Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Tetrahedron Letters
Tetrahedron Letters 47 (2006) 3085–3089

Enzymatic formation of pyrrole-containing novel cyclic polyprenoids by bacterial squalene:hopene cyclase

Hideya Tanaka,[a] Hisashi Noma,[a] Hiroshi Noguchi[a] and Ikuro Abe[a,b,*]

[a]*School of Pharmaceutical Sciences and the 21st Century COE Program, University of Shizuoka, 52-1 Yada, Shizuoka 422-8526, Japan*
[b]*PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan*

Received 6 February 2006; revised 20 February 2006; accepted 27 February 2006

Abstract—A convergent synthesis provided two pyrrole-containing squalene analogues, in which a C_{20} isoprene unit is connected to pyrrole, 2-(geranylgeranyl)pyrrole and 2-(farnesyldimethylallyl)pyrrole. When incubated with recombinant squalene:hopene cyclase from *Alicyclobacillus acidocaldarius*, 2-(farnesyldimethylallyl)pyrrole was enzymatically converted to a 10:1 mixture of a tricyclic and a bicyclic unnatural novel polyprenoids, whereas 2-(geranylgeranyl)pyrrole was not a substrate for the enzyme.
© 2006 Elsevier Ltd. All rights reserved.

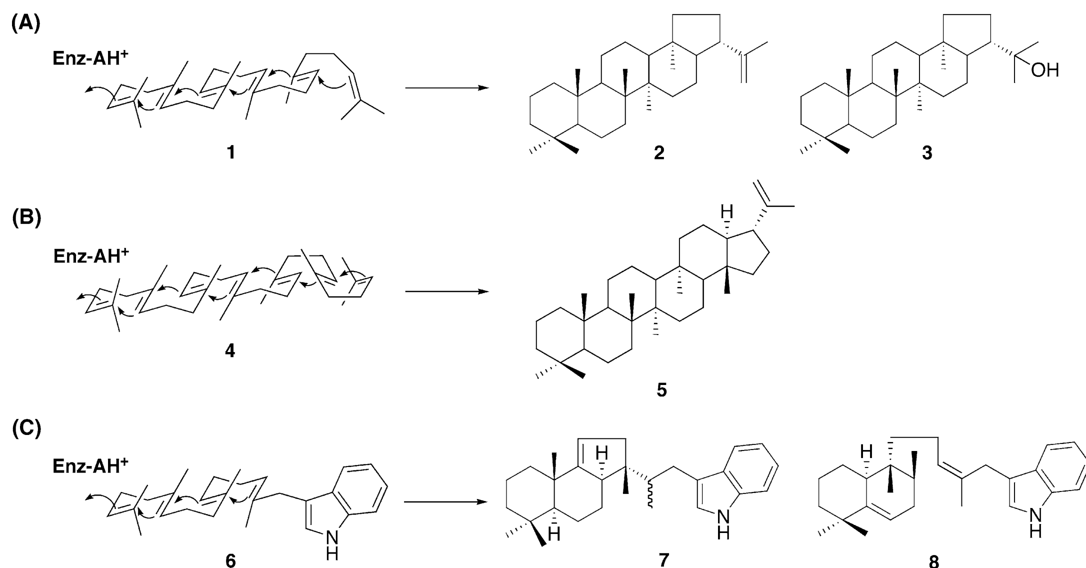
The broad substrate tolerance and catalytic potential of squalene:hopene cyclase (SHC) (E.C. 5.4.99.7) from a thermoacidophilic bacteria *Alicyclobacillus acidocaldarius* is remarkable.[1–3] The enzyme normally catalyzing cyclization of squalene (**1**) into a 5:1 mixture of hop-22(29)-ene (**2**) and hopan-22-ol (**3**) (Scheme 1A) accepts a variety of substrate analogues (C_{15}–C_{35}) and efficiently performs sequential ring-forming reactions to produce a series of unnatural cyclic polyprenoids. One of the most impressive examples is the cyclization of a C_{35} squalene analogue (**4**) into a 'supra-natural' hexacyclic polyprenoid (**5**) with a 6.6.6.6.6.5-fused ring system (Scheme 1B).[3a] Further, we have recently reported formation of indole-containing novel cyclic polyprenoids (**7** and **8**) from 3-(farnesyldimethylallyl)indole (**6**) (Scheme 1C).[3b] Here we now describe synthesis and enzymatic cyclization of two new substrate analogues in which a C_{20} isoprene unit is connected to pyrrole. It was anticipated that incorporation of the less bulky pyrrole ring, instead of indole, would not significantly disturb the cyclization reaction and thus lead to the formation of novel unnatural cyclic polyprenoids with highly fused ring systems. In addition, interesting biological activities of the cyclization products were also expected due to the presence of the heteroaromatic ring moiety.

The convergent synthesis of the substrate analogues, 2-(geranylgeranyl)pyrrole (**9**)[4] and 2-(farnesyldimethylallyl)pyrrole (**10**),[5] involved regioselective alkylation of pyrrole at C-2 by geranylgeranyl bromide or farnesyldimethylallyl bromide, which was mediated by zinc triflate as described before (Scheme 2).[3b] When incubated with purified recombinant *A. acidocaldarius* SHC,[6] 2-(geranylgeranyl)pyrrole (**9**) was completely inactive and did not afford any reaction products (Scheme 3A), which was confirmed by HPLC and GLC analysis. This was consistent with the previous observation that 3-(geranylgeranyl)indole was not a substrate for *A. acidocaldarius* SHC,[3b] suggesting that the enzyme is particularly sensitive to structural changes on the *pro*-C14β face and thus fails to bind the substrate analogue (Scheme 3A). Presumably, α-orientation of the *pro*-C14 methyl group is crucial for the correct folding of the substrate as in the case of the folding of squalene (Scheme 1A).[3a–d] The alternative β-orientation of the *pro*-C14 methyl group may interact repulsively with the *pro*-C10 methyl or a nearest neighbor from the substrate binding pocket of the enzyme; possibly with the residue I261 on the basis of the crystal structure of *A. acidocaldarius* SHC[1,2] (Fig. 1).

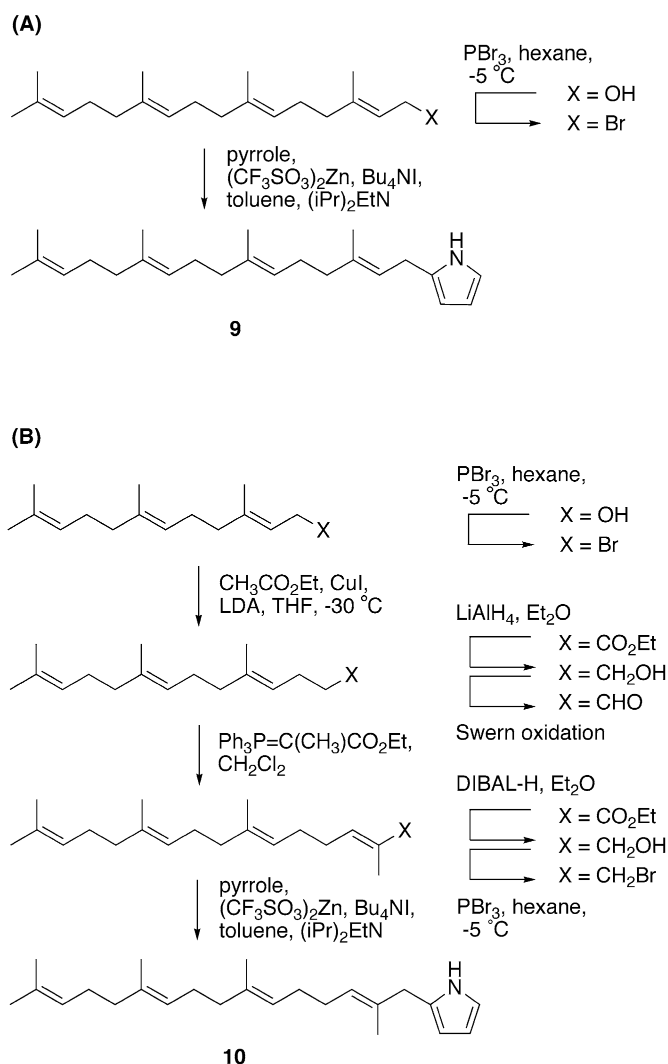
In contrast, 2-(farnesyldimethylallyl)pyrrole (**10**), in which a farnesyl C_{15} unit is connected in a head-to-head fashion to a dimethylallyl C_5 unit, was converted to a 10:1 mixture of the two novel tri- and bicyclic products **11** and **12** (Scheme 3B), whose structures were

Keywords: Squalene cyclase; Triterpene synthase; Unnatural polyprenoids; Pyrrole.
* Corresponding author. Tel./fax: +81 54 264 5662; e-mail: abei@ys7.u-shizuoka-ken.ac.jp

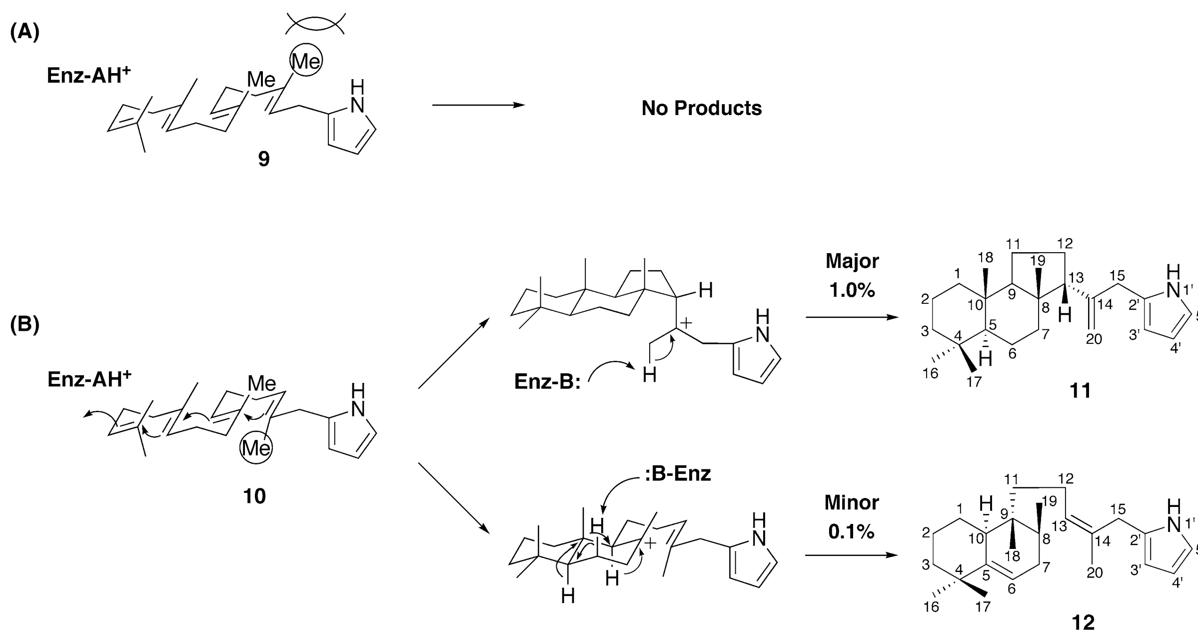
0040-4039/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2006.02.151



Scheme 1. Proposed mechanism for the conversion of (A) squalene (**1**) to hop-22(29)-ene (**2**) and hopan-22-ol (**3**); (B) a C_{35} squalene analogue (**4**) into a supra-natural hexacyclic polyprenoid (**5**); (C) 3-(farnesyldimethylallyl)indole (**6**) to indole-containing novel cyclic polyprenoids **7** and **8** by recombinant *Alicyclobacillus acidocaldarius* SHC.



Scheme 2. Synthesis of the pyrrole-containing substrate analogues.



Scheme 3. Proposed mechanism for the enzymatic conversion of the pyrrole-containing substrate analogues.

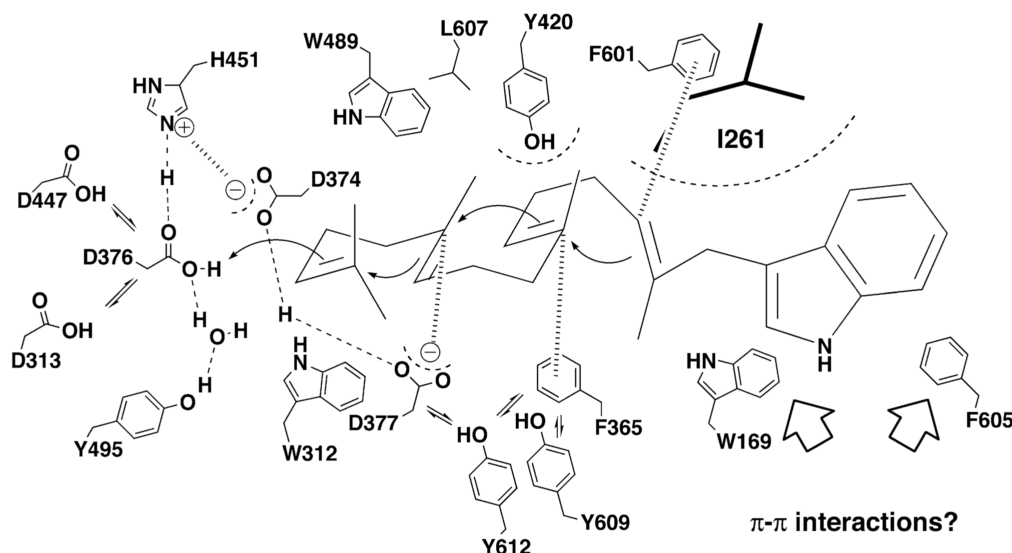


Figure 1. Proposed interactions between the active-site amino acid residues of *Alicyclobacillus acidocaldarius* SHC and 3-(farnesyldimethylallyl)indole (**6**) (adapted from Ref. 1c).

determined by NMR (^{1}H and ^{13}C NMR, HMQC, HMBC, NOESY, and differential NOE) and MS spectroscopy.[7,8] Thus, the ^{1}H NMR spectrum of the major product **11** (1.0 mg, 1.0% yield) revealed the presence of four methyl singlets (δ 0.64, 0.76, 0.77, and 0.80) and two vinylic protons (δ 4.71 and 4.82), in addition to the signals due to the pyrrole ring. A structure with the 6.6.5-fused tricyclic ring system was uniquely consistent with both biogenetic reasoning and the heteronuclear correlation spectroscopy (HMQC and HMBC). Further, the ring junctions and the stereochemistry of ring substituents were established by NOE experiments. The stereochemistry of C-13 was confirmed by NOEs observed between Me-18 and Me-19, but absent between Me-19 and H-20. Interestingly, the core structure of **11** with the $\Delta^{14(20)}$ double bond was different from that of the previously reported indole-containing 6.6.5-tricyclic product **7** with the $\Delta^{9(11)}$ double bond (Scheme 1C).[3b] On the other hand, the ^{1}H NMR of the minor product **12** (0.1 mg, 0.1% yield) indicated the presence of one methyl doublet (δ 0.85), three methyl singlets (δ 0.61, 1.02, and 1.06), one vinylic methyl (δ 1.56), and one vinylic proton (δ 5.95), which was in good agreement with those of the indole-containing minor product **8** with a 6.6-fused ring system (Scheme 1C).[3b] Indeed, NMR data (HMQC, HMBC, and NOE) unambiguously established the bicyclic structure with the Δ^{5} double bond. This was the first demon-

stration of the enzymatic formation of pyrrole-containing cyclic polyprenoids.

As in the case of 3-(farnesyldimethylallyl)indole (**6**),[3b] the enzymatic cyclization of 2-(farnesyldimethylallyl)-pyrrole (**10**), folded in all *pre*-chair conformation, was interrupted at the bicyclic or tricyclic intermediate cation (Scheme 3B). From the 6.6.5-tricyclic tertiary cation, proton elimination at H-20 yielded **11** as the major product (Scheme 3B), whereas, in the case of the indole-containing analogue **6**, an alternative backbone rearrangement (H-13β→14, CH_3-8β→13β, H-9α→8α) with elimination of H-11β afforded **7** with the $\Delta^{9(11)}$ double bond (Scheme 1C).[3b] Presumably, the presence of the bulky indole moiety caused perturbation in the folding conformation of the intermediate cation, which resulted in the irregular rearrangement reaction and the proton abstraction at the different position. On the other hand, from the 6.6-bicyclic tertiary intermediate cation, a backbone rearrangement (H-9α→8α, CH_3-10β→9β, H-5α→10α) with elimination of H-6β yielded the minor product **12** (Scheme 3B). In both cases, it is remarkable that the stereochemistry of the cyclization reactions of the unnatural substrates was strictly controlled by the enzyme in a regio- and stereo-specific manner.

Here it should be noted that the less bulky pyrrole-containing analogue **10** was not as a good substrate as the indole-containing analogue **6**. The yield of the pyrrole product **11** was five times less than that of the indole product **7**. This may be due to the stereoelectronic effect of the heteroaromatic ring. Thus, the π-electron rich indole ring moiety fits better into the active-site of the enzyme because of more efficient π–π interactions with the active-site aromatic residues; possibly with W169 and/or F605 lining the active-site cavity of *A. acidocaldarius* SHC[1,2] (Fig. 1). Manipulation of the enzyme reaction by combination of newly designed substrate analogues and structure-based modulation of the active-site geometry of the enzyme would lead to further production of structurally distinct 'supra-natural steroids', which is now in progress in our laboratories.

Acknowledgements

This work was in part supported by the COE21 Program, and Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by Grant-in-Aid from The Sankyo Foundation of Life Science, Japan. H.T. is a recipient of the JSPS Fellowship for Young Scientist (No. 175479).

Supplementary data

Supplementary data (Complete set of spectroscopic data (^{1}H and ^{13}C NMR, HMQC, HMBC, and NOE) of the substrate analogues (**9** and **10**) and enzyme reaction products (**11** and **12**) (nine pages)) associated with this article can be found, in the online version, at doi:10.1016/j.tetlet.2006.02.151.

References and notes

1. For reviews, see: (a) Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189–2206; (b) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812–2833; (c) Hoshino, T.; Sato, T. *Chem. Commun.* **2002**, 291–301.
2. For crystal structure of *A. acidocaldarius* SHC see: (a) Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811–1815; (b) Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J. Mol. Biol.* **1999**, *286*, 175–187.
3. For enzymatic conversion of squalene analogs by bacterial squalene cyclase, see: (a) Abe, I.; Tanaka, H.; Noguchi, H. *J. Am. Chem. Soc.* **2002**, *124*, 14514–14515; (b) Tanaka, H.; Noguchi, H.; Abe, I. *Org. Lett.* **2005**, *7*, 5873–5876; (c) Tanaka, H.; Noguchi, H.; Abe, I. *Org. Lett.* **2004**, *6*, 803–806; (d) Tanaka, H.; Noguchi, H.; Abe, I. *Tetrahedron Lett.* **2004**, *45*, 3093–3096; (e) Rohmer, M.; Anding, C.; Ourisson, G. *Eur. J. Biochem.* **1980**, *112*, 541–547; (f) Bouvier, P.; Berger, Y.; Rohmer, M.; Ourisson, G. *Eur. J. Biochem.* **1980**, *112*, 549–556; (g) Rohmer, M.; Bouvier, P.; Ourisson, G. *Eur. J. Biochem.* **1980**, *112*, 557–560; (h) Renoux, J.-M.; Rohmer, M. *Eur. J. Biochem.* **1986**, *155*, 125–132; (i) Abe, I.; Rohmer, M. *J. Chem. Soc., Perkin Trans. 1* **1994**, 783–791; (j) Abe, I.; Dang, T.; Zheng, Y. F.; Madden, B. A.; Feil, C.; Poralla, K.; Prestwich, G. D. *J. Am. Chem. Soc.* **1997**, *119*, 11333–11334; (k) Robustell, B.; Abe, I.; Prestwich, G. D. *Tetrahedron Lett.* **1998**, *39*, 957–960; (l) Robustell, B.; Abe, I.; Prestwich, G. D. *Tetrahedron Lett.* **1998**, *39*, 9385–9388; (m) Zheng, Y. F.; Abe, I.; Prestwich, G. D. *J. Org. Chem.* **1998**, *63*, 4872–4873; (n) Sato, T.; Abe, T.; Hoshino, T. *Chem. Commun.* **1998**, 2617–2618; (o) Hoshino, T.; Kondo, T. *Chem. Commun.* **1999**, 731–732; (p) Hoshino, T.; Ohashi, S. *Org. Lett.* **2002**, *4*, 2553–2556; (q) Hoshino, T.; Nakano, S.; Kondo, T.; Sato, T.; Miyoshi, A. *Org. Biomol. Chem.* **2004**, *2*, 1456–1470; (r) Nakano, S.; Ohashi, S.; Hoshino, T. *Org. Biomol. Chem.* **2004**, *2*, 2012–2022; (s) Hoshino, T.; Kumai, Y.; Kudo, I.; Nakano, S.; Ohashi, S. *Org. Biomol. Chem.* **2004**, *2*, 2650–2657.
4. 2-((2*E*,6*E*,10*E*)-3,7,11,15-Tetramethylhexadeca-2,6,10,14-tetraenyl)-pyrrole (**9**): ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.95 (br s, 1H), 6.65 (m, 1H), 6.14 (m, 1H), 5.92 (m, 1H), 5.36 (m, 1H), 5.12 (m, 3H), 3.36 (d, 2H, $J = 7.3$ Hz), 2.05 (m, 12H), 1.70 (s, 3H), 1.62 (s, 3H), 1.61 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 137.4, 135.4, 135.0, 131.2, 124.4, 124.1, 124.1, 120.9, 116.1, 108.5, 104.9, 39.7, 39.6, 26.7, 26.6, 26.6, 26.4, 26.3, 25.7, 21.0, 17.7, 16.1, 16.0, 16.0. HRMS (FAB): found for $[C_{24}H_{37}N]^+$ 339.2898; calcd. 339.2926.
5. 3-((2*E*,6*E*,10*E*)-2,7,11,15-Tetramethylhexadeca-2,6,10,14-tetraenyl)-pyrrole (**10**): ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.90 (br s, 1H), 6.66 (m, 1H), 6.13 (m, 1H), 5.94 (m, 1H), 5.29 (m, 1H), 5.13 (m, 3H), 3.28 (s, 2H), 2.00 (m, 12H), 1.69 (s, 3H), 1.63 (s, 3H), 1.61 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 135.5, 135.0, 135.0, 133.6, 131.2, 126.5, 124.4, 124.2, 124.1, 116.4, 108.3, 106.1, 39.7, 39.7, 38.3, 28.2, 28.1, 26.8, 26.6, 25.7, 17.7, 16.1, 16.0, 15.8. HRMS (FAB): found for $[C_{24}H_{37}N]^+$ 339.2938; calcd 339.2926.
6. The recombinant *A. acidocaldarius* SHC was prepared as described in the previous papers.[3a–d] Reaction mixture contained the substrate analog **9** (or **10**) (20 mg) and purified recombinant SHC (120 mg) in 400 ml of 50 mM Na citrate, pH 6.0, 0.1% Triton X-100, was incubated at 60 °C for 16 h. The incubations were stopped by freezing and lyophilization, followed by extraction with hexane (3 × 200 ml). The combined extracts were evaporated to dryness, separated on SiO_2 TLC (20% EtOAc in hexane) and reverse-phase HPLC (Phenomex Gemini 5m ODS column, 250 × 4.6 mm, 5% THF in MeCN, 0.5 ml/min) to

give **11** (0.2 mg) and **12** (0.02 mg) (from **10**). In order to get enough amount of products for NMR analysis, the experiments were repeated five times.

7. Major product (**11**): ^{1}H NMR (500 MHz, CD_3OD): δ 6.48 (m, 1H, H-5′), 5.87 (m, 1H, H-4′), 5.66 (m, 1H, H-3′), 4.82 (s, 1H, H-20), 4.71 (s, 1H, H-20), 3.25 (s, 2H, H-15), 1.99 (t, 1H, $J = 6.8$ Hz, H-13), 0.80 (s, 3H, Me-18), 0.77 (s, 3H, Me-16), 0.76 (s, 3H, Me-17), 0.64 (s, 3H, Me-19). ^{13}C NMR (125 MHz, CD_3OD): δ 149.3 (C-14), 130.0 (C′-2), 117.3 (C′-5), 112.5 (C-20), 108.2 (C′-4), 107.0 (C′-3), 64.7 (C-9), 58.8 (C-5), 57.3 (C-13), 43.7 (C-3), 41.3 (C-1), 37.9 (C-15), 34.0 (C-16), 34.0 (C-10), 30.8 (C-7), 26.7 (C-12), 21.8 (C-17), 20.7 (C-11), 20.5 (C-2), 19.4 (C-6), 15.9 (C-18), 15.8 (C-19). The NMR assignments were performed according to data from H–H COSY, HMQC, HMBC, NOESY, differential NOE experiments. LRMS (FAB^{+}): *m*/*z* 69, 95, 134, 191, 231, 340 (M^{+}). HRMS (FAB^{+}): found for [$C_{24}H_{37}N$] 339.2923; calcd 339.2926. $[\alpha]_D^{25}$ −21.7 (*c* 0.10 in $CHCl_3$).
8. Minor product (**12**): ^{1}H NMR (400 MHz, CD_3OD): δ 6.57 (m, 1H, H-5′), 5.95 (m, 1H, H-4′), 5.77 (m, 1H, H-3′), 5.47 (m, 1H, H-6), 5.21 (m, 1H, H-13), 1.56 (s, 3H, Me-20), 1.06 (s, 3H, Me-17), 1.02 (s, 3H, Me-16), 0.85 (d, 3H, $J = 6.8$ Hz, Me-19), 0.61 (s, 3H, Me-18). ^{13}C NMR (100 MHz, CD_3OD): δ 149.2 (C-5), 134.5 (C-14), 117.3 (C′-5), 127.1 (C-13), 117.4, 117.3, 108.2, 106.5, 42.2 (C-3), 41.2 (C-10), 39.1, 37.3, 34.7 (C-4), 32.7 (C-8), 31.2, 30.3 (C-17), 29.5 (C-16), 29.2, 28.7, 23.3, 22.4, 22.1, 19.8, 16.5 (C-18), 15.5 (C-19). The NMR assignments were performed according to data from H–H COSY, HMQC, HMBC, differential NOE experiments. LRMS (FAB^{+}): *m*/*z* 69, 80, 119, 134, 176, 191, 340 (M^{+}). HRMS (FAB^{+}): found for [$C_{24}H_{37}N$] 339.2925; calcd 339.2926. $[\alpha]_D^{25}$ 16.9 (*c* 0.10 in $CHCl_3$).