Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

β-Amyrin synthase from *Euphorbia tirucalli*. Steric bulk, not the π-electrons of Phe, at position 474 has a key role in affording the correct folding of the substrate to complete the normal polycyclization cascade

Ryousuke Ito, Yukari Masukawa, Chika Nakada, Kanako Amari, Chiaki Nakano, and Tsutomu Hoshino*

Graduate School of Science and Technology, and Department of Applied Biological Chemistry, Faculty of Agriculture, Niigata University, Niigata 950-2181, Japan

*Corresponding author. Prof. Dr T. Hoshino; E-mail to hoshitsu@agr.niigata-u.ac.jp

Electronic Supporting Information for this article is available on the WWW under…….
Amino acid alignment, GC analyses of products generated by the various mutants and spectroscopic data of products **9**-**17** are described including ^{1}H-, ^{13}C-NMR, 2D NMR and EIMS. The enzymatic activities measured in vivo, the homology modeling, CD spectra and TTCED database are also included.

Abstract

β-Amyrin, a triterpene, is widely distributed in plants and its glycosides confer important biological activities. Mutagenesis studies on β-amyrin synthase are very limited as compared with those of squalene-hopene cyclase and lanosterol synthase. This study was conducted to elucidate the function of the F474 residue of *Euphorbia tirucalli* β-amyrin cyclase, which is highly conserved in the superfamily of oxidosqualene cyclases. Nine site-specific variants with Gly, Ala, Val, Leu, Met, Tyr, Trp, His, and Thr were constructed. We isolated 9 products from these mutants in addition to β-amyrin and determined the chemical structures. The Gly and Ala mutants produced significantly larger amounts of the bicyclic products and a decreased amount of β-amyrin, indicating that the F474 residue was located near the B-ring formation site. Surprisingly, the Ala variant produced (9β*H*)-polypoda-7,13,17,21-tetraen-3β-ol and (9β*H*)-polypoda-8(26),13,17,21-tetraen-3β-ol, which are generated from a chair-boat folding conformation. This is the first report describing the conformational change from the chair-chair into the chair-boat folding conformation among the reported mutagenesis studies of oxidosqualene cyclases. Substitution with aliphatic amino acids lacking π-electrons such as Val, Leu, and Met led to a significantly decreased production of bicyclic compounds, and in turn exhibited a higher production of β-amyrin. Furthermore, the Leu and Met variants exhibited high enzymatic activities: ca. 74% for Leu and ca. 91% for Met variants as compared to the wild-type. These facts unambiguously demonstrate that the major role of Phe474 is not to stabilize the transient cation via cation-π interaction, but is to confer the appropriate steric bulk near the B-ring formation site, leading to the completion of the normal polycyclization pathway without accumulation of abortive cyclization products.

Introduction

Triterpenes are one of the most abundant natural products and have important biological activities. Specifically, the glycosides, i.e., saponins, found in plants have been used as medicines.[1, 2] The structural diversity of triterpenes is remarkable as more than 100 different carbon skeletons are known.[1] Prokaryotes (bacteria) produce a 6,6,6,6,5-fused pentacyclic triterpene called hopene and hopanol. Animal and fungi also produce a 6,6,6,5-fused tetracyclic lanosterol skeleton. On the other hand, plant triterpenes exhibit remarkable structural diversity. Among plant triterpenes, the following are well known: cycloartenol (a primary metabolite) with a 6,6,6,5-fused tetracycle, β-amyrin consisting of a 6,6,6,6,6-fused pentacycle, and lupeol possessing a 6,6,6,6,5-fused skeleton. All of the triterpene scaffolds are constructed by enzymatic ring-forming reactions of squalene or (3*S*)-2,3-oxidosqualene (**1**), catalyzed by squalene cyclase and oxidosqualene cyclase (OSC), respectively.[3] To date, squalene-hopene cyclase (SHC)[3-5] from *Alicyclobacillus*

acidocaldarius and oxidosqualene-lanosterol cyclases from *Saccharomyces cerevisiae*[3e, 6] and hog liver[7] have been studied in detail by the protocols of the mutagenesis and/or substrate analogs. The hopene skeleton is biosynthesized from all pre-chair folding conformation.[1,3] Lanosterol is constructed by the *chair-boat-chair-chair* conformation of **1**.[7e, 8] β-Amyrin **2** is widely distributed in the plant kingdom, and constitutes the fundamental core of glycyrrhizin, a major bioactive compound derived from the underground parts of *Glycyrrhiza* (licorice) plant, which possesses a wide range of pharmacological properties. β-Amyrin scaffold is constructed by the folding of **1** in the *chair-chair-chair-boat-boat* conformation (Scheme 1).[9] However, the mutagenesis studies on β-amyrin synthase are very limited as compared with those of SHC and lanosterol synthase. In fact, only three papers [10a-c] have been published until now. The functions of the active site residues have been assigned only in light of the product distribution, but no enzymatic activity has been reported for the site-specifically mutated OSCs. By assessing the in vivo activities, we reported

<Scheme 1>

the experimental evidence indicating that the Asp of the $D^{485}C^{486}TA$ motif, which is highly conserved in OSCs, triggers the initiation of the polycyclization reaction and that the C564 is involved in hydrogen bond formation with the carboxyl residue of D485, resulting in an acidity enhancement.[10a] Scheme 1 shows the cyclization pathway to afford **2**. In a preceding paper, we addressed the function of Phe728 by mutagenesis experiments and found that the main role was to stabilize cationic intermediates by the π-electrons of Phe728, i.e., through cation-π interaction, which was inferred from the assessment of the enzymatic activities.[10b] Structures of the accumulated triterpenes by some of the site-directed variants indicated that Phe728 stabilizes the 6,6,6,6-fused baccharenyl **6** and oleanyl secondary cations **8**. Furthermore, we successfully isolated the triterpenes derived from malabaricanyl **4**, dammarenyl **5**, lupanyl **7,** and oleanyl cation **8**. Herein, we report the functional analysis of the Phe474 moiety of *Euphorbia tirucalli* β-amyrin synthase (EtAS), which is highly conserved in OSCs and SHCs (see Supporting Information Fig. S1). Mutations at position 474 afforded 9 new enzymatic products **9**-**17** in addition to β-amyrin **2**. The Ala and Gly mutants afforded a large amount of bicyclic products **9**-**13**, generated from bicyclic cations **3** and **3'**, suggesting that F474 is located in proximity to the B-ring formation site. Very interestingly, the F474A mutant afforded the 6,6-fused bicyclic products **9** and **10** with 9β-H stereochemistry, the configuration of which indicated that **1** was folded in a *chair-boat* conformation, differing from the normal *chair-chair* conformation that leads to 9α-H stereochemistry (see cation **3** in Scheme 1). Thus, we succeeded in altering the conformation by mutation, and this is the first report to induce an energetically unfavorable boat conformation from the favorable chair conformation (chair→boat). Seven other variants, i.e., Val, Leu, Met, Thr, His, Tyr, and Trp mutants, were constructed to address the exact function of the F474 residue. The relative enzyme activity measured *in vivo* and the product distribution ratio suggested that steric bulk at position 474 was more critical to the completion of the normal polycyclization pathway

rather than the π-electrons of Phe, i.e., the stabilization of the carbocation intermediate *via* cation-π interaction. Herein, we describe the structural determination of products generated by the mutants and the polycyclization pathway leading to the products, and discuss the function of the F474 residue based on the enzymatic activities measured *in vivo*.

Results and Discussion

Identification of triterpene products generated by site-specific mutants

Nine site-directed variants were constructed by substituting Phe474 with Gly, Ala, Val, Leu, Met, Thr, His, Tyr, and Trp. Mutagenesis of F474 in the wild type pYES2-EtAS/CT, which was constructed by us,[10a,b] was performed using the QuikChange site-directed mutagenesis method, and then transformed into *Saccharomyces cerevisiae* GIL77 lacking the gene encoding lanosterol synthase. The products were monitored with GC (Supporting Information Fig. S2). The GCMS analyses showed that the Ala mutant produced 8 triterpenes other than β-amyrin **2**. Large-scale cultures (50 L) enabled us to isolate all of the triterpene products. The mutant yeast cells collected by centrifugation were suspended in 1 L of 15% KOH/MeOH and refluxed 3 times, followed by extraction of the lipophilic materials with *n*-hexane (1 L × 9). The cyclic triterpene-rich fraction was obtained by partial purification on an SiO_2 column to remove oxidosqualene **1**, dioxidosqualene, and non-triterpene impurities using a mixture of hexane:EtOAc as the eluent (100:5 or 100:20). The triterpene-containing fraction (ca 320 mg) was acetylated with Ac_2O/py and further purified on an SiO_2 column with hexane:EtOAc (100:0.5) to completely remove the dioxidosqualene. Figure 1 shows the GC profile of the acetate fraction from the Ala mutant.

<Figure 1>

The acetate derivatives were subjected to repeated HPLCs (Inertsil ® Sil-100A) using a mixture of hexane and THF (100:0.1 or 100:0.07) as eluents to afford the pure triterpene products **9**-**16**. Structures of all of the enzymatic products were determined by detailed NMR analyses including DEPTs, ^{1}H-^{1}H COSY, HOHAHA, NOESY, HSQC, and HMBC. Substrate **1** has 8 methyl groups. The ^{1}H-NMR spectrum of product **9** (600 MHz, $CDCl_3$) showed the presence of 5 vinylic methyls (each 3H, s) at δ_H (ppm) 1.67, 1.61, 1.61, 1.60, and 1.68. Four olefinic protons were observed at δ_H (ppm) 5.22 (1H, bs) and 5.12 (3H, m), indicating that 4 double bonds were present in compound **9**. This finding suggested that **9** had a bicyclic framework. In the HMBC spectrum, the two methyl groups Me-23 (δ_H 0.865, 3H, s) and Me-24 (δ_H 0.934, 3H, s,) had correlations with C-3 (δ_C 81.28, d), C-4 (δ_C 37.51, s), and C-5 (δ_C 41.84, d). Me-25 (δ_H 0.902, 3H, s) also showed HMBC cross peaks with C-5, C-1 (δ_C 33.75, t), and C-9 (δ_C 53.80, d). Furthermore, the vinylic Me-26 (δ_H 1.67, 3H, s) exhibited HMBC correlations with C-9 and C-7 (δ_C 118.8, d). These findings indicated that **9** had a bicyclic scaffold. To our surprise, a definitive NOE was observed between H-9 (δ_H 1.23, m)

and Me-25 (Supporting Information Fig. S3.1.7), indicating the 9β-H configuration. Thus, **9** was determined to be (9β*H*)-polypoda-7,13,17,21-tetraen-3β-ol, which has never been reported as a natural product before. This structure clearly suggested that the boat folding conformation (Scheme 2) was allocated to the B-ring construction of **9**, which is in sharp contrast to the chair folding conformation for the B-ring construction of β-amyrin scaffold leading to the 9α-H configuration (see cation **3** in Scheme 1). Product **10** (600 MHz, $CDCl_3$) showed 4 olefinic methyl groups (δ_H 1.57, 1.60, 1.60, and 1.68; 3H, s for each of the Me signals) and 3 olefinic protons (δ_H 5.12, m, 3H) in addition to one methylidene group (CH_2-26: δ_H 4.54, 1H, bs; 4.71, 1H, bs), indicating that **10** was also a bicyclic product. H-26 had HMBC correlations with C-8 (δ_C 148.5, s), C-9 (δ_C 57.44, d), and C-7 (δ_C 30.94, t). A strong NOE between Me-25 (δ_H 0.930, 3H, s) and H-9 (δ_H 1.59, m) was confirmed (Supporting Information Fig. S3.2.7), verifying that **10** also had (9β*H*)-stereochemistry. Thus, **10** was also generated from the chair-boat conformation for the construction of the A/B-fused ring system. The polypodatetraene skeleton is assigned for bicyclic products with the (9α*H*)-configuration.[12] Compound **10** was reported to be one of the enzymatic products by *S. cerevisiae* lanosterol synthase Tyr707 mutants.[6e] The detailed 2D NMR analyses indicated that products **11** and **12** were also bicyclic triterpenes. However, an NOE was not found between Me-25 and H-9, but an unambiguous NOE was observed between H-5 and H-9 (see Supporting Information Fig. S3.3.7, S3.3.8, S3.4.7, S3.4.8), demonstrating that **11** and **12** had 9α-H configurations and were produced in the chair-chair conformation in accordance with the biosynthetic pathway of β-amyrin. Products **11** and **12** were isolated from *Cratoxylum cochinchinens*.[11] The 3-deoxy-derivatives of **11** and **12** were first isolated as fern constituents from *polypodiaceous* and *aspidiaceous* plants.[12] The ^{1}H NMR spectrum of product **13** (600 MHz, C_6D_6) showed 5 vinylic methyl groups (δ_H 1.70, 1.72, 1.76, 1.80, 1.80; 3H, s, for each of the signals). In the HMBC spectrum, Me-26 (δ_H 1.72, 3H, s) had cross peaks with C-8 (δ_C 126.2, s) and C-9 (δ_C 139.9, s). Therefore, the structure of **13** could be assigned as polypoda-8(9)-13, 17, 21-tetraen-3β-ol. Lodeiro et al. pointed out that **11**-**13** could be produced as artifacts from **1** with a prolonged treatment with SiO_2.[13a] However, the high production of **11** and **12** negated the possibility of the artifacts. In addition, there was no indication that **9** and **10** were produced as artifacts,[13a] and they were produced in a relatively high yield (Table 1). Thus, **9**-**12** were enzymatic products, not artifacts. Product **13** was isolated as the enzymatic product of the *Arabidopsis thaliana* At5g42600 gene product.[13b] Products **14** and **15** were previously isolated from the triterpene mixture of the F728 variant[10b] and the structures of **14** and **15** were identified by NMR and GC-MS, which were completely identical to those of the triterpenes isolated in the genome mining study of *Oryza sativa*.[14] The ^{1}H NMR spectrum of product **16** (400 MHz, C_6D_6) showed no vinylic Me group, indicating that **16** was the fully cyclized product of **1**. Me-27 (δ_H 1.115, 3H, s) had HMBC correlations with C-14 (δ_C 158.1, s), C-18 (δ_C 49.13, d), and C-12 (δ_C 33.97, t). Me-26 (δ_H 1.199, 3H, s) also had a HMBC cross peak with C-14. Thus, the double bond was located at C14-C15. The HMBC cross peaks from Me-28 (δ_H 0.933, 3H, s) were definitively found for C-16 (δ_C 38.02, t)

and C-22 (δ_C 35.42, t). Both Me-29 and Me-30 showed HMBC correlations with C-20 (δ_C 28.99, s), C-19 (δ_C 33.43, t), and C-21 (δ_C 37.00, t). A clear NOE was observed between Me-28 (δ_H 0.933, 3H, s) and H-18 (δ_H 1.17, 1H, ms). The detailed NMR analyses indicated that the structure of **16** was taraxerol as shown in Scheme 2. Recently, the gene encoding taraxerol synthase was cloned from *Kalanchoe daigremontiana*.[15] The Met variant produced **17**, which was not found in any other mutants. The yeast encoding EtAS F474M was cultured (24 L). As described above, the cell pellets collected by centrifugation were suspended in 15% KOH/MeOH and refluxed, followed by extraction of lipophilic materials with *n*-hexane. The obtained non-saponifiable lipids (645 mg) were subjected to column chromatography on SiO_2 (hexane: EtOAc=100:3-5) to yield the triterpene fraction including dioxidosqualene, followed by acetylation with Ac_2O/py. Repeated chromatography on SiO_2 (hexane:EtOAc=100:0.4) removed dioxidosqualene and gave the enriched fraction of the triterpene acetates. Normal phase HPLC (hexane:THF=100:0.025) successfully afforded the pure acetate of **17** (10.9 mg). In the ^{1}H NMR spectrum (600 MHz, $CDCl_3$), no olefinic methyl proton was observed, indicating that a pentacyclic skeleton could be assigned to **17**. Me-26 (δ_H1.070, 3H, s) had HMBC correlations with C-8 (δ_C 147.6, s), C-13 (δ_C 37.02, s), C-14 (δ_C 41.60, s), and C15 (δ_C 31.66, t). In the HOHAHA spectrum, H-7 (δ_H 5.46, 1H, br s) had cross peaks with H-6 (δ_H 1.98,1H, m; 2.12, 1H, m) and H-5 (δ_H 1.36,1H, m). These data indicated the double bond was positioned at C7-C8. Me-27 (δ_H 1.082, 3H, s) had HMBC correlations with C-12 (δ_C 36.06, t), C-13, C-14, and C18 (δ_C 46.84, d). Me-28 (δ_H 1.055, 3H, s) showed HMBC cross peaks for C-16 (δ_C 36.08, t), C-17 (δ_C 30.94, s), C-18, and C-22 (δ_C 36.57, t). From both Me-29 and Me-30 (δ_H 0.976, 3H, s, and 0.966, 3H, s, respectively), clear HMBCs were found for C-19 (δ_C 34.60, t), C-20 (δ_C 28.23, s), and C-21 (δ_C 33.87, t). In the NOESY spectrum, the following correlations were found: H-5/H-9, Me-25/Me26, Me-26/H-18, Me-28/H-18, and H-18/Me-30. Thus, product **17** was determined to be multiflorenol as shown in Scheme 2. Kushiro and coworkers reported **17** as one of the enzymatic products of *Arapidopsis thaliana* LUP 5.[16] The structures of all triterpenes **9**–**17** (Scheme 2) were confirmed by detailed 2D NMR analyses (Supporting Information Fig. S3.1~S3.9).

Polycyclization pathway of 1 to form products 9–17

Scheme 2 shows the cyclization pathway of **1** to yield each of the enzymatic products **9**-**17**. (3*S*)-**1** was folded in a chair-chair conformation to give bicyclic cation **3**, while the chair-boat conformation afforded cation **3'**. Elimination of H-7 in the cationic intermediates **3'** and **3** could produce **9** and **12**, respectively. Deprotonation of Me-26 in **3'** and **3** afforded **10** and **11**, respectively. Cation **3** was further cyclized to give malabaricanyl cation **4** with a 6,6,5-fused tricyclic ring system, which underwent a ring enlargement to provide a 6,6,6-fused tricyclic cation, followed by a further cyclization to give 6,6,6,5-fused tetracyclic cation **5** (damamrenyl cation). During this cyclization process, two different folding conformations were possible: a chair-chair-chair-*boat* conformation led to cation **5**, while a chair-chair-chair-*chair* conformation

furnished **5'**(17-*epi*-dammarenyl cation). The anti-periplanar 1,2-migrations of the hydride and methyl group, followed by elimination of H-7, led to the production of **14** and **15**, which could be produced from the false intermediate **5'** and the actual intermediate **5**, respectively. According to Scheme 1, cation **5** underwent a ring expansion to give cation **6** which then cyclized to provide **7**. Further ring expansion afforded **8** with a C-19 cation. A successive antiparallel 1,2-shifts of the hydrides gave **8** with a C-13 cation: H-18α to C-19 and H-13β to C-18. Removal of 12α-H gave **2**. The shift of Me-27 to the C-13 cation and deprotonation of 15β-H (in anti-periplanar fashion) gave **16**. The antiparallel 1,2-shifts of Me-27 to C-13 and Me-26 to C-14, followed by the elimination of 7α-H, provided **17**.

<Scheme 2>

Product distribution ratio generated by the site-specific variants

Table 1 shows the product distribution ratios (%) generated by the mutants. We reported that the wild-type EtAS produced tetracyclic products **14** and **15** in very small amounts (3.3%), in addition to a large amount of **2** (96.7%).[10a,b] The Gly and Ala variants afforded a much higher production of bicyclic compounds (**9**~**12**) up to 75~94%, but the production of **2** was significantly decreased (5.9~21%), indicating that the polycyclization reaction terminated at the premature bicyclic stage by these mutations and that F474 was located near the B-ring formation site of the reaction cavity, which was further supported by homology modeling of β-amyrin synthase prepared based on the X-ray crystal structure of human lanosterol cyclase (Supporting Information Fig. S4). This finding suggests that the π-electrons of the F474 residue stabilizes the transit C-8 carbocation intermediate via cation-π interaction, leading to the complete polycyclization reaction. However, the substitution with aliphatic amino acids lacking π-electrons, such as Val, Leu, and Met, gave significantly decreased amounts of the bicyclic compounds, and in turn exhibited a higher production of **2** (Table 1). This finding strongly indicated that the appropriate steric size, nearly equivalent to Phe (wild-type), was more critical to the completion of the polycyclization reaction of **1** than the π-electron density (see van der Waal's volume [17] shown in the legend of Figure 2). The importance of the steric volume was further supported by the finding that the Ala variant was superior to the Gly mutant in terms of the production of **9** and **10** (chair-boat conformation). The Thr variant also afforded bicyclic product **9** (from a chair-boat folding, 1.8%) and products **11** and **12** (from a chair-chair folding, 36.6%), together with a decreased production of pentacycles **2** and **16** (55.3%). The high production of the bicyclic compounds by the Thr mutant would be ascribed to the decreased steric bulk as found for the Gly and the Ala mutants.

<Table 1>

Thus, by altering the steric bulk and taking into account the product distribution ratios, it can be concluded that an appropriate steric volume is required for the completion of the polycyclization to form **2**. The substitutions of Phe with other aromatic amino acids such as His, Tyr, and Trp led to significant differences in the product distribution ratios. The Tyr variant showed a similar product distribution as the wild-type (Phe), albeit with a small amount of **16** (4.9%). Furthermore, a bicyclic compound was not produced. The steric bulk of Tyr is the similar as that of Phe (the wild-type), but the former is slightly larger than the latter (see van der Waal's volume in Figure 2). Thus, the Tyr mutation would have led to a slight production of **16** via a shunt pathway. The Trp mutant has a largest steric bulk among the natural amino acids. This would have guided the high production of bicyclic compound **11** (30 %) and the relatively decreased formation of **2** (70%). This finding also provides additional evidence that the appropriate steric volume at this position is essential to acquiring the normal polycyclization pathway. As for the His variant, a small amount of bicyclic **12** (3.0 %) and a high production of **2** (ca 90%) was observed, thus the product distribution was not much different from that of the Met or Leu variant (Table 1). Indeed, the steric volumes of the three amino acid residues Leu, Met, and His are similar (see van der van der Waal's volume in Figure 2).[17] It is noticeable that alteration of the steric bulk, irrespective of increase or decrease in steric size, afforded both bicycles **9**~**12** and pentacycles **16** and**17**. The mutations at position 474 could bring about the local change of the protein architecture (namely, the B-ring formation site). Thus, a somewhat improper orientation of **1** near the B-ring could lead to the production of the bicycles. The inappropriate disposition of **1** inside the reaction cavity could further perturb the D/E-ring formation sites and the deprotonation sites could be altered to form **16** and **17**.

Enzymatic activities of the mutants relative to that of the wild-type enzyme.

With the exception of our preceding papers,[10a, b] few reports describing the expression levels of OSC mutants are available. In order to assess the *in vivo* enzymatic activities, the quantities of the OSC protein expressed and the triterpene products accumulated in the yeast GIL77 must be estimated. The expressed amounts of OSC are frequently found to be none or significantly decreased. In previous papers,[10a, b] we introduced a western blot protocol to quantitatively measure the amounts of the expressed EtAS enzymes. All of the variants produced substantial amounts of EtAS enzymes, as shown in Supporting Information Fig. S5.1. The amounts of triterpenes were quantified by GC analysis using geranylgeraniol as the internal standard (see Fig. S5.2 for total amounts of **2**, **16** and **17**; Fig. S5.4 for those of **14** and **15**; Fig. S5.6 for those of **11** and **12**; Fig. S5.8 for those of **9** and **10**). The enzymatic activities were estimated as follows: the amounts of triterpene produced were divided by the amounts of the expressed enzymes. Fig. S5.3, Fig. S5.5, Fig. S5.7 and Fig. S5.9 depict the enzymatic activities for the production of **2**, **16** and **17**, those for **14** and **15**, those for **11** and **12**, and those for **9** and **10**, respectively. Figure 2 depicts the enzymatic activities of each variant for the production of all the triterpenes, which are plotted against the activity of the wild-type (100%). The relative activities of the F474H and the F474W variants were

ca 25% and ca 5% of the wild-type, respectively. On the other hand, the Tyr mutant showed ca 92 % activity. Mecozzi et al. reported that the cation-π electron binding energies of the aromatic rings with Na^+ were as follows: 21.0 for His, 27.1 for Phe, 26.9 for Tyr, and 32.6 kcal/mol for Trp,[18] thus it is envisaged that the activity of the Tyr mutant is nearly equivalent to that of the wild-type (Phe) and that the Trp mutant has a most significantly enhanced activity, but the activity of the Trp variant was modest (Figure 2). The significantly decreased activity of the Trp mutant was also found in the F365W mutant of *Alicyclobacillus acidocaldarius* SHC (Supporting Information, Fig. S1),[4e, q] and also in the F728W mutant of EtAS,[10b] despite the two Phe residues having been demonstrated to stabilize the transient cation through cation-π interaction. When the activities of the His, Tyr and Trp mutants were evaluated, it seemed that the aromatic side chains likely acted as π-electron donors to stabilize the transient cation **3** or **3'** via cation-π interaction, as discussed in the functional analysis of F728.[10b] However, the mutants replaced with aliphatic amino acids, especially Leu and Met variants possessing no aromatic π-electrons, also exhibited high activities to yield pentacycles from oleanyl cation **8**; ca 74% for Leu and ca 91% for Met variants. This fact unambiguously demonstrated that the major role of EtAS Phe474 is not to stabilize the transient cation via cation-π interaction, but is to confer the appropriate steric bulk near to the B-ring formation site, so that the normal polycyclization pathway was achieved without accumulation of abortive cyclization products. The van der Waal's volume of Tyr (0.6115 nm^3) is not much different from that of Phe (0.55298 nm^3), although Tyr is slightly larger than Phe. This would have given rise to nearly the same activity between the Tyr mutant (ca 92%) and the wild-type EtAS. The markedly larger volume of Trp (0.79351 nm^3) than that of Phe could give a significantly decreased activity (5% of the wild-type) and the concomitant production of bicyclic compound **11**. This could be explained as follows:

<Figure 2>

the Trp residue occupied a large volume in the active site and caused a local conformational change at the B-ring formation site. This idea was supported also by the functional analyses of EtAS F728W[10b] and SHC F365W mutants.[4e, q] When the enzymatic activities for the production of **2**, **16,** and **17** were analyzed, the amounts of the pentacylces produced by each of the mutants increased in proportion to the van der Waal's volume, i.e., Gly<Ala<Val<Leu<Met<Tyr mutants (see the arrow shown in Figure 2). However, the production of **2**, **16**, and **17** by the Thr and His variants were significantly decreased, although the van der Waal's volume of Thr is between those of Ala and Val. Furthermore, the volume of His is nearly the same as that of Leu and Met. This finding suggests that a different mechanism(s) is involved in the cyclization reaction of the Thr and His variants. The basic imidazole ring readily forms the hydrogen-bonded network with the neighboring amino acid residues, possibly resulting in significant perturbation of the active site. Thus, the marked decrease in the activity was observed. The polar hydroxyl group in Thr also has a similar function, i.e., hydrogen bonding with other amino acids, such as the imidazole ring, thus

leading to a marked decrease in the activity. Why did the Gly and Ala mutants give a larger amount of the bicyclic compounds? The B-ring formation site of the Gly and Ala mutants would be loosely packed due to the introduction of significantly decreased steric volumes. Thus, substrate **1** may have been improperly situated around the B-ring formation site in the reaction cavity. The quantities of the bicyclic products were gradually decreased when the van der Waal's volumes came close to the steric size of Phe. The volume of Thr is larger than that of Ala, but smaller than that of Val, which may have thus conferred somewhat higher or decreased activity of the Thr variant compared to that of the Val or the Ala mutants, respectively (13.4 % for the Thr, 6.8% for the Val and 34% for the Ala mutants, see Figure 2 and the sum of the values shown in Figs. S5.7 and S5.9, Supporting Information). We measured the circular dichroism spectra (CD) of the mutants in order to inspect whether the protein structure of the mutated EtAS was altered (Supporting Information Fig. S6). However, the change in the CD spectra was very small if at all visible, probably suggesting that the local change of the protein architecture could not be detected by CD. In summary, the enzymatic activities estimated here demonstrated that the steric volume at position 474 was critical to the completion of the normal polycyclization pathway, but the π-electrons of Phe do not play a major role in the catalytic action. This is the first report indicating the more crucial role of the steric bulk of the aromatic active site rather than the π-electrons in the polycyclization cascade by triterpene cyclases.

Conclusions

The EtAS F474 residue corresponds to the F365 moiety of SHC in the amino acid alignment (Supporting Information, Fig. S1). We have definitively demonstrated that the F365 of SHC plays a key role in the stabilization of the bicyclic cation (3-deoxy-**3**) through cation-π interaction.[4e, q] This suggests that the π-electrons of F474 may also have a critical role for the enzymatic catalysis via cation-π interaction, but this study demonstrated that the steric bulk at position 474 has a more important role for β-amyrin construction, rather than π-electron densities. In the case of SHC, the enzyme activities of the F365Y and the F365W mutants were significantly decreased, despite the π-electron densities of Tyr and Trp being increased, compared to that of Phe (the wild-type).[4q] The B-ring formation site would be tightly packed. Thus, introduction of Tyr and Trp with larger steric bulks into the F365 position induced the local conformational change, which led to the decreased enzymatic activities resulting from the loose binding of the squalene substrate.[4e] By the substitution of F365 with unnatural mono-, di-, trifluorophenylalanines, we first succeeded in demonstrating that the higher electron density of the π-electrons gave rise to higher enzymatic activities by determining the K_m and k_{cat}.[4e] A fluorine atom is extremely electronegative but has a similar van der Waal's radius compared to that of a hydrogen atom. Thus, the function of F365 could be assigned to stabilization of the cationic intermediate through cation-π interaction. On the other hand, cation-π interaction cannot be assigned to the function of F474, because the Leu and Met variants with no aromatic π-electrons had significantly enhanced activities (oleanane-type skeleton: 60% and 82% of the wild type,

respectively, see Supporting Information Fig. S5.3). Furthermore, oleanane-type triterpenes were produced in proportion to the van der Waal's volume (Figure 2). This finding implies that β-amyrin synthases might exist in nature where the aliphatic amino acid(s) is substituted at the position corresponding to the F474 of EtAS. *Arabidopsis thaliana* LUP2 (Supporting Information Fig. S1) produces β-amyrin, α-amyrin and lupeol (55:30:15).[19] Recently, we elucidated the function of the triterpene cyclase from rice plant (AK070534), which produces achilleol B.[14] Achilleol B is a seco-type triterpene[20] and is biosynthesized from oleanyl cation **8** through twice C-C bond cleavages (Grob fragmentations).[14] Recently, Qi et al. reported the β-amyrin synthase (AsbAS1) from *Avena strigosa*.[21,22] In the three triterpene cyclases, Leu residue is replaced rather than Phe at this position (Supporting Information Fig. S1). Recently, sequence and structure information on 639 triterpene cyclases were assembled by Racolta et al., who reported the Triterpene Cyclase Engineering Database (abbreviated as TTCED).[23] The Phe474 residue of EtAS is highly conserved in 85% of OSC superfamily, but this position is substituted with Leu in 13% of the putative OSCs (see Table S1 in Supporting Information). These findings suggest that aliphatic Leu residue at this position also work for the normal polycyclization reaction as exemplified in Figure 2. In this study, we succeeded in the preparation of (9βH)-polypodatetraenes (**9** and **10**), which were formed through the chair-boat conformation. This is the first report to describe the conformational change from chair-*chair* into chair-*boat* in triterpene biosyntheses. Wu et al. also reported the alteration of from the chair-*boat*-chair conformation into the chair-*chair*-chair conformation by mutating *S. cerevisiae* lanosterol cyclase Phe 699, yielding the tricyclic malabaricatriene core (boat→chair for B-ring).[6c] In the mutagenesis experiments of F445 in *S. cerevisiae* lanosterol synthase, which corresponds to EtAS F474 in the amino acid alignment (see Fig. S1), the following triterpenes were isolated: isomalabaricatriene, lanosta-7, 24-diene-3β-ol and parkeol, but no production of bicyclic compounds was reported, [6g] despite EtAS F474 mutants creating substantial amounts of bicyclic compounds **9**~**12**. This finding suggests that the protein structure of EtAS is somewhat different from that of *S. cerevisiae* lanosterol synthase. It was assumed that the function of the F445 in the lanosterol synthase is to stabilize the C14 cation of the tricyclic intermediate and to direct the final deprotonation site during lanosterol biosynthesis. [6g] However, the major role of EtAS F474 is not to stabilize the transient cation through cation-π interaction, but is to guide the normal polycyclization pathway to yield β-amyrin where the deprotonation site is not altered. In conclusion, the enzyme architecture, the location of EtAS F474 in the reaction cavity and the function for the catalysis would be somewhat different from those of *S. cerevisiae* lanosterol synthase and *Alicyclobacillus acidocaldarius* SHC.

Experimental Section

Instruments

NMR spectra were recorded in $CDCl_3$ or C_6D_6 with a Bruker DMX 600 or DPX400 spectrometer. The chemical shifts (δ) are given in ppm relative to the residual solvent peak as the internal reference (δ_H=7.26 and δ_C=77.0 ppm for $CDCl_3$; δ_H=7.28 and δ_C=128.0 ppm for C_6D_6). GC analyses were done on a Shimadzu GC-2014 chromatograph fitted with a flame ionization detector (J&W DB-1 capillary column, 0.25 mm x 30 m). GC/MS spectra were obtained with a JMS-Q1000 GC K9 (JEOL) instrument under electronic impact at 70 eV with a Zebron ZB-5ms capillary column (0.25 mm x 30 m), the oven temperature being elevated from 220 to 270 ºC (3 ºC/min).

Mutagenesis experiments

Mutagenesis of Phe474 in wild-type pYES2–EtAS/CT was performed with the QuikChange site-directed mutagenesis method. The following oligonucleotide primers were used; substitutions are underlined.

For the F474G variant: sense primer, 5’-GGATCATGGACTGGCTCTGATCAGGATCATGGTTGG -3’; and antisense primer, 5’-CCAACCATGATCCTGATCAGAGCCAGTCCATGATCC -3’.

For the F474A variant: sense primer, 5’-CCAAAGGATCATGGACTGCCTCTGATCAGGATC -3’; and antisense primer, 5’-GATCCTGATCAGAGGCAGTCCATGATCCTTTGG -3’.

For the F474V variant: sense primer, 5’-GGATCATGGACTGTCTCTGATCAGGATCATGGTTG -3’; and antisense primer, 5’-CAACCATGATCCTGATCAGAGACAGTCCATGATCC-3’.

For the F474L variant: sense primer, 5’-CCAAAGGATCATGGACTTTGTCTGATCAGGATC -3’; and antisense primer, 5’-GATCCTGATCAGACAAAGTCCATGATCCTTTGG -3’.

For the F474M variant: sense primer, 5’-CCAAAGGATCATGGACTATGTCTGATCAGGATCATGGTTGG -3’: and antisense primer, 5’-CCAACCATGATCCTGATCAGACATAGTCCATGATCCTTTGG-3’.

For the F474Y variant: sense primer, 5’-CCAAAGGATCATGGACTTACTCTGATCAGGATC -3’; and antisense primer, 5’-GATCCTGATCAGAGTAAGTCCATGATCCTTTGG-3’.

For the F474W variant: sense primer, 5’-CCAAAGGATCATGGACTTGGTCTGATCAGGATC-3’: and antisense primer, 5’-GATCCTGATCAGAACCAGTCCATGATCCTTTGG-3’.

For the F474H variant: sense primer, 5’-CCAAAGGATCATGGACTCACTCTGATCAGGATC-3’: and antisense primer, 5’-GATCCTGATCAGAGTGAGTCCATGATCCTTTGG-3’.

For the F474T variant: sense primer, 5’-CCAAAGGATCATGGACTACCTCTGATCAGGATC-3’: and antisense primer, 5’-GATCCTGATCAGAGGTAGTCCATGATCCTTTGG-3’.

PCRs were conducted with a 16-cycle program: 98 °C for 0.5 min, 60 °C for 1 min, 68 °C for 9 min, and a final extension at 68 °C for 9 min. KOD-plus DNA polymerase (Toyobo) was used with dNTPs (0.2 mM), dimethylsulfoxide (5%) and $MgSO_4$ (0.1 mM) in a final volume of 50 μL. Mutations were confirmed by DNA sequencing with the dideoxy chain-termination method and a Beckman Coulter CEQ8000 (Beckman Coulter).

The constructions and the cultural conditions of the mutants.

The detailed methods were reported in the previous papers. [10a,b]

Measurements of *in vivo* protein expression levels and CD measurement.

These experimental protocols were described in the preceding paper.[10a,b]

Spectroscopic data of enzymatic products 9 ~17 acetates.

(9β*H*)-Polypoda-7,13,17,21-tetraen-3β-ol, i.e., (9β*H*)-γ-Polypodatetraen-3β-ol acetate.
^{1}H-NMR (600 MHz, $CDCl_3$) δ0.902 (3H, s, Me-25), 1.18 (1H, m, H-11), 1.23 (1H, m, H-9), 1.28 (1H, m, H-1), 1.38 (1H, dd, *J*=11.4, .5.4 Hz, H-5), 1.51(1H, m, H-11), 1.599 (3H, Me-29), 1.606 (6H, s, Me-27 & Me-28), 1.666 (3H, bs, Me-26), 1.678 (3H, s, Me-30), 1.70 (2H,m, H-2), 1.71 (1H, m, H-1), 1.86 (1H, m, H-6), 1.96 (1H, m, H-6), 1.98 (4H, m, H-15 & H-19), 2.02 (2H, m, H-12), 2.063 (3H, s, C$\underline{H}_3$-CO), 2.07 (4H, m, H-16 & H-20), 4.50 (1H, dd, *J*=8.4, 6.6 Hz, H-3), 5.12 (3H, m, H-13, H-17 & H-21), 5.22 (1H, bs, H-7). ^{13}C-NMR (150 MHz, $CDCl_3$) δ 15.96 (q, C-24), 15.96 (q, C-27 or C-28), 16.16 (q, C-27 or C-28), 17.61 (q, C-29), 170.8 (s, C-31), 21.23 (q, $\underline{C}H_3$-CO), 22.11 (q, C-25), 23.28 (q, C-26), 23.57 (t, C-6), 23.94 (t, C-2), 25.61 (q, C-30), 26.78 (2C, t, C-16 & C-20), 27.77 (q, C-23), 29.94 (t, C-12), 31.77 (t, C-11), 33.75 (t, C-1), 36.46 (s, C-10), 37.51 (s, C-4), 39.71 (2C, t, C-15 & C-19), 41.84 (d, C-5),. 53.80 (d, C-9), 81.28 (d, C-3), 118.8 (d, C-7), 124.3 (d, C-21), 124.4 (d, C-13 or C-17), 124.5 (d, C-13 or C-17), 131.1 (s, C-22), 134.9 (s, C-14 or C-18), 135.2 (s, C-14 or C-18), 136.9 (s, C-8). 170.8 (s, CH_3-$\underline{C}$O) EIMS: see Fig. S3.1.1. HREIMS: calcd. 468.39673; found 468.39715. $[\alpha]_D^{25}$= -30.0 (c=0.05, $CHCl_3$).

(9β*H*)-Polypoda-8(26),13,17,21-tetraen-3β-ol, i.e., (9β*H*)-α-polypodatetraen-3β-ol acetate.
^{1}H-NMR (600 MHz, $CDCl_3$) δ0.848 (3H, s, Me-24), 0.873 (3H, s, Me-23), 0.930 (3H, s, Me-25), 1.11 (1H, bd, *J*=12.6 Hz, H-1), 1.33 (1H, m, H-11), 1.38 (2H, m, H-5 & H-6), 1.52 (1H, m, H-11), 1.570 (3H, s, Me-27), 1.59 (1H, m, H-9), 1.602 (6H, s, Me-28 & Me-29), 1.62 (1H, m, H-6), 1.678 (3H, s, Me-30), 1.68 (2H, m, H-2), 1.71 (1H, m, H-12), 1.78 (1H, m, H-1), 1.88(1H, m, H-12), 1.98 (4H, m, H-15 & H-19), 2.052 (3H, s, C$\underline{H}_3$-CO), 2.07 (4H, m, H-16 & H-20), 2.08(1H, m, H-7), 2.17 (1H, bd, *J*=12.0 Hz, H-7), 4.47(1H, dd, *J*=9.0, 6.6 Hz, H-3), 4.54(1H, bs, H-26), 4.71(1H, bs, H-26), 5.12 (3H, m, H-13, H-17&H-21). ^{13}C-NMR (150 MHz, $CDCl_3$) δ16.01 (q, C-27 or C-28), 16.09 (q, C-27 or C-28), 16.65(q, C-24), 17.68 (q, C-29), 21.30 (q, $\underline{C}H_3$-CO), 22.47 (q, C-25), 22.99 (t, C-6), 24.14 (t, C-2), 25.68 (q, C-30), 26.51 (t, C-11 or C-12), 26.60 (t, C-11 or C-12), 26.71 (t, C-16 or C-20), 26.77 (t, C-16 or C-20), 28.20(q, C-23), 30.94 (t, C-7), 33.94 (t, C-1), 37.59(s, C-10), 37.75(s, C-4), 39.74 (2C, t, C-15 & C-19), 45.12(d, C-5), 57.44(d, C-9), 81.09 (d, C-3), 109.7 (t, C-26), 124.3 (d, C-21), 124.4 (d, C-13 or C-17), 124.5 (d, C-13 or C-17), 131.2 (s, C-22), 134.9 (s, C-14 or C-18), 135.1 (s, C-14 or C-18), 148.5 (s, C-8), 170.9 (s, CH_3-$\underline{C}$O). EIMS: see Fig. S3.2.1. HREIMS: calcd. 468.39673; found: 468.39627. $[\alpha]_D^{25}$=-8.08 (c=0.07, $CHCl_3$); no

report of $[\alpha]_D$ for this compound. To confirm the stereochemistry of 9β-H, the NMR spectra of this compound were measured again in the C_6D_6 solution (see Fig. S3.2.10 and Fig. S3.2.11, NOESY spectrum, in Supporting Information).

Acetates of α-and γ-polypodatetraen-3β-ols. The detailed NMR data, $[\alpha]_D$ and HREIMS data were described in Supporting Information Fig. S3.3 and Fig. S3.4.

Polypoda-8(9)-13,17,21-tetraen-3β-ol acetate. ^{1}H-NMR (600 MHz, C_6D_6) δ1.031 (3H, s, Me-23), 1.049 (3H, s, Me-24), 1.064 (3H, s, Me-25), 1.21 (1H, bd, *J*=12.3 Hz, H-5), 1.47 (1H, m, H-1), 1.50 (1H, m, H-6), 1.62 (1H, m, H-6), 1.696 (3H, s, Me-29), 1.722 (3H, s, Me-26), 1.758 (3H, s, Me-27 or Me-28), 1.76 (1H, m, H-2), 1.805 (6H, s, Me-30 and either Me-27 or Me-28), 1.85 (1H, m, H-1), 1.881 (3H, s, C$\underline{H_3}$-CO), 1.97 (1H, m, H-2), 1.99 (1H, m, H-7), 2.02 (1H, m, H-7), 2.06 (1H, m, H-11), 2.26 (4H, m, H-15 & H-19), 2.28 (1H, m, H-11), 2.28 (2H, m, H-12), 2.33 (2H, m, H-16), 4.84 (1H, dd, *J*=11.8, 4.4 Hz, H-3), 5.38 (bt, *J*=6.9 Hz, H-21), 5.46 (2H, m, H-13 & H-17). ^{13}C-NMR (150 MHz, C_6D_6) δ16.14 (q, C-27 or C-28), 16.23 (q, C-27 or C-28), 16.92(q, C-24), 17.72 (q, C-29), 18.97(t, C-6), 19.59 (q, C-26), 20.24 (q, C-25), 20.83 (q, $\underline{C}H_3$-CO), 24.53 (t, C-2), 25.83 (q, C-30), 27.10 (t, C-16 or C-20), 27.24 (t, C-16 or C-20), 28.20 (q, C-23), 28.80 (t, C-11), 29.54 (t, C-12), 33.88 (t, C-7), 34.94 (t, C-1), 37.94(s, C-4), 38.88(s, C-10), 40.20 (2C, t, C-15 & C-19), 51.15(d, C-5), 80.57 (d, C-3), 124.7 (d, C-13 or C-17), 124.9 (d, C-21), 125.4 (d, C-13 or C-17), 126.2 (s, C-8), 131.1 (s, C-22), 135.1 (s, C-14), 139.9(s, C-9), 169.9 (s, CH_3-$\underline{C}$O). EIMS: see Fig. S3.5.1. HREIMS: calcd. 468.39673; found: 468.39657 $[\alpha]_D^{25}$= + 13.9 (c=0.04, $CDCl_3$).

Taraxerol acetate. ^{1}H-NMR (400 MHz, C_6D_6) δ0.79 (1H, m, H-5), 0.83 (1H, ddd, *J*=13.2, 13.2, 3.6 Hz, H-1), 0.933 (3H, s, Me-28), 0.944 (3H, s, Me-25), 1.019 (3H, s, Me-23), 1.051 (3H, s, Me-24), 1.115 (3H, s, Me-27), 1.155 (3H, s, Me-30), 1.16 (1H, m, H-21), 1.17 (1H, m, H-18), 1.185 (3H, s, Me-29), 1.199 (3H, s, Me-26), 1.20 (1H, m, H-22), 1.44 (3H, m, H-7 & H-9 &H-11), 1.44(1H, m, H-6), 1.46 (1H, m, H-19), 1.48 (1H, m, H-1), 1.51 (1H, m, H-21), 1.57 (2H, m, H-6 & H-22), 1.59 (1H, m, H-19), 1.64 (1H, m, H-11), 1.67 (2H, m, H-12), 1.71(1H, m, H-2), 1.84 (2H, m, H-2 & H-16), 1.884 (3H, s, C$\underline{H_3}$-CO), 2.07 (1H, dd, *J*=9.6, 3.2 Hz, H-7), 2.17 (1H, dd, *J*=14.4, 3.2 Hz, H-16), 4.78 (1H, dd, *J*=12.0, 4.8 Hz, H-3), 5.75 (1H, dd, *J*=8.4, 2.8 Hz, H-15). ^{13}C-NMR (100 MHz, C_6D_6) δ15.59 (q, C-25), 16.88 (q, C-24), 17.75 (t, C-11), 18.84 (t, C-6), 20.86 (q, $\underline{C}H_3$-CO), 21.48 (q, C-27), 23.83 (t, C-2), 26.13(q, C-26), 27.99 (q, C-23), 28.99 (s, C-20), 30.09 (2C, q, C-28 and either C-29 or C-30), 33.43 (t, C-19), 33.97 (2C, q, C-12 and either C-29 or C-30), 35.42 (t, C-22), 36.06 (s, C-17), 37.00 (t, C-21), 37.34 (t, C-1), 37.76 (s, C-10 or C-13), 37.82 (s, C-4), 37.92 (s, C-10 or C-13), 38.02 (t, C-16), 39.19 (s, C-8), 41.31 (t, C-7), 49.13 (2C, d, C-9 & C-18), 55.50 (d, C-5), 81.17 (d, C-3), 117.3 (d, C-15), 158.1 (s, C-14), 170.0 (s, CH_3-$\underline{C}$O). EIMS: see Fig. S3.8.1. HREIMS : calcd. 468.39673; found: 468.39691. $[\alpha]_D^{25}$= + 20.1 (c=0.65, $CDCl_3$).

Multiflorenol acetate. ^{1}H-NMR (600 MHz, $CDCl_3$) δ 0.761 (3H, s, Me-25), 0.855 (3H, s, Me-23), 0.87 (1H, m, H-12), 0.934 (3H, s, Me-24), 0.965 (3H, m, Me-30), 0.976 (3H, s, Me-29), 1.055 (3H, s, Me-28), 1.070 (3H, s, Me-26), 1.082 (3H, s, Me-27), 1.19 (1H, m, H-1), 1.24 (1H, m, H-21), 1.36 (1H, m, H-5), 1.37 (1H, m, H-16), 1.38 (2H, m, H-19), 1.43 (1H, m, H-11), 1.46 (1H, m, H-21), 1.48 (2H, m, H-22), 1.50 (1H, m, H-18), 1.54 (1H, m, H-11), 1.62 (2H, m, H-12 & H-16), 1.64 (3H, m, H-2 & H-15), 1.70 (1H, m, H-1), 1.77(1H, m, H-15), 1.98 (1H, m, H-6), 2.051(3H, s, C$\underline{H}_3$-CO), 2.12 (1H, m, H-6), 2.16 (1H, m, H-9), 4.51(1H, dd, *J*=11.4; 4.2 Hz, H-3), 5.46 (1H, bs, H-7). ^{13}C-NMR (150 MHz, $CDCl_3$) δ 13.20 (q, C-25), 15.96 (q, C-24), 17.09 (t, C-11), 21.32 (q, $\underline{C}H_3$-CO), 23.94 (t, C-6), 24.20 (t, C-2), 26.16 (q, C-27), 27.09 (q, C-26), 27.64 (q, C-23), 28.23 (s, C-20), 30.94 (q, C-28), 30.94 (s, C-17), 31.66 (t, C-15), 33.67 (q, C-29), 33.87 (t, C-21), 34.09 (q, C-30), 34.60 (t, C-19), 35.06 (s, C-10), 36.06 (t, C-12 or C-16), 36.08 (t, C-12 or C-16), 36.57 (t, C-22), 36.80 (t, C-1), 37.02 (s, C-13), 37.69 (s, C-4), 41.60 (s, C-14), 46.84 (d, C-18), 48.72(d, C-9), 50.24 (d, C-5), 81.17 (d, C-3), 111.5 (d, C-7), 147.6 (s, C-8), 171.0 (s, CH_3-$\underline{C}$O). EIMS: see Fig. S3.9.1. HREIMS: calcd. 468.39673; found: 468.39555. $[\alpha]_D^{25}$= -16.9 (c=0.094, $CDCl_3$).

Acknowledgements

This work was supported in part by Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Nos. 25450150 and 18380001)

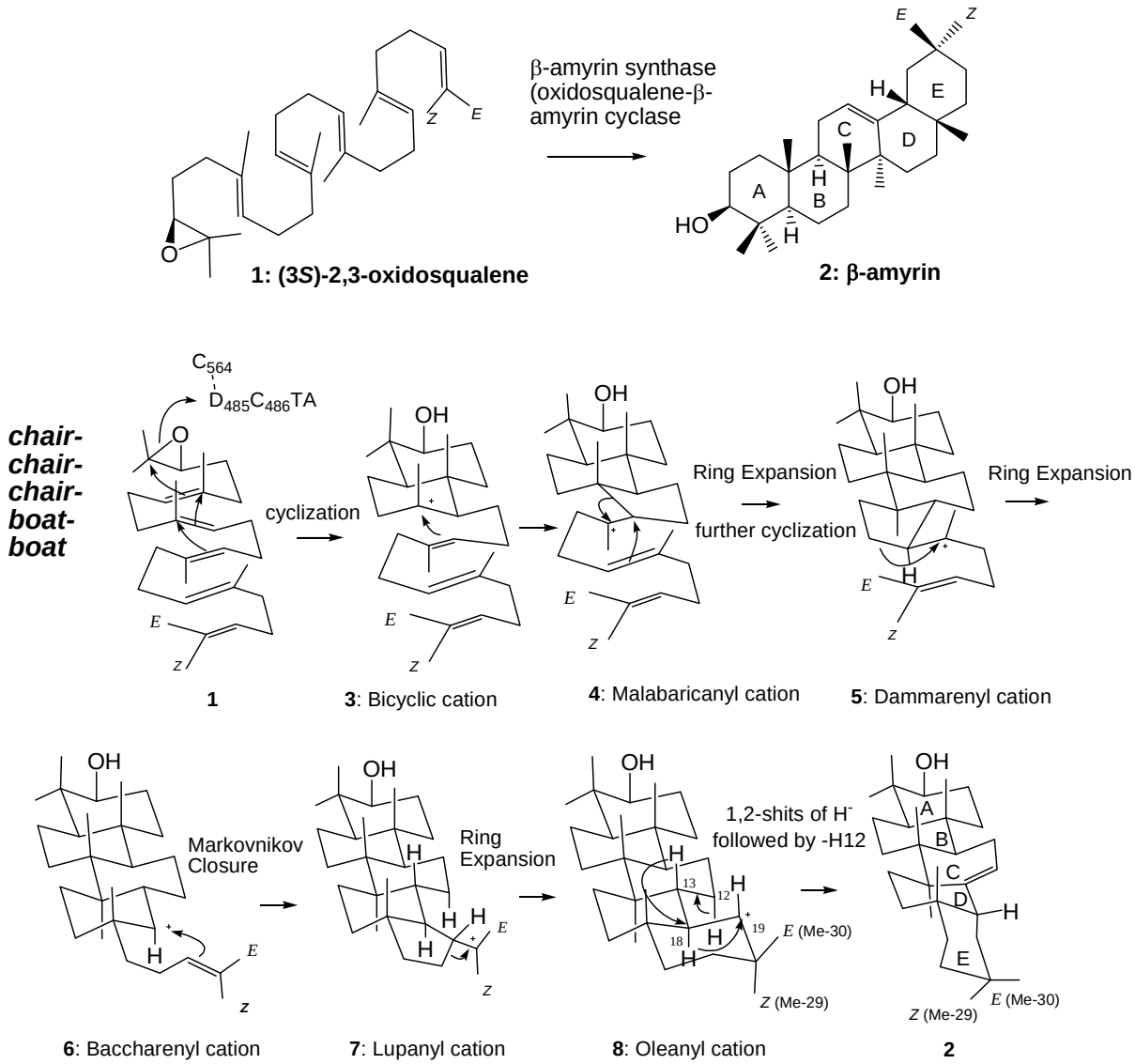
References

1. R. Xu, G. C. Fazio, S. P. T. Matsuda, *Phytochemistry*, 2004, **65**, 261-291.
2. P. M. Dewick, in Medicinal natural products: a biosynthetic approach, 3rd ed., John Willey & Sons, 2009.
3. Reviews for the polycyclization of triterpene cyclases: (a) T. Hoshino and T. Sato, *Chem. Commun.* 2002, 291-301; (b) K. U. Wend, G. E. Schulz, E. J. Corey and D. R. Liu, *Angew. Chem.Int. Ed.* 2000, **39**, 2812-2833; (c) G. Siedenburg and D. Jendrossek, *Appl. Environ. Microbiol.*, 2011, **77**, 3905-3915; (d) R. A. Yoder and J. N. Johnston, *Chem. Rev*. 2005, **105**, 4730-4756; (e) T-K.Wu, C-H. Chang, Y-T. Liu and T-T. Wang, *Chem. Rec*., 2008, **8**, 302-305; (f) I. Abe, *Nat. Prod. Rep*., 2007, **24**, 1311-1331. (g) W. D. Nes, *Chem. Rev*. 2011, **111**, 6423-6451.
4. Our studies. see (a) Y. Yonemura, T. Ohyama, and T. Hoshino, *Org. Biomol. Chem*. 2012, **10**, 440-446; (b) T. Hoshino, Y. Kumai and T. Sato, *Chem. Eur. J*. 2009, **15**, 2091-2100; (c) J. Cheng and T. Hoshino, *Org. Biomol. Chem*., 2009, ***7***, 1689-1699; (d) T. Hoshino, Y. Yonemura, T. Abe and Y. Sugino, *Org. Biomol. Chem.*, 2007, **5**, 792-801; (e) N. Morikubo, Y. Fukuda, K. Ohtake, N. Shinya, D. Kiga, K. Sakamoto, M. Asanuma, H. Hirota, S. Yokoyama and T. Hoshino, *J. Am. Chem. Soc*. 2006, **128**, 13184-13194; (f) T. Abe and T. Hoshino, *Org. Biomol. Chem*., 2005, **3**,

3127-3139; (g) T. Hoshino, K. Shimizu and T. Sato, *Angew. Chem. Int. Ed*. 2004, **43**, 6700-6703, (h) T. Hoshino, S. Nakano, T. Kondo, T. Sato and A. Miyoshi, *Org. Biomol. Chem.* 2004, **2**, 1456-1470; (i) T. Hoshino, Y. Kumai, I. Kudo, S. Nakano and S. Ohashi, *Org. Biomol. Chem.* 2004, **2**, 2650-2657; (j) S. Nakano, S. Ohashi and T. Hoshino, *Org. Biomol. Chem.*, 2004, **2**, 2012-2022; (k) T. Sato, M, Kouda and T. Hoshino, *Biosci Biotechnol Biochem.* 2004, **68**, 728-738; (l) T. Sato, S. Sasahara, T. Yamakami and T. Hoshino, *Biosci. Biotechnol. Biochem.*, 2002, **66**, 1660-1670; (m) T. Hoshino and S. Ohashi, *Org. Lett.*, 2002, **4**, 2553-2556; (n) T. Sato and T. Hoshino, *Biosci. Biotechnol. Biochem*. 2001, **65**, 2233-2242. (o) T. Hoshino, T. Abe and M. Kouda, *Chem. Commun*., 2000, 441-442; (p) T. Hoshino, M. Kouda, T. Abe and T. Sato, *Chem. Commun*., 2000, 1485-1486; (q) T. Hoshino and T. Sato, *Chem. Commun*., 1999, 2005-2006; (r) T. Sato and T. Hoshino, *Biosci. Biotechnol. Biochem*., 1999, **63**, 2189-2198; (s) T. Hoshino, M. Kouda, T. Abe and S. Ohashi, *Biosci. Biotechnol. Biochem*., 1999, **63**, 2038-2041; (t) T. Hoshino and T. Sato, *Chem. Commun*., 1999, 2005-2006; (u) T. Sato and T. Hoshino, *Biosci. Biotechnol. Biochem*., 1999, **63**, 1171-1180; (v) T. Sato, Y. Kanai, T. Hoshino, *Biosci. Biotechnol. Biochem*., 1998, **62**, 407-411.

5. Poralla and coworker's studies, see: (a) S. Schmitz, C. Füll, T. Glaser, K. Albert and K. Poralla, *Tetrahedron Lett.* 2001, **42**, 883-885; (b) C. Full and K. Poralla, *FEMS Microbiol Lett.* 2000, **183**, 221-224; (c) C. Pale-Grosdemange, T. Merkofer, M. Rohmer and K. Poralla, *Tetrahedron Lett.* 1999, **40**, 6009-6012; (d) T. Merkofer, C. Pale-Grosdemange, K. U. Wendt, M. Rohmer and K. Poralla, *Tetrahedron Lett.* 1999, **40**, 2121-2124; (e) D. Ochs, C. Kaletta, K. D. Entian, A. Beck-Sickinger and K. Poralla, *J. Bacteriol*. 1992, **174**, 298-302.
6. Wu and coworker's studies, see (a) C. H. Chang, H.Y. Wen, W. S. Shie, C. T. Liu, M. E. Li, Y. T. Liu, W-H Li and T. K. Wu, *Org. Biomol. Chem*., 2013, **11**, 4214-4219; (b) C. H. Chang, Y. C. Chen, S. W. Tseng, Y.T. Liu, H.Y. Wen, W.H. Li, C. Y. Huang, C. Y. Ko, T. T. Wang and T. K. Wu, *Biochimie* 2012, **94**, 2376-2381; c) T. K. Wu, C. H. Chang, H. Y. Wen, Y. T. Liu, W. H. Li, T. T. Wang and W. S. Shie, *Org. Lett***,** 2010, **12**, 500-503; (d) T. K. Wu, W. H. Li, C. H. Chang, H. Y. Wen, Y.T. Liu and Y. C. Chang, *Eur. J. Org. Chem*. 2009, 5731-5737; (e) T. K. Wu, T. T. Wang; C. H. Chang, Y. T. Liu and W. S. Shie, *Org. Lett*. 2008, **10**, 4959-4962; (f) T. K.Wu, H.Y. Wen, C. H. Chang, Y. T. Liu, *Org. Lett.* 2008, **10**, 2529-2532. (g) T. K. Wu, Y. T. Liu, F.H. Chiu and C. H. Chang, *Org. Lett*., 2006, **8**, 4691-4694.
7. The enzymatic reactions of the substrate analogs with hog liver cyclase. (a) T. Hoshino, A. Chiba, N. Abe, *Chem. Eur. J*. 2012, **18**, 13108-13116; (b) T. Hoshino and Y. Sakai, *Tetrahedron Lett.* 2001, **42**, 7319-7323; (c) T. Hoshino and Y. Sakai, *Chem. Commun*., 1998, 1591-1592; (d) T. Hoshino, E. Ishibashi and K. Kaneko, *J. Chem. Soc., Chem. Commun.*, 1995, 2401-2402; (e) E. J. Corey, S. C. Virgil, *J. Am. Chem. Soc.* 1991, **113**, 4025-4026; (f) E. J. Corey, S. C. Virgil, S. Sarshar, *J. Am. Chem. Soc.* 1991, **113**, 8171–8172; (g) E. J. Corey, A. Krief and H. Yamamoto, *J. Am. Chem. Soc.* 1971, **93**, 1493–1494; (h) E. E. van Tamelen, R. P. Hanzlik, K. B. Sharpless, R. B. Clayton, W. J. Richter and A. L. Burlingame, *J. Am. Chem. Soc*. 1968, **90**, 3284–3286.

8. (a) A. Eschenmoser and D. Arigoni, *Helv. Chim. Acta*, 2005, **88**, 3011-3050. (b) J-L. Giner, S. Rocchetti, S. Neunlist, M. Rohmer and D. Arigoni, *Chem. Commun.* 2005, 3089-3091
9. (a) A. Eschenmoser, L. Ruzicka,O. Jeger and D. Arigoni, *Helv. Chim. Acta*. 1955, **38**, 1890-1940; (b) T. Suga and T. Shishibori, *Phytochemistry*, 1975, **14**, 2411-2417.
10. (a) R. Ito, Y. Masukawa and T. Hoshino, *FEBS J*. 2013, **280**,1267-1280; (b) R. Ito, I Hashimoto, Y. Masukawa and T. Hoshino, *Chem. Eur. J*. 2013, **19**, 17150-17158; (c) T. Kushiro, M. Shibuya, K. Masuda and Y. Ebizuka, *J. Am. Chem. Soc.* 2000, **122**, 6816-6824.
11. L. H. D. Nguyen and L. J. Harrison, *Phtochemistry*, 1994, **50**, 471-476.
12. K. Shiojima, Y.Arai, K. Masuda, T. Kamada and H. Ageta, *Tetrahedron Lett.* 1983, **24**, 5733-5736.
13. (a) S. Lodeiro, Q. Xiong, W. K. Wilson, M. D. Kolesnikova, C. S. Onak and S. P. T. Matsuda, *J. Am. Chem. Soc*., 2007, **129**, 11213-11222; (b) Q. Xiong, W. K. Wilson and S.P.T. Matsuda, *Angew*. *Chem. Int. Ed.*, 2006, **45**, 1285-1288.
14. R. Ito, K. Mori, I. Hashimoto, C. Nakano, T. Sato and T. Hoshino, *Org. Lett.*, 2011, **13**, 2678-2681
15. Z. Wang, T. Yeats, H. Han and R. Jetter, *J. Biol. Chem*. 2010, **285**, 29703-29712.
16. T. Kushiro, M. Shibuya, K. Masuda and Y. Ebizuka, *Tetrahedron Lett.* 2000, **41**, 7705-7710.
17. Z-h Lin, H-x Long, Zhu Bo, Y-q Wang and Y-z Wu, *Peptides*, 2008, **29**, 1798-1805.
18. S. Mecozzi, A. P. West, JR. and D. A. Dougherty, *Proc. Natl. Acad. Sci., USA*, 1996, **93**, 10566-10571.
19. T. Husselstein-Muller, H. Schallerand P. Benveniste, *Plant Mol. Biol.*, 2001, **45**, 75-92.
20. (a) A. F. Barrero, R. E. A. Manzaneda, R. R. A. Manzaneda; S. Arseniyadis and E. Guittet, *Tetrahedron* 1990, **46**, 8161-8168; (b) J. F. Arteaga, V. Domingo, J. F. Quilez del Moral and A. F. Barrero, *Org. Lett*. 2008, **10**, 1723-1726.
21. X.Qi, S. Bakht, M. Leggett, C. Maxwell, R. Melton, and A. Osbourn, *Proc. Natl. Acad. Sci. USA.*, **2004**, *101*, 8233-8238.
22. Z. Xue, L. Duan, D. Liu, J. Guo, S. Ge, J. Dicks, P. O. Maille, A. Osbourn, and X. Qi, *New Phytologist*, 2012, **193**, 1022-1038.
23. S. Racolta, P. B. Juhl, D. Sirim, and J. Pleiss, *Proteins*, 2012, **80**, 2009-2019; see http://www.ttced.uni-stuttgart.de/ for TTCED.



Scheme 1. Cyclization pathway of (3*S*)-2,3-oxidosqualene **1** to generate β-amyrin triterpene **2**. The substrate **1** is folded in a *chair-chair-chair-boat-boat* conformation in the enzyme cavity and the proton released from the DCTA motif attacks the epoxide ring, leading to the sequential ring-forming reactions and the construction of the 6,6,6,6,6-fused pentacyclic ring scaffold **8** via carbocationic intermediates **3-7**. Oleanyl cation **8** undergoes the rearrangement reactions of 1,2-hydride shifts and deprotonation of the axial-oriented H-12 to yield β-amyrin **2**.

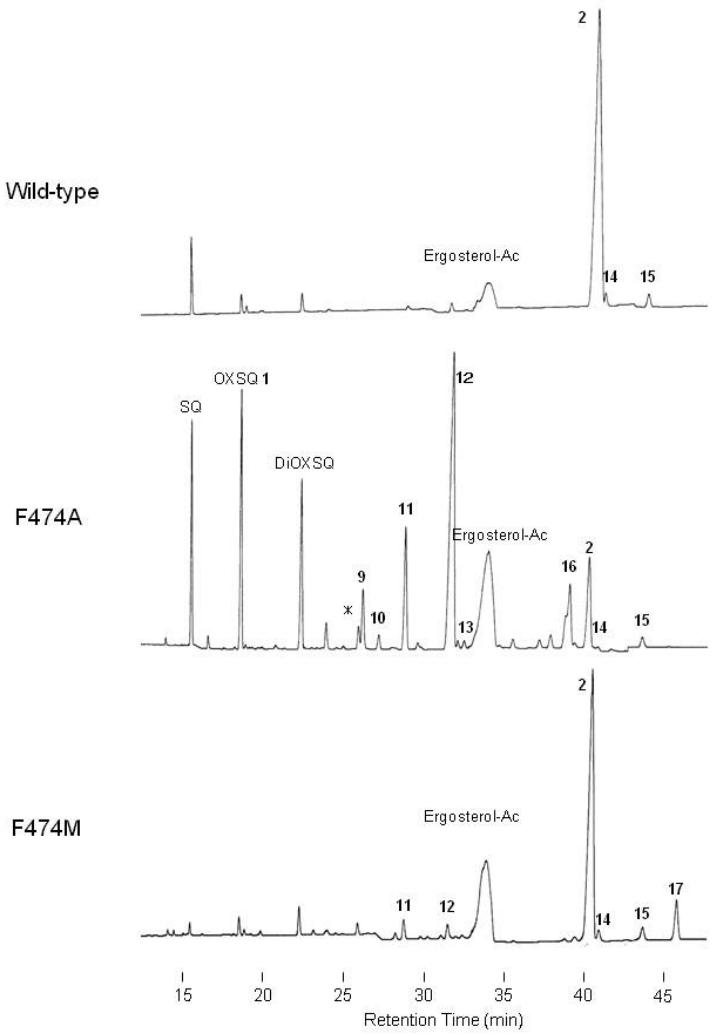
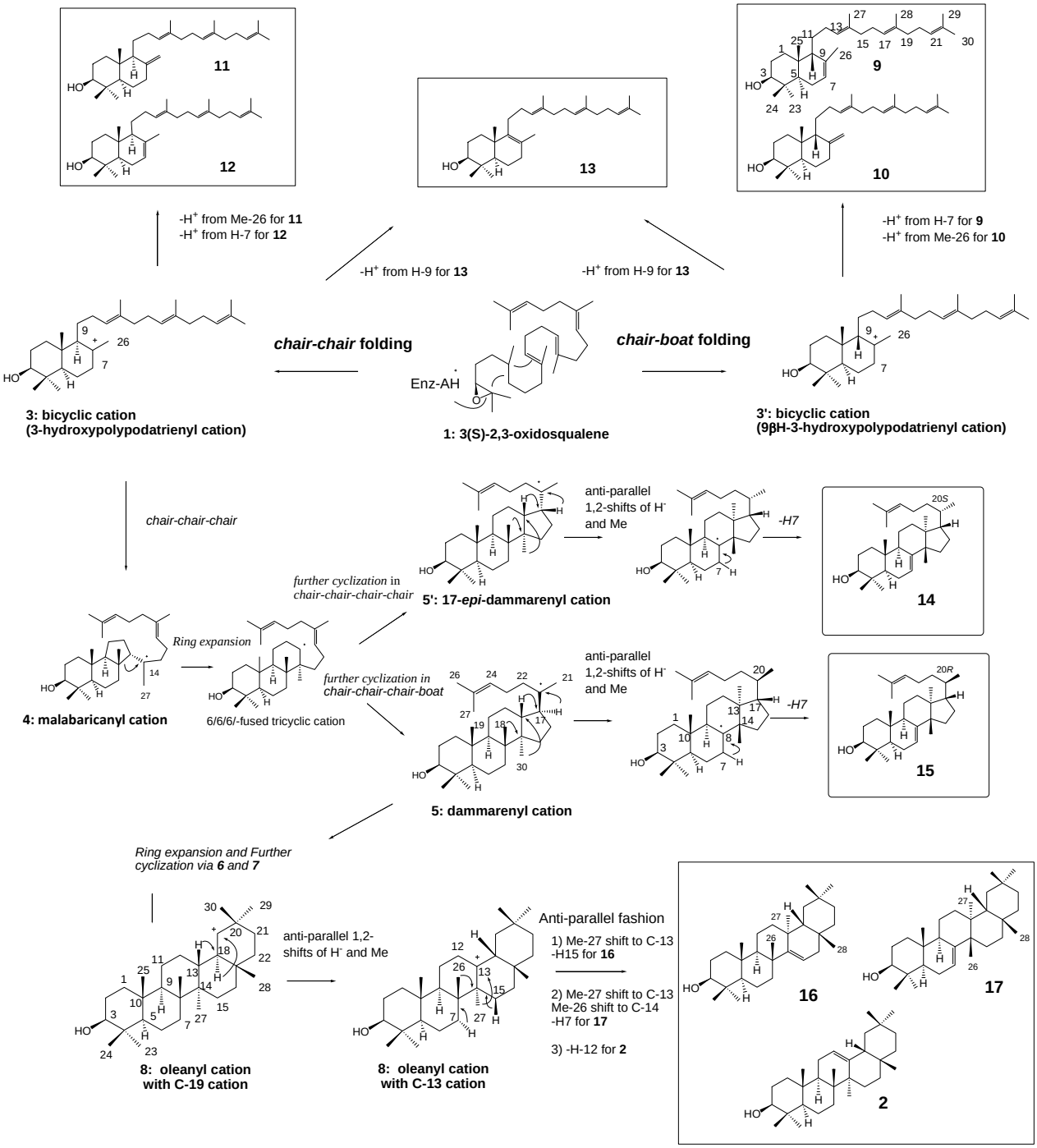


Figure 1. GC trace of the lipophilic materials obtained from F474A and F474M mutants.
Each of the mutants was cultured. The yeast cells collected by centrifugation were saponified by heating with 15%KOH/MeOH, followed by extraction of the lipophilic materials with hexane. The hexane-extract was evaporated to dryness. The acetylated residues thus obtained were subjected to GC analyses. GC column: J&W, DB-1, capillary (Length 30 m, I.D.0.32 mm, Film Thickness 0.25 µm), Injection temp. 300°C; Column temp.:190-250°C (10°C/min), 250-270°C (0.35°C/min).
*: impurities. Ergosterol-Ac: ergosterol acetate.
Products **9**~**17** were numbered by the order of elution in GC analyses.



Scheme 2. Cyclization pathway of substrate **1** to form bicyclic **9**~**13**, tetracyclic **14** and **15**, and pentacyclic products **2**, **16,** and **17**. Products **8**~**20** were numbered by the order of elution in GC analyses (Figure 1). Compound Names: **9**: (9β*H*)-polypoda-7,13,17,21-tetraen-3β-ol, i.e., (9β*H*)-γ-polypodatetraen-3β-ol; **10**: (9β*H*)-polypoda-8(26), 13,17,21-tetraen-3β-ol, i.e., (9β*H*)-α-polypodatetraen-3β-ol; **11**: polypoda-8(26), 13,17,21-tetraen-3β-ol, i.e., α-polypodatetraen-3β-ol; **12**: polypoda-7,13,17,21-tetraen- 3β-ol, i.e.,γ-polypodatetraen-3β-ol; **13**: polypoda-8(9),13,17,21-tetraen-3β-ol; **14**: tirucalla-7, 24-diene-3β-ol; **15**: butyrospermol; **16**: taraxerol (13α−methyl-27-norolean-14-en-3β-ol); **17**: multiflorenol (13α-methyl-14β-methyl-26-norolean-7-en-3-β-ol).

Table 1. Distribution ratio of the products (%) generated by each of the mutants. Each of the values represents the percentage of the triterpenes produced. Products **9**~**17** were numbered by the order of elution in GC analyses (Figure 1).

	Bicycle (C-B type) from cation **3'**			Bicycle (C-C type) from cation **3**			Bicycle from **3** or **3**'	Tetracycle			Pentacycle from cation **8**			
Product Number	**9**	**10**	Total % of cation **3'**	**11**	**12**	Total % of cation **3**	**13**	**14** from cation **5'**	**15** from cation **5**	Total % of **5'**and 5.	**2**	**16**	**17**	Total % of cation **8**
Wild	-	-	-	-	-	-	-	1.5	1.8	3.3	96.7	-	-	96.7
F474G	1.6	-	1.6	43.2	49.3	**92.5**	-	-	-	-	5.9	-	-	**5.9**
F474A	7.5	1.9	**9.4**	12.8	53.5	**66.3**	0.3*	0.5	2.5	3.0	12.5	8.7	-	**21.2**
F474V	-	-	-	-	14.2	14.2	-	4.2	4.4	8.6	71.3	5.9	-	77.2
F474L	-	-	-	-	12.0	12.0	-	2.0	4.1	6.1	81.9	-	-	81.9
F474M	-	-	-	1.1	0.9	2.0	-	3.8	4.6	8.4	80.5	1.1	8.0	89.6
F474T	1.8	-	1.8	8.8	27.8	**36.6**	-	2.0	4.3	6.3	48.0	7.3	-	**55.3**
F474H	-	-	-	-	3.0	3.0	-	2.2	3.7	5.9	89.6	1.5	-	91.1
F474Y	-	-	-	-	-	-	-	1.9	3.1	5.0	90.1	4.9	-	95.0
F474W	-	-	-	30.1	-	**30.1**	-	-	-	-	69.9	-	-	**69.9**

C-B and C-C types of bicyclic products stand for the chair-boat and chair-chair conformations, respectively.

Product distribution ratio for the F474A variant was calculated as follows: 1.49 (Fig. S5.8) + 10.5 (Fig. S5.6) + 0.48 (Fig. S5.4) + 3.36 (Fig. S5.2) = 15.83 (total amounts produced, mg/L-culture), thus,1.49/15.83x100=9.4% from cation **3'**, 10.5/05.83x100=66.3% from cation **3**, 0.48/15.83=3.0% from **5'** and **5**, 3.36/15.83=21.2% from cation **8**, the total % was calculated to be 99.9%. * Product **13** was excluded from the calculation of the product distribution ratio, because a negligible amount was produced. The bolded numbers highlight the function of each of the mutants.

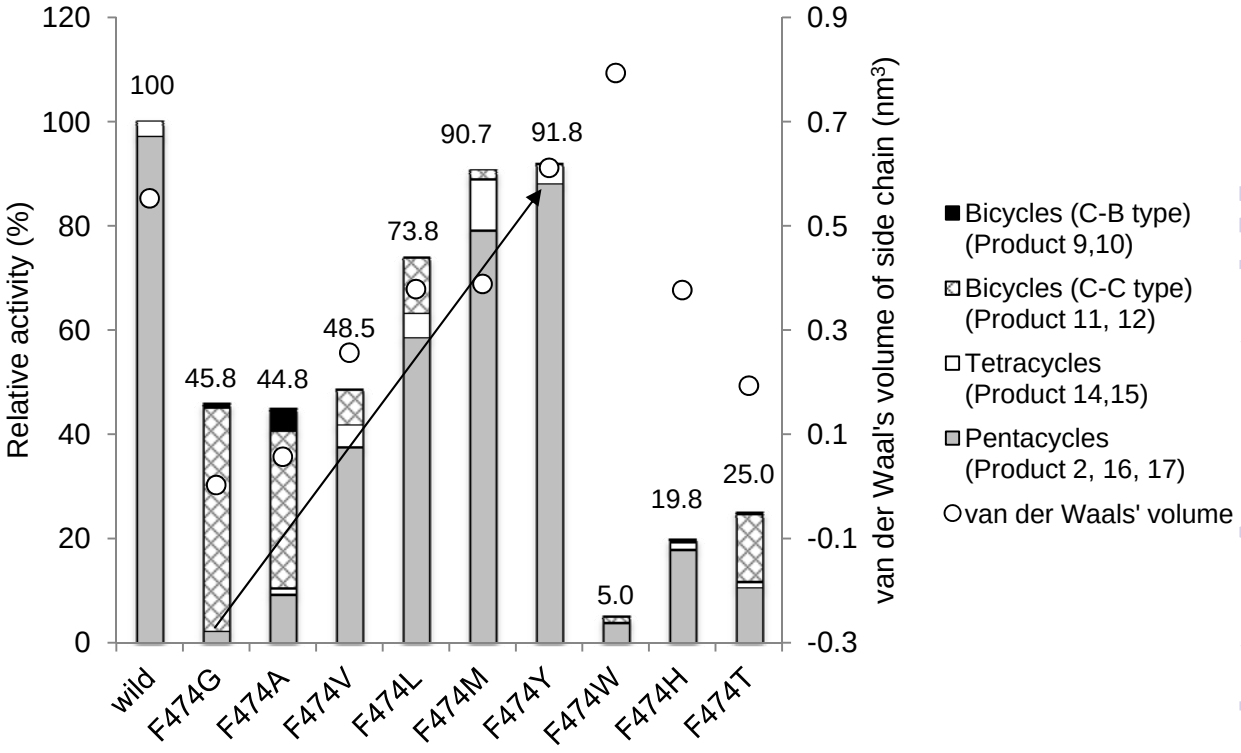
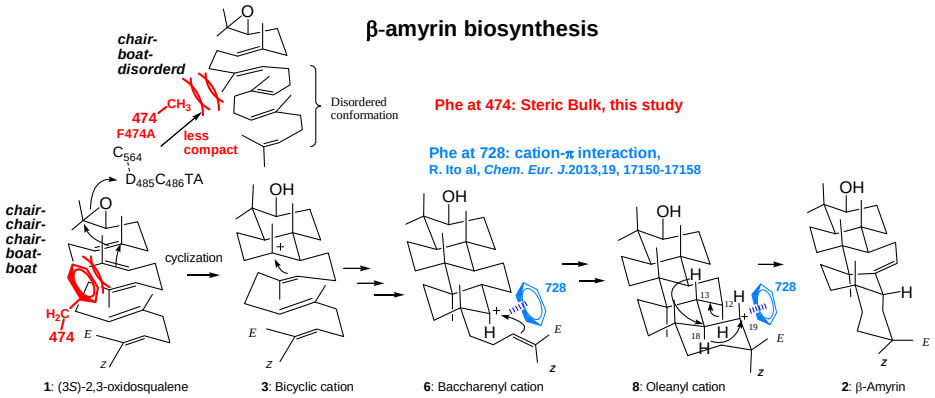


Figure 2. Enzymatic activities of the mutants relative to that of the wild type EtAS. The wild-type activity (100%) indicates the sum of the relative activities shown in Fig. S5.3 (production of **2**) and Fig. S5.5 (production of **14** and **15**). The activities for each of mutant represent the total sum of the relative activities (%) estimated for respective mutants, which are shown in Fig. S5.3, Fig. S5.5, Fig. S5.7, and Fig. S5.9 (Supporting Information). The total activities for the mutants thus obtained were evaluated against that of the wild-type EtAS (100%). Production of the pentacyclic oleanane skeleton derived from cation **8** roughly increased in proportion to the increase in the van der Waal's volume (see the arrow directed up to the right). The van der Waal's volumes (nm^3) for the side residues are as follows: Gly, 0.00279; Ala, 0.05702; Val, 0.25674; Leu, 0.37876; Met, 0.38872; Thr, 0.19341; His, 0.37694; Tyr, 0.6115, Trp, 0.79351; Phe (wild-type), 0.55298. The values were obtained from reference 17.

Graphical Abstract



This paper describes the importance of the steric bulk at 474residue for completion of the cyclization cascade, but not π-electrons of the Phe residue, which was in a sharp contract to the function of Phe728 residue (cation-π interaction, previously revealed by us).