JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Identification of Pyridine Synthase Recognition Sequences Allows a Modular Solid-Phase Route to Thiopeptide Variants

Walter J. Wever,[†] Jonathan W. Bogart,[†] and Albert A. Bowers*

Division of Chemical Biology and Medicinal Chemistry, University of North Carolina at Chapel Hill, Eshelman School of Pharmacy, Chapel Hill, North Carolina 27599, United States

Supporting Information

ABSTRACT: Thiopeptides are structurally complex, bioactive natural products derived from ribosomally synthesized and post-translationally modified peptides. A remarkable set of enzymes were recently revealed to catalyze the formation of the core trithiazolylpyridine of thiopeptides via a formal [4 + 2] cycloaddition. These pyridine synthases typically act late in thiopeptide biosynthesis to affect macrocyclization and cleavage of the N-terminal leader peptide, making them potentially useful biocatalysts for preparation of new thiopeptide variants. Herein we investigate the leader peptide requirements for TclM from thiocillin biosynthesis in *Bacillus cereus* ATCC 14579. Through a series of truncations, we define a minimum recognition sequence (RS) that is necessary and sufficient for TclM activity. This RS can be readily synthesized and ligated to linear thiopeptide cores prepared via solid-phase peptide synthesis (SPPS), giving an efficient and modular route to thiopeptide variants. We exploit this strategy to define C-terminal core peptide requirements and explore the differences in promiscuity of two pyridine synthases, TclM and TbtD, ultimately examining their ability to access new structural variants.

Thiazolyl peptides, or thiopeptides (e.g., micrococcin P1 (**1**) and thiomuracin GZ (**2**); Figure 1a), are an intriguing group of ribosomally encoded and post-translationally modified peptide natural products (RiPPs) that have undergone intensive investigation since they were first discovered nearly 70 years ago.[1] Part of the reason for the strong interest in thiopeptides is their pronounced biological activity. These compounds have long been known to have strong activity against drug-resistant bacteria, such as methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *enterococci* (VRE), and penicillin-resistant *Streptococcus pneumoniae* (PRSP).[1] This activity is probably best exemplified by the semisynthetic thiopeptide analogue LFF571 (an orthologue of **2**), which was developed and carried through phase-II clinical trials by Novartis Group.[2] LFF571 has nanomolar activity against bacterial EF-Tu and has progressed for treatment of intestinal infections caused by *Clostridium difficile*. More recently, several thiopeptides have been shown to exhibit antiplasmodial and anticancer activity.[3,4]

In view of the potential therapeutic importance of thiopeptides, a number of efforts have focused on preparation of analogues. Two major routes have been explored: chemical synthesis and biosynthetic pathway engineering. Innovative

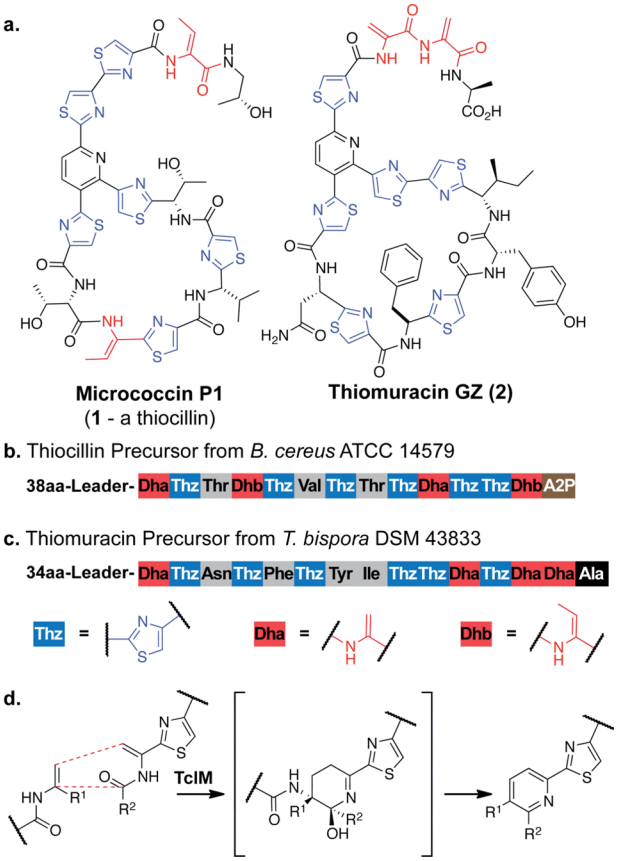

Figure 1. Structure and biosynthetic macrocyclization in thiopeptide biosynthesis. (a) Thiocillin (left) and thiomuracin D (right) and (b, c) sequences of respective core peptides that undergo an enzyme-catalyzed formal [4 + 2] cycloaddition (d) to make their central pyridine rings (A2P = amino-2-propanol).

chemistries developed by Nicolaou, Ciufolini, Moody, and Bach have led to total syntheses of several thiopeptides.[5−10] In principle, synthesis provides access to thiopeptide analogues with expanded chemical diversity by the introduction of non-proteinogenic building blocks. However, the synthetic challenge of the central trisubstituted pyridine, which typically must be installed early, limits the flexibility of many synthetic routes. In contrast, biosynthetic pathway engineering has proven tremen-

Received: May 25, 2016
Published: August 30, 2016

DOI: 10.1021/jacs.6b05389
J. Am. Chem. Soc. 2016, 138, 13461−13464

dously versatile. Because thiopeptides are made from gene-encoded peptides (Figure 1b,c), gene replacement combined with site-directed mutagenesis has given facile access to substantially modified thiopeptide variants.[3,4,11] Still, these methods are largely limited by the 20 canonical amino acids as well as the promiscuity of multiple biosynthetic enzymes, several of which are capable of modifying similar residues.[12]

We recently described the in vitro characterization of a single enzyme, TclM, that catalyzes a long-hypothesized formal [4 + 2] cycloaddition to form the trisubstituted pyridine in the core of the thiocillins (e.g., **1**) from *Bacillus cereus* ATCC 14579 (Figure 1d).[13] This cycloaddition is nonspontaneous: it can be accomplished on a synthetic scale on simple substrates only if they are predisposed by forced tautomerization of the amide to its imidate isostere and heated in a microwave. TclM catalyzes the mild late-stage cyclization on a substrate that already displays a heavily modified core region, decorated with thiazoles (Thzs) and dehydroalanines (Dhas), and a 38-residue N-terminal leader peptide (Figure 1b). The latter leader sequence is removed during the cycloaddition reaction, presumably during aromatization of the central pyridine. We demonstrated that TclM carries out its chemistry without the need for a cofactor or accessory enzyme but that the leader peptide does appear to be necessary.

Given the dramatic late-stage chemistry done by TclM, we sought to exploit this enzyme class as mild, broadly applicable biocatalysts to access new thiopeptide variants. We envisioned that late-stage enzymatic cyclization could allow the best of both biosynthesis and total synthesis: a modular, linear, and diversifiable route to variants not possible in previous synthetic schemes as well as defined access to non-natural structural motifs that are inaccessible to natural biosynthetic pathways. More specifically, we sought to develop a solid-phase route to variant substrate cores that could then be coupled to a leader peptide and cyclized by TclM or a homologue. Key to these efforts would be the identification of a concise, synthetically tractable pyridine synthase recognition sequence (RS) within the leader peptide that would significantly simplify the burden of peptide synthesis in this approach.

We began the development of this strategy by exploring TclM's leader peptide requirements through a series of leader peptide truncation experiments. Our previously reported solution-phase synthesis of an artificial thiocillin core was adapted to solid-phase peptide synthesis (SPPS). Substantially optimized yields were obtained in a microwave-assisted Fmoc-based synthesis using H-Rink Amide ChemMatrix resin and HATU as the main coupling reagent (see the Supporting Information). This allowed us to split material between a series of six leader peptide truncates to test substrate requirements: 20-residue [TclE(19−38)], 10-residue [TclE(29−38)], nine-residue [TclE(30−38)], eight-residue [TclE(31−38)], seven-residue [TclE(32−38)], and six-residue [TclE(33−38)] leader peptides. The five smaller truncates were prepared by extension of the core in a single solid-phase synthesis. The 20-residue conjugate was prepared by native chemical ligation onto the corresponding thioester, obtained from its intein fusion as previously described.[13] The remaining free cysteines were then converted to Dhas according to the method of Davis and co-workers.[14] Subsequent incubations with TclM yielded cyclized product from all of the conjugates except the shortest six-residue truncate and minimal amounts in the seven-residue truncate (Figure 2), as indicated by absorbance at both 254 and 350 nm and confirmed by MS/MS. This suggests that residues 32−38 of the leader peptide are necessary and sufficient for processing by

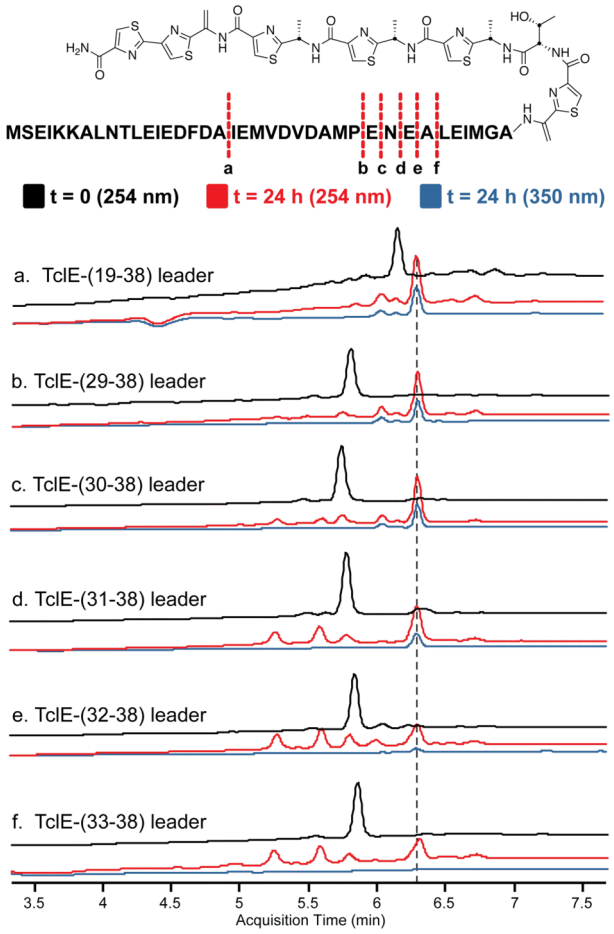

Figure 2. UV traces of TclM assays with leader peptide truncates containing (a) 20, (b) 10, (c) 9, (d) 8, (e) 7, and (f) 6 residues.

TclM. Interestingly, cyclization seemed most efficient with the 10-residue leader truncate, potentially because of increased solubility of the substrate.

Several byproducts with absorbance at 254 nm begin to appear in reactions with the eight-, seven-, and six-residue truncates (Figure 2f). At present, the masses do not correlate to obvious byproducts of the proposed mechanism, although several masses may suggest further enzymatic truncation of the leader peptide fragment (Figure S3). Addition of protease inhibitors and EDTA failed to inhibit the formation of the byproducts (Figure S3); further spectroscopic characterization will be required in order to accurately determine the identities of these compounds.

We next sought to exploit this minimal recognition motif in an expedited route to variant substrates. We envisioned a modular four-step route to thiopeptide variants (Figure 3a): (1) Fmoc-based SPPS of the modified cores (Figure 3b), (2) on-bead ligation of the RS, (3) cleavage and installation of Dhas, and (4) enzymatic macrocyclization. We initially screened a number of non-native RS linkages in order to simplify the synthetic scheme. In these efforts, however, TclM seemed sensitive to having the native peptide backbone at the RS−core juncture. We therefore chose to use the *N*-acyl-*N*′-methylacylurea (MeNbz) chemistry developed by Dawson and co-workers for Fmoc-based solid-phase synthesis of leader peptide thioesters.[15] This chemistry allowed the facile preparation of large-scale quantities of leader peptide thioester and efficient on-bead conjugation of crude leader peptide to linear thiazole-containing cores after orthogonal deprotection of the cysteine trityl ethers with 0.1 N

DOI: 10.1021/jacs.6b05389
J. Am. Chem. Soc. 2016, 138, 13461−13464

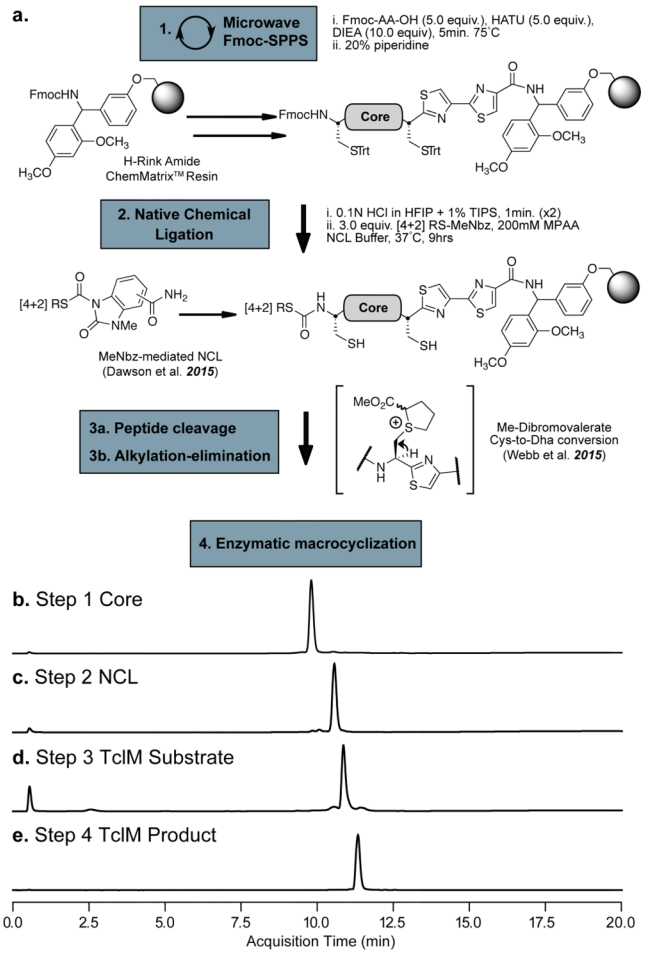

Figure 3. (a) Modular four-step route to thiopeptide variants and (b–e) UV chromatograms for each of the individual steps.

hydrochloric acid in hexafluoroisopropanol.[16] Subsequent cleavage of the ligated products provided clean linear thiopeptides with only one intermediate chromatographic purification (Figure 3c). Importantly, the use of methyl 2,5-dibromovalerate as developed by Webb and co-workers completely suppressed the formation of undesired intramolecular "staples" by bis-*S*-alkylation and provided quick, clean conversion to the desired substrates (Figure 3d).[17] Simple diethyl ether extraction removed excess alkylation–elimination reagent that might otherwise interfere with enzymatic cyclization. Substrates prepared in this manner could be cleanly converted to the resultant thiopeptides by incubation with TclM (Figure 3e).

Previous in vivo mutagenesis has highlighted the broad substrate promiscuity of TclM, but permissivity at residues directly flanking the formal 4π-cycloaddition partner was not completely clear because of differential flux through other enzymes in the pathway.[18,19] We therefore sought to deploy our streamlined substrate route to probe potential C-terminal substrate requirements of TclM and the recently characterized homologue TbtD from the thiomuracin GZ pathway as a prelude to more in-depth comparative efforts.[20] We chose to extend this investigation first to TbtD to compare pyridine synthases from pathways that make differently sized macrocycles with distinct biological activities (Figure 1a).[11,20] The leader peptides for TclE and TbtA are largely divergent. While a minimal leader sequence had not yet been demonstrated for TbtD, we anticipated that similar rules might apply for this TclM homologue, so the corresponding 10-residue leader peptide fragment from the thiomuracin pathway was used.

We prepared 13 cores using SPPS, which were separately ligated to the two different leader fragments in a kind of combinatorial synthesis of a series of chimeras. The cores were designed to combine synthetic simplicity and similarity to natural products while testing defined iterative changes to the C-terminus. These ligation products were then cleaved, converted to Dhas, and incubated with enzyme; the assay results are summarized in Table 1, and several useful insights are discussed

Table 1. Results of Enzymatic Assays with Leader/Core Chimeras[a]

		Cyclization	
Entry	Core	TclE(29-38)/TclM	TbtA(25-34)/TbtD
1	Dha Thz Thr Dha Thz Ala Thz Ala Thz Dha Thz Thz Ala	✓	✗
2	Dha Thz Thr Ala Thz Ala Thz Ala Thz Dha Thz Thz	✓	✗
3	Dha Thz Thr Ala Thz Ala Thz Ala Thz Dha Thz Dha	✗	✗
4	Dha Thz Thr Ala Thz Ala Thz Ala Thz Dha Thz	✗	✗
5	Dha Thz Thr Ala Thz Ala Thz Ala Ala Thz Dha Thz Thz	✓	✗
6	Dha Thz Thr Ala Ser Ala Ser Ala Thz Dha Thz Thz	✓	✗
7	Dha Thz Asn Thz Phe Thz Tyr Val Thz Thz Dha Thz Dha	✗	✓
8	Dha Thz Asn Thz Phe Thz Tyr Val Thz Thz Dha Thz	✗	✗
9	Dha Thz Asn Thz Val Thz Tyr Val Thz Thz Dha Thz Dha	✗	✓
10	Dha Thz Asn Thz Val Thz Tyr Val Thz Thz Dha Thz	✗	✗
11	Dha Thz Asn Thz Val Thz Gly Val Thz Thz Dha Thz Thz	✗	✓
12	Dha Thz Asn Thz Phe Thz Val Thz Thz Dha Thz Thz Dha	✗	✗
13	Dha Thz Asn Ser Phe Ser Tyr Val Thz Thz Dha Thz Dha	✗	✗

[a]A green checkmark indicates that the pyridine product was observed in the crude assay; a red "x" indicates that only the starting material was observed.

in detail below. Additionally, antibiotic activity of the two "nearest natural" cyclic products (entries 1 and 6) was confirmed by disk diffusion assay; as a control, the all-Ala product from entry 2 was inactive in these assays, as anticipated on the basis of prior mutagenesis (Figure S1).[19]

Similar to TclM, the 10-residue leader peptide is sufficient for TbtD cyclization (entries 7, 9, and 11). A logo analysis of the C-terminal 10 amino acids from the 13 leader peptides from known pyridine-containing thiopeptides shows substantial sequence divergence, despite a strong proclivity for β-branched and aliphatic patches, mixed with acidic residues toward the N-terminus (Figure S2). The location of a pyridine synthase RS may be conserved in the C-terminal fragments of thiopeptide leader peptides, although a larger sampling will be needed to confirm this trend as key residues are not immediately obvious from informatic analysis.

Both TclM and TbtD are highly sensitive to C-terminal constituents. For example, TclM requires the exocyclic bisthiazole downstream of the 4π-Dha; truncation to the monothiazole or replacement of the Cys-12 thiazole with a Dha prevents observable enzymatic cyclization (entry 1 or 2 vs entries 3 and 4).[19] Similarly, TbtD is sensitive to the placement of a monothiazole-Dha motif in this exocyclic position: removal of the C-terminal Dha, prevents cyclization. This effect is seen in multiple substrates for TbtD (entries 7 vs 8 and 9 vs 10). Interestingly, TbtD is tolerant of the replacement of the C-terminal monothiazole-Dha motif with a bisthiazole (entry 11). Thus, TbtD can modify both endocyclic (thiomuracin-like) and exocyclic (thiocillin-like) bisthiazole substrates.

DOI: 10.1021/jacs.6b05389
J. Am. Chem. Soc. 2016, 138, 13461–13464

TclM and TbtD exhibit different requirements for composition, size, and extent of functionalization of the internal (core) substituents. TclM modifies a substrate with a larger macrocycle (entry 5) as long as the C-terminal bisthiazole is present, but neither enzyme accepts the shortened and likely severely constrained TbtD substrate lacking a tyrosine (entry 12). Replacement of the two internal thiazoles with serines in both the thiocillin and thiomuracin series results in substrates that can be cyclized by TclM but not TbtD (entries 6 and 13). Bacterial strains were not capable of producing comparable substrates in earlier mutasynthetic efforts. The cyclization result suggests that TclM has potentially broader permissivity with respect to internal residues than TbtD. Given the conformational constraint of macrocyclization, pyridine synthases may also exhibit some amount of substrate control.

Leader peptides have proven key to RiPP enzymology in general, and understanding their role has led to significant innovations in the use of RiPP biosynthetic pathways for biotechnology.[21−27] RiPP leader peptides have been shown to interact with the biosynthetic enzymes and are thought to both activate the enzyme and guide the subsequent post-translational modifications on the core substrate. Pyridine synthases may well contain a new RiPP recognition element that interacts with leader peptide C-terminal recognition sequences. Additionally, our experiments indicate that the pyridine synthases exhibit selectivity for C-terminal core functionality, which is potentially necessary for "setting the register" in identifying the dehydroalanine pair to undergo cyclization, presumably via the intermediate imidate tautomer (Figure 4).

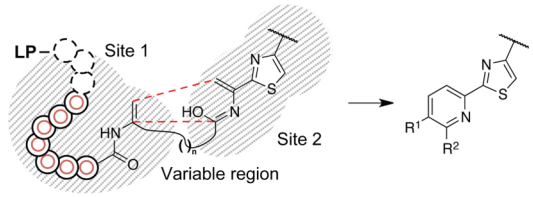

Figure 4. Proposed model of pyridine synthase enzyme−substrate interaction. TclM likely recognizes both C- and N-terminal motifs, including a C-terminal fragment of the leader peptide.

In conclusion, we have demonstrated the first application of microwave-assisted Fmoc-SPPS to generate thiazole-containing peptides and developed a new chemoenzymatic route to access two classes of thiopeptides, thiocillin and thiomuracin. The development of this strategy was facilitated by the discovery that a truncated seven-residue leader peptide is necessary and sufficient to coordinate cycloaddition by TclM and that the location of such an RS seems conserved in the substrate of the pyridine synthase TbtD. The solid-phase chemoenzymatic route should allow a more extensive exploration of these requirements in a systematic manner across multiple such pyridine synthases; the results presented here represent a preliminary sampling. We anticipate that this strategy could be extended to other classes of RiPPs and generate diverse libraries of analogues not accessible by strictly chemical or biosynthetic means.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b05389.

Procedures, synthetic schemes, and figures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*abower2@email.unc.edu

Author Contributions

†W.J.W. and J.W.B. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank B. Li and C. Neumann for informative discussions. A.A.B. is a Beckman Young Investigator and acknowledges support by the Arnold and Mabel Beckman Foundation.

■ REFERENCES

(1) Bagley, M. C.; Dale, J. W.; Merritt, E. A.; Xiong, X. *Chem. Rev.* **2005**, *105*, 685.
(2) LaMarche, M. J.; Leeds, J. A.; Dzink-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.
(3) Schoof, S.; Pradel, G.; Aminake, M. N.; Ellinger, B.; Baumann, S.; Potowski, M.; Najajreh, Y.; Kirschner, M.; Arndt, H.-D. *Angew. Chem., Int. Ed.* **2010**, *49*, 3317.
(4) Hegde, N. S.; Sanders, D. A.; Rodriguez, R.; Balasubramanian, S. *Nat. Chem.* **2011**, *3*, 725.
(5) Ciufolini, M. A.; Lefranc, D. *Nat. Prod. Rep.* **2010**, *27*, 330.
(6) Aulakh, V. S.; Ciufolini, M. A. *J. Am. Chem. Soc.* **2011**, *133*, 5900.
(7) Lefranc, D.; Ciufolini, M. A. *Angew. Chem., Int. Ed.* **2009**, *48*, 4198.
(8) Gross, S.; Nguyen, F.; Bierschenk, M.; Sohmen, D.; Menzel, T.; Antes, I.; Wilson, D. N.; Bach, T. *ChemMedChem* **2013**, *8*, 1954.
(9) Nicolaou, K. C. *Angew. Chem., Int. Ed.* **2012**, *51*, 12414.
(10) Nicolaou, K. C.; Zou, B.; Dethe, D. H.; Li, D. B.; Chen, D. Y.-K. *Angew. Chem., Int. Ed.* **2006**, *45*, 7786.
(11) Just-Baringo, X.; Albericio, F.; Alvarez, M. *Angew. Chem., Int. Ed.* **2014**, *53*, 6602.
(12) Luo, X.; Zambaldo, C.; Liu, T.; Zhang, Y.; Xuan, W.; Wang, C.; Reed, S. A.; Yang, P.-Y.; Wang, R. E.; Javahishvili, T.; Schultz, P. G.; Young, T. S. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 3615.
(13) 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.
(14) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernández-González, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem. Sci.* **2011**, *2*, 1666.
(15) Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197.
(16) Palladino, P.; Stetsenko, D. A. *Org. Lett.* **2012**, *14*, 6346.
(17) Morrison, P. M.; Foley, P. J.; Warriner, S. L.; Webb, M. E. *Chem. Commun.* **2015**, *51*, 13470.
(18) Bowers, A. A.; Acker, M. G.; Young, T. S.; Walsh, C. T. *J. Am. Chem. Soc.* **2012**, *134*, 10313.
(19) Bowers, A. A.; Acker, M. G.; Koglin, A.; Walsh, C. T. *J. Am. Chem. Soc.* **2010**, *132*, 7519.
(20) Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A. *J. Am. Chem. Soc.* **2015**, *137*, 16012.
(21) 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.
(22) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. *Nat. Chem. Biol.* **2015**, *11*, 564.
(23) Sardar, D.; Pierce, E.; McIntosh, J. A.; Schmidt, E. W. *ACS Synth. Biol.* **2015**, *4*, 167.
(24) Goto, Y.; Ito, Y.; Kato, Y.; Tsunoda, S.; Suga, H. *Chem. Biol.* **2014**, *21*, 766.
(25) Weiz, A. R.; Ishida, K.; Makower, K.; Ziemert, N.; Hertweck, C.; Dittmann, E. *Chem. Biol.* **2011**, *18*, 1413.
(26) Oman, T. J.; van der Donk, W. A. *Nat. Chem. Biol.* **2010**, *6*, 9.
(27) Cheung, W. L.; Pan, S. J.; Link, A. J. *J. Am. Chem. Soc.* **2010**, *132*, 2514.

DOI: 10.1021/jacs.6b05389
J. Am. Chem. Soc. 2016, 138, 13461−13464