



Published in final edited form as:

J Am Chem Soc. 2012 May 23; 134(20): 8625–8632. doi:10.1021/ja301735a.

Bacterial biosynthesis and maturation of the didemnin anticancer agents

Ying Xu[†], Roland D. Kersten[‡], Sang-Jip Nam[‡], Liang Lu[†], Abdulaziz M. Al-Suwailem[§], Huajun Zheng^{||}, William Fenical[‡], Pieter C. Dorrestein^{‡,⊥}, Bradley S. Moore^{‡,⊥}, and Pei-Yuan Qian[†]

[†]KAUST Global Collaborative Research, Division of Life Science, School of Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

[‡]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093, USA

[§]The Coastal and Marine Resources Core Lab, Red Sea Research Center, 4700 King Abdullah University of Science and Technology, Thuwal, Makkah 23955-6900, Kingdom of Saudi Arabia

^{||}Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China

[⊥]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92093, USA

Abstract

The antineoplastic agent didemnin B from the Caribbean tunicate *Trididemnum solidum* was the first marine drug to be clinically tested in humans. Because of its limited supply and its complex cyclic depsipeptide structure, considerable challenges were encountered during didemnin B's development that continue to limit aplidine (dehydrodidemnin B), which is currently being evaluated in numerous clinical trials. Herein we show that the didemnins are bacterial products produced by the marine α -proteobacteria *Tistrella mobilis* and *Tistrella bauzanensis* via a unique post-assembly line maturation process. Complete genome sequence analysis of the 6,513,401 bp *T. mobilis* strain KA081020-065 with its five circular replicons revealed the putative didemnin biosynthetic gene cluster (*did*) on the 1,126,962 bp megaplasmid pTM3. The *did* locus encodes a 13-module hybrid nonribosomal peptide synthetase-polyketide synthase enzyme complex organized in a co-linear arrangement for the synthesis of the fatty acylglutamine ester derivatives didemnins X and Y rather than didemnin B as first anticipated. Imaging mass spectrometry of *T. mobilis* bacterial colonies captured the time-dependent extracellular conversion of the didemnin X and Y precursors to didemnin B in support of an unusual post-synthetase activation mechanism. Significantly, the discovery of the didemnin biosynthetic gene cluster may provide a long-term solution to the supply problem that presently hinders this group of marine natural products and pave the way for the genetic engineering of new didemnin congeners.

Correspondence to: Bradley S. Moore; Pei-Yuan Qian.

Corresponding authors boqianpy@ust.hk and bsmoore@ucsd.edu.

Author Contribution Y.X and R.D.K contributed the same to the work.

Supporting Information

Copies of the ¹H NMR, annotation of MS^D of the didemnins and hydrolysis experiment of didemnin X/Y are included. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Introduction

Sessile marine invertebrates such as sponges and tunicates are intimately linked with microbes that are consumed as food and hosted for their protective chemistry. These marine animals are rich sources of structurally diverse collections of unique bioactive molecules that in some cases have been developed into drugs for the benefit of human health.^{1,2} A major challenge in the drug development of marine natural products, however, has historically been the difficulty in supplying sufficient material for pre-clinical and clinical evaluation due to natural limitations of the source organism.³ Marine products are frequently endowed with complex organic structures that complicate their total syntheses in large scale to meet the demands of a drug development program.⁴ Microbes are commonly suggested as the actual producers of many molecules originally isolated from invertebrate tissues, and thus microbial fermentation offers the promise of a renewable supply of important marine drug candidates.⁵ However, very few examples exist in which clinically relevant compounds originally extracted from marine invertebrates are produced by cultured microbes. Herein we show that the didemnin family of anticancer agents, including didemnin B as the first marine natural product clinically tested in humans,⁶ is produced by marine bacteria of the genus *Tistrella* through an unexpected biosynthetic pathway.

Invertebrate-derived natural products, especially those with complex polyketide and nonribosomal peptide structures, are generally proposed to be microbial in origin since the molecular basis governing their biosyntheses resides exclusively in microbes.⁷⁻⁹ Metagenomic studies have revealed the microbial production of noteworthy bioactive natural products from diverse marine invertebrates, including the pederin group of polyketides from sponges,¹⁰ the patellamides and ecteinascidin-743 (Yondelis) from tunicates,^{11,12} and the bryostatin macrolides from bryozoans.¹³ In each case, uncultivated bacterial symbionts are thought to be the natural producers. Free-living marine microbes are often not implicated in the synthesis of marine invertebrate natural products, although a field-collected marine cyanobacterium was shown to harbor the sponge macrolide swinholide A.¹⁴ As a consequence, molecular engineering approaches involving whole pathway expression are envisaged to produce complex molecules from marine invertebrate symbionts in laboratory microbial strains in order to develop a renewable supply of the compound and allow for the rational engineering of new chemical entities to probe structure-function relationships.¹⁵ To date, only the cyanobactin group of ribosomal peptides, which includes the patellamides, has been heterologously produced and engineered.¹⁶ Complex polyketides, nonribosomal peptides and hybrids thereof, on the other hand, present an unmet technical challenge due to their large biosynthetic gene clusters that often exceed 100 kb.

Didemnins (Fig. 1) are cyclic depsipeptides with remarkable antitumor, antiviral and immunosuppressive properties⁶ that were first discovered in the Caribbean tunicate *Trididemnum solidum* in 1981.¹⁷ The most potent analog didemnin B was the first marine natural product to enter clinical trials as an anticancer agent, yet ultimately failed to demonstrate effective antitumor activity while displaying cardiac and neuromuscular toxicities^{18,19} A closely related compound, dehydrodidemnin B (aplidine), was isolated from the Mediterranean tunicate *Aplidium albicans* and shown to be more potent and less toxic despite its minor structural change.^{20,21} Aplidine has since replaced didemnin B in clinical trials and is currently in multiple phase II and III trials for the treatment of various cancers.⁶ Due to their cyclic depsipeptide structures with highly modified amino acid residues, didemnins have long been suspected as microbial products assembled by a hybrid nonribosomal peptide synthetase–polyketide synthase (NRPS–PKS) enzymatic pathway.¹⁵ In this study, we report that the didemnins are not only the products of the α -proteobacterium *Tistrella mobilis*, which was also recently reported by Tsukimoto and co-workers,²² but also of *Tistrella bauzanensis*. More importantly, we show through the

complete genome sequencing of *T. mobilis* and MALDI-imaging MS technique that didemnin B is biosynthesized by an unanticipated rare mechanism involving a post assembly activation of lipoglutamine congeners, didemnins X and Y.

Results and Discussion

Isolation of didemnins from *Tistrella* species

We isolated the α -proteobacterium *T. mobilis* KA081020-065 from the Red Sea and measured a very potent cytotoxicity against HeLa cells. Following a bioassay-guided fractionation of the ethyl acetate extract, we isolated two metabolites that based on high-resolution MS and NMR characterization and comparison with authentic standards, were identical to the cyclic depsipeptides didemnin B (Fig. S1) and nordidemnin B (Fig. S2) originally isolated from the Caribbean tunicate *Trididemnum solidum*.¹⁷ Since this result showed that the didemnins are bacterial products, we examined additional *Tistrella* strains to explore the extent of didemnin chemistry in this poorly characterized genus. To our surprise, all strains tested, including two Pacific Ocean *T. mobilis* isolates and a *T. bauzanensis* strain, also produced didemnin B and nordidemnin B at varying levels (Fig. 2). During the preparation of this manuscript, Tsukimoto and co-workers independently reported their isolation from a *T. mobilis* strain isolated from marine sediments in Japan.²² Based on these preliminary observations, the didemnins appear to be commonly associated with this recently described bacterial genus whose secondary metabolic potential has not yet been reported.²³ To further explore the natural product biosynthetic capacity of *Tistrella* and to deduce the molecular basis governing didemnin assembly, we sequenced the genome of the Red Sea isolate *T. mobilis* KA081020-065.

Genome characteristics of *T. mobilis* KA081020-065

The complete genome sequence of *T. mobilis* strain KA081020-065 totaling 6,513,401 bp was established by 454 pyrosequencing to reveal five replicons comprising a 3,919,492 bp circular chromosome and four circular plasmids ranging in size from 83,885 bp (pTM4) to 1,126,962 bp (pTM3) (Fig. 3). Its genome organization with three megaplasmids greater than 600 kb each is reminiscent of several other α -proteobacteria members such as the rice plant endophyte *Azospirillum* sp. B510 that harbors six plasmids, five of which in excess of 500 kb.²⁴ General features of the *T. mobilis* genome, which has a high average G+C content of 68%, are summarized in Table S1.

Based on the chemical structure of the didemnin depsipeptides that contains the two ketide-extended amino acid residues isostatine (Ist) and α -(α -hydroxyisovaleryl)-propionic acid (Hip),¹⁷ we hypothesized their biosynthesis by a modular hybrid nonribosomal peptide synthetase-polyketide synthase (NRPS-PKS) pathway reminiscent of nostopeptolide biosynthesis.²⁵ Bioinformatic analysis of the *Tistrella* genome revealed just one large NRPS-PKS biosynthetic gene cluster (*did*) located on pTM3 (TMO_c0602–c0611) that was consistent with the proposed biosynthesis of the didemnin family of cyclic depsipeptides. We detected only one other multimodular NRPS-based gene cluster in the genome, a 34,665 bp locus on pTM2 (TMO_b0323–b0325), which rather encodes a putative eight-residue peptide product inconsistent with didemnin assembly (Fig. S3).

Bioinformatic analysis of the didemnin biosynthetic gene cluster reveals an unexpected post-assembly processing mechanism

In silico analysis of the putative didemnin biosynthetic gene cluster (*did*) revealed eight NRPS and two PKS-encoding genes spanning from *didA* to *didJ* (Fig. S4), comprising 13 modules in total (Fig. 4). Based on the enzymatic logic of thiotemplate-mediated assembly line biosynthesis, the projected product of the DidB–J megasynthetase perfectly correlates

with the co-linear synthesis of didemnin B. The DidB adenylation (A) domain A3 is predicted to load pyruvate and reduce it *in cis* via the adjacent ketoreductase (KR) domain to give lactate bound to the DidB thiolation (T) domain in a reaction reminiscent to that proposed in valinomycin assembly.²⁶ Further processing by the monomodular DidC and the trimodular DidD NRPSs are next proposed to assemble the Pro-DMeLeu-Thr-Ile tetrapeptide fragment. The *N*-methylated leucine residue is the lone D -amino acid in the didemnin B molecule, which correlates nicely with the epimerase (E) and methyltransferase (MT) domains that reside in the Leu-specific module 5 of DidD. Malonate extension of the DidD product by the PKS DidE, which lacks an integrated acyltransferase (AT) domain reminiscent of *trans*-AT PKSs,²⁷ yields the ketide-extended β -hydroxy γ -amino acid residue Ist. Analysis of the genes bordering *didA–J* and elsewhere in the *T. mobilis* genome, however, did not reveal a discrete AT from *trans*-AT systems, suggesting that the didemnin polyketide synthase may co-opt the fatty acid synthase AT FabD as reported recently for the biosynthesis of FK228.²⁸

Subsequent activation of 2-oxoisovaleric acid by the DidF A domain, whose sequence is similar to the *N*-terminal HetE A domain (43% identity at amino acid level) that is also responsible for its selection in hectochlorin biosynthesis,²⁹ sets up the incorporation of the α -hydroxy acid 2-hydroxyisovaleric acid and a second round of malonate extension by the DidG PKS. The DidG MT domain putatively incorporates the α -methyl group in the Hip residue. Further chain elongation of the DidG product by the sequential addition of the residues Leu, Pro, and Me₂Tyr via the monomodular DidH–J NRPSs completes the assembly of the linear didemnin B molecule. The *C*-terminal thioesterase (TE) domain of DidJ is hypothesized to offload didemnin B as the cyclic depsipeptide product. In the case of co-produced nordidemnin B, the DidD A7 domain is anticipated to exhibit substrate flexibility toward aliphatic residues to also activate and incorporate Val. [Hysp]didemnin B³⁰ is an additionally characterized didemnin B derivative from *T. mobilis* extract (Fig. S5) which further indicates substrate promiscuity in DidF domain A9 for α -hydroxy acids 2-hydroxyisovaleric acid (didemnin B) and 2-hydroxy-3-methylvaleric acid ([Hysp]didemnin B. While the biosynthetic model described above is clearly consistent with the observed didemnin B and nordidemnin B chemistry of *T. mobilis*, this proposal ignores the DidA bimodular NRPS that should function as the priming synthetase prior to DidB catalysis. Inspection of both DidA A domains revealed their preference for Gln. Furthermore, sequence analysis of the DidA *N*-terminal condensation (C) domain suggested its enzymatic role in the fatty acylation of the first amino acid residue to yield a lipopeptide product (Fig. 4) consistent with other important nonribosomal peptides such as surfactin.³¹ Among the natural didemnin congeners previously identified from the tunicate *T. solidum*, only didemnins X and Y have both Gln residues and a lipid component (Fig. 1).³² Interestingly, although these two didemnin derivatives contain Gln residues, they each have more than two residues as predicted by the bimodular DidA if each module were to function just once in peptide elongation. We thus explored the possibility that the immediate product of the didemnin biosynthetic pathway is not (nor)didemnin B, but rather an acylglutamine ester such as (nor)didemnin X or Y in which the ester side chain facilitates the cellular export of the cyclic depsipeptide from the Gram-negative bacterium for extracellular unmasking of the didemnin B byproduct.

Conversion of didemnins X and Y to didemnin B captured by imaging mass spectrometry

Repeated attempts were unsuccessful to isolate additional didemnin congeners by altering the culture medium in order to evaluate the activation hypothesis. We thus turned to MALDI-TOF MS to analyze the bacterial culture extracts, whereupon we successfully detected low abundance masses consistent with didemnins X and Y in addition to two novel derivatives, nordidemnins X and Y (Fig. S6). We confirmed their structures by MSⁿ analysis

and high resolution Fourier-transform ion cyclotron resonance MS as reported in the Supporting Information (Fig. S7). The production of didemnins X and Y not only suggests that they are the fully elaborated biosynthetic product of the didemnin synthetase but also indicates that the DidA bimodular NRPS functions iteratively to install the three to four Gln residues in didemnins X and Y, respectively (Fig. 4).

In order to explore the biosynthetic connection between these molecules, we next used MALDI-imaging MS^{33,34} to study the temporal and spatial distribution of the didemnin compounds. We interrogated *T. mobilis* colonies over a time course ranging from 6 hours to 6 days (Fig. 5), which strikingly revealed that didemnins X and Y are produced and excreted from the colony during the early growth phase of the bacterium peaking at 24 hours. Concomitant with their disappearance by day 4 was the emergence of didemnin B as the primary didemnin metabolite that steadily increased over time. This time-lag difference between didemnins X and Y and didemnin B was similarly observed in the nordidemnin series (Fig. 5) and is consistent with the proposed model in which the lipoglutamine side chain facilitates the cellular export of the mature didemnin molecule for extracellular processing.

We next set out to determine whether the maturation of didemnins X and Y involved a secreted factor. To this end, we performed a hydrolysis assay of purified didemnin precursors and filtered growth media cultured with and without the bacterium. The assay clearly showed that didemnin X/Y turnover to didemnin B was dependent on the addition of a protein extract from a 1 day *T. mobilis* culture supernatant (Fig. S8). While studies are ongoing to identify the secreted hydrolytic factor involved in the peptide activation mechanism, sequence analysis of the open reading frames adjacent to the *didA–J* operon revealed a number of genes encoding membrane-associated transport proteins (*orf3*, *orf6*) and hydrolytic enzymes (*orf2*, *orf14*) (Fig. S4), further supporting the export-hydrolysis model. Recently an activation mechanism was established in the nonribosomal peptide synthesis of the antibiotic xenocoumacin from the γ -proteobacterium *Xenorhabdus nematophila* in which a precursor molecule is processed by a membrane-bound and D-Asn-specific peptidase.³⁵ Although the characterized pathway in xenocoumacin processing parallels the proposed pathway for didemnin maturation, the *did* cluster does not harbor a homolog of the xenocoumacin ABC transporter-transpeptidase XcnG. In addition, the maturation of xenocoumacin takes place in the periplasmic space rather than the extracellular space for didemnins, which may suggest an alternative mechanism in *T. mobilis*.

Conclusion

The discovery of the bacterial synthesis of the didemnins raises the question about their origin in the Caribbean tunicate *T. solidum* where they were first characterized over 30 years ago by Rinehart and coworkers.¹⁷ Based on the findings reported here, we suspect that the tunicate-derived didemnins are similarly bacterial in origin. A recent metagenomic profile of the tropical Pacific tunicate *Lissoclinum patella* revealed that the distinctive patellamide ribosomal peptides and the antitumor polyketide patellazole are biosynthetically linked to different bacteria within the tunicate microbiome.³⁶ Similarly, the approved drug Yondelis (ET-743) from the tropical tunicate *Ecteinascidia turbinata* was recently shown to be bacterial in origin.¹² With the *T. mobilis* genomic information now in hand, molecular methods could be used to examine didemnin-associated tunicates for the presence of *Tistrella* bacteria or other microbes that carry the *did* genes. The finding that the didemnin biosynthetic gene cluster resides on a plasmid is also intriguing since other bacteria may acquire the metabolic pathway by horizontal gene transfer. The *did*-containing megaplasmid pTM3 contains several essential elements required for conjugation, including a relaxase (TMO_c0026) needed for DNA transfer. The bacterial chromosome additionally contains a

trb operon which encodes the components of a bacterial mating apparatus, (TMO_2187–2198), another important component in Gram-negative bacterial conjugation.³⁷

With a bacterial production of the didemnins now in place, the supply of these structurally complex and potent anticancer agents may no longer be limiting. While clinical trials with didemnin B have been suspended, phase II trials with aplidine (dehydrodidemnin B; plitidepsin) from the Mediterranean tunicate *Aplidium albicans* are actively underway.³⁸ Presently, multi-step total synthesis is employed to produce aplidine for the clinic due to its limited natural supply.^{6, 39} Biotechnology now may offer an alternative solution. The *T. mobilis* DidB tetradomain gene product is putatively responsible for the selection and attachment of pyruvate, its ketoreduction to lactate, and its ensuing condensation with the DidA fatty acylglutamine residue (Fig. 4). Gene inactivation of the *didB* KR domain would maintain the oxidation state of the pyruvate unit and thereby prevent the esterification reaction, effectively switching the didemnin B biosynthetic pathway to directly produce aplidine. While seemingly straightforward once a genetics system is established in *Tristella*, production levels may be significantly impacted since the activation and export mechanism that governs didemnin B synthesis from the lipopeptides didemnins X and Y may no longer be effective in such a mutant for aplidine. We are actively pursuing answers to the detailed mechanism of didemnin B biosynthesis and whether we can reengineer *T. mobilis* to directly produce aplidine. A provisional patent has also been filed on the didemnin biosynthetic gene cluster (U.S. provisional application number 61/537,416).

Experimental Section

Strain, Fermentation and Isolation of didemnin compounds

Tristella mobilis KA081020-065 was isolated from seawater collected from the Red Sea during a 2009 research cruise. Its crude extract showed remarkable cytotoxicity on HeLa cells in a bioactive compound screening of Red Sea bacteria. *T. mobilis* KA081020-065 was grown in GYP medium [10 g glucose, 4 g yeast extract, 2 g peptone, and 17 g sea salts per l deionized water] using 50-liter stirred fermenters at 25 °C. After 72 hours, ethyl acetate was added to the culture to extract the metabolites. The crude extract was fractionated by reversed phase C18 liquid chromatography and eluted with increasing amounts of methanol in water. The active fraction was further purified by semi-preparative reversed phase (RP) HPLC with 63% MeCN in water to give didemnin B and nordidemnin B at 0.2 mg l⁻¹ and 0.1 mg l⁻¹, respectively. For MSⁿ and FTMS analysis, crude extracts were separated by RP-HPLC (10-100% MeCN in 0.1% trifluoroacetic acid (TFA), 40 min gradient). For hydrolysis assay, didemnin X and Y were purified from crude extract by Sephadex LH20 (GE Healthcare) gel filtration with MeOH as a mobile phase and RP-HPLC (25-75% MeCN in 0.1% TFA, 60 min gradient).

MS sample preparation and MS analysis of didemnins

For MALDI-TOF MS analysis, extract or peptide samples were dissolved in MeOH, mixed 1:1 with saturated Universal MALDI Matrix and spotted on a Bruker MSP96 anchor plate. MALDI-TOF MS analysis was performed on a Microflex Bruker Daltonics mass spectrometer outfitted with Compass 1.3 software suite. For MSⁿ and FTMS analysis, HPLC-purified samples were redissolved in 50 µl MeOH/water/formic acid (50:49:1) and injected by a nanomate-electrospray ionization robot (Advion) for consecutive electrospray into the MS inlet of a LTQ 6.4T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (Thermo Finnigan). MS and MSⁿ data were acquired in the positive ion mode. FTMS data were acquired in 400–2,000 m/z scans. Selected peptide mass signals were manually isolated and fragmented by CID. MSⁿ data were collected either in ion trap or FT

detection mode. All data were analyzed using QualBrowser, which is part of the Xcalibur LTQ-FT software package (Thermo Fisher).

Genome sequencing, annotation, and analysis

We used a massively parallel pyrosequencing technology (Roche 454 GS FLX)⁴⁰ to sequence the *T. mobilis* KA081020-065 genome. We obtained 315,496 reads with an average length of 334 bp, which provided a 16.2-fold coverage of the 6.4 Mb genome. From these reads, we assembled 112 contigs (>500 bp) using the Newbler software of the 454 suite package. To enhance sequence quality and to construct scaffolds, we sequenced a paired-end Illumina library of an additional 2,625,640 bp with average length of 115 bp, which were then mapped to the genome sequence. All contig relationships within scaffolds were then validated by PCR, while the relationships among scaffolds were determined by multiplex PCR.⁴¹ Gaps were filled by sequencing PCR products. The final sequence assembly was carried out using the phred/phrap/consed package (<http://www.genome.washington.edu>) in which we resequenced the low sequence quality regions. The final error rate of the genome sequence was 0.28 per 100,000 bases. The sequences of the five replicons were deposited in the GenBank database as accession numbers CP003236–CP003240.

Putative protein-coding sequences (CDS) were determined by combining the prediction results of the Glimmer 3.02⁴² and Z-Curve programs.⁴³ Functional annotations of CDS were performed by searching the KEGG and the NCBI non-redundant protein databases. tRNA genes were predicted with tRNAScan-SE (v1.23).⁴⁴ Protein domain predictions and COG assignments⁴⁵ were performed by RPS-BLAST using the NCBI CDD library.⁴⁶

In silico analysis of the didemnin biosynthetic gene cluster and other NRPSs in *T. mobilis* KA081020-065

The putative roles of the proteins in the didemnin gene cluster were assigned using protein–protein BLAST and Pfam analyses. NRPS A domain specificities were predicted using the online program NRPSpredictor2 and antiSMASH.^{47–49}

Time-course study of didemnin production in *T. mobilis* by MALDI-imaging MS

One μ l of fresh overnight *T. mobilis* culture was spotted onto GYP agar plates and allowed to grow for 6 hours to 6 days. Colonies with different growing times were transferred to Bruker MSP 96 anchor plates and analyzed concurrently using the IMS technique as previously described.^{33,50}

Didemnin precursor hydrolysis assay

50 μ g HPLC-purified didemnin X and Y were each dissolved in 5 μ l water and mixed with (a) 45 μ l sterile GYP medium (negative control) or (b) 45 μ l protein filtrate (10 kDa cutoff, washed 2x with GYP medium) of 1 ml of a 1 d *T. mobilis* KA081020-065 liquid culture (28 °C) and incubated for 24 h at 25 °C. Hydrolysis samples were taken and spotted on a MALDI-TOF MS plate for subsequent MS analysis at time points 5 min, 1 h, 12 h and 24 h.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Z. Shao for providing the Pacific Ocean *Tistrella* strains and N. Ziemert, J. Yang and P. Lai for assistance with data analysis. Financial support was provided by King Abdullah University of Science and Technology (SA-

C0040/UK-C0016 to P.Y.Q.), China Ocean Mineral Resources Research and Development Association (DY125-15-T-02 to P.Y.Q.), and the National Institutes of Health (GM085770 to B.S.M., CA044848 to W.F., GM086283 and S10RR029121 to P.C.D.).

References

1. Molinski TF, Dalisay DS, Lievens SL, Saludes JP. *Nat. Rev. Drug. Discov.* 2009; 8:69. [PubMed: 19096380]
2. Gerwick WH, Moore BS. *Chem. Biol.* 2012; 19:85. [PubMed: 22284357]
3. Radjasa OK, Vaske YM, Navarro G, Vervoort HC, Tenney K, Linington RG, Crews P. *Bioorg. Med. Chem.* 2011; 19:6658. [PubMed: 21835627]
4. Florence GJ, Gardner NM, Paterson I. *Nat. Prod. Rep.* 2008; 25:342. [PubMed: 18389141]
5. Piel J. *Curr. Med. Chem.* 2006; 13:39. [PubMed: 16457638]
6. Lee J, Currano JN, Carroll PJ, Joullié MM. *Nat. Prod. Rep.* 2012; 29:404. [PubMed: 22270031]
7. Faulkner DJ. *Nat. Prod. Rep.* 2000; 17:1. [PubMed: 10714897]
8. Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S. *Proc. Natl. Acad. Sci. U.S.A.* 2004; 101:16222. [PubMed: 15520376]
9. Lane AL, Moore BS. *Nat. Prod. Rep.* 2011; 28:411. [PubMed: 21170424]
10. Fisch KM, Gurgui C, Heycke N, van der Sar SA, Anderson SA, Webb VL, Taudien S, Platzer M, Rubio BK, Robinson SJ, Crews P, Piel J. *Nat. Chem. Biol.* 2009; 5:494. [PubMed: 19448639]
11. Donia MS, Hathaway BJ, Sudek S, Haygood MG, Rosovitz MJ, Ravel J, Schmidt EW. *Nat. Chem. Biol.* 2006; 2:729. [PubMed: 17086177]
12. Rath CM, Janto B, Earl J, Ahmed A, Hu FZ, Hiller L, Dahlgren M, Kreft R, Yu F, Wolff JJ, Kweon HK, Christiansen MA, Håkansson K, Williams RM, Ehrlich GD, Sherman DH. *ACS Chem. Biol.* 2011; 6:1244. [PubMed: 21875091]
13. Sudek S, Lopanik NB, Waggoner LE, Hildebrand M, Anderson C, Liu H, Patel A, Sherman DH, Haygood MG. *J. Nat. Prod.* 2007; 70:67. [PubMed: 17253852]
14. Andrianasolo EH, Gross H, Goeger D, Musafija-Girt M, McPhail K, Leal RM, Mooberry SL, Gerwick WH. *Org. Lett.* 2005; 7:1375. [PubMed: 15787510]
15. Salomon CE, Magarvey NA, Sherman DH. *Nat. Prod. Rep.* 2004; 21:105. [PubMed: 15039838]
16. Donia MS, Ravel J, Schmidt EW. *Nat. Chem. Biol.* 2008; 4:341. [PubMed: 18425112]
17. Rinehart KL Jr, Gloer JB, Hughes RG Jr, Renis HE, McGovern JP, Swynenberg EB, Stringfellow DA, Kuentzel SL. *Science.* 1981; 212:933. [PubMed: 7233187]
18. Shin DM, Holoye PY, Murphy WK, Forman A, Papasozomenos SC, Hong WK, Raber M. *Cancer Chemother. Pharmacol.* 1991; 29:145. [PubMed: 1662119]
19. Williamson SK, Wolf MK, Eisenberger MA, O'Rourke M, Brannon W, Crawford ED. *Invest. New Drugs.* 1995; 13:167. [PubMed: 8617581]
20. Ding X, Vera MD, Liang B, Zhao Y, Leonard MS, Joullié MM. *Bioorg. Med. Chem. Lett.* 2001; 11:231. [PubMed: 11206466]
21. Le Tourneau C, Raymond E, Faivre S. *Curr. Pharm. Des.* 2007; 13:3427. [PubMed: 18045196]
22. Tsukimoto M, Nagaoka M, Shishido Y, Fujimoto J, Nishisaka F, Matsumoto S, Harunari E, Imada C, Matsuzaki T. *J. Nat. Prod.* 2011; 74:2329. [PubMed: 22035372]
23. Shi BH, Arunpairojana V, Palakawong S, Yokota A. *J. Gen. Appl. Microbiol.* 2002; 48:335. [PubMed: 12682872]
24. Kaneko T, Minamisawa K, Isawa T, Nakatsukasa H, Mitsui H, Kawaharada Y, Nakamura Y, Watanabe A, Kawashima K, Ono A, Shimizu Y, Takahashi C, Minami C, Fujishiro T, Kohara M, Katoh M, Nakazaki N, Nakayama S, Yamada M, Tabata S, Sato S. *DNA. Res.* 2010; 17:37. [PubMed: 20047946]
25. Hoffmann D, Hevel JM, Moore RE, Moore BS. *Gene.* 2003; 311:171. [PubMed: 12853152]
26. Cheng YQ. *ChemBioChem.* 2006; 7:471. [PubMed: 16511823]
27. Piel J. *Nat. Prod. Rep.* 2010; 27:996. [PubMed: 20464003]
28. Wesener SR, Potharla VY, Cheng YQ. *Appl. Environ. Microbiol.* 2011; 77:1501. [PubMed: 21183648]

29. Ramaswamy AV, Sorrels CM, Gerwick WH. *J. Nat. Prod.* 2007; 70:1977. [PubMed: 18001088]
30. Banaigs B, Mansour EA, Bonnard I, Boulanger A, Francisco C. *Tetrahedron.* 1999; 55:9559.
31. Rausch C, Hoof I, Weber T, Wohlleben W, Huson DH. *BMC. Evol. Biol.* 2007; 16:78. [PubMed: 17506888]
32. Sakai R, Stroh JG, Sullins DW, Rinehart KL. *J. Am. Chem. Soc.* 1995; 117:3734.
33. Yang YL, Xu Y, Straight P, Dorrestein PC. *Nat. Chem. Biol.* 2009; 5:885. [PubMed: 19915536]
34. Yang YL, Xu Y, Kersten RD, Liu WT, Meehan M, Moore BS, Bandeira N, Dorrestein PC. *Angew. Chem. Int. Ed. Engl.* 2011; 50:5839. [PubMed: 21574228]
35. Reimer D, Pos KM, Thines M, Grün P, Bode HB. *Nat. Chem. Biol.* 2011; 7:888. [PubMed: 21926994]
36. Donia MS, Fricke WF, Partensky F, Cox J, Elshahawi SI, White JR, Phillippy AM, Schatz MC, Piel J, Haygood MG, Ravel J, Schmidt EW. *Proc. Natl. Acad. Sci. U.S.A.* 2011; 108:E1423. [PubMed: 22123943]
37. Zatyka M, Thomas CM. *FEMS Microbiol. Rev.* 1998; 21:291.
38. Muñoz-Alonso MJ. *Curr. Opin. Investig. Drugs.* 2009; 10:536.
39. Jou G, González I, Albericio F, Lloyd-Williams P, Giralt E. *J. Org. Chem.* 1997; 62:354. [PubMed: 11671410]
40. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. *Nature.* 2005; 437:376. [PubMed: 16056220]
41. Tettelin H, Radune D, Kasif S, Khouri H, Salzberg SL. *Genomics.* 1999; 62:500. [PubMed: 10644449]
42. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. *Nucleic Acids Res.* 1999; 27:4636. [PubMed: 10556321]
43. Guo FB, Ou HY, Zhang CT. *Nucleic Acids Res.* 2003; 31:1780. [PubMed: 12626720]
44. Lowe TM, Eddy SR. *Nucleic Acids Res.* 1997; 25:955. [PubMed: 9023104]
45. Tatusov RL, Galperin MY, Natale DA, Koonin EV. *Nucleic Acids Res.* 2000; 28:33. [PubMed: 10592175]
46. Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH. *Nucleic Acids Res.* 2007; 35:D237. [PubMed: 17135202]
47. Röttig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O. *Nucleic Acids Res.* 2011; 39:W362. [PubMed: 21558170]
48. Rausch C, Weber T, Kohlbacher O, Wohlleben W, Huson DH. *Nucleic Acids Res.* 2005; 33:5799. [PubMed: 16221976]
49. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. *Nucleic Acids Res.* 2011; 39:W339. [PubMed: 21672958]
50. Kersten RD, Yang YL, Xu Y, Cimermancic P, Nam SJ, Fenical W, Fischbach MA, Moore BS, Dorrestein PC. *Nat. Chem. Biol.* 2011; 7:794. [PubMed: 21983601]

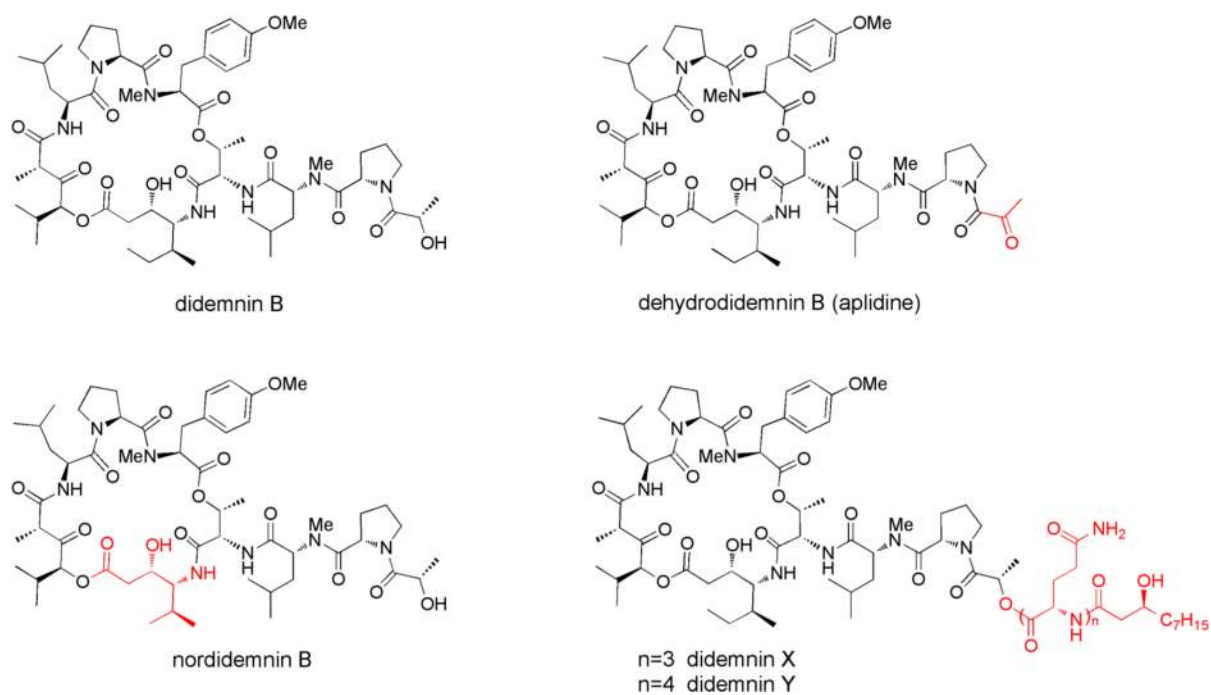


Figure 1.
Structures of didemnin B and other representative didemnin analogs in which structural differences are colored red.

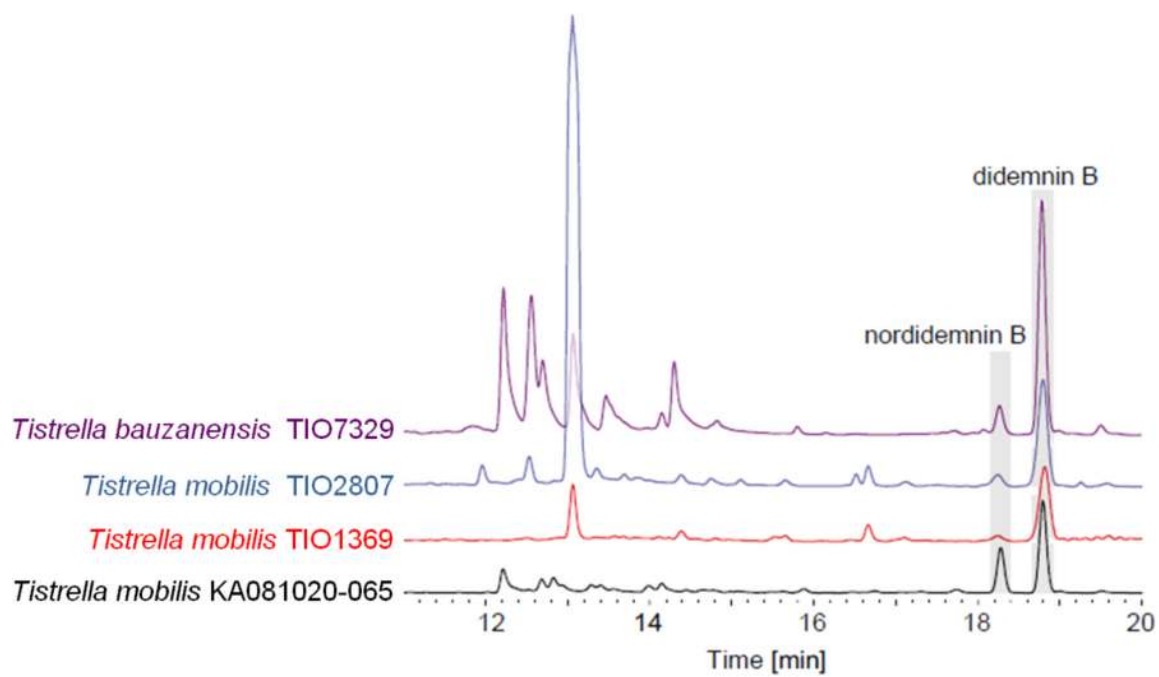


Figure 2.
HPLC profiles of didemnin B and nordidemnin B in extracts of four different *Tistrella* strains.

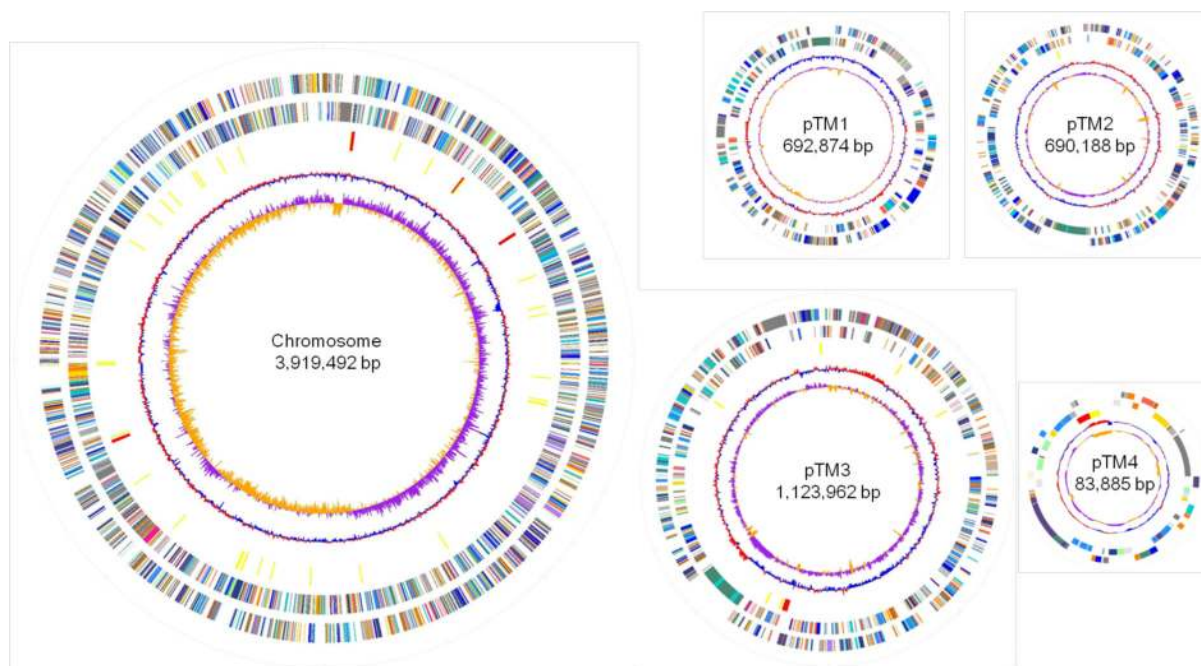


Figure 3.

Genome maps of *T. mobilis* KA081020-065. From the outside inward, the first and second circles show predicted coding regions transcribed on the forward and reverse strand, respectively, colored by functional categories according to COG classification. The third circle shows RNAs (yellow for tRNA and red for rRNA). The fourth and fifth circles show G+C content and GC skew $(G-C)/(G+C)$ calculated using a 5 kb window (purple for values >1 , orange for values <1), respectively.

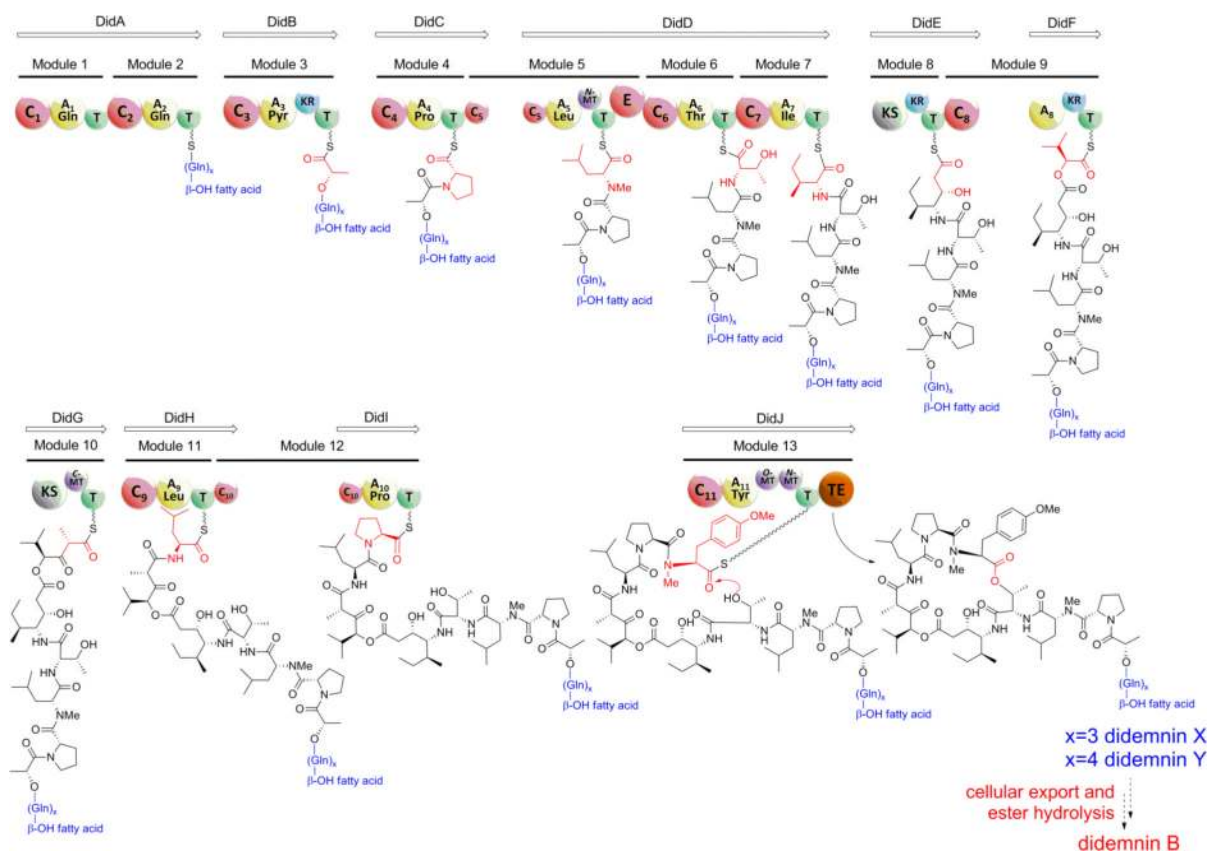


Figure 4. Proposed biosynthetic pathway of didemnin and domain organization of the biosynthetic *did* gene cluster in *T. mobilis* KA081020-065. Domain notation: C, condensation; A, adenylation (substrate included); E, epimerase; KR, ketoreductase; KS, b-ketoacyl synthase; MT, methyltransferase; T, thiolation; TE, thioesterase.

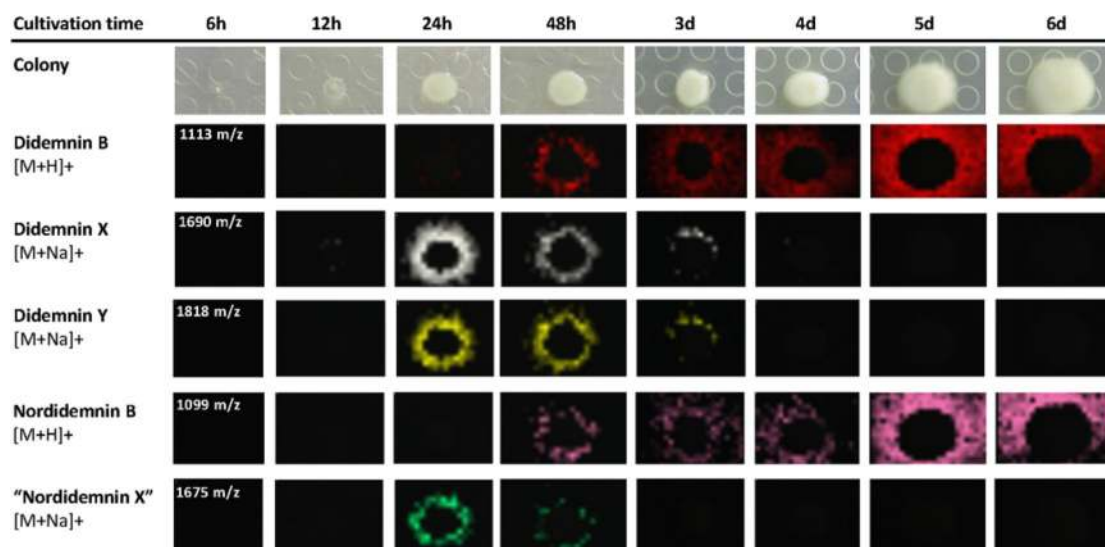


Figure 5.
Time-course production of didemnins by *T. mobilis* KA081020-065 detected by MALDI-imaging mass spectrometry.