Discovery of the Tancilactone Antibiotics by Genome Mining of Atypical Bacterial Type II Diterpene Synthases

Dr. Liao-Bin Dong^{#a}, Dr. Jeffrey D. Rudolf^{#a}, Dr. Ming-Rong Deng^a, Dr. Xiaohui Yan^a, and Prof. Dr. Ben Shen^{a,b}

^[a]Department of Chemistry, The Scripps Research Institute Jupiter, FL 33458 (USA), shenb@scripps.edu.

^[b]Department of Molecular Medicine, Natural Products Library Initiative at The Scripps Research Institute, The Scripps Research Institute, Jupiter, FL 33458 (USA).

[#] These authors contributed equally to this work.

Abstract

Although genome mining has advanced the identification, discovery, and study of microbial natural products, the discovery of bacterial diterpenoids continues to lag behind. Herein, we report the identification of 66 putative producers of novel bacterial diterpenoids, and the discovery of the tancilactone (TNL) family of antibiotics, by genome mining of type II diterpene synthases that do not possess the canonical DXDD motif. The TNLs, which are broad-spectrum antibiotics with moderate activities, are produced by both *Streptomyces* sp. CB03234 and *Streptomyces* sp. CB03238 and feature a highly functionalized diterpenoid skeleton that is further decorated with chloroan-thranilate and γ -butyrolactone moieties. Genetic manipulation of the *tnl* gene cluster resulted in TNL congeners, which provided insights into their biosynthesis and structure–activity relationships. This work highlights the biosynthetic potential that bacteria possess to produce diterpenoids and should inspire continued efforts to discover terpenoid natural products from bacteria.

Keywords

antibiotics; biosynthesis; terpenoids; genomics; natural products

Terpenoids are the largest, most structurally diverse family of natural products with over 70000 known compounds (<http://dnpc.chemnetbase.com>). Terpenoids of bacterial origin, however, are relatively rare, with diterpenoids epitomizing this disparity.^[1] Although recent genomics, bioinformatics, and experimental studies support that terpene synthases (TSs) are widely distributed in bacteria,^[2] only a handful of highly functionalized bacterial diterpenoids are known. The structures and gene clusters of platensimycin (PTM) and platencin (PTN),^[3] the gibberellins,^[4] phenalinolactone (PLA),^[5] and brasiliocardin (BRA)^[6]

Conflict of Interest

The authors declare no conflict of interest.

exemplify the biosynthetic potential of bacteria to produce unique, biologically active, diterpenoid natural products (Scheme 1 and Figure S1 in the Supporting Information).

The genomic era has seen a rejuvenation in the discovery of natural products. Genome mining for novel natural products is prevalent among polyketides,^[7] nonribosomal peptides,^[8] hybrid polyketide–nonribosomal peptides,^[9] and ribosomally synthesized and post-translationally modified peptides.^[10] Yet, the discrepancy in numbers of diterpenoids found in plants and fungi versus bacteria suggests inefficient means for diterpenoid discovery in bacteria. Given that bacteria indeed possess the biosynthetic machinery needed for diterpenoid production,^[1, 11] genome mining for complex diterpenoid natural products should be achievable.

TSs utilize carbocation chemistry to catalyze their diverse cyclization reactions and are divided into two subtypes,^[11, 12] Type I TSs catalyze the heterolytic cleavage of the diphosphate moiety, that is, ionization, on the prenyl diphosphate substrate; type II TSs protonate a double bond or epoxide, that is, protonation, while leaving the diphosphate moiety intact (Figure S2). Canonical types I and II TSs possess signature sequence motifs, DDxxD and NSE/DTE, and DxDD, respectively, that are associated with initial carbocation formation and have been utilized as probes to identify uncharacterized TSs.^[11] Diterpene synthases (DTSs) with atypical sequence motifs are known, but rare, with the type I *ent*-atiserene synthase PtmT1/PtnT1 (two DxxxD motifs) from PTN biosynthesis^[3b] and the type II DTSs PlaT2 (DSAN) and Bra4 (ESAE) from PLA and BRA biosynthesis, respectively, as examples.^[5a, 6b] The atypical sequence motifs of PlaT2 and Bra4 suggest that they protonate an epoxide, resembling the eukaryotic tri-TS oxidosqualene cyclase (OSC; Figure S2).^[13]

We targeted gene clusters containing genes that encoded atypical type II DTSs. A sequence similarity network (SSN) of type II TSs, built from the Structure-Function Linkage Database at UCSF,^[14] revealed that known type II DTSs only populated four small clusters at an *e* value of 10^{-5} (Figure 1A). The largest cluster contains canonical DTSs from actinobacteria, including, but not limited to, the copalyl diphosphate synthases PtmT2 and CPS, the terpenedienyl diphosphate synthase CYC1, and the labdadienyl diphosphate synthase SCLAV_p0490.^[15] The three smaller clusters contain the atypical DTSs, including PlaT2 and Bra4,^[5a, 6b] BjCPS from the proteobacterium *Bradyrhizobium japonicum*, and halimadienyl diphosphate synthase Rv3377c from *Mycobacterium tuberculosis*.^[16] Interestingly, Lon15, a tetra-TS involved in longestin (LON) biosynthesis/^[17] clusters together with PlaT2 and Bra4. Although tetra-TS Lon15 is included in this subfamily, for simplicity, we use term “DTS” to refer to the subfamily or members therein.

An SSN of PlaT2/Bra4-like DTSs was constructed to differentiate these atypical type II DTSs.^[18] This network, which consists of 66 unique homologues from actinobacteria (Table S1), is comprised of three major subfamilies at an *e* value of 10^{-142} : 44 Bra4-like proteins, 12 PlaT2-like proteins, and six Lon15-like proteins (Figure 1B). Three of these DTSs were from strains in the Actinomycetales strain collection at The Scripps Research Institute. Two DTSs, TnIT2 from CB03234 and CB03238,^[7a, 19] clustered with PlaT2; the third, from *Micromonospora* sp. TSRI0369,^[7a] clustered with Bra4 (Figure 1B). At a threshold of

10⁻¹⁵⁸, Bra4 separates from a cluster that mainly consists of DTSs from *Micromonospora* spp. (Figure 1B). A phylogenetic tree of these 66 DTSs supports the SSN subfamilies (Figure S3). None of these DTSs contain the canonical type II DxDD motif, or retain the central Asp as in OSCs;^[12] instead, they possess a consensus motif of (E/D)(T/S)xE, in which x is an aliphatic amino acid (Figures 1C and S4). A genome neighborhood network (GNN) of the 66 DTSs was constructed to differentiate, annotate, and prioritize the gene clusters for natural-product discovery (Figures S5 and S6).^[20] By using a neighborhood size of 20 genes (41 total genes per DTS) and an *e* value of 10⁻⁵, 60 of the 66 type II DTS-encoding genes are in genetic proximity to genes that encode a polyprenyl diphosphate epoxidase and a UbiA-like prenyltransferase (Table S1). These two proteins were previously predicted to epoxidize the terminal double bond on the polyprenyl diphosphate and catalyze a type I TSs-like extension of the terpene scaffold, respectively.^[5a, 6b] Furthermore, 59 of these 60 gene clusters also encode at least one polyprenyl synthase (Table S1). Together, these four core genes represent the genetic foundation for this family of bacterial diterpenoids. Gene clusters that do not contain all four genes have homologues elsewhere in the genome, although it is unclear if they are involved in the biosynthesis of these natural products (Table S2). In addition, significant numbers of diverse tailoring enzymes encoded by genes in proximity to these core genes indicate that the nascent terpenoid skeletons undergo extensive modifications to afford structural diversity in this family of natural products.

We prioritized CB03234 and CB03238 for diterpenoid discovery based on 1) their presence in our Actinomycetales strain collection (Figure S7); 2) differences in the gene clusters for tiancilactone (TNL), PLA, BRA, and LON (Scheme 1B, Figure S5, and Tables S3–S6); and 3) the fact that CB03234 is genetically amenable.^[7a] We first generated the *tnlT2* mutant strain SB23001 from CB03234 (Tables S7–S9 and Figure S8).^[21] Both CB03234 and CB03238 wild-type (WT) strains were then fermented under conditions known for diterpenoid production,^[22] with SB23001 as a negative control. HPLC analysis revealed three distinct metabolites from the CB03234 and CB03238 WT strains that were abolished in SB23001 (Figure 2A); this confirmed that they were encoded by the *tnl* and *tnl'* gene clusters. A large scale (24 L) fermentation of CB03234 was then performed, which resulted in the isolation of eight natural products, TNLs A–H (**1–8**; Figure 2B). The structures of **1–8** (Scheme 2A) were determined based on a combination of high-resolution ESIMS (HRESIMS), 1D and 2D NMR spectroscopy, chemical transformation, and computational methods (Tables S10–S12 and Figures S9–S71).

TNL A (**1**), the major product of CB03234, had a molecular formula of C₃₂H₄₄ClNO₈ based on HRMS (ESI) analysis (*m/z* 606.2836 [*M*+H]⁺; Figure S9) and the 3:1 ratio of the peak heights for a Cl-containing natural product (Figure S10). The NMR spectra (Table S10) of **1** showed resonances that were consistent with those of PLA CD2, with two key differences (Figure S1).^[5a] The C-12—C-13 double bond in PLA CD2 was lost and the C-12 position in **1** was oxygenated, as supported by the ¹H NMR doublet of CH₃-16 and HMBC correlations of CH₃-16 with C-12 and C-13 (Scheme 2B). The second difference was the moiety attached at C-18. The low-field signals of δ_{H} = 7.79 (d), 7.31 (dd), and 6.68 ppm (d) suggested the presence of a 2,4,5-trisubstituted benzoate moiety. In addition, HMBC correlations of H-3'

with C-1', C-2', C-4', and C-7' suggested that this 2,4,5-trisubstituted benzoate was a C-4'-chloroanthranilate (Scheme 2A and B).

The structures of **2–7** were similarly elucidated (see the Supporting Information). TNL B (**2**) was the deschloro congener of **1**; TNL C (**3**) possessed the chloroanthranilate moiety on C-3 instead of C-18; TNLs D (**4**) and E (**5**) had modified γ -butyrolactone moieties; TNL F (**6**) lost a carboxyl group on the γ -butyrolactone moiety of **1**, which resulted in a malondialdehyde side chain; TNL G (**7**) had the γ -butyrolactone moiety of **1** replaced by an ethanol tail; and TNL H (**8**), which was not oxidized at C-18 or C-12 and did not possess the chloroanthranilate and γ -butyrolactone moieties, was an ether-containing (C-11 and C-15) diterpenoid skeleton with a carboxylic acid side chain (Scheme 2A).

The relative stereochemistry of **1** was elucidated by analysis of the ROESY spectra. Correlations between H-3 and H-1b, H-3 and H-5, H-5 and H-9, H-9 and CH₃-17, and CH₃-20 and CH₃-19 established the 3*S**, 4*R**, 5*R**, 8*S**, 9*R**, and 10*R** configurations (Scheme 2B), the result of which was an *anti/anti/syn*-perhy-drophenanthrene scaffold. The key ROESY correlations between CH₃-20 and H-12, and H-12 and H-14, along with a correlation between H-13 and CH₃-17, established the 12*S**, 13*R**, and 14*S** configurations (Scheme 2B). Because each of these key ROESY correlations was also detected in **2–7**, their relative stereochemistries were assigned to be the same as that of **1**.

Based on the ROESY correlations observed, the stereochemistries of C-3, C-5, C-8, C-9, and C-10 in **8** are identical to those in **1** (Scheme 2C). Furthermore, ROESY correlations between H-12 and CH₃-20, together with correlations between H-15 and CH₃-17, and H-15 and H-9, supported the 11*R** and 15*S** configurations of **8**. The absolute stereochemistry of C-3 was determined to be *S* by means of the modified Mosher method (Scheme 2C and Figures S68–S71).^[23] Taken together, the absolute stereochemistry of **8**, and thus, the TNLs, was established (Scheme 2A). This assignment was further supported by density functional theory calculations with the optical rotation (OR) of **8** (see the Supporting Information): experimental, -108.3° ; calculated, -128.2° .

Using the GNN, we predicted the genetic boundaries of the *tnl* gene cluster and generated the *tnlorf1* and *tnlorf3* mutant strains SB23002 and SB23003, respectively, from CB03234 (Figures S5, S72, and S73).^[21] Fermentation of SB23002 and SB23003 revealed that neither mutation affected TNL production (Figure S74); this confirmed that *orf1* and *orf3* were not required for TNL biosynthesis. We also inactivated two methyltransferase-encoding genes, *tnlM1* and *tnlP5*, which resulted in mutant strains SB23004 and SB23005, respectively (Tables S7–S9 and Figures S75 and S76). Fermentation of SB23004 and SB23005 resulted in the accumulation of three new metabolites, TNLs I (**9**), J (**10**), and K (**11**; Scheme 2A, Figures 2A, and S77–S97, as well as Table S13). The disappearance of the CH₃-24 signal and an upfield migration of the chemical shift of C-12 in **9** supported the structure of **9** and the function of TnlM1 as an O-methyltransferase. Supported by a new signal at $\delta_C = 216.5$ ppm and a molecular weight 2 Da less than that of **9**, the structure of **10** had a ketone at C-12. TNL K (**11**), clearly lost the signal of CH₃-8' ($\delta_H = 2.86$ ppm) in **1**, which confirmed the demethylated structure of **11** and TnlP5 as an N-methyltransferase.

The discovery of the TNLs and genetic studies of the *tnl* gene cluster, along with previous studies on the PLA and BRA biosynthetic pathways,^[5a, b, 6b, c] provides a logical biosynthetic pathway for the TNLs (Scheme 3). After geranylgeranyl diphosphate (GGPP) is biosynthesized and terminally epoxidized by TnlT4 and TnlT1, respectively, TnlT2 catalyzes the cyclization of epoxyGGPP to form a (3S)-hydroxydodecahydrophenanthrene scaffold, which presumably retains the diphosphate group at C-15. TnlT3 then catalyzes a prenylation-like addition of three carbon atoms to the C-15 carbocation, possibly in the form of phosphoenolpyruvate;^[5a] these four carbon atoms eventually form the γ -butyrolactone moiety. Based on the isolation of PLA CD6, which contains an enone in the C-15 side chain (Figure S1),^[5a] TNL is likely to undergo a similar oxidation prior to C-21 hydroxylation and γ -butyrolactone formation by TnlO1 (Scheme 3A). The intermediacy of this oxidized 23-carbon scaffold is supported by the structure of **8**, which is a decarboxylated and C-11 oxidized (the ether might arise from a nonenzymatic oxo-Michael addition) shunt product that retains the C-12—C-13 double bond and unoxidized C-18 methyl group (Scheme 2A). The enzyme responsible for desaturation at C-15—C-21 is still unknown given that the previously proposed dehydrogenase in PLA biosynthesis, PlaZ,^[5a] is not conserved in the TNL gene clusters.

Four tailoring steps complete the biosynthesis of **1** (Scheme 3). Hydroxylation at C-18 and C-12 is required prior to acylation and methylation. Based on analysis by using a cytochrome P450 SSN from *Streptomyces* (Figure S98),^[25] these hydroxylations are expected to be catalyzed by the TnlO4 and TnlO3 P450s, respectively; sequence similarities between PlaO2 and TnlO2 and the lack of a third, clear P450 transformation suggest that TnlO2 may be nonfunctional.^[5b] Based on the presence of genes encoding the tryptophan dioxygenase TnlP2 and tryptophan halogenase TnlP3, and the isolation of the deschloro congener **2**, the chloroanthranilate moiety is derived from tryptophan via (chloro)kynurenine. Chloroanthranilate is first activated by the adenosine triphosphate (ATP)-dependent synthetase TnlP4, which is transferred to the peptidyl carrier protein TnlP1, and finally transferred to the C-18 hydroxy group. C-18 acylation is likely to be catalyzed by TnlP6 (Scheme 3B), based on its similarity to the acyltransferase CloN7 from clorobiocin biosynthesis.^[26] N-Methylation on the anthranilate moiety is catalyzed by TnlP5; however, it is unclear if this occurs prior to or after anthranilate installation given that desmethyl congener **11** has no other structural changes, relative to **1** (Scheme 2A). Finally, O-methylation, which is catalyzed by TnlM1 and supported by the isolation of **9**, completes the biosynthesis of **1** (Scheme 3). The C-12 oxo moiety of **10** is likely to result from extraneous oxidation of the C-12 hydroxy group. TNLs C-G (**3–7**) are presumably shunt products or isolation artifacts that are either the result of nonenzymatic transformations, that is, **3–5**, or degradation of the γ -butyrolactone moiety, that is, **6** and **7**.

The TNLs showed moderate broad-spectrum antibacterial activities. A disk diffusion assay showed clear inhibition zones for **1** and **3–5** against *Staphylococcus aureus*, *Kocuria rhizophila*, *Mycobacterium smegmatis*, and *Pseudomonas putida* (Figure S99). Inhibition zones were also found for **1** and **3–5** against *Bacillus subtilis* and *Escherichia coli*, but were not as clear. Minimum inhibitory concentrations (MICs) of TNLs **1–5** and **8** were determined by using the broth dilution method.^[24] Although none of the compounds showed

activity against *B. subtilis* or *E. coli* (i.e., MIC > 64 $\mu\text{g mL}^{-1}$), TNLs A–E (**1–5**) exhibited MICs ranging from 8–64 $\mu\text{g mL}^{-1}$ against *S. aureus*, *K. rhizophila*, *M. smegmatis*, and *P. putida* (Table S14). Due to limited quantities, TNLs F (**6**) and G (**7**) were not tested. No inhibition zones were found for **9–11** against any of the tested strains, which suggested that the C-12 methoxy group and/or the N-methyl group of the chloroanthranilate moiety were indispensable for antibiotic activity. In comparison, the PLAs exhibited MIC values of 3–30 $\mu\text{g mL}^{-1}$ against Gram-positive bacteria, but had no activity against Gram-negative strains; [5c] BRA showed no antibacterial activity.[6a]

In summary, we identified a family of 66 putative diterpenoid gene clusters, 63 of which were new, by searching for gene clusters that encoded type II DTSs without the canonical DxDD motif. This study challenges the current paradigm of the rarity of terpenoids in bacteria, and provides an extraordinary opportunity to discover and study new biosynthetic pathways that result in highly functionalized bacterial terpenoids. The unique structures of the TNLs, composed of a diterpene-derived *anti/anti/syn*-perhydrophenanthrene skeleton attached to γ -butyrolactone and chloroanthranilate moieties, support future screening efforts of these diterpenoids against an array of biological targets. Furthermore, differences in the activities of the TNLs, PLAs, and BRA suggest distinct modes of action. The *anti/anti/syn* and *anti/syn/anti* configurations of the tricyclic cores in TNL and BRA implicate that the DTSs or polyprenyl epoxidases in these gene clusters do not all generate identical stereoisomers (Figures S3 and S100); this further expands the potential for structural diversity. Finally, although canonical enzymes, domains, or motifs are typically used in genome mining, the discovery of the TNLs by identifying biosynthetic machineries with atypical motifs is applicable to all classes of natural products.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work is supported, in part, by National Institutes of Health grant GM114353 (to B.S.). J.D.R. is supported, in part, by a National Institutes of Health Pathway to Independence Award GM124461. M.-R.D. is supported, in part, by the Guangdong Institute of Microbiology, Guangzhou, Guangdong, P.R. China, and a scholarship (2017GDASCX-0502) from the Guangdong Academy of Science, Guangzhou, Guangdong, P.R. China. We thank Dr. Qin-Shi Zhao from the Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, P.R. China, for access to Discovery Studio. We thank the High Performance Computing Core at The Scripps Research Institute for providing HPC resources. This is manuscript #29668 from The Scripps Research Institute.

References

- [1]. Smanski MJ, Peterson RM, Huang S-X, Shen B, Curr. Opin. Chem. Biol 2012, 16, 132–141. [PubMed: 22445175]
- [2]. Yamada Y, Kuzuyama T, Komatsu M, Shin-ya K, Omura S, Cane DE, Ikeda H, Proc. Natl. Acad. Sci. USA 2015, 112, 857–862. [PubMed: 25535391]
- [3]. a) Rudolf JD, Dong L-B, Shen B, Biochem. Pharmacol 2017, 133, 139–151; [PubMed: 27865713]
b) Smanski MJ, Yu Z, Casper J, Lin S, Peterson RM, Chen Y, Wendt-Pienkowski E, Rajski SR, Shen B, Proc. Natl. Acad. Sci. USA 2011, 108, 13498–13503. [PubMed: 21825154]
- [4]. Hedden P, Sponsel V, J. Plant Growth Regul. J. 2015, 34, 740–760.

- [5]. a) Dürr C, Schnell H-J, Luzhetskyy A, Murillo R, Weber M, Welzel K, Vente A, Bechthold A, Chem. Biol 2006, 13, 365–377; [PubMed: 16632249] b) Daum M, Schnell H-J, Herrmann S, Guenther A, Murillo R, Mueller R, Bisel P, Mueller M, Bechthold A, ChemBioChem 2010, 11, 1383–1391; [PubMed: 20540056] c) Gebhardt K, Meyer SW, Schinko J, Bringmann G, Zeeck A, Fiedler H-P, J. Antibiot 2011, 64, 229–232. [PubMed: 21179045]
- [6]. a) Komaki H, Nemoto A, Tanaka Y, Takagi H, Yazawa K, Mikami Y, Shigemori H, Kobayashi J. i., Ando A, Nagata Y, J. Antibiot 1999, 52, 13–19; [PubMed: 10092191] b) Hayashi Y, Matsuura N, Toshima H, Itoh N, Ishikawa J, Mikami Y, Daiiri T, J. Antibiot 2008, 61, 164–174; [PubMed: 18503195] c) Schwarz PN, Roller L, Kulik A, Wohlleben W, Stegmann E, Buchmann A, Gross H, Gross H, Wohlleben W, Stegmann E, Biotechnol. J 2018, 13, 1700527.
- [7]. a) Yan X, Ge H, Huang T, Hindra, Yang D, Teng Q, Crnovcic I, Li X, Rudolf J, R. D, Lohman J, Gansemans Y, Zhu X, Huang Y, Zhao L-X, Jiang Y, Van Nieuwerburgh F, Rader C, Duan Y, Shen B, mBio 2016, 7, e02104–16; [PubMed: 27999165] b) Helfrich EJN, Reiter S, Piel J, Curr. Opin. Biotechnol 2014, 29, 107–115; [PubMed: 24762576] c) Boddy CN, J. Ind. Microbiol. Biotechnol 2014, 41, 443–450. [PubMed: 24174214]
- [8]. Baltz RH, J. Ind. Microbiol. Biotechnol 2018, <https://doi.org/10.1007/s10295-017-1999-8>.
- [9]. Pan G, Xu Z, Guo Z, Hindra, Ma M, Yang D, Zhou H, Gansemans Y, Zhu X, Huang Y, Zhao L-X, Jiang Y, Cheng J, Van Nieuwerburgh F, Suh J-W, Duan Y, Shen B, Proc. Natl. Acad. Sci. USA 2017, 114, E11131–E11140. [PubMed: 29229819]
- [10]. Hetrick KJ, van der Donk WA, Curr. Opin. Chem. Biol 2017, 38, 36–44. [PubMed: 28260651]
- [11]. Dickschat JS, Nat. Prod. Rep 2016, 33, 87–110. [PubMed: 26563452]
- [12]. Christianson DW, Chem. Rev 2006, 106, 3412–3442. [PubMed: 16895335]
- [13]. Wendt KU, Schulz GE, Corey EJ, Liu DR, Angew. Chem. Int. Ed 2000, 39, 2812–2833; Angew. Chem 2000, 112, 2930–2952.
- [14]. Akiva E, Brown S, Almonacid DE, Barber II AE, Custer AF, Hicks MA, Huang CC, Lauck F, Mashiyama ST, Meng EC, Mischel D, Morris JH, Ojha S, Schnoes AM, Stryke D, Yunes JM, Ferrin TE, Holliday GL, Babbitt PC, Nucleic Acids Res. 2014, 42, D521–D530. [PubMed: 24271399]
- [15]. a) Rudolf JD, Dong L-B, Cao H, Hatzos-Skintges C, Osipiuk J, Endres M, Chang C-Y, Ma M, Babnigg G, Joachimiak A, Phillips GN, Shen B, J. Am. Chem. Soc 2016, 138, 10905–10915; [PubMed: 27490479] b) Ikeda C, Hayashi Y, Itoh N, Seto H, Daiiri T, J. Biochem 2007, 141, 37–45; [PubMed: 17148547] c) Xu M, Hillwig ML, Lane AL, Tiernan MS, Moore BS, Peters RJ, J. Nat. Prod 2014, 77, 2144–2147; [PubMed: 25203741] d) Hamano Y, Kuzuyama T, Itoh N, Furihata K, Seto H, Daiiri T, J. Biol. Chem 2002, 277, 37098–37104; [PubMed: 12138123] e) Yamada Y, Komatsu M, Ikeda H, J. Antibiot 2016, 69, 515–523. [PubMed: 26814669]
- [16]. a) Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ, FEBS Lett. 2009, 583, 475–480; [PubMed: 19121310] b) Nakano C, Okamura T, Sato T, Daiiri T, Hoshino T, Chem. Commun 2005, 1016–1018.
- [17]. Hayashi Y, Onaka H, Itoh N, Seto H, Daiiri T, Biosci. Biotechnol. Biochem 2007, 71, 3072–3081. [PubMed: 18071268]
- [18]. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL, Biochim. Biophys. Acta Proteins Proteomics 2015, 1854, 1019–1037.
- [19]. Hindra D, Yang Q, Teng L-B, Dong I, Crnovcic T, Huang H, Ge B, Shen, Org. Lett 2017, 19, 1386–1389. [PubMed: 28256838]
- [20]. Gerlt JA, Biochemistry 2017, 56, 4293–4308. [PubMed: 28826221]
- [21]. Gust B, Challis GL, Fowler K, Kieser T, Chater KF, Proc. Natl. Acad. Sci. USA 2003, 100, 1541–1546. [PubMed: 12563033]
- [22]. Smanski MJ, Peterson RM, Rajski SR, Shen B, Antimicrob. Agents Chemother 2009, 53, 1299–1304. [PubMed: 19164156]
- [23]. Ohtani I, Kusumi T, Kashman Y, Kakisawa H, J. Am. Chem. Soc 1991, 113, 4092–4096.
- [24]. Wiegand I, Hilpert K, Hancock REW, Nat. Protoc 2008, 3, 163–175. [PubMed: 18274517]
- [25]. Rudolf JD, Chang C-Y, Ma M, Shen B, Nat. Prod. Rep 2017, 34, 1141–1172. [PubMed: 28758170]

- [26]. Anderle C, Alt S, Bringmann G, Gust B, Heide L, Arch. Microbiol 2007, 187, 227–237.
[PubMed: 17308937]

Figure 1.

Genome mining of atypical bacterial type II DTSs. A) Representative SSN of type II DTSs displayed at an e value of 10^{-50} . Large nodes depict characterized type II DTSs in bacteria. Colors of nodes represent the source of the protein; see inset legend. B) SSN of the 66 type II DTS homologues found in the TnIT2 subfamily displayed at e values of 10^{-142} and 10^{-158} . Shapes and colors of nodes represent bacterial genera and associated biosynthetic pathway, respectively; see inset legend. C) Sequence logo of the consensus motif of (E/D)(T/S)xE.

Figure 2. Discovery of the TNLs. A) Metabolite profiles of *Streptomyces* strains upon HPLC analysis: I) CB03234; inset shows a magnification of part of the chromatogram; II) CB03238; III) SB23001 (CB03234 *tnlT2*); IV) SB23004 (CB03234 *tnlMI*); and V) SB23005 (CB03234 *tnlP5*). B) Representative metabolite profile of the diterpene-enriched fractions of large-scale fermentation of CB03234 upon LC-MS analysis (total ion current).

Scheme 1.

Structures and gene clusters of four terpenoids found within the 66 putative diterpenoid producers. A) The TNLs were discovered in this study; PLA, BRA, and LON (KS-505a) were known. The absolute stereochemistries of BRA and TNL were determined; only the relative stereochemistries of PLA and LON were reported. B) Genetic organizations of the *tnl* (CB03234), *tnl'* (CB03238), *pla*, *bra*, and *lon* biosynthetic gene clusters, highlighting the genetic and potential structural diversities found within this family natural products. Genes are colored based upon their annotated Pfam IDs (Tables S3–S6). GenBank accession numbers for *tnl*/CB03234 (BK010469) and *tnl*/CB03238 (BK010470).

Scheme 2.

Structural characterization of the TNLs. A) Structures of TNLs A–K (1–11, respectively). B) Structural and relative stereochemical determination of **1** by means of 2D NMR spectroscopy. C) Absolute stereochemical determination of **8** by means of ROESY correlations and the modified Mosher method.

Scheme 3.

Proposed biosynthetic pathway for TNL A (**1**) supported by the isolation of congeners or shunt metabolites **4–11**, as indicated by the dashed arrows. A) The four core proteins, TnlT1–T4, construct the diterpenoid scaffold and elongate the side chain. Then, a series of oxidative reactions form the γ -butyrolactone moiety and hydroxylate the C-18 and C-12 positions before acylation of the chloroanthranilate moiety. Finally, the N- and O-methylations, catalyzed by TnlP5 and TnlM1, respectively, complete the biosynthesis, although their timing is still unknown. IPP: isopentenyl diphosphate, DMAPP: dimethylallyl diphosphate, GGPP: geranylgeranyl diphosphate. B) Biosynthesis of the chloroanthranilate moiety from L-tryptophan follows that of tryptophan degradation through the kynurenine pathway. Chemical transformations are colored red; for multiple transformations in one step, red and blue are used.