Polyketide Biosynthesis

DOI: 10.1002/anie.200905467

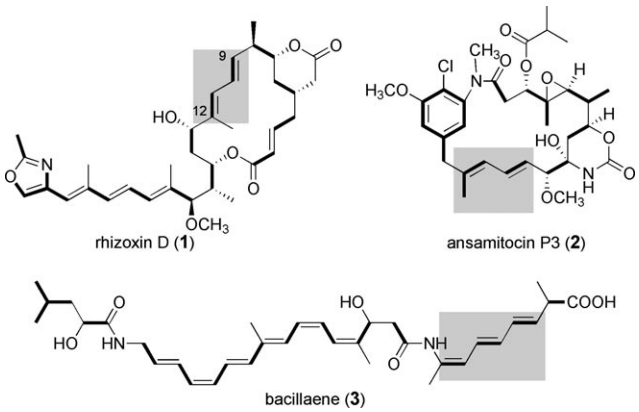
Functionally Distinct Modules Operate Two Consecutive α,β→β,γ Double-Bond Shifts in the Rhizoxin Polyketide Assembly Line**

*Björn Kusebauch, Benjamin Busch, Kirstin Scherlach, Martin Roth, and Christian Hertweck**

Dedicated to Professor Wilhelm Boland on the occasion of his 60th birthday

Complex bacterial polyketides, such as macrolides, polyethers, and polyenes, are a major source of medicinally relevant compounds. Their structural diversity results from a number of programmed events that are controlled by type I polyketide synthases.[1,2] These giant multimodular enzymatic assembly lines catalyze the condensation of malonyl-derived building blocks and the subsequent optional processing of the β keto groups into alcohol, alkenyl, and alkyl moieties. Recent chemical, biochemical, and structural studies have revealed the enzymatic mechanisms underlying the stereochemical course of the ketoreduction, elimination, and enoylreduction reactions.[3–5] As the overall spatial arrangement of the polyketide molecule is shaped profoundly by the unsaturation in the carbon backbone, the dehydratase-catalyzed formation of *E* and *Z* double bonds plays a key role in many pathways.[1,2]

A direct consequence of the *anti* elimination of water from a β-hydroxy-substituted intermediate is that double bonds are generally located in between the incorporated acetate units. However, in various polyketides, double bonds are placed at atypical positions, mainly shifted from α,β- to β,γ-positions. Shifted double-bond systems are found in the structures of various cyclic and acyclic complex polyketides (Scheme 1). The biosynthetic pathways of these compounds give rise to irregular single unsaturations, as in vicenistatin,[6] diene moieties, as found in the antitumor agents rhizoxin D (**1**)[7] and ansamitocin P3 (**2**),[8] and even a triene stretch in bacillaene (**3**).[9] Likewise, α,β→β,γ double-bond shifts play a major role in the formation of polyunsaturated fatty acids by desaturase-independent PUFA synthases (PUFAS).[10] Although the requisite biosynthetic genes have been analyzed successfully, the biochemical basis of the double-bond shifts has remained a mystery. Only recently, feeding experiments with synthetic surrogates provided indirect evidence for the timing of double-bond migration during ansamitocin biosynthesis.[11] Herein we show that the diene moiety of rhizoxin is shifted sequentially by two distinct polyketide synthase (PKS) modules, one of which features a novel type of domain similar to a dehydratase (DH) domain.



Scheme 1. Structures of selected polyketides featuring shifted double bonds.

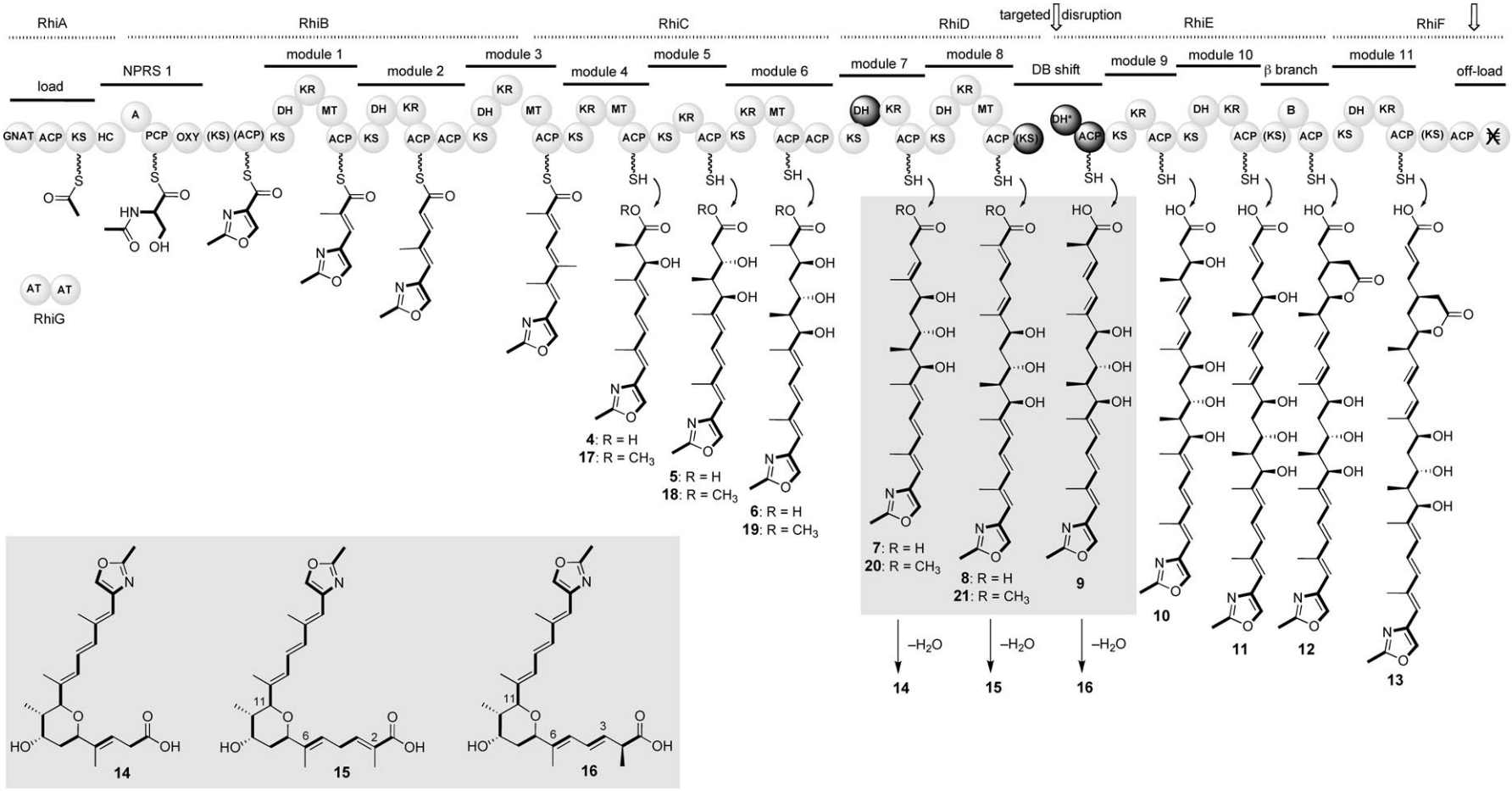
To analyze the molecular basis of rhizoxin biosynthesis we recently cloned and sequenced the entire *rhi* gene cluster[7] from the genome of the bacterial endosymbiont *Burkholderia rhizoxinica*[12] of the rice-seedling-blight fungus *Rhizopus microsporus*.[13–15] A related locus (*rzx*) that codes for rhizoxin biosynthesis was also identified in the genome of the plant commensal *Pseudomonas fluorescens* Pf-5.[16] Analyses of the *rhi* locus revealed a giant thiotemplate system composed of *trans*-acyltransferase (AT) PKS and nonribosomal peptide synthetase (NRPS) modules (Scheme 2).[7,17] Through deletion of the thioesterase (TE) domain[18] we obtained three late-pathway intermediates, **11**–**13**, with a shifted diene moiety.[19] This finding strongly suggests that the unusual isomerization takes place during and not after polyketide-chain assembly. We initially assumed that a conjugated diene moiety generated by modules 7 and 8 would be shifted in a single step. We believed that the split module downstream of module 8 composed of a nonfunctional ketosynthase (KS) domain (KS°) at the C terminus of RhiD as well as a DH-like domain (DH*) and an acyl carrier protein (ACP) at the N terminus of RhiE would be a good candidate for a "double-bond-shift module".

To validate this hypothesis, we examined the metabolic profile of the ΔTE knockout mutant by HPLC/MS and

[*] B. Kusebauch, B. Busch, Dr. K. Scherlach, Dr. M. Roth, Prof. Dr. C. Hertweck
Leibniz-Institut für Naturstoff-Forschung und Infektionsbiologie e.V., Hans-Knöll-Institut
Abteilung Biomolekulare Chemie und Biotechnikum
Beutenbergstrasse 11a, 07745 Jena (Germany)
Fax: (+49) 3641-532-0804
E-mail: christian.hertweck@hki-jena.de
Prof. Dr. C. Hertweck
Friedrich-Schiller-Universität, Jena (Germany)

[**] This project was supported financially by the DFG. We thank A. Perner and F. Rhein for assistance with MS and NMR measurements, respectively. We are also grateful to C. Heiden, K.-D. Menzel, and K. Perlet for support in fermentations.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200905467.



Scheme 2. Model of the multimodular rhizoxin PKS–NRPS biosynthesis assembly line and deduced structures of prematurely released intermediates **4**–**13** and dehydration products **14**–**16**. Abbreviations not defined in the text: GNAT, GCN5-related *N*-acetyltransferase domain; HC, heterocyclization domain; A, adenylation domain; PCP, peptidyl carrier protein domain; OXY, oxygenase domain; KR, ketoreductase domain; MT, methyltransferase domain; B, β-branch domain.

searched for molecular masses corresponding to the proposed intermediates.

In the chromatographic profile we identified a composite peak *m*/*z* 459 corresponding to the isomeric intermediates **8** and **9**. (Figure 1a). These metabolites, however, proved to be inseparable and not sufficiently stable to enable full structural analysis. Intermediates **8** and **9** are prone to the elimination of

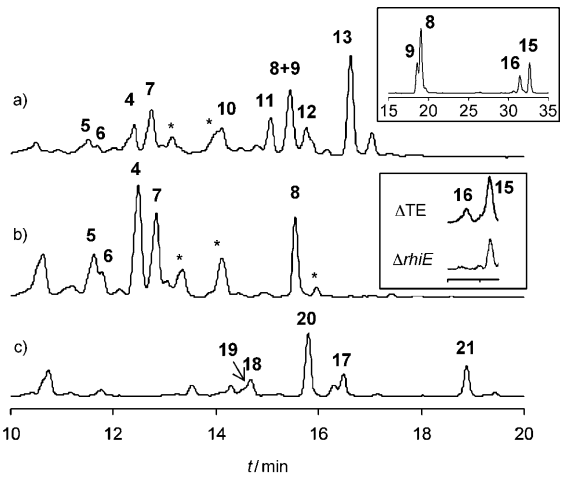


Figure 1. HPLC profiles (UV detection at $\lambda = 310$ nm) of the metabolites produced by *B. rhizoxinica* mutants lacking a) the TE domain, b) DH* (Δ*rhiE* mutant), and c) DH*, after treatment of the extract with $TMSCHN_2$. The inset in (a) shows the analytical HPLC profile of a fraction containing **8** and **9** under optimized conditions for the separation of **15** and **16**. The use of trifluoroacetic acid (TFA) during the isolation process catalyzes the formation of **15** and **16**. The inset in (b) shows the extracted ion chromatogram of the metabolites derived from *B. rhizoxinica* TE and *rhiE* knockout mutants for *m*/*z* 442 $[M+H]^+$. The Δ*rhiE* mutant does not produce **16**. Peaks marked with an asterisk are not related to rhizoxin biosynthesis.

water under only slightly acidic conditions with the formation of two isomers, **15** and **16**. These isomers are more stable and could be separated by preparative HPLC. Upon the addition of 1% TFA/H_2O, the native precursors **8** and **9** were fully converted into **15** and **16**. After careful optimization of the method, we could isolate the two isomers by preparative HPLC from a larger-scale fermentation (200 L) of the ΔTE mutant.

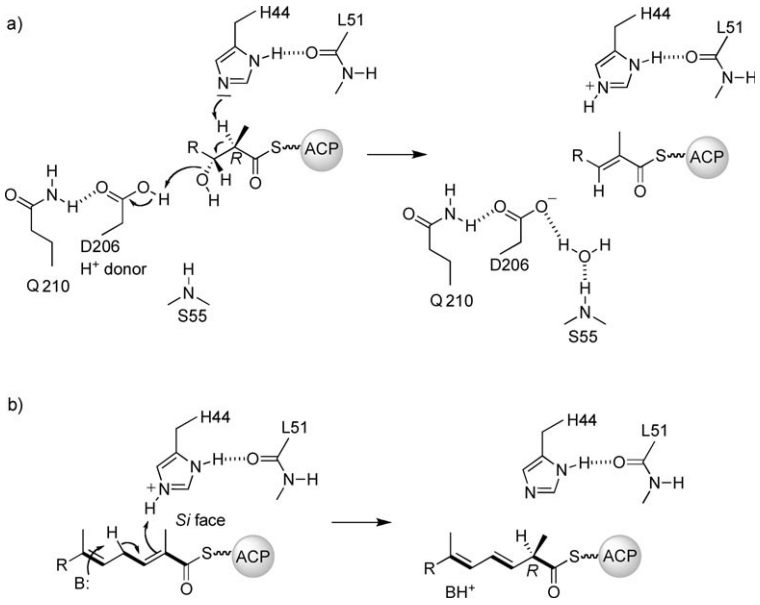
From high-resolution ESIMS data, the molecular formula $C_{26}H_{35}NO_5$ was deduced for both dehydration products. Initially, we expected that the isolated compounds would bear a lactone motif or an additional double bond, as indicated by the loss of water. Instead, HMBC long-range couplings between C7 and H11 and between C11 and H7 revealed the presence of a tetrahydropyran (THP) ring in **15** and **16** (Scheme 2). NMR spectroscopic data also indicates a conjugated diene system between C3 and C6 for **16**. In contrast, **15** has a skipped diene between C2 and C6. As in **16**, the second double bond is already located between C5 and C6. This surprising finding indicates that only the C2–C3 double bond has been shifted downstream of module 8, and that the diene moiety is assembled and shifted sequentially. To determine the exact timing of the first double-bond migration in the rhizoxin pathway, we next isolated the corresponding THP derivative **14** of the native intermediate **7** generated by module 7. Two-dimensional NMR correlations demonstrate that the double bond in **14** has already been shifted from the α,β- to the β,γ-position. The related α,β-unsaturated regioisomer was absent in the crude extract.

The structures of the derivatized intermediates provide insight into the exact timing of the double-bond shifts. Accordingly, the DH domain of module 7 would be involved in the formation of the C11–C12 double bond of rhizoxin D (**1**), whereas the unusual KS°–DH*–ACP module would mediate the second isomerization. To test the role of this noncanonical DH* domain in double-bond migration, we attempted the mutation of the genomic region that codes for the N terminus of RhiE comprising the DH* and ACP domains. Genetic manipulation of the fragile cultured endosymbionts was a challenge; however, we eventually succeeded in generating an insertional disruption (see the Supporting Information). Metabolic profiling indicated that the resulting mutant was incapable of producing the late-pathway intermediates **10**–**13** (Figure 1b), which is in full accord with the colinear arrangement of the thiotemplate system. More importantly, this mutant exclusively produced isomer **8**, in which the second double-bond shift has not taken place. By transforming the intermediates into the corresponding THP derivatives and analyzing the mixture by HPLC/MS, we could identify **15** with the skipped double-bond system, whereas **16** was not detected (inset in Figure 1b). To rule out the possibility that the shifted double bond is an artifact of THP formation and hence not catalyzed by DH7, we confirmed our results by a second analytical method: treatment of the crude extract of a 50 L fermentation of the *rhiE* knockout mutant with trimethylsilyldiazomethane ($TMSCHN_2$) resulted in the formation of methyl esters **17**–**21** of the biosynthetic intermediates **4**–**8** (Figure 1c). Notably, no cyclization was observed, since the workup did not require the addition of an acid. One- and two-dimensional NMR experiments revealed that the double bond is already shifted in **20**, as in **14**, and that **21** contains a skipped-diene system like **15**.

In conjunction with the structures of the identified shunt products, the mutational analysis clearly showed the *rhi* PKS modules involved in double-bond migrations. Interestingly, two individual modules mediate the double-bond shifts by two different operations. Module 7 introduces the unsaturation with a concomitant double-bond shift, which corresponds to a formal β,γ-dehydration. This scenario has been proposed for the ansamycin pathway on the basis of indirect evidence[11] and is analogous to that found for the bacillaene pathway.[20] However, in rhizoxin biosynthesis, the formation and shifting of the second double bond is separated in time. In this case, module 8 first catalyzes chain elongation and β-keto processing to the stage of the substituted acryloyl thioester; then, the double bond is shifted by the downstream "shift module" harboring the DH* domain.

A closer bioinformatic examination of the DH7 and DH* domains from the *rhi* PKS indicated a good overall similarity to DH domains from other modular type I PKS. However, according to a phylogenetic analysis (see the Supporting Information), DH domains fall into particular clades. There is

a clade for fatty acid synthases (FAS), another for PUFAS, and a major branch formed by type I PKS DH domains from a variety of bacteria. Interestingly, within these clades, the RhiE DH* domain forms a clearly distinct clade with BaeR DH* and DifK DH* from the bacillaene and difficidin pathways, respectively. Multiple sequence alignment showed that the sequences of DH* domains deviate from those of classical DH domains, as conserved motifs are either missing or mutated: Hx_3Gx_4P is mutated to Hx_9, and the GYxYGPxF and DxxxQ/H motifs are fully absent (Scheme 3). Whereas



Scheme 3. Comparison of the models a) for dehydration, as deduced from the DEBS DH domain,[21] and b) for a double-bond shift mediated by the DH* domain in the *rhi* PKS.

the H44 residue is conserved in DH and DH* domains, DH* domains feature a noticeable D206-to-N206 mutation (nomenclature for the DH domain of 6-deoxyerythronolide B synthase (DEBS)) in the DxxxQ/H motif, which serves as a proton donor in standard dehydration reactions.[21] Notably, an analogous D-to-N exchange in *E. coli* ubiquinol oxidase eliminates proton-pumping activity.[22] Given the absence of the catalytic DH motifs, it is clear that the DH* domain cannot function as a classical dehydratase, in which H44 acts as a catalytic base to deprotonate at the α position, and D206 promotes the elimination.[21] Instead, we postulate that H44 serves as a proton donor, and a yet unknown basic residue assists in the double-bond migration. Threading of the RhiE DH* sequence with the recently solved DEBS DH crystal structure[21] showed an overall similar double-hotdog fold and indicated that H44 is located at the same position. In the biosynthetic model, double-bond migration would involve an enantioselective proton transfer to the *Si* face of the sp^2-hybridized carbon atom (C2) of the α,β-unsaturated thioester, or to the corresponding enol tautomer. The course of the reaction is in fact in full agreement with the observed *R* absolute configuration at C8 in the final macrolide ring. Future in vitro experiments will shed more light on the exact enzymatic mechanism of the double-bond shifts.

Double-bond migrations are a hallmark of desaturase-independent PUFA pathways, and they are also observed in a variety of complex polyketide pathways. Even so, direct evidence for the timing and the underlying biochemical operators is lacking. In this study we investigated the biogenesis of the shifted diene moiety of the antimitotic macrolide rhizoxin D (**1**) from the endofungal bacterium *B. rhizoxinica*. The isolation and full structure elucidation of complex biosynthetic intermediates that are released in minute amounts from a mutant blocked in the off-loading mechanism revealed the timing of desaturation and double-bond migration. The analysis of the two isomeric intermediates **8** and **9**, which could only be separated through THP formation (to give **15** and **16**), led to the surprising discovery that the conjugated double bonds of **1** are shifted sequentially. According to the modular architecture of the *rhi* PKS, the first migration takes place concomitant with desaturation in module 7. In stark contrast, the second double bond is first generated by module 8 and then shifted by a novel type of "shift module" (KS°–DH*–ACP). The function of the unusual DH* domain was corroborated by gene disruption and detailed analysis of the metabolites produced by the mutant. Bioinformatic analyses did not show any obvious characteristic of the DH7 domain. However, the DH* domain lacks critical motifs for dehydration and seems to be designed for α,β→β,γ double-bond shifts. The predicted course of DH*-mediated proton migration is supported by the absolute configuration at C8 in the macrolide. The insight into the molecular basis of double-bond migration in polyketide biosynthesis gained in this study was fully unexpected. It may aid in the analysis of related fatty acid and polyketide synthases and provide a basis for the rational engineering of noncanonical polyene systems.

Received: September 29, 2009
Published online: December 23, 2009

Keywords: biosynthesis · dehydration · fatty acids · macrolides · polyketides

[1] C. Hertweck, *Angew. Chem.* **2009**, *121*, 4782–4811; *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716.
[2] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468–3496.
[3] J. Wu, T. J. Zaleski, C. Valenzano, C. Khosla, D. E. Cane, *J. Am. Chem. Soc.* **2005**, *127*, 17393–17404.
[4] M. M. Alhamadsheh, N. Palaniappan, S. Daschouduri, K. A. Reynolds, *J. Am. Chem. Soc.* **2007**, *129*, 1910–1911.
[5] D. H. Kwan, Y. Sun, F. Schulz, H. Hong, B. Popovic, J. C. C. Sim-Stark, S. F. Haydock, P. F. Leadlay, *Chem. Biol.* **2008**, *15*, 1231–1240.
[6] Y. Ogasawara, K. Katayama, A. Minami, M. Otsuka, T. Eguchi, K. Kakinuma, *Chem. Biol.* **2004**, *11*, 79–86.

[7] L. P. Partida-Martinez, C. Hertweck, *ChemBioChem* **2007**, *8*, 41 – 45.
[8] T. W. Yu, L. Bai, D. Clade, D. Hoffmann, S. Toelzer, K. Q. Trinh, J. Xu, S. J. Moss, E. Leistner, H. G. Floss, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7968 – 7973.
[9] R. A. Butcher, F. C. Schroeder, M. A. Fischbach, P. D. Straight, R. Kolter, C. T. Walsh, J. Clardy, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1506 – 1509.
[10] U. Kaulmann, C. Hertweck, *Angew. Chem.* **2002**, *114*, 1947 – 1950; *Angew. Chem. Int. Ed.* **2002**, *41*, 1866 – 1869.
[11] F. Taft, M. Brünjes, T. Knobloch, H. G. Floss, A. Kirschning, *J. Am. Chem. Soc.* **2009**, *131*, 3812 – 3813.
[12] L. P. Partida-Martinez, I. Groth, M. Roth, I. Schmitt, K. Buder, C. Hertweck, *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 2583 – 2590.
[13] L. P. Partida-Martinez, C. Hertweck, *Nature* **2005**, *437*, 884 – 888.
[14] K. Scherlach, L. P. Partida-Martinez, H.-M. Dahse, C. Hertweck, *J. Am. Chem. Soc.* **2006**, *128*, 11529 – 11536.
[15] L. P. Partida-Martinez, S. Monajembashi, K. O. Greulich, C. Hertweck, *Curr. Biol.* **2007**, *17*, 773 – 777.
[16] N. Brendel, L. P. Partida-Martinez, K. Scherlach, C. Hertweck, *Org. Biomol. Chem.* **2007**, *5*, 2211 – 2213.
[17] T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck, J. Piel, *Nat. Biotechnol.* **2008**, *26*, 225 – 233.
[18] J. Moldenhauer, X.-H. Chen, R. Borriss, J. Piel, *Angew. Chem.* **2007**, *119*, 8343-8354; *Angew. Chem. Int. Ed.* **2007**, *46*, 8195 – 8197.
[19] B. Kusebauch, B. Busch, K. Scherlach, M. Roth, C. Hertweck, *Angew. Chem.* **2009**, *121*, 5101 – 5104; *Angew. Chem. Int. Ed.* **2009**, *48*, 5001 – 5004.
[20] See the following Communication: J. Moldenhauer, D. C. G. Götz, C. R. Albert, S. K. Bischof, K. Schneider, R. D. Süßmuth, M. Engeser, H. Gross, G. Bringmann, J. Piel, *Angew. Chem.* **2009**, DOI: 10.1002/ange.200905468; *Angew. Chem. Int. Ed.* **2009**, DOI: 10.1002/anie.200905468.
[21] A. Keatinge-Clay, *J. Mol. Biol.* **2008**, *384*, 941 – 953.
[22] J. W. Thomas, A. Puustinen, J. O. Alben, R. B. Gennis, M. Wikström, *Biochemistry* **1993**, *32*, 10923 – 10928.

