Organic LETTERS

Letter

pubs.acs.org/OrgLett

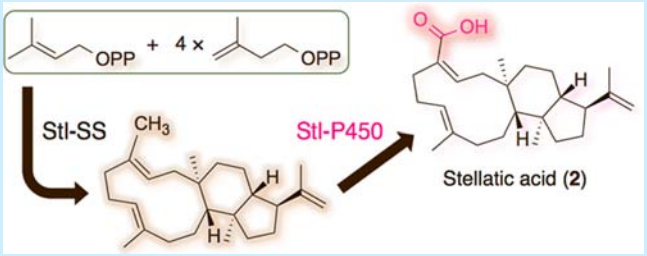
Molecular Basis for Stellatic Acid Biosynthesis: A Genome Mining Approach for Discovery of Sesterterpene Synthases

Yudai Matsuda,[†] Takaaki Mitsuhashi,[†] Zhiyang Quan, and Ikuro Abe*

Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Supporting Information

ABSTRACT: The search for a new sesterterpene synthase in the genome of *Emericella variecolor*, which reportedly produces diverse sesterterpenoids, is described. One gene product (a chimeric protein with prenyltransferase and terpene cyclase domains) led to the synthesis of a novel tricyclic sesterterpene, stellata-2,6,19-triene (**1**), from DMAPP and IPP, and the hydrocarbon was further transformed into stellatic acid (**2**) by cytochrome P450 monooxygenase encoded by the gene adjacent to the sesterterpene synthase gene.



Terpenoids, derived from five-carbon isoprene units, are the largest class of natural products, and over 50000 terpenoid compounds with diverse carbon skeletons have been reported to date. Their structural diversity is primarily due to the existence of a variety of terpenoid cyclases, which can synthesize (poly)cyclic molecules with multiple chiral centers from a linear polyisoprenoid substrate in a single reaction. Thus, numerous terpenoid cyclases have been identified and characterized to understand the basis for the production of such diverse molecular architectures.[1,2]

The C_{25} sesterterpenoids derived from five isoprene units, however, are relatively rare species among terpenoid natural products,[3] and the genetic and molecular bases for their biosyntheses have remained elusive for a long time. The first sesterterpene synthase, AcOS, was identified from the genome of the fungus *Aspergillus clavatus* in 2013 and found to be responsible for the biosynthesis of ophiobolin F.[4] Interestingly, AcOS is a fusion protein containing prenyltransferase (PT) and terpene cyclase (TC) domains. In AcOS, the PT domain generates the linear substrate geranylfarnesyl pyrophosphate (GFPP) from dimethylallyl pyrophosphate (DMAPP) and isopentenyl pyrophosphate (IPP), and the TC domain cyclizes GFPP into ophiobolin F. Similar bifunctionality was initially found and is also employed in the fungal diterpene synthases, such as fusicoccadiene synthase and phomopsene synthase.[5,6] Subsequently, the second example of a sesterterpene synthase, Bcl-TS, was discovered from the bacterium *Bacillus clausii*, which transforms GFPP into the linear sesterterpene hydrocarbon β-geranylfarnesene.[7,8] At present, these two enzymes are the only known sesterterpene synthases. Since our knowledge about sesterterpene biosynthesis is very limited, the discovery and functional analysis of other enzymes catalyzing sesterterpene biosynthesis would facilitate our understanding of how such structurally diverse sesterterpenoids are produced and accelerate the identification of novel enzymes.

Aspergillus variecolor and its perfect state *Emericella variecolor* reportedly produce many sesterterpenoids such as variecolin, ophiobolins, astellatol, variculanol, and stellatic acid (Figure 1),[9−14] and therefore, the fungi were expected to be rich sources of sesterterpene synthases. We previously performed the genome sequencing analysis of *E. variecolor* NBRC 32302 in the course of our biosynthetic study on the meroterpenoid anditomin.[15] In this study, we searched for novel sesterterpene synthase genes in the *E. variecolor* genome and identified the gene involved in the biosynthesis of stellatic acid. We also analyzed the function of the cytochrome P450 monooxygenase gene adjacent to the sesterterpene synthase gene and elucidated that the P450 oxidizes the sesterterpene hydrocarbon into stellatic acid.

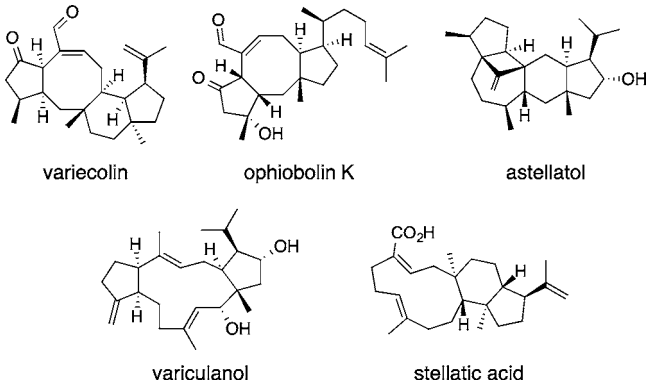


Figure 1. Sesterterpenoids isolated from *A. variecolor* and *E. variecolor*.

By searching for a gene encoding a homologous protein to AcOS, nine possible sesterterpene or diterpene synthase genes, all encoding chimeric proteins with PT and TC domains, were found in the *E. variecolor* genome. Among these candidates, we focused on the gene encoding the product exhibiting the

Received: August 20, 2015
Published: September 9, 2015

DOI: 10.1021/acs.orglett.5b02404
Org. Lett. 2015, 17, 4644−4647

highest identity with AcOS (~40% identity). The characteristic DDXXD motif[16] for binding a trinuclear Mg^{2+} cluster is conserved in both the N-terminal TC and C-terminal PT domains as ^{92}DDVTD^{96} and ^{473}DDVED^{477}, respectively. In contrast, another Mg^{2+}-binding motif in the TC domain is not completely conserved, since the asparagine/aspartate residue in the NSE/DTE motif[16] is substituted with histidine, as ^{223}**H**DLAS**W**EK**E**^{231}. The recently found novel effector triad,[17] consisting of arginine, aspartate, and glycine, was also present in the TC domain, as ^{179}**R**IQ**D**I**G**^{184} (Figure S1).

To analyze the function of the putative sesterterpene synthase gene, it was heterologously expressed in *Aspergillus oryzae* NSAR1, a quadruple auxotrophic mutant strain (*niaD*$^-$, *sC*$^-$, Δ*argB*, *adeA*$^-$)[18] that has been successfully utilized for biosynthetic studies on fungal natural products.[15,19,20] The transformant was cultivated in the inductive medium, and the mycelial extract was analyzed by GC−MS. As a result, a single new product **1** with *m*/*z* 340 $[M]^+$ was detected (Figure 2A and Figure S2), and therefore, it was expected that the new metabolite is a sesterterpene hydrocarbon. This observation indicated that the N/D to H substitution in the NSE/DTE motif did not cause the inactivation of the enzyme, although a previous study showed that a diterpene synthase-like protein from *Salvia sclarea*, with a mutation similar to that of histidine in the motif, lacked terpene synthase activity in both in vitro and in vivo reconstitution experiments.[21]

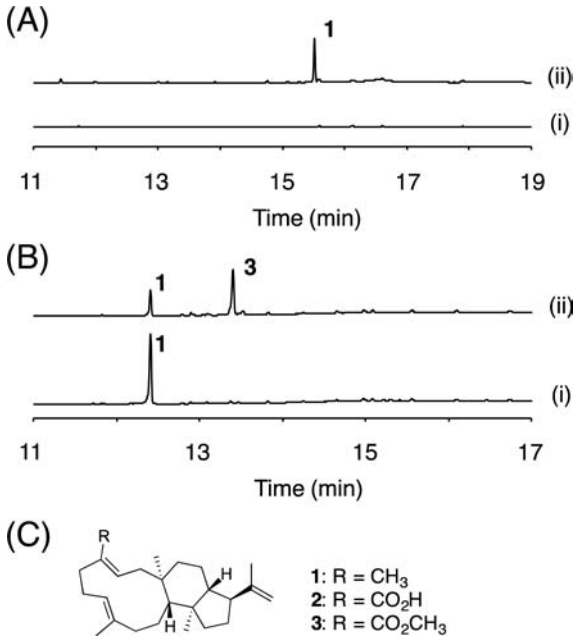


Figure 2. Functional analyses of Stl-SS and Stl-P450. (A) GC−MS chromatograms of mycelial extracts from *A. oryzae* transformants harboring (i) only empty vector; (ii) Stl-SS. (B) GC−MS chromatograms of mycelial extracts methylated by TMS-diazomethane from *A. oryzae* transformants harboring (i) Stl-SS and (ii) Stl-SS and Stl-P450. (C) Structures of compounds **1**−**3**. Note that inconsistent retention times for **1** in (A) and (B) are due to the different analytical conditions.

The molecular formula of **1** was established to be $C_{25}H_{40}$ by HR-FAB-MS, indicating six degrees of unsaturation. The complete structure of **1** was further revealed by NMR analyses (Tables S2 and S3 and Figures S5−S16). The ^{13}C NMR spectrum (in C_6D_6) revealed 25 signals, which consist of five singlet methyls, 10 methylenes including one exomethylene (C-24), two olefinic (C-2 and C-6), and three aliphatic methines, and three sp^2 (C-3, C-7, and C-19) and two sp^3 quaternary carbons. Thus, the results indicated that **1** possesses a tricyclic chemical structure. Interpretation of the ^{1}H−^{1}H COSY spectrum then revealed the spin systems of H-1/H-2, H-4/H-5/H-6, H-8/H-9, H-12/H-13, and H-16/H-17/H-18 (Figure 3A). Additionally, HMBC correlations from the singlet methyl

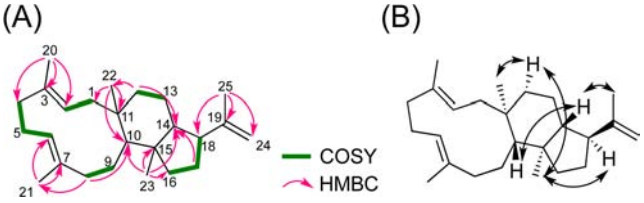


Figure 3. (A) ^{1}H−^{1}H COSY and key HMBC correlations in **1**. (B) Key NOESY correlations in **1**.

signals were observed, as follows: H_3-20 (δ 1.46) to C-2, C-3, and C-4; H_3-21 (δ 1.44) to C-6 and C-7; H_3-22 (δ 0.85) to C-1, C-10, and C-11; H_3-23 (δ 0.78) to C-10, C-14, C-15, and C-16; and H_3-25 (δ 1.67) to C-18, C-19, and C-24. These cross peaks confirmed the connections of C-3 to C-2, C-4, and C-20, C-7 to C-6 and C-21, C-11 to C-1, C-10, and C-22, C-15 to C-10, C-14, C-16, and C-23, and C-19 to C-18, C-24, and C-25 (Figure 3A). Furthermore, the HMBC correlations of H-8 (δ 2.00) to C-21 and C-10, H-12 (δ 1.69) to C-22, H-12 (δ 1.27) to C-14, and H-17 (δ 1.87) to C-14 established the connections of C-7 to C-8, C-9 to C-10, C-11 to C-12, C-13 to C-14, and C-14 to C-18, respectively (Figure 3A). The geometries of the two double bonds were both determined to be *E*, based on the higher field chemical shifts of C-20 and C-21 (δ 16.8 and 16.9, respectively), which provided the planar structure of **1** consisting of an 11−6−5 membered tricyclic ring system (Figure 2C). The relative configuration of **1** was determined by NOSEY correlations (Figure 3B): H-12α (δ 1.27) and H-22/H-23, H-18 and H-23, and H-14 and H-10/H-25 (obtained from the NOESY spectrum in $CDCl_3$). The carbon skeleton of **1**, including the relative configurations, completely matched that of stellatic acid (**2**),[13] and therefore, we reasoned that **1** is the biosynthetic precursor of **2** and has the same absolute configurations as those of **2**. The hydrocarbon **1** was hereby named stellata-2,6,19-triene, and the sesterterpene synthase was designated as Stl-SS (*s*esterterpene *s*ynthase).

To obtain further insight into the Stl-SS-catalyzed reaction, an in vitro assay with recombinant Stl-SS heterologously expressed in *Escherichia coli* was performed. The reaction proceeded to yield **1** in the presence of Mg^{2+}, IPP, and DMAPP, confirming that the single enzyme not only synthesizes the C_{25} linear substrate GFPP from five-carbon starter units but also cyclizes GFPP into **1**, as seen in the reaction with another sesterterpene synthase, AcOS (Figure S4).[4]

As the sesterterpene hydrocarbon **1** appeared to be the biosynthetic intermediate of **2**, we then decided to elucidate the pathway leading to **2** from **1**. Investigation of the flanking regions of the Stl-SS gene revealed a gene encoding a cytochrome P450 monooxygenase located next to the Stl-SS gene, therefore suggesting that the P450 is involved in the stellatic acid biosynthesis (DDBJ/EMBL/GenBank accession no. LC073704). To determine the function of the P450, which

DOI: 10.1021/acs.orglett.5b02404
Org. Lett. 2015, 17, 4644−4647

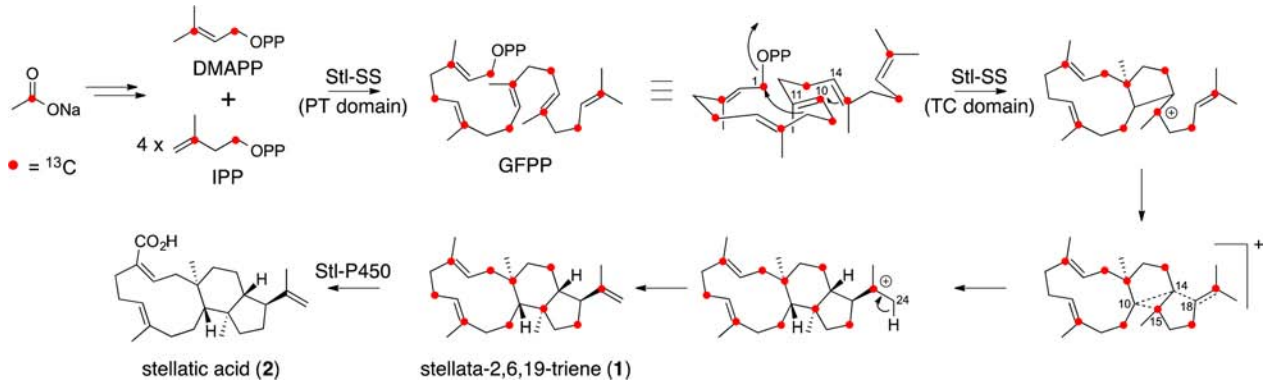


Figure 4. Proposed biosynthetic mechanism leading to stellatic acid (**2**).

was designated as Stl-P450, the coexpression system of Stl-SS and Stl-P450 was constructed using *A. oryzae*. The mycelial extract of the transformant was methylated with TMS-diazomethane prior to the GC−MS analysis. The chromatogram revealed one new peak **3** with *m*/*z* 384 $[M]^+$ (Figure 2B and Figure S3), which corresponds to the molecular weight of methyl stellatate. After the large-scale cultivation and isolation of the new metabolite, we confirmed that the Stl-P450 product is indeed identical to stellatic acid, by MS and NMR analyses as well as a comparison of its specific rotation, $[\alpha]^{24}_D$ + 14.6 (*c* 0.83, $CHCl_3$), with the reported value[13] (Table S4 and Figures S17 and S18) and thus established the molecular basis for stellatic acid biosynthesis (Figure 4). Stl-P450 is multifunctional and catalyzes three successive oxidation reactions on the C-20 methyl group to generate the carboxylic acid, as seen in the artemisinin pathway in which the P450 CYP71AV1 converts the sesquiterpene amorphadiene into artemisinic acid.[22] In terms of bioactivity, **2** exhibited moderate cytotoxicity against P388 murine leukemia cells, with an IC_{50} value of 24.7 μM.

To obtain insight into the Stl-SS-catalyzed cyclization, an isotope-incorporation experiment was performed using [1-^{13}C]-sodium acetate, which specifically labels the C-1 and C-3 positions of DMAPP and IPP (Figure 4). The resultant metabolite, stellata-2,6,19-triene (**1**), was thus labeled with ^{13}C at 10 distinct positions with ca. 15% ^{13}C incorporation: C-1, −3, −5, −7, −9, −11, −13, −15, −17, and −19 (Figure 4 and Figure S19). This observation indicated that the biogenesis of **1** follows the isoprene rule,[23] and provided a plausible mechanism for the cyclization. The reaction is initiated by the elimination of the pyrophosphate group to generate an allylic cation, which is followed by the sequential cyclization at C-1/C-11 and C-10/C-14 to yield a tertiary carbocationic species with an 11−5 fused bicyclic ring system. The cyclopentylcarbinyl carbocation would then undergo the ring expansion, with participation by the terminal π-bond, to generate the sesterterpene scaffold with an 11−6−5 fused tricyclic ring system:[24] C−C bond migration occurs to form the C-10/C-15 connection, and a C−C bond at C-14/C-18 is newly formed. Finally, the deprotonation from the C-24 methyl group completes the reaction to afford **1** (Figure 4). The cyclization reaction by Stl-SS thus differs from that by AcOS in that the carbon skeleton of the cyclized moiety of the AcOS product is similar to that of the diterpene fusicoccadiene. Thus, the terminal isoprene unit of GFPP does not participate in the AcOS-catalyzed cyclization reaction.[4] In contrast, the reaction by Stl-SS utilizes the terminal dimethylallyl group of GFPP for the polyene cyclization, and therefore, the scaffold with an 11−6−5 fused tricyclic ring system cannot be found in the diterpenoid natural products.

In conclusion, we identified the bifunctional sesterterpene synthase Stl-SS, responsible for the biosynthesis of the new tricyclic sesterterpene stellata-2,6,19-triene (**1**) with the unique 11−6−5 fused ring system, by a genome mining approach. As mentioned above, *E. variecolor* NBRC 32302 possesses eight other putative sesterterpene/diterpene synthases, and therefore, functional analyses of these enzymes could provide the scaffolds of other known sesterterpenoids or completely novel molecular architectures. Since homologous proteins of Stl-SS and AcOS are present in many fungal genomes, functional characterizations of these cryptic terpene synthases will also be quite interesting. Furthermore, we elucidated the role of the P450 monooxygenase encoded by the gene adjacent to the Stl-SS gene and thus determined the complete biosynthetic pathway of stellatic acid (**2**). Considering the structures of the known fungal sesterterpenoids, their biosyntheses appear to require a relatively small number of tailoring enzymes, as we demonstrated with the stellatic acid pathway. Thus, it would be easy to perform not only the genome mining of a sesterterpene gene but also the full reconstitution of the biosynthetic pathway to obtain the natural product derived from a cryptic gene cluster. We believe that our approach would facilitate the discovery of numerous new sesterterpene synthases and probably biologically active useful molecules for drug discovery.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details, supplementary figures, and spectral data. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.5b02404.

Experimental details, supplementary figures, and spectral data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: abei@mol.f.u-tokyo.ac.jp.

Author Contributions

†Y.M. and T.M. contributed equally.

Notes

The authors declare no competing financial interest.

DOI: 10.1021/acs.orglett.5b02404
Org. Lett. 2015, 17, 4644−4647

■ ACKNOWLEDGMENTS

We thank Prof. K. Gomi (Tohoku University) and Prof. K. Kitamoto (The University of Tokyo) for kindly providing the expression vectors and the fungal strain. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and the Kobayashi International Scholarship Foundation.

■ REFERENCES

(1) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412.
(2) Gao, Y.; Honzatko, R. B.; Peters, R. J. *Nat. Prod. Rep.* **2012**, *29*, 1153−1175.
(3) Wang, L.; Yang, B.; Lin, X.-P.; Zhou, X.-F.; Liu, Y. *Nat. Prod. Rep.* **2013**, *30*, 455−473.
(4) Chiba, R.; Minami, A.; Gomi, K.; Oikawa, H. *Org. Lett.* **2013**, *15*, 594−597.
(5) Toyomasu, T.; Tsukahara, M.; Kaneko, A.; Niida, R.; Mitsuhashi, W.; Dairi, T.; Kato, N.; Sassa, T. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 3084−3088.
(6) Toyomasu, T.; Kaneko, A.; Tokiwano, T.; Kanno, Y.; Kanno, Y.; Niida, R.; Miura, S.; Nishioka, T.; Ikeda, C.; Mitsuhashi, W. *J. Org. Chem.* **2009**, *74*, 1541−1548.
(7) Sato, T.; Yamaga, H.; Kashima, S.; Murata, Y.; Shinada, T.; Nakano, C.; Hoshino, T. *ChemBioChem* **2013**, *14*, 822−825.
(8) Ueda, D.; Yamaga, H.; Murakami, M.; Totsuka, Y.; Shinada, T.; Sato, T. *ChemBioChem* **2015**, *16*, 1371−1377.
(9) Hensens, O. D.; Zink, D.; Williamson, J. M.; Lotti, V. J.; Chang, R. S.; Goetz, M. A. *J. Org. Chem.* **1991**, *56*, 3399−3403.
(10) Wei, H.; Itoh, T.; Kinoshita, M.; Nakai, Y.; Kurotaki, M.; Kobayashi, M. *Tetrahedron* **2004**, *60*, 6015−6019.
(11) Sadler, I. H.; Simpson, T. J. *J. Chem. Soc., Chem. Commun.* **1989**, 1602−1604.
(12) Singh, S. B.; Reamer, R.; Zink, D.; Schmatz, D.; Dombrowski, A.; Goetz, M. *J. Org. Chem.* **1991**, *56*, 5618−5622.
(13) Qureshi, I. H.; Husain, S. A.; Noorani, R.; Murtaza, N.; Iitaka, Y.; Iwasaki, S.; Okuda, S. *Tetrahedron Lett.* **1980**, *21*, 1961−1962.
(14) Liangsakul, J.; Pornpakakul, S.; Sangvichien, E.; Muangsin, N.; Sihanonth, P. *Tetrahedron Lett.* **2011**, *52*, 6427−6430.
(15) Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I. *J. Am. Chem. Soc.* **2014**, *136*, 15326−15336.
(16) Shishova, E. Y.; Di Costanzo, L.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2007**, *46*, 1941−1951.
(17) Baer, P.; Rabe, P.; Fischer, K.; Citron, C. A.; Klapschinski, T. A.; Groll, M.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2014**, *53*, 7652−7656.
(18) Jin, F. J.; Maruyama, J.; Juvvadi, P. R.; Arioka, M.; Kitamoto, K. *FEMS Microbiol. Lett.* **2004**, *239*, 79−85.
(19) Matsuda, Y.; Iwabuchi, T.; Wakimoto, T.; Awakawa, T.; Abe, I. *J. Am. Chem. Soc.* **2015**, *137*, 3393−3401.
(20) Liu, C.; Tagami, K.; Minami, A.; Matsumoto, T.; Frisvad, J. C.; Suzuki, H.; Ishikawa, J.; Gomi, K.; Oikawa, H. *Angew. Chem., Int. Ed.* **2015**, *54*, 5748−5752.
(21) Caniard, A.; Zerbe, P.; Legrand, S.; Cohade, A.; Valot, N.; Magnard, J.-L.; Bohlmann, J.; Legendre, L. *BMC Plant Biol.* **2012**, *12*, 119.
(22) Ro, D. K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, M. C. Y.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D. *Nature* **2006**, *440*, 940−943.
(23) Ruzicka, L. *Experientia* **1953**, *9*, 357−367.
(24) Tantillo, D. J. *Nat. Prod. Rep.* **2011**, *28*, 1035−1053.

DOI: 10.1021/acs.orglett.5b02404
Org. Lett. 2015, 17, 4644−4647