OL | Organic Letters

pubs.acs.org/OrgLett Letter

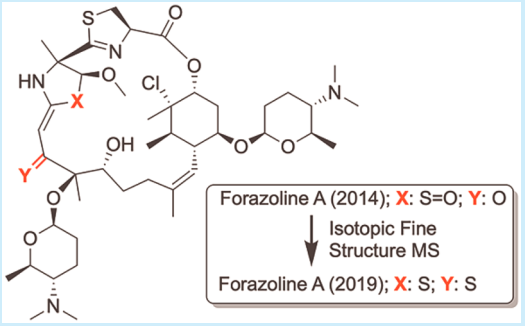
MS-Derived Isotopic Fine Structure Reveals Forazoline A as a Thioketone-Containing Marine-Derived Natural Product

Fan Zhang, Thomas P. Wyche, Yanlong Zhu, Doug R. Braun, Jia-Xuan Yan, Shaurya Chanana, Ying Ge, Ilia A. Guzei, Marc G. Chevrette, Cameron R. Currie, Michael G. Thomas, Scott R. Rajski, and Tim S. Bugni*

Cite This: https://dx.doi.org/10.1021/acs.orglett.9b04535 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Forazoline A is a structurally complex PKS−NRPS hybrid produced by marine-derived *Actinomadura* sp. During the course of studies highlighting the application of IFS analysis as a powerful tool for natural products analysis, we were alerted to an earlier misinterpretation with respect to forazoline A structure elucidation. In particular, IFS reveals that forazoline A contains a thioketone moiety rarely seen in secondary metabolites and, thus, constitutes an even more intriguing structure than originally thought.



Unambiguous assignment of elemental composition is critical to structure determination initiatives for new molecules.[1] Unfortunately, for most systems, the correct formulas cannot be determined solely on the basis of mass accuracy of monoisotopic MS signals.[2] A number of initiatives have endeavored to devise more efficient and accurate approaches for determining elemental composition by mass spectrometry; these include the development of the "seven golden rules" to restrict molecular formula space,[2] the advent of false discovery rate analyses,[3] and high-resolution direct infusion-based mass spectrometry in combination with whole ^{13}C metabolome isotope labeling.[4] However, despite these advances, an efficient and effectively foolproof means of assigning elemental compositions remains elusive.

Recent advances in MS have relied heavily on the idea that mass defects shed insight into molecular composition with a level of precision not achievable by other means. In particular, isotopic fine structure (IFS) is a unique MS signature attributable to the different mass defects of isotopic contributions and their natural abundance.[5] In theory, the number of atoms for each element (such as C, N, O, and S) within a given compound can be readily calculated based on the relative isotopic abundance and natural abundance of each isotope. Importantly, every element has a unique mass defect (Table S1).[6] As a result, unique IFS patterns are available for validating elemental composition assignments with ultrahigh resolution provided by fourier transform ion cyclotron resonance mass spectrometers (FT-ICR MS). We endeavored recently to apply IFS analyses to a panel of complex natural products whose structural elucidations proved particularly challenging prior to the advent and proliferation of FT-ICR MS technology.

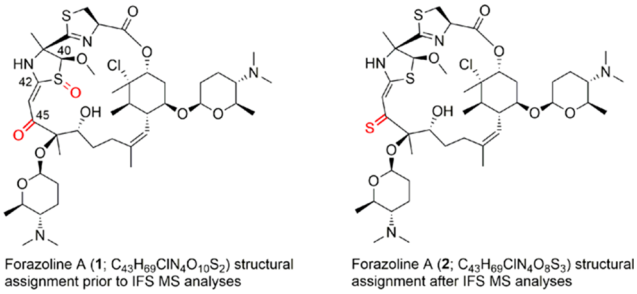
We discovered forazoline A (**1**), a novel halogenated and polythiolated polyketide from marine-derived *Actinomadura* sp., in 2014.[7] Importantly, forazoline A exerts antifungal activity *in vivo* through a previously unreported mechanism, making it an excellent antifungal drug lead. The structural complexity of forazoline A required a multifaceted strategy en route to structure elucidation. Besides the usual battery of standard NMR data acquisitions, $^{13}C-^{13}C$ gCOSY NMR experiments enabled us to establish carbon−carbon connectivities and $^{13}C-^{15}N$ HMQC NMR data unveiled direct carbon−nitrogen connectivities. Additionally, DFT NMR calculations and NOE experiments proved vital to structure elucidation efforts as did mass spectrometry. Particularly valuable, inductively coupled plasma atomic emission spectroscopy (ICP-AES) studies shed significant insight into the number of sulfur atoms contained within the forazoline scaffold. These collective data, obtained using state of the art instrumentation at the time, enabled us to assign forazoline A as structure **1** shown in Scheme 1.

However, more recent IFS studies in our laboratory suggested an apparent discrepancy with this assignment. In initial IFS studies, we employed FT-ICR MS using negative ion

Received: December 18, 2019

https://dx.doi.org/10.1021/acs.orglett.9b04535
Org. Lett. XXXX, XXX, XXX−XXX

Scheme 1. Original (1) and Revised (2) Structures for Forazoline A as Illuminated by FT-ICR MS and Subsequent IFS Analyses



mode which returned a monoisotopic mass for forazoline A of m/z = 899.3894 ($[M - H]^-$), suggesting a molecular formula of $C_{43}H_{69}ClN_4O_8S_3$ within 1 ppm error. These data contrasted with the originally noted $C_{43}H_{69}ClN_4O_{10}S_2$ that had been determined by HRMS and ICP-AES clearly signaling the need for more in depth analysis.[7]

To reinvestigate the elemental composition of forazoline A, a careful analysis of IFS was carried out. FTMS data initially collected at a resolving power of ~1 M for the ($[M - H]^-$ at m/z 899 (Figure 1b) revealed poorly resolved ^{34}S and ^{37}Cl

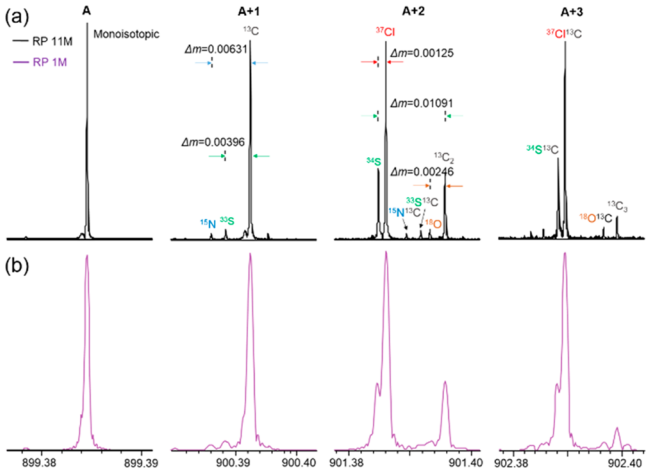


Figure 1. IFS of forazoline A with a monoisotopic peak at m/z 899.3892 ($[M - H]^-$). (a) IFS of forazoline A acquired with 11 M resolving power. (b) IFS of forazoline A acquired with 1 M resolving power.

isotopologues; therefore, a narrowband method[8] was implemented to increase the resolving power to 11 M at which point all isotopologues were well resolved (Figure 1a). The IFS of the A + 1 peaks revealed the presence of ^{15}N, ^{33}S, and ^{13}C isotopic peaks, with a difference of 0.00631 Da between the ^{13}C and ^{15}N isotopologues and a difference of 0.00396 Da between the ^{13}C and ^{33}S isotopologues. The A + 2 isotopic peaks were composed of ^{34}S, ^{37}Cl, $^{13}C^{15}N$,$^{13}C^{33}S$, ^{18}O, and $^{13}C_2$ isotopologues, where the mass differences between ^{34}S and ^{37}Cl, $^{13}C_2$ and ^{34}S, as well as $^{13}C_2$ and ^{18}O were 0.00125, 0.01091, and 0.00246 Da, respectively (Figure 1a). On the basis of these IFS data, the ratio of Cl:S was determined to be 1:3; the relatively higher natural abundance for ^{37}Cl and ^{34}S played a key role in these calculations.

Validation of these data was achieved by comparing simulated versus actual IFS data sets. Simulated IFS data sets for the two candidate formulas ($C_{43}H_{68}ClN_4O_8S_3^-$ and $C_{43}H_{68}ClN_4O_{10}S_2^-$) were generated using Bruker DataAnalysis software and then compared to experimentally observed IFS data corresponding to $[M - H]^-$. Notably, the experimental IFS data set for forazoline A was in good agreement with the simulated IFS of the molecular formula $C_{43}H_{68}ClN_4O_8S_3^-$ (Figure 2). In contrast, the A + 1, A + 2, and A + 3 isotopic

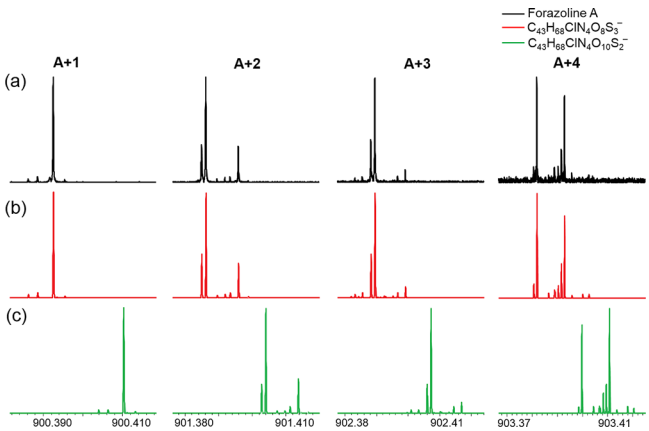


Figure 2. Comparison of IFS of forazoline A with calculated patterns of previously published and revised molecular formulas for forazoline A. (a) Experimental IFS of forazoline A ($[M - H]^-$). (b) Calculated IFS for $C_{43}H_{68}ClN_4O_8S_3^-$. (c) Calculated IFS for $C_{43}H_{68}ClN_4O_{10}S_2^-$.

peaks for the simulated IFS of $C_{43}H_{68}ClN_4O_{10}S_2^-$ differed dramatically from the experimental IFS signature. Hence, IFS data unambiguously revealed the true molecular formula of forazoline A to be $C_{43}H_{69}ClN_4O_8S_3$, as reflected by structure **2** (Scheme 1).

Beyond the application of IFS data in negative mode, we also acquired positive ion FTMS and FT MS/MS data and found two diagnostically important fragment ions at m/z 760.2886 and 619.1732 (Figure S2). Precise determination of the elemental composition of the m/z 760.2886 fragment was achieved via IFS analyses, which also clearly indicated the presence of three S atoms in forazoline A, further validating revision of the molecular formula of forazoline A to $C_{43}H_{69}ClN_4O_8S_3$.

These advances in both instrumentation and structural knowledge made clear the need to substitute two O atoms (in the previous assignment) with one S atom in the correctly assigned structure (**2**, Scheme 1). The sulfoxide in **1** was corrected to a thioether, based, in part, on the need to remove two O atoms. Although the presence of the sulfoxide (in **1**) had been suggested earlier by DFT calculations using the method B3LYP/6-31G(d,p),[7] we now posit that this was likely an inappropriate algorithm to use for this specific calculation. Forazoline A is a sulfur-containing molecule, and thus, diffuse functions are needed to correctly compute the ^{13}C NMR chemical shifts.[9−11] Additionally, the α,β-unsaturated ketone C-45 was originally assigned on the basis of a carbon resonance at 212.0 ppm, which was noted to be further downfield than expected. However, having accounted for all carbon substituents, assignment of C-45 as a ketone seemed a logical and not altogether unreasonable task. Interestingly, given the revised structure **2**, C-45 logically can be securely assigned as the shown α,β-unsaturated thioketone.

Structure revision of forazoline A was confirmed by single-crystal X-ray diffraction analysis (Figures 3 and S3). The

https://dx.doi.org/10.1021/acs.orglett.9b04535
Org. Lett. XXXX, XXX, XXX−XXX

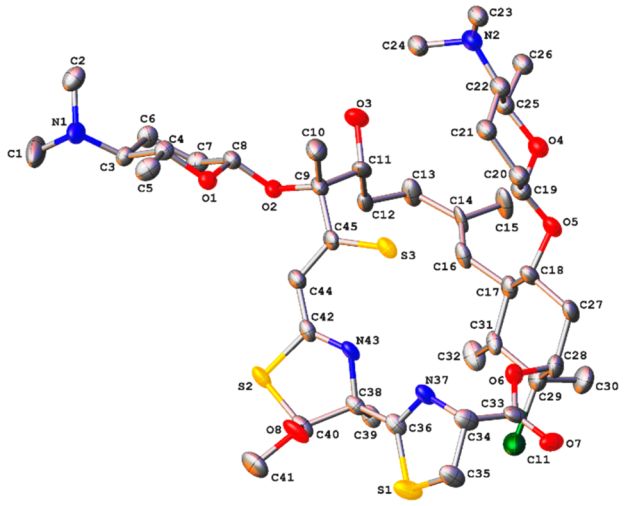


Figure 3. A molecular drawing of forazoline A (**2**) shown with 40% probability ellipsoids. All H atoms and the minor component of the disorder are omitted.

crystallographic data are consistent with the newly substantiated molecular formula ($C_{43}H_{69}ClN_4O_8S_3$) and further supported the C-45 as a component of the α,β-unsaturated thioketone. Beyond validation of the thioketone, careful analysis of the X-ray data allowed the absolute configuration of forazoline A unequivocally to be assigned as 3*S*, 4*R*, 8*S*, 9*R*, 11*R*, 17*R*, 18*R*, 19*R*, 22*S*, 25*R*, 28*R*, 29*S*, 31*R*, 34*R*, 38*S*, and 40*S*.

The thioketone group in forazoline A (**2**) is highly unusual among natural products. In attempting to understand the genetic origins of thioketone installation, we sequenced the *Actinomadura* sp. WMMB-499 genome using PacBio technology (Supporting Information). Having the complete WMMB-499 genome at our disposal (NCBI accession number CP044407), we applied antiSMASH 5.0[12] analyses to identify two PKS−NRPS hybrid gene clusters in the genome; of these, only one was deemed a viable candidate for forazoline production. This BGC (Figure 4) showed 66% similarity

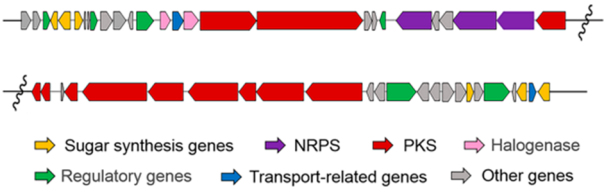


Figure 4. Layout of tentative PKS−NRPS BGC of forazoline A. Note the separation of two PKS blocks (red) by the NRPS coding genes (purple). Genes possibly associated with C-45 thiolation are either obscure or absent from this cluster, leading to the hypothesis that thioketone installation occurs *in trans*.

with the forosamine cluster,[13] whereas the other identified PKS−NRPS BGC showed no signs of being able to carry out either sugar synthesis or chlorination chemistries. Notably, two forosamine sugars are present in **2**.

At present, the issue of thioketone incorporation at C-45 remains an open question. We have found no evidence of thioketone-installing genes within the forazoline A cluster and have never noted any evidence of what we envision as a likely thioketone precursor, namely, the C-45 ketone. Proposals invoking cysteine as the source of the thioketone S atom are not altogether discounted, since the forazoline cluster does, in fact, encode the use of cysteine as a building block, albeit for the two thiazole rings.

To begin to address the issue of thioketone installation, we carried out isotopic labeling studies with ^{13}C-labeled building blocks ($^{13}C_3$-L-cysteine, 1-^{13}C-sodium propionate, and 1-^{13}C-sodium acetate). Samples of forazoline A isolated from ^{13}C-labeled and unlabeled media were analyzed by ^{13}C NMR. The carbon skeleton of **2** was predicted to incorporate five units of malonyl-CoA and, thus, five units of acetate. Indeed, C-11, C-16, C-18, C-31, and C-42 were enriched by 1-^{13}C-acetate relative to unlabeled material (Figures 5 and S6), indicating

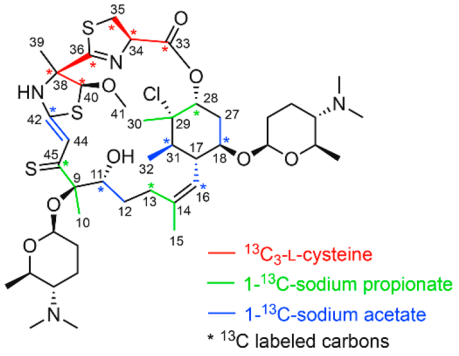


Figure 5. Isotopic incorporation of ^{13}C-enriched starting materials into the revised forazoline A structure (**2**).

that these carbons in **2** originate directly from malonyl-CoA. Next, a high level of ^{13}C enrichment was clearly discernible at C-28, C-31, and C-45 following 1-^{13}C-sodium propionate supplementations (Figures 5 and S5). Accordingly, it appeared that methylmalonyl-CoA was utilized three times during the course of forazoline biosynthesis. In parallel, WMMB-499 fermentations supplemented with $^{13}C_3$-L-cysteine gave rise to **2** with significant isotopic incorporations at C-33, C-34, C-35, C-36, C-38, and C-40, suggesting that the two thiazole-like moieties are derived from L-cysteine (Figures 5 and S4).

Isotopic incorporation studies make it clear that forazoline A assembly relies on both NRPS and PKS machineries. Most intriguing is that the thioketone C-45 appears to derive from PKS-driven incorporation (Figure 5). This finding supports our hypothesis that a likely precursor to intact **2** is the C-45 keto variant (Figure S7, intermediate **8**) prior to conversion to its thiolated product. In effect, we envision that thioketone incorporation at C-45 likely represents a tailoring step. The absence of any clear candidates for the execution of this tailoring chemistry within the forazoline cluster (Figure 4, Table S2) suggests that thioketone installation very likely needs to be carried out *in trans* by an enzyme coded for beyond the forazoline cluster boundaries. This poses an exceedingly rare and unique opportunity to discover new biosynthetic chemistry. *In trans* tailoring chemistries in their own right are rare. Although significant progress has been made to understand the biosynthesis of several thioamide-containing natural products,[14−18] the extent to which these advances can translate to the current forazoline system are unknown. Although we have identified a number of genes that code for *S*-handling enzymes well beyond the forazoline cluster (Table S3), the involvement of any of these in forazoline assembly remains unknown. Efforts to clarify these issues are ongoing in our laboratories.

https://dx.doi.org/10.1021/acs.orglett.9b04535
Org. Lett. XXXX, XXX, XXX−XXX

In sum, the structure of forazoline A has been revised unambiguously via the application of IFS and crystallographic data. The assignment of forazoline A as structure **2** (Scheme 1), given the complexities of the overall scaffold and key functional group similarities, could not have been readily achieved without the newly presented IFS data. This example highlights the potential of IFS in the context of expedient natural products structure elucidation. As directly shown here with forazoline A, IFS also promises to dramatically enhance the precision and speed with which new molecules are correctly elucidated. These attributes of IFS also promise to make more obvious, natural product structural features that inspire new biosynthetic initiatives to uncover otherwise covert chemistries.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.orglett.9b04535.

A complete description of experimental methods including forazoline A production, isolation, and sample preparation as well as isotopic incorporation, crystallographic data, and biosynthetic projections with supporting genomics data (PDF)

Accession Codes

The complete genome of WMMB-499 has been sequenced and deposited with NCBI accession number CP044407. CCDC 1976898 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, or by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033.

■ AUTHOR INFORMATION

Corresponding Author

Tim S. Bugni − *Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States;* orcid.org/0000-0002-4502-3084; Email: tim.bugni@wisc.edu

Authors

Fan Zhang − *Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States*

Thomas P. Wyche − *Merck & Co., Inc., Cambridge, Massachusetts 02141, United States*

Yanlong Zhu − *Department of Cell and Regenerative Biology and Human Proteomics Program, School of Medicine and Public Health, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States*

Doug R. Braun − *Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States*

Jia-Xuan Yan − *Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States*

Shaurya Chanana − *Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States*

Ying Ge − *Department of Cell and Regenerative Biology, Human Proteomics Program, School of Medicine and Public Health, and Department of Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States;* orcid.org/0000-0001-5211-6812

Ilia A. Guzei − *Department of Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States;* orcid.org/0000-0003-1976-7386

Marc G. Chevrette − *Department of Genetics and Department of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States;* orcid.org/0000-0002-7209-0717

Cameron R. Currie − *Department of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States*

Michael G. Thomas − *Department of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States*

Scott R. Rajski − *Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.orglett.9b04535

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by funding from (i) the University of Wisconsin—Madison School of Pharmacy and the Graduate School at the University of Wisconsin, (ii) NIH Grants U19AI109673 and U19AI142720, and (iii) NIGMS Grant R01GM104192. M.G.T. acknowledges funding from the National Science Foundation under Grant MCB-1716594; any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We also thank the Analytical Instrumentation Center at the School of Pharmacy, University of Wisconsin—Madison, for the facilities to acquire spectroscopic data. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH Grant P41GM103399 (NIGMS).

■ REFERENCES

(1) Kind, T.; Fiehn, O. *Bioanal. Rev.* **2010**, *2*, 23−60.
(2) Kind, T.; Fiehn, O. *BMC Bioinf.* **2007**, *8*, 105.
(3) Matsuda, F.; Shinbo, Y.; Oikawa, A.; Hirai, M. Y.; Fiehn, O.; Kanaya, S.; Saito, K. *PLoS One* **2009**, *4*, No. e7490.
(4) Giavalisco, P.; Hummel, J.; Lisec, J.; Inostroza, A. C.; Catchpole, G.; Willmitzer, L. *Anal. Chem.* **2008**, *80*, 9417−9425.
(5) Miladinović, S. M.; Kozhinov, A. N.; Gorshkov, M. V.; Tsybin, Y. O. *Anal. Chem.* **2012**, *84*, 4042−4051.
(6) Sleno, L. *J. Mass Spectrom.* **2012**, *47*, 226−236.
(7) Wyche, T. P.; Piotrowski, J. S.; Hou, Y.; Braun, D.; Deshpande, R.; McIlwain, S.; Ong, I. M.; Guzei, I. A.; Westler, W. M.; Andes, D. R.; Bugni, T. S. *Angew. Chem., Int. Ed.* **2014**, *53*, 11583−11586.
(8) Caprioli, R. M.; Malorni, A.; Sidona, G. *Mass Spectrometry in Biomolecular Sciences*; NATO ASI Ser. C, Vol. *475*; Kluwer: Dordrecht, The Netherlands, 1996.
(9) Papajak, E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 597−601.
(10) Greer, E. M.; Aebisher, D.; Greer, A.; Bentley, R. *J. Org. Chem.* **2008**, *73*, 280−283.

(11) Denis, P. A. *J. Chem. Theory Comput.* **2005**, *1*, 900−907.
(12) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2019**, *47*, W81−W87.
(13) Hong, L.; Zhao, Z.; Melançon, C. E., III; Zhang, H.; Liu, H. *J. Am. Chem. Soc.* **2008**, *130*, 4954−4967.
(14) Liu, J.; Lin, Z.; Li, Y.; Zheng, Q.; Chen, D.; Liu, W. *Org. Biomol. Chem.* **2019**, *17*, 3727−3731.
(15) Litomska, A.; Ishida, K.; Dunbar, K. L.; Boettger, M.; Coyne, S.; Hertweck, C. *Angew. Chem., Int. Ed.* **2018**, *57*, 11574−11578.
(16) Dunbar, K. L.; Dell, M.; Molloy, E. M.; Kloss, F.; Hertweck, C. *Angew. Chem., Int. Ed.* **2019**, *58*, 13014−13018.
(17) Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 3030−3035.
(18) Mahanta, N.; Szantai-Kis, D. M.; Petersson, E. J.; Mitchell, D. A. *ACS Chem. Biol.* **2019**, *14*, 142−163.

https://dx.doi.org/10.1021/acs.orglett.9b04535
Org. Lett. XXXX, XXX, XXX−XXX