Natural Products Hot Paper

DOI: 10.1002/anie.201404572

# Paenilamicin: Structure and Biosynthesis of a Hybrid Nonribosomal Peptide/Polyketide Antibiotic from the Bee Pathogen *Paenibacillus larvae***

*Sebastian Müller, Eva Garcia-Gonzalez, Andi Mainz, Gillian Hertlein, Nina C. Heid, Eva Mösker, Hans van den Elst, Herman S. Overkleeft, Elke Genersch, and Roderich D. Süssmuth**

*Dedicated to Professor Eckhard Ottow*

**Abstract:** *The spore-forming bacterium Paenibacillus larvae is the causative agent of American Foulbrood (AFB), a fatal disease of honey bees that occurs worldwide. Previously, we identified a complex hybrid nonribosomal peptide/polyketide synthesis (NRPS/PKS) gene cluster in the genome of P. larvae. Herein, we present the isolation and structure elucidation of the antibacterial and antifungal products of this gene cluster, termed paenilamicins. The unique structures of the paenilamicins give deep insight into the underlying complex hybrid NRPS/PKS biosynthetic machinery. Bee larval co-infection assays reveal that the paenilamicins are employed by P. larvae in fighting ecological niche competitors and are not directly involved in killing the bee larvae. Their antibacterial and antifungal activities qualify the paenilamicins as attractive candidates for drug development.*

Honey bees (*Apis mellifera*) are instrumental to pollination in wild ecosystems and are the single most important cultured species used in agriculture for crop pollination. Our food supply relies heavily on healthy honeybee colonies, and the growing pollination demand in agriculture combined with the decline in the global stock of both managed honey bees and wild pollinators is perceived as a major threat to human civilisation.[1,2] Infectious species that cause disease in honey bees are therefore under close scrutiny and a number of epidemic diseases threatening to depopulate large areas of agricultural land from honey bee colonies have caused major concern in recent years.[3] One such pathogen is the bacterium *Paenibacillus larvae*, which is the causative agent of the epizootic American Foulbrood (AFB),[4] the most destructive bacterial disease affecting honey bees. Despite the enormous impact of AFB, the underlying molecular mechanisms remain largely elusive. This lack of knowledge makes AFB difficult to fight and it is therefore imperative that biological pathways characteristic for *P. larvae* are mapped and evaluated for points of interference.

One attractive strategy to gain insight into the molecular physiology of pathogenic bacteria is to investigate metabolic pathways unique to the species of interest. Often such pathways render secondary metabolites employed by the bacterium for either host infection or defence against competing microorganisms. In such a strategy, biochemical pathways are charted that are potentially essential for the bacterium and could thus be exploited as targets for pesticide development. The genome of *P. larvae* harbours gene clusters coding for the synthesis of peptide-based metabolites.[5,6] In this framework, we previously unearthed a remarkably complex hybrid nonribosomal peptide synthetase (NRPS)/polyketide synthase (PKS) gene cluster (*pam*) that is characteristic for *P. larvae* and linked this cluster to an intriguing mixture of natural products that possess both antibacterial and antifungal bioactivity.[5] Herein, we present the isolation and structure elucidation of four bioactive nonribosomal peptide (NRP)/polyketide hybrids assigned to the *pam* gene cluster, hereafter named paenilamicins A1, A2, B1 and B2, by means of HPLC–ESI-MS, GC–MS, and NMR spectroscopy.

With the aim of identifying the secondary metabolites of the *pam* gene cluster from *P. larvae* DSM25430, culture supernatant was fractionated by XAD-16 adsorption chromatography and the fractions were tested for antibacterial activity against *Bacillus megaterium*. The bioactive fractions showed dominant mass peaks at $m/z$ 1009.67285 $[M+H]^+$ (paenilamicin A1) and 1037.68042 $[M+H]^+$ (paenilami-

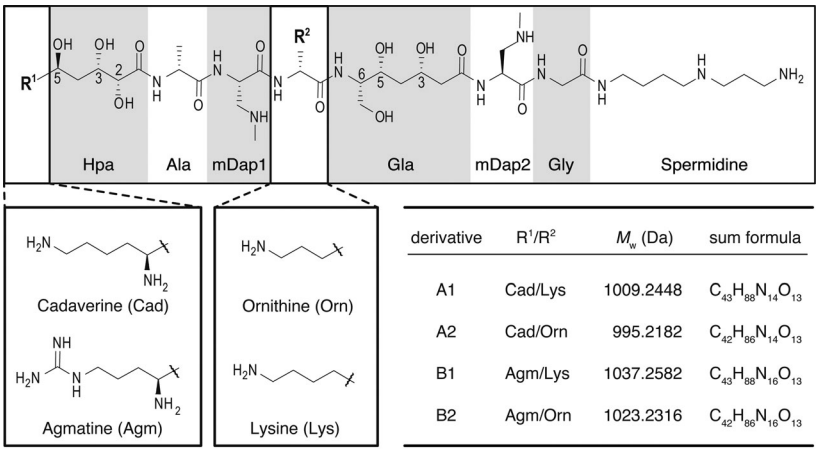
[*] S. Müller, Dr. A. Mainz, E. Mösker, Prof.Dr. R. D. Süssmuth
Institut für Chemie, Technische Universität Berlin
10623 Berlin (Germany)
E-mail: suessmuth@chem.tu-berlin.de
Homepage: http://www.biochemie.tu-berlin.de
Dr. E. Garcia-Gonzalez, G. Hertlein, N. C. Heid,
Priv.-Doz. Dr. E. Genersch
Institute for Bee Research
Department of Molecular Microbiology and Bee Diseases
16540 Hohen Neuendorf (Germany)
H. van den Elst, Prof. Dr. H. S. Overkleeft
Dept. of Bio-organic Synthesis, Leiden University
Einsteinweg 55, 2333 CC Leiden (The Netherlands)
Priv.-Doz. Dr. E. Genersch
Institute of Microbiology and Epizootics, Freie Universität Berlin
Robert-von-Ostertag-Str. 7-13, 14163 Berlin (Germany)

[**] This work was supported by the Cluster of Excellence "Unifying Concepts in Catalysis" funded by the DFG and coordinated by the TU Berlin, as well as by grants from the Ministries for Agriculture from Brandenburg and Sachsen-Anhalt, Germany, and through the German Research Foundation (Graduate School 1121).

Supporting information for this article (including experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201404572.

cin B1), and two less abundant peaks with *m/z* 995.65837 $[M+H]^+$ (paenilamicin A2) and 1023.66402 $[M+H]^+$ (paenilamicin B2) in LC–MS experiments (Table S1 in the Supporting Information). Subsequently these compounds were purified by preparative reversed phase (RP)-HPLC (see the Supporting Information). GC–MS analysis revealed the presence of D-Ala, Gly, D-Lys/D-Orn and spermidine (Figures S2–S4 in the Supporting Information). Full structure elucidation of the four derivatives was achieved through multidimensional NMR spectroscopy (Figure S1 and Table S2) to reveal a unique core structure composed of the building blocks 2,3,5-trihydroxy pentanoic acid (Hpa), Ala, *N*-methyl-diaminopropionic acid (mDap), galantinic acid (Gla), Gly, and 4,3-spermidine (Spe; Figure 1). This structure is reminiscent of galantin, an antibiotic isolated from a poorly characterized environmental isolate belonging to the family Bacillaceae.[7,8] We unambiguously identified two positions at which the four derivatives differ: the *N*-terminal fragment is either occupied by a decarboxylated Lys (cadaverine, Cad) moiety or by a decarboxylated Arg (agmatine, Agm) moiety, thus giving rise to the paenilamicin A and paenilamicin B series, respectively. Each of these may then contain an internal Lys (paenilamicins A1 and B1) or Orn residue (paenilamicins A2 and B2; Figure 1). NMR spectroscopic evidence for the differing moieties in the paenilamicins comes from fingerprint $^{1}H$-$^{13}C$ correlation spectra (Figure 2A,B). The presence of agmatine residues in the paenilamicin B series is further supported by the observation of guanidine-Nε-Hε correlation peaks with a characteristic $^{15}N$ chemical shift of 84.3 ppm (Figure 2D). *N*-methylation of both diaminopropionic acid residues in each of the paenilamicins was confirmed by the detection of two resolved *N*-methyl signals in the 1D-$^{13}C$ spectra (Figure 2C).




| derivative | $R^1/R^2$ | $M_w$ (Da) | sum formula |
|---|---|---|---|
| A1 | Cad/Lys | 1009.2448 | $C_{43}H_{88}N_{14}O_{13}$ |
| A2 | Cad/Orn | 995.2182 | $C_{42}H_{86}N_{14}O_{13}$ |
| B1 | Agm/Lys | 1037.2582 | $C_{43}H_{88}N_{16}O_{13}$ |
| B2 | Agm/Orn | 1023.2316 | $C_{42}H_{86}N_{16}O_{13}$ |

***Figure 1.*** Structures of the paenilamicins. Top: main scaffold of the paenilamicins with the modulated fragments $R_1$ and $R_2$ highlighted. The structures of the corresponding cadaverine, agmatine, ornithine, and lysine moieties are shown in the bottom left panels. The different combinations of these building blocks are summarized at the bottom right together with the corresponding molecular masses and formulae.

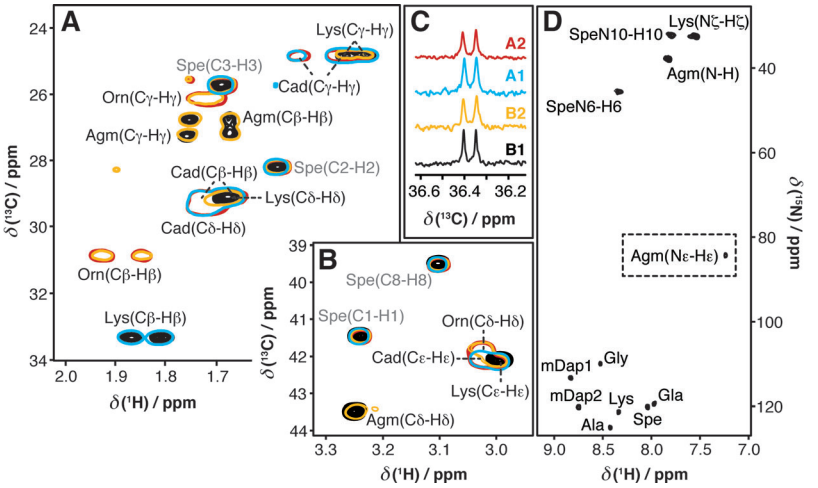



***Figure 2.*** NMR spectroscopic assignment of differences between the paenilamicins. A, B) Sections of the $^{1}H$-$^{13}C$ HSQC spectra showing side-chain resonances of the ornithine (Orn), lysine (Lys), cadaverine (Cad), and agmatine (Agm) residues in paenilamicin A1 (cyan), A2 (red), B1 (black), and B2 (yellow). Residual signal intensity of lysine cross-peaks in the spectra of paenilamicins A2 and B2 originate from minor contaminations with paenilamicins A1 and B1, respectively. C) A section of 1D $^{13}C$ spectra that demonstrates the presence of two *N*-methyl signals (mDap) in each of the four derivatives [color code as in (A)]. D) A natural abundance $^{1}H$-$^{15}N$ SOFAST-HMQC spectrum of paenilamicin B1 with the Nε-Hε signal of the guanidine moiety of agmatine highlighted with a dashed box. Resonance assignments for the backbone amides ($\delta(^{15}N) > 100$ ppm) and free amino groups ($\delta(^{15}N) < 50$ ppm) are also indicated.

According to genome data and gene inactivation experiments,[5,6] all four paenilamicins are synthesized by the same NRPS/PKS biosynthesis assembly line (Figure 3). The structure elucidation of the dedicated products of the *pam* gene cluster enabled us to fully reconstruct the paenilamicin biosynthesis pathway (Figure 3). Based on the structural data for paenilamicins A1 and B1, NRPS1 (*pam*A) shows relaxed substrate specificity and is thus capable of activating both Lys (paenilamicin A series) and Arg (paenilamicin B series). The substrate is then further processed through PKS1-mediated Claisen condensation with malonyl-CoA followed by ketoreductase (KR)-mediated formation of a secondary alcohol. The PamB PKS/NRPS hybrid then performs another Claisen condensation (PKS2) with hydroxymalonyl-CoA and subsequent ketoreduction. NMR analysis revealed a (2*R*,3*S*,5*S*) configuration for 2,3,5-trihydroxy pentanoic acid. This extended building block is condensed with D-Ala by NRPS2. No epimerization domain for the stereochemical conversion of L-Ala was found by annotation of NRPS2. Nevertheless, we identified an alanine racemase (NCBI ID:YP_008966234.1) in the genome

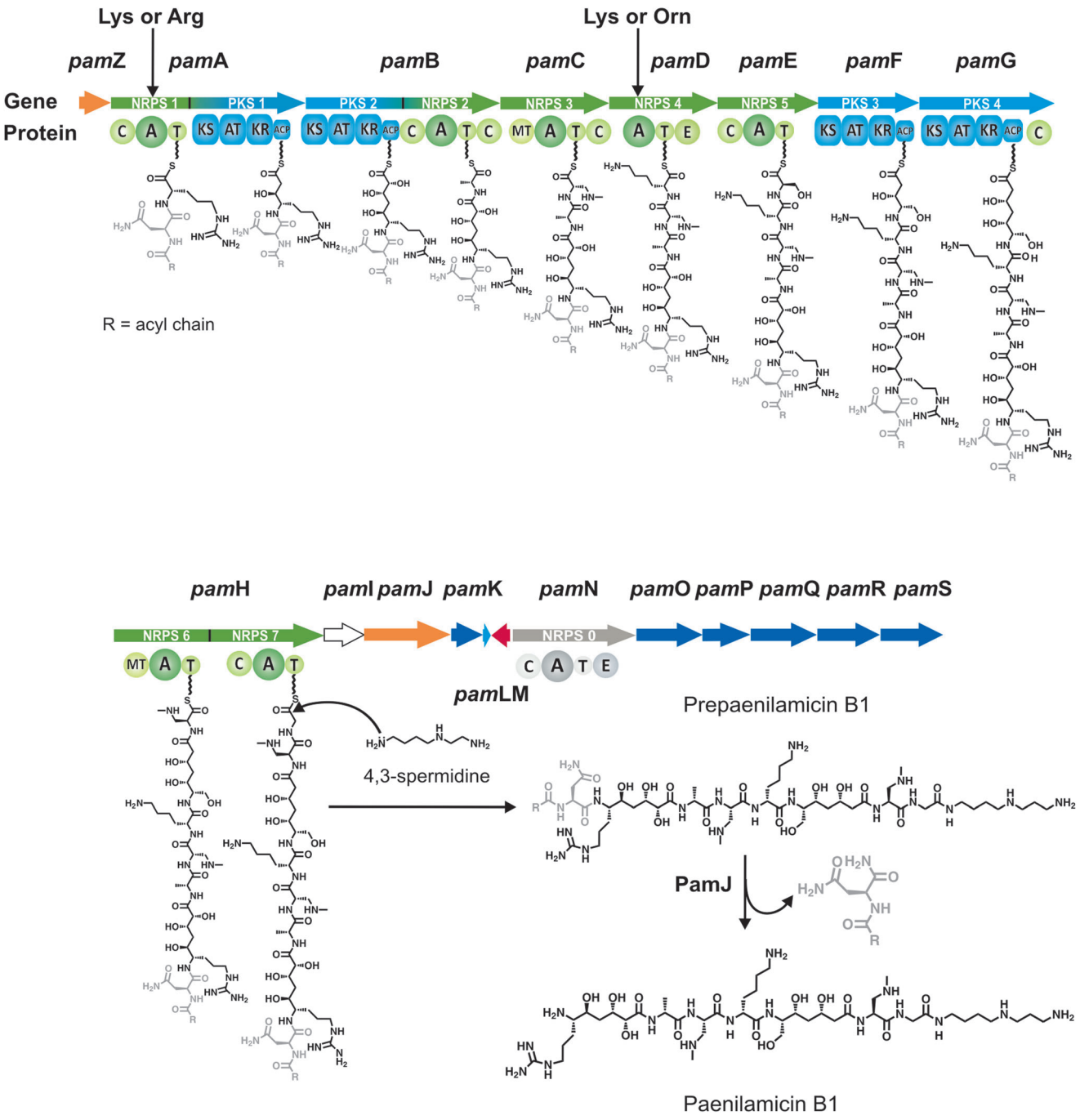



***Figure 3.*** The *pam* gene cluster from *Paenibacillus larvae* DSM25430 (NCBI ID: NC_023134.1) and biosynthesis of the paenilamicins. A proposed assembly line for the polyketide/peptide backbone of prepaenilamicin B1 and its activation to give paenilamicin B1 is shown. This process starts with the activation of acyl-D-Asn (grey; R = fatty acid) by NRPS0 (PamN). Prepaenilamicin B1 is released in a TE-independent fashion through the nucleophilic condensation of 4,3-spermidine to give Gly. Prepaenilamicin is potentially activated by the transmembrane transporter/peptidase PamJ. Arrows above NRPS1 and NRPS4 indicate substrate specificity. Genes responsible for precursor biosynthesis (dark blue), resistance/transport (orange), and transcriptional regulation (red), as well as genes of unknown function/hypothetical proteins (white) are indicated. C = condensation domain; A = adenylation domain; T = thiolation domain; TE = thioesterase domain; E = epimerization domain; MT = methylation domain; KS = ketosynthase domain; AT = acyl transferase domain; KR = ketoreductase domain; ACP = acyl carrier protein.

of *P. larvae*, which could be responsible for the generation of the incorporated D-Ala.[6] Alanine racemases catalyze the pyridoxal phosphate (PLP)-mediated conversion of L-Ala into D-Ala, which is an important component of bacterial petidoglycans.[9] Despite the absence of epimerization domains in the synthetases of cyclosporin[10] and HC toxin,[11] these natural products contain D-Ala, which is likewise provided by alanine racemases.[12,13]

The biosynthesis of the paenilamicins continues with the condensation of L-mDap, D-Lys/D-Orn, and L-Ser in reactions

catalyzed by the NRPS3–5 gene products PamC, PamD, and PamE. The incorporation of L-mDap is genetically supported by the presence of an *N*-methyltransferase (*N*-MT) domain in NRPS3 (PamC), as well as in NRPS6 (PamH). Although *N*-methylation is the most common methylation reaction in nonribosomal peptide synthesis,[14] it usually occurs at the peptide bond. Rare examples of side-chain *N*-methylation include pristinamycins $I_A$ und $I_B$, in which the nitrogen atom of the 4-amino-L-phenylalanine side chain is singly or doubly methylated.[15] The A-domain of NRPS4 (PamD) shows relaxed substrate specificity and is able to process D-Lys (paenilamicin A1 and B1) and D-Orn (paenilamicin A2 and B2).

After the condensation of L-Ser, the next step in the paenilamicin biosynthesis involves PKS3 (PamF) and PKS4 (PamG), which perform a pair of sequential Claisen condensations with malonyl-CoA. Each condensation step is followed by ketoreduction, which according to NMR data, is *R*-selective at position 5 and *S*-selective at position 3. Next, the bimodular NRPS PamH effects the attachment of a second L-mDap (NRPS6) and subsequently a Gly (NRPS7) residue. The presence of a C-terminal 4,3-spermidine residue (Figure S4 and S5) suggests chain termination by nucleophilic cleavage from the NRPS/PKS assembly line. The gene product of *pam*I shows homology to the BtrH protein family (Table S3). Interestingly, we found the same domain at the C-terminus of PamH, thus suggesting a role in the transpeptidase reaction, which transfers the peptide from the T-domain to the spermidine residue. In the biosynthesis of butirosin, the BtrH protein is described as an aminoglycoside 1-*N*-acyltransferase responsible for the ligation of the aminoglycoside side chain of butirosin.[16] This points to a potentially new peptide-release mechanism in thioesterase (TE)-independent NRPSs, a topic that is currently under investigation in our group.

The incorporation of (2*R*)-hydroxymalonyl by PKS1 is supported by the presence of *pam*P, *pam*O, *pam*K, and *pam*Q, which encode for the biosynthesis of (2*R*)-hydroxymalonyl-CoA from 1,3-bisphosphoglycerate (Figure S6 A). Furthermore, the gene products of *pam*S and *pam*R assist in the biosynthesis of L-Dap from *O*-phospho-L-Ser and L-Glu (Figure S6 B). Sequence alignments predict PamZ to be an acetyltransferase with homology to the zwittermicin-resistance protein ZmaR.[17]

The given functional assignment of gene products left out the role of PamN (NRPS0), which is a single NRPS module with an A-domain specific for Asn and an E-domain for conversion into D-Asn. This arrangement is highly reminiscent of the biosynthesis of the prodrug forms of xenocoumacins.[18] The xenocoumacin gene cluster harbors a gene coding for a bifunctional protein with a periplasmatic peptidase and three transmembrane helices (XcnG; type I architecture) that is involved in the resistance mechanism. The peptidase subunit is responsible for the specific cleavage of the acylated D-Asn moiety of the pre-xenocoumacins to produce the bioactive xenocoumacins.[18] Although PamJ shows structural differences to XcnG, a comparable biosynthesis and prodrug-activation mechanism for paenilamicin appears very likely (Figure 3).

We reported that paenilamicin is involved in the time-course of infection but does not influence total larval mortality in vivo.[5] This led us to assume that the main function of paenilamicin is not to kill bee larvae. Hence, we established co-infection assays with *P. alvei*, in which larvae were co-infected with *P. larvae* and *P. alvei*. After larval death, the remains were tested for *P. alvei* survival (Figure 4).

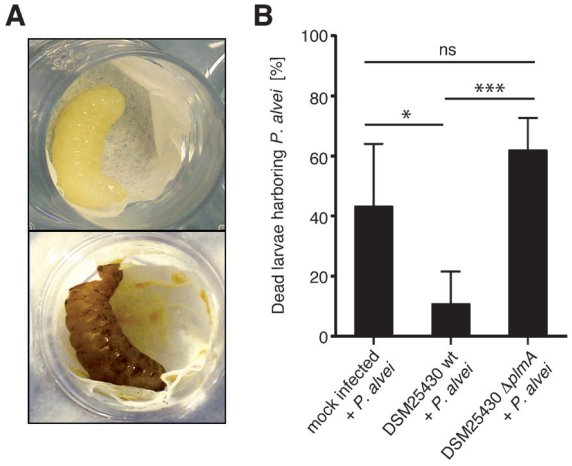



***Figure 4.*** The in vivo effect of paenilamicin. A) Mock infected healthy larva (top) and a moribund larva infected with *P. larvae* (bottom) of same age. B) Larvae reared on a diet containing *P. alvei* spores were mock infected or infected with either *P. larvae* wt or *P. larvae* Δ*pam*A. The remains of dead larvae were analyzed for the growth of *P. alvei* and the proportion of dead larvae that allow the recovery of *P. alvei* was determined. The bars show the mean + SD of 3 independent infection assays with 30 larvae each and were analyzed by Student's t-test; **p*-value < 0.05, ****p*-value < 0.001, SD = standard deviation.

Uninfected control groups showed the robustness of the assay (mortality < 20%). Control groups infected with *P. alvei* showed a recovery rate for this bacterium of $43.3 \pm 20.8\,\%$. When larvae were co-infected with *P. larvae* DSM25430 wild-type (wt), the recovery rate of *P. alvei* decreased significantly to $10.9 \pm 10.8\,\%$ (Student's t-test *p*-value = 0.042). Interestingly, larvae co-infected with *P. alvei* and *P. larvae* DSM25430 Δ*pam*A gave a higher recovery of *P. alvei* ($62.0 \pm 10.9\,\%$), a value comparable to that for the group infected with *P. alvei* only (non-significant difference, *p*-value > 0.05), but significantly higher than the value obtained from larvae co-infected with *P. alvei* and *P. larvae* DSM25430 wt (Student's t-test *p*-value = 0.0006). These results clearly show that paenilamicin production impacts potentially competing bacterial populations present in the larval gut.

In conclusion, we have elucidated the structures of four paenilamicins and assigned their biosynthesis to a complex NRPS/PKS hybrid assembly line, thereby revealing some fascinating biosynthetic features. Co-infection assays in combination with gene inactivation shed light on the in vivo function of these compounds, the virulence mechanisms of *P. larvae*, and the etiopathogenesis of AFB. The interesting bioactivity of paenilamicins could make them lead structures with potential for human or veterinary drug development programs.

Received: April 29, 2014
Published online: July 30, 2014

**Keywords:** antibacterial agents · nonribosomal peptide synthesis · *Paenibacillus larvae* · polyketide synthesis · virulence factors

[1] M. A. Aizen, L. A. Garibaldi, S. A. Cunningham, A. M. Klein, *Curr. Biol.* **2008**, *18*, 1572 – 1575.
[2] M. A. Aizen, L. D. Harder, *Curr. Biol.* **2009**, *19*, 915 – 918.
[3] E. Genersch, *Appl. Microbiol. Biotechnol.* **2010**, *87*, 87 – 97.
[4] E. Genersch, E. Forsgren, J. Pentikainen, A. Ashiralieva, S. Rauch, J. Kilwinski, I. Fries, *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 501 – 511.
[5] E. Garcia-Gonzalez, S. Müller, G. Hertlein, N. Heid, R. D. Süssmuth, E. Genersch, *Microbiologyopen*, DOI: 10.1002/mbo3.195.
[6] M. Djukic, E. Brzuszkiewicz, A. Fünfhaus, J. Voss, K. Gollnow, L. Poppinga, H. Liesegang, E. Garcia-Gonzalez, E. Genersch, R. Daniel, *PLoS One* **2014**, *9*, e90914.
[7] J. Shoji, R. Sakazaki, Y. Wakisaka, K. Koizumi, M. Mayama, S. Matsuura, *J. Antibiot.* **1975**, *28*, 122 – 125.
[8] N. Sakai, Y. Ohfune, *J. Am. Chem. Soc.* **1992**, *114*, 998 – 1010.
[9] H. Hayashi, H. Wada, T. Yoshimura, N. Esaki, K. Soda, *Annu. Rev. Biochem.* **1990**, *59*, 87 – 110.
[10] G. Weber, K. Schorgendorfer, E. Schneider-Scherzer, E. Leitner, *Curr. Genet.* **1994**, *26*, 120 – 125.
[11] J. S. Scott-Craig, D. G. Panaccione, J. A. Pocard, J. D. Walton, *J. Biol. Chem.* **1992**, *267*, 26044 – 26049.
[12] Y. Q. Cheng, J. D. Walton, *J. Biol. Chem.* **2000**, *275*, 4906 – 4911.
[13] K. Hoffmann, E. Schneider-Scherzer, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **1994**, *269*, 12710 – 12714.
[14] M. Z. Ansari, J. Sharma, R. S. Gokhale, D. Mohanty, *BMC Bioinf.* **2008**, 9:454, DOI: 10.1186/1471-2105-9-454.
[15] V. de Crécy-Lagard, V. Blanc, P. Gil, L. Naudin, S. Lorenzon, A. Famechon, N. Bamas-Jacques, J. Crouzet, D. Thibaut, *J. Bacteriol.* **1997**, *179*, 705 – 713.
[16] N. M. Llewellyn, Y. Li, J. B. Spencer, *Chem. Biol.* **2007**, *14*, 379 – 386.
[17] E. A. Stohl, S. F. Brady, J. Clardy, J. Handelsman, *J. Bacteriol.* **1999**, *181*, 5455 – 5460.
[18] D. Reimer, K. M. Pos, M. Thines, P. Grün, H. B. Bode, *Nat. Chem. Biol.* **2011**, *7*, 888 – 890.