ORGANIC LETTERS

2011
Vol. 13, No. 10
2678–2681

Triterpene Cyclases from *Oryza sativa* L.: Cycloartenol, Parkeol and Achilleol B Synthases

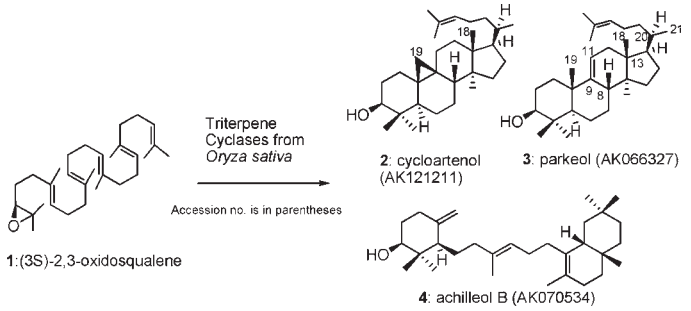
Ryousuke Ito, Kouya Mori, Ippei Hashimoto, Chiaki Nakano, Tsutomu Sato, and Tsutomu Hoshino*

Department of Applied Biological Chemistry, Faculty of Agriculture, and Graduate School of Science and Technology, Niigata University, Ikarashi 2-8050, Nishi-ku, Niigata, Japan 950-2181

hoshitsu@agr.niigata-u.ac.jp

Received March 24, 2011

ABSTRACT



The gene products of AK121211, AK066327, and AK070534 from *Oryza sativa* encode cycloartenol, parkeol, and achilleol B synthases, respectively. Parkeol synthase is a unique enzyme that affords parkeol as a single product. Achilleol B synthase is the third *seco*-type triterpene cyclase identified to date, and triterpenes produced by this synthase include achilleol B (90%), tetracyclic (5.12%) and pentacyclic scaffolds (4.37%), and unidentified triterpenes (0.51%). The pathway for achilleol B biosynthesis is proposed.

Naturally occurring cyclic triterpenes are biosynthesized by the enzymatic reaction of squalene or (3*S*)-2,3-oxidosqualene as the substrates. The structural diversity of cyclic triterpenes is remarkable. In addition to sterols (lanosterol and cycloartenol), approximately 100 additional triterpenoid scaffolds have been identified to date.[1] Lanosterol and cycloartenol can serve as precursors to membrane sterol and steroid hormones. Many genes encoding squalene-hopene cyclases (SHCs)[2] and oxidosqualene cyclases (OSCs)[3] have been elucidated; investigations of OSCs from *Arabidopis thaliana* have provided substantial insight into plant OSCs. Specifically, OSCs are generally classified into the following three categories: accurately functional OSC (to give a single or dominant product),[3,4] multifunctional OSC (to produce a variety of triterpene skeletons by a single OSC, none of which are dominant),[3,4] and a *seco*-type triterpene synthase OSC (C−C bond cleavage(s), i.e. Grob fragmentation). Only two examples of a *seco*-type triterpene synthase have been identified to date: marneral synthase (At5g42600, PEN5) from *A. thaliana*[5] and *seco*-β- or *seco*-α-amyrin synthase (At1g78500, PEN6).[6] It has generally been assumed that lanosterol is only produced by vertebrates and fungi and that it is not produced by

(1) Xu, R.; Fazio, G. C.; Matsuda, S. P. T. *Phytochemsitry* **2004**, *65*, 261–291.

(2) (a) Hoshino, T.; Sato, T. *Chem. Commun.* **2002**, 291–301. (b) Tippelt, A.; Jahnke, L.; Poralla, K. *Biochem. Biophys. Acta* **1998**, *1391*, 223–232.

(3) Phillips, D. R.; Rasbery, J. M.; Bartel, B.; Matsuda, S. P. T. *Curr. Opin. Plant Biol.* **2006**, *9*, 305–314.

(4) Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Kolesnikova, M. D.; Onak, C. S.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **2007**, *129*, 11213–11222.

(5) Xiong, Q.; Wilson, W. K.; Matsuda, S. P. T. *Angew. Chem., Int. Ed.* **2006**, *45*, 1285–1288.

(6) Shibuya, M.; Xiang, T.; Katsube, Y.; Otsuka, M.; Zhang, H.; Ebizuka, Y. *J. Am. Chem. Soc.* **2007**, *129*, 1450–1455.

(7) (a) Kolesnikova, M. D.; Xiong, Q.; Lodeiro, S.; Hua, L.; Matsuda, S. P. T. *Arch. Biochem. Biophys.* **2006**, *447*, 87–95. (b) Suzuki, M.; Xiang, T.; Ohyama, K.; Seki, H.; Saito, K.; Muranaka, T.; Hayashi, H.; Katsube, Y.; Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Plant Cell Physiol.* **2006**, *47*, 565–571.

10.1021/ol200777d
Published on Web 04/28/2011

plants. However, genome mining of *A. thaliana* triterpene cyclases has led to the identification of lanosterol synthase (At3g45130, LAS1).[7a,7b] Additionally, a plant lanosterol synthase was identified in *Lotus japonicus*, which is a dicotyledonous plant.[8] Recently, Muranaka et al. demonstrated that phytosterols are produced via both cycloartenol and lanosterol pathways (dual pathway).[9]

Rice is one of the most important cereals in the world. Some phytosterols (cycloartenol, stigmasterol, sitsterol, and campesterol) have been isolated from rice bran;[10] however, these triterpene derivatives including γ-oryzanol are only biosynthesized from a cycloartane scaffold. In 1999, Akihisa et al.[11] reported that monocyclic achilleol A[12] and *seco*-type graminol A[11] were present in rice bran. However, to the best our knowledge, no other cyclic triterpene skeletons have been isolated from *Oryza sativa*. In 2004, the genome project of *Oryza sativa* was completed.[13] We selected the following six gene candidates encoding triterpene cyclases based on a homology search: AK121211, AK066327, AK068026, AK067451, AK072702, and AK070534 (DDBJ accession no.). These genes have approximately 30% homology with *Sacchromyces cerevisiae* lanosterol synthase and *A. thaliana* cycloartenol synthase [Supporting Information (SI), Figure S1]. Herein, we describe the functional analyses of these six candidates. AK121211 encoded cycloartenol synthase. AK066327 encoded an enzyme that was defined as parkeol synthase. Each of the two cyclases gave rise to a single product. The product of AK070534 was achilleol B synthase, which yielded multiple products. No cyclic triterpene product was detected from the constructs of the other three gene candidates, AK068026,[14] AK067451,[14] and AK072702,[14] suggesting that the gene expression may have been unsuccessful in *S. cerevisiae* GIL77 or that these candidates are pseudogenes.

cDNA clones were obtained from the Rice Genome Resource Center at the National Institute of Agrobiological Sciences, Japan. The putative *osc* genes were amplified by PCR. The subcloned pYES2 vector harboring each of the *osc* genes was transformed into *S. cerevisiae* GIL77, which lacks lanosterol synthase (*gal2 hem3−6 erg7 ura3−167*),[15] after which the transformed yeasts were cultured and the triterpene cyclases were induced by adding galactose according to a previously published protocol.[15] Next, 15% KOH/MeOH was added to the harvested cells, and the samples were then refluxed for 1 h, after which the lipophilic materials were extracted with hexane and analyzed by GC/MS and SiO_2-TLC (SI, Figures S2 and S3).

The transformant carrying the pYES2 vector harboring the AK121211 gene was cultured and the lipophilic materials from the yeast cells were analyzed by GC/MS. Three triterpene products, **2**, **5**, and **6**, were identified in the GC trace at a ratio of 45.9:32.8:21.3. Purification was conducted as follows. The hexane extracts from the cultured cells were subjected to SiO_2 column chromatography and then eluted with hexane/EtOAc (100:1−3) to obtain the triterpene-enriched fraction, which was acetylated with Ac_2O/pyridine. The samples were then purified by 5% $AgNO_3$−SiO_2 column chromatography (hexane/EtOAc = 100:0.5−1.0) to give a pure acetate of cycloeucalenol **6** and a mixture of acetates of **2** and **5**. These acetates were then purified by normal phase HPLC (hexane/2-ProOH = 100:0.01).

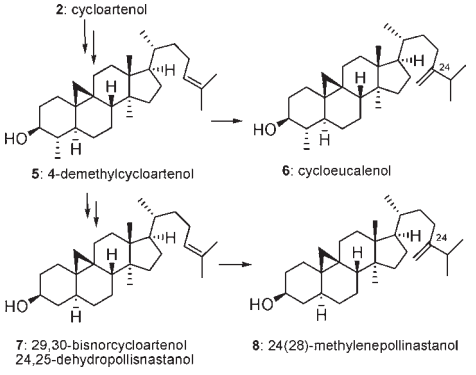


Figure 1. Metabolic pathway of cycloartenol **2** by *S. cerevisiae*. Two different pathways may exist: **2**→**5**→**6** or **2**→**5**→**7**→**8**.[16]

The structural determinations of these products are described in Figures S4−S6 (SI). Cyclopropane ring protons were found in the ^{1}H NMR spectra of **2**, **5**, and **6** [δ_H 0.54 (1H, J = 3.9 Hz), 0.25 (1H, J = 3.9 Hz) for **2**], which clearly indicated that this gene encodes cycloartenol synthase. Previously, Venkatramesh and Nes[16] reported that **2** is metabolized in *S. cerevisiae* into 29,30-bisnorcycloartenol **7** and 24(28)-methylene-pollinastanol **8**. This indicates that the dimethyl groups at C-4 are removed through oxidation and decarboxylation to give **7**, which is then converted into **8** with the aid of *S*-adenosylmethionine. However, **7** and **8** were not detected in our

(8) Sawai, S.; Akashi, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Ayabe, S.; Aoki, T. *Plant Cell Physiol.* **2006**, *47*, 673–677.

(9) Ohyama, K.; Suzuki, M.; Kikuchi, J.; Saito, K.; Muranaka, T. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 725–730.

(10) (a) Fang, N.; Yu, S.; Badger, T. M. *J. Agric. Food Chem.* **2003**, *51*, 3260–3267. (b) Luo, H.-F.; Li, Q.; Yu, S.; Badger, T. M.; Fang, N. *J. Nat. Prod.* **2005**, *68*, 94–97. (c) Akihisa, T.; Yasukawa, K.; Yamaura, M.; Ukiya, M.; Kimura, Y.; Shimizu, N.; Arai, K. *J. Agric. Food Chem.* **2000**, *48*, 2313–2319.

(11) Akihisa, T.; Koike, K.; Kimura, Y.; Sashida, N.; Matsumoto, T.; Ukiya, M.; Nikaido, T. *Lipids* **1999**, *34*, 1151–1157.

(12) Barrero, A. F.; Alvarez-Manzaneda, E. J.; Alvarez-Manzaneda, R. *Tetrahedron Lett.* **1989**, *30*, 3351–3352.

(13) DNA data bank of Japan http://www.ddbj.nig.ac.jp/.

(14) The AK068026 gene is N-terminally truncated by 88 amino acid residues when compared to the AK066327 ORF, but all amino acid residues in AK068026 are included in those of AK066327. The motif of DCTA, which is responsible for the proton attack on the epoxide ring, is not highly conserved in AK072702 or AK067451 (DTTA for AK067451, while DTLA for 072702). Despite the second amino acid T having been mutated into C, no production of triterpene was observed. Furthermore, a Kozak sequence (accatg, the consensus sequence for initiation of translation) was introduced into each of the three genes to improve the expression levels, but no cyclic triterpene was produced. For the Kozak sequence, see: Kozak, M. *Cell* **1986**, *44*, 283–292.

(15) (a) Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Eur. J. Biochem.* **1998**, *256*, 238–244. (b) Kawano, N.; Ichinose, K.; Ebizuka, Y. *Biol. Pharm. Bull.* **2002**, *25*, 477–482.

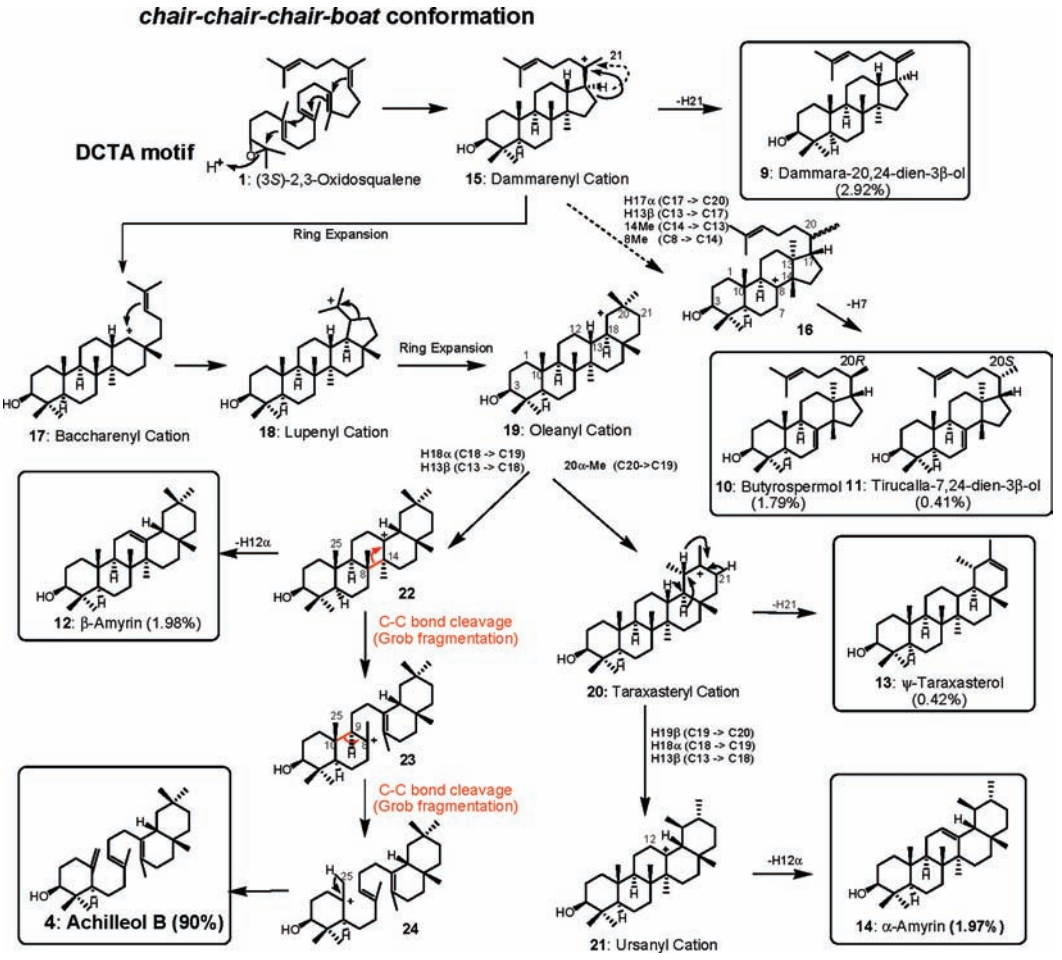


Figure 2. Products of yeast harboring the AK070534 gene and the metabolic pathways.The parentheses show the production ratio.

experiments. Instead, we isolated **5** and **6**, which indicates that methylation at C-24 occurred from **5**, but not from **7**. Taken together, these findings indicate that two different methabolic pathways of **2** may exist in *S. cerevisiae* (Figure 1).

GC/MS analysis confirmed that the yeast transformant harboring the AK066327 gene produced only one product (SI, Figure S2). The hexane extracts from 5 L cultured cells were subjected to SiO_2 column chromatography and elution with hexane/EtOAc (100:1−3), after which the partially purified product was acetylated with Ac_2O/pyridine. The acetate fraction was further purified by SiO_2 column chromatography using hexane/EtOAc (100:0.5−2), which gave the pure acetate of **3** (10.1 mg). Substrate **1** contains five double bonds and six allylic methyl groups in the molecule. Two double bonds and two allylic Me groups were identified on the ^1H and ^{13}C NMR spectra ($CDCl_3$): δ_H 5.22 (1H, bd, J = 5.6 Hz, H-11), 5.09 (1H, bt, J = 6.9 Hz, H-24); δ_H 1.60 (3H, s) and 1.67 (3H, s); δ_C 148.1 (s, C-9), 130.9 (s, C-25), 125.1 (d, C-24), 115.2 (d, C-11), suggesting that **3** had a tetracyclic skeleton. The following HMBC cross peaks were clearly observed: Me-19 (δ_H 1.06, 3H, s)/C-9 and H-11/C-13 (δ_C 44.27, s). These findings indicate that the double bond is located at C9−C11. A clear NOE between Me-19 and H-8 (δ_H 2.16, bm, 1H) supported the β-orientation for H-8. Thus, this product was assigned as parkeol (SI, Figure S7). The chemical shift of Me-21 in $CDCl_3$ was δ_H 0.857 (d, J = 6.0 Hz), which is very similar to the previously reported value of (20*R*)-parkeol (δ_H 0.87).[17]

The amounts of the triterpenes produced by the transformant harboring the AK070534 gene were very poor (see SI, Figure S2). GC/MS analysis of the hexane extract revealed a relatively large peak that showed m/z 426 as the molecular ion at the shorter retention time (rt, 16.8 min). In addition, several other peaks showing m/z 426 were also found at longer rt's (24−26 min), although they were present in negligible amounts. To isolate these products and determine the structures by NMR spectra, we cultivated the transformant on a large scale (150 L culture). SiO_2 column chromatography (hexane/EtOAc = 100:0−100:5) of the hexane extract (lipophilic materials)

(16) Venkatramesh, M.; Nes, W. D. *Arch. Biochem. Biophys.* **1995**, *324*, 189–199.

(17) Khalid, S. A.; Varga, E.; Szendrei, K.; Duddeck, H. *J. Nat. Prod.* **1989**, *52*, 1136–1138.

obtained after saponification revealed separation between the major product **4** and the minor products. The fraction containing **4** was acetylated and then completely purified by normal phase HPLC (hexane/2-PrOH = 100/0−100/2), which resulted in a yield of 4.2 mg. The detailed 2D NMR analyses (SI, Figure S8) supported the presence of monocyclic and bicyclic ring systems connected by the acyclic linear moiety, which led to the proposal of the structure of **4** (achilleol B) that was isolated from *Achillea odorata*.[18] Acetates of the minor products were subjected to normal phase HPLC (hexane/$CHCl_3$ = 100:0.1), which yielded six pure products (acetates **9**−**14**), and then to NMR analyses (SI, Figures S10−S15) to determine the structures. GC analyses (SI, Figures S2 and S9) showed that the production ratios of each product were as follows: **4:9:10:11:12:13:14** = 90:2.9:1.8:0.41:1.98:0.42:1.97.

The biosynthetic pathways for these products are shown in Figure 2. Proton attack of the epoxide ring of **1** triggered by the Asp residue of DCTA motif gave dammarenyl cation **15**. Proton elimination from Me-21 afforded dammara-20,24-dien-3β-ol **9**. A succesive migration of the hydride and Me groups in **15** gave **16** possessing the C8 cation, where two isomers of 20*R*- and 20*S*-configurations were generated, thus yielding butyrospermol **10** and tirucalla-7,24-dien-3β-ol **11** as a result of the removal of H-7. The tetracyclic cation **15** underwent ring expansion to afford pentacyclic lupenyl cation **18** via baccharenyl cation **17**. Further ring expansion led to oleanyl cation **19**. A 1,2-shift of C20α-Me to the C19 cation gave taraxasteryl cation **20**. Deprotonation of H-21 of **20** afforded Ψ-taraxasterol **13**. 1,2-Shifts of hydrides in an antiperiplanar fashion gave ursanyl cation **21**, which was followed by deprotonation of H-12α of **21** to yield α-amyrin. Hydride shifts of H-18α and H-13β of **19** gave cation **22**. Proton elimination of H-12α of **22** afforded β-amyrin **12**. The C−C bond cleavages (Grob fragmentations) of C8−C14 and C10−C9 gave **24**, followed by deprotonation from Me-25, which yielded achilleol B **4** (*seco*-type triterpene). The finding of the production of β-amyrin **12** (or pentacyclic oleanyl cation **19**) by the AK070534 enzyme could lead to the credible proposal for the formation mechanism of **4** that is shown in Figure 2.

Gene AF169966 (GenBank accession no.) from *Oryza sativa* has been suggested as the putative cycloartenol synthase gene;[3,15b,19] however, no functional evidence of this has been provided to date. We demonstrated here that the AK121211 gene, which is equivalent to AF169966[20] (SI, Figure S1), encodes cycloartenol synthase. Pearson et al.[21] reported that *Gemmata obscuriglobus* OSC produces both lanosterol and parkeol in an approximately 1:3 ratio, but triterpene cyclase that affords parkeol as a single product has not been reported to date; thus, the AK066327 gene is the first report of a gene encoding parkeol synthase.[22] An AK070534-encoded enzyme produces achilleol B as a main product (ca. 90%) in addition to minor products of six other triterpenes (ca. 10%). To date, only two genes that afford a *seco*-type triterpene have been identified from *A. thaliana*,[5,6] as described above. Thus, AK070534 is the third *seco*-type OSC. Folding of **1** in a *chair*−*boat*−*chair*−*boat* conformation in the enzyme cavity produces cycloartenol and parkeol via the protosteryl cation, while a *chair*−*chair*−*chair*−*boat* conformation of **1** affords **4** and **9**−**14** via dammarenyl cation **15**. Phylogenetic analyses of the six OSC genes (SI, Figure S16) showed that AK121211, AK066327, and AK068026 are grouped as a protosteryl cation intermediate type, while AK070534, AK072702, and AK067451 are of a dammarenyl cation type. These findings coincide with the results obtained from the functional analyses described herein. It is not surprising that Achilleol B synthase shares the same clade as β-amyrin synthase from *Avena strigosa* (monocot),[19a] because achilleol B synthase affords a negligible amount of β-amyrin (Figure 2).

In summary, we demonstrated here that the rice plant produces a variety of triterpene scaffolds in addition to a cycloartane skeleton, including parkeol, achilleol B, α- and β-amyrin, Ψ-taraxasterol, butyrospermol, tirucalla-7,24-dien-3β-ol, and dammara-20,24-dien-3β-ol. We did not recover achilleol A and graminol A, despite these compounds being identified in rice in a study conducted by Akihisa et al.[11] Further studies may be necessary to overview the triterpene cyclases of *Oryza sativa*.

Acknowledgment. We thank Professors Y. Ebizuka and M. Shibuya of The University of Tokyo and Niigata College of Pharmacy, respectively, for the kind gift of *S. cerevisiae* GIL77 and valuable suggestions regarding this investigation. A portion of this research was financially supported by a Grant-in-Aid for Scientific Research (B) (No. 18380001) to T.H. from the Japanese Society for the Promotion of Science.

Supporting Information Available. Experimental procedures, NMR, MS spectra, and phylogenetic tree. This material is available free of charge via the Internet at http://pubs.acs.org.

(18) (a) Barrero, A. F.; Manzaneda, R. E. A.; Manzaneda, R. R. A.; Arseniyadis, S.; Guittet, E. *Tetrahedron* **1990**, *46*, 8161–8168. (b) Arteaga, J. F.; Domingo, V.; Quilez del Moral, J. F.; Barrero, A. F. *Org. Lett.* **2008**, *10*, 1723–1726.

(19) (a) Haralampidis, K.; Bryan, G.; Qi, X.; Papadopoulou, K.; Bakht, S.; Melton, R.; Osbourn, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13431–13436. (b) Qi, X.; Bakht, S.; Leggett, M.; Maxwell, C.; Melton, R.; Osbourn, A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8233–8238.

(20) AK121211 has a full length ORF, but AF169966 lacks the two amino acid residues of MetTrp at the N-terminal site of AK121211 (see SI, Figure S1).

(21) Pearson, A.; Budin, M.; Brocks, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15352–15357.

(22) No detectable amount of lanosterol was found in the GC/MS chromatogram, despite the biosynthetic processes of parkeol and lanosterol being identical except for the final deprotonation site (C11 for parkeol, C8 for lanosterol).