

In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin

Graham A. Hudson,^{†,‡} Zhengan Zhang,^{†,‡} Jonathan I. Tietz,[†] Douglas A. Mitchell,^{*,†,‡} and Wilfred A. van der Donk^{*,†,‡}

[†]Department of Chemistry, University of Illinois at Urbana–Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, United States

[‡]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, 1206 West Gregory Drive, Urbana, Illinois 61801, United States

S Supporting Information

ABSTRACT: Thiopeptides are potent antibiotics that inhibit protein synthesis. They are made by a remarkable post-translational modification process that transforms a linear peptide into a polycyclic structure. We present here the in vitro biosynthesis of the core scaffold of thiomuracin catalyzed by six proteins. We show that cyclodehydration precedes dehydration, and that dehydration is catalyzed by two proteins in a tRNA^{Glu}-dependent manner. The enzyme that generates the pyridine core from two dehydroalanines ejects the leader peptide as a C-terminal carboxamide. Mutagenesis studies of the enzyme TbtD identified important residues for a formal [4+2] cycloaddition process. The core structure of thiomuracin exhibits similar antimicrobial activity to other known congeners, illustrating that in vitro biosynthesis is a viable route to potent antibiotics that can be explored for the rapid and renewable generation of analogues.

The genome sequencing programs of the past decade have revealed that ribosomally synthesized and post-translationally modified peptides (RiPPs) form a large class of natural products that are produced in all domains of life.¹ These compounds are made from a linear precursor peptide that is composed of a leader peptide (LP), which is removed in the late stages of maturation, and a core peptide, which is converted into the RiPP. The LP carries important recognition elements for key biosynthetic enzymes, which in turn allows these enzymes a high degree of tolerance with respect to sequence variation in the core peptide.² As a result, these pathways are highly evolvable.³ Recent studies have illustrated that LP recognition in diverse classes of RiPPs is achieved in a common manner, relying on structurally conserved domains that are present in otherwise highly divergent biosynthetic enzymes.^{4–6} This finding offers the tantalizing opportunity to engineer hybrid RiPPs, not only within RiPP classes^{7–9} but also between different RiPP classes. Indeed, Nature already uses this hybrid strategy in the biosynthesis of the thiopeptides. These compounds are made by utilizing Ser/Thr dehydratases that produce dehydro amino acids alongside cyclodehydratases that convert Ser, Thr, and Cys residues to the corresponding azolines.^{10–13} The presence of both types of post-translational modifications facilitated the evolution of new chemistry,

including a formal [4+2] cycloaddition that forms the six-membered, nitrogen-containing heterocycles that are characteristic of thiopeptides.¹⁴ This Communication presents the in vitro reconstitution of the activities of six proteins that together biosynthesize the core structure of thiomuracin via 22 chemical transformations (Figure 1A), which provides insights into the order of events and requirements for catalysis.

The thiopeptide antibiotics block bacterial protein translation by interacting with either the 50S ribosomal subunit or the

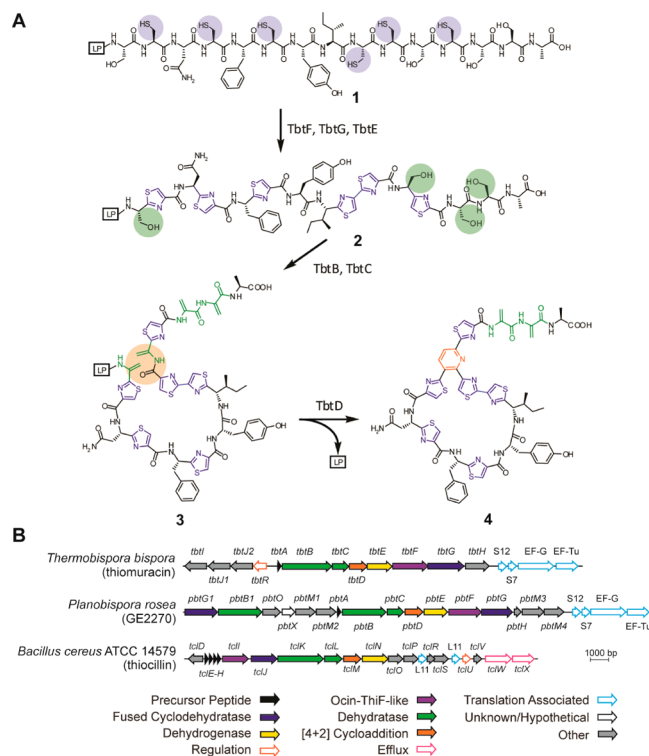


Figure 1. (A) Biosynthetic route to the thiomuracin core scaffold 4. (B) Gene clusters for the biosynthesis of thiomuracin, GE2270A, and thiocillin. Sequence of the His₆-tagged leader peptide: PHHHHH-HHSQVDLNDLPMDVFEADSGVAVESLTAGHGMTEVGA.

Received: September 29, 2015

Published: December 17, 2015

elongation factor Tu.¹⁴ Thiopeptides exhibit potent activity toward multi-drug-resistant pathogenic bacteria. Unfortunately, poor physicochemical properties have plagued this otherwise promising class of antibiotics; thus far, only two agents are in commercial use, with both restricted to animal usage (thiostrepton for skin infections; nosiheptide for growth promotion).¹⁵ Impressive efforts in the semi-synthetic derivatization of GE2270A led to LFF571,¹⁶ which recently completed a phase II trial for moderate *Clostridium difficile* infections.^{17,18} This trial underscores the opportunities for improving the pharmacological properties of thiopeptides.

Recent advances in our mechanistic understanding of RiPP cyclodehydratases^{19–22} and dehydratases^{4,23} prompted us to attempt the in vitro biosynthetic reconstitution of a thiopeptide. This class of natural products leverages both RiPP biosynthetic platforms in addition to a formal [4+2] cycloaddition enzyme that converts two dehydroalanine (Dha) moieties to the class-defining central pyridine [or (dehydro)piperidine].^{24,25} The thiopeptide family is comprised of >100 members, with many compounds adorned with a second macrocycle (e.g., thiostrepton).¹⁴ To begin this effort, we restricted our thiopeptide in vitro reconstitution targets to mono-macrocytic compounds whose gene clusters and bioactivities were firmly established. Our top candidates thus became GE2270A and thiomuracin (Figure 1).¹² Although mono-macrocytic thiopeptides are more “minimalistic”, they harbor formidable chemical complexity and biological activity. *Planobispora rosea* ATCC53733 and *Nonomuraea* sp. Bp3714-39 are the canonical producers of GE2270A and thiomuracin, respectively.¹² We did not have access to a thiomuracin-producing strain of *Nonomuraea*, but genome-mining efforts predicted that *Thermobispora bispora* DSM 43833 would produce an identical molecule owing to the presence of a nearly identical biosynthetic gene cluster (Figure 1B). Derived from a thermophilic organism, the *T. bispora* proteins were envisioned to have the desired stability for efficient heterologous expression in *Escherichia coli*.

All genes encoding the biosynthetic proteins from *T. bispora* (*tbt*) were codon optimized for *E. coli* expression (Table S1). Based on previous reports, genes for F-protein-dependent cyclodehydratase (*tbtF*, *tbtG*), thiazoline dehydrogenase (*tbtE*), and the [4+2] cycloaddition enzyme (*tbtD*) were subcloned into a previously described plasmid that fuses maltose-binding protein (MBP) to the N-terminus of the protein of interest.²² In contrast, genes encoding the precursor peptide (*tbtA*) and the split LanB-type dehydratase⁴ (*tbtB*, *tbtC*) were expressed as N-terminally His₆-tagged proteins. The purity of all proteins used in the study was assessed visually by Coomassie-stained SDS-PAGE analysis (Figure S1).

The precursor peptide TbtA was initially expressed and purified as an N-terminally His₆-tagged fusion protein (1) (Figure 2A). Reaction of TbtA with purified TbtB/C under conditions previously reported for the NisB dehydratase^{4,23} did not result in any chemical transformation (Figure S2). In contrast, reaction of 1 with the cyclodehydratase TbtG, the dehydrogenase TbtE, and TbtF, an orthologue of HcaF that was recently shown to function as an auxiliary LP binding protein,²² resulted in a mass shift of –120 Da (Figure 2B). This observation is consistent with successful cyclodehydration and dehydrogenation of the six cysteines in the core peptide to the corresponding thiazoles. When TbtE was omitted, a mass consistent with six thiazoline modifications was observed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS, Figure S3). Further, when

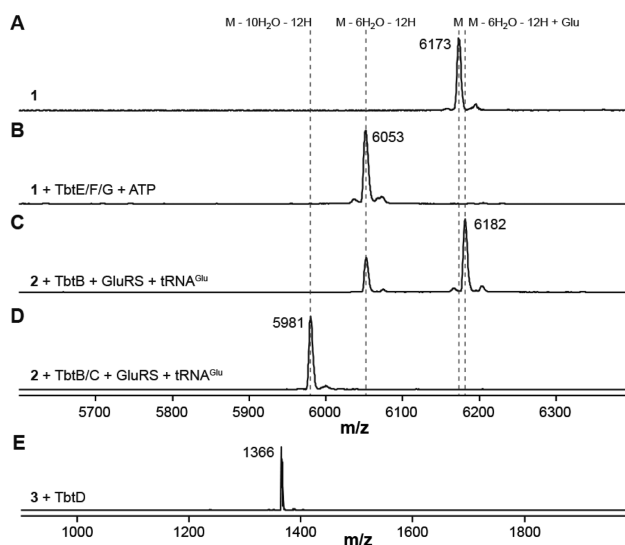


Figure 2. MALDI-TOF-MS of (A) His₆-TbtA (1), calcd *m/z* 6173; (B) hexazole 2, calcd *m/z* 6053; (C) glutamylated 2, calcd *m/z* 6182; (D) tetra-dehydrated hexazole 3, calcd *m/z* 5981; and (E) thiomuracin GZ 4, calcd *m/z* 1366. For the reactions shown in panels C and D, ATP, Mg²⁺, *E. coli* GluRS, and *T. bispora* tRNA^{Glu}(CUC) were present.

TbtF or ATP was omitted from the reaction, no product formation was detected.

TbtB is an orthologue of the glutamylation domain of the NisB dehydratase that uses glutamyl-tRNA to glutamylate Ser residues.⁴ TbtC is a member of the SpaB_C family of proteins (PF14028) and has sequence homology to the second domain of NisB that eliminates the glutamate from the glutamylated Ser to install Dha. Reaction of the purified hexazole-containing intermediate 2 with TbtB in the presence of Glu, ATP, *E. coli* tRNA^{Glu}, and *E. coli* glutamate tRNA transferase (GluRS) did not result in any change in mass (Figure S4). We reasoned that sequence differences between *E. coli* and *T. bispora* tRNA^{Glu} might contribute to the lack of activity. Hence, the two tRNA^{Glu} iso-acceptors encoded in the *T. bispora* genome were prepared by in vitro transcription. Reaction of 2 with Glu, ATP, *T. bispora* tRNA^{Glu}(CUC), and *E. coli* GluRS resulted in glutamylation of 2 (Figure 2C). When TbtC was added to an otherwise identical reaction, four-fold dehydration of 2 was observed (Figure 2D). Use of *T. bispora* tRNA^{Glu}(UUC) resulted in only partial dehydration (Figure S5); this result may indicate either that TbtB preferentially uses *T. bispora* tRNA^{Glu}(CUC) or that *E. coli* GluRS is less efficient in charging *T. bispora* tRNA^{Glu}(UUC). Using the same conditions with 1 did not result in an observable reaction (Figure S2), demonstrating for the first time that dehydration activity requires prior cyclodehydration. Further, all assay components were required for activity (Figure S5). These experiments establish the first activity of a split LanB and extend the use of Glu-tRNA^{Glu} to dehydration during thiopeptide biosynthesis.

Treatment of the hexazole-containing, tetra-dehydrated His₆-TbtA (3) with TbtD resulted in the consumption of 3 and appearance of two new masses consistent with the excised LP and the desired macrocycle (Figures 2E and S6). The identity of the LP was confirmed by high-resolution electrospray ionization tandem MS (HR-ESI-MS/MS, Figure 3). The data indicated that the LP terminates in a carboxamide rather than a carboxylic acid, providing the first evidence that the cycloaddition enzyme uses elimination chemistry, rather than

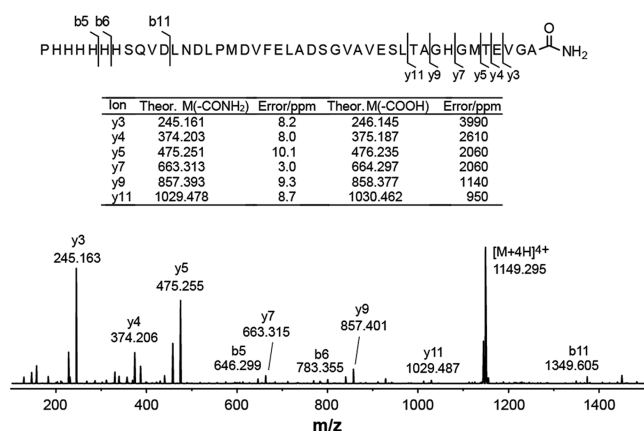


Figure 3. ESI-MS/MS of the His₆-TbtA leader peptide (*m/z* 1149) generated by macrocyclization.

hydrolysis, to release the LP. The generation of the macrocyclic core scaffold of thiomuracin by the TbtD-orthologue TclM using a semi-synthetic substrate was also recently reported.²⁶ Interestingly, like TbtC, TbtD (and TclM) is a member of the SpaB_C protein family, suggesting that elimination reactions may be common in PF14028.⁴ Removal of the N-terminal 31 (of 34) amino acids of the LP from **2** by treatment with endoproteinase GluC resulted in a peptide that was no longer a substrate for TbtD (Figure S6). In NisB, where the glutamylation enzyme is covalently linked to the SpaB_C elimination domain, LP binding occurs in the glutamylation domain.^{4,6} To probe whether TbtB, which contains this binding domain, would accelerate catalysis by TbtD, TbtB was added to the cycloaddition reaction. Analysis of macrocycle formation suggested that TbtB afforded no rate enhancement (Figure S7). The same was observed upon addition of TbtF, which also contains a LP-binding domain.²² Thus, it appears that TbtD independently engages the LP by an unknown strategy.

To obtain the quantities of **4** required for structure confirmation, the entire in vitro process was optimized and scaled up using MBP-TbtA as the starting material (see Supporting Information). The purity of **4** was assessed by HPLC and ESI-MS (Figure S8), the molecular formula of **4** was deduced by HR-ESI-MS (Figure S9), and its structure was confirmed by NMR spectroscopy (Figures S10–S17 and Table S3). The overall chemical shift values were highly consistent with those of the closely related thiomuracin series reported by Morris et al.¹² Correlations obtained via ¹H–¹H COSY, ¹H–¹H TOCSY, ¹H–¹H NOESY, ¹H–¹³C HSQC, and ¹H–¹³C HMBC experiments (Figure 4) allowed assignment of all ¹H and ¹³C resonances in the molecule, with the exception of the carboxylic acid proton, which was not observed. All ¹³C resonances were indirectly detected through the HSQC and HMBC experiments. A combination of NOESY and HMBC correlations allowed positional assignment of each of the six thiazole and two Dha moieties. Importantly, the presence of the pyridine moiety resulting from the formal [4+2] cycloaddition was evidenced by a pair of characteristic doublets at 8.38 and 8.51 ppm and the accompanying HSQC and HMBC correlations; the thiazole protons were likewise observed as sharp singlets at 7.55, 8.13, 8.17, 8.27, 8.55, and 8.67 ppm. Due to the high similarity of chemical shifts and coupling constants of **4** to thiomuracin D, as well as the biosynthetic logic of the enzymes employed, all amino acids were assigned L-stereochemistry.

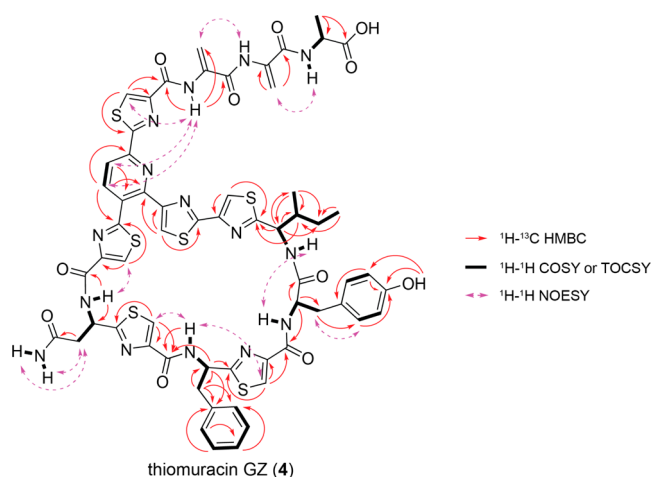


Figure 4. NMR correlations used to assign the structure of **4**. For all spectroscopic data, see Figures S10–S17 and Table S3.

To assess the bioactivity of **4**, we determined the minimal inhibitory concentration (MIC) toward a panel of Gram-positive, Gram-negative, and fungal species by microbroth dilution (Table 1). While **4** was inactive against the tested

Table 1. Bioactivity of Thiomuracin GZ

organism	MIC ^a (μg/mL)	MIC ^a (μM)
<i>Staphylococcus aureus</i> USA300	0.25	0.18
<i>Enterococcus faecium</i> U503	0.063	0.046
<i>Bacillus anthracis</i> str. Sterne	2	1.5
<i>Escherichia coli</i> MC4100	>64	>47
<i>Pseudomonas aeruginosa</i> PA01	>64	>47
<i>Mycobacterium smegmatis</i> B-14616	>64	>47
<i>Aspergillus niger</i>	>128	>94
<i>Saccharomyces cerevisiae</i>	>128	>94
<i>Talaromyces stipitatus</i>	>128	>94

^aMIC, minimum inhibitory concentration.

Gram-negative bacteria and fungi, growth inhibition was found for all Gram-positive bacteria in the panel. The most potent activity was toward vancomycin-resistant *Enterococcus faecium* U503 (VRE), followed by methicillin-resistant *Staphylococcus aureus* USA300 (MRSA). These drug-resistant bacteria are increasingly problematic human pathogens. Despite **4** lacking other post-translational modifications including C-methylation of thiazole, β-hydroxylation of Phe, and oxidation/cyclization of Ile, the bioactivity reported here is on par with that of fully modified thiomuracin A, which has MICs of 0.25 and 0.5 μg/mL against VRE and MRSA, respectively.^{12,27} We term the structure of **4** thiomuracin GZ.

GE2270A has a macrocyclic structure similar to that of thiomuracin (Figure S18), differing in three of the four amino acids that are present in the macrocycle.¹² We therefore wondered whether the TbtD orthologue in the GE2270A biosynthetic gene cluster (PbtD) would also be able to catalyze the non-cognate [4+2] cycloaddition of **3**. Reaction of PbtD with **3** indeed resulted in the clean formation of **4** with concomitant elimination of the LP (Figure S18). This observation is promising with respect to thiopeptide engineering for structure–activity relationship efforts, because it illustrates that the [4+2] enzyme is tolerant of substantial changes in the amino acid sequence of the macrocycle (Phe,

Tyr, Ile in thiomuracin are Val, Gly, Phe in GE2270A, respectively).

To investigate the importance of conserved residues within the [4+2] cycloaddition clade of the SpaB_C superfamily, a multiple sequence alignment was generated, and 23 residues were replaced with alanine (Figure S19). The activity of these mutant enzymes was qualitatively assessed by an endpoint MALDI-TOF-MS assay under two reaction conditions. In the first, high concentrations of substrate (100 μ M) and enzyme (5 μ M) were used to probe for any residual enzymatic activity. Only two mutant proteins (R332A and Y319A) were diminished in activity under these conditions (Table S4). Lowering the reactant concentrations to 15 μ M substrate and 1 μ M enzyme implicated an additional set of residues (W47, E105, H191, H290, and S287) as possibly involved in substrate binding or catalysis. The relative activities of these mutants is provided in Table S5.

In summary, this work presents the first in vitro biosynthesis of a thiopeptide core scaffold. We establish the functions of the split LanB proteins, demonstrate that the use of Glu-tRNA^{Glu} and the auxiliary LP docking domain is not confined to lanthipeptide and heterocycloanthracin biosynthesis, respectively, and show that thiomuracin GZ harbors potent antibacterial activity, despite the lack of ancillary tailoring. A compelling application of the data and methods presented herein would be to generate unnatural thiopeptide analogues for improvement of pharmacological properties.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b10194.

Experimental details; supporting Tables S1–S5 and Figures S1–S19 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*dougasm@illinois.edu

*vddonk@illinois.edu

Author Contributions

[†]G.A.H. and Z.Z. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health (NIH Director's New Innovator Award Program DP2 OD008463 to D.A.M.; R01 GM058822 to W.A.V.), the David and Lucile Packard Fellowship for Science and Engineering (to D.A.M.), and the Robert C. and Carolyn J. Springborn Endowment (to J.I.T.). The Bruker UltrafleXtreme MALDI TOF/TOF mass spectrometer was purchased in part with a grant from NIH (S10 RR027109 A). The authors thank Dr. Ben Ramirez (UIC) for assistance in collecting the high-field NMR data. The 900 MHz NMR spectrometer used was purchased by funds provided by GM068944.

■ REFERENCES

(1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.;

Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süßmuth, R. E.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. *Nat. Prod. Rep.* **2013**, *30*, 108.

(2) Oman, T. J.; van der Donk, W. A. *Nat. Chem. Biol.* **2010**, *6*, 9.

(3) Melby, J. O.; Nard, N. J.; Mitchell, D. A. *Curr. Opin. Chem. Biol.* **2011**, *15*, 369.

(4) Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van der Donk, W. A.; Nair, S. K. *Nature* **2015**, *517*, 509.

(5) Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W. E.; Jaspars, M.; Naismith, J. H. *Nat. Chem. Biol.* **2015**, *11*, 558.

(6) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. *Nat. Chem. Biol.* **2015**, *11*, 564.

(7) van Heel, A. J.; Mu, D.; Montalban-Lopez, M.; Hendriks, D.; Kuipers, O. P. *ACS Synth. Biol.* **2013**, *2*, 397.

(8) Houssen, W. E.; Bent, A. F.; McEwan, A. R.; Pieiller, N.; Tabudravu, J.; Koehnke, J.; Mann, G.; Adaba, R. I.; Thomas, L.; Hawas, U. W.; Liu, H.; Schwarz-Linek, U.; Smith, M. C.; Naismith, J. H.; Jaspars, M. *Angew. Chem., Int. Ed.* **2014**, *53*, 14171.

(9) Sardar, D.; Lin, Z.; Schmidt, E. W. *Chem. Biol.* **2015**, *22*, 907.

(10) Kelly, W. L.; Pan, L.; Li, C. *J. Am. Chem. Soc.* **2009**, *131*, 4327.

(11) Wieland Brown, L. C.; Acker, M. G.; Clardy, J.; Walsh, C. T.; Fischbach, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 2549.

(12) Morris, R. P.; Leeds, J. A.; Naegeli, H. U.; Oberer, L.; Memmert, K.; Weber, E.; Lamarche, M. J.; Parker, C. N.; Burren, N.; Esterow, S.; Hein, A. E.; Schmitt, E. K.; Krastel, P. *J. Am. Chem. Soc.* **2009**, *131*, 5946.

(13) Liao, R.; Duan, L.; Lei, C.; Pan, H.; Ding, Y.; Zhang, Q.; Chen, D.; Shen, B.; Yu, Y.; Liu, W. *Chem. Biol.* **2009**, *16*, 141.

(14) Bagley, M. C.; Dale, J. W.; Merritt, E. A.; Xiong, X. *Chem. Rev.* **2005**, *105*, 685.

(15) Just-Baringo, X.; Albericio, F.; Alvarez, M. *Mar. Drugs* **2014**, *12*, 317.

(16) LaMarche, M. J.; Leeds, J. A.; Amaral, A.; Brewer, J. T.; Bushell, S. M.; Deng, G.; Dewhurst, J. M.; Ding, J.; Dzik-Fox, J.; Gamber, G.; Jain, A.; Lee, K.; Lee, L.; Lister, T.; McKenney, D.; Mullin, S.; Osborne, C.; Palestrant, D.; Patane, M. A.; Rann, E. M.; Sachdeva, M.; Shao, J.; Tiamfook, S.; Trzasko, A.; Whitehead, L.; Yifru, A.; Yu, D.; Yan, W.; Zhu, Q. *J. Med. Chem.* **2012**, *55*, 2376.

(17) Mullane, K.; Lee, C.; Bressler, A.; Buitrago, M.; Weiss, K.; Dabovic, K.; Praestgaard, J.; Leeds, J. A.; Blais, J.; Pertel, P. *Antimicrob. Agents Chemother.* **2015**, *59*, 1435.

(18) Bhansali, S. G.; Mullane, K.; Ting, L. S.; Leeds, J. A.; Dabovic, K.; Praestgaard, J.; Pertel, P. *Antimicrob. Agents Chemother.* **2015**, *59*, 1441.

(19) Dunbar, K. L.; Melby, J. O.; Mitchell, D. A. *Nat. Chem. Biol.* **2012**, *8*, 569.

(20) Dunbar, K. L.; Mitchell, D. A. *J. Am. Chem. Soc.* **2013**, *135*, 8692.

(21) Dunbar, K. L.; Chekan, J. R.; Cox, C. L.; Burkhart, B. J.; Nair, S. K.; Mitchell, D. A. *Nat. Chem. Biol.* **2014**, *10*, 823.

(22) Dunbar, K. L.; Tietz, J. I.; Cox, C. L.; Burkhart, B. J.; Mitchell, D. A. *J. Am. Chem. Soc.* **2015**, *137*, 7672.

(23) Garg, N.; Salazar-Ocampo, L. M.; van der Donk, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 7258.

(24) Walsh, C. T.; Malcolmson, S. J.; Young, T. S. *ACS Chem. Biol.* **2012**, *7*, 429.

(25) Li, C.; Kelly, W. L. *Nat. Prod. Rep.* **2010**, *27*, 153.

(26) Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A. *J. Am. Chem. Soc.* **2015**, *137*, 3494.

(27) LaMarche, M. J.; Leeds, J. A.; Dzik-Fox, J.; Gangl, E.; Krastel, P.; Neckermann, G.; Palestrant, D.; Patane, M. A.; Rann, E. M.; Tiamfook, S.; Yu, D. *J. Med. Chem.* **2012**, *55*, 6934.