Published in final edited form as:

Org Lett. 2017 April 07; 19(7): 1772–1775. doi:10.1021/acs.orglett.7b00545.

Linear peptides are the major products of a biosynthetic pathway that encodes for cyclic depsipeptides

Thomas P. Wyche^{†,‡,§}, Antonio C. Ruzzini^{†,§}, Christine Beemelmans^{†,||,§}, Ki Hyun Kim^{†,○}, Jonathan L. Klassen^{⊥,¥}, Shugeng Cao^{†,Ω}, Michael Poulsen^{⊥,€}, Tim S. Bugni[‡], Cameron R. Currie[⊥], and Jon Clardy^{†,*}

[†]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA, 02115

[‡]Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA, 53705

[⊥]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA, 53706

Abstract

Three new dentigerumycin analogs are produced by *Streptomyces* sp. M41, a bacterium isolated from a South African termite, *Macrotermes natalensis*. The structures of the complex nonribosomal peptide synthetase–polyketide synthase (NRPS-PKS) hybrid compounds were determined by 1D- and 2D-NMR spectroscopy, high-resolution mass spectrometry, and circular dichroism (CD) spectroscopy. Both cyclic and linear peptides are reported, and the genetic organization of the NRPS modules within the biosynthetic gene cluster accounts for the observed structural diversity.

Graphical Abstract

Bacteria living in close association with insect hosts are increasingly appreciated as productive sources of biologically active small molecules. Many of these bacteria belong to

*Corresponding Author: jon_clardy@hms.harvard.edu.

^{||}Leibniz Institute for Natural Product Research and Infection Biology, Hans-Knöll Institute, 07745, Jena, Germany

[○]School of Pharmacy, Sungkyunkwan University, Suwon, Gyeonggi-do, 440–746, Republic of Korea

[¥]Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, 06269, USA

^ΩDaniel K. Inouye College of Pharmacy, University of Hawaii at Hilo, HI, 968720, USA

[€]Center for Social Evolution, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Denmark

[§]T.P.W., A.C.R., and C.B. contributed equally to this work.

Author Contributions

The manuscript was written through contributions of all authors.

Supporting Information

Procedures for isolation and characterization of **1–3**, dereplication method, compound purification, and HRMS and NMR spectra and assignments. The Supporting Information is available free of charge on the ACS Publications website.

genera that have a rich history in the field of small molecule discovery, and their existence in well-defined ecological niches helps identify their likely functional roles. Prospecting for naturally occurring small molecules from these sources represents an artisanal approach to natural product discovery – one based on the premise that these bacteria and their chemical defenses have been selected to promote survival.

We recently started to investigate bacterially-produced natural products from fungus-growing termites in the genus *Macrotermes* using two complementary approaches. The targeted discovery of antifungal compounds using phenotypic screens that began from ecologically-paired microorganisms resulted in the isolation and characterization of a new ansamycin named natalamycin A¹ and new glycosylated macrolactams named macrotermycin A–D.² An unbiased high-throughput mass spectrometry-based pipeline has also yielded new natural products, including the cyclic depsipeptides microtermolide A and B.³ Here, we continue to establish the microbial associates of fungus-growing termites as a productive source of chemical diversity using this unbiased investigation of the metabolites from Actinobacteria isolated from termites. To our surprise, this investigation expanded the membership of a familiar nonribosomal peptide synthetase/polyketide synthase (NRPS/PKS) scaffold that has also been reported from fungus-growing ants, a distinct bacterial and insect population that also practices fungiculture. The two fungus-growing systems are also geographically distinct with the ants in the New World Tropics and the termites in Sub-Saharan Africa.

As part of an ongoing program to catalog the chemical potential of termite nest-derived bacteria, we applied a high-throughput dereplication strategy that is based on principal component analyses of high-performance liquid chromatography–high-resolution mass spectrometry (HPLC-HRMS) data.⁴ Using this platform, we recently analyzed a total of 41 Actinobacteria that were isolated from *Macrotermes natalensis* termites and focused on the metabolites produced by *Amycolatopsis* sp. M39.² An additional isolate, *Streptomyces* sp. M41 (hereafter, M41), was prioritized based on its unique metabolomic profile, and the results are presented here. A compound with m/z 767.4270 $[M + H]^+$ was particularly distinctive from our initial principal component analysis (PCA) (Figure S1); analysis in negative ion mode (m/z 783.4276 $[M - H]^-$) revealed a metabolite with the predicted molecular formula of $C_{35}H_{60}N_8O_{12}$. A large scale fermentation of M41 revealed an additional analog with another previously unreported molecular formula: $C_{35}H_{61}N_9O_{11}$, predicted based on the ion m/z 782.4455 $[M - H]^-$. These two potentially novel metabolites were isolated using standard chromatographic procedures (SI).

The ¹H NMR spectrum of the first isolated metabolite **1** (m/z 783.4276 $[M - H]^-$) showed typical features of a peptide-derived compound, including six α proton signals (δ_H 5.58, 5.46, 5.44, 4.70, 4.55, and 4.22). The ¹³C NMR and gHSQC spectra were similar to dentigerumycin and contained six amide or ester carbonyl carbons, four oxygen-bearing carbons, and eight nitrogen-bearing carbons. Interpretation of gCOSY, TOCSY, gHSQC, and gHMBC NMR spectra allowed us to assign five amino acid-derived and one polyketide-derived substructure, which included three piperazic acid units (Pip-1, Pip-2, and Pip-3), β -hydroxyleucine (β -OH-Leu), *N*-hydroxyglycine (*N*-OH-Gly), and a pyran-bearing acyl side chain (Table S1). In particular, the presence of the *N*-OH-Gly was confirmed by upfield

shifts of C-12 (δ_C 50.2 and δ_H 4.70 and 4.22) compared to those of glycine. To validate our *N*-OH assignment, we performed a Ti(III)Cl₃-mediated *N*-OH reduction as described previously,⁵ and confirmed the identity of the reduced product by NMR and HRMS (Table S3). Based on the predicted molecular formula and NMR data analysis, compound **1** was a linear peptide with a terminal carboxylic acid. The sequence of the amino acid units and the acyl side chain were established by analyses of gHMBC correlations, completing the planar structure of **1**. The stereochemistry was determined using a combination of spectroscopic analysis, degradative reactions, and derivatizations. The relative configuration of the pyran ring in the acyl side chain was determined by ROESY correlations, and that of the two consecutive stereogenic centers of β -OH-Leu was established as 19*S** and 20*R** by *J*-based configuration analysis. Marfey's method – in combination with a comparison of ¹H and ¹³C NMR shifts for dentigerumycin – allowed for the determination of the absolute configurations of the amino acid residues as 2*S*, 7*R*, 14*R*, 19*S*, and 20*R*.

We next isolated and characterized metabolite **2** (C₃₅H₆₁N₉O₁₁), which based on HRMS data differed from **1** by the addition of NH and loss of an oxygen. Indeed, analysis of 1D and 2D NMR data confirmed that compound **2** was very similar to **1**; we could assign five amino acid-derived substructures and one polyketide-derived substructure including three piperazic acid units (Pip-1, Pip-2, and Pip-3), β -hydroxyleucine (β -OH-Leu), and the same pyran-bearing acyl side chain as found for **1** (Table S2). The hydroxyl peak of the *N*-OH-Gly was also present in the ¹H NMR spectrum. However, two additional hydrogens (δ_H 7.58, 7.26) were present in the ¹H NMR spectrum of **2** that were likely attributed to an amide. Analysis of ¹⁵N HSQC and ¹⁵N HMBC data from ¹⁵N-labeled **2** confirmed the presence of a terminal amide (Figure 1), resulting in a linear peptide. Instead of a terminal carboxylic acid, compound **2** contained a terminal amide. The sequence of amino acid units and the acyl side chain was again determined by key correlations observed in the gHMBC NMR spectrum, and the stereochemistry of **2** was established by the same methods as described for **1** using a combination of spectroscopic analysis, degradative reactions, and derivatizations.

The chemical structures of **1** and **2** are reminiscent of dentigerumycin, a cyclic depsipeptide that was originally isolated from an Actinobacterial symbiont of fungus-growing ants.^{5,6} A closer inspection of minor products in the organic extract revealed another potentially new metabolite with the predicted molecular formula C₃₉H₆₅N₉O₁₄ based on a positive ion of *m/z* 906.4573 [M + Na]⁺. We isolated and characterized this third low abundance metabolite, compound **3** (C₃₉H₆₅N₉O₁₄). Analysis of ¹H and ¹³C NMR spectra revealed that **3** was also remarkably similar to dentigerumycin. The minor M41 product **3** and dentigerumycin differ by a single methylene unit on the pyran ring side chain and at three amino acids: the *N*-OH-Ala and ester-forming Ala of dentigerumycin are replaced by an *N*-OH-Gly and *N*-OH-Thr, respectively in **3**, which also lacks a hydroxylation on Pip-1. We established the sequence of the amino acid units and the acyl side chain of **3** by extensive analyses of gHMBC NMR data, which completed the planar structure. Shielded NMR signals for the Thr C α (δ_C 65.3 and δ_H 4.71) suggested *N*-hydroxylation of the amino acid, which was confirmed by Ti(III)Cl₃-mediated *N*-OH reduction (Table S3). The stereochemistry was determined using a combination of spectroscopic analysis, degradative reactions, and derivatizations, as was discussed for **1** (SI). Finally, the relative configuration

of stereogenic centers in *N*-hydroxythreonine was established as 25 S^* and 26 S^* by ROESY experiment and *J*-based configuration analysis ($^3J_{25,26} = 7.0$ Hz).

The three new metabolites were named based on their structural similarity to dentigerumycin (Figure 1). The cyclic depsipeptide **3** most similar to dentigerumycin was named dentigerumycin B, and the linear peptides **1** and **2** were named dentigerumycin C and D, respectively.

The chemical diversity that we observed for the three dentigerumycin analogs, in particular the difference between the linear and cyclic peptide, prompted further characterization of the producing strain. We therefore sequenced the genome of M41 using PacBio SMRT sequencing technology⁷ in order to resolve the genetic basis of dentigerumycin analog diversity. PacBio reads were assembled to 5 replicons (Table S4) using the hierarchical genome assembly process⁸, and a single biosynthetic locus responsible for dentigerumycin production was readily detected using antiSMASH⁹ prediction software (Figure 2). The predicted biosynthetic gene cluster (BGC) is ~100 kb, including the requisite polyketide synthases and 4 multimodular NRPSs genes. The genetic organization of the NRPS genes helps to explain the observed structural diversity. All four of the NRPS genes are required for the production of **3** whereas **1** and **2**, despite being the major fermentation products, require only three of the four multimodular NRPS genes encoded within the BGC (Figure 2). A number of the remaining genes in the BGC might explain the abundant formation of **1** and **2**, which on the basis of the BGC are incomplete products. For example, a number of stand-alone enzymes with sequence similarity to thioesterases may be employed to generate these alternative natural products by prematurely releasing the growing polypeptide prior to the incorporation of the final peptide (L-Thr) and cyclization. An alternative hypothesis is that the transfer of the growing polypeptide from the third to the final NRPS is simply inefficient. This scenario is rationalized by the highly modular nature of the biosynthesis of these molecules. At a multigene level, the results may reflect the manner in which large biosynthetic pathways evolve: through the combination of smaller compatible parts.¹⁰ Here, for instance, the compatibility of the observed biosynthetic parts may be reduced.

The dentigerumycins belong to a rather large family that includes a number of distinct amino acid constituents at each position within the cyclic depsipeptide end products. This family of related hybrid PKS/NRPS natural products, which includes but is not limited to azinotricin,¹¹ the polyoxypeptins,^{12,13} and verucopeptin,^{14,15} are typically characterized by a 19-membered hexapeptide macrocycle that includes at least one piperazic acid and a highly conserved β -OH-leucine. Due to their nanomolar potencies against cancer cell lines,¹⁶ a number of these molecules have been the subject of total synthesis.^{17,18} The structures of dentigerumycin A and B are remarkable within the family in that three of the six amino acids that form the 19-membered macrocycle are piperazic acids, including two that occur in tandem. The β -OH-Leu is critical to the molecular structure of this hybrid family of natural products as the β -OH participates in cyclization of the peptide macrocycle and the amino group forms a peptide bond with the scaffold's polyketide-derived appendage. Interestingly, a family of closely-related structures that include the substitution of the β -OH-leucine for an α -ketoacid results in a very similar 18-membered macrocycle that much like the dentigerumycins contain three piperazic acids.^{6,16}

The structures of dentigerumycin B and C are incomplete products of the organism's biosynthetic gene cluster. The basis for the observed deviation in structure is one that follows the genetic organization of the NRPS modules. In the case of the dentigerumycin C and D structures, an entire multimodular NRPS is bypassed to afford the linear polypeptide structure whereas all four of the NRPSs typically operate to produce the 19-membered macrocycle. The results presented here are not unusual in that these types of so-called shunt or incomplete metabolites are frequently reported, including within this family of natural products. Aurantimycin D, for example, is a PKS-derivatized dipeptide that possesses a C-terminal amide resembling that of dentigerumycin C.¹⁹ The natural products GE3 and GE3B have also been reported as a macrocyclic depsipeptide and its inactive linear analog, respectively.²⁰ What is unusual about our results is the ratio at which these metabolites are produced: the linear or shunt metabolites are the dominant products of the bacterium whereas the complete macrocyclic antibiotic is produced sparingly. The results further highlight the structural diversity that can be generated from a single biosynthetic gene cluster. In this case, a functional NRPS is apparently underemployed during the biosynthesis of the major cyclic depsipeptides isolated from M41.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful for financial support through the NIH (R01-GM086258 to J.C., RC4-GM096347 to J.C. and C.R.C., and R01-GM104192 to T.S.B.) and through the German National Academy of Sciences Leopoldina for a postdoctoral fellowship to C.B. (LPDS 2011-2). This research was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, & Future Planning (2015R1C1A1A02037383). A.C.R. was supported by a PDF from the Canadian Institutes of Health Research.

References

1. Kim KH, Ramadhar TR, Beemelmans C, Cao S, Poulsen M, Currie CR, Clardy J. *Chem Sci*. 2014; 5:4333–4338. [PubMed: 25386334]
2. Beemelmans C, Kim KH, Klassen J, Cao S, Wyche TP, Hou Y, Ramadhar TR, Poulsen M, Bugni TS, Currie CR, Clardy J. *Org Lett*. 2017; 19:1000–1003. [PubMed: 28207275]
3. Carr G, Poulsen M, Klassen JL, Hou Y, Wyche TP, Bugni TS, Currie CR, Clardy J. *Org Lett*. 2012; 14:2822–2825. [PubMed: 22591554]
4. Hou Y, Braun DR, Michel CR, Klassen JL, Adnani N, Wyche TP, Bugni TS. *Anal Chem*. 2012; 84:4277–4283. [PubMed: 22519562]
5. Oh DC, Poulsen M, Currie CR, Clardy J. *Nat Chem Biol*. 2009; 5:391–393. [PubMed: 19330011]
6. Sit CS, Ruzzini AC, Van Arnam EB, Ramadhar TR, Currie CR, Clardy J. *Proc Natl Acad Sci U S A*. 2015; 112:13150–13154. [PubMed: 26438860]
7. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. *Science*. 2009; 323:133–138. [PubMed: 19023044]

8. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. *Nat Methods*. 2013; 10:563–569. [PubMed: 23644548]
9. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Muller R, Wohlleben W, Breitling R, Takano E, Medema MH. *Nucleic Acids Res*. 2015; 43:W237–W243. [PubMed: 25948579]
10. Medema MH, Cimermancic P, Sali A, Takano E, Fischbach MA. *PLoS Comput Biol*. 2014; 10:e1004016. [PubMed: 25474254]
11. Maehr H, Liu CM, Palleroni NJ, Smallheer J, Todaro L, Williams TH, Blount JF. *J Antibiot (Tokyo)*. 1986; 39:17–25. [PubMed: 3753967]
12. Umezawa K, Nakazawa K, Uemura T, Ilccda Y, Kondo S, Naganawa H, Kinoshita N, Hashizume H, Hamada M, Takeuchi T, Ohba S. *Tetrahedron Lett*. 1998; 39:1389–1392.
13. Umezawa K, Nakazawa K, Ikeda Y, Naganawa H, Kondo S. *J Org Chem*. 1999; 64:3034–3038. [PubMed: 11674399]
14. Sugawara K, Toda S, Moriyama T, Konishi M, Oki T. *J Antibiot (Tokyo)*. 1993; 46:928–35. [PubMed: 8344874]
15. Yoshimura A, Nishimura S, Otsuka S, Hattori A, Kakeya H. *Org Lett*. 2015; 17:5364–5367. [PubMed: 26484856]
16. Miller ED, Kauffman CA, Jensen PR, Fenical W. *J Org Chem*. 2007; 72:323–330. [PubMed: 17221946]
17. Hale KJ, Manaviazar S, George J. *Chem Commun (Camb)*. 2010; 46:4021–4042. [PubMed: 20405066]
18. Li W, Gan J, Ma D. *Angew Chem Int Ed Engl*. 2009; 48:8891–8895. [PubMed: 19839021]
19. Schlegel R, Ritzau M, Ihn W, Stengel C, Graefe U. *Nat Prod Lett*. 1995; 6:171–176.
20. Agatsuma T, Sakai Y, Mizukami T, Saitoh Y. *J Antibiot (Tokyo)*. 1997; 50:704–708. [PubMed: 9315088]

Figure 1.
The chemical structure of dentigerumycin B – D.

Figure 2.

Gene map of the dentigerumycin BGC in *Streptomyces* sp. M41. The multimodular NRPSs that build the dentigerumycins are highlighted in blue and red, and the amino acid specificity of the adenylation domains are listed. One hypothesis regarding the production of **1** and **2** is that the growing nonribosomal peptide is intercepted by at least one of the four genes annotated as thioesterases (green), or a fifth belonging to the same enzyme superfamily, the α/β -hydrolases (orange), before reaching the final NRPS (blue).