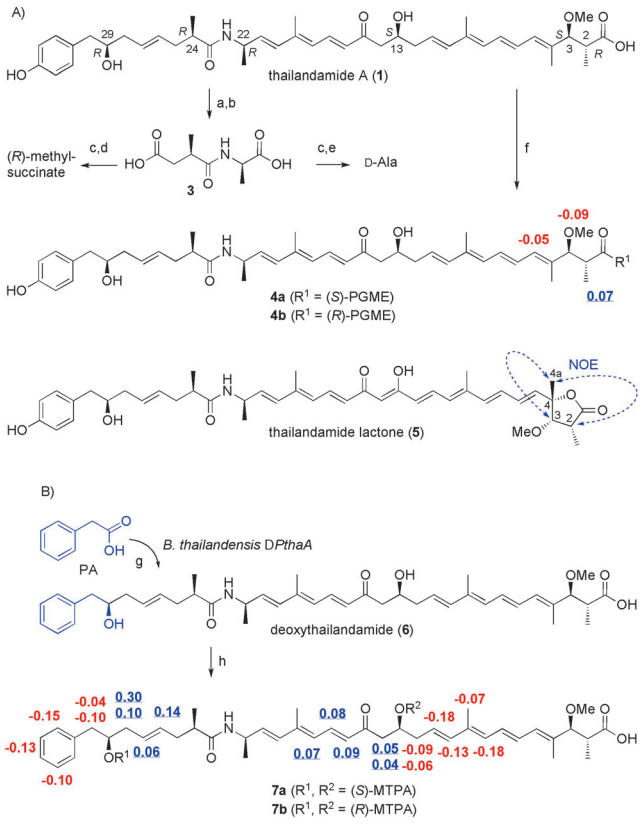
Polyketides

DOI: 10.1002/anie.201200067

Assembly and Absolute Configuration of Short-Lived Polyketides from *Burkholderia thailandensis***

*Keishi Ishida, Thorger Lincke, and Christian Hertweck**

Complex polyketides from bacteria comprise a broad range of macrolides, polyenes, and polyethers, many of which play eminent roles for the development of therapeutics.[1] Typically, these compounds feature multiple chiral centers that render them challenging benchmarks for organic synthesis. In light of this, it is impressive to learn how swiftly complex polyketides are assembled in microorganisms, utilizing only simple building blocks such as activated acyl and malonyl units.[2,3] The types, number, and processing of these basic blocks are programmed by the architecture of modular polyketide synthases (PKS) which are encoded in the microbial genomes. One most remarkable discovery was that the order of catalytic domains involved in chain elongation and side-chain processing is typically colinear with the final product.[4] This principle has not only opened the door to rationally engineering polyketides, but also enabled forecasting PKS architectures from polyketide structures and vice versa.[5] However, through recent full genome sequencing projects it has become more and more obvious that most of the genes coding for biosynthetic assembly lines remain silent under standard laboratory conditions.[6] Furthermore, PKS expression and polyketide formation may only occur transiently, thus suggesting that the metabolites function as temporary signals.[7] This remarkable case has been observed for *Burkholderia thailandensis*, a bacterium that serves as a model for the pathogenic relatives *B. mallei* and *B. pseudomallei*. Mining the genome of *B. thailandensis* revealed a giant *trans*-acyltransferase (*trans*-AT) PKS[8] gene cluster, and a bioinformatic structure prediction aided in the discovery of a hitherto unknown polyketide, thailandamide A (**1**, Scheme 1 A).[7] However, the absolute configuration of the labile polyene **1** and the structure of a short-lived congener (thailandamide B; **2**; see Figure 2 for structure), have remained elusive. Yet this very structural information would serve as a stepping-stone to clarify the biological function of this enigmatic, quorum-sensing-regulated metabolite. Herein we report a multidisciplinary approach involving bioinformatics, chemical degradation, and precursor-directed biosynthesis to rigorously determine the absolute configuration of thailandamide A. On the basis of the stereochemical assignments, in silico predictions, and mutational pathway dissection, we also unveil the structure of the immediate, thermolabile PKS product thailandamide B.



Scheme 1. Structure of thailandamide A (**1**) and experiments revealing the absolute configuration of **1**. A) Degradation and derivatization. a) O_3, −78 °C; b) 35% H_2O_2, AcOH, 60 °C; c) 6 M HCl, 105 °C; d) *S*-PGME, PyBOP; e) L-FDAA; f) (*S*,*R*)-PGME, PyBOP. B) Synthesis of deoxythailandamide. g) *B. thailandensis* Δ*PthaA*, precursor-directed biosynthesis; h) (*S*,*R*)-MTPA chloride, *pyridine*. B. $\Delta\delta^{SR}$ values in ppm.

Thailandamide features six chiral centers, at C2, C3, C13, C22, C24, and C29. We initiated the stereochemical analysis by chemical degradation and derivatization. For the isolation of preparative amounts of thailandamide A (**1**) we employed a genetically engineered strain, the quorum sensing mutant *B. thailandensis* Δ*PthaA*, which produces thailandamide constitutively.[9] From an upscaled fermentation (10 L) of the mutant we succeeded in isolating sufficient amounts of **1** (36.9 mg) for several degradation and derivatization experi-

[*] Dr. K. Ishida, T. Lincke, Prof. Dr. C. Hertweck
Leibniz Institute for Natural Product Research and Infection Biology, HKI, Dept. of Biomolecular Chemistry, and Bio Pilot Plant
Beutenbergstr. 11a, 07745 Jena (Germany)
E-mail: christian.hertweck@hki-jena.de
Prof. Dr. C. Hertweck
Chair for Natural Product Chemistry, Friedrich Schiller University Jena (Germany)

[**] We thank A. Perner, H. Heinecke, and F. Rhein for MS and NMR measurements. This research was financially supported by the Pakt für Wissenschaft und Innovation.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201200067.

ments. First, to elucidate the absolute configuration at positions C22 and C24, the polyene was subjected to ozonolysis (Scheme 1 A). The resulting oxidative degradation product (**3**) was hydrolyzed using 6 M HCl and derivatized using Marfey's reagent for Ala, and (*S*)-PGME (phenylglycine methylester) for (*R*)-methylsuccinate. Since the RP-HPLC analysis of the L-FDAA (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide) derivative of the hydrolysate revealed the identity of D-Ala, we could unequivocally assign an *R* configuration to position C22. The di-(*S*)-PGME amide of hydrolyzed **3** showed the same retention time as the (*R*)-methylsuccinate di-(*S*)-PGME amide on the HPLC column, thus indicating the *R* configuration at C24.

To elucidate the absolute configuration at C2, again the formation of a PGME amide provided valuable insights. From the $\Delta\delta^{SR}$ values of the thailandamide A-derived (*S*)- and (*R*)-PGME amides **4a** and **4b**, respectively, we deduced the *R* configuration of position C2. The configuration of C3, however, could not be determined in this way. Therefore, we inspected the NOE correlations of the γ-butyrolactone ring in the thailandamide lactone **5**, a congener and derivative of **1** produced by *B. thailandensis* Δ*PthaA*. From these correlations we were able to unequivocally deduce the *S* configuration at C3. The determination of the absolute configurations at positions C13 and C29 proved to be most challenging, since no reference compounds were available for fragments obtained by ozonolysis. Furthermore, treatment with Mosher's reagent (MTPA; α-methoxy-α-(trifluoromethyl)phenylacetic acid) chloride gave mixed products, including the mono-MTPA ester at the phenolic position of **1**. Unexpectedly, attempts to selectively protect the phenolic group yielded several by-products because of the temperature sensitivity of **1**. To circumvent these limitations we aimed at generating a thailandamide variant lacking the phenolic OH group. Since a synthetic method for achieving this was out of reach, we attempted producing this derivative by precursor-directed biosynthesis. Analysis of the polyketide backbone suggested that *p*-hydroxphenylacetate (PHPA) served as a starter unit which is loaded onto the first PKS module. The deoxy variant of **1** would in principle result from the incorporation of phenylacetate (PA). After optimization of production parameters, we eventually achieved this goal by supplementing a *B. thailandensis* Δ*PthaA* culture with the nonnatural starter unit in minimal medium (Scheme 1 B). HPLC-HRMS monitoring of the fermentation indicated that the mutant indeed incorporated phenylacetate (Figure 1, trace c).

An isolation/purification workflow as established for **1** yielded sufficient amounts of deoxythailandamide A (**6**; (Scheme 1 B), which could be readily esterified by MTPA chloride without the formation of any side products. Finally, the $\Delta\delta^{SR}$ values of the deoxythailandamide A-derived di-(*S*)- and di-(*R*)-MTPA esters **7a** and **7b**, respectively, revealed an *S* configuration for C13 and an *R* configuration for C29. To complement the stereochemical analysis, we analyzed the enzymatic domains encoded in the *tha* gene locus (Scheme 2). According to NRPSPredictor2,[10] the adenylation (A) domain of ThaH has the amino acid code DMPQLGMVWK, which indicates a preference for small amino acids such as

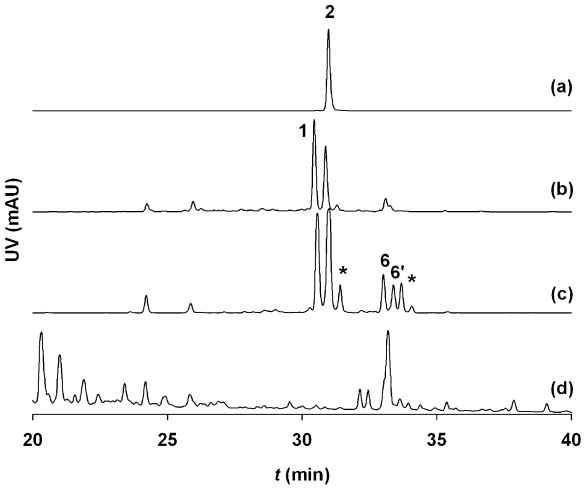
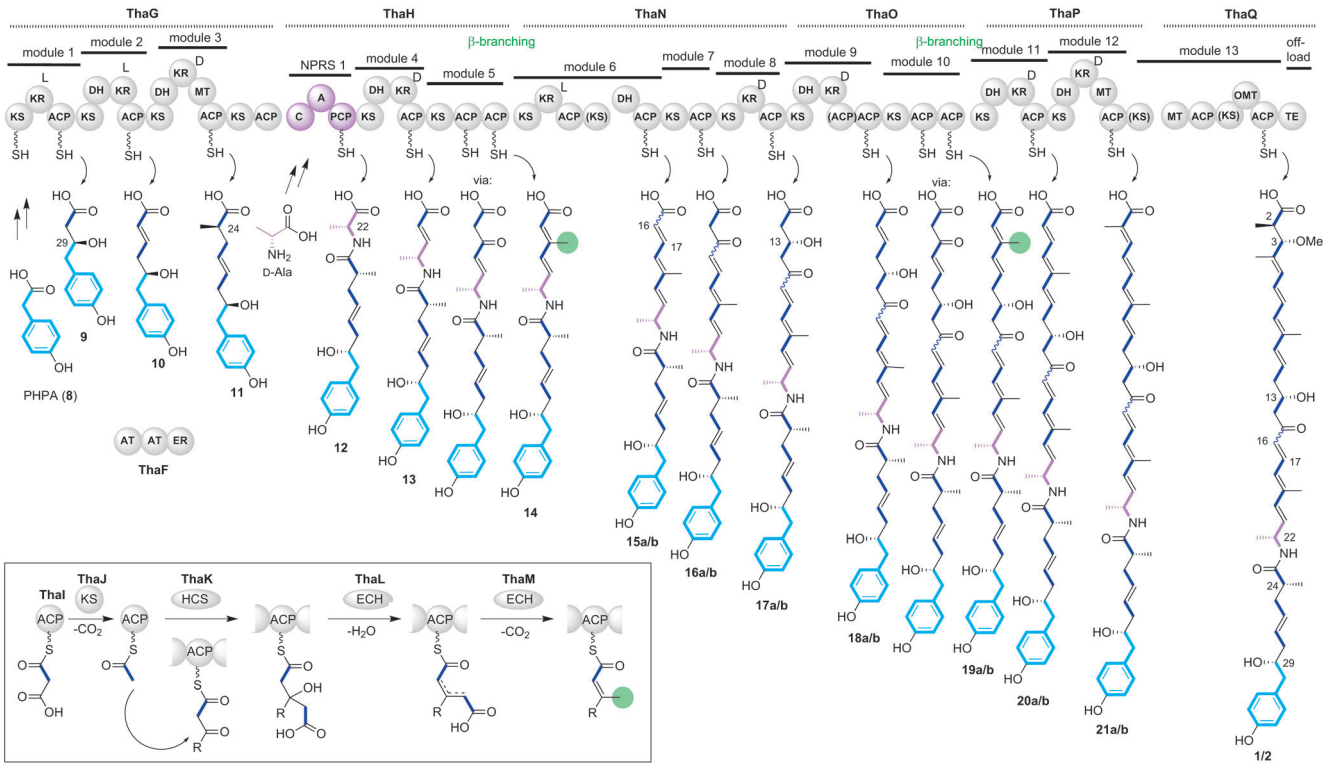


Figure 1. HPLC profiles of a sample of purified thailandamide B [**2**; (a)], extracts from cultures of *B. thailandensis* Δ*PthaA* (b), extracts from cultures of *B. thailandensis* Δ*PthaA* supplemented with phenylacetate in minimal medium (c), and d) extracts from cultures of *B. thailandensis* Δ*PthA*/Δ*TE* double mutant. UV absorbance at $\lambda = 317$ nm (a–c) and $\lambda = 220$ nm (d). **6′**: Putative deoxythailandamide B; *: unidentified unstable (deoxy)thailandamides.

Gly, Ala, and Val. Since the chemical analyses showed that **1** is derived from D-Ala, it is surprising that the condensation (C) domain does not have a dual condensation/epimerization function.[11] Consequently, this seems to be a rare scenario where an A domain incorporates a D-amino acid, as for example in the cyclosporin synthetase.[12] Next, we predicted all ketoreductase (KR) domain specificities[13–15] in the thailandamide PKS. For the hydroxy-substituted positions, the KR finger printing fully matched with the detected configurations at C29(L/*R*) and C13 (D/*S*; see Table S7 in the Supporting Information). Interestingly, the *S* configuration of the methoxy-substituted carbon atom (C3) could not be predicted since the product is not colinear with the architecture of module 13, which lacks a KR domain. The domain organization could be suggestive for a route involving enolization and methylation, followed by hydrogenation.[8] A likely candidate is the *trans*-enoylreductase (*trans*-ER) domain in ThaF, which controls the absolute configuration at C24. While stereocontrol of enoylreductases has been unveiled for *cis*-AT PKS systems,[16–18] the predictions are not adaptable for the *trans*-ER in *trans*-AT PKS systems, which are more similar to PfaD in polyunsaturated fatty acid (PUFA) biosynthesis.[19] However, one may assume that the stereochemical course of enoyl reduction would be comparable for C24 and C2/3, but no clear mechanistic analogies could be drawn. Thus, a more plausible alternative would be the iterative use of a KR domain upstream of module 13. Indeed all vicinal KR domains produce D-hydroxy groups, which is in accord with the observed *S* configuration at C3. Furthermore, by using HPLC-HRMS we were able to detect the exact mass for predicted hydroxy-substituted intermediate in a Δ*TE*-mutant (see below, and the Supporting Information).

As to the double-bond architectures, previous work suggested that the dehydratase (DH) domain mediates an



Scheme 2. Model for thailandamide assembly on the modular *tha trans*-AT PKS, based on domain analysis and metabolic profiling of a Δ*TE* mutant. KS: ketosynthase, AT: acyl transferase, ACP: acyl carrier protein, MT: methyltransferase, KR: ketoreductase, DH: dehydratase, ER: enoyl reductase, TE: thioesterase. Insert: mechanism of β branching (green circles highlight methyl groups derived from β-branching). ECH: enoyl-CoA hydratase/crotonase, HCS: HMG-CoA synthase. Structure numbers are intermediates detected by HR-LCMS.

anti elimination of water in D-3-hydroxyacyls to yield *trans* double bonds, whereas *cis* double bonds result from the corresponding L-isomers.[2,20] However, this prediction did not match with the observed *E* configuration at C16/17. This inconsistency implied that thailandamide A (**1**) might not represent the true primary product of the *tha* PKS, and we assumed that the yet enigmatic, unstable congener of **1**, thailandamide B (**2**) would be the product released from the megasynthase. However, we repeatedly experienced difficulties in isolating the isomer as it readily degraded and isomerized in the course of purification, even in the absence of light. The resulting mixtures gave complex NMR data which evaded interpretation because of overlapping signals of various stereoisomers. Finally, we established a low-temperature isolation and purification protocol, thus carefully maintaining temperatures below room temperature to obtain pure **2**. NMR measurements were also conducted below 10 °C to provide high quality ^{1}H-^{1}H COSY, HSQC, and HMBC NMR data (see the Supporting Information) which allowed the full structural elucidation of **2**.

Interestingly, the ^{1}H and ^{13}C NMR spectra of **2** were almost identical with those of **1** except for positions C16 to C18 (position 16, **1**; $\delta_H = 6.24$ ppm, $\delta_C = 130.8$ ppm, **2**; $\delta_H = 6.11$ ppm, $\delta_C = 125.0$ ppm, position 17, **1**; $\delta_H = 7.62$ ppm, $\delta_C = 140.5$ ppm, **2**; $\delta_H = 6.93$ ppm, $\delta_C = 139.3$ ppm, position 18, **1**; $\delta_H = 6.19$ ppm, $\delta_C = 129.9$ ppm, **2**; $\delta_H = 7.42$ ppm, $\delta_C = 128.6$ ppm). This deviation suggested that the double bond at C16/17 was *Z* (*cis*) in **2** compared to that of *E* (*trans*) in **1**. The diagnostic coupling constant 11.2 Hz between H16 and H17 provided further support for the *Z* configuration at C16–C17. According to coupling constants and ROESY data, all other geometrical constitutions were identical with those in **1** (Figure 2 and Figure S11 in the Supporting Information). The architecture of **2** is in full agreement with the deduced biosynthetic code, thus implying that **2** would be the immediate product released from the modular assembly line. The *cis* double bond could be prone to rapid isomerization to give **1**. Alternatively, both *cis* and *trans* isomers could form concomitantly because of some degree of flexibility of the DH domain. To address this, we first tested whether **2** would readily isomerize in vitro to yield **1**.

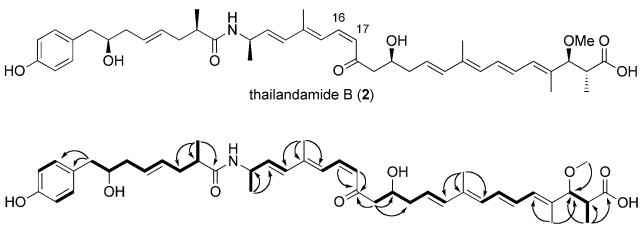


Figure 2. Structure of thailandamide B (**2**) and selected correlations observed by 2D NMR spectroscopy. Bold line: ^{1}H-^{1}H COSY correlations; arrows: HMBC correlations.

However, only small amounts of **1** are formed by incubating **2** at room temperature with or without light, with formation of many yet uncharacterized isomerization or degradation products, a scenario that does not reflect co-occurrence of **1** and **2** found in the culture broth. Thus, to pinpoint the timing of *cis*/*trans* double-bond formation, we next aimed at dissecting the biosynthetic pathway in vivo. For this purpose, we deleted the gene region coding for the off-loading thioesterase (TE) domain at the C-terminus of ThaQ. This genetic manipulation was believed to result in a Δ*TE*-mutant accumulating and releasing premature polyketide intermediates as in the bacillaene[21,22] and rhizoxin[23,24] pathways. Indeed, by HPLC-MS and HRMS (Exactive) measurements we were able to detect the masses for the entire range of predicted structures. Through SIM (selected ion monitoring) we obtained snapshots of the entire pathway (Scheme 2). In this way, it was not only possible to deduce the course chain propagation and branching events (Scheme 2, inset), but also to trace the onset of *Z*/*E* isomerization. We found that up to module 5 only single isomers are formed, whereas products generated by module 6 and all additional downstream modules appeared as two main isomers (e.g. **15a**,**b**, **16a**,**b**; Scheme 2). It is highly unlikely that an as yet unidentified isomerase would be promiscuous enough to act equally efficiently on all of these different substrates. Thus, the results taken together, we reason that the DH domain of split module 6 unexpectedly yields *cis* and *trans* double bonds, either through catalyzing both *anti* and *syn* eliminations or through isomerization of the double bond of the tethered substrate.

Elucidating the absolute configuration of complex polyketides with multiple chirality centers is often an ambitious undertaking which may require multidisciplinary approaches.[26–29] In this work we describe a particularly challenging case of a stereochemical analysis to unravel the structure of a short-lived polyene, thailandamide A (**1**). We succeeded in the unequivocal determination of all six stereocenters of this fragile molecule (2*R*, 3*S*, 13*S*, 22*R*, 24*R*, 29*R*), which was only feasible by a combination of chemical degradation, derivatization, 1D/2D NMR analyses, and precursor-directed biosynthesis to yield the deoxy variant (**6**) of **1**. The stereochemical elucidation in conjunction with complementary bioinformatic analyses implied the unexpected and quite unusual direct incorporation of D-Ala into the thailandamide backbone. Furthermore, PKS domain analyses (KR fingerprinting) guided us to the presumed immediate product of the *tha* PKS. Using a low temperature isolation protocol we indeed succeeded in isolating and fully characterizing the highly unstable thailandamide B (**2**). However, monitoring pathway dissection by HRMS and using an engineered Δ*TE* mutant eventually revealed that *E*/*Z* isomers are generated online by module 6, which seems to lack rigid stereocontrol. This finding is particularly noteworthy since the predicted stereochemical course of β-keto processing in many other cases could not be correlated with the ultimate double bond geometry in *trans*-AT PKS products. Finally, the stereochemical analysis also provides strong support for a model involving an iterative use of a KR domain. In sum, these unexpected results not only point out intriguing stereochemical aspects of polyketide biosynthesis, but also endorse the power of synergizing chemical and biological methods in natural product research.

Experimental Section

See the Supporting Information.

Received: January 4, 2012
Revised: March 5, 2012
Published online: ■■ ■■, ■■■■

Keywords: bioinformatics · biosynthesis · polyketides · stereochemistry · structure elucidation

[1] D. O'Hagan, *The Polyketide Metabolites*, Ellis Horwood, Chichester, **1991**.
[2] C. Hertweck, *Angew. Chem.* **2009**, *121*, 4782 – 4811; *Angew. Chem. Int. Ed.* **2009**, *48*, 4688 – 4716.
[3] C. Khosla, *J. Org. Chem.* **2009**, *74*, 6416 – 6420.
[4] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380 – 416.
[5] K. J. Weissman, P. F. Leadlay, *Nat. Rev. Microbiol.* **2005**, *3*, 925 – 936.
[6] J. M. Winter, S. Behnken, C. Hertweck, *Curr. Opin. Chem. Biol.* **2011**, *15*, 22 – 31.
[7] 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.
[8] J. Piel, *Nat. Prod. Rep.* **2010**, *27*, 996 – 1047.
[9] K. Ishida, T. Lincke, S. Behnken, C. Hertweck, *J. Am. Chem. Soc.* **2010**, *132*, 13966 – 13968.
[10] M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, *Nucleic Acids Res.* **2011**, *39*, W362 – W367.
[11] C. J. Balibar, F. H. Vaillancourt, C. T. Walsh, *Chem. Biol.* **2005**, *12*, 1189 – 1200.
[12] K. Hoffmann, E. Schneider-Scherzer, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **1994**, *269*, 12710 – 12714.
[13] P. Caffrey, *ChemBioChem* **2003**, *4*, 649 – 662.
[14] A. Baerga-Ortiz, B. Popovic, A. P. Siskos, H. M. O'Hare, D. Spiteller, M. G. Williams, N. Campillo, J. B. Spencer, P. F. Leadlay, *Chem. Biol.* **2006**, *13*, 277 – 285.
[15] R. Castonguay, W. He, A. Y. Chen, C. Khosla, D. E. Cane, *J. Am. Chem. Soc.* **2007**, *129*, 13758 – 13769.
[16] 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.
[17] D. H. Kwan, P. F. Leadlay, *ACS Chem. Biol.* **2010**, *5* 829 – 838.
[18] D. H. Kwan, F. Schulz, *Molecules* **2011**, *16*, 6092 – 6115.
[19] S. B. Bumpus, N. A. Magarvey, N. L. Kelleher, C. T. Walsh, C. T. Calderone, *J. Am. Chem. Soc.* **2008**, *130*, 11614 – 11616.
[20] J. Wu, T. J. Zaleski, C. Valenzano, C. Khosla, D. E. Cane, *J. Am. Chem. Soc.* **2005**, *127*, 17393 – 17404.
[21] J. Moldenhauer, X.-H. Chen, R. Borriss, J. Piel, *Angew. Chem.* **2007**, *119*, 8343 – 8345; *Angew. Chem. Int. Ed.* **2007**, *46*, 8195 – 8197.
[22] J. Moldenhauer, D. C. G. Götz, C. R. Albert, S. K. Bischof, K. Schneider, R. Süßmuth, M. Engeser, H. Groß, G. Bringmann, J. Piel, *Angew. Chem.* **2010**, *122*, 1507 – 1509; *Angew. Chem. Int. Ed.* **2010**, *49*, 1465 – 1467.
[23] B. Kusebauch, B. Busch, K. Scherlach, M. Roth, C. Hertweck, *Angew. Chem.* **2009**, *121*, 5101 – 5104; *Angew. Chem. Int. Ed.* **2009**, *48*, 5001 – 5004.
[24] B. Kusebauch, B. Busch, K. Scherlach, M. Roth, C. Hertweck, *Angew. Chem.* **2010**, *122*, 1502 – 1506; *Angew. Chem. Int. Ed.* **2010**, *49*, 1460 – 1464.

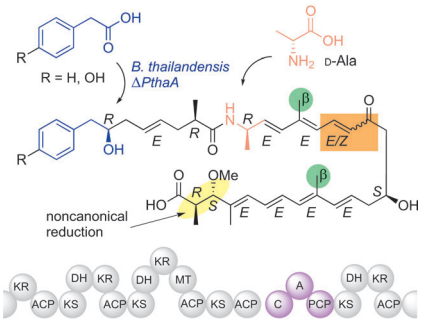
[25] C. T. Calderone, *Nat. Prod. Rep.* **2008**, *25*, 845 – 853.
[26] D. Menche, *Nat. Prod. Rep.* **2008**, *25*, 905 – 918.
[27] M. Bock, K. Buntin, R. Müller, A. Kirschning, *Angew. Chem.* **2008**, *120*, 2341 – 2344; *Angew. Chem. Int. Ed.* **2008**, *47*, 2308 – 2311.
[28] D. Menche, F. Arikan, O. Perlova, N. Horstmann, W. Ahlbrecht, S. C. Wenzel, R. Jansen, H. Irschik, R. Müller, *J. Am. Chem. Soc.* **2008**, *130*, 14234 – 14243.
[29] D. Janssen, D. Albert, R. Jansen, R. Müller, M. Kalesse, *Angew. Chem. Int. Ed.* **2007**, *46*, 4898 – 48901.

Communications

Polyketides

K. Ishida, T. Lincke,
C. Hertweck* ——— ■■■■–■■■■

Assembly and Absolute Configuration of Short-Lived Polyketides from *Burkholderia thailandensis*



Decoded before decay: Cryptic and highly unstable polyketide metabolites, thailandamides A and B, were isolated from *Burkholderia thailandensis*, and their absolute configurations fully elucidated using a combination of chemical degradation, bioinformatics, NMR spectroscopy, precursor-directed biosynthesis, and analysis of pathway intermediates (see scheme).