Organic LETTERS

Letter

pubs.acs.org/OrgLett

Isolation, Structure Elucidation, Biosynthesis, and Synthesis of Antalid, a Secondary Metabolite from *Polyangium species*

Thomas Tautz,[†] Judith Hoffmann,[‡] Thomas Hoffmann,[‡] Heinrich Steinmetz,[§] Peter Washausen,[§] Brigitte Kunze,[§] Volker Huch,[∥] Andreas Kitsche,[⊥] Hans Reichenbach,[§] Gerhard Höfle,[§] Rolf Müller,*[,‡] and Markus Kalesse*[,†,§]

[†]Institute for Organic Chemistry, Leibniz Universität Hannover, Schneiderberg 1B, D-30167 Hannover, Germany

[‡]Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, Building E8.1, D-66123 Saarbrücken, Germany

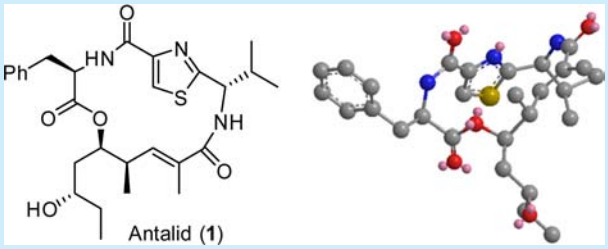
[§]Helmholtz Centre for Infection Research (HZI), Inhoffenstr. 7, D-38124 Braunschweig, Germany

[∥]Institute for Inorganic Chemistry, Saarland University, Building B2.2, D-66123 Saarbrücken, Germany

[⊥]Institute for Biostatistics, Leibniz Universität Hannover, Herrenhäuser Straße 2, D-30419 Hannover, Germany

Supporting Information

ABSTRACT: The isolation, structure elucidation, and synthesis of antalid (**1**), a novel secondary metabolite from *Polyangium sp.*, is described herein. The structure elucidation of **1** was performed with the aid of mass spectrometry, high field NMR experiments, and crystal structure analysis. The absolute configuration of antalid was confirmed through the Mosher ester method and ultimately by total synthesis. In addition, the biosynthetic origin of this hybrid PKS-NRPS natural product was unraveled by the *in silico* analysis of its biosynthetic gene cluster.



Myxobacteria are known to be a reliable source of novel natural products with often unique structural features that are rarely observed in other organisms.[1,2] To date, more than 100 secondary metabolites with different core structures and various biological activities have been isolated from myxobacterial sources.[3] A majority of these metabolites derive from non-ribosomal peptide synthetases, polyketide synthases, or are hybrids thereof.[4,5]

Already in 2001 we isolated antalid from *Polyangium sp.*, strain Pl 4620, and determined its constitution.[6] In the course of our current HPLC−MS based screening project, the crude extract from a liquid culture of the same *Polyangium sp.* was re-evaluated and showed a secondary metabolite with m/z = 528.2518 for the $[M + H]^+$ ion. Subsequent dereplication based on retention time, accurate mass, and isotope pattern using our in-house database for myxobacterial compounds indicated that the respective peak could be identical with antalid. Consequently, we set out to isolate this secondary metabolite, antalid (**1**), by fermentation of *P. spec.* Pl 4620 in 90 L of probion medium supplemented with 1% (v/v) XAD-16 adsorber resin. After several extraction steps and purification using RP-MPLC, a total of 30 mg of the pure compound (**1**) was isolated (see Supporting Information, SI). The accurate m/z value together with the observed isotope pattern indicated a tentative molecular formula of $C_{28}H_{37}N_3O_5S$, which would require 12 double bond equivalents. The exact two-dimensional structure was elucidated using 1D- and 2D-NMR spectroscopy (Figure 1a). Two methine protons in the range of amino acid α-protons (19-H, m, 4.57 ppm, and 12-H, dd, 5.23 ppm, ^3J = 9.2, 3.8 Hz) in combination with the corresponding COSY and HMBC correlations indicated the presence of a Phe and a Val residue. HMBC signals from 12-H, 13-H, 19-H, and NH_{Phe}, and a downfield shifted proton (s, 8.34 ppm) to three different quaternary carbons suggested the presence of a heteroaromatic five-membered ring, which was identified as a thiazole in direct neighborhood to a carbonyl group. The PKS-derived part of the molecule as inferred from COSY and HMBC couplings is terminated by a triple substituted double bond (8-H, dd, 6.60 ppm, ^3J = 8.5, 1.3 Hz). Finally, HMBC correlations from 5-H/19-H to 20-C and 10 H/12-H to 11-C determined the cyclic structure of antalid. Since we had been successful in predicting the configuration of polyketides based on analysis of keto

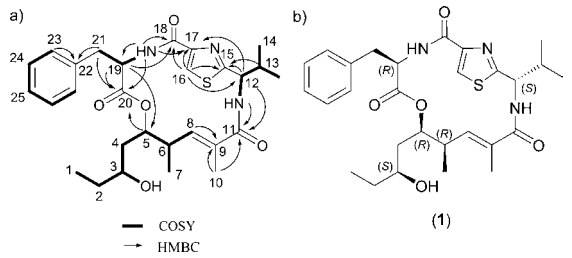


Figure 1. Structure of antalid (**1**): (a) planar structure and key 2D NMR correlations; (b) absolute configuration of the stereogenic centers.

Received: March 28, 2016
Published: May 24, 2016

DOI: 10.1021/acs.orglett.6b00810
Org. Lett. 2016, 18, 2560−2563

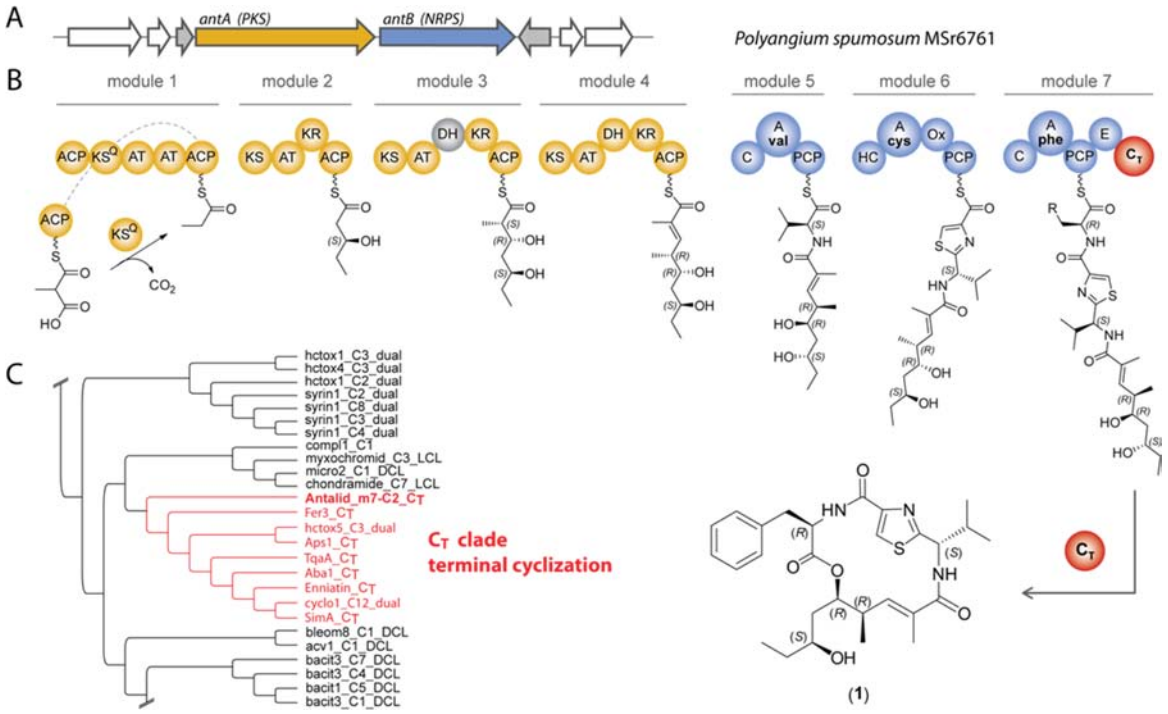
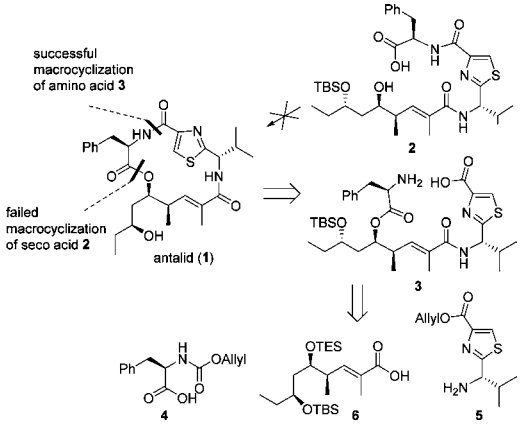
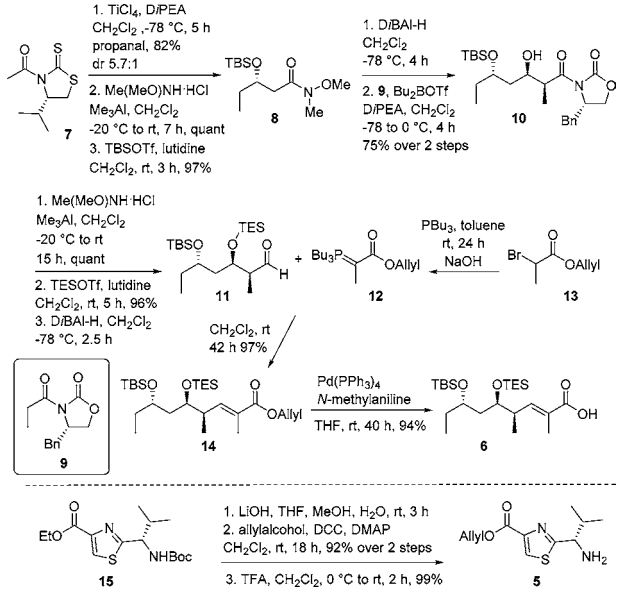


Figure 2. (A) Organization of the *ant* biosynthetic gene cluster in *Polyangium spumosum* MSr6761. (B) Proposed biosynthetic route to antalid formation in *P. spumosum*. The PKS modules 1−4 are located on *ant*A, whereas the NRPS modules 5−7 are found on *antB*. (C) Phylogenetic analysis of the terminal C domain found in the antalid cluster indicates homology to a group of special fungal C domains (C_T), which catalyze scaffold release from the megasynthetase. A, adenylation domain; ACP, acyl-carrier-protein domain; AT, acyltransferase domain; C, condensation domain; DH, dehydratation domain; E, epimerase domain; KR, ketoreductase domain; KS, ketosynthase domain; KS^Q, decarboxylating KS domain (^{Q}refers to the replacement of the active site cys to glu, which is often observed in such KS); PCP, peptidyl-carrier-protein domain.

Scheme 1. Retrosynthetic Disconnections of Antalid



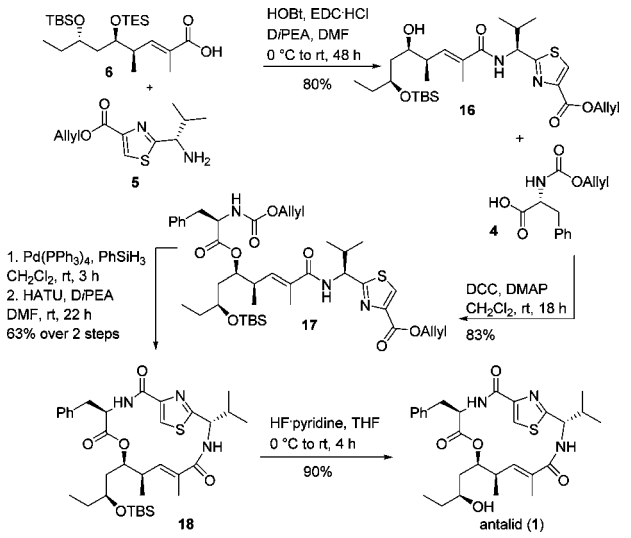
Scheme 2. Syntheses of Segments 6 and 5



reductase and enoyl reductase domains in the past,[7] we identified and sequenced the biosynthetic gene cluster (Ant-m2-KR and Ant-m3-KR). Examination of the configuration at hydroxyl branches (3-C, 5-C) was performed using our statistical method[10] since typical identifiers such as the LDD motif or the indicative tryptophan that are required according to Caffrey[8] and McDaniel[9] were missing (SI). Our statistical method calculated a score difference value of 7.2 for the configuration at 3-C, which is too low in order to be reliable. The score difference value of 85.98 for the configuration at 5-C, however, was of such a large absolute value that the D-configuration could be deduced with very high confidence. Based on the D-configuration of the alcohol at 5-C the statistical analysis of the configuration of the methyl branch at 6-C gave a score difference value of −124.7, indicating an L-configuration with very high confidence. This analysis was confirmed by a crystal structure analysis, which established the relative configuration of all five stereogenic centers unambiguously. In order to finalize the structure, the absolute configuration of 3-C was clarified by preparing (*S*)- and (*R*)-Mosher esters of the alcohol at the respective carbon, which unambiguously revealed the 3D structure of antalid as shown in Figure 1b.[11] Detailed preparation protocols and all MS and NMR data are included in the SI.

DOI: 10.1021/acs.orglett.6b00810
Org. Lett. 2016, 18, 2560−2563

Scheme 3. Fragment Coupling and Endgame in the Synthesis of Antalid (1)



In the course of the project we used our myxobacterial crude extract LC−MS database to identify several other *P. spec.*, which also produce antalid.[12] Since we had achieved sequencing of the genome of one of the alternative antalid producers, *P. spumosum* MSr6761, we were able to search for the respective biosynthetic gene cluster. In line with retrobiosynthetic considerations we expected a PKS-NRPS hybrid biosynthetic gene cluster featuring a characteristic heterocyclization domain to form the thiazole moiety of antalid.[13] By using a known myxobacterial heterocyclization domain as query sequence we identified a modular PKS-NRPS cluster with the necessary modules spanning 36,889 bp (Figure 2A).[14] The cluster is organized in two genes, *antA* and *antB*, with the PKS modules located on *antA* and the NRPS modules on *antB*. Homologues of a serine proteinase inhibitor upstream to *antA* and an active transporter-related protein downstream to *antB* may be involved in resistance and transport, respectively. With respect to the biosynthesis, only three PKS modules would be necessary to form the PKS subunit of antalid, whereas four modules are found on AntA (Figure 2B). However, analyzing the first ketosynthase (KS) reveals that the active site cysteine is substituted by a serine. Such KS frequently decarboxylate the extender unit without actually catalyzing the condensation step (SI).[15] Consequently, AntA-m1-KS most likely acts by decarboxylating an initially loaded methylmalonate to form the starter unit, an ACP-bound propionate. In line with this the double-AT could be a remnant of a former biosynthetic pathway version that had an elongating KS domain rather than a decarboxylating one. Biosynthesis continues with two methylmalonyl-CoA extensions to form the PKS backbone of antalid. The dehydratase (DH) of module 3 is inactive as inferred from the structure of antalid (**1**). It should be noted though, that the peptide sequence of AntA-m3-DH has no obvious deviation in the typical motifs that would indicate an inactive domain.[16] The forthcoming NRPS modules incorporate valine, cysteine, and phenylalanine with the cysteine of module 6 being cyclized and oxidized to form a thiazole. An epimerase in module 7 is in agreement with the (*R*)-Phe moiety found in antalid (**1**). The intriguing aspect of antalid biosynthesis is, however, the lack of a terminal thioesterase (TE) domain, which releases the product. Instead, the cluster ends with a condensation (C) domain. In particularly some fungal NRPS are known to release a cyclized product by special terminal C domains (C_T).[17] Bacterial C domains of similar function have been reported for the pipecolate-containing compounds FK520 and rapamycin.[18] In addition, a recently characterized new type of terminal C domain can release linear carboxylic acids from the assembly line.[19] Protein sequence alignment using the NaPDoS web tool resulted in a clear grouping of the pipecolate-related C_T domains in a single clade, while the terminal antalid C domain is located among the fungal C_T (Figure 2C and SI).[20] This finding is supported by a more comprehensive alignment, which revealed distinct groups for all three atypical C domains and even indicated a subgrouping of the antalid domain slightly apart from the fungal C_T (SI).

Similar to the biosynthetic route, our initial strategy for the synthesis of antalid (**1**) was to establish the macrocyclization of seco acid **2** through an ester linkage as the pivotal step (Scheme 1). Despite extensive experimentation with various established macrocyclization conditions (Yamaguchi,[21] Shiina,[22] Keck-Boden,[23,24] etc.) we were not able to obtain the desired cyclization product. As an alternative, we envisioned to establish the macrocycle through peptide bond formation of intermediate **3**. In order to liberate the amino and acid functionalities we relied on the alloc protecting group and an allyl ester, which could be removed selectively under the same reaction conditions in a single transformation.

Since **14** exhibits a typical motif that could in principal be constructed via a vinylogous Mukaiyama aldol reaction, we initially investigated this conversion. However, only decomposition or elimination products were observed under a variety of conditions. In the forward direction the synthesis of the polyketide fragment started with a Nagao aldol reaction using reagent **7** and propanal.[25,26] Subsequent conversion to the Weinreb amide followed by TBS protection provided compound **8**.[27] Reduction to the corresponding aldehyde[28] and Evans aldol reaction established the remaining two chiral centers of the polyketide segment. Transformation to aldehyde **11** was achieved with the established sequence of Weinreb amide formation, TES protection, and D*i*BAl-H reduction. Based on our experience, saponification of α-substituted unsaturated esters is often problematic. Thus, we decided to introduce an allyl ester. Therefore, we prepared Wittig reagent **12** from α-bromo allylester **13**[29] and tri-*n*-butylphosphine, which reacted satisfyingly in the olefination reaction that led to the polyketide fragment.

The allyl ester allowed mild deprotecting conditions and established acid **6** in high yields through deprotection with palladium and *N*-methylaniline (Scheme 2).

The synthesis of thiazole **5** started out from known ester **15**[30,31] and introduced an allyl moiety to selectively deprotect both the acid and the amine under mild conditions at the end of the synthesis. Next, acid **6** was brought to reaction with thiazole **5** employing EDC and HOBt, which removed the TES group simultaneously. Compound **16** could therefore directly be used in the condensation step with alloc-protected D-phenylalanine (**4**).[32,33] Now the stage was set for the removal of both protecting groups (allyl ester and alloc) and subsequent macrocycle formation. The removal of both protecting groups was a clean transformation, and we therefore subjected the corresponding amino acid directly to cyclization conditions utilizing HATU. The desired product was obtained in 63% yield and was transformed to antalid (**1**) by removing the remaining TBS group with HF-pyridine in 90% (Scheme 3). All spectroscopic

DOI: 10.1021/acs.orglett.6b00810
Org. Lett. 2016, 18, 2560−2563

data of synthetic and authentic antalid were in good agreement and thus confirm the proposed structure.

In summary, we have isolated and characterized the novel secondary metabolite antalid (**1**) from *Polyangium spec.* and *P. spumosum* Pl 4620. Following retrobiosynthetic considerations we identified the PKS-NPRS hybrid biosynthetic gene cluster by mining the genome of the antalid (**1**) producer *P. spumosum* MSr6761 for thiazole forming modules. Finally, we confirmed the assigned structure by total synthesis, which provides antalid (**1**) in a longest linear sequence of 15 steps. Worth mentioning is the fact that the macrocyclization could not be achieved by joining both the polyketide and the peptidic portion but through amide formation between both amino acids, which were liberated through simultaneous allyl ester and alloc cleavage.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.6b00810.

Experimental procedures and spectral data of compounds described herein; X-ray analysis of antalid (**1**) (PDF)

Accession Codes

The biosynthetic gene cluster has been deposited in the MIBiG repository with accession number BGC0001235. The GenBank accession number for the nucleotide sequence is KU245058. A CID-based MS/MS spectrum of antalid has been uploaded to the GNPS library, which is accessible under http://gnps.ucsd.edu.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: Markus.Kalesse@oci.uni-hannover.de.
*E-mail: Rolf.Mueller@helmholtz-hzi.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank B. Trunkwalter for helpful technical assistance and C. Kakoschke (both Helmholtz Centre for Infection Research, Braunschweig, Germany) for measuring NMR spectra.

■ REFERENCES

(1) Weissman, K. J.; Müller, R. *Nat. Prod. Rep.* **2010**, *27*, 1276.
(2) Reichenbach, H. *J. Ind. Microbiol. Biotechnol.* **2001**, *27*, 149.
(3) (a) Wenzel, S. C.; Müller, R. *Mol. BioSyst.* **2009**, *5*, 567. (b) Schäberle, T. F.; Lohr, F.; Schmitz, A.; König, G. M. *Nat. Prod. Rep.* **2014**, *31*, 953.
(4) Bode, H. B.; Müller, R. Secondary metabolism in myxobacteria. In *Myxobacteria: Multicellularity and differentiation*; Whitworth, D., Ed.; ASM Press: Chicago, 2007; p 259.
(5) Weissman, K. J.; Müller, R. *Bioorg. Med. Chem.* **2009**, *17*, 2121.
(6) Kunze, B.; Washausen, P.; Reichenbach, H.; Höfle, G. Unpublished data.
(7) (a) Hartmann, O.; Kalesse, M. *Angew. Chem., Int. Ed.* **2014**, *53*, 7335. (b) Jahns, C.; Hoffmann, T.; Müller, S.; Gerth, K.; Washausen, P.; Höfle, G.; Reichenbach, H.; Kalesse, M.; Müller, R. *Angew. Chem., Int. Ed.* **2012**, *51*, 5239. (c) Essig, S.; Bretzke, S.; Müller, R.; Menche, D. *J. Am. Chem. Soc.* **2012**, *134*, 19362. (d) Brodmann, T.; Janssen, D.; Kalesse, M. *J. Am. Chem. Soc.* **2010**, *132*, 13610. (e) Bock, M.; Buntin, K.; Müller, R.; Kirschning, A. *Angew. Chem., Int. Ed.* **2008**, *47*, 2308. (f) Menche, D.; Arikan, F.; Perlova, O.; Horstmann, N.; Ahlbrecht, W.; Wenzel, S. C.; Jansen, R.; Irschik, H.; Müller, R. *J. Am. Chem. Soc.* **2008**, *130*, 14234. (g) Janssen, D.; Albert, D.; Jansen, R.; Müller, R.; Kalesse, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 4898.
(8) (a) Caffrey, P. *ChemBioChem* **2003**, *4*, 654−657. (b) Caffrey, P. *Chem. Biol.* **2005**, *12*, 1060.
(9) Reid, R.; Piagentini, M.; Rodriguez, E.; Ashley, G.; Viswana-than, N.; Carney, J.; Santi, D. V.; Hutchinson, C. R.; McDaniel, R. *Biochemistry* **2003**, *42*, 72.
(10) Kitsche, A.; Kalesse, M. *ChemBioChem* **2013**, *14*, 851. The online profile Hidden Markov Model analysis tool can be found at: https://akitsche.shinyapps.io/profileHMM_App.
(11) Hoye, T. R.; Jeffrey, C. S.; Shao, F. *Nat. Protoc.* **2007**, *2*, 2451.
(12) Krug, D.; Müller, R. *Nat. Prod. Rep.* **2014**, *31*, 768.
(13) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. *BMC Evol. Biol.* **2007**, *7*, 78.
(14) Buntin, K.; Irschik, H.; Weissman, K. J.; Luxenburger, E.; Blöcker, H.; Müller, R. *Chem. Biol.* **2010**, *17*, 342.
(15) (a) Bisang, C.; Long, P. F.; Cortés, J.; Westcott, J.; Crosby, J.; Matharu, A. L.; Cox, R. J.; Simpson, T. J.; Staunton, J.; Leadlay, P. F. *Nature* **1999**, *401*, 502. (b) Long, P. F.; Wilkinson, C. J.; Bisang, C. P.; Cortes, J.; Dunster, N.; Oliynyk, M.; McCormick, E.; McArthur, H.; Mendez, C.; Salas, J. A.; Staunton, J.; Leadlay, P. F. *Mol. Microbiol.* **2002**, *43*, 1215.
(16) Keatinge-Clay, A. *J. Mol. Biol.* **2008**, *384*, 941.
(17) Gao, X.; Haynes, S. W.; Ames, B. D.; Wang, P.; Vien, L. P.; Walsh, C. T.; Tang, Y. *Nat. Chem. Biol.* **2012**, *8*, 823.
(18) (a) Gatto, G. J., Jr.; McLoughlin, S. M.; Kelleher, N. L.; Walsh, C. T. *Biochemistry* **2005**, *44*, 5993. (b) Schwecke, T.; Aparicio, J. F.; Molnar, I.; König, A.; Khaw, L. E.; Haydock, S. F.; Oliynyke, M.; Caffrey, P.; Cortes, J.; Lester, J. B.; Bohm, G. A.; Staunton, J.; Leadley, P. F. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7839.
(19) Müller, S.; Rachid, S.; Hoffmann, T.; Surup, F.; Volz, C.; Zaburannyi, N.; Müller, R. *Chem. Biol.* **2014**, *21*, 855.
(20) Ziemert, N.; Podell, S.; Penn, K.; Badger, J. H.; Allen, E.; Jensen, P. R. *PLoS One* **2012**, *7*, e34064.
(21) Inanaga, J.; Hirata, K.; Saeki, H.; Katsuki, T.; Yamaguchi, M. *Bull. Chem. Soc. Jpn.* **1979**, *52*, 1989.
(22) Shiina, I.; Kubota, M.; Ibuka, R. *Tetrahedron Lett.* **2002**, *43*, 7535.
(23) Boden, E. P.; Keck, G. E. *J. Org. Chem.* **1985**, *50*, 2394.
(24) Kageyama, M.; Tamura, T.; Nantz, M. H.; Roberts, J. C.; Somfai, P.; Whritenour, D. C.; Masamune, S. *J. Am. Chem. Soc.* **1990**, *112*, 7407.
(25) Hodge, M. B.; Olivo, H. F. *Tetrahedron* **2004**, *60*, 9397.
(26) Nagao, Y.; Hagiwara, Y.; Kumagai, T.; Ochiai, M.; Inoue, T.; Hashimoto, K.; Fujita, E. *J. Org. Chem.* **1979**, *51*, 2391.
(27) Gupta, S.; Lakshmanan, V.; Kim, B. S.; Fecik, R.; Reynolds, K. A. *ChemBioChem* **2008**, *9*, 1609.
(28) DeShong, P.; Lin, M.-T.; Perez, J. J. *Tetrahedron Lett.* **1986**, *27*, 2091.
(29) von Werne, T.; Patten, T. E. *J. Am. Chem. Soc.* **2001**, *123*, 7497.
(30) Aguilar, E.; Meyers, A. I. *Tetrahedron Lett.* **1994**, *35*, 2473.
(31) Bruno, P.; Peña, S.; Just-Baringo, X.; Albericio, F.; Álvarez, M. *Org. Lett.* **2011**, *13*, 4648.
(32) Mori, M.; Somada, A.; Oida, S. *Chem. Pharm. Bull.* **2000**, *48*, 716.
(33) Nishimura, T.; Yamada, K.; Takebe, T.; Yokoshima, S.; Fukuyama, T. *Org. Lett.* **2008**, *10*, 2601.

DOI: 10.1021/acs.orglett.6b00810
Org. Lett. 2016, 18, 2560−2563