Organic LETTERS

Letter

Cite This: Org. Lett. 2018, 20, 5238–5241

pubs.acs.org/OrgLett

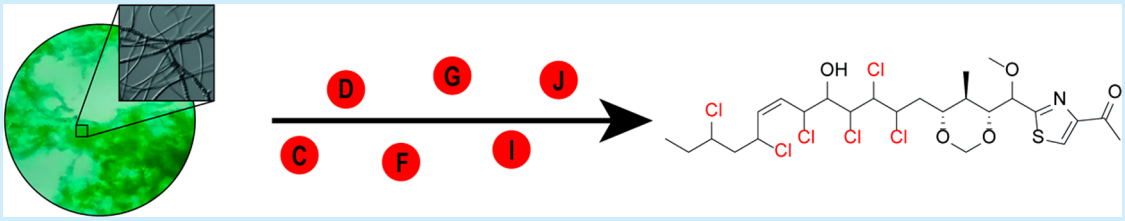
# Aranazoles: Extensively Chlorinated Nonribosomal Peptide–Polyketide Hybrids from the Cyanobacterium *Fischerella* sp. PCC 9339

Philipp Moosmann,[†] Reiko Ueoka,[†] Muriel Gugger,[‡] and Jörn Piel*[,†]

[†]Institut für Mikrobiologie, ETH Zürich, Vladimir-Prelog-Weg 4, 8093 Zürich, Switzerland

[‡]Institut Pasteur, Collection des Cyanobactéries, Département de Microbiologie, 75015 Paris, France

Supporting Information




**ABSTRACT:** An analysis of cyanobacterial genomes revealed an architecturally unique biosynthetic gene cluster with an unusually high number of genes encoding predicted iron(II)/α-ketoglutarate-dependent halogenases. Mass spectrometry-guided identification of the corresponding metabolites yielded the aranazoles, extensively halogenated nonribosomal peptide–polyketide hybrids. Their chlorine-bearing fatty acyl-like moiety is reminiscent of the hyperhalogenated chlorosulfolipids, natural products of unknown enzymatic origin that were previously isolated from eukaryotic algae and mussels.



Enzymatic halogenation, a process almost exclusively known from specialized metabolism, is involved in the biosynthesis of numerous structurally diverse natural products. Examples with important biological activities are the antibiotic vancomycin, the anticancer drug candidate salinosporamide A, and the hormone triiodothyronine.[1−3] In most cases, such molecules contain few halogens, but exceptions of more extensively halogenated compounds are known (Figure 1, **1**–**3**).[4,5] In the pathways characterized to date, multiple halogen substituents are usually attached by one to few halogenases acting on several substrate positions.[6−9] The most extensively halogenated compounds known are the chlorosulfolipids isolated from eukaryotic algae and mussels (Figure 1, **4**), which carry up to 11 halogen atoms introduced by as yet unknown biosynthetic enzymes.[10−12]

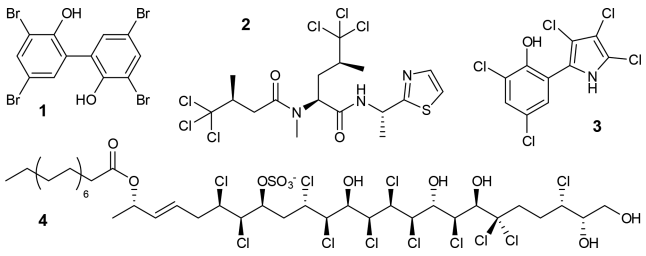



**Figure 1.** Examples of extensively halogenated natural products: a brominated biphenyl (**1**), dysidenin (**2**), pentachloropseudiline (**3**), and a chlorosulfolipid (**4**).

Various halogenating enzymes have independently evolved that utilize halogen and carbon substrates with distinct reactivities.[13] Halogenation of electron-rich moieties involving a reactive hypohalite intermediate occurs in heme-dependent haloperoxidases,[14,15] vanadium-dependent haloperoxidases,[16] and flavin-dependent halogenases.[8,17] Unactivated alkyl moieties are halogenated by Fe(II)/α-ketoglutarate-dependent halogenases (αKGHs) that use a radical mechanism for C−H functionalization.[7,18,19] *S*-Adenosyl methionine (SAM)-dependent chlorinases and fluorinases generate a halogenated 5′-deoxyadenosine product from SAM, which is then further metabolized.[20,21]

Halogenating enzymes are present in many bacteria and can be detected on the basis of sequence features, thus offering opportunities for the genomic prediction and discovery of novel halogenated natural products.[7] Using such a strategy, we here report the identification of a cyanobacterial biosynthetic gene cluster encoding an unprecedented number of αKGH homologues. Mass spectrometry (MS) guided natural product isolation yielded the hyperhalogenated aranazoles. The genetic data and comparative structural features of four different characterized congeners suggest that the halogenases act on a fatty acid derived starter moiety of a polyketide–nonribosomal peptide hybrid precursor.

Previous studies on cyanobacterial natural products document an impressive biosynthetic diversity, of which much

**Received:** July 13, 2018
**Published:** August 24, 2018




DOI: 10.1021/acs.orglett.8b02193
Org. Lett. 2018, 20, 5238–5241

remains uncharacterized.[22−27] This potential was also suggested by a phylum-wide genome analysis of cyanobacteria at the PCC collection, in which 80% of the biosynthetic gene clusters (BGCs) could not be assigned to any known compounds.[27] As examples that support the richness of this resource, we recently reported a large and entirely untapped diversity of cyanobacterial peptide natural products with unprecedented posttranslational modifications present in a large percentage of the PCC strains.[28,29] During comparative genome analyses on these pathways, we noticed the presence of a BGC in *Fischerella* sp. PCC 9339 that was unique among all 436 sequenced genomes (status February 2018) and was predicted to encode a hybrid polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS) system (Figure 2). This locus, termed the *arz* cluster, spans 43 kbp and

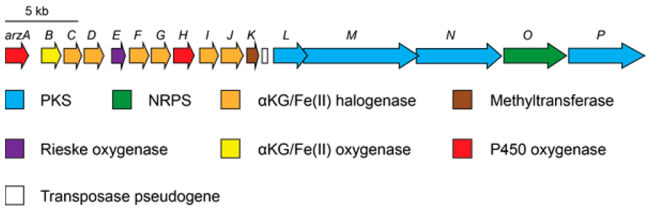



**Figure 2.** *arz* gene cluster identified in the genome of *Fischerella* sp. PCC 9339 with predicted gene functions shown in color.

contains 16 genes. Among these are seven genes for highly similar (>77% mutual amino acid identity) putative αKG enzymes. This is a remarkable number for a single biosynthetic pathway, suggesting a highly unusual natural product.

Many Fe(II)/αKG-dependent enzymes contain a characteristic HXD/E motif in their active site and act as oxygenases, while enzymes in which the D/E is replaced by small aliphatic residues usually catalyze halogenations.[30] Closer analysis of the αKGH enzymes suggested that ArzB is an oxygenase, while the other six homologues would be bona fide halogenases (Figure S1). To our knowledge, this would be the largest number of genes encoding halogenases in a single gene cluster reported so far. Since the genomic scaffold harboring the *arz* genes contained a gap, the missing sequence information was obtained by PCR, confirming physical connection of the genes (Figure 2, Table S1).[31,32] Apart from the PKS, NRPS, and putative halogenase genes in the cluster, we identified three further gene candidates for modifications via radical biochemistry (the putative Rieske oxygenase ArzE and the cytochrome P450 enzymes ArzA and ArzH). A further protein, encoded by *arzK*, resembles SAM-dependent methyltransferases. ArzO, the single NRPS module of the PKS/NRPS assembly line ArzLMNOP, contains an adenylation domain with predicted specificity for cysteine combined with a cyclization (Cy) and an oxidation (Ox) domain, which suggested a thiazole moiety in the metabolite (Figure 3, Table S1).

Chlorinated and brominated compounds, as expected for αKGH biochemistry, provide a convenient handle for MS-guided isolation based on isotope patterns. To identify the product(s) of the *arz* cluster, the slow-growing strain was cultivated for 4−5 months. Analysis of organic cell extracts by liquid chromatography/high-resolution MS (LC/HRMS) revealed multiple peaks in the chromatogram with associated isotopic patterns that were consistent with polychlorination (Figure 3, Figures S5−7). Compound **5** was purified first to

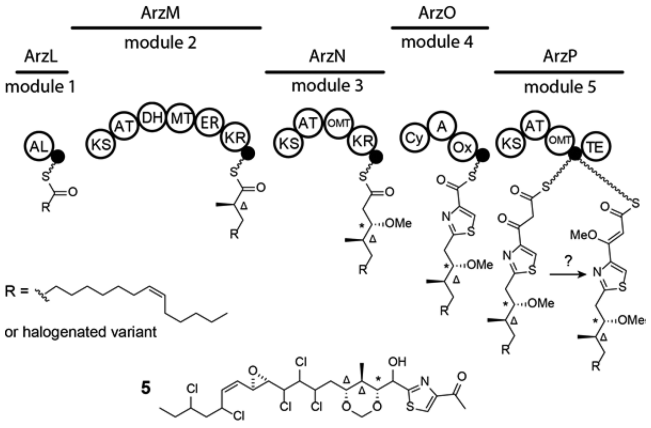



**Figure 3.** *arz* assembly line and proposed aranazole biosynthesis based on the NRPS/PKS architecture and structure of isolated aranazole A (**5**). The absolute configuration of stereocenter indicated with an asterisk (*) was proposed based on sequence analysis of the KR domain of module 3; stereocenters marked with a triangle (Δ) were inferred from NMR data on the relative stereochemistry. Key: AL, acyl-CoA/carrier protein ligase; KS, ketosynthase; AT, acyltransferase; DH, dehydratase; MT, methyltransferase; ER, enoylreductase; KR, ketoreductase; OMT, *O*-methyltransferase; A, adenylation domain; Cy, cyclization domain; Ox, oxidase domain; TE, thioesterase. Small unlabeled circles represent acyl carrier proteins.

yield around 100 μg from 1 L of culture broth. The LC/HRMS data ($m/z$ = 622.0507, calcd 622.0517) suggested $C_{24}H_{32}Cl_5NO_5S$ as molecular formula for **5**. Most of the carbon skeleton could be determined by COSY correlations (Figure 4). The $^1$H NMR and HSQC data showed characteristic signals for an epoxide, which was connected to the carbon backbone by COSY correlations. A double bond was determined as *Z*-configured via proton coupling constants ($^3J_{H16,H17}$ = 10.9 Hz). In addition, we observed signals for a thiazole ring that was localized by HMBC correlations between C-5 and H-6/OH-6. Furthermore, a dioxane ring was identified, for which NOEs suggested a relative stereochemistry with axial protons. HMBC correlations of H-23/23′ to C-7, C-8, and C-9, as well as the lack of COSY correlations, suggested the position of the methylene in the heterocycle. $^1$H NMR and HSQC data were indicative of an additional methyl group and a ketone carbon. HMBC correlations of H-1 with C-2 connected the two groups, and correlations of H-1 with C-3 and H-4 with C-2 connected them to the thiazole ring.

The oxygen atoms were assigned with COSY interactions or characteristic proton and carbon chemical shifts, followed by assigning the chlorine atoms (Figure 4). Determination of the chlorine-containing stereocenters, attempted with HECADE experiments, remained ambiguous. Analysis of the KR domain of module 3 suggests that an (*S*)-configured alcohol is introduced (labeled with an asterisk (*) in Figures 3 and 5, Figure S2).[33] Based on this feature, we propose that the stereocenters at C-7, C-8, and C-9 all have the (*R*) configuration. Taken together, the data suggested the structure of **5**, a polyhalogenated new nonribosomal peptide−polyketide hybrid named aranazole A.

The structure of **5** contains several features that are consistent with the architecture of the *arz* cluster (Figure 3), including a large number of chlorine atoms, an *O*,*O*-methylene unit that matches an O-methyltransferase (OMT) domain in module 3, and the adjacent thiazole moiety that was predicted



DOI: 10.1021/acs.orglett.8b02193
*Org. Lett.* 2018, 20, 5238−5241

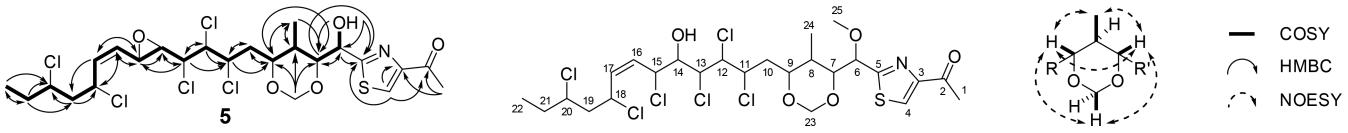



**Figure 4.** Observed couplings of aranazoles A.

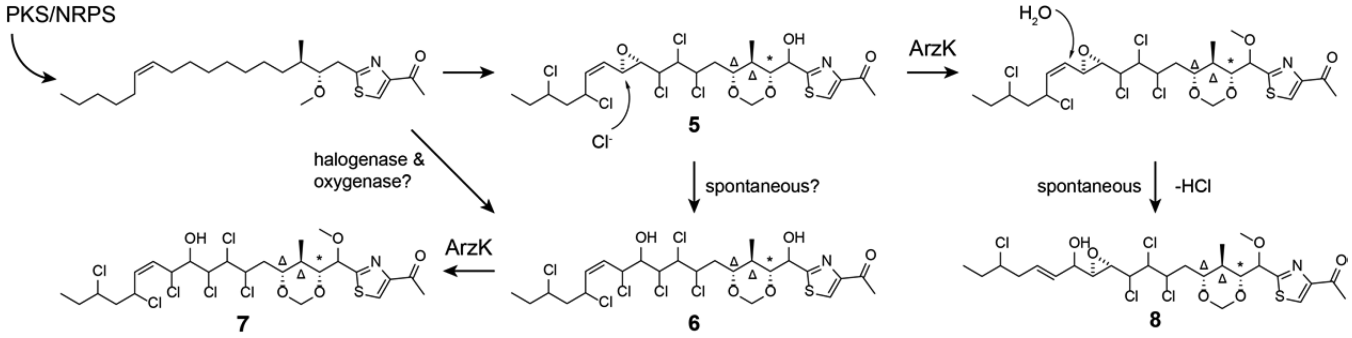



**Figure 5.** Proposed downstream modifications of aranazoles based on the isolated congeners aranazoles B–D (**6**–**8**). The sixth chloride might be introduced either by attack of a chloride on an epoxide or by halogenation and hydroxylation. The absolute configurations of stereocenters indicated with an asterisk (*) were proposed based on sequence analysis of the KR domain of module 3, and stereocenters marked with a triangle (Δ) were inferred by the determination of the relative stereochemistry.

from module 4. However, the number of αKGH genes in the *arz* cluster suggested an even higher number of chlorine atoms than are present in **5**. Indeed, LC–MS data suggested additional polychlorinated compounds, of which three congeners were isolated in sufficiently high yields for NMR analysis. Compounds **6** and **7** contained an additional chlorine atom as compared to **1**. They differ in the presence of an *O*-methyl group, which in **7** was located to C-6 by HMBC correlations with H-25. Compound **8** features the same *O*-methylated Eastern moiety as **7**, while the Western part contains only four chlorines, an additional hydroxyl group at C-16, located by COSY correlations, and a shifted *E*-configured double bond ($^3J_{H17,H18}$ = 15.5 Hz, Figure 4). NOESY and HECADE experiments with the congeners did not provide additional insight into the unassigned stereocenters.

The characterized compounds differ in the presence or absence of an *O*-methyl group at C-6, likely introduced by the methyltransferase ArzK, and in diverse chlorine and oxygen patterns of the Western moiety. For **5** and **8** containing five or less chlorines, this moiety contains an epoxide. The additional chlorine at C-15 in **6** and **7** might therefore either be introduced by nucleophilic epoxide opening or result from an epoxidase side reaction, perhaps involving one of the αKGHs. Compound **8** differs from the other structures by a shifted double bond with *E* configuration. This double bond and modified oxygen and chlorine pattern could be the result of an $S_N2'$ reaction by attack of water at C-16 and an allylic chloride moiety as a leaving group, as present in the other congeners (Figure 5). A monounsaturated fatty acid derived starter, such as myristoleic acid loaded by ArzL, is a likely origin of the polychlorinated Western portion. Module 2 would then elongate the starter with a methyl-branched building block. Since the domain architecture of this module suggests that the β-carbon is completely reduced, the 1,3-dioxane moiety might be generated by post-PKS hydroxylation by one of the four oxygenases. Interestingly, the last PKS module carrying an OMT domain is architecturally identical to the terminal module of the haliangicin PKS.[34] This module generates a β-ketoacyl moiety and converts it via *O*-methylation to the corresponding methyl enolether. It is unknown whether the OMT domain of the *arz* module is inactive or if the enolether is first generated and then reacts to the ketones **5**–**8** (Figure 3, Figure S3).

Aranazole A was tested for bioactivity against HeLa cells, *E. coli*, *Salmonella typhimurium*, bacilli, *Saccharomyces cerevisiae*, and the lymphocytic choriomeningitis virus, but no significant activity could be detected. In contrast, chlorosulfolipids exhibit diverse antimicrobial, cytotoxic, and protein kinase-inhibiting activities. Based on the architecture of the terminal PKS module, we considered whether the true enzymatic end product of the *arz* pathway might be the carboxylated methyl enolether discussed above. Such a compound is expected to be highly reactive and might be nonenzymatically degraded to the observed compounds by enol ether hydrolysis and decarboxylation. In MS analyses of extracts from the aranazole producer, we did not obtain evidence for the presence of enol ethers or β-ketoacids. However, the months-long cultivation times required to produce submilligram amounts of **5**–**8** might exceed the half-life of such unstable congeners. It would be interesting to see how synthetic versions of the enolethers behave in bioassays.

In conclusion, we identified four structurally unusual polychlorinated compounds from a cyanobacterial PKS/NRPS pathway involving an unprecedented number of halogenase homologues. The results contribute to the appreciation of cyanobacteria as a rich and underexplored resource of novel chemistry. The structural similarities of the halogenated moieties in aranazoles and chlorosulfolipids raise the question of whether similar enzymatic processes are involved in their production and whether cyanobacterial producers exist for the latter compounds.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.8b02193.



DOI: 10.1021/acs.orglett.8b02193
*Org. Lett.* 2018, 20, 5238–5241

Experimental procedures, supplementary figures and tables, characterization data, and NMR spectra of all new compounds (PDF)

## AUTHOR INFORMATION


### Corresponding Author

*E-mail: jpiel@ethz.ch.

### ORCID

Jörn Piel: 0000-0002-2282-8154


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank U. Greczmiel (ETH Zurich) for the viral bioactivity assays. J.P. acknowledges funding by the SNF (IZLSZ3_149025) and the Helmut Horten Foundation. R.U. was the recipient of an Alexander von Humboldt postdoctoral fellowship. N. Sassoon (PCC, Institut Pasteur) is thanked for helping with cyanobacterial cultures and biomass production. The PCC is funded by the Institut Pasteur.

## REFERENCES


(1) Harris, C. M.; Kannan, R.; Kopecka, H.; Harris, T. M. *J. Am. Chem. Soc.* **1985**, *107*, 6652.
(2) Groll, M.; Huber, R.; Potts, B. C. M. *J. Am. Chem. Soc.* **2006**, *128*, 5136.
(3) Bianco, A. C.; Salvatore, D.; Gereben, B.; Berry, M. J.; Larsen, P. R. *Endocr. Rev.* **2002**, *23*, 38.
(4) Kazlauskas, R.; Lidgard, R. O.; Wells, R. J. *Tetrahedron Lett.* **1977**, *18*, 3183.
(5) Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V. A.; Willis, C. L.; Sherman, D. H. *Gene* **2002**, *296*, 235.
(6) Unson, M. D.; Faulkner, D. J. *Experientia* **1993**, *49*, 349.
(7) Flatt, P.; Gautschi, J.; Thacker, R.; Musafija-Girt, M.; Crews, P.; Gerwick, W. *Mar. Biol.* **2005**, *147*, 761.
(8) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. *Nat. Chem. Biol.* **2014**, *10*, 640.
(9) Wynands, I.; van Pee, K. H. *FEMS Microbiol. Lett.* **2004**, *237*, 363.
(10) Elovson, J.; Vagelos, P. R. *Proc. Natl. Acad. Sci. U. S. A.* **1969**, *62*, 957.
(11) Ciminiello, P.; Dell'Aversano, C.; Fattorusso, E.; Forino, M.; Magno, S.; Di Rosa, M.; Ianaro, A.; Poletti, R. *J. Am. Chem. Soc.* **2002**, *124*, 13114.
(12) Nilewski, C.; Carreira, E. M. *Eur. J. Org. Chem.* **2012**, *2012*, 1685.
(13) Neumann, C. S.; Fujimori, D. G.; Walsh, C. T. *Chem. Biol.* **2008**, *15*, 99.
(14) Hager, L. P.; Morris, D. R.; Brown, F. S.; Eberwein, H. *J. Biol. Chem.* **1966**, *241*, 1769.
(15) Sundaramoorthy, M.; Terner, J.; Poulos, T. L. *Structure* **1995**, *3*, 1367.
(16) Butler, A.; Carter-Franklin, J. N. *Nat. Prod. Rep.* **2004**, *21*, 180.
(17) Dong, C. J.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pée, K. H.; Naismith, J. H. *Science* **2005**, *309*, 2216.
(18) Koehntop, K. D.; Emerson, J. P.; Que, L., Jr. *JBIC, J. Biol. Inorg. Chem.* **2005**, *10*, 87.
(19) Galonic, D. P.; Vaillancourt, F. H.; Walsh, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 3900.
(20) O'Hagan, D.; Schaffrath, C.; Cobb, S. L.; Hamilton, J. T.; Murphy, C. D. *Nature* **2002**, *416*, 279.
(21) Dong, C.; Huang, F.; Deng, H.; Schaffrath, C.; Spencer, J. B.; O'Hagan, D.; Naismith, J. H. *Nature* **2004**, *427*, 561.
(22) Engene, N.; Rottacker, E. C.; Kastovsky, J.; Byrum, T.; Choi, H.; Ellisman, M. H.; Komárek, J.; Gerwick, W. H. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1171.
(23) Jiménez, J. I.; Scheuer, P. J. *J. Nat. Prod.* **2001**, *64*, 200.
(24) Cardellina, J. H.; Marner, F. J.; Moore, R. E. *Science* **1979**, *204*, 193.
(25) Moore, R. E.; Cheuk, C.; Patterson, G. M. L. *J. Am. Chem. Soc.* **1984**, *106*, 6456.
(26) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. *J. Am. Chem. Soc.* **2001**, *123*, 5418.
(27) Calteau, A.; Fewer, D. P.; Latifi, A.; Coursin, T.; Laurent, T.; Jokela, J.; Kerfeld, C. A.; Sivonen, K.; Piel, J.; Gugger, M. *BMC Genomics* **2014**, *15*, 977.
(28) Morinaka, B. I.; Lakis, E.; Verest, M.; Helf, M. J.; Scalvenzi, T.; Vagstad, A. L.; Sims, J.; Sunagawa, S.; Gugger, M.; Piel, J. *Science* **2018**, *359*, 779.
(29) Morinaka, B. I.; Vagstad, A. L.; Helf, M. J.; Gugger, M.; Kegler, C.; Freeman, M. F.; Bode, H. B.; Piel, J. *Angew. Chem., Int. Ed.* **2014**, *53*, 8503.
(30) Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. *Nat. Chem. Biol.* **2016**, *12*, 636.
(31) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. *Nucleic Acids Res.* **2005**, *33*, 5799.
(32) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res.* **2011**, *39*, W362.
(33) Caffrey, P. *ChemBioChem* **2003**, *4*, 654.
(34) Sun, Y.; Feng, Z.; Tomura, T.; Suzuki, A.; Miyano, S.; Tsuge, T.; Mori, H.; Suh, J. W.; Iizuka, T.; Fudou, R.; Ojika, M. *Sci. Rep.* **2016**, *6*, 22091.




DOI: 10.1021/acs.orglett.8b02193
*Org. Lett.* 2018, 20, 5238−5241