Chimeric Triterpene Synthase. A Possible Model for Multifunctional Triterpene Synthase

Tetsuo Kushiro, Masaaki Shibuya, and Yutaka Ebizuka*

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

Received August 21, 1998

Abstract: Two triterpene synthases, β-amyrin synthase (EC 5.4.99.−) from *Panax ginseng* and lupeol synthase (EC 5.4.99.−) from *Arabidopsis thaliana*, were used to construct a series of chimeric proteins between these two enzymes in order to investigate the region important for product specificity. Functional expression in yeast and analysis of the synthase products have revealed that chimera 1, in which the N-terminal half is β-amyrin synthase and the C-terminal half is lupeol synthase, produced both β-amyrin and lupeol in a 3:1 ratio. By dividing the whole sequence into four regions, all the possible combinations of the two synthases were constructed. Among them, chimera 7, in which only region B (the second quarter from the N-terminus) is β-amyrin synthase, produced β-amyrin and lupeol in a 4:1 ratio, indicating the importance of region B in β-amyrin formation. Another chimera, which was created by the mixed PCR method, produced β-amyrin and lupeol in a 1:4 ratio, indicating that the sequence which is important for product distribution lies within 80 amino acid residues in region B. The incorporation experiment of [1,2-$^{13}C_2$]acetate showed that, during the formation of lupeol, the final proton abstraction takes place from either of the two *gem*-dimethyl groups in a 1:1 ratio. This is the first demonstration of the scrambling of methyl groups during the biosynthesis of any terpenoids. On the other hand, no scrambling of methyl groups was observed during β-amyrin formation, indicating that the isopropyl group of the lupenyl cation must be held tightly by β-amyrin synthase protein.

Introduction

Cyclization of 2,3-oxidosqualene into sterols and triterpenes is one of the most remarkable and fascinating biotransformations found in nature. The reaction is catalyzed by oxidosqualene cyclases (OSCs) which create tetra- and pentacyclic carbon frameworks, thereby generating a number of asymmetric stereocenters in a single biotransformation.[1] This fascinating reaction together with its importance in biological systems has attracted many researchers toward understanding the mechanism of this catalysis. The cyclization products of OSCs are widely distributed among eukaryotic organisms and play important roles in biological systems. In animals, fungi, and yeast, lanosterol is the ubiquitous cyclization product which serves as a precursor for various sterols and steroid hormones.[2] In plants, cycloartenol instead of lanosterol is the cyclization product which serves as a precursor of plant sterols.[3] Many of the mechanistic studies on these OSCs have been focused on those which are responsible for sterol biosynthesis. Up to now, cDNAs for five lanosterol synthases[4] and three cycloartenol synthases[5,6] have been cloned, and extensive studies such as affinity labeling, reaction with substrate analogues, and site-directed mutagenesis were carried out using either recombinant or purified enzymes to uncover some aspects of the complex cyclization mechanism.[7]

On the other hand, very little is known about the triterpene-producing OSCs. Triterpenoids are one of the most abundant natural products commonly occurring in plants and exhibit a wide range of structural diversity.[8] They are often glycosylated after some oxidative modifications and stored in underground parts as saponins. These triterpene saponins are often active constituents of important crude drugs used as traditional medicines, and are still considered as promising candidates for new drug development.[9] In some plants, they are found in epicuticular wax as triterpene alcohols and their esters without further structural modifications. These triterpenoids are regarded

* To whom correspondence should be addressed. Phone/Fax: +81-3-5689-2962. E-mail: yebiz@mol.f.u-tokyo.ac.jp.

(1) Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189−2206.

(2) Nes, W. D. *Isopentenoids and Other Natural Products*, ACS Symposium Series 562; American Chemical Society: Washington, DC, 1994.

(3) Goodwin, T. W. In *Biosynthesis of Isoprenoid Compounds*; Porter, J. W., Spurgeon, S. L., Eds.; Wiley & Sons: New York, 1980; pp 443−480.

(4) (a) Buntel, C. J.; Griffin, J. H. *J. Am. Chem. Soc*. **1992**, *114*, 9711−9713 (*Candida albicans*). (b) Corey, E. J.; Matsuda, S. P. T.; Bartel, B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2211−2215. Shi, Z.; Buntel, C. J.; Griffin, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7370−7374 (*Saccharomyces cerevisiae*). (c) Corey, E. J.; Matsuda, S. P. T.; Baker, C. H.; Ting, A.; Cheng, H. *Biochem. Biophys. Res. Commun*. **1996**, *219*, 327−331 (*Schizosaccharomyces pombe*). (d) Kusano, M.; Shibuya, M.; Sankawa, U.; Ebizuka, Y. *Biol. Pharm. Bull. Jpn*. **1995**, *18*, 195−197. Abe, I.; Prestwich, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9274−9278 (*Rattus norvegicus*). (e) Baker, C. H.; Matsuda, S. P. T.; Liu, D. R.; Corey, E. J. *Biochem. Biophys. Res. Commun*. **1995**, *213*, 154−160. Sung, C. K.; Shibuya, M.; Sankawa, U.; Ebizuka, Y. *Biol. Pharm. Bull. Jpn*. **1995**, *18*, 1459−1461 (*Homo sapiens*).

(5) (a) Corey, E. J.; Matsuda, S. P. T.; Bartel, B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11628−11632 (*Arabidopsis thaliana*). (b) Morita, M.; Shibuya, M.; Lee, M.-S.; Sankawa, U.; Ebizuka, Y. *Biol. Pharm. Bull. Jpn.* **1997**, *20*, 770−775 (*Pisum sativum*).

(6) Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Eur. J. Biochem*. **1998**, *256*, 238−244.

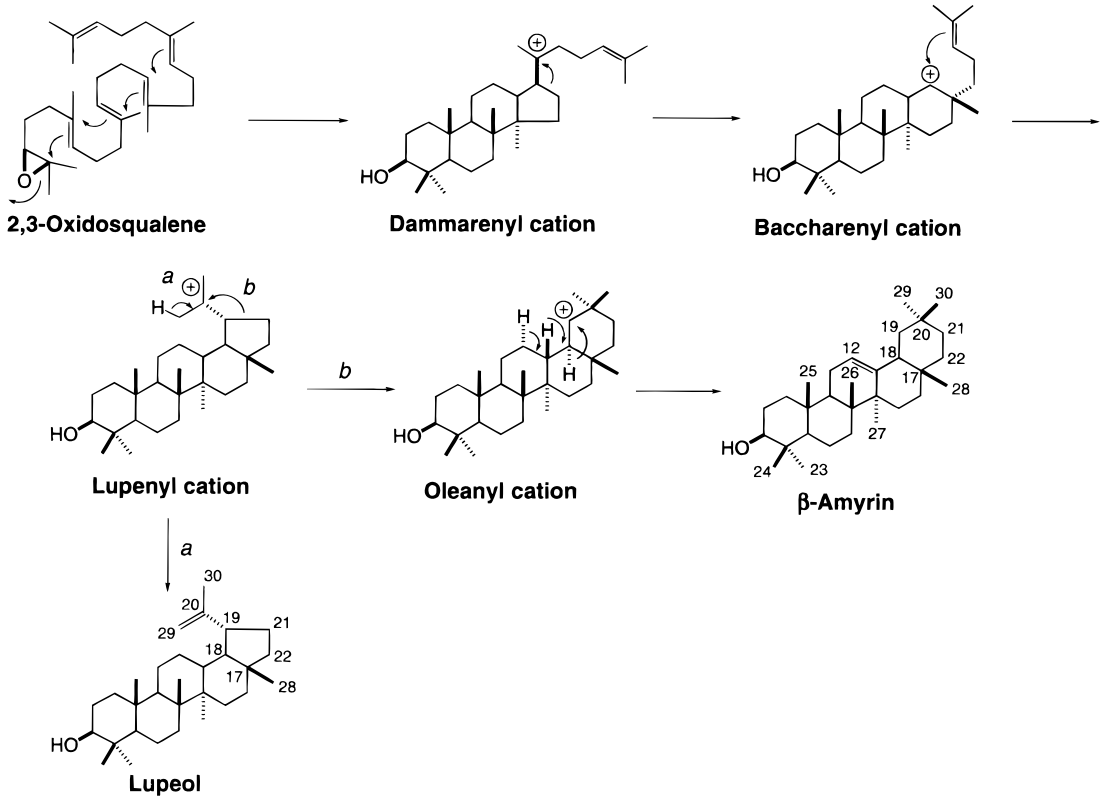
(7) (a) Abe, I.; Prestwich, G. D. *J. Biol. Chem*. **1994**, *269*, 802−804. (b) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1277−1288. Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc*. **1997**, *119*, 1289−1296.

(8) Das, M. C.; Mahato, S. B. *Phytochemistry* **1983**, *22*, 1071−1095.

(9) Waller, G. R.; Yamasaki, K. *Saponins Used in Traditional and Modern Medicine*; Plenum Press: New York, 1996.

10.1021/ja983012h CCC: $18.00
Published on Web 02/02/1999

Scheme 1. Cyclization of 2,3-Oxidosqualene into β-Amyrin and Lupeol



as secondary metabolites as they are found species specifically and their function in the producing plant is yet unknown. Generally, the cyclization mechanism of 2,3-oxidosqualene into various triterpenes is more complex than sterol biosynthesis as they, in most cases, involve additional ring closure to generate a pentacyclic carbon skeleton, and more extensive methyl and hydride shifts.[1] To understand how each triterpene synthase controls the complex reaction leading to the defined cyclization product would be a major issue for the origin of the structural diversity among these natural products and for the future engineering of rationally designed triterpenoid biosynthesis. Furthermore, it is of great interest to know whether there exist a corresponding number of product specific triterpene synthases, since it is quite common that one plant species produces more than one skeletal type of triterpenes.

Recently, two triterpene synthase cDNAs, β-amyrin synthase from *Panax ginseng* and lupeol synthase from *Arabidopsis thaliana*, have been cloned by our group[6] and by Matsuda et al.,[10] allowing us to investigate the mechanism controlling the product specificity. Comparison of the deduced amino acid sequences of these two OSCs revealed surprisingly high 70.4% identity, even though they are derived from plants belonging to different plant families and giving rise to different cyclization products. From this high level of sequence identity, we speculated that the difference of only a few amino acid residues should be responsible for the product specificity. Therefore, in this study, we set out a domain swapping study between these two OSCs to investigate which region of polypeptide sequences is responsible for product specificity. The feasibility of this approach was demonstrated by Chappell et al. for two sesquiterpene cyclases.[11]

As shown in Scheme 1, the cyclization mechanisms leading to lupeol and β-amyrin have the early part in common. The reaction is initiated by ring opening of epoxide, followed by cation-π cyclization via the pre-chair−chair−chair conformation to give the tetracyclic dammarenyl cation. Ring expansion to the baccharenyl cation and fifth ring closure give the tertiary lupenyl cation. The two products branch at this point, as proton abstraction from one of the *gem*-dimethyl groups gives lupeol, while ring expansion to the oleanyl cation followed by two 1,2-hydride shifts and proton abstraction from C-12 gives β-amyrin.

Construction of a chimeric enzyme and analysis of its cyclization products, if any, would give detailed information regarding the product specificity of the two OSCs, and hence should provide a detailed understanding of the mechanism of oxidosqualene cyclization.

Results and Discussion

Construction of Chimeric Clones and Expression in Yeast. The lupeol synthase used in this study was cloned by the PCR method. *A. thaliana* cDNA was prepared by reverse transcription of mRNA extracted from the *A. thaliana* total plant. Specific N-terminal and C-terminal primers containing the *Kpn* I site immediately upstream of the start codon and the *Xho* I site immediately downstream of the stop codon were designed on the basis of the reported cDNA sequence.[10] PCR using these primers and *A. thaliana* cDNA as a template produced a 2.3 kb full length fragment. Subcloning into the *Kpn* I and *Xho* I sites of yeast expression vector pYES2 (Invitrogen) under the *GAL1* promoter gave plasmid $pOSC_{LUP}$. Yeast mutant GIL77 (*gal2 hem3-6 erg7 ura3-167*), which lacks lanosterol synthase activity,[12] harboring $pOSC_{LUP}$, was cultured, and recombinant protein expressed as described previously.[6] Analysis of the extract of

(10) Reyna, J. B.; Bartel, B.; Matsuda, S. P. T. GenBank/EMBL data bank Acc. No. U49919.

(11) Back, K.; Chappell, J. *Proc. Natl. Acad. Sci. U. S.A.* **1996**, *93*, 6841−6845.

(12) Gollub, E. G.; Liu, K.; Dayan, J.; Adlersberg, M.; Sprinson, D. B. *J. Biol. Chem.* **1977**, *252*, 2846−2854.

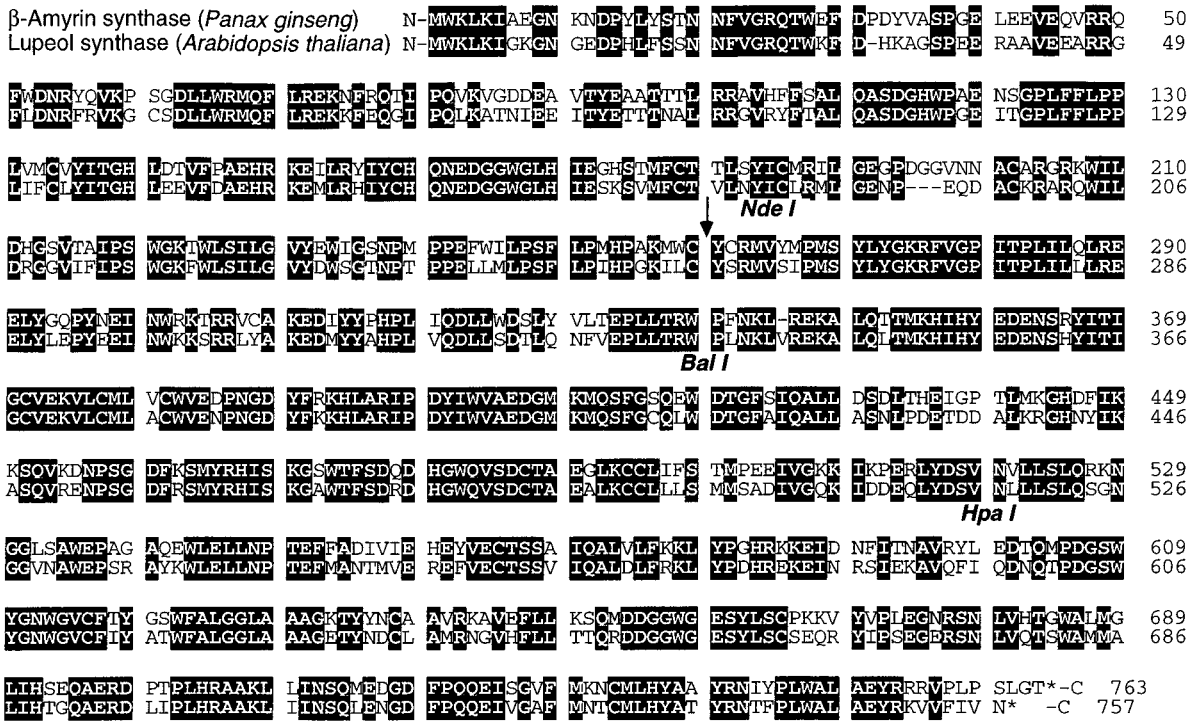


Figure 1. Alignment of amino acid sequences of *P. ginseng* β-amyrin synthase and *A. thaliana* lupeol synthase. Hyphens were introduced to maximize homology. Identical amino acid residues are boxed. A total of 537 of the 763 amino acids are identical (70.4%). Three restriction enzyme sites employed for chimera construction are noted. The vertical arrow indicates the position of scrambling for the pChi-mix6 clone.

this transformant confirmed the production of lupeol by means of reverse phase HPLC and ^{1}H and ^{13}C NMR analysis. The amount of lupeol produced was about half of the β-amyrin produced by a transformant harboring pOSC$_{PNY}$ described below. The obtained *A. thaliana* lupeol synthase clone was sequenced in both strands. However, four amino acid residues were different from the reported sequence. These are isoleucine instead of methionine at position 399, valine instead of glycine at 503, glutamate instead of lysine at 582, and isoleucine instead of tyrosine at 698. It is not clear whether these differences are due to the presence of an isomeric clone within the plant or mistakes caused by PCR.

For *P. ginseng* β-amyrin synthase, pOSC$_{PNY}$ which has been reported was used.[6]

To set up the domain swapping study, we divided the whole sequence into four regions with roughly equal length (Figure 1). Region A contains the well-conserved QW motif (sequence DGGWGLH 164−170),[13] while region D contains several conserved QW motifs. Region C contains the DCTAE motif which has been predicted to lie in the active site of lanosterol synthase,[7a] while region B contains a sequence corresponding to the WWVHTR (231−236) sequence of *Saccharomyces cerevisiae* lanosterol synthase which also has been demonstrated to be in the substrate binding pocket.[7b]

In our initial experiment, the chimeric enzymes were constructed using the common restriction site. Both β-amyrin and lupeol synthases contain the common *Bal* I site (nucleotide number 1018, amino acid number 340 in β-amyrin synthase, nucleotide number 1006, amino acid number 336 in lupeol synthase) which is located nearly in the center of the whole sequence. Thus, the *Bal* I and *Xho* I fragment of pOSC$_{LUP}$ was ligated into the *Bal* I and *Xho* I site of pOSC$_{PNY}$ to construct pChi 1 in which the N-terminal half is β-amyrin synthase and the C-terminal half is lupeol synthase, and *vice versa* for the construction of pChi 2. Introduction of each plasmid into yeast mutant GIL77 and functional expression were done as described before.[6]

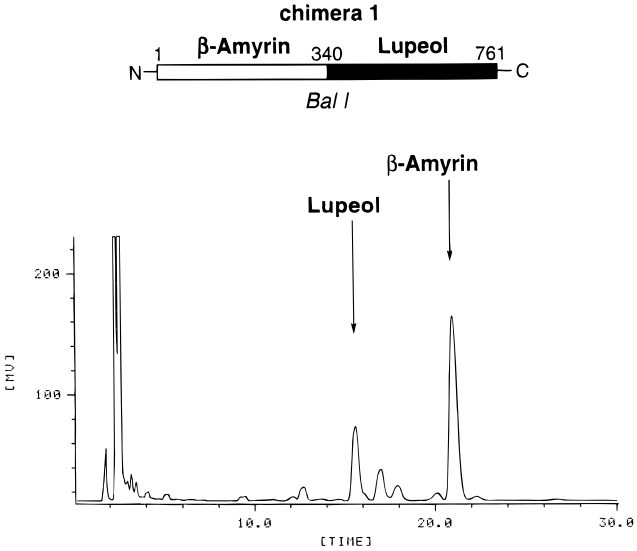


Figure 2. HPLC profile of the extract from chimera 1. The horizontal line indicates the retention time in minutes.

The yeast transformant with pChi 1 was cultured, expression induced by galactose, and the product extracted with hexane. Separation by TLC and analysis of 4,4-dimethyl sterol fraction by reverse phase HPLC were conducted as described before.[6] Surprisingly, two peaks corresponding to β-amyrin and lupeol in retention time were detected, with the ratio of about 3:1 (Figure 2). The identity of each product as β-amyrin and lupeol was confirmed by ^{1}H and ^{13}C NMR and by EI-MS analysis as well. The production level of the sum of two products was about the same as that of the native β-amyrin synthase. Generation of two products by a single chimeric protein is quite interesting in that each half of the polypeptides still reserves enough

(13) Poralla, K.; Hewelt, A.; Prestwich, G. D.; Abe, I.; Reipen, I.; Sprenger, G. *Trends Biochem. Sci.* **1994**, *19*, 157−158.

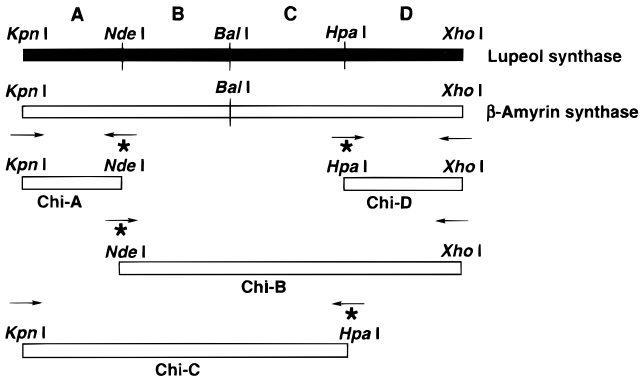


Figure 3. Schematic representation of the construction of chimeric clones. Asterisks indicate the positions of mutations on primers for restriction enzyme sites.

character to produce each cyclization product. In addition, as indicated in the ratio of the two products, the N-terminal half contributes more to the production of β-amyrin. It is not clear why this chimeric enzyme produced more than one cyclization product. One possible explanation for this phenomenon is that protein expressed in yeast might take two different conformations and each one is responsible for one product formation. Another possibility is that the chimeric enzyme forms a rather loose active site which leads to the formation of two products. It has been well documented in terpenoid synthases such as monoterpene and sesquiterpene synthases that a single enzyme is responsible for multiproduct formation.[14] At the moment, we cannot distinguish between these possibilities; however, the present result demonstrates the feasibility of this chimeric approach toward studies on the product specificity in triterpene synthases.

Interestingly, this chimeric enzyme also produced minor triterpene alcohols (as determined on the basis of EI-MS analysis to give *m*/*z* 426) which can be seen at R_t 17 and 18 min on HPLC. However, the amounts of these products were not sufficient to determine their structures (which will be reported in due course).

On the other hand, the yeast transformant harboring pChi 2 gave none of detectable cyclization products, indicating that this chimeric enzyme is inactive. This might be due to aberrant folding of the expressed chimeric protein; however, the exact reason for this observation is not clear. This result also shows some limitation to the chimeric studies in that proper combination of amino acid residues at the N-terminal and C-terminal halves is necessary for maintaining the active three-dimensional structure of the protein.

To further carry out the domain swapping study, we chose two additional restriction enzyme sites for chimera construction. These are the *Nde* I site (nucleotide number 549, amino acid number 184) and the *Hpa* I site (nucleotide number 1546, amino acid number 516) of lupeol synthase. Since β-amyrin synthase contains none of these sites, two restriction sites were generated by introducing point mutations in primers without changing the coded amino acids. Four primers, PNY-Nde-S, PNY-Nde-A, PNY-Hpa-S, and PNY-Hpa-A, were designed in both sense and antisense directions carrying the above-mentioned point mutations. With $pOSC_{PNY}$ as a template, PCR using these primers and either of the N-terminal or C-terminal β-amyrin synthase specific primers generated four fragments (pChiA-D) carrying the appropriate restriction sites (Figure 3). These fragments were sequenced in both strands to confirm that only the desired

(14) Croteau, R. *Chem. Rev.* **1987**, *87*, 929−954.

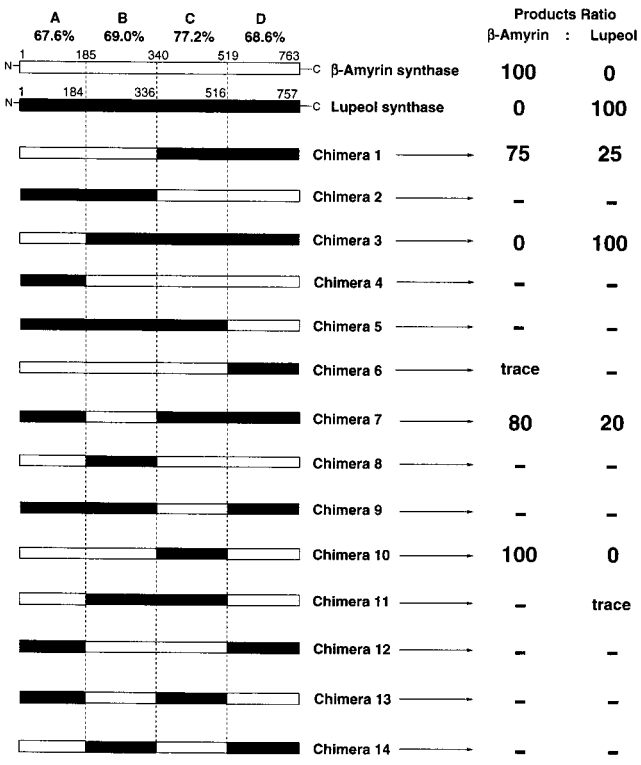


Figure 4. Chimera studies between β-amyrin and lupeol synthases. The identity of each domain is noted on the top.

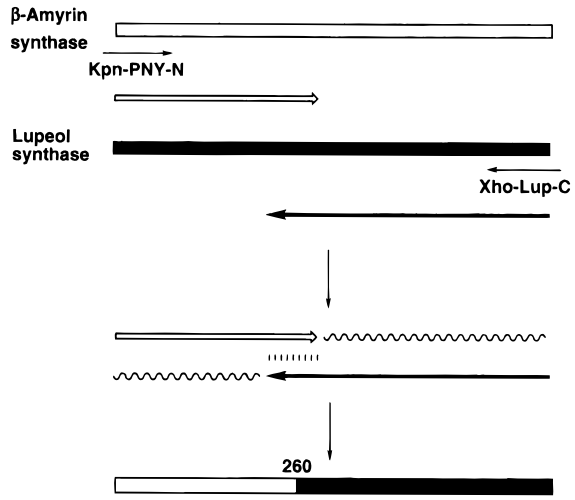
mutations had occurred and no other mutation due to mistakes in PCR had taken place. Restriction enzyme mediated chimera construction was carried out to obtain 12 chimeric clones, pChi 3 through pChi 14. These correspond to all the possible combinations of the two enzymes.

Expression in yeast revealed that seven of these chimeric clones were inactive, giving none of the cyclization product. The results from pChi 6 and pChi 11 showed diminished levels of production of each compound (about 1/10 of the active chimeric clone as judged by HPLC). However, the remaining three clones gave interesting results (Figure 4). pChi 3, in which only region A is β-amyrin synthase, produced lupeol as a sole product. On the other hand, pChi 7 in which only region B is β-amyrin synthase, produced both compounds with β-amyrin as a major product. These results strongly indicate that region B is extremely important for β-amyrin formation. pChi 10, in which only region C is lupeol synthase, produced β-amyrin only, indicating that region C is not important for product differentiation and again supporting the importance of region B for β-amyrin formation. The result of pChi 3 shows that region A is not important for product specificity; however, together with the fact that pChi 4 was inactive, this result indicates the importance of proper combination of amino acid residues between the N-terminal and the C-terminal ends.

As for the lupeol formation, it seems that regions C and D must be intact lupeol type as judged by the results of pChi 1, pChi 3, and pChi 7. This is further supported by the result from pChi 11 in that region D is replaced by β-amyrin synthase, giving diminished levels of lupeol. Comparison of the results of pChi 3 with pChi 1 and pChi 7 suggests that region B also alters the production of lupeol. This implies that region B reserves some characters for lupeol formation as well.

Chimeric Clone by Mixed PCR. We have also sought another method for generating chimeric clones between β-amyrin and lupeol synthases. Since both enzymes exhibit extremely high level of DNA sequence identity in some regions, we have

Scheme 2. Construction of the Chimeric Clone by Mixed PCR



expected a scrambling in PCR to take place at roughly the center of the whole sequence if we use both plasmids as a template and primers designed from each end of the sequences (Scheme 2). With the mixture of $pOSC_{PNY}$ and $pOSC_{LUP}$ as a template, primers Kpn-PNY-N designed for the 5′-end of β-amyrin synthase and Xho-Lup-C designed for the 3′-end of lupeol synthase were used for the following PCR. Each primer contains a *Kpn* I site immediately upstream of the start codon and an *Xho* I site immediately downstream of the stop codon in order to engineer the full length insert into yeast expression vector pYES2. PCR was carried out with the first 10 cycles with 94 °C, 1 min, 58 °C, 10 s, 72 °C, 10 s, and then 20 cycles with 94 °C, 1 min, 58 °C, 1.5 min, 72 °C, 1.5 min with final extension of 72 °C, 10 min. A band corresponding to a full length of 2.3 kb was detected on agarose gel electrophoresis. Control experiments using either $pOSC_{PNY}$ or $pOSC_{LUP}$ as a sole template gave no band corresponding to this position, indicating that scrambling has actually taken place. This full length fragment was subcloned into pYES2, and six colonies were picked up and then introduced into yeast mutant GIL77 for functional expression. Among these six clones, only one clone pChi-mix6 gave products comparable in amount to those of active chimeric clones. Both lupeol and β-amyrin were produced, but this time lupeol was the major product (lupeol:β-amyrin = 4:1). This clone pChi-mix6 was sequenced in order to figure out the position of scrambling. As a result, scrambling had taken place within region B, just at the sequence of MWCYC (258−262) of β-amyrin synthase, and ILCYS (254−258) of lupeol synthase (Figure 1). No other mutation due to PCR was detected. Unexpectedly, the scrambling took place at rather narrow region which is located further toward the N-terminus, indicating that antisense chain extension was faster during the initial PCR cycle.

The present results together with pChi 1 and pChi 3 revealed that a particular sequence in region B is extremely important for determining the product ratio for lupeol and β-amyrin. That is, the only 80 amino acid long sequence between Cys 260 and Trp 340 of β-amyrin synthase drastically alters the product ratio. pChi 1, in which the whole region B is β-amyrin synthase, produced β-amyrin and lupeol in a 3:1 ratio while pChi-mix6, in which these 80 amino acid residues are lupeol synthase, produced β-amyrin and lupeol in a 1:4 ratio. Comparison with the result from pChi 3 suggests that the sequence of region B up to Cys 260 (Ile 185 to Cys 260) of β-amyrin synthase still retains some residues in favor of β-amyrin formation.

The above results all together have indicated the importance of region B in β-amyrin formation and determining the product ratio between lupeol. These results agree well with the finding by Corey et al. that His 234 of *S. cerevisiae* lanosterol synthase is important for catalysis and might be located near the C-20 protosteryl cation in the active site of the enzyme,[7b] since branching into lupeol and β-amyrin formation takes place at the lupenyl cation stage and any factor from triterpene synthase proteins which affects the formation of β-amyrin or product distribution should be located near the E-ring of lupenyl and/or oleanyl cation. The location of the cation in the E-ring of these triterpenes is in close proximity to the position of the C-20 protosteryl cation during the formation of lanosterol if we assume the same orientation of substrate binding relative to the enzyme protein during cyclization.

Fate of Methyl Groups during the Biosynthesis of Lupeol and *β*-Amyrin. As the amino acid sequence responsible for the product specificity of β-amyrin and lupeol synthases has been narrowed down, we were intrigued by the actual mechanism leading to each different cyclization product. As shown in Scheme 1, the branch point between lupeol and β-amyrin formation is at the lupenyl cation stage. Either proton abstraction from one of the methyl groups to form *exo*-methylene, or migration of the carbon−carbon bond for ring expansion into a six-membered ring followed by successive 1,2-hydride shifts and H-12α proton abstraction, differentiates the two products. From the model, it is clear that H-12α and one of the terminal methyl groups in lupenyl cation are quite close in spacial distance. Therefore, it might be possible that the same basic amino acid residue in both β-amyrin and lupeol synthases is responsible for the final proton abstraction to terminate the reaction. To test this hypothesis, it is first important to determine from which of the prochiral methyl groups of lupenyl cation the proton is abstracted, since (*E*)-methyl of oxidosqualene, and thus of baccharenyl cation, which originates from C-2 of mevalonate, must be in the closer side to H-12α during the formation of lupenyl cation.

To investigate the origin of C-29 and C-30 of lupeol, a feeding experiment with [1,2-$^{13}C_2$]acetate was carried out using the yeast mutant GIL77 harboring $pOSC_{LUP}$. Since C-6 of mevalonate originates from intact incorporation of acetate, the biosynthetic origin of C-29 and C-30 could be easily distinguished, as one derived from C-6 of mevalonate should appear with a doublet while one from C-2 of mevalonate should appear as an enriched singlet. Doubly labeled acetate was fed to yeast culture during the galactose induction and further at the resting period with glucose as a carbon source. After extraction and isolation of lupeol, ^{13}C NMR was measured. The labeling patterns were all consistent with the expected mevalonate pathway. Unexpectedly, the signals due to C-30 (19.3 ppm) and C-29 (109.3 ppm) were both accompanied by doublets (J = 42.7 and 72.5 Hz, respectively) (Figure 5). In addition, the signal due to C-20 (151.0 ppm) was also accompanied by two sets of doublets (J = 42.0 and 72.5 Hz) (Figure 5). The scrambling of two methyl groups at the stage of DMAPP and/or IPP can be ruled out as signals of C-24 (15.4 ppm) and C-23 (28.0 ppm), the other side of the terminal *gem*-dimethyl group, exhibited ^{13}C signals with and without accompanying doublet, respectively. This result indicates that scrambling between C-29 and C-30 has taken place. That is, the final proton abstraction from lupenyl cation takes place from either of the two methyl groups in a 1:1 ratio, demonstrating that recombinant *A. thaliana* lupeol synthase does not control the proton abstraction from either of the two methyl groups. As far as we know, this is the first

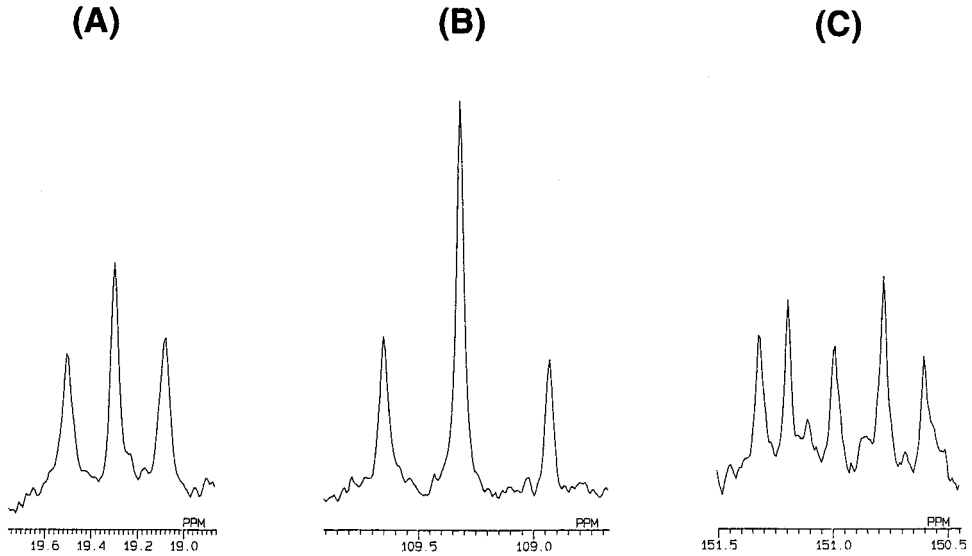


Figure 5. Partial ^{13}C NMR spectra (100 MHz, $CDCl_3$) of lupeol obtained from incubation with [1,2-$^{13}C_2$]acetate: (A) C-30 (19.3 ppm), (B) C-29 (109.3 ppm), and (C) C-20 (151.0 ppm).

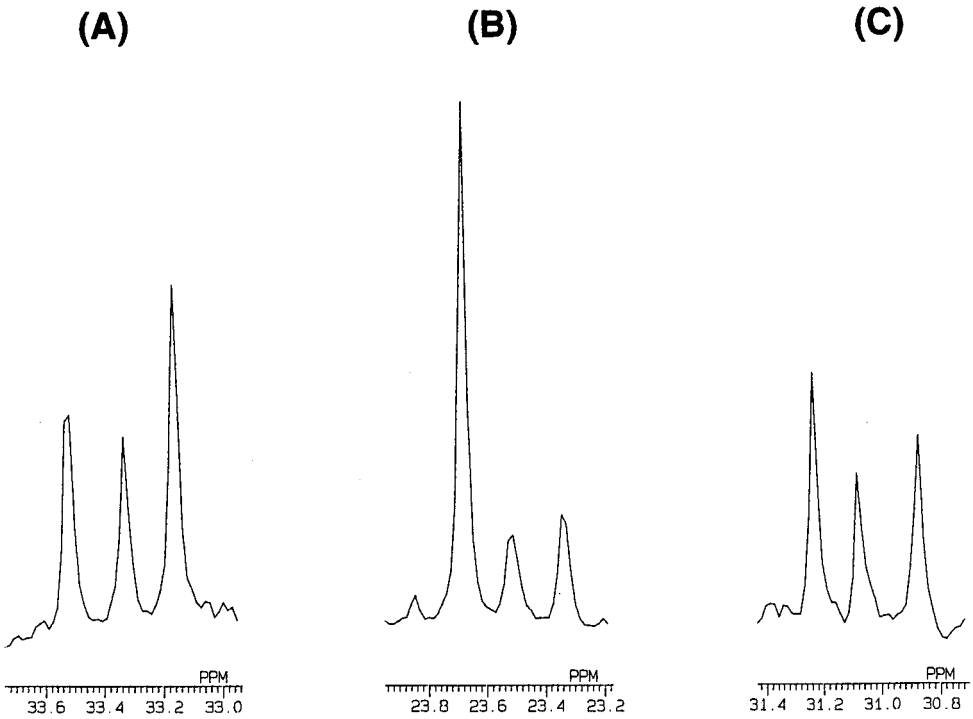


Figure 6. Partial ^{13}C NMR spectra (100 MHz, $CDCl_3$) of β-amyrin obtained from incubation with [1,2-$^{13}C_2$]acetate: (A) C-29 (33.3 ppm), (B) C-30 (23.7 ppm), and (C) C-20 (31.1 ppm). The C-30 signal was overlapped with the satellite peak of C-11 at 23.5 ppm.

demonstration of the scrambling of methyl groups during the biosynthesis of any terpenoids.

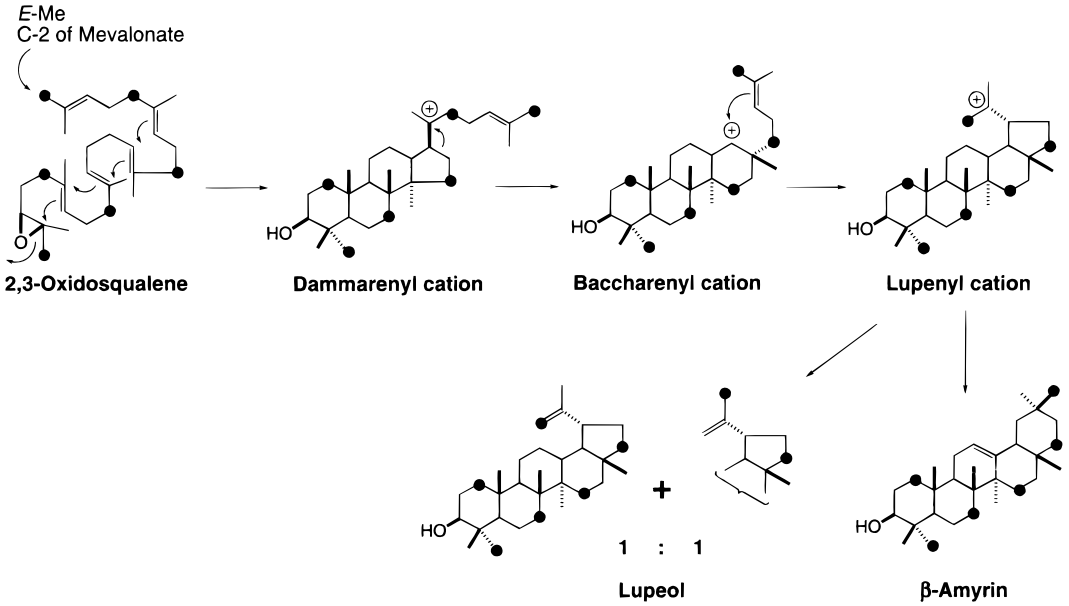
The result from lupeol synthase led us to test whether this scrambling also takes place during the formation of β-amyrin and see if this observation is general to all triterpenes or only specific to lupeol biosynthesis. Although C-29 and C-30 of β-amyrin have already been reported to originate from C-6 and C-2 of mevalonate, respectively, in cell suspension cultures of *Rabdosia japonica*,[15] the information regarding the scrambling of methyl groups, if any, is not available.

Feeding of [1,2-$^{13}C_2$]acetate was carried out in the same manner as described above using yeast mutant GIL77 harboring $pOSC_{PNY}$. The cells of yeast transformant were extracted with hexane, β-amyrin isolated, and measured for ^{13}C NMR.[16] The labeling pattern was again consistent with the expected mevalonate pathway. The signals which correspond to C-29 (33.3 ppm) and C-30 (23.7 ppm) appeared as one with an accompanying doublet (J = 35.1 Hz) and an enriched singlet, respectively (Figure 6), indicating that no scrambling had taken place. The signal due to C-20 (31.1 ppm) appeared with only one set of doublets (J = 35.1 Hz), supporting the above result. The present study indicates that, for the biosynthesis of β-amyrin, no scrambling of methyl groups takes place and that the ring expansion from lupenyl cation to oleanyl cation is

(15) Seo, S.; Yoshimura, Y.; Uomori, A.; Takeda, K.; Seto, H.; Ebizuka, Y.; Sankawa, U. *J. Am. Chem. Soc.* **1988**, *110*, 1740−1745.

(16) The assignment of ^{13}C signals due to C-2 and C-15 in the literature (Seo, S.; Tomita, Y.; Tori, K. *Tetrahedron Lett.* **1975**, 7) should be reversed in that the signal at δ 28.3 appeared with an accompanying doublet while the signal at δ 27.0 appeared as a singlet. Therefore, the signal for C-2 is δ 28.3 and that for C-15 is δ 27.0.

Scheme 3. Fate of Methyl Groups during β-Amyrin and Lupeol Formation (Dots Indicate the Carbons Originating from C-2 of Mevalonate)



completely controlled by the enzyme. The fact that C-30 of β-amyrin originates from C-2 of mevalonate demonstrates that the isopropyl moiety of lupenyl cation rotates in a clockwise direction (viewed from C-19) before C19−C21 bond migration takes place.

Whether these phenomena also operate in the chimeric clone which produces both β-amyrin and lupeol was further tested. Thus, feeding of [1,2-$^{13}C_2$]acetate to the yeast mutant GIL77 harboring pChi 1, in which β-amyrin and lupeol were producd in the ratio of 3:1, was carried out. The obtained β-amyrin and lupeol gave the same labeling patterns as above when analyzed with ^{13}C NMR (data not shown). Thus, for β-amyrin, the signal due to C-29 appeared with an accompanying doublet while C-30 appeared as an enriched singlet. For lupeol, the signals of both C-29 and C-30 were accompanied by doublets, indicating the scrambling of these methyl groups. These results show that, during β-amyrin formation, stereospecific ring expansion takes place in both native β-amyrin synthase and the chimeric clone, while for lupeol formation, the final proton abstraction is not controlled at all in both native lupeol synthase and the chimeric clone.

The results obtained here using labeled acetate have pointed out that, during lupeol biosynthesis, the final proton abstraction takes place from both terminal *gem*-dimethyl groups in equal ratios (Scheme 3). This raises a question whether the final proton abstraction is mediated by any specific amino acid residue or by a solvent water molecule within the active site of the enzyme. The results from β-amyrin synthase that the origins of carbons C-29 and C-30 are strictly controlled is reasonable, since ring expansion at the lupenyl cation stage requires a juxtaposition in which the empty p-orbital of the tertiary cation becomes almost parallel to the migrating C19−C21 σ bond. Such stereospecific ring expansion is only possible if the isopropyl cation moiety is held tightly by the enzyme protein. The observed formation of both β-amyrin and lupeol in some of the chimeric enzymes might be due to the lack of this tight holding of the isopropyl moiety caused by the change in some of the amino acid residues.

The results from the chimeric enzymes pointed out the importance of an 80 amino acid long sequence, spanning from Cys 260 to Trp 340 of β-amyrin synthase, in the formation of β-amyrin and controlling the product ratio between lupeol. Within these 80 amino acid residues, only 20 residues are different between these OSCs. Therefore, it would be of great interest to see which of these amino acid residues are responsible for the above-mentioned mechanism that controls the ring expansion step. Furthermore, the identification of the residue which is in charge of final deprotonation would give a more detailed insight into the mechanism of product specificity exhibited by these triterpene synthases. Our current efforts are directed to answer these important issues.

Conclusions

The domain swapping studies described here show for the first time that the product specificity of triterpene synthases is due to only a limited number of amino acid residues in the restricted region of the entire sequences. Considering the facts that a huge variety of triterpene skeletons, which are presumed to be formed by a slight difference in cyclization mechanism, occur in nature and that one plant species often produces several triterpenes of different skeletal types, it is rather questionable to believe the presence of the corresponding number of product specific triterpene synthases. For example, in *Ixeris chinensis*, 20 different triterpene alcohols were detected.[17] According to the recent report, two sesquiterpene cyclase clones isolated from *Abies grandis* catalyzed the cyclization of farnesyl diphosphate into 34 and 52 different sesquiterpenes, respectively, and these account for most of the sesquiterpenes found in the original plant.[18] Taking these observations into account, it would be possible to speculate such a multifunctional triterpene synthase also exists in nature. As described in this paper, product specificity of triterpene synthases is governed by only a few amino acid residues in a restricted region of polypeptide. It is now tempting to speculate that several point mutations during the course of evolution might have generated a multifunctional enzyme such as the chimeric ones described in this study. Our efforts toward cloning of several different triterpene synthases are now underway.

(17) Shiojima, K.; Suzuki, H.; Kodera, N.; Ageta, H.; Chang, H.-C.; Chen, Y.-P. *Chem. Pharm. Bull*. **1996**, *44*, 509−514.

(18) Steele, C. L.; Crock, J.; Bohlmann, J.; Croteau, R. *J. Biol. Chem*. **1998**, *273*, 2078−2089.

Experimental Section

Synthesis of oligo DNA was carried out by Nihon Bioservice (Saitama, Japan). Yeast strain GIL 77 (*gal2 hem3-6 erg7 ura3-167*) was constructed by crossing GL7 (*MAT***a** *gal2 hem3-6 erg7*)[12] to INVSC2 (*MAT*α *his3-D200 ura3-167*) (Invitrogen), sporulating the resultant diploid and selecting a segregant with the desired phenotype.[19] Yeast GIL77 was maintained on YEPD medium supplemented with ergosterol (20 μg/mL) and Tween 80 (5 mg/mL). ^{1}H and ^{13}C NMR spectra were recorded on a JEOL GSX400 spectrometer. Chloroform-*d* (99.8% atom ^{2}H, ISOTEC) was used as a solvent with solvent signal δ 7.26 for ^{1}H and δ 77.0 (t) for ^{13}C as references for chemical shifts. Merck 60 F_{254} (0.25 mm thickness, 20 × 20 cm) was used for silica gel TLC. All the other reagents were of reagent grade unless otherwise noted.

Cloning of Full Length *A. thaliana* Lupeol Synthase. Total RNA was extracted from the whole plant of *A. thaliana* by the phenol−SDS method and lithium chloride precipitaion. RNA (50 μg) was reverse transcribed using 0.5 μg of oligo(dT) primer (RACE 32, 5′-GACTCGAGTCGACATCGATTTTTTTTTTTTTT-3′) as in the literature,[20] and reverse transcriptase (Superscript II, BRL) with dNTP (0.2 mM) in a total volume of 20 μL for 2 h at 37 °C according to the manufacturer's protocol. The resulting cDNA mixture was diluted with 80 μL of Tris/EDTA (10 mM Tris−HCl, 1 mM EDTA, pH 7.4) and directly used as a template for the PCR. A sense primer was designed at the N-terminal region and an antisense primer was designed at the C-terminal region, according to the database sequence of *A. thaliana* lupeol synthase.[10] The *Kpn* I site and *Xho* I site were introduced immediately upstream of the ATG codon and downstream of the stop codon, respectively. These are 5′-GTAC**GGTACC**ATGTGGAAGTTGAAGATAGGA-3′ (*Kpn* I site in bold face) and 5′-AATAAGT**CTCGAG**TTAATTAACGATAAACAC-3′ (*Xho* I site in bold face). PCR was carried out with primers (1 μg each) using *Ex-Taq* DNA polymerase (Takara Shuzo) with dNTP (0.2 mM) in a final volume of 100 μL according to the manufacturer's protocol. The reaction was carried out for 30 cycles using Gradient 40 (Stratagene) with a program (94 °C, 1 min, 58 °C, 2 min, 72 °C, 3 min, and final extension at 72 °C, 10 min). The 2.3 kb PCR product was separated on agarose gel electrophoresis and purified using a Wizard PCR Preps Kit (Promega). This DNA fragment was digested with *Kpn* I and *Xho* I, ligated into the corresponding position of pYES2 (Invitrogen) and subcloned into *E. coli* strain NovaBlue (Novagen). Plasmid DNA was purified using a Wizard Midipreps Kit (Promega) to give $pOSC_{LUP}$. For sequencing of lupeol synthase, the full length clone was subcloned into pT7Blue (Novagen) and sequenced using a Thermo Sequenase Cycle Sequencing Kit (Aloka).

Functional Expression of *A. thaliana* Lupeol Synthase. $pOSC_{LUP}$ was introduced into yeast strain GIL77 using the lithium acetate method[19] and plated onto synthetic complete medium without uracil (SC−U) supplemented with ergosterol (20 μg/mL), hemin (13 μg/mL), and Tween 80 (5 mg/mL) and cultured at 30 °C for selecting the desired transformant. The transformant yeast was grown in liquid culture (1000 mL) for 2 days at 30 °C with shaking (220 rpm). Cells were collected and resuspended in SC−U medium without glucose (1000 mL), supplemented with ergosterol (20 μg/mL), hemin (13 μg/mL), and Tween 80 (5 mg/mL), and 2% galactose was added for induction at 30 °C for 10 h. Cells were again collected and resuspended in 0.1 M potassium phosphate buffer (pH 7.0) (580 mL) supplemented with 3% glucose and hemin (13 μg/mL) and cultured for 24 h at 30 °C. Cells were refluxed with 50 mL of 20% KOH/50% EtOH(aq) for 1 h, and the mixture was extracted with hexane three times, combined, and concentrated. The extract was purified by silica gel column chromatography with benzene as an eluent to give ca. 1 mg of lupeol. HPLC analysis was carried out using a SUPER-ODS column (4.6 × 200 mm) (TOSOH) with 95% CH_3CN(aq) as a solvent (flow rate 1.0 mL/min, detection UV 202) at 40 °C (retention time for lupeol: 16 min). ^{1}H NMR (400 MHz, $CDCl_3$) δ 0.76 (3H, s), 0.79 (3H, s), 0.83 (3H, s), 0.94 (3H, s), 0.97 (3H, s), 1.03 (3H, s), 1.68 (3H, s), 1.92 (1H, m), 2.37 (1H, m), 3.19 (1H, dd, J = 11, 5 Hz), 4.56 (1H, s), 4.68 (1H, s). EI-MS (JEOL JMS-SX 102A): m/z 426 [M^+] (90%); 408 [$M^+ - H_2O$] (40); 393 [$M^+ - H_2O - CH_3$] (20); 365 (40); 207 (80); 189 (100).

Construction of Chimera 1 and 2. $pOSC_{PNY}$ was digested with *Bal* I and *Xho* I and ligated with the 1.3 kb C-terminal half fragment of lupeol synthase obtained by digesting $pOSC_{LUP}$ with *Bal* I and *Xho* I to give pChi 1 in which the N-terminal half is β-amyrin synthase and the C-terminal half is lupeol synthase. pChi 2 was constructed similarly by digestion of $pOSC_{LUP}$ with *Bal* I and *Xho* I and ligation with the 1.3 kb C-terminal half fragment of β-amyrin synthase.

Construction of Chimera 3−14. All chimera construction was carried out using plasmid DNA pT7Blue, since pYES2 contained *Nde* I and *Hpa* I restriction sites within the plasmid. $pOSC_{PNY}$ and $pOSC_{LUP}$ were digested with *Kpn* I and *Xba* I and ligated into pT7Blue digested with the same restriction enzyme to give $pOSC_{PNY}$-T and $pOSC_{LUP}$-T, respectively.

Four primers were designed in order to generate fragments containing point mutations at nucleotides 555 and 1560 in β-amyrin synthase. These are PNY-Nde-S, 5′-ACTCTTAGCTA**CATATG**TATGCGTT-3′; PNY-Nde-A, 5′-ACGCATA**CATATG**TAGCTAAGAGTT-3′ (*Nde* I site in bold face and mutated nucleotide underlined); PNY-Hpa-S, 5′-GTATGATTCT**GTTAAC**GTGCTACTT-3′; PNY-Hpa-A, 5′-GTAGCAC**GTTAAC**AGAATCATACAG-3′ (*Hpa* I site in bold face and mutated nucleotide underlined).

PCR was carried out with $pOSC_{PNY}$-T (30 ng) as a template and the following pair of primers (120 pmol each) with dNTP (0.2 mM), $MgCl_2$ (1 mM), and *KOD* DNA polymerase (TOYOBO) according to the manufacturer's protocol; Kpn-PNY-N (N-terminal specific primer for β-amyrin synthase described before[6]) and PNY-Nde-A gave the 500 bp fragment Chi-A, PNY-Nde-S and Xho-PNY-C (C-terminal specific primer for β-amyrin synthase described before[6]) gave the 1.6 kb fragment Chi-B, Kpn-PNY-N and PNY-Hpa-A gave the 1.5 kb fragment Chi-C, and PNY-Hpa-S and Xho-PNY-C gave the 750 bp fragment Chi-D. PCR was carried out for 20 cycles with a program (98 °C, 1 min, 60 °C, 1 min, 74 °C, 1 min, and final extension at 74 °C, 10 min). All of the fragment was subcloned into pT7Blue and sequenced in both strands to confirm the desired mutation.

Chimera 3 was constructed by digesting $pOSC_{LUP}$-T with *Kpn* I and *Nde* I and ligated with the 500 bp Chi-A fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 3.

Chimera 4 was constructed by digesting $pOSC_{LUP}$-T with *Nde* I and *Xho* I and ligated with the 1.6 kb Chi-B fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 4.

Chimera 5 was constructed by digesting $pOSC_{LUP}$-T with *Hpa* I and *Xho* I and ligated with the 750 bp Chi-D fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 5.

Chimera 6 was constructed by digesting $pOSC_{LUP}$-T with *Kpn* I and *Hpa* I and ligated with the 1.5 kb Chi-C fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 6.

Chimera 7 was constructed by digesting pChi 4 with *Bal* I and *Xho* I and ligated with the 1.3 kb *Bal* I and *Xho* I C-terminal half fragment of lupeol synthase to give pChi 7.

Chimera 8 was constructed by digesting pChi 2 with *Kpn* I and *Nde* I and ligated with the Chi-A fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 8.

Chimera 9 was constructed by digesting $pOSC_{LUP}$ with *Bal* I and *Xho* I and ligated with the *Bal* I and *Xho* I C-terminal half fragment of pChi 6.

Chimera 10 was constructed by digesting pChi 1 with *Hpa* I and *Xho* I and ligated with the Chi-D fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 10.

Chimera 11 was constructed by digesting pChi 3 with *Hpa* I and *Xho* I and ligated with the Chi-D fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 11.

(19) Rose, M. D.; Winston, F.; Hieter, P. *Methods in Yeast Genetics*; Cold Spring Harbor Laboratory Press: New York, 1990.

(20) Frohman, M. A.; Dush, M. K.; Martin, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8998−9002.

Chimera 12 was constructed by digesting pChi 4 with *Bal* I and *Xho* I and ligated with the *Bal* I and *Xho* I C-terminal half fragment of pChi 6.

Chimera 13 was constructed by digesting pChi 7 with *Hpa* I and *Xho* I and ligated with the Chi-D fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 13.

Chimera 14 was constructed by digesting pChi 9 with *Kpn* I and *Nde* I and ligated with the Chi-A fragment digested with the same restriction enzymes. The resulting plasmid was digested with *Kpn* I and *Xho* I and introduced into pYES2 to give pChi 14.

All the chimeric clone was introduced into yeast strain GIL77 and expressed similarly as with the lupeol synthase case, and the cyclization products were analyzed by reverse phase HPLC. From a 20 mL scale culture for each sample, the following production level was measured: (chimera 1) β-amyrin 17 nmol, lupeol 5 nmol, (chimera 3) lupeol 11 nmol, (chimera 6) β-amyrin 1.7 nmol, (chimera 7) β-amyrin 28 nmol, lupeol 7 nmol, (chimera 10) β-amyrin 17 nmol.

Construction of the Chimeric Clone by Mixed PCR. PCR was carried out using both $pOSC_{PNY}$ and $pOSC_{LUP}$ (1 μg each) as a template with Kpn-PNY-N and Xho-Lup-C as primers (1 μg each) with dNTP (0.2 mM) and *Ex-Taq* DNA polymerase (Takara Shuzo) in a final volume of 100 μL according to the manufacturer's protocol. PCR was carried out for 10 cycles with a program (94 °C, 1 min, 58 °C, 10 s, 72 °C, 10 s), and then for 20 cycles with a program (94 °C, 1 min, 58 °C, 1.5 min, 72 °C, 1.5 min and final extension at 72 °C, 10 min). The resulting 2.3 kb full length fragment was digested with *Kpn* I and *Xho* I and subcloned into pYES2 to obtain pChi-mix6. Introduction into yeast and expression were done as described above. The production level was lupeol 10 nmol and β-amyrin 2.5 nmol from a 20 mL scale culture.

Feeding of [1,2-$^{13}C_2$]Acetate. [1,2-$^{13}C_2$]Sodium acetate (90% atom ^{2}H, MSD ISOTOPES) 85 mg was mixed with 165 mg of nonlabeled sodium acetate, dissolved in 2 mL of sterile water, and added to the transformed yeast culture (1000 mL) during the galactose induction and resting period. The culture condition and extraction procedure was the same as described for lupeol synthase. For pChi 1, β-amyrin, and lupeol were separated by reverse phase HPLC using an ODS-80T_M column (7.8 × 300 mm) (TOSOH) with 95% CH_3CN(aq) as a solvent (retention time for lupeol, 56 min; β-amyrin, 72 min).

Lupeol: ^{13}C NMR (100 MHz, $CDCl_3$) δ 14.5 (d, J = 36.6 Hz), 15.4 (d, J = 35.9 Hz), 16.0 (d, J = 36.6 Hz), 16.1 (d, J = 35.9 Hz), 18.0 (d, J = 35.9 Hz), 18.3 (d, J = 35.1 Hz), 19.3 (d, J = 42.7 Hz), 20.9 (d, J = 34.3 Hz), 25.1 (d, J = 34.3 Hz), 27.4 (d, J = 37.4 Hz), 27.4 (s), 28.0 (s), 29.8 (d, J = 32.0 Hz), 34.3 (s), 35.6 (s), 37.1 (d, J = 35.1 Hz), 38.0 (d, J = 34.3 Hz), 38.7 (s), 38.9 (d, J = 31.3 Hz), 40.0 (s), 40.8 (d, J = 36.6 Hz), 42.8 (d, J = 37.4 Hz), 43.0 (d, J = 35.1 Hz), 48.0 (d, J = 32.0 Hz), 48.3 (s), 50.4 (d, J = 35.1 Hz), 55.3 (d, J = 35.1 Hz), 79.0 (d, J = 36.6 Hz), 109.3 (d, J = 72.5 Hz), 151.0 (dd, J = 72.5, 42.0 Hz).

β-Amyrin: ^{13}C NMR (100 MHz, $CDCl_3$) δ 15.5 (d, J = 36.6 Hz), 15.6 (d, J = 35.1 Hz), 16.8 (d, J = 38.1 Hz), 18.4 (d, J = 35.1 Hz), 23.5 (d, J = 33.5 Hz), 23.7 (s), 26.0 (d, J = 33.5 Hz), 26.1 (s), 26.9 (s), 27.2 (d, J = 36.6 Hz), 28.1 (s), 28.4 (d, J = 36.6 Hz), 31.1 (d, J = 35.1 Hz), 32.5 (d, J = 33.6 Hz), 32.6 (s), 33.3 (d, J = 35.1 Hz), 34.7 (s), 36.9 (d, J = 38.2 Hz), 37.1 (s), 38.6 (s), 38.8 (d, J = 36.7 Hz), 39.8 (d, J = 36.6 Hz), 41.7 (d, J = 35.0 Hz), 46.8 (s), 47.2 (s), 47.6 (d, J = 35.1 Hz), 55.1 (d, J = 35.1 Hz), 79.0 (d, J = 36.6 Hz), 121.7 (d, J = 71.7 Hz), 145.2 (d, J = 71.7 Hz).

Acknowledgment. We are grateful to Professor W. D. Nes of Texas Tech University for providing us with a yeast mutant GL7. This work was in part supported by the Research for the Future Program (Grant JSPS-RFTF96I00302 to Y.E.) of The Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (A) (Grant 10358015 to Y.E.), and Grant-in-Aid for JSPS Fellows to T.K. from The Ministry of Education, Science, Sports and Culture, Japan.

JA983012H