Natural Products

Deutsche Ausgabe: DOI: 10.1002/ange.201601448
Internationale Ausgabe: DOI: 10.1002/anie.201601448

Astellifadiene: Structure Determination by NMR Spectroscopy and Crystalline Sponge Method, and Elucidation of its Biosynthesis

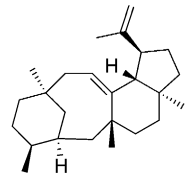
Yudai Matsuda[+], Takaaki Mitsuhashi[+], Shoukou Lee[+], Manabu Hoshino, Takahiro Mori, Masahiro Okada, Huiping Zhang, Fumiaki Hayashi, Makoto Fujita, and Ikuro Abe**

Abstract: *Genome mining of a terpene synthase gene from Emericella variecolor NBRC 32302 and its functional expression in Aspergillus oryzae led to the production of the new sesterterpene hydrocarbon, astellifadiene (**1**), having a 6-8-6-5-fused ring system. The structure of **1** was initially investigated by extensive NMR analyses, and was further confirmed by the crystalline sponge method, which established the absolute structure of **1** and demonstrated the usefulness of the method in the structure determination of complex hydrocarbon natural products. Furthermore, the biosynthesis of **1** was proposed on the basis of isotope-incorporation experiments performed both in vivo and in vitro. The cyclization of GFPP involves a protonation-initiated second cyclization sequence, 1,2-alkyl migration, and 1,5-hydride shift to generate the novel scaffold of **1**.*

The last decade witnessed an unprecedented accumulation of genome sequences from many microorganisms, thus demonstrating the microbial potential to produce many more compounds than previously expected. The availability of these genomic data has allowed not only the rapid identification of biosynthetic genes for known compounds but also the genome-based discovery of new natural products. Thus, numerous reports have described the methodologies and applications for the isolation of cryptic molecules.[1]

Along with the acquisition of many natural products, the development of advanced technologies that allow the rapid structure determination of natural molecules is required. NMR- and MS-based characterizations can be performed routinely, and often provide the planar structure and relative configuration of a compound with high reliability.[2] However, the determination of the absolute stereochemistry can still remain challenging.[3] The crystalline sponge method is an X-ray technique which enables the crystallographic analysis of even oily compounds in minute quantities, and thus has served well in the determinations of the absolute structures of natural products and their derivatives.[4]

Meanwhile, genome mining and heterologous expression of fungal di- or sesterterpene synthases have recently led to the discovery of new terpenoids.[5] These terpene synthases consist of a prenyltransferase (PT) domain and a terpene cyclase (TC) domain. The PT synthesizes C_{20} geranylgeranyl pyrophosphate or C_{25} geranylfarnesyl pyrophosphate (GFPP), while the TC transforms the polyprenyl substrates into cyclized products. Importantly, the terpenes obtained through this genome mining approach include compounds with novel scaffolds.[5a,c,d] However, the structure determination of complex terpenoids is often difficult with only conventional methodologies, particularly when the oxidation level of the terpenoid is low. Herein we report the production and rapid structure characterization of the novel sesterterpene astellifadiene (**1**), and the use of NMR analyses combined with the crystalline sponge method facilitated the unambiguous determination of its structure. Additionally, the mechanistic study of the formation of **1** revealed a unique cyclization mode.



astellifadiene (**1**)

To obtain a new terpene compound, we focused on *Emericella variecolor* NBRC 32302 as a potential source of terpene synthase genes, since we previously characterized two new bifunctional terpene synthases from the fungus,[5b,d] and the strain is rich in possible uncharacterized terpene synthase genes. Among several candidates, we selected a gene encoding a 716 amino-acid protein (see Figure S1 in the Supporting Information). According to the previous successful functional analyses of this class of terpene synthases,[5b–d,6] the gene was heterologously expressed in *Aspergillus oryzae* NSAR1[7] (Figure 1 A). The mycelial extract of the transformant harboring the gene was analyzed by GC-MS, and consequently, the new metabolite **1** with m/z 340 $[M]^+$ was observed (Figures 1 B ; see Figure S3A). We speculated that **1** is a new sesterterpene hydrocarbon based on its mass spectrum, and **1** was thus isolated for structure determination.

The molecular formula of **1** was established as $C_{25}H_{40}$ by HR-MS, thus indicating six degrees of unsaturation. The structure of **1** was investigated by NMR spectroscopy (see Figures S9–S20 and Tables S2 and S3). The ^{13}C NMR spectrum revealed 25 signals including four olefinic carbon atoms, thus suggesting a tetracyclic backbone for **1**. A thorough analysis of the ^{1}H-^{1}H COSY, HMBC, and NOESY correlations established the planar structure of **1** as an unprece-

---

[*] Dr. Y. Matsuda,[+] T. Mitsuhashi,[+] Dr. T. Mori, Dr. M. Okada, Prof. Dr. I. Abe
Graduate School of Pharmaceutical Sciences, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033 (Japan)
E-mail: abei@mol.f.u-tokyo.ac.jp

S. Lee,[+] Dr. M. Hoshino, Prof. Dr. M. Fujita
Graduate School of Engineering, The University of Tokyo, and JST-ACCEL, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656 (Japan)
E-mail: mfujita@appchem.t.u-tokyo.ac.jp

Dr. H. Zhang, Dr. F. Hayashi
RIKEN Center for Life Science Technologies
Yokohama, Kanagawa 230-0045 (Japan)

[+] These authors contributed equally to this work.

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201601448.

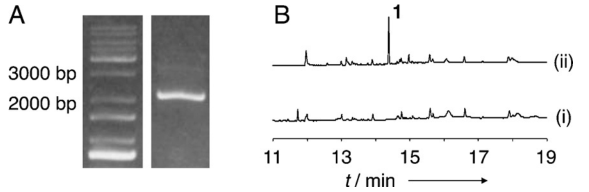


Figure 1. A) Confirmation of the expression of *EvAS* in the *A. oryzae* transformant. B) GC-MS chromatograms of mycelial extracts derived from *A. oryzae* transformants harboring (i) only empty vector and (ii) *EvAS*.

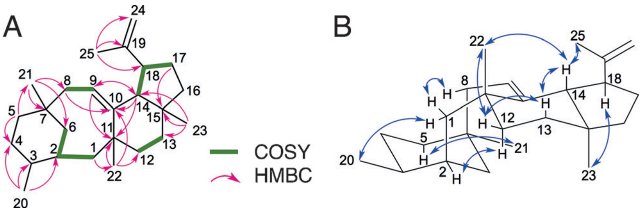


Figure 2. A) ^{1}H-^{1}H COSY and key HMBC correlations in **1**. B) Key NOESY correlations in **1**.

dented 6-8-6-5-membered tetracyclic ring system, and its relative stereochemistry as 2*S**, 3*S**, 7*R**, 11*R**, 14*R**, 15*S**, 18*S** (Figure 2, and see Supporting Information for the detailed scheme).

Next, we aimed to establish the absolute configuration of **1**, but the unfunctionalized nature of **1** hampered the rapid determination of its complete structure by conventional techniques. The X-ray Bijvoet method provides direct access to information about the chirality of molecules, but **1** exists as an oil and is not suitable for common diffraction studies. To the best of our knowledge, the crystalline sponge method is the only way to directly determine the absolute configuration of a compound without crystallization.[4,8] Thus, we chose to apply this method to determine the absolute structure of a chiral hydrocarbon, for the first time.

A single crystal of the porous complex $[(\mathrm{ZnI_2})_2(\mathrm{tpt})_3 \cdot x(\mathrm{solvent})]_n$ [crystalline sponge; tpt = tris(4-pyridyl)-1,3,5-triazine][9] was treated with **1**. The guest-absorbed crystal was subjected to X-ray crystallographic data collection. The space-group change from centrosymmetric *C*2/*c* to non-centrosymmetric *C*2 was confirmed (see Figure S5), and is required for the absolute structure determination. The crystal structure of the **1**/crystalline sponge complex was solved and two independent molecules of **1** per asymmetric unit were clearly observed.[10] The absolute configuration of **1** was thus established as 2*S*, 3*S*, 7*R*, 11*R*, 14*R*, 15*S*, 18*S*, based on the Flack [0.071 (6)] and the Hooft [0.056 (6)] parameters. Significantly, the crystal structure confirmed not only the configuration but also the conformation of **1**: the six-membered A and C rings both adopt the chair conformation, while the eight-membered B ring displays the boat-boat conformation of a cyclooctene (Figure 3). The sesterterpene **1** and the terpene synthase were hereby designated as astellifadiene and EvAS (DDBJ/EMBL/GenBank accession number: LC113889), respectively.

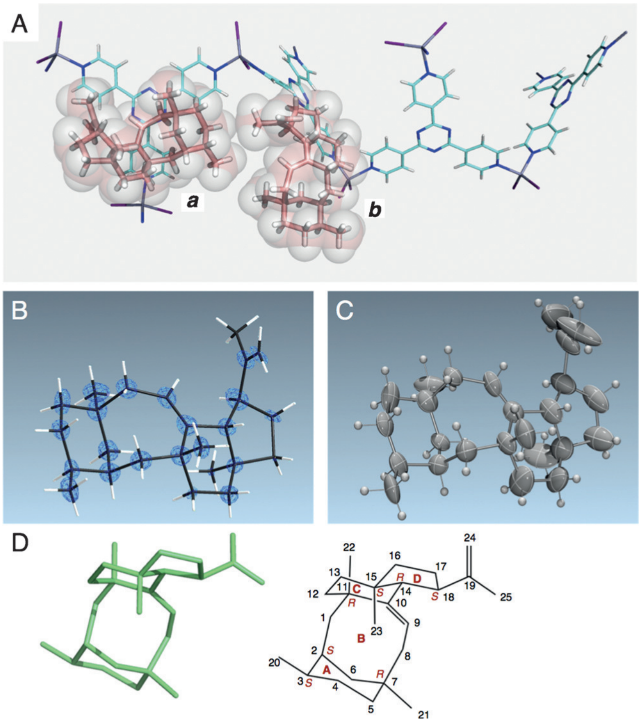


Figure 3. Crystal structure of the complex of **1** and the crystalline sponge. A) Structure of an asymmetric unit containing two crystallographically independent molecules of **1** (***a*** and ***b***). Solvent molecules have been omitted for clarity. B) Crystal structure of guest ***a*** superimposed with the Fourier electron density map (σ level = 0.8) for ***a***. C) Ortep drawing with 50% probability for guest ***a***. D) Another view of the structure of ***a***.

We then sought to elucidate the cyclization mechanism leading to **1**, and performed an isotope-incorporation study with [1-^{13}C]sodium acetate. Since fungi employ the mevalonate pathway, both DMAPP and IPP derived from [1-^{13}C]sodium acetate are labeled at C1 and C3, thus generating GFPP labeled with ^{13}C at ten positions (Figure 4A). The resulting sesterterpene **1** was labeled at C1, C3, C5, C7, C9, C11, C13, C15, C17, and C19 (Figure 4A; see Figure S21). This labeling pattern does not follow the isoprene rule,[11] and thus indicates the occurrence of C−C bond migration(s) during the biogenesis of **1**. On the basis of these observations, a plausible cyclization mechanism to yield **1** can be proposed as follows (Figure 4B). After the elimination of the pyrophosphate group of GFPP, the initial cyclization occurs at C1/C11 and C10/C14 to provide **2**$^+$, which consists of an 11-5-fused ring system. Subsequently, the ring expansion proceeds with the aid of the terminal π-bond to yield **3**$^+$ with an 11-6-5-fused ring system, and involves C−C bond rearrangement to form the C13−C15 bond and the new bond C14−C18. The deprotonation from one of the methyl groups of the isopropyl cation neutralizes **3**$^+$ to afford **4**, thus completing the first half of the reaction. The reprotonation at C3 initiates the second cyclization sequence, thus generating the C2−C6 bond to yield the 5-8-6-5-fused tetracyclic carbocation **5**$^+$. **5**$^+$ then undergoes the bond migration to yield **6**$^+$ with the bicyclo[5.3.1]undecane moiety. Finally, the 1,5-hydride shift from C10 to C6, followed by the deproto-

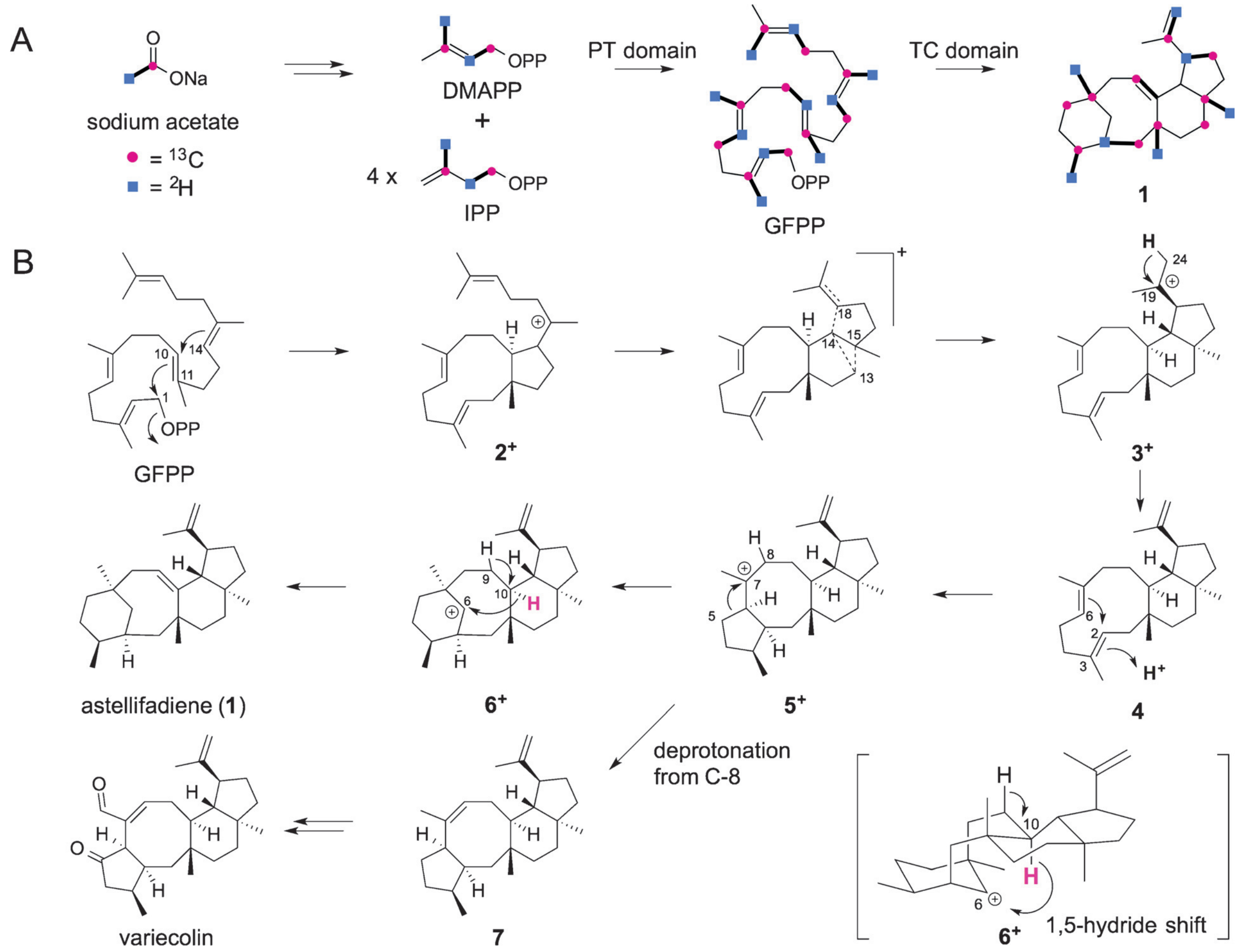


Figure 4. A) Summary of the labeling experiments of **1** using either [1-^{13}C] or [1-^{13}C,^{2}H$_3$]sodium acetate. Bold lines indicate the acetate unit. B) Proposed reaction mechanism for the generation of **1**.

nation from C9, completes the reaction to afford **1**. H10 of **6**$^+$ appears to be sufficiently close to C6 (Figure 4B), thus supporting the occurrence of the hydride shift between these two positions.

Additionally, we incorporated either [1-^{13}C, ^{2}H$_3$]sodium acetate or [2-^{13}C, ^{2}H$_3$]sodium acetate into **1**, and the resultant labeling patterns were found to be consistent with the above-proposed mechanism (Figure 4A; see Figures S7, S22, and S23 and Tables S4 and S5; see Supporting Information for the detailed interpretation of the NMR spectra). The results also revealed that C24, not C25, is labeled with ^{2}H when [1-^{13}C, ^{2}H$_3$]sodium acetate is used, thus illuminating the fact that the first deprotonation occurs at the methyl group originating from C5 of DMAPP. Furthermore, an investigation of the fate of the hydrogen atom derived from C10 is important, since this hydrogen is eliminated from C10 and is proposed to be transferred to C6. The present labeling experiments showed that C9 is not labeled with the deuterium originally bound to the adjacent carbon atom (C10), and is compatible with the proposed 1,5-hydride shift between C10 and C6.

To obtain further insight into the cyclization mechanism, we performed in vitro assays with recombinant EvAS expressed in *Escherichia coli*. After confirming its activity to synthesize **1** in the presence of DMAPP, IPP, and Mg^{2+} (see Figure S4A), the reaction was then performed in ^{2}H$_2$O to investigate whether the reprotonation of **4** occurs, according to the previous study in which ^{2}H$_2$O was utilized to probe similar reprotonation steps.[12] As expected, the incorporation of one atom of deuterium into **1** was observed (see Figures S3B and S4B), thus proving the existence of the second round of cyclization initiated by the protonation.

The cyclization reaction is initiated by the formation of the 11-5-fused bicyclic carbocationic species, which is the same as those for the fusicoccadiene, phomopsene, ophiobolin F, and stellatatriene synthases.[5a,b,6,13] This observation is consistent with the fact that the TC domain of EvAS as well as those of the other enzymes mentioned above, belongs to clade B (see Figure S2).[5c] Intriguingly, the absolute configuration of **1** is in good agreement with that of variecolin.[14] Considering their structural similarity, the predicted precursor of variecolin, the hydrocarbon **7**, appears to be synthesized by a shared pathway with **1**, with **5**$^+$ as a common intermediate (Figure 4B), as the deprotonation of **5**$^+$ from C8 generates **7** with a 5-8-6-5-fused ring system.

Finally, homology modeling studies on the basis of the recently reported crystal structure of fusicoccadiene synthase (PaFS), a diterpene synthase from *Phomopsis amygdali*,[15] suggested that the amino-acid residues constituting the bottom surface of the active-site cavity of the TS are uniquely substituted in EvAS (see Figure S8). The large to small substitutions of F65, W225, and M309 in PaFS significantly increase the volume and change the shape of the active site in EvAS. These differences might explain the substrate and product specificities of the terpene cyclases.

In summary, we isolated and characterized the sesterterpene astellifadiene (**1**), having a novel scaffold, and proposed a plausible mechanism for the cyclization to yield **1**. The structure determination of a natural product is sometimes challenging, especially with a complex hydrocarbon containing many chiral centers. In this study, the crystalline sponge method, coupled with an NMR study, readily established the absolute structure of a natural product immediately after its isolation. It should be emphasized that the combination of two reliable analysis methods completely eliminated any ambiguity in the other method (e.g., assumed conformation in NMR and heavy atom assignment in X-ray analysis). Until recently, X-ray crystallography, the most definitive method for structure determination, has not been applicable for minute amounts of oily compounds. Given that **1** is an unfunctionalized hydrocarbon and represents one of the most difficult classes of compounds for absolute structure determination, the rapid, nondestructive, and complete structure analysis of a minute quantity (10 μg) of oily **1** provides a particularly important example in natural product structure determination. Meanwhile, we also investigated the cyclization mechanism to synthesize **1** by isotope-labeling experiments, which provided a reasonable scenario for the construction of the 6-8-6-5-fused ring system. In conclusion, the genome mining of many cryptic biosynthetic enzymes, combined with the crystalline sponge method, will rapidly expand our structural knowledge of diverse natural products.

Acknowledgments

We thank Prof. K. Gomi (Tohoku University) and Prof. K. Kitamoto (The University of Tokyo) for providing the expression vectors and the fungal strain. We are grateful to Prof. H. Oikawa (Hokkaido University) and Prof. T. Kuzuyama (The University of Tokyo) for providing the doubly labeled acetates. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology.

Keywords: biosynthesis · crystal growth · natural products · structure elucidation · terpenoids

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 5785–5788
Angew. Chem. **2016**, *128*, 5879–5882

[1] a) T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Bruccoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben, R. Breitling, E. Takano, M. H. Medema, *Nucleic Acids Res.* **2015**, *43*, W237–W243; b) Y. Tsunematsu, K. i. Ishiuchi, K. Hotta, K. Watanabe, *Nat. Prod. Rep.* **2013**, *30*, 1139–1149; c) E. J. Helfrich, S. Reiter, J. Piel, *Curr. Opin. Biotechnol.* **2014**, *29*, 107–115; d) A.-C. Letzel, S. J. Pidot, C. Hertweck, *Nat. Prod. Rep.* **2013**, *30*, 392–428.

[2] a) R. C. Breton, W. F. Reynolds, *Nat. Prod. Rep.* **2013**, *30*, 501–524; b) G. T. Carter, *Nat. Prod. Rep.* **2014**, *31*, 711–717.

[3] K. Ling-Yi, W. Peng, *Chin. J. Nat. Med.* **2013**, *11*, 193–198.

[4] a) Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen, M. Fujita, *Nature* **2013**, *495*, 461–466; b) N. Zigon, M. Hoshino, S. Yoshioka, Y. Inokuma, M. Fujita, *Angew. Chem. Int. Ed.* **2015**, *54*, 9033–9037; *Angew. Chem.* **2015**, *127*, 9161–9165; c) S. Urban, R. Brkljača, M. Hoshino, S. Lee, M. Fujita, *Angew. Chem. Int. Ed.* **2016**, *55*, 2678–2682; *Angew. Chem.* **2016**, *128*, 2728–2732.

[5] a) T. Toyomasu, A. Kaneko, T. Tokiwano, Y. Kanno, Y. Kanno, R. Niida, S. Miura, T. Nishioka, C. Ikeda, W. Mitsuhashi, *J. Org. Chem.* **2009**, *74*, 1541–1548; b) Y. Matsuda, T. Mitsuhashi, Z. Quan, I. Abe, *Org. Lett.* **2015**, *17*, 4644–4647; c) Y. Ye, A. Minami, A. Mandi, C. Liu, T. Taniguchi, T. Kuzuyama, K. Monde, K. Gomi, H. Oikawa, *J. Am. Chem. Soc.* **2015**, *137*, 11846–11853; d) B. Qin, Y. Matsuda, T. Mori, M. Okada, Z. Quan, T. Mitsuhashi, T. Wakimoto, I. Abe, *Angew. Chem. Int. Ed.* **2016**, *55*, 1658–1661; *Angew. Chem.* **2016**, *128*, 1690–1693.

[6] R. Chiba, A. Minami, K. Gomi, H. Oikawa, *Org. Lett.* **2013**, *15*, 594–597.

[7] F. J. Jin, J. Maruyama, P. R. Juvvadi, M. Arioka, K. Kitamoto, *FEMS Microbiol. Lett.* **2004**, *239*, 79–85.

[8] M. Hoshino, A. Khutia, H. Xing, M. Fujita, *IUCrJ* **2016**, *3*, 139–151.

[9] K. Biradha, M. Fujita, *Angew. Chem. Int. Ed.* **2002**, *41*, 3392–3395; *Angew. Chem.* **2002**, *114*, 3542–3545.

[10] CCDC 1446744 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre.

[11] L. Ruzicka, *Experientia* **1953**, *9*, 357–367.

[12] a) J. Jiang, X. He, D. E. Cane, *J. Am. Chem. Soc.* **2006**, *128*, 8128–8129; b) D. J. Miller, J. Gao, D. G. Truhlar, N. J. Young, V. Gonzalez, R. K. Allemann, *Org. Biomol. Chem.* **2008**, *6*, 2346–2354.

[13] T. Toyomasu, M. Tsukahara, A. Kaneko, R. Niida, W. Mitsuhashi, T. Dairi, N. Kato, T. Sassa, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3084–3088.

[14] O. D. Hensens, D. Zink, J. M. Williamson, V. J. Lotti, R. S. Chang, M. A. Goetz, *J. Org. Chem.* **1991**, *56*, 3399–3403.

[15] M. Chen, W. K. Chou, T. Toyomasu, D. E. Cane, D. W. Christianson, *ACS Chem. Biol.* **2016**, DOI: 10.1021/acschembio.1025b00960.

Received: February 10, 2016