is catalyzed by the C-domain. A termination domain at the end of the module that incorporates the last amino acid catalyzes liberation or cyclization of the linear intermediate (Mootz and Marahiel, 1997). The assembly of the multifunctional proteins of the peptide synthetases is reflected in its genetic organization following the colinearity rule (Duitman et al., 1999).

Besides direct antagonism of phytopathogens, cyclic lipopeptides, such as surfactin and fengycin, can also interact with plant cells as bacterial determinants for turning on an immune response through the stimulation of the induced systemic resistance phenomenon (Ongena et al., 2007).

*Bacillomycin D* is a member of the iturin family that comprises iturin A, C, D, and E, bacillomycin F and L, bacillopeptin, and mycosubtilin. Members of the iturin family contain one β-amino fatty acid and seven D-amino acids (Chen et al., 2009b). The peptide moiety of the iturin lipopeptides contains a tyrosine in the D-configuration at the second amino acid position and two additional D-amino acids at positions 3 and 6 (Fig. 83.5). The members of the iturin family exhibit strong fungicidal activity and bacillomycin D has been identified as the main antifungal activity directed against fungal plant pathogens in *B. amyloliquefaciens* plantarum strains FZB42 and C06. Mycelium growth and spore germination are suppressed in
Fusarium oxysporum, Rhizoctonia solani, and Monilinia fructicola (Koumoutsi et al., 2004; Liu et al., 2011).

The gene cluster involved in bacillomycin D synthesis covers 37,293 bp and consists of four genes (Koumoutsi et al., 2004). The first open reading frame (ORF), bmyD, encodes a putative malonyl coenzyme A transacylase, which participates in fatty acid synthesis. The ORFs encoding the hybrid PKS/NRPS enzyme, BmyA, and the NRPS enzymes, BmyB and BmyC, are organized like their respective counterparts in the iturin A and mycosubtilin operons. Seven amino acid-activating modules are distributed in BmyA (A1), BmyB (B1-4), and BmyC (C1-2). In addition to the canonical A, C, and T domains, modules B1 (d-Tyr), B2 (d-Asn), and C1 (d-Ser) contain epimerization domains, directing conversion of the respective amino acids in a d-configuration. The last domain of this multienzyme system is a thioesterase domain, which is required for release and circularization of the synthesized lipopeptide molecule. Transcription of the bacillomycin D gene cluster is directly controlled by the global regulator DegU. A transmembrane protein of unknown function, YczE, is also necessary for the synthesis of bacillomycin D (Koumoutsi et al., 2007). While the majority of B. amyloliquefaciens plantarum strains were found to contain a gene cluster encoding bacillomycin D, strain CAU B946 was found to synthesize iturin A, which is controlled by the ituA operon located at the same site as the bmyD gene cluster in FZB42 (Blom et al., 2012).

Fengycin (synonymous to plipastatin) is a cyclic lipodecapeptide containing a β-hydroxy fatty acid with a side chain of 16–19 carbon atoms. Four d-amino acids and one nonproteinogenic ornithine residue have been identified in the peptide portion of fengycin. Fengycin is active against filamentous fungi and is known for inhibiting phospholipase A2. Similar to bacillomycin D, toxicity against pathogenic fungi relies mainly on their membrane permeabilization properties. Owing to its high productivity in synthesizing fengycin, biocontrol exerted by strain C06 relies rather on fengycin than on bacillomycin D (Liu et al., 2011).

The fen five-gene cluster occurring in FZB42 is related to the pps operon in B. subtilis 168 and is situated at the same locus as this strain, about 25 kb downstream from the bmyD gene cluster. It is detectable in the genomes of the B. amyloliquefaciens plantarum strains but occurs only fragmentarily in B. amyloliquefaciens-type strain DSM7T (Rueckert et al., 2011). The fen gene cluster covers the same size as the bmyD operon, although the five NRPS enzymes encoded by fenA-E contain 10 modules specifying a dekapptide. Modules A2, B2, C2, and D3 contain additional epimerization domains responsible for conversion to D-amino acids. FenE contains a thioesterase protein liable for release and cyclization of the lipopeptide.

Surfactin is a heptapeptide with an LLDLLDL chiral sequence linked by a β-hydroxy fatty acid consisting of 13–15 carbon atoms to form a cyclic lactone ring structure. Surfactin is surface active (biotenside) and acts hemolytic, antiviral, and antibacterial by altering membrane integrity (Peipoux et al., 1999). The biological role of surfactin is thought to be as supporting colonization.
of surfaces and acquisition of nutrients through its surface wetting and detergent properties. Mutants of *B. amyloliquefaciens*, blocked in surfactin biosynthesis, have been shown to be impaired in biofilm formation (Chen et al., 2007).

The giant gene cluster involved in the synthesis of surfactin consists of four genes covering 26 kb. Three of them encode the NRPS SrfAA, SrfAB, and SrfAC, while the terminal SrfAD is highly homologous to external thioesterases of type II. Modules present in SrfAA assemble the first three amino acids Glu, Leu, and Leu within the nascent peptide. The third module converts i-Leu into the \( l \)-configured isomer by its C-terminal epimerization domain. SrfAB incorporates Val, Asp, and p-Leu, while SrfAC is responsible for the activation and incorporation of the last leucine residue and catalysis of product release by cyclization.

While nonribosomal synthesis of the lipopeptide bacilysin does not, Bacilysin \([ l \text{-alanyl}-2,3\text{-epoxy-} \text{clohexamone}]-\text{d}-\text{alanine}\) contains an \( l \)-dipeptide bacilysin does not. Bacilysin \([ p\text{-pantyl} \text{-amino}-\text{acetic} \text{acid}]-\text{d}-\text{alanine}\) proceeds in a nonribosomal mode, catalyzed by amino acid ligase DhbE. Bacilysin is active in a wide range of bacteria and against the yeast, *Candida albicans*, due to the antipicacin moiety, which is released after uptake into susceptible cells. Culture fluids of a *B. amyloliquefaciens* strain suppress *Erwinia amylovora*, the causative agent of fire blight in *Erwinia amylovora* after uptake into susceptible cells. Culture fluids of a *B. amyloliquefaciens* strain are active in a wide range of bacteria and against the yeast, *Candida albicans*, due to the antipicacin moiety, which is released after uptake into susceptible cells. Culture fluids of a *B. amyloliquefaciens* strain suppress *Erwinia amylovora*, the causative agent of fire blight in *Erwinia amylovora* after uptake into susceptible cells.

While nonribosomal synthesis of the lipopeptide bacilysin does not, Bacilysin \([ l \text{-alanyl}-2,3\text{-epoxy-} \text{clohexamone}]-\text{d}-\text{alanine}\) contains an \( l \)-dipeptide bacilysin does not. Bacilysin \([ p\text{-pantyl} \text{-amino}-\text{acetic} \text{acid}]-\text{d}-\text{alanine}\) proceeds in a nonribosomal mode, catalyzed by amino acid ligase DhbE. Bacilysin is active in a wide range of bacteria and against the yeast, *Candida albicans*, due to the antipicacin moiety, which is released after uptake into susceptible cells. Culture fluids of a *B. amyloliquefaciens* strain suppress *Erwinia amylovora*, the causative agent of fire blight in *Erwinia amylovora* after uptake into susceptible cells.

Bacterial polyketides are a large family of secondary metabolites that include many bioactive compounds with antibacterial, immunosuppressive, antitumor, or other physiologically relevant properties. Bacterial polyketides are synthesized by type I PKSs, modularly organized assembly lines, starting from acyl CoA precursors by decarboxylative Claisen condensations. In general, their biosynthetic pathway follows the same logic as in nonribosomally synthesized peptides and requires at least three domains: an acyltransferase (AT), a ketosynthase (KS), and an acyl-carrier-protein (ACP) corresponding to PCP in NRSP. Usually the order of the modules dictates the sequence of biosynthetic events (Chen et al., 2009c).

A few years ago, a special class of PKSs that lack the cognate AT domain and require a separate AT enzyme acting iteratively in trans was detected in bacilli and other bacteria (Chen et al., 2007; Shen, 2003). In the genome of FZB42\(^1\), three giant PKS gene clusters are located at sites approximately 1.4 Mbp (pks2, mlu), 1.7 Mbp (pks1, bae), and 2.3 Mbp (pks3, dfn) clockwise distant from the origin of replication. Notably, none of the PKSs encoded by the three pks gene clusters harbor modules with cognate AT domains. Instead, one or more genes encoding single AT proteins were detected upstream from the megsynthases encoding genes in all three pks gene clusters of FZB42. Discrete ATs act iteratively by loading malonyl CoA onto all PKS modules during polyketide synthesis. Another notable feature observed in analysis of *Bacillus* polyketide gene clusters is the presence of noncanonical bimodular sequence KS KR ACP KS D D ACP (KS being a nonelongating KS) with a split downstream module on two proteins (Moldenhauer et al., 2007). With the help of cassette mutagenesis in combination with advanced mass spectrometric techniques, two polyketides, bacillaene and difficidin/oxydifficidin, have been identified, which are encoded by gene clusters pks1 (bae) and pks3 (dfn) (Chen et al., 2006). Gene cluster pks2 was assigned for the synthesis of macrolactin and the cluster has been renamed mlu. Bacillaene, but not difficidin, and macrolactin are produced in *B. subtilis* and *B. mojavensis*. The bacillaene gene cluster is also present in the closely related *B. amyloliquefaciens* subspecies *amyloliquefaciens* (Schneider et al., 2007).

**Bacillaene** is, due to its molecular structure, a highly unstable inhibitor of prokaryotic protein synthesis and does not have effects on eukaryotic organisms (Patel et al., 1995). NMR studies of partially purified extracts from *B. subtilis* revealed bacilaene as an open-chain, unsaturated enamine acid with an extended polyene system (Butcher et al., 2007). Recently, special features of bacillaene synthesis, the archetype of trans-AT PKS, have been uncovered in FZB42, and bacilaene II bearing a glucosyl moiety has been identified as the final product of the bae pathway (Moldenhauer et al., 2007, 2010).

Organization and genomic localization of the bae gene cluster is very similar to that of the pkX gene cluster of *B. subtilis* 168 (Fig. 83.6), but a putative regulator PksR, a homolog of *B. subtilis* PksA was detected as a discrete gene in a region far upstream from bae (Chen et al., 2007). Using strain OKB105, a derivative of *B. subtilis* 168 harboring an intact copy of sfp encoding P-pant-transferase, we assigned bacilaene as the synthesis product of the pkX gene cluster in *B. subtilis* 168 (Chen et al., 2006). The observation that the baeM gene of FZB42\(^1\) does not contain the superfluous last module present in the homologous pkM gene of *B. subtilis* suggests that this module is skipped during evolution. The presence of three discrete ATs that are iteratively
acting in trans, BaeC, BaeD, and BaeE is an interesting trait unique for the bae gene cluster. Our unpublished results performed with knock-out mutant strains (Chen XH and Borriss R) have revealed that the action of those trans-ATs is not restricted on bacillaene synthesis but has also some impact on synthesis of the two other polyketides produced by FZB42\textsuperscript{27}, difficidin and macrolactin, respectively. The final gene of the operon, baeS, was shown to encode a P450-like enzyme involved in the desaturation of bacillaene (Moldenhauer et al., 2007). Difficidin and oxydifficidin were identified as products of the dfn gene cluster in FZB42\textsuperscript{27} (Chen et al., 2006). Difficidin has been shown to inhibit protein biosynthesis (Zweerink and Edison, 1987), but the exact molecular target remains unknown. The polyketides are highly unsaturated 22-membered macrocyclic polyene lactone phosphate esters (Wilson et al., 1987) and are by far the most effective antibacterial compounds produced by FZB42\textsuperscript{27}. Notably, difficidin is efficient in suppressing plant pathogenic bacterium Erwinia amylovora, which causes fire blight disease in orchard trees (Chen et al., 2009a). Our studies have revealed that the production of difficidin is restricted to the representatives of B. amyloliquefaciens plantarum and the giant gene cluster encoding the respective megasynthases occurs only in that plant-associated subspecies of B. amyloliquefaciens (Schneider et al., 2007; Borriss et al., 2011b).

Unlike bae, but similar to the third pks gene cluster encoding macrolactin synthesis, the dfn operon harbors only one trans-AT, which presumably acts iteratively in transfer of the acyl moiety. A model for the biosynthesis of difficidin following largely the collinearity rule of KS module structure and chain elongation has been established (Chen et al., 2006). A candidate for the likely final step in difficidin biosynthesis, phosphorylation, is DifB, which resembles a hypothetical bacterial protein kinase and contains typical kinase-like sequence profiles (smart00220, COG0661).

Macrolactins are the biosynthesis product of the pks\textsubscript{2} (mln) gene cluster in FZB42\textsuperscript{27} and have been characterized as an inhibitor of peptide deformylase (Yoo et al., 2006). Macrolactins, originally detected in an unclassified deep-sea bacterium, contain three separate diene structure elements in a 24-membered lactone ring (Gustafson et al., 1989). 7-O-malonyl macrolactin induces disruptions of cell division, thereby inhibiting the growth of multiresistant enterococci (Romero-Tabarez et al., 2006). Macrolactin A and 7-O-succinyl macrolactin displays antibacterial properties against vancomycin-resistant enterococci and methicillin-resistant S. aureus (Kim et al., 2011). In the culture fluid of FZB42\textsuperscript{T}, four macrolactins were identified—macrolactin A and D as well as 7-O-malonyl- and 7-O-succinyl-macrolactin (Schneider et al., 2007).

The domain structure of the mln gene cluster of FZB42\textsuperscript{27} reflects the pathway, whereby the macrolactin skeleton is synthesized by extension of an acetyl starter unit by successive Claisen condensations with malonyl CoA. The polyketide megasynthase encoded by mlnB to mlnH comprises 11 modules, each containing at least the three basic domains—a KS, a ketoreductase (KR), and an ACP—but lacking an integrated AT domain. MlnA displays striking similarity to malonyl-CoA-specific trans-ATs and its activity should lead to incorporation of
the same extender unit in all modules. The completed polyketide is released and cyclized by the thioesterase at the C-terminus of the last module present in MlsH (Chen et al., 2007).

83.4.2 Ribosomal Synthesis of Bacteriocins

There are two different pathways for the synthesis of bioactive peptides, the nonribosomal mechanism (see 4.1) and the ribosomal synthesis of linear precursor peptides that are post-translationally modified and processed (lantibiotics and bacteriocin-like peptides). While numerous secondary metabolites (lipopeptides and polypeptides), synthesized independently from ribosomes in plant associated B. amyloliquefaciens, are known (see above), antimi-

83.4 Thirteen Gene Clusters are Involved in Biocontrol

B. amyloliquefaciens plantarum. Lantibiotics (lantionine containing antibiotics), a special group of bacteriocins, are amphiphilic peptides of bacterial origin and are nearly exclusively produced by Gram-positive bacteria. They contain unusual constituents such as nonproteinogenic didehydroamino acids and lanthionines (Sahl and Bierbaum, 1998). Interresidual thioether bonds are unique features of lantibiotics (Stein, 2005). Lanthionine Subtilin was the first lantibiotic isolated from a streptococcus and was discovered in 1970 (Lee et al., 2008). This cluster encodes a small gene cluster comprising genes for synthesis and modification of the prepeptide, and a transporter with an associated protease domain, mrsT. Genes encoding proteins for self-protection, mrsF, mrsE, and mrsG, were also found in the chromosome of FZB42T. Likewise, the two-component system MrbR2/K2 involved in ensuring “immunity” is present in FZB42T, while the single regulator MrsR1, essential for mersacidin biosynthesis, is without counterpart in FZB42T. The lantibiotic mersacidin is produced in YAU B946, which is more closely related to FZB42T than strain 10 (Stein et al., 2002). We characterized strain A1/3, a member of the B. amyloliquefaciens plantarum group (Borriess et al., 2005), and therefore, in the kind immunity genes of the operon (Borriess et al., 2011). The same is true for mersacidin, a type B lantibiotic and at first detected in Bacillus sp. HIL Y-85 (Chat-

893
structure conservation (Haft et al., 2010). The novel TOMM discovered from FZB42\textsuperscript{2} has a molecular mass of 1335 Da and has been named \textit{plantazolicin} (PZN) (Scholz et al., 2011). The chemical structure of \textit{plantazolicin} \textit{A} (R=CH) and \textit{B} (R=H) has been elucidated by resolving ESI-MSMS, 2D \textit{H}-\textit{C}-correlated NMR spectroscopy, as well as \textit{H}-\textit{N}-HMOC/\textit{H}-\textit{N}-HMBC NMR experiments. \textit{\textsuperscript{15}N}-labeling before the experiments facilitated the structure determination, unveiling a hitherto unusual number of thiazoles and oxazoles formed from a linear 14mer precursor peptide (Kalyon et al., 2011). PZN exhibited a highly selective antibiotic activity toward \textit{Bacillus anthracis}, but no other tested human pathogen (Molohon et al., 2011).

As in the case of mersacidin, biosynthetic \textit{pzn} genes are located in a variable part of the genome within GI 3, together with unique genes involved in restriction and modification of DNA. They are transcribed into two poly-cistronic mRNAs (\textit{pznFGHI} and \textit{pznCDBEL}) and a monocistronic mRNA for \textit{pznA} (Fig. 83.8) as revealed by reverse transcriptase polymerase chain reaction (RT-PCR) (Scholz et al., 2011). It is assumed that the leader peptide of \textit{PznA} serves as recognition site for the \textit{PznABC} synthetase complex (see above). Methyltransferase \textit{PznL} performs \textit{N,N}-dimethylation presumably after proteolytic processing by the protease \textit{PznE}. The complete \textit{pzn} gene cluster was also detected in the genome of \textit{B. amyloliquefaciens} amyloliquefaciens or \textit{B. subtilis}, a highly conserved homolog of the \textit{pzn} gene cluster was detected in the more distantly related \textit{B. pumilus}.

Last but not the least a very recent finding obtained after mariner transposon mutagenesis of the FZB42 mutant strain RS06 should be mentioned here. Besides mersacidin and \textit{plantazolicin}, a circular bacteriocin with similarity to \textit{uberolysin}, named \textit{amylocyclizin}, AZN, was identified (Scholz, 2011). The peptide suppressed growth of the plant pathogenic actinobacterium \textit{Clavibacter michiganensis} and of several Gram-positive bacteria. An \textit{azn} operon consisting of six genes was proven to be responsible for the synthesis of the bacteriocin, but further investigations are necessary to reveal its exact molecular structure.

### 83.5 Conclusions and Outlook

During the last 10 years, a considerable knowledge base about Gram-positive PGPR has been accumulated (see Chapter 53). Especially a group of root-surface associated bacilli, closely related to industrial \textit{B. amyloliquefaciens} strains, has gained increasing attention. These bacteria have been recognized as members of subspecies YAU B9601 contained only a small part of the gene cluster directing immunity against \textit{mersacidin} in FZB42\textsuperscript{2} shown at the bottom.
Acknowledgments

“plantarum” clearly distinct from the type strain \textit{B. amyloliquefaciens} DSM7\textsuperscript{T} (Borriss et al., 2011b). \textit{B. amyloliquefaciens} FZB42\textsuperscript{T} has been proposed as type strain for the new subspecies. The strain is distinguished by several features:

1. FZB42\textsuperscript{T} and its genomic sequence (CP000560) are freely available for scientific investigations from the BGSC stock center (strain number 10A6) and DSMZ (strain number DSM23117). It is freely available despite the fact that the strain is commercialized by ABiTEP GmbH Berlin and successfully used in agri- and horticulture applications (http://www.abitep.de/en/Research.html; see also Chapter 64).

2. In contrast to most environmental \textit{Bacillus} strains, \textit{B. amyloliquefaciens} FZB42\textsuperscript{T} is naturally competent and amenable for genetic transformation. In order to assign unknown gene functions, a large mutant strain collection is available. Some are already deposited at BGSC, and the others can be obtained after request from the Research department of ABiTEP GmbH.

3. A GFP-labeled derivative of FZB42 useful for root colonization studies is also available.

4. A transposon library of FZB42 using the mariner transposon TnYLB-1 has been generated and can be used to screen for genes involved in plant growth promotion and biocontrol.

5. Microarrays with oligonucleotides, representing the complete set of FZB42 genes and numerous intergenic sequences harboring candidate small (regulatory) RNAs for transcriptomic studies, are available.

6. \textit{B. amyloliquefaciens} FZB42\textsuperscript{T} is closely related to other known PGP bacilli used commercially (QST713, GB03, A1/3). Studies performed with FZB42\textsuperscript{T} are valuable for a general understanding of the mode of action of that interesting group of bacteria.

As already proposed (Borriss, 2011), it is hoped that the community of scientists, interested on exploiting plant growth-promoting (PGP) bacilli, will choose FZB42 as a paradigm for research on that group of bacteria with a high potential to substitute or at least to replace, in part, harmful agrochemicals.

ACKNOWLEDGMENTS

The author thanks all the former members of his laboratory at Humboldt University and the many colleagues.
and friends with whom he worked together in the exciting field of plant growth promotion and biocontrol during the last decade. Their fruitful collaboration and innovative work made it possible to write this chapter. Special thanks to the German Ministry of Education and Research (BMBF) for financial support given in the framework of the Chinese–German collaboration program, the GenoMik and GenoMikPlus network, and recently in the PATHCONTROL project.

REFERENCES


References

897


Chapter 83 Comparative Analysis of the Complete Genome Sequence of the Plant Growth