

Sesquiterpenoids

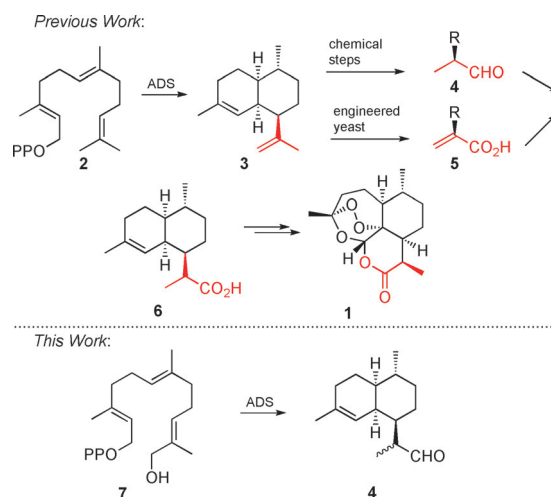
International Edition: DOI: 10.1002/anie.201609557
German Edition: DOI: 10.1002/ange.201609557

An Efficient Chemoenzymatic Synthesis of Dihydroartemisinic Aldehyde

Melodi Demiray, Xiaoping Tang, Thomas Wirth, Juan A. Faraldos, and Rudolf K. Allemann*

Abstract: Artemisinin from the plant *Artemisia annua* is the most potent pharmaceutical for the treatment of malaria. In the plant, the sesquiterpene cyclase amorphadiene synthase, a cytochrome-dependent CYP450, and an aldehyde reductase convert farnesyl diphosphate (FDP) into dihydroartemisinic aldehyde (DHAAL), which is a key intermediate in the biosynthesis of artemisinin and a semisynthetic precursor for its chemical synthesis. Here, we report a chemoenzymatic process that is able to deliver DHAAL using only the sesquiterpene synthase from a carefully designed hydroxylated FDP derivative. This process, which reverses the natural order of cyclization of FDP and oxidation of the sesquiterpene hydrocarbon, provides a significant improvement in the synthesis of DHAAL and demonstrates the potential of substrate engineering in the terpene synthase mediated synthesis of high-value natural products.

The sesquiterpenoid endoperoxide artemisinin (**1**) is widely used as a first-line treatment for malaria in combination therapy.^[1] Although elegant organic syntheses of artemisinin have been published,^[2] the worldwide supply of **1** predominantly relies on extraction from the plant *Artemisia annua*.^[3] The demand for artemisinin is mainly from the developing world, which requires the drug to be produced at low cost. Currently the most efficient way to synthesize artemisinin is to combine biosynthesis with chemical steps. Central to the biosynthesis of **1** (Scheme 1) is the class I sesquiterpene cyclase amorphadiene synthase (ADS), which catalyzes the conversion of (*E,E*)-farnesyl diphosphate (FDP, **2**) into amorpha-4,11-diene (**3**). In this complex reaction cascade, two 6-membered rings, four stereocentres, and two double bonds are formed with exquisite regio- and stereochemical control in one step.^[4] Dihydroartemisinic aldehyde (DHAAL, **4**) can be made from amorpha-4,11-diene (**3**) either through a three-step chemical synthesis or by combining a biooxidation with two chemical steps.^[5] Compound **4** can then be converted into **1** chemically or enzymatically in four well-



Scheme 1. Top: Current routes to artemisinin. Bottom: One-step ADS-catalyzed synthesis of dihydroartemisinic aldehyde (DHAAL, **4**) from 12-hydroxyfarnesyl diphosphate (**7**).

established steps.^[2a,6] It is noteworthy that an elegant semi-synthetic pathway has been developed that uses ADS and five other enzymes in yeast to produce artemisinic acid (**5**), which is then converted into dihydroartemisinic acid (**6**) by transition metal-catalyzed hydrogenation. The pharmaceutical company Sanofi scaled up this process in 2014 but the manufacture was discontinued owing to strong market forces,^[7] thus highlighting the need for ecologically friendly, low-cost alternatives for the production of artemisinin.

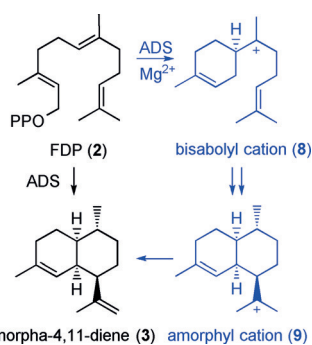
Herein, we report a novel chemoenzymatic process that exploits the substrate promiscuity of ADS to convert the hydroxylated FDP analogue **7** into the synthetic intermediate DHAAL (**4**). In contrast to existing procedures, this process, which reverses the natural order of cyclization of FDP and oxidation of the sesquiterpene hydrocarbon, significantly shortens the synthesis of dihydroartemisinic aldehyde (**4**), in that it uses only one enzyme and requires a single oxidation step that occurs prior to the ADS-catalyzed cyclization to **4**. The process avoids several redox steps after the cyclization since it bypasses the formation of the intermediate amorpha-diene (**3**) altogether (Scheme 1).

Amorphadiene synthase (ADS) catalyzes the Mg²⁺-dependent conversion of its natural substrate FDP (**2**; see Figure S13 in the Supporting Information) along a complex reaction path that involves isomerization of the 2,3-double bond from the *E* to the *Z* configuration and 1,6-cyclization to generate a bisabolyl cation (**8**), followed by a 1,3-hydride shift from C1 to C7 and 1,10-cyclization to the amorphyll cation (**9**). Finally, deprotonation at C12 or C13 generates amorpha-4,11-diene (**3**; Scheme 2 and Figures S4, S22, S23).^[8]

[*] M. Demiray, Dr. X. Tang, Prof. Dr. T. Wirth, Dr. J. A. Faraldos, Prof. Dr. R. K. Allemann
School of Chemistry, Cardiff University
Main Building, Park Place, Cardiff CF10 3AT (Great Britain)
E-mail: allemannrk@cf.ac.uk
Homepage: <http://blogs.cardiff.ac.uk/allemann>

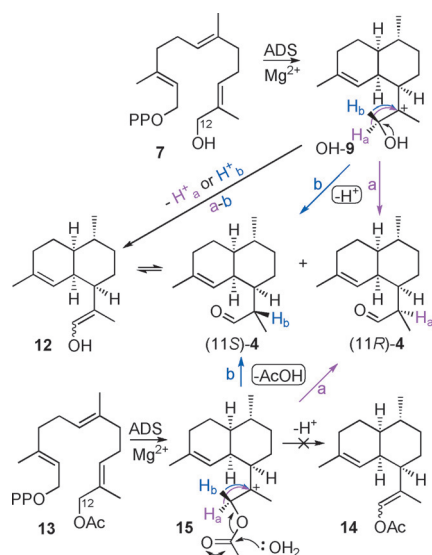
Supporting information for this article can be found under:
<http://dx.doi.org/10.1002/anie.201609557>.

© 2017 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



Scheme 2. Synthesis of amorphadiene (**3**) from FDP (**2**) in the plant *A. annua*.

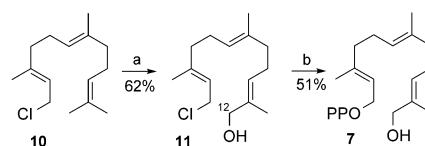
Many sesquiterpene synthases display some degree of substrate promiscuity, and are able to convert methylated and fluorinated farnesyl diphosphate analogues into modified terpenoids.^[9–11] In this work, we investigated for the first time the potential of a sesquiterpene synthase to accept a hydroxylated FDP analogue as a substrate. In particular, we asked whether ADS^[11b] could convert 12-hydroxyfarnesyl diphosphate (**7**;^[12] Scheme 1) into dihydroartemisinic aldehyde (**4**) via a 12-hydroxyamorphyl cation (OH-**9**; Scheme 3). It is



Scheme 3. The ADS-catalyzed synthesis of **4** from 12-hydroxyfarnesyl diphosphate (**7**) and 12-acetoxymethyl farnesyl diphosphate (**13**) via carbocations OH-**9** and **15**, respectively.

known that α -hydroxylated carbocations such as OH-**9** can isomerize under acidic conditions to aldehydes.^[13] The electrophilic nature of terpene synthase chemistry combined with the inherent reactivity of α -hydroxylated carbocations such as OH-**9** should therefore lead to aldehyde **4**.

Diphosphate **7** was synthesized in two steps from commercially available (*E,E*)-farnesyl chloride (**10**; Scheme 4). The short synthesis involved low-temperature acid-catalyzed selenium dioxide oxidation of **10**^[14] and diphosphorylation of the resulting chloride **11**^[15] (Tables S1, S2 and Figures S13–



Scheme 4. Synthesis of 12-hydroxyfarnesyl diphosphate (**7**) from commercially available farnesyl chloride (**10**). a) SeO₂ (0.3 equiv), salicylic acid (0.3 equiv), *t*BuOOH (5 equiv), CH₂Cl₂, 0°C. b) (Bu₄N)₃HP₂O₇ (2 equiv), CH₃CN, room temperature.

S16). For ¹H NMR and GC–MS comparison, an authentic sample of aldehyde **4** was prepared from commercially available (11*R*)-dihydroartemisinic acid (11*R*)-**6** (Figures S7, S8, S26, S27).^[6b,16]

The ¹H-NMR spectrum of the purified enzymatic products (Figure 1 and Figures S24, S25) from incubations of **7** with ADS showed the presence of a 3:2 mixture of aldehydes

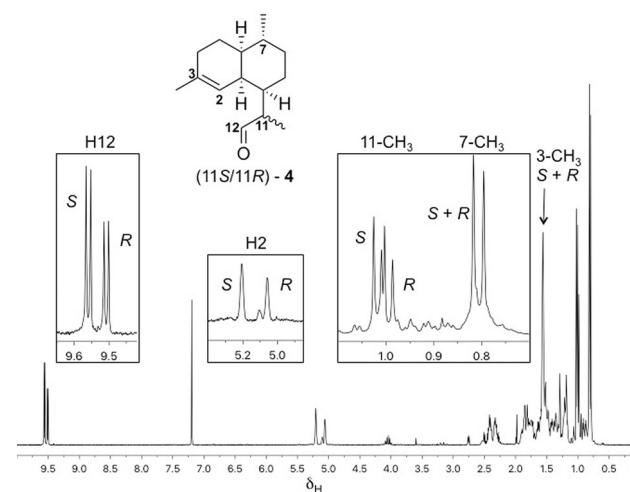


Figure 1. ¹H NMR (500 MHz, CDCl₃) of the pentane-extractable products generated by ADS-catalyzed turnover of **7** after silica-gel purification, with inserts showing the diagnostic peaks assigned to H12, H2, 11-CH₃, and 7-CH₃.

as clearly revealed by their diagnostic NMR signals at δ_{H} 9.62 ppm (1H, d, $J = 3.5$ Hz, CHO, major) and 9.57 ppm (1H, d, $J = 3.5$ Hz, CHO, minor). The identity of the minor product was determined to be (11*R*)-**4** based on ¹H-NMR data and comparison with an authentic sample of **4** (Figures S26, S27).^[17a] The structure of the major aldehyde was deduced to be the (11*S*)-**4** epimer based on its diagnostic ¹H-NMR signals at δ_{H} 5.27 (s, H2), 1.64 (s, 3-Me), 1.07 (d, $J = 6.8$ Hz, 11-Me), and 0.87 ppm (d, $J = 6.5$ Hz, 7-Me), which are complementary to those of (11*R*)-**4** at δ_{H} 5.13 (s, H2), 1.64 (s, 3-Me), 1.06 (d, $J = 6.9$ Hz, 11-Me), and 0.87 ppm (d, $J = 6.5$ Hz, 7-Me; Figure 1). In addition, the epimeric nature of (11*R*)-**4** and (11*S*)-**4** is further supported by the epimerization of the corresponding esters that is described later in Scheme 6. Analysis of the NMR sample by GC–MS showed a 7:2 mixture of (11*S*)-**4** and (11*R*)-**4**, in addition to approximately 10% of a minor product (Figure 2) that we tentatively propose to be the enol form (**12**) of aldehyde **4**

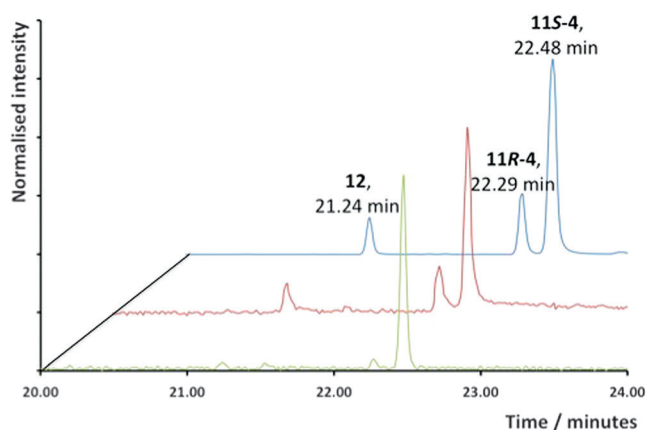
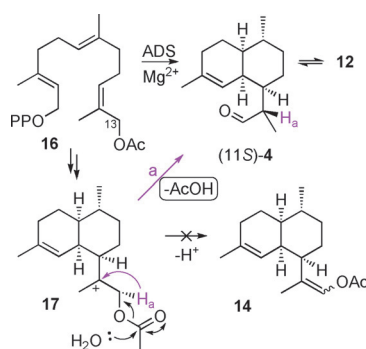


Figure 2. GC chromatogram (TIC) of the pentane-extractable products from an incubation of diphosphate **7** (blue), **13** (red), and **16** (green) with ADS. GC peaks were assigned as putative enol **12** (21.24 min), and aldehydes (11*R*)-**4** (22.29 min), and (11*S*)-**4** (22.48 min).

based on its identical MS spectrum (Figure S3) with that of (11*S*)-**4**. The MS spectra as well as the GC elution times of (11*R*)-**4** (faster) and (11*S*)-**4** are consistent with those reported previously.^[17] The reaction mechanism for the conversion of FDP to amorphadiene (Scheme 2) suggests that enol **12** may initially be formed from carbocation OH-**9** through deprotonation of either H_a or H_b (path a-b in Scheme 3) before **12** equilibrates to the observed mixture of epimers of **4** outside the active site of ADS. However incubations of diphosphate **7** with ADS in D₂O led to the exclusive production of deuterated (11*S*)-**4** (Figure S10); no deuterated (11*R*)-**4** was detected by GC-MS. This observation indicates that only (11*S*)-**4** can be formed via enol **12**.

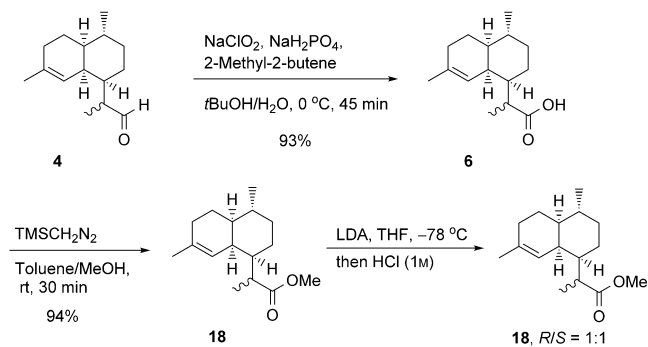
To further investigate the capability of ADS to act on unnatural substrates, 12-acetoxyfarnesyl diphosphate (**13**) was prepared in three steps (32% overall yield) from known 1-*O*-THP-protected 12-hydroxyfarnesol^[14] (Figures S17,S18) and incubated with ADS. GC-MS analysis revealed that the same three products were obtained as from incubations with diphosphate **7** (Figure 2 and Figure S9). The observation that no vinyl acetate **14** is produced from **13** (Scheme 3) suggests that the enol **12** is not a product of ADS catalysis but is formed exclusively from (11*S*)-**4** after its release from the active site. Control experiments excluded the presence of ester hydrolase activity in our enzyme preparations (see the Supporting Information). Remarkably, the incubation of 13-acetoxyfarnesyl diphosphate (**16**, Figure S19) with ADS (Scheme 5) yielded (11*S*)-**4** (93%) and enol **12** (7%; Figure 2 and Figures S5,S6). The absence of significant amounts of the (11*R*)-epimer of **4** implies that enol **12** does not revert to (11*R*)-**4** under the reaction conditions (pH ≈ 7.5; Scheme 3). These results strongly suggest that only one pathway operates to yield the observed epimeric mixture of aldehydes **4**. Hence the formation of **4** from **7** (and **13**) most likely proceeds through an oxygen-assisted non-stereospecific intramolecular [1,2]-hydride shift from carbocation OH-**9** (or **15** for **13**; paths a and b in Scheme 3). Since the stable vinyl acetate **14** is not formed from 13-acetoxyfarnesyl diphosphate (**16**), deprotonation of **17** via enol **12** to (11*S*)-**4** appears unlikely. Hence a stereospecific [1,2]-hydride shift of H_a from



Scheme 5. ADS-catalyzed production of (11*S*)-**4** from 13-acetoxyfarnesyl diphosphate (**16**) via carbocation **17**.

carbocation **17** most likely accounts for the exclusive formation of (11*S*)-**4** (path a in Scheme 5) from diphosphate **16**.

To further advance this novel chemoenzymatic approach towards the production of artemisinin, the 2:3 mixture of (11*R*) and (11*S*)-**4** was oxidized to the corresponding dihydroartemisinic acids (**6**) in 93% yield (Scheme 6),^[2b] and



Scheme 6. Synthesis of dihydroartemisinic acid methyl ester (**18**).

subsequently converted into the corresponding dihydroartemisinic methyl esters (**18**) in 94% yield (Figure S11).^[18] Treatment of the esters with LDA under kinetic control resulted in a 1:1 mixture of (11*R*)- and (11*S*)-methyl esters (Figure S28). (11*R*)-dihydroartemisinic methyl ester is the desired intermediate for the synthesis of artemisinin.^[19]

In conclusion, we have developed an efficient chemoenzymatic route to dihydroartemisinic aldehyde (**4**), a major intermediate in the production of artemisinin (**1**), the most important drug for the treatment of malaria. We have for the first time shown that hydroxylated FDP analogues can be accepted as substrates by sesquiterpene synthases. Hence our work offers a novel “reversed biosynthetic” approach for the synthesis of functionally diversified hydroxylated terpenoids. The chemical synthesis of such products is often difficult owing to the need to make several small rings with high stereo- and regiocontrol. The relatively high substrate promiscuity and the templating effect of the active site allow terpene synthases to chaperone unnatural substrates along well-defined reaction paths to specific products with high fidelity. This design offers a promising approach for the production of high-value terpenoids and terpene alkaloids^[11a]

that is complementary to conventional synthetic and biosynthetic procedures.

Acknowledgements

This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) through grants (BB/H01683X/1, BB/M022463/1, BB/N012526/1) and by the Engineering and Physical Sciences Research Council (EPSRC) through grant EP/M013219/1. We thank Dr. Veronica Gonzalez (Cardiff University) for helpful discussions. The help of Dr. Rob Jenkins and Thomas Williams with mass spectrometry and NMR spectroscopy is gratefully acknowledged.

Conflict of interest

The authors declare no conflict of interest.

Keywords: artemisinin · chemoenzymatic synthesis · sesquiterpenoids · substrate engineering · synthetic biology

How to cite: *Angew. Chem. Int. Ed.* **2017**, *56*, 4347–4350
Angew. Chem. **2017**, *129*, 4411–4415

- [1] a) D. L. Klayman, *Science* **1985**, *228*, 1049–1055; b) N. J. White, *Trends Parasitol.* **2004**, *113*, 1084–1092; c) S. Krishna, L. Bustamante, R. K. Haynes, H. M. Staines, *Trends Pharmacol. Sci.* **2008**, *29*, 520–527; d) P. G. Kremsner, S. Krishna, *Lancet* **2004**, *364*, 285–294.
- [2] a) F. Lévesque, P. H. Seeberger, *Angew. Chem. Int. Ed.* **2012**, *51*, 1706–1709; *Angew. Chem.* **2012**, *124*, 1738–1741; b) J. S. Yadav, B. Thirupathiah, P. Srihari, *Tetrahedron* **2010**, *66*, 2005–2009; c) M. A. Avery, W. K. M. Chong, C. Jennings-White, *J. Am. Chem. Soc.* **1992**, *114*, 974–979.
- [3] a) T. Townsend, V. Segura, G. Chigeza, T. Penfield, A. Rae, D. Harvey, D. Bowles, I. A. Graham, *PLoS One* **2013**, *8*, e61989; b) T. R. Larson, C. Branigan, D. Harvey, T. Penfield, D. Bowles, I. A. Graham, *Ind. Crops Prod.* **2013**, *45*, 1–6; c) W. K. Milhous, P. J. Weina, *Science* **2010**, *327*, 279–280; d) I. A. Graham, et al., *Science* **2010**, *327*, 328–331.
- [4] a) H. J. Bouwmeester, T. E. Wallaart, M. H. A. Janssen, B. van Loo, B. J. M. Jansen, M. A. Posthumus, C. O. Schmidt, J.-W. De Kraker, W. A. König, M. C. R. Franssen, *Phytochemistry* **1999**, *52*, 843–854; b) P. Mercke, M. Bengtsson, H. J. Bouwmeester, M. A. Posthumus, P. E. Brodelius, *Arch. Biochem. Biophys.* **2000**, *381*, 173–180; c) Y. J. Chang, S. H. Song, S. H. Park, S. U. Kim, *Arch. Biochem. Biophys.* **2000**, *383*, 178–184.
- [5] a) K. H. Teoh, D. R. Polichuk, D. W. Reed, G. Nowak, P. S. Covello, *FEBS Lett.* **2006**, *580*, 1411–1416; b) K. H. Teoh, D. R. Polichuk, D. W. Reed, P. S. Covello, *Botany* **2009**, *87*, 635–642; c) K. Yang, R. S. Monafared, H. Wang, A. Lundgren, P. E. Brodelius, *Plant Mol. Biol.* **2015**, *88*, 325–340; d) P. Muangphrom, M. Suzuki, H. Seki, E. O. Fukushima, T. Muranaka, *Plant Biotechnol.* **2014**, *31*, 483–491.
- [6] a) J. A. Dietrich, Y. Yoshikuni, K. J. Fisher, F. X. Woolard, D. Ockey, D. J. McPhee, N. S. Renninger, M. C. Y. Chang, D. Baker, J. D. Keasling, *ACS Chem. Biol.* **2009**, *4*, 261–267; b) P. J. Westfall, et al., *Proc. Natl. Acad. Sci. USA* **2012**, *E111*–E118; c) C. J. Paddon et al., *Nature* **2013**, *496*, 528–532; d) D.-K. Ro, et al., *Nature* **2006**, *440*, 940–943.
- [7] M. Peplow, *Nature* **2016**, *530*, 389–390.
- [8] a) S. Picaud, P. Mercke, X. He, X. O. Sterner, M. Brodelius, D. E. Cane, P. E. Brodelius, *Arch. Biochem. Biophys.* **2006**, *448*, 150–155; b) S.-H. Kim, K. Heo, Y.-J. Chang, S.-H. Park, S.-K. Rhee, S.-U. Kim, *J. Nat. Prod.* **2006**, *69*, 758–762.
- [9] D. E. Cane, *Chem. Rev.* **1990**, *90*, 1089–1103.
- [10] a) S. Touchet, K. Chamberlain, C. M. Woodcock, D. J. Miller, M. A. Birkett, J. A. Pickett, R. K. Allemann, *Chem. Commun.* **2015**, *51*, 7550–7553; b) V. Gonzalez, S. Touchet, D. J. Grundy, J. A. Faraldos, R. K. Allemann, *J. Am. Chem. Soc.* **2014**, *136*, 14505–14512; c) O. Cascón, S. Touchet, D. J. Miller, V. Gonzalez, J. A. Faraldos, R. K. Allemann, *Chem. Commun.* **2012**, *48*, 9702–9704; d) J. A. Faraldos, D. J. Miller, V. Gonzalez, Z. Yoosuf-Aly, O. Cascon, A. Li, R. K. Allemann, *J. Am. Chem. Soc.* **2012**, *134*, 5900–5908; e) D. J. Miller, F. L. Yu, R. K. Allemann, *ChemBioChem* **2007**, *8*, 1819–1825; f) J. A. Faraldos, Y. Zhao, P. E. O'Maille, J. P. Noel, R. M. Coates, *ChemBioChem* **2007**, *8*, 1826–1833.
- [11] a) K. A. Rising, C. M. Crenshaw, H. J. Koo, T. Subramanian, K. A. H. Chehade, C. Starks, K. D. Allen, D. A. Andres, H. P. Spielmann, J. P. Noel, *ACS Chem. Biol.* **2015**, *10*, 1729–1736; b) R. Lauchli, K. S. Rabe, K. Z. Kalbarczyk, A. Tata, T. Heel, R. K. Kitto, F. H. Arnold, *Angew. Chem. Int. Ed.* **2013**, *52*, 5571–5574; *Angew. Chem.* **2013**, *125*, 5681–5684.
- [12] a) M. Nagaki, T. Ichijo, R. Kobashi, Y. Yagihashi, T. Musashi, J. Kawakami, N. Ohya, T. Gotoh, H. Sagami, *J. Mol. Catal. B* **2012**, *80*, 1–6; b) M. Nagaki, M. Nakada, T. Musashi, J. Kawakami, N. Ohya, M. Kurihara, Y. Maki, T. Nishino, T. Koyama, *Biosci. Biotechnol. Biochem.* **2007**, *71*, 1657–1662.
- [13] a) H. B. Henbest, T. I. Wrigley, *J. Chem. Soc.* **1957**, 4596–4604; b) J. M. Coxon, M. P. Hartshorn, D. N. Kirk, *Tetrahedron* **1964**, *20*, 2531–2545; c) J. M. Coxon, M. P. Hartshorn, D. N. Kirk, *Tetrahedron* **1964**, *20*, 2547–2554; d) S. Julia, J.-P. Lavaux, *Bull. Soc. Chim. Fr.* **1963**, 1238–1241.
- [14] M. W. Rose, N. D. Rose, J. Boggs, S. Lenevich, J. Xu, G. Barany, M. D. Distefano, *J. Pept. Res.* **2005**, *65*, 529–537.
- [15] a) V. J. Davison, A. B. Woodside, T. R. Neal, K. E. Stremmler, M. Muehlbacher, C. D. Poulter, *J. Org. Chem.* **1986**, *51*, 4768–4779; b) Z. Huang, C. D. Poulter, *Org. Synth.* **1988**, *66*, 211–215; c) R. K. Keller, R. Thompson, *J. Chromatogr.* **1993**, *645*, 161–167.
- [16] M. Jung, D. A. Bustos, H. N. Elshohly, J. D. McChesney, *Synlett* **1990**, 743–744.
- [17] a) C. M. Berteau, et al., *Planta Med.* **2005**, *71*, 40–47; b) Y. Zhang, K. H. Teoh, D. W. Reed, L. Maes, A. Goossens, D. J. H. Olson, A. R. S. Ross, P. S. Covello, *J. Biol. Chem.* **2008**, *283*, 21501–21508.
- [18] E. Kühnel, D. D. P. Laffan, G. C. Lloyd-Jones, T. M. Campo, I. R. Shepperson, J. L. Slaughter, *Angew. Chem. Int. Ed.* **2007**, *46*, 7075–7078; *Angew. Chem.* **2007**, *119*, 7205–7208.
- [19] a) M. P. Feth, K. Rossen, A. Bugard, *Org. Process Res. Dev.* **2013**, *17*, 282–293; b) J. Turconi, F. Griolet, R. Guevel, G. Oddon, R. Villa, A. Geatti, M. Hvala, K. Rossen, R. Goller, A. Burgard, *Org. Process Res. Dev.* **2014**, *18*, 417–422; c) A. Burgard, T. Gieshoff, A. Peschl, D. Hoerstermann, C. Keleschovsky, R. Villa, S. Michelis, M. P. Feth, *Chem. Eng. J.* **2016**, *294*, 83–96.

Manuscript received: September 29, 2016

Revised: January 18, 2017

Final Article published: March 13, 2017