A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Alicyclobacillus acidocaldarius squalene-hopene cyclase: the critical role of the steric bulk at Ala306 and the first enzymatic synthesis of epoxydammarane from 2,3-oxidosqualene

Authors: Natsumi Ideno, Shikou Umeyama, Takashi Watanabe, Mami Nakajima, Tsutomu Sato, and Tsutomu Hoshino

To be cited as: *ChemBioChem* 10.1002/cbic.201800281

Link to VoR: http://dx.doi.org/10.1002/cbic.201800281

A Journal of ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

10.1002/cbic.201800281

Alicyclobacillus acidocaldarius squalene-hopene cyclase: the critical role of steric bulk at Ala306 and the first enzymatic synthesis of epoxydammarane from 2,3-oxidosqualene

Natsumi Ideno, Shikou Umeyama, Takashi Watanabe, Mami Nakajima, Tsutomu Sato and Tsutomu Hoshino*

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

Tsutomu Hoshino: Prof. Dr.
*Corresponding author
E-mail: hoshitsu@agr.niigata-u.ac.jp

10.1002/cbic.201800281

Abstract

The acyclic molecule squalene (**1**) is cyclized into a 6,6,6,6,5-fused pentacyclic hopene (**2**) and hopanol (**3**) (ca. 5:1) by *Alicyclobacillus acidocaldarius* squalene-hopene cyclase (*Aa*SHC). The polycyclization reaction proceeds with regio- and stereochemical specificity under precise enzymatic control. This pentacyclic hopane skeleton is generated by folding **1** into an all-chair conformation. The Ala306 in *Aa*SHC is conserved in known SHCs; however, increases in steric bulk (A306T and A306V) led to the accumulation of 6,6,6,5-fused tetracyclic scaffolds possessing 20*R* stereochemistry in high yield (94% for A306V). The production of the 20*R* configuration indicated that **1** had been folded in a chair-chair-chair-boat conformation; in contrast, the normal chair-chair-chair-chair conformation affords the tetracycle with 20*S* stereochemistry, but the yield produced by the A306V mutant was very low (6%). Consequently, bulk at position 306 significantly affects the stereochemical fate during the polycyclization reaction. SHC also accepts (3*R*) and (3*S*)-2,3-oxidosqualenes (OXSQs) to generate 3α, β-hydroxyhopenes and 3α, β-hydroxyhopanols through a polycyclization reaction initiated at the epoxide ring. However, the Val and Thr mutants generated epoxydammarane scaffolds from (3*R*)-OXSQ, which indicated that the polycyclization cascade started at the terminal double bond position. This work is the first to report the polycyclization of oxidosqualene starting from the terminal double bond.

Introduction

The cyclization mechanisms by which acyclic squalene (**1**) and (3*S*)-2,3-oxidosqualene form polycyclic triterpenes have attracted much attention from chemists and biochemists for over a half a century.[1,2] *Alicyclobacillus acidocaldarius* squalene-hopene cyclase (*Aa*SHC) catalyzes the conversion of **1** into hopene (**2**) and hopanol (**3**) (ca 5:1), pentacyclic triterpenes consisting of 6,6,6,6,5-fused A/B/C/D/E-ring systems, as shown in Scheme 1A.[2a,b] The X-ray crystallographic structure of the SHC from *A. acidocaldarius* (*Aa*SHC) was first reported in 1997[3] and further refined later.[4] Compound **1** is folded into an all-prechair conformation inside the reaction cavity and is cyclized in a regio- and stereospecific fashion via a series of carbocationic intermediates, leading to the formation of 5 new C-C bonds and 9 stereocenters. This reaction consists of 8 reaction steps[2b] (Scheme 1B): (1) the acyclic molecule **1** is folded into a chair conformation to give the intermediary monocyclic cation **4**; (2) a second ring formation reaction occurs to afford cation **5**, consisting of a 6,6-fused bicyclic cation; (3) a third ring formation reaction produces the 6,6,5-fused tricyclic cation **6**; (4) cation **6** undergoes a ring expansion to generate the secondary cation **7** with a 6,6,6-

fused A/B/C-ring system; (5) a further ring closure gives the tertiary cation **8** with a 6,6,6,5-fused tetracycle (17-*epi*-dammarenyl cation); (6) cation **8** undergoes a second ring expansion reaction to provide the secondary cation **9** (prohopanyl cation); (7) the final ring formation reaction furnishes the 6,6,6,6,5-fused pentacyclic cation **10** (hopanyl cation); and (8) proton elimination from Me-29 introduces a double bond to produce **2**, and a water molecule attacks the C-22 cation of **10** to give **3**. Contrary to this step-wise mechanism, there are reports that the polycyclization reactions proceed in a concerted manner.[5a, b] (3*S*)-2,3-Oxidosqualene (**11**) is also accepted as a substrate by SHC, yielding 3β-hydroxyhopene (**12**) and 3β-hydroxyhopanol (**13**) through an all-chair conformation in a similar cyclization pathway to that of **1** (Scheme 1C). (3*R*)-2,3-Oxidosqualene (**14**) is also converted into 3α-hydroxyhopene (**15**) and 3α-hydroxyhopanol (**16**) through a boat-chair-chair-chair-chair conformation.[6a-d]

<Scheme 1>

To identify the active site residues of *Aa*SHC and their functions, site-specific mutagenesis experiments were conducted. The central Asp residue (Asp376) in the $D^{374}XD^{376}D^{377}$ motif functions as the proton donor to initiate the polycyclization reaction.[7a,b] His451 hydrogen bonds with Asp376 to enhance the motif acidity.[2a,6a] Asp377 likely stabilizes cation **4**.[7a] Trp312 likely binds with the terminal methyl group via a hydrophobic interaction (possibly CH-π interaction) to place **1** in close proximity to the DXDD motif, thus facilitating the initiation of the polycyclization reaction.[8] Phe365 stabilizes cation **5**.[9] Phe601 likely stabilizes cations **6** and **7**,[10] and Phe605 stabilizes the intermediary cations **9** and **10**.[11] All the Phe residues mentioned above function to stabilize transient cations via cation/π interactions.[12] The main function of Y420 is to stabilize the bicyclic cation **5** by pointing towards the active site cavity and providing additional negative electrostatic potential.[2a,4,13] The steric bulk at the Leu607 position is critical for the correct chair-folding around the B-ring formation site.[13] The steric bulk of another aliphatic residue, Ile261, is also crucial for enforcing the correct folding around the D/E-ring formation sites.[14] The I261A variant generates cations **6** and **7** with 13α-H and 17α-H configurations, respectively, which are opposite to the configurations of the genuine intermediary cations **6** (13β-H) and **7** (17β-H); thus, the steric bulk at this position directs the stereochemical fate. There have been many recent reports on the use of SHC mutants as biocatalysts, by which many known and unknown compounds have been successfully created.[15]

Herein, we report the function of the Ala306 residue in *Aa*SHC, which is widely conserved in known SHCs (Fig. S1, Supporting Information). The X-ray crystal

10.1002/cbic.201800281

structure of *Aa*SHC complexed with a modeled hopene molecule indicates that Ala306 is located near the squalene molecule (Fig. S2).[3,4] To investigate the function of Ala306 of *Aa*SHC in hopene biosynthesis, we constructed site-directed mutants in which the steric bulk at this position is altered, incubated **1**, **11**, and **14** with these variants, and determined the structures of the resulting enzymatic products. We provide clear evidence that the Ala306 residue is located in close proximity to the D-ring formation site and that the steric bulk at this position is a key determinant of the stereochemical fate during the polyolefin cyclization reaction, similar to the Ile216 residue. Furthermore, we report the first case in which the polycyclization reaction of oxidosqualene was initiated from the terminal double bond instead of the terminal epoxide ring, resulting in the production of a 17-*epi*-epoxydammarane scaffold.

Results

Isolation of products generated from the incubation of 1 with A306X mutants

Fig. 1A shows the GC profile obtained by incubating **1** with the cell-free homogenates of the A306V variant. Squalene **1** (200 mg) was incubated (pH 6.0, 45°C, 16 h) with the cell-free extracts (300 mL) obtained from a 6 L culture of the A306V variant. The incubation mixture was freeze-dried, and the lipophilic materials, including the enzymatic products, were extracted with hexane. These were then subjected to SiO_2 column chromatography by eluting with 100% hexane to remove the Triton X-100 present in the incubation mixture and to partially separate the products into several fractions. Only a small amount of **1** was recovered, indicating that the reaction by the variant was almost complete. The purity of the partially separated fractions was monitored by GC-MS. The fraction containing a mixture of **17**, **19**, and **20** was obtained by partial fractionation with 100% hexane as the eluent. Complete separation of the products was attained by column chromatography with activated SiO_2 (heated in an oven at 180°C for 2 h), yielding 2.2 mg of **17**, 1.0 mg of **19**, and 1.5 mg of **20** in a pure state. However, separation of the other products was more difficult. The fraction containing products **18**, **21**, **22**, **23**, and **24** was chromatographed on a 5% $AgNO_3$-impregnated SiO_2 column, and pure **23** was obtained in a yield of 1.4 mg by eluting with hexane/EtOAc (100:0.2), but the separation of **18** and **21** failed. This separation was achieved by C_{18} reverse-phase HPLC (THF/H_2O=60:40), which yielded 0.8 mg of **18** and 1.3 mg of **21** in a pure state. The fraction obtained by eluting with hexane/EtOAc (100:1) contained a mixture of **21**, **22**, **23**, and **24**. Pure **24** (2.1 mg) was obtained by reverse-phase HPLC (THF/H_2O=70:30), but the separation of **21**, **22**, and **23** was not successful. The fractions containing **18**, **21**, **22**, and **23**, which were not separable by 5%

argentation chromatography, were combined and subjected to 10%-$AgNO_3$ SiO_2 column chromatography, and eluted with hexane/EtOAc (100:0.1) to yield a mixture of **18**, **21**, and **22**, and then C_{18}-HPLC (THF/H_2O=60:40) successfully gave pure **22** in a yield of 0.6 mg. The isolation yields of **18**, **21**, **22**, and **23** were very low, despite the large peaks detected by GC (Fig. 1A). The low yields could be ascribed to both the separation difficulty and the use of an $AgNO_3$-SiO_2 column (as the recovery from $AgNO_3$ chromatography is generally low). In addition, there were many remaining fractions that have yet to be purified. Fig. S3.1 depicts the GC trace from the incubation of **1** with the A306T variant. Products **25**–**27**, which were not detected as products of the Val mutant, were newly revealed in the GC traces, and **28** was identified as a new product of the A306F mutant (Fig. S3.2). Products **25**–**28** were all identified by comparing their EIMS fragment patterns and retention times to those of previously isolated authenticated compounds.[2b]

<Figure 1>

Isolation of products generated from the incubation of 11 and 14 with A306X mutants

Fig. 1B shows the GC chromatogram of the reaction mixture obtained by incubating a mixture of **11** and **14** with the A306V mutant. To isolate the enzymatic products **29**–**33**, a racemic mixture of **11** and **14** (100 mg) was incubated (pH 6.0, 45°C, 20 h) with the cell-free extract (300 mL) from a 6 L culture of the A306V mutant. After lyophilization of the reaction mixture, the reaction products were extracted with hexane, which was passed through a short SiO_2 column by elution with hexane/EtOAc (100:20) to remove Triton X-100. The residues obtained after removing Triton X-100 were subjected to SiO_2 column chromatography in a stepwise gradient fashion (hexane/EtOAc=100:1 to 100:2). Product **33** was isolated in a pure state (6.0 mg), but the other products could not be separated. The inseparable fractions were divided into the following 2 fractions: a mixture (22.4 mg) of **29**, **31**, and **33**, and a mixture (14.7 mg) of **30** and **32**. Products **29** and **31** were successfully separated by normal-phase HPLC (hexane/2-propanol=100:0.1) to yield **29** (2.3 mg) and **31** (1.0 mg) in a pure state. Products **30** and **32** were separated by normal-phase HPLC (hexane/THF=100:1.0) with yields of 2.0 mg and 2.3 mg, respectively. The A306F variant generated products **34**, **35**, and **36** from a mixture of **11** and **14**, as shown in Fig. S4. Fifty mg of a mixture of **11** and **14** was incubated with the A306F mutant. SiO_2 column chromatography (hexane/EtOAC=100:1 to 100:5) afforded pure **36** (14.5 mg) and a mixture (6.0 mg) of **34** and **35**, which were then acetylated with Ac_2O/Py (1:1).

Normal-phase HPLC (hexane:2-PrOH=100:0.02) successfully isolated the pure acetates of **34** and **35** in yields of 0.6 mg and 0.8 mg, respectively. The isolation yields are not consistent with the production yields because some fractions still contained inseparable products.

Structural determination of products 17–36.

Fig. 2 lists the structures of enzymatic products that were identified. Products **17**–**28**, with the exception of **23**, were previously isolated from the site-specific variants of *Aa*SHC.[2b]. Detailed NMR analyses of **23,** a new product generated from *Aa*SHC-mutants, are described below. The EIMS and NMR spectra of **17**–**22** and **24** are shown in Figs. S5–S10, and brief descriptions for proposing the structures are provided in the Supporting Information. We isolated **17** from the L607F,[13] Y420A,[13], and Y609F variants.[8] Compound **18** is identical to the product isolated from variants I261A[14] and F605A.[11] Product **19** was isolated from mutants Y609A,[8] Y612A[8], and F365A.[9] Product **20** is identical to the compound generated by the F601A mutant.[10] Compound **21** was isolated from the W489 variant.[2b, 16, 17] Product **22** was previously isolated from the F605A mutant.[11] Product **24** was isolated from the W169F, W169H, and W489F variants.[17] The EIMS spectra of minor products **25**–**28** are shown in Fig. S13. The spectra and GC retention times of **25**–**28** were the same as those of samples we isolated previously: **25** and **27** from F605A,[11] **26** from P263G, [19], and **28** from F365A.[9] Accordingly, these structures are illustrated in Fig. 2.

The detailed NMR analyses of product **23,** including 2D NMR spectra (600 MHz, C_6D_6, Fig. S11.9), revealed that **23** had the same tetracyclic skeleton as **22**. Furthermore, **22** and **23** had identical EIMS spectra (Figs. S10.1 and S11.1), but their GCMS retention times were different (Fig. 1A), indicating that the 2 compounds are stereoisomers. The relative configurations of **23**, which is present in the A/B/C/D-fused tetracyclic scaffolds, were determined to be identical to those of **22** by detailed analyses of the NOESY spectrum (Fig. S11.6). Thus, the C-20 stereochemistry of **23** is opposite to that of **22** and can be assigned as the 20*S* configuration for **23**. Consequently, **23** is 20*S*-tirucalla-7(8),24-diene, as shown in Fig. 2. The δ_H of Me-21 of 20*S*-tirucalla-7(8),24-diene in $CDCl_3$ was reported to be δ_H 0.895 (d, J=6.8 Hz), [18] which was close to that of **23** (δ_H 0.885, J=6.8 Hz) in $CDCl_3$ solution. Thus, the 20*S* stereochemistry of **23** is credibly assigned.

Next, the structural analyses of the enzymatic products **29**–**36**, which were generated by incubating a mixture of **11** and **14**, are described. We have not reported the enzymatic reactions of **11** and **14** by the *Aa*SHC variants; all the products are new SHC-

metabolites, and the NMR analyses are described below.

The ^{1}H-NMR spectrum of **29** (400 MHz, C_6D_6, Fig. S14.2) showed the presence of an isopropylidene residue: δ_H 1.83 (s, 3H, Me-26) and 1.75 (s, 3H, Me-27). The ^{13}C NMR spectrum (100 MHz, C_6D_6, Fig. S14.4) revealed the presence of a tetrasubstituted double bond (δ_C 139.7, s; 134.9, s). Detailed HMBC analyses (Fig. S14.10) indicated that **29** consists of a 6,6,6,5-fused tetracyclic system and that the double bonds of **29** exist at C-13–C-17 and C-24–C25. The multiplicity of H-3 (δ_H 3.29) showed a broad singlet, indicating that the hydroxyl group is α-oriented. The doublet Me was located at C-20 because Me-21 (δ_H 1.14, d, *J*=7.2 Hz, 3H) showed a strong HMBC cross peak for C-17. The chemical shift of Me-21 (δ_H 0.918 d, *J*=6.8 Hz) in the $CDCl_3$ solution (see Fig. S14.3) was very close to the reported value (δ_H 0.910 d, *J*=6.9 Hz) for 20*R*-dammara-13(17),24-diene.[18] Thus, **29** was determined to be 3α-hydroxy-20*R*-dammara-13(17)-24-diene, as shown in Fig. 2.

The EIMS spectrum of **30** (Fig. S15.1) is almost identical to that of **29** (Fig. S14.1), indicating that **30** is a stereoisomer of **29**. Detailed NMR analyses (Fig. S15.9) showed that dammara-13(17),24-diene can be assigned to **30** as described in the structural determination of **29**. The C-20 stereochemistry of **30** was the same as that of **29** because the Me-21 of **30** showed a δ_H 0.919 (d, *J*=6.8 Hz, 3H) identical to that of **29** in $CDCl_3$ solution (cf. in the C_6D_6 solution, δ_H 1.14 for Me-21 of **29**; δ_H 1.15 for Me-21 for **30**; both compounds had the same chemical shifts in C_6D_6). The OH was determined to be β-oriented because H-3 (δ_H 3.16) gave a dd splitting pattern (dd, *J*=10.2, 6.1 Hz). The entire structure of **30** can be deemed 3β-hydroxy-20*R*-dammara-13(17),24-diene, as shown in Fig. 2.

The ^{1}H-NMR (400 MHz, C_6D_6) and ^{13}C-NMR (100 MHz, C_6D_6) spectra of **31** (Fig. S16.3 and S16.4) revealed the presence of 2 double bonds: δ_H 5.41 (2H: m, 1H, H-12; m, 1H, H-24) and δ_C 125.5 (d, C-24), 130.9 (s, C-25), 115.9 (d, C-12) and 149.2 (s, C-13). Two Me groups (δ_H 1.88, s, 3H, Me-26; 1.73, s, 3H, Me-27) had clear HMBC cross peaks with C-25 and C-24, indicating that the isopropylidene moiety remained without participation of the cyclization reaction, and thus, **31** is a tetracyclic product. Me-30 (δ_H 1.13, s, 3H) showed a distinct HMBC correlation with C-13; Me-21 (δ_H 1.05, d, *J*=6.8 Hz, 3H) had a clear HMBC contour with C-17 (δ_C 48.94, d); and the multiplicity of C-17 (d) indicated the presence of a proton at C-17. Thus, another double bond is positioned at C-12–C-13. Indeed, the chemical shifts of H-11 [δ_H 1.92 (m) and 2.16 (m)] clearly indicated that H-11 is located at the allylic position. The OH group was α-oriented because the multiplicity of H-3 (δ_H 3.29) showed a broad singlet. Detailed NMR analyses (Fig. S16.10) indicated that **31** possesses 3α-hydroxy-17-*epi*-dammara-

12(13),24-diene with *R* stereochemistry at C-20 because the chemical shift of Me-21 in $CDCl_3$ solution was δ_H 0.789 (d, *J*=6.4 Hz, see Fig. S16.2.2), which was very close to the reported value (δ_H 0.784, d, *J*=6.8 Hz),[18] and to that of Me-21 (δ_H 0.793, d, *J*=6.8 Hz, measured in $CDCl_3$, see Fig. S9.9) in product **21**.

Detailed NMR analyses (Fig. S17.9) indicated that product **32** also has a 17-*epi*-dammara-12(13),24-diene scaffold, which was further confirmed by the fact that the EIMS spectra of **31** (Fig. S16.1) and **32** (Fig. S17.1) were identical. The OH function was oriented at the β-position, as the multiplicity of H-3 (δ_H 3.14) was a double doublet (*J*=11.4, 4.7 Hz). Furthermore, the chemical shift of Me-21 (δ_H 1.06, d, *J*=6.7 Hz, Fig. S17.2.2) of **32** in C_6D_6 solution was the same as that (δ_H 1.05, d, *J*=6.8 Hz, Fig. S16.3.2) of **31** (cf. δ_H 0.784 for **32** and δ_H 0.789 for **31**, in $CDCl_3$ solution, Fig. S17.2.1). Thus, **32** was proposed to be 3β-hydroxy-17-*epi*-20*R*-dammara-12(13),24-diene.

No olefinic protons and no double bonds were found in the ^{1}H-NMR (600 MHz, acetone d_6, Fig. S18.2) and ^{13}C-NMR spectra (150 MHz, acetone d_6, Fig. S18.3) of product **33**, indicating that the full cyclization reactions occurred. Three carbons connected to the oxygen atom were found: δ_C 86.06 (s, C-20), 88.23 (d, C-24) and 70.43 (s, C-25). Three Me groups had distinct HMBC correlations with these 3 carbons: Me-21 (δ_H 1.19, s, 3H)/C-20 and both Me-26 (δ_H 1.13, s, 3H) and Me-27 (δ_H 1.05, s, 3H)/C-24/C-25, where the chemical shift assignments of Me-26 and Me-27 are exchangeable. Me-21 exhibited further HMBC contours with C-22 (δ_C 38.70, t) and C-17 (δ_C 49.30, d), and H-24 (δ_H 3.71, dd, *J*=10.6, 5.2 Hz) was spin-coupled with H-23 (δ_H 1.75, m, 1H; 1.85, m, 1H). Thus, the moiety 2-(tetrahydrofuran-2-yl)propan-2-ol is connected to C-17. This THF moiety was further confirmed by EIMS, in which the fragment ion *m/z* 143 was observed as the base peak (Fig. 18.1). More detailed HMBC analyses revealed the presence of a 6,6,6,5-fused tetracyclic ring, in which H-17 is positioned in a β-orientation because H-17 had a clear NOE with H-13 (δ_H 2.09, m, 1H). Further analyses, including 2D NMR (Fig. S18.9), indicated that product **33** is 20,24-epoxydammarane-25-ol with 17β-H (17-*epi*-epoxydammarane skeleton). No NOE was found between Me-21 and H-24, indicating that either (20*R*, 24*R*) or (20*S*, 24*S*) can be assigned for the THF ring. We have reported that epoxydammarane scaffolds are created by the SHC-mediated reactions of non-natural (3*R*)- and (3*S*)-2,3-squalene diols and 2,3:22,23-dioxidosqualenes,[20] but this work is the first to report the production of the epoxydammarane from oxidosqualene. The ^{1}H- and ^{13}C-NMR data of **33** (Fig. S18.9) were indistinguishable from those of (20*R*, 24*R*)-17-*epi*-epoxydammarane (Fig. S18.10), which was obtained from the reaction of (3*R*)-2,3-squalene diol with native SHC.[20] (20*S*, 24*S*)-17-*epi*-Epoxydammarane and (20*R*, 24*R*)-17-*epi*-epoxydammarane

are diastereomers, thus the NMR spectra of the 2 diastereomers must be different, which allowed us to propose (20*R*, 24*R*)-stereochemistry for **33**.

Product **34**-acetate showed peaks for 5 vinylic Me groups: δ_H 1.69 (s, 3H, Me-29); δ_H 1.81 (s, 3H, Me-30); δ_H 1.72 (s, 6H, Me-26 and Me-28); and δ_H 1.78 (s, 3H, Me-27). One tetrasubstituted double bond (δ_C 126.2, s, C-8 and δ_C 140.4, s, C-9) was found. Furthermore, 3 aliphatic Me signals were observed: δ_H 1.05 (s, 3H, Me-25), δ_H 0.892 (s, 3H, Me-24), and δ_H 1.05 (s, 3H, Me-23). These findings are suggestive of a bicyclic ring system. Me-25 and Me-26 had distinct HMBC cross peaks with C-9 and C-8, respectively, thus the tetrasubstituted double bond is arranged at C-8 and C-9. The H-3 (δ_H 5.03) showed a broad singlet, indicating α-oriented OH. Consequently, the structure of **35** is polypoda-8(9),13,17,21-tetraen-3α-ol, as shown in Fig. 2.

Detailed NMR analyses of **35**-acetate (Fig. S20.9) and **36** (Fig. S21.9) revealed that the structures of **35** and **36** are polypoda-7(8),13,17,21-tetraen-3α-ol and polypoda-7(8),13,17,21-tetraen-3β-ol, respectively.

<Figure 2>

Discussion

Mechanistic insights into the production of 17-36 from substrates 1, 11, and 14.

Scheme 2 shows the formation mechanisms of 6,6-fused bicyclic products. Substrates **1** and **11** were folded into a chair-chair conformation to give 6,6-fused bicyclic cations **5** and **37**, respectively. Substrate **14** was folded into a boat-chair conformation to afford the same bicyclic skeleton **38**, but the OH-orientation (α-arrangement) from **14** was opposite to that (β-orientation) from **11**. This finding is consistent with the cyclization mechanisms of **1**, **11**, and **14** (Scheme 1). Proton elimination from Me-26 gave **17** (path *a*). The deprotonation of 7α-H produced **19**, **35**, and **36** (path *b*). The elimination of H-9 generated **34** (path *d*), which was obtained only from substrate **14**. Product **28** was produced by the following successive antiparallel 1,2-shifts: H-9→C-8 cation, Me-25→C-9, H-5→C-10, and the deprotonation of 6β-H to give the double bond at C-5 and C-6 (path *c*).

Scheme 3A shows the formation mechanism of tricyclic **20**. This compound was produced only from **1**. Substrate **1** was folded into a chair-chair-boat conformation (**39**) to furnish the 6,6,5-fused A/B/C-ring system with 13α-H (**40**). On the other hand, the chair-chair-chair conformation (**41**) gives the 6,6,5-fused tricycle with 13β-H (**42**), which is identical to cation **7**. Deprotonation from Me-27 of cation **40** afforded product **20** (path *a*).

Scheme 3B depicts the cyclization mechanisms leading to the pentacyclic products

10.1002/cbic.201800281

25 and **26**, the skeletons of which were produced only from **1** and not from **11** and **14**. Squalene **1** was folded into an all-chair conformation to afford the final hopanyl cation **10**. The following successive rearrangement reactions of **10** with the C-22 cation afforded **25** (path *b*): H-21 to C-22 cation, H-17 to C-21, Me-28 to C-17, followed by deprotonation of H-13. H-21 was eliminated to give **26** (path *c*).

Scheme 3C illustrates the formation mechanism of product **33**, which is produced by the A306V and the A306T mutants (Fig. 1B and Fig. S4). The (20*R*, 24*R*) stereochemistry of **33** indicates that this compound was produced from (3*R*)-oxidosqualene **14**, but not from **11**. Substrate **14** was folded into an all-chair conformation. Proton attack on the terminal double bond, but not on the epoxide ring, triggered successive ring formation reactions to yield the 17-*epi*-6,6,6,5-fused dammarenyl cation **43**. The epoxide ring of **43** attacked the C-20 cation (*Re*-face attack), and then the nucleophilic attack of a water molecule on the C-25 cation afforded **33**. Intriguingly, the cyclization reaction started from the terminal double bond position and not from the epoxide ring. This study is the first to report a polycyclization reaction that initiated from the terminal double bond, despite the epoxide ring being present in the substrate. As shown in Scheme 1B, the SHC-mediated cyclization reaction starts from the epoxide ring side but not from the terminal double bond side. This phenomenon has been clearly explained in terms of the higher nucleophilicity of epoxide than that of double bond π-electrons towards the DXDD motif, a proton donor that can initiate the polycyclization reaction. [7] However, why epoxydammarane compounds such as **33** are never produced from **11** and **14** by the native SHC, even in small amounts, is unclear. An increase in the steric bulk at position 306 (A306V and A306T variants) may have altered the geometry of the substrate inside the reaction cavity, and the terminal double bond may have been positioned nearer to the DXDD motif than the epoxide ring. In the case of the wild type SHC, the epoxide ring may be in closer contact with the DXDD motif, and the more highly nucleophilic nature of the epoxide ring may further facilitate the polycyclization reaction. We have reported that native SHC can accept both (3*R*)- and (3*S*)-2,3-squalene diols to afford 24*R*-and 24*S*-epoxydammarane, respectively. [20] Furthermore, the four different 2,3;22,23-dioxidosqualenes (3*R*,*S* and 22*R*,*S* isomers) are also converted into four 3-hydroxy-epoxydammarane scaffolds with four different configurations (3*R*, 24*R*; 3*R*, 24*S*; 3*S*, 24*R*; 3*S*, 24*S*); [20] that is, native SHC can accept both 22*R*- and 22*S*-isomers as substrates, which correspond to (3*R*)- and (3*S*)-isomers in the case of the polycyclization reactions of 2,3-oxidosqualene, respectively, resulting in the production of both 24*R*- and 24*S*-epoxydammaranes. This fact that native SHC can accept both

10.1002/cbic.201800281

(3*R*)-**14** and (3*S*)-**11** raises the question why (3*S*)-**11** was not cyclized by the A306V and A306T variants to produce 24*S*-epoxydammarane, while both (3*R*)-**14** and (3*S*)-**11** underwent cyclization reactions from the epoxide ring to generate 3α- and 3β-hydroxy-tetracycle **29**–**32** and bicycle **34**–**36**. It is also unclear why the mutated SHCs cannot accept the monocyclic and bicyclic products of the premature cyclization step to yield further cyclized products through reentry of the mon- and/or bicyclic product into the reaction cavity after exiting. The substrate uptake and the product release occur through the same channel,[3,4] which may be very compact; thus, the truncated cyclization products cannot reenter the reaction cavity. In contrast, tetraprenyl-β-curcumene cyclase from *Bacillus megaterium* (BmeTC) can accept 3-deoxyachilleol A (monocyclic triterpene) to yield (+)-ambrein as a result of the re-entry of this monocyclic product into the reaction cavity.[21] Further studies are necessary to provide detailed insight into the cyclization mechanism underlying how **33** is produced from **14**.

<Scheme 2>

Scheme 4 shows the formation mechanisms of tetracyclic products from **1**, **11**, and **14**. Two types of formation mechanisms are proposed: (A) the chair-chair-chair-boat folding conformation **44** from **1**, leading to the tetracycle **45** with 17α-H; (B) the chair-chair-chair-chair folding conformation **48** from **1**, resulting in the tetracycle **49** with 17β-H. Substrates **11** and **14** were also cyclized according to the folding conformation of Type A, yielding the 3β-hydroxytetracyclic cation **46** and 3α-hydroxytetracyclic cation **47**, respectively. By contrast, Type B is limited to the cyclization reactions of **1**. The Newman projection of cations **45**, **46**, and **47** is shown in Scheme 4A. A minimal motion (60° rotation) in the reaction cavity leads to 20*R* stereochemistry, but a large rotation (120°) gives the 20*S* configuration. A minimal motion is preferable to a large motion in the reaction cavity and is thus allowed to afford 20*R* stereochemistry. This theory was proposed by Abe and Rohmer.[18] By contrast, the minimal motion (60° rotation) of cation **49** with 17β-H affords a 20*S* configuration (allowed, Scheme 4B), but the large motion (120° rotation) results in the generation of 20*R* products (not allowed, Scheme 4B). Thus, cation **49** led to products with 20*S* stereochemistry. Cation **45** underwent further structural modifications. A hydride shift of H-17 to the C-20 cation, followed by the deprotonation of H-13, furnished **18**, **29**, and **30** (path *a*). 1,2-Shifts of H-17 to the C-20 cation and H-13 to C-17, followed by the removal of H-12α, gave **21**, **31**, and **32** (path *b*). The following consecutive 1,2-rearrangement reactions occurred: H-17→C-20 cation, H-13→C-17, Me-30→C-13, Me-18→C14, and the subsequent deprotonation of H-7α generated product **22** (path *c*). Proton elimination of

Me-21 of cation **49** provided product **24** (path *d*). Cation **49** underwent sequential 1,2-shift reactions (path *c*): H-17→C-20, H-13→C-17, Me-30→C-13, Me-18→C14. The subsequent deprotonation reaction of H-7α furnished **23** (path *c*). The nucleophilic attack of a water molecule on the C-20 cation gave **27** (path *e*).

<Scheme 3>

Distribution ratio of products generated by altering the steric bulk at position 306

Table 1 shows the product distribution ratio obtained by the reactions of squalene **1**. To address the steric bulk, we employed the van der Waals volumes (nm^3) reported by Lin et al.: Gly, 0.00279; Ala, 0.05702; Thr, 0.19341; Val, 0.25674; Leu, 0.37876; and Phe, 0.55298.[22] These values indicate that the steric bulk increases as follows: Gly<Ala<Thr<Val<Leu<Phe. The Gly mutant, with decreased steric bulk, afforded the normal cyclization products **2** and **3** in large quantities and the aberrant cyclization products **25** and **26** (pentacycle) in small amounts, but did not produce any bi-, tri-, or tetracyclic products. The Thr variant, with increased bulk, afforded decreased amounts of **2** and **3** (25%), while the tetracyclic products were generated in increased quantities (57%). The generation of the tetracyclic products suggests that the Ala306 residue is located near the D-ring formation site. A further increase in steric bulk (A306V variant) afforded small amounts of **2** (2%) and of bicyclic (4%) and tricyclic products (2%), but large quantities of tetracyclic products were produced (90%). Furthermore, almost all the tetracyclic products possessed a 20*R* configuration (94%) derived from cation **45**, while only 6% were 20*S* tetracycles derived from cation **49**. In the case of the A306T variant, 63% of the products consisted of **18**, **21**, and **22**, derived from cation **45** (Scheme 4A), and 37% consisted of **23**, **24**, and **27**, derived from cation **49** (Scheme 4B). Therefore, an increase in the steric bulk facilitates the formation of cation **45**. Cation **8**, an intermediary tetracyclic cation (Scheme 1A), is the same as cation **49** (Scheme 4B), indicating that 20*S* tetracyclic products should be generated from the intermediary cation **8** or **49** during the normal cyclization process. However, increased steric bulk at the 306 position drastically increased the production of 20*R* tetracycles, strongly indicating that the size of the residue at 306 is critical to the normal polycyclization pathway. The Leu mutant furnished bicyclic products with a significantly high yield (43%), and an almost equal amount of tetracyclic products (39%), as shown in Table 1 and Fig. S3.3. The Leu and Val residues have the same hydrophobic properties, but possess no polar OH and π-electrons; the difference between them is only the steric size (Leu>Val). Replacing Phe with bulk larger than Val and Leu generated only bicyclic products and no tetracyclic products. This finding

10.1002/cbic.201800281

indicates that the introduction of significantly larger substituents at the 306 position altered the orientation of substrate **1** inside the reaction cavity. Consequently, the steric bulk at position 306 significantly affects the stereochemical fate and the truncation steps during the polycyclization cascade.

Table 2 shows the product distribution ratios obtained by the incubation of **11** and **14**. The wild type produced **12** and **13** from **11,** and **15** and **16** from **14**, but no other product was produced, as shown in Scheme 1B. The A306T variant produced tetracyclic products in larger amounts (82%), in addition to **12** and **15** (8%). This result is consistent with the outcomes from the incubation of **1** (Table 1), but a small amount (8%) of epoxydammarane **33** was found by GC (Fig. S4). The A306V variant also produced tetracyclic products with a 20*R* configuration (47%), and no products with 20*S* stereochemistry were present above the limit of GC detection. Interestingly, a larger amount of **33** was produced (20%) than that by the A306T variant. The A306F mutant produced only bicyclic products **34**–**36**, which was consistent with the results of squalene incubation with the A306F variant. Notably, products **29**–**32** were all generated from cations **46** and **47** (equivalent to **45**), which resulted from a chair-chair-chair-boat folding conformation (Scheme 4A). In addition, **33** was produced only from **14** but not from **11**. The higher production of **33** by the Val mutant than by the Thr mutant indicates that the production of **33** is also dependent on the steric bulk at position 306. In fact, the larger Phe variant did not generate any **33** (Table 2). Thus, an appropriate steric size is required for the formation of **33**.

This study clearly demonstrates that the appropriate steric bulk at position 306 of *Aa*SHC is critical for directing the normal polycyclization cascade. We describe the enzymatic reactions of various substrate analogs with SHC, e.g., norsqualenes lacking Me-group [6a, 23] and homosqualenes bearing ethyl groups on the squalene backbone, [24] and we provide insight into how the polycyclization reaction is affected by the specific modifications. The steric bulk at the Me-positions of squalene substrate significantly altered the polycyclization pathway, e.g., alteration of the folding conformation leading to different stereochemistry and a truncation of the ring-forming cascade, which in some cases resulted in the generation of novel triterpene scaffolds.[23b] The results of the site-specific mutagenesis experiments provide further evidence that the appropriate steric size of the active site residues is critical to correct folding and the full cyclization pathway.

Conclusions

10.1002/cbic.201800281

The following conclusions can be drawn from our study: (1) the appropriate steric bulk of the Ala306 residue directs the normal polycyclization reaction without premature termination; (2) the Ala side chain is in close proximity to the D-ring formation site, as the A306V mutant produced large amounts of tetracyclic products with the 20*R* configuration; and (3) this is the first report describing the initiation of the polycyclization reaction from the terminal double bond rather than from the epoxide ring. As discussed above, the production of **33** raises the question why the native SHC does not produce epoxydammarane structure(s) from **11** or **14**, specifically why the proton attack occurs only at the epoxide ring, but not the terminal double bond, despite both the terminal double bond and the epoxide ring entering the reaction cavity through the substrate channel. The A306V and A306T variants are able to accept both heads of oxidosqualene **14** because products **29**–**32** are generated by a proton attack on the epoxide ring, while product **33** is produced by a proton attack on the terminal double bond.

Experimental

General analytical methods: NMR spectra were recorded in C_6D_6 on a Bruker DMX 600 or DPX 400 spectrometer, the chemical shifts (δ) given in ppm relative to the residual solvent peak δ_H 7.280 and δ_C128.0 as the internal reference for ^{1}H- and ^{13}C-NMR spectra, respectively. In the case of $CDCl_3$ solutions, the chemical shifts are given in ppm relative to the solvent peak (δ_H 7.26; δ_C 77.0). The coupling constants *J* are given in Hz. GC analyses were performed on a Shimadzu GC-8A chromatograph equipped with a flame ionization detector (DB-1 capillary column, 0.53 mm × 30 m). GC-MS spectra were performed on a JEOL SX 100 spectrometer under electronic impact at 70 eV with a DB-1 capillary column (0.32 mm x 30 m). HR-EIMS was performed by direct inlet system. Specific rotation values were measured at 25°C with a Horiba SEPA-300 polarimeter.

Site-directed mutagenesis

The *shc* genes were mutated via oligonucleotide-directed *in vitro* mutagenesis using the Pharmacia Unique Site Elimination Mutagenesis kit (Pfizer, New York, NY, USA). Plasmid pKS, a pUC119 derivative containing a 1 kb *Kpn*I-*Sac*I fragment, was used as a template for all mutagenesis reactions. The synthetic oligonucleotides for the mutations (with changes to the wild type sequence shown in bold type) are as follows:

 10.1002/cbic.201800281

A306G: 5'-GGATGTTTCAG**GGG**TCGATATCGCCGGTGTG-3'
A306T: 5'-GGATGTTTCAG**ACT**TCGATATCGCCGGTGTG-3'
A306V: 5'-GGATGTTTCAG**GTT**TCGATATCGCCGGTGTG-3'
A306L: 5'-GGATGTTTCAG**TTG**TCGATATCGCCGGTGTG-3'
A306F: 5'-GGATGTTTCAG**TTT**TCGATATCGCCGGTGTG-3'

Silent mutations were also used to screen the desired mutants by restriction fragment analysis (underlined sequences). To ascertain that the desired mutation had been achieved, the entire region of the inserted DNA was sequenced.

Incubation conditions and GC analysis

The SHC described here is from *A. acidocaldarius*.[7-14, 16-17, 19,20] Standard culture of the cloned *Escherichia coli* and incubation conditions were performed according to our published protocols.[7] The cell-free extract was prepared as follows. One liter of *E. coli* culture encoding the native and mutated SHCs (pET3a) was harvested by centrifugation 50 mL of citrate buffer solution (pH 6.0, 50 mM) was added to the pellets, and ultrasonication was then performed to disrupt the cells. The supernatant was used for incubations after removal of all cell debris by centrifugation. One mL of the supernatant contained *ca.* 200 μg of the pure SHCs. One mg of the substrates, **1** and a mixture of **11** and **14**, and 20 mg of Triton X-100 were emulsified with 1.0 mL dH_2O and 3.0 mL Na-citrate buffer solution (pH 6.0, 0.5 M). To this solution, 2.0 mL of the cell-free extract was added, followed by incubation at 45°C for 12 h. To terminate the enzyme reaction, 6 mL of 15% KOH/MeOH was added and heated at 80°C for 30 min. The lipophilic enzyme products and the substrate analogs, which remained unreacted, were extracted 4 times with hexane (5 mL) from the incubation mixtures, and the quantities of the products and the starting materials were estimated by GC analyses with a DB-1 capillary column (30 m in length, J&W Scientific, Folsom, CA, USA). GC-EIMS product profiles and substrate distributions, prepared after the saponification treatment (quenching reaction), were identical to those obtained prior to the saponification reaction, and hence there was no change in product distribution and structural modification, or the substrates produced by the quenching conditions. The epoxide ring was also unmodified by this quenching treatment. GC analyses shown in Fig. 1A were performed under the following conditions: injection temperature, 290°C; column temperature, 290°C; N_2 carrier pressure, 0.75-1.5 kg/cm^2. GC analyses shown in Fig. 1B were performed on the same GC instrument and GC column as (A), but the N_2 carrier pressure was 0.5 kg/cm^2.

Spectroscopic data of enzymatic products.

Almost the enzymatic products obtained from squalene incubation were reported in our previous papers, as described in Text.[2b] To validate the identification, we isolated products **17—24** and determined by the detailed NMR analyses: see the Figs. S5—S12 in Supporting Information. Production of **23** by the mutated SHCs has not been reported before. Production yields of **25—28** were significantly small, thus identifications were made by GC-EIMS, as described in Fig. S13. NMR and EIMS data of products **29—36**, which have not been reported before. Therefore, the spectroscopic data of **23** and **29—36** are given here.

Product 23: $[\alpha]_D^{25}$ = -0.565 (*c*=0.018, EtOH); ^{1}H NMR (600 MHz, C_6D_6): δ=0.964 (s, 3H, Me-19), 0.967 (s, 3H, Me-28), 1.03 (s, 3H, Me-29), 1.04 (m, 1H, H-1), 1.05 (s, 3H, Me-18), 1.12 (d, *J*=6.4 Hz, Me-21), 1.18 (s, 3H, Me-30), 1.27 (m, 2H, H-22 and H-3), 1.42 (m, 1H, H-16), 1.48 (dd, *J*=12.1, 5.4 Hz, 1H, H-5), 1.52 (m, 2H, H-2 and H-3), 1.58 (m, 1H, H-20), 1.60 (m, 1H, H-11), 1.62 (m, 1H, H-15), 1.63 (m, 1H, H-17), 1.66 (m, 1H, H-11), 1.68 (m, 1H, H-2), 1.72 (m, 1H, H-22), 1.74 (m, 1H, H-1), 1.75 (s, 3H, Me-27), 1.77 (m, 2H, H-12 and H-15), 1.83 (s, 3H, Me-26), 1.92 (m, 1H, H-12), 2.00 (m, 1H, H-6), 2.09 (m, 1H, H-16), 2.14 (m, 1H, H-23), 2.25 (m, 1H, H-6), 2.33 (m, 1H, H-23), 2.42 (m, 1H, H-9), 5.42 (bt, *J*=6.2 Hz, H-24), 5.47(bd, *J*=2.6 Hz, 1H, H-7) ; ^{13}C NMR (150 MHz, C_6D_6): δ=13.36 (q, C-19), 17.72 (q, C-27), 18.48 (t, C-11), 18.62 (q, C-21), 19.44 (t, C-2), 21.50 (q, C-29), 22.25 (q, C-18), 24.70 (t, C-6), 25.53 (t, C-23), 25.88 (q, C-26), 27.56 (q, C-30), 28.60 (t, C-16), 33.16 (q, C-28), 33.29 (s, C-4), 34.29 (t, C-12), 34.46 (t, C-15), 35.39 (s, C-10), 36.37 (d, C-20), 36.64 (t, C-22), 39.26 (t, C-1), 42.69 (t, C-3), 43.82 (s, C-13), 49.50 (2C, d, C-9 and s, C-14), 51.66 (d, C-5), 53.42 (d, C-17), 118.5 (d, C-7), 125.7 (d, C-24), 130.7 (s, C-25), 146.1 (s, C-8). Me-21 showed δ_H 0.885 (d, *J*=6.8 Hz) in the $CDCl_3$ solution. The assignments of C-13 and C-14 may be exchangeable. MS(EI): *m/z*: 69 (30%), 109 (28%), 395 (100%), 410 (27%); [M^+]; HRMS (EI): *m/z*: calcd. for $C_{30}H_{50}$: 410.39125; found: 410.39187.

Product 29: $[\alpha]_D^{25}$= -17.15 (*c*=0.206, EtOH); ^{1}H NMR (400 MHz, C_6D_6): δ=0.872 (s, 3H, Me-29), 0.949 (s, 3H, Me-19), 1.045 (s, 3H, Me-28), 1.105 (s, 3H, Me-18), 1.141 (d, *J*=7.2 Hz, 3H, Me-21), 1.253 (s, 3H, Me-30), 1.39 (m, 1H, H-11), 1.40 (m, 1H, H-15), 1.43 (m, 1H, H-6), 1.47 (m, 2H, H-1), 1.48 (m, 1H, H-7), 1.53 (2H, m, 1H, H-22; m, 1H, H-5), 1.54 (m, 1H, H-6), 1.55 (m, 1H, H-2), 1.62 (m, 1H, H-22), 1.67 (m, 1H, H-11), 1.70 (m, 1H, H-7), 1.71 (m, 1H, H-9), 1.752 (s, 3H, Me-27), 1.828 (s, 3H, Me-26), 1.93 (m, 1H, H-2), 2.02 (m, 1H, H-12), 2.07 (m, 1H, H-15), 2.10 (m, 1H, H-23), 2.25 (m, 1H, H-23), 2.28 (m, 1H, H-16), 2.46 (m, 1H, H-16), 2.59 (dd, *J*=13.6, 3.6 Hz, 1H, H-12), 2.72 (m, 1H, H-20), 3.29 (br s, 1H, H-3), 5.42 (t, *J*=6.4 Hz, H-24); ^{13}C NMR (100 MHz, C_6D_6): δ=16.50 (q, C-19), 17.53 (q, C-18), 17.79 (q, C-27), 18.53 (t, C-6),

10.1002/cbic.201800281

20.18 (q, C-21), 22.09 (t, C-11), 22.39 (q, C-29), 23.23 (q, C-30), 23.30 (t, C-12), 25.86 (q, C-26), 25.94 (t, C-2), 27.18 (t, C-23), 28.59 (q, C-28), 29.47 (t, C-16), 31.05 (t, C-15), 32.24 (d, C-20), 33.90 (t, C-1), 35.72 (t, C-7), 36.02 (t, C-22), 37.74 (s, C-10), 37.79 (s, C-4), 41.60 (s, C-8), 49.71 (d, C-5), 51.82 (d, C-9), 56.94 (s, C-14), 75.80 (d, C-3), 125.5 (d, C-24), 130.9 (s, C-25), 134.9 (s, C-17), 139.7 (s, C-13). The assignments of C-4 and C-10 may be exchangeable. MS (EI) see Fig. S14.1. HRMS (EI): *m*/*z*: calcd. for $C_{30}H_{50}O$: 426.38616; found: 410.38896.

Product 30: $[\alpha]_D^{25}$= -10.9 (*c*=0.278, EtOH); ^{1}H NMR (600 MHz, C_6D_6): δ=0.796 (dd, *J*=11.8, 1.4 Hz, 1H, H-5), 0.909 (s, 3H, Me-19), 0.912 (s, 3H, Me-29), 0.92 (m, 1H, H-1), 1.070 (s, 3H, Me-18), 1.149 (s, 3H, Me-28), 1.152 (d, *J*=6.7 Hz, 3H, Me-21), 1.275 (s, 3H, Me-30), 1.40 (2H, m, 1H, H-11; m, 1H, H-15), 1.44 (m, 1H, H-6), 1.46 (m, 1H, H-7), 1.47 (m, 1H, H-9), 1.53 (m, 1H, H-22), 1.56 (m, 2H, H-2), 1.62 (2H: m, 1H, H-22; m, 1H, H-7), 1.63 (m, 1H, H-11), 1.64 (m, 1H, H-6), 1.69 (m, 1H, H-1), 1.753 (s, 3H, Me-27), 1.831 (s, 3H, Me-26), 2.02 (m, 1H, H-12), 2.07 (m, 1H, H-15), 2.09 (m, 1H, H-23), 2.24 (m, 1H, H-23), 2.28 (m, 1H, H-16), 2.47 (m, 1H, H-16), 2.59 (ddd, *J*=12.5, 4.7, 1.8 Hz, 1H, H-12), 2.72 (m, 1H, H-20), 3.16 (dd, *J*=10.2, 6.1 Hz, 1H, H-3), 5.41 (t, *J*=7.0 Hz, H-24); ^{13}C NMR (150 MHz, C_6D_6): δ=15.83 (q, C-29), 16.61 (q, C-19), 17.52 (q, C-18), 17.79 (q, C-27), 18.59 (t, C-6), 20.18 (q, C-21), 22.17 (t, C-11), 23.11 (q, C-30), 23.31 (t, C-12), 25.86 (q, C-26), 27.17 (t, C-23), 27.89 (t, C-2), 28.35 (q, C-28), 29.47 (t, C-16), 31.06 (t, C-15), 32.26 (d, C-20), 35.84 (t, C-7), 36.01 (t, C-22), 37.62 (s, C-10), 39.15 (s, C-4), 39.26 (t, C-1), 41.34 (s, C-8), 52.02 (d, C-9), 56.29 (d, C-5), 56.83 (s, C-14), 78.46 (d, C-3), 125.4 (d, C-24), 130.9 (s, C-25), 135.1 (s, C-17), 139.5 (s, C-13). MS (EI) see Fig. S15.1. HRMS (EI): *m*/*z*: calcd. for $C_{30}H_{50}O$: 426.38616; found: 410.38648.

Product 31: $[\alpha]_D^{25}$= +6.3 (*c*=0.074, EtOH); ^{1}H NMR (400 MHz, C_6D_6): δ=0.900 (s, 3H, Me-29), 1.011 (s, 3H, Me-18), 1.030 (s, 3H, Me-19), 1.047 (d, *J*=6.8 Hz, 3H, Me-21), 1.065 (s, 3H, Me-28), 1.131 (s, 3H, Me-30), 1.32 (2H: m, 1H, H-1; m, 1H, H-15), 1.44 (m, 1H, H-22), 1.46 (m, 1H, H-7), 1.50 (m, 1H, H-6), 1.51 (m, 1H, H-2), 1.52 (m, 1H, H-1), 1.57 (2H; m, 1H, H-5; m, 1H, H-6), 1.62 (m, 1H, H-22), 1.63 (m, 2H, H-16), 1.730 (s, 3H, Me-27), 1.77 (m, 1H, H-15), 1.80 (m, 1H, H-7), 1.85 (m, 1H, H-2), 1.884 (s, 3H, Me-26), 1.92 (2H: m, 1H, H-11; m, 1H, H-9), 1.98 (m, 1H, H-20), 2.16 (m, 1H, H-11), 2.23 (m, 2H, H-23), 2.59 (m, 1H, H-17), 3.29 (bs, 1H, H-3), 5.41 (2H: m, 1H, H-12; m, 1H, H-24); ^{13}C NMR (100 MHz, C_6D_6): δ=15.48 (q, C-21), 15.71 (q, C-19), 17.05 (q, C-18), 17.74 (q, C-27), 18.69 (t, C-6), 22.44 (q, C-29), 22.85 (q, C-30), 23.86 (t, C-11), 24.49 (t, C-16), 25.76 (t, C-2), 25.89 (q, C-26), 26.53 (t, C-23), 28.59 (q, C-28), 32.07 (t, C-15), 33.46 (t, C-1), 34.71 (d, C-20), 35.06 (t, C-7), 36.35 (t, C-22),

37.60 (s, C-10), 37.61 (s, C-4), 38.52 (s, C-8), 48.36 (d, C-9), 48.94 (d, C-17), 49.75 (d, C-5), 50.57 (s, C-14), 75.77 (d, C-3), 115.9 (d, C-12), 125.5 (d, C-24), 130.9 (s, C-25), 149.2 (s, C-13). The assignments of C-4 and C-10 may be exchangeable. MS (EI) see Fig. S16.1. HRMS (EI): *m*/*z*: calcd. for $C_{30}H_{50}O$: 426.38616; found: 410.38432.

Product 32: $[\alpha]_D^{25}$= +8.0 (*c*=0.213, EtOH); ^{1}H NMR (600 MHz, C_6D_6): δ=0.865 (dd, *J*=12.0, 10.3 Hz, 1H, H-5), 0.92 (m, 1H, H-1), 0.947 (s, 3H, Me-29), 0.973 (s, 3H, Me-18), 0.991 (s, 3H, Me-19), 1.061 (d, *J*=6.7 Hz, 3H, Me-21), 1.148 (s, 3H, Me-30), 1.157 (s, 3H, Me-28), 1.30 (dd, *J*=6.3, 11.8 Hz, 1H, H-15), 1.42 (m, 1H, H-7), 1.47 (m, 1H, H-22), 1.50 (m, 1H, H-6), 1.52 (m, 1H, H-2), 1.56 (2H: m, 1H, H-1; m, 1H, H-2), 1.62 (m, 2H, H-16), 1.63 (m, 1H, H-22), 1.64 (m, 1H, H-6), 1.68 (m, 1H, H-7), 1.72 (m, 1H, H-9), 1.732 (s, 3H, Me-27), 1.78 (m, 1H, H-15), 1.832 (s, 3H, Me-26), 1.89 (m, 1H, H-11), 1.99 (m, 1H, Me-20), 2.08 (m, 1H, H-11), 2.24 (m, 2H, H-23), 2.59 (m, 1H, H-17), 3.14 (dd, *J*=11.4, 4.7 Hz, 1H, H-3), 5.39 (bs, 1H, H-12), 5.47 (t, *J*=7.4 Hz, H-24); ^{13}C NMR (150 MHz, C_6D_6): δ= 15.48 (q, C-21), 15.88 (2C; q, C-19; q, C-29), 16.97 (q, C-18), 17.73 (q, C-27), 18.72 (t, C-6), 22.83 (q, C-30), 23.91 (t, C-11), 24.50 (t, C-16), 25.88 (q, C-26), 26.53 (t, C-23), 27.71 (t, C-2), 28.32 (q, C-28), 32.08 (t, C-15), 34.72 (d, C-20), 35.08 (t, C-7), 36.34 (t, C-22), 37.48 (s, C-10), 38.30 (s, C-8), 38.96 (t, C-1), 39.06 (s, C-4), 48.57 (d, C-9), 48.94 (d, C-17), 50.47 (s, C-14), 56.32 (d, C-5), 78.50 (d, C-3), 115.9 (d, C-12), 125.5 (d, C-24), 130.9 (s, C-25), 149.3 (s, C-13). MS (EI) see Fig. S17.1. HRMS (EI): *m*/*z*: calcd. for $C_{30}H_{50}O$: 426.38616; found: 410.38648.

Product 33: $[\alpha]_D^{25}$= -20.8 (*c*=0.239, EtOH); ^{1}H NMR (600 MHz, acetone d_6): δ=0.817 (s, 3H, Me-29), 0.82 (m, 1H, H-5), 0.849 (s, 3H, Me-28), 0.86 (m, 1H, H-1), 0.861 (s, 3H, Me-19), 0.969 (s, 3H, Me-18), 1.025 (s, 3H, Me-30), 1.050 (s, 3H, Me-27), 1.06 (m, 1H, H-15), 1.135 (s, 3H, Me-26), 1.14 (m, 1H, H-3), 1.188 (s, 3H, Me-21), 1.21 (m, 1H, H-11), 1.26 (m, 1H, H-7), 1.32 (m, 1H, H-3), 1.38 (m, 1H, H-2), 1.39 (m, 1H, H-6), 1.43 (2H: m, 1H, H-15; m, 1H, H-9), 1.49 (m, 1H, H-12), 1.52 (m, 1H, H-6), 1.53 (m, 1H, H-16), 1.58 (m, 1H, H-11), 1.59 (m, 1H, H-22), 1.61 (2H: m, 1H, H-2; m, 1H, H-7), 1.65 (m, 1H, H-1), 1.75 (m, 1H, H-23), 1.76 (m, 1H, H-16), 1.85 (m, 1H, H-23), 1.92 (2H: m, 1H, H-12; m, 1H, H-22), 2.09 (m, 1H, H-13), 2.24 (m, 1H, H-17), 3.71 (dd, *J*=10.6, 5.2 Hz, 1H, H-24); ^{13}C NMR (150 MHz, acetone d_6): δ= 16.33 (q, C-18), 16.64 (q, C-19), 16.94 (q, C-30), 19.37 (t, C-6), 19.47 (t, C-2), 21.86 (q, C-29), 23.21 (t, C-11), 25.83 (q, C-27), 26.40 (t, C-23), 26.84 (t, C-12), 27.37 (q, C-26), 28.29 (t, C-16), 28.91 (q, C-21), 33.12 (t, C-15), 33.76 (q, C-28), 33.95 (s, C-4), 36.09 (t, C-7), 38.19 (s, C-10), 38.70 (t, C-22), 41.35 (t, C-1), 41.68 (s, C-8), 42.89 (t, C-3), 44.60 (d, C-13), 49.36 (d, C-17), 49.85 (s, C-14), 51.71 (d, C-9), 57.79 (d, C-5), 70.43 (s, C-25), 86.06 (s, C-20), 88.23 (d, C-24). MS (EI) see Fig. S18.1. HRMS (EI): *m*/*z* 429 (M^+-Me):

calcd. for $C_{29}H_{49}O_2$: 426.37325; found: 429.37229.

Product 34-acetate: $[\alpha]_D^{25}$= -9.5 (*c*=0.014, EtOH); ^{1}H NMR (600 MHz,C_6D_6): δ=0.892 (s, 3H, Me-24), 1.051 (s, 3H, Me-25), 1.052 (s, 3H, Me-23), 1.50 (m, 1H, H-6), 1.62 (m, 1H, H-6), 1.68 (m, 1H, H-1), 1.691 (s, 3H, Me-29), 1.724 (s, 3H, Me-26), 1.724 (s, 3H, Me-28), 1.779 (s, 3H, Me-27), 1.806 (s, 3H, Me-30), 1.82 (s, 3H, Me-32, acetyl Me), 1.85 (m, 1H, H-1, 1.88 (2H, m, H-2), 1.92 (m, 1H, H-5), 2.01(dd, *J*=6.0, 18.2 Hz, 1H, H-7), 2.14 (m, 1H, H-12), 2.16 (m, 1H, H-7), 2.24 (4H: m, H-15 and H-19), 2.32 (m, 1H, H-12), 2.33 (m, 4H, H-16 and H-20), 2.38 (m, 2H, H-11), 5.03 (bs, 1H, H-3), 5.37 (t, *J*=6.8 Hz, 1H, H-21), 5.42 (t, *J*=6.8 Hz, 1H, H-13), 5.43 (t, *J*=6.8 Hz, 1H, H-17); ^{13}C NMR (150 MHz, C_6D_6): δ= 16.12 (q, C-28), 16.21 (q, C-27), 169.7 (s, C-31, acetyl <u>C</u>O), 17.72 (q, C-29), 18.87 (t, C-6), 19.70 (q, C-26), 20.02 (q, C-25), 20.76 (q, C-32, acetyl <u>C</u>H_3), 21.84 (q, C-24), 23.71 (t, C-2), 25.83 (q, C-30), 27.20 (2C: t, C-16 and t, C-20), 27.88 (q, C-23), 28.74 (t, C-12), 29.65 (t, C-11), 30.75 (t, C-1), 33.66 (t, C-7), 36.99 (s, C-4), 38.98 (s, C-10), 40.19 (t, C-15), 40.19 (t, C-19), 46.52 (d, C-5), 77.48 (d, C-3), 124.7 (d, C-17), 124.9 (d, C-21), 125.3 (d, C-13), 126.2 (s, C-8), 131.1 (s, C-22), 134.8 (s, C-18), 135.1 (s, C-14), 140.4 (s, C-9). The assignments of C-14 and C-18 may be exchangeable. MS (EI) see Fig. S19.1. HRMS (EI): *m*/*z* 408 (M^+-CH_3CO): calcd. for $C_{30}H_{48}$: 408.37560; found: 408.37590.

Product 35-acetate: $[\alpha]_D^{25}$= -20.0 (*c*=0.025, EtOH); ^{1}H NMR (600 MHz,C_6D_6): δ=0.867 (s, 3H, Me-25), 0.893 (s, 3H, Me-24), 0.957 (s, 3H, Me-23), 1.11 (m, 1H, H-1), 1.44 (m, 1H, H-11), 1.56 (m, 1H, H-1), 1.61 (m, 1H, H-11), 1.691 (s, 3H, Me-29), 1.740 (s, 6H, Me-27 and Me-28), 1.806 (s, 3H, Me-30), 1.83 (3H: m, 2H, H-2; m, 1H, H-5), 1.88 (m, 1H, H-9), 1.889 (s, 3H, Me-26), 1.890 (s, 3H, Me-32, acetyl Me), 1.96 (very broad, 2H, H-6), 2.15 (m, 1H, H-12), 2.24 (m, 4H, H-15 and H-19), 2.33 (m, 4H, H-16 and H-20), 2.34 (m, 1H, H-12), 4.99 (t, *J*=2.4 Hz, H-3), 5.37 (t, *J*=6.8 Hz, H-21), 5.42 (t, *J*=6.8 Hz, H-13), 5.43 (t, *J*=6.8 Hz, H-17), 5.45 (very br s, 1H, H-7) ; ^{13}C NMR (150 MHz, C_6D_6): δ= 13.57 (q, C-25), 16.10 (q, C-28), 16.28 (q, C-27), 169.6 (s, acetyl <u>C</u>O), 17.72 (q, C-29), 20.77 (q, acetyl <u>C</u>H_3), 21.90 (q, C-24), 22.36 (q, C-26), 23.08 (t, C-2), 23.59 (t, C-6), 25.83 (q, C-30), 27.20 (2C: t, C-16 and C-20), 27.67 (2C: q, C-23; t, C-11), 30.61 (t, C-12), 32.29 (t, C-1), 36.54 (s, C-4), 36.58 (s, C-10), 40.19 (t, C-19), 40.22 (t, C-15), 44.99 (d, C-5), 54.12 (d, C-9), 77.90 (d, C-3), 122.4 (d, C-7), 124.7 (d, C-17), 124.9 (d, C-21),125.3 (d, C-13), 131.1 (s, C-22), 135.1 (s, C-18), 135.2 (s, C-14), 135.6 (s, C-8). The assignments may be exchangeable between C-4 and C-10, between C-14 and C-18, between C-27 and C-28 and C-15 and C-19. MS (EI) see Fig. S20.1. HRMS (EI): *m*/*z* 408 (M^+-CH_3CO): calcd. for $C_{30}H_{48}$: 408.37560; found: 408.37759.

Product 36: $[\alpha]_D^{25}$= -34.4 (*c*=1.697, EtOH); ^{1}H NMR (400 MHz,C_6D_6): δ=0.878 (s,

3H, Me-25), 0.975 (s, 3H, Me-24), 1.084 (s, 3H, Me-23), 1.11 (m, 1H, H-1), 1.22 (t, *J*=8.4 Hz, H-5), 1.40 (m, 1H, H-11), 1.41 (m, 1H, H-2), 1.55 (m, 1H, H-2), 1.60 (m, 1H, H-11), 1.69 (m, 1H, H-9), 1.694 (s, 3H, Me-29), 1.750 (s, 3H, Me-28), 1.768 (s, 3H, Me-27), 1.807 (s, 3H, Me-30), 1.835 (s, 3H, Me-26), 1.87 (m, 1H, H-1), 2.02 (very b r, 2H, H-6), 2.16 (m, 1H, H-12), 2.27 (m, 4H, H-15 and H-19), 2.27-2.38 (m, 4H, H-16 and H-20), 2.38 (m, 1H, H-12), 3.15 (dd, *J*=10.8, 4.8 Hz, H-3), 5.37 (t, *J*=6.8 Hz, 1H, H-21), 5.42 (t, *J*=6.8 Hz, 1H, H-13), 5.44 (t, *J*=6.8 Hz, 1H, H-17), 5.53 (very br s, 1H, H-7); ^{13}C NMR (100 MHz, C_6D_6): δ=13.71 (q, C-25), 131.1 (s, C-22), 16.12 (q, C-28), 16.26 (q, C-27), 17.73 (q, C-29), 22.27 (q, C-26), 23.81 (t, C-6), 25.30 (q, C-24), 25.84 (q, C-30), 27.13 (t, C-20), 27.23 (t, C-16), 27.70 (t, C-11), 27.93 (t, C-2), 28.12 (q, C-23), 30.66 (t, C-12), 36.78 (s, C-10), 37.54 (t, C-1), 38.83 (s, C-4), 40.22 (2C, t, C-15 and C-19), 49.84 (d, C-5), 54.37 (d, C-9), 78.76 (d, C-3), 122.5 (d, C-7), 124.7 (d, C-17), 124.9 (d, C-21), 125.3 (d, C-13), 135.1 (2C, s, C-14 and C-18), 135.2 (s, C-8). The assignments may be exchangeable between C-14 and C-18, between C-16 and C-20 and between C-27 and C-28. MS (EI) see Fig. S21.1. HRMS (EI): *m*/*z*: calcd. for $C_{30}H_{50}O$: 426.38616; found: 410.38354.

Acknowledgements

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

Keywords: squalene; cyclization, enzyme catalysis; hopene; epoxydammarane

References

[1] a) R. B. Woodward , K. Bloch, *J. Am. Chem. Soc.*, **1953**, *75*, 2023-2024. b) A. Eschenmoser, L. Ruzicka, O. Jeger , D. Arigoni, *Helv. Chim. Acta* **1955**, *38*, 1890–1904. c) J. W. Cornforth, R. H. Cornforth , C. Donniger, G. Popjak, Y. Shimizu, S. Ichii, E. Forchielli , E. Caspi, *J. Am. Chem. Soc.*, **1965**, *87*, 3224-3228.

[2] For a review of squalene, oxidosqualene cyclases, see: a) K. U. Wendt, G. E. Schulz, E. J. Corey, D. R. Liu, *Angew. Chem., Int. Ed.* **2000**, *39*, 2812–2833; b) T. Hoshino, T. Sato, *Chem. Commun.* **2002**, 291–301; c) R. A. Yoder, J. N. Johnston, *Chem. Rev.* **2005**, *105*, 4730–4756; d) M. J. R. Segura, B. E. Jackson, S. P. T. Matsuda, *Nat. Prod. Rep.*, **2003**, *20*, 304-317; e) I. Abe, *Nat. Prod. Rep.* **2007**, *24*, 1311–1331; f) T. K. Wu, C. H. Chang, Y. T. Liu, T. T. Wang, *Chem. Rec.* **2008**, *8*, 302–325; g) W.

D. Nes, *Chem. Rev.* **2011**, *111*, 6423–6451; h) G. Siedenburg, D. Jendrossek, *Appl. Environ. Microbiol.*, **2011**, *77*, 3905-3915; i) P.-O. Syrén, S. Henche, A. Eichler, B. M. Nestl, B. Hauer, *Curr. Opin. Struct. Biol.* **2016**, *41*, 73-82; j) T. Hoshino, *Org. Biomol. Chem.*, **2017**, *15*, 2869-2891.

[3] K. U. Wendt, K. Poralla, G. E. Schulz, *Science*, **1997**, *277*, 1811-1815.

[4] a) K. U. Wendt, A. Lenhart, G. E. Schulz, *J. Mol. Biol.*, **1999**, *286*, 175-187; b) D. J. Reinert, G. Balliano, G. E. Schulz, *Chem. Biol.*, **2004**, *6*, 1717-1720.

[5] a) R. Rajamani, J. Gao, *J. Am. Chem. Soc.*, **2003**, *125*, 12768-12781; b) L. Smentek, B. A. Hess, Jr., *J. Am. Chem. Soc.*, **2010**, *132*, 17111-17117.

[6] a) T. Hoshino, S. Nakano, T. Kondo, T. Sato, A. Miyoshi, *Org. Biomol. Chem.*, **2004**, *2*, 1456-1470; b) M. Rohmer, C. Anding, G. Ourisson, *Eur. J. Biochem.*, **1980**, *112*, 541-547; c) P. Bouvier, Y. Berger, M. Rohmer, G. Ourisson, *Eur. J. Biochem.*, **1980**, *112*, 549-556; d) R. Rohmer, P. Bouvier, G. Ourisson, *Eur. J. Biochem.*, **1980**, *112*, 557-560.

[7] a) T. Sato, T. Hoshino, *Biosci, Biotechnol. Biochem.*, **1999**, *63*, 2189-2198; (b) T. Dang, G. D. Prestwich, *Chem. Biol.*, **2000**, *7*, 643–649.

[8] T. Sato, T. Hoshino, *Biosci, Biotechnol. Biochem.*, **2001**, *65*, 2233-2142.

[9] T. Hoshino, T. Sato, *Chem. commun.*, **1999**, 2005-2006.

[10] T. Hoshino, M. Kouda, T. Abe , S. Ohashi, *Biosci, Biotechnol. Biochem.*, **1999**, *63*, 2038-2041.

[11] T. Hoshino, M. Kouda, T. Abe , T. Sato, *Chem. Commun.*, **2000**, 1485-1486.

[12] M. Morikubo, Y. Fukuda, K. Ohtake, N. Shinya, D. Kiga, K. Sakamoto, M. Asanuma, H. Hirota, S. Yokoyama , T. Hoshino, *J. Am. Chem. Soc.*, **2006**, *128*, 13148-13194.

[13] T. Sato, S. Sasahara, T. Yamakami, T. Hoshino, *Biosci, Biotechnol. Biochem.*, **2002**, *66*, 1660-1670.

[14] T. Hoshino, T. Abe, M. Kouda, *Chem. Commun.*, **2000**, 441-442.

[15] a) L. C. Kühnel, B. M. Nestl, B. Hauer, *ChemBioChem*, **2017**, *18*, 2222-2225; b) S. A. Bastian, S. C. Hammer, N. Kreß, B. M. Nestle, B. Hauer, *ChemCatChem.*, **2017**, *9*, 4364-4368; c) S. C. Hammer, A. Marjanovic, J. M. Dominicus, B. M. Nestle, B. Hauer, *Nat. Chem. Biol.*, **2015**, *11*, 121-126; d) M. Seitz, P. O. Syrén, M. Seitz, B. M. Nestle, B. Hauer, *ChemBioChem*, **2013**, *14*, 436-439.

[16] T. Sato and T. Hoshino, *Biosci, Biotechnol. Biochem.*, **1999**, *63*, 1171-1180.

[17] T. Sato, T. Abe, T. Hoshino, *Chem. Commun.*, **1998**, 2617-2618.

[18] I. Abe , M. Rohmer, *J. Chem. Soc., Perkin Trans 1*, **1994**, 783-791.

[19] T. Sato, M. Kouda, T. Hoshino, *Biosci., Biotechnol. Biochem.*, **2004**, *68*, 728-738.

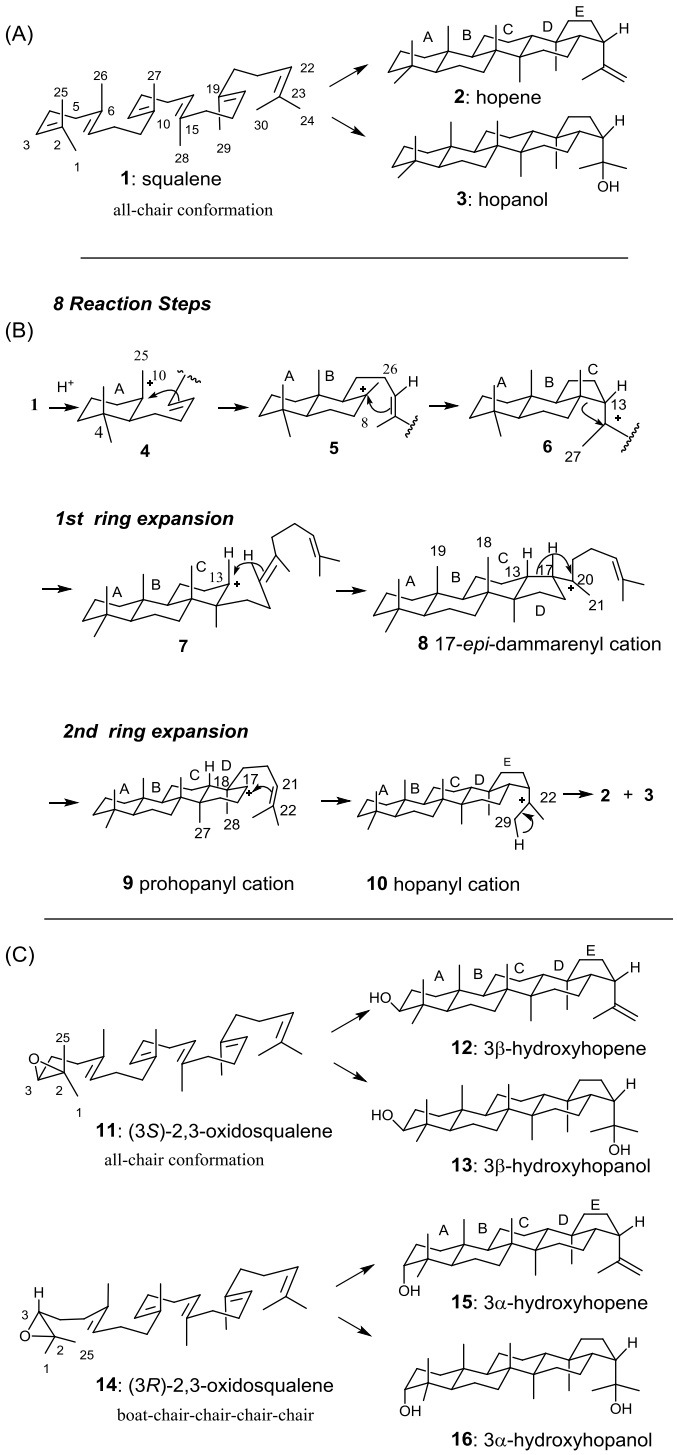
[20] T. Hoshino, Y. Yonemura, T. Abe , Y. Sugino, *Org. Biomol. Chem.*, **2007**, *5*, 792-801.

10.1002/cbic.201800281

[21] D. Ueda, T. Hoshino, T. Sato, *J. Am. Chem. Soc.*, **2013**, *135*,18335-18338.
[22] Z-h Lin, H-x Long, Z. Bo, Y-q Wang, , Y-z Wu, *Peptides*, **2008**, *29*, 1798-1805.
[23] a) T. Hoshino and T. Kondo, *Chem. Commun*. **1999**, 731-732; b) T. Hoshino and S. Ohashi , *Org. Lett.*, **2002**, *4*, 2553-2556; c) S. Nakano, S. Ohashi and T. Hoshino, *Org. Biomol. Chem.*, **2004**, *2*, 2012-2022.
[24] a) K. Takahashi, Y. Sasaki, T. Hoshino, *Eur. J. Org. Chem.* **2018**, 1477-1490; b) I. Kaneko, Y. Terasawa, T. Hoshino, *Chem. Eur. J.*, in press: doi.org/10.1002/chem.201801668

10.1002/cbic.201800281

Figure, Scheme, and Table Legends



Scheme 1. (A) Polycyclization pathways of squalene (**1**) into hopene (**2**) and hopanol (**3**) by squalene-hopene cyclase. Squalene **1** folds in an all-prechair conformation in the

10.1002/cbic.201800281

reaction cavity. (B) The cyclization consists of 8 reaction steps. (C) Polycyclization reactions of (3*S*)-2,3-oxidosqualene (**11**) and (3*R*)-2,3-oxidosqualene (**14**). Here, **11** folds into an all-prechair structure similar to that in the polycyclization reaction of **1**, leading to 3β-hydroxyhopene (**12**) and 3β-hydroxyhopanol (**13**). However, **14** folds into a boat-chair-chair-chair-chair conformation, resulting in the production of 3α-hydroxyhopene (**15**) and 3α-hydroxyhopanol (**16**).

10.1002/cbic.201800281

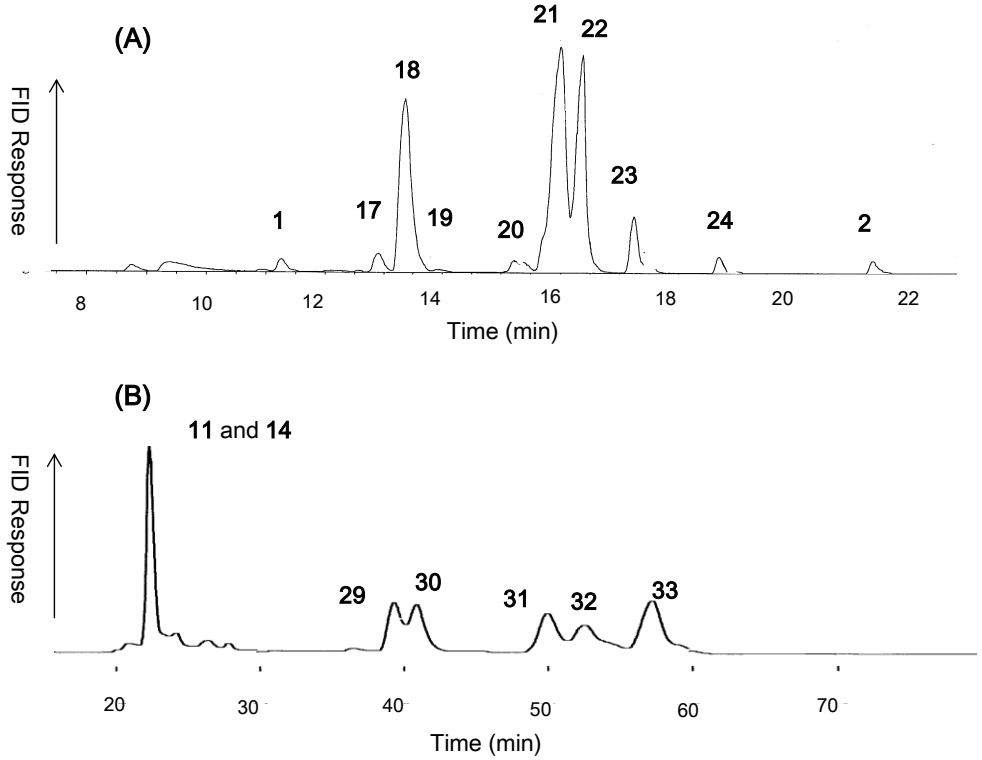


Fig. 1. (A) Incubation mixture of squalene (**1**) with the A306V variant. (B) Incubation mixture of 2,3-oxidosqualene (a mixture of **11** and **14**) with the A306V variant. The cell-free extracts (5 mL) were prepared from a 100 mL culture of the A306V variant. Incubation of 1.0 mg of **1** and that of a mixture of **11** and **14** (1.0 mg) with 2.0 mL of cell-free homogenates was carried out at 45°C for 24 h. The lipophilic materials were extracted with hexane. Triton X-100 present in the hexane extract was removed by passage through a short SiO_2 column using hexane/EtOAc (100:20) as eluent. The lipophilic materials thus collected were subjected to GC measurements. The GC conditions are described in the Experimental section.

10.1002/cbic.201800281

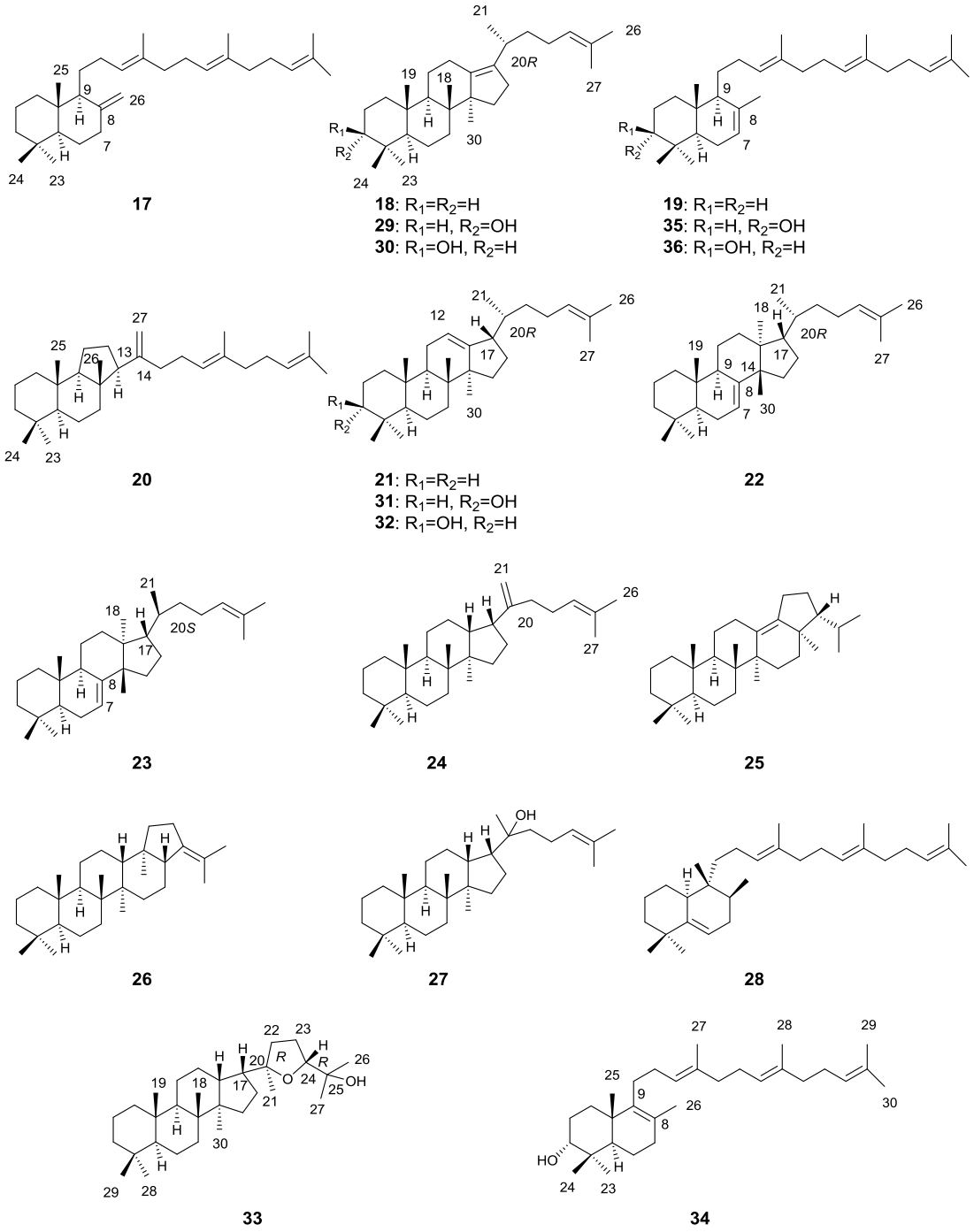


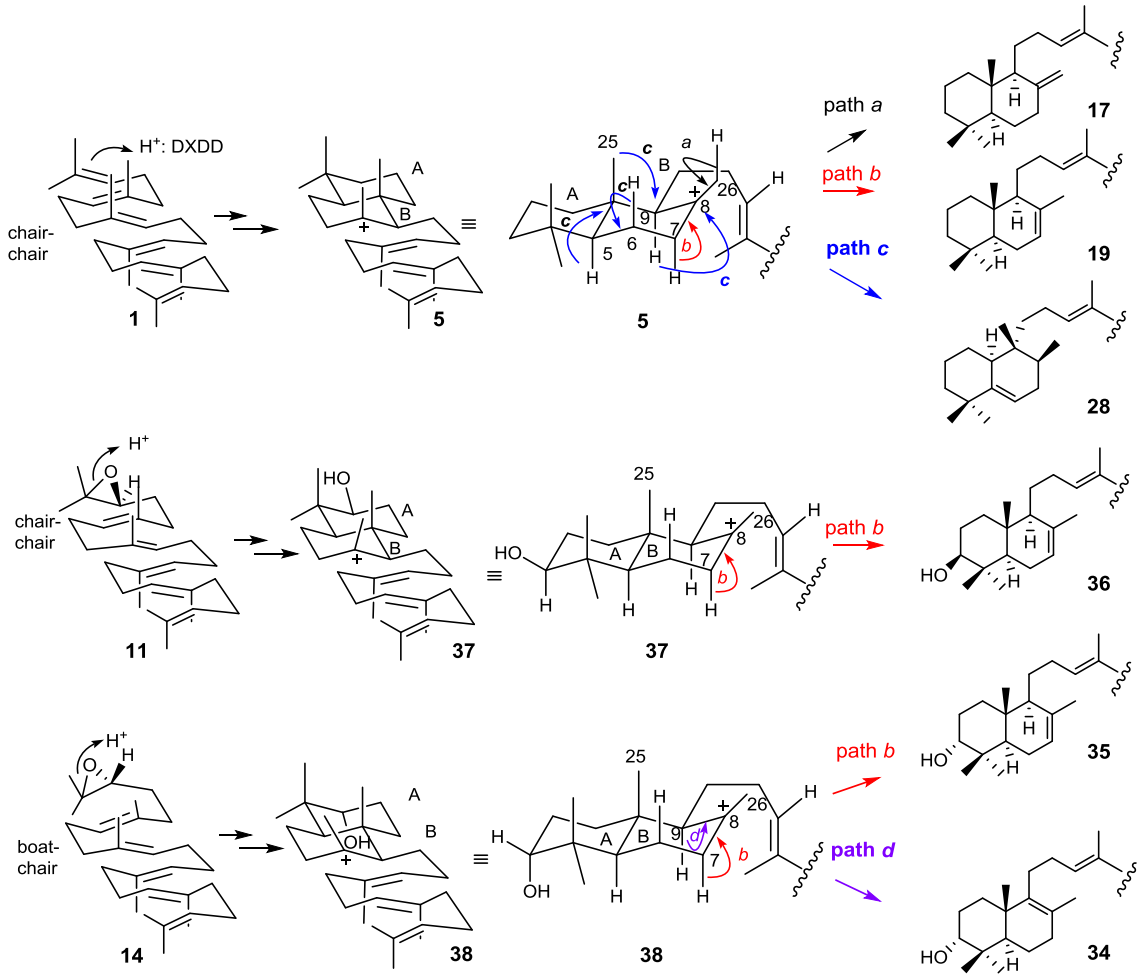
Fig. 2. Structures of the enzymatic products produced by A306X variants. Products **17**–**24** were produced from the incubation of **1** with the A306V mutant. Products **18**, **21**–**27**

10.1002/cbic.201800281

were generated by the incubation of **1** with the A306T variant (Fig. S3.1). Products **17**, **19**, and **28** were produced by the reaction of **1** with the A306F mutant (Fig. S.1.2). Fig. S4 shows the GC profiles of the incubation mixtures of a racemic mixture of **11** and **14** with the A360T and A306F variants. Products **29**–**33** were produced by the A306T mutant. Products **34**–**36** were produced by the A306F mutant.

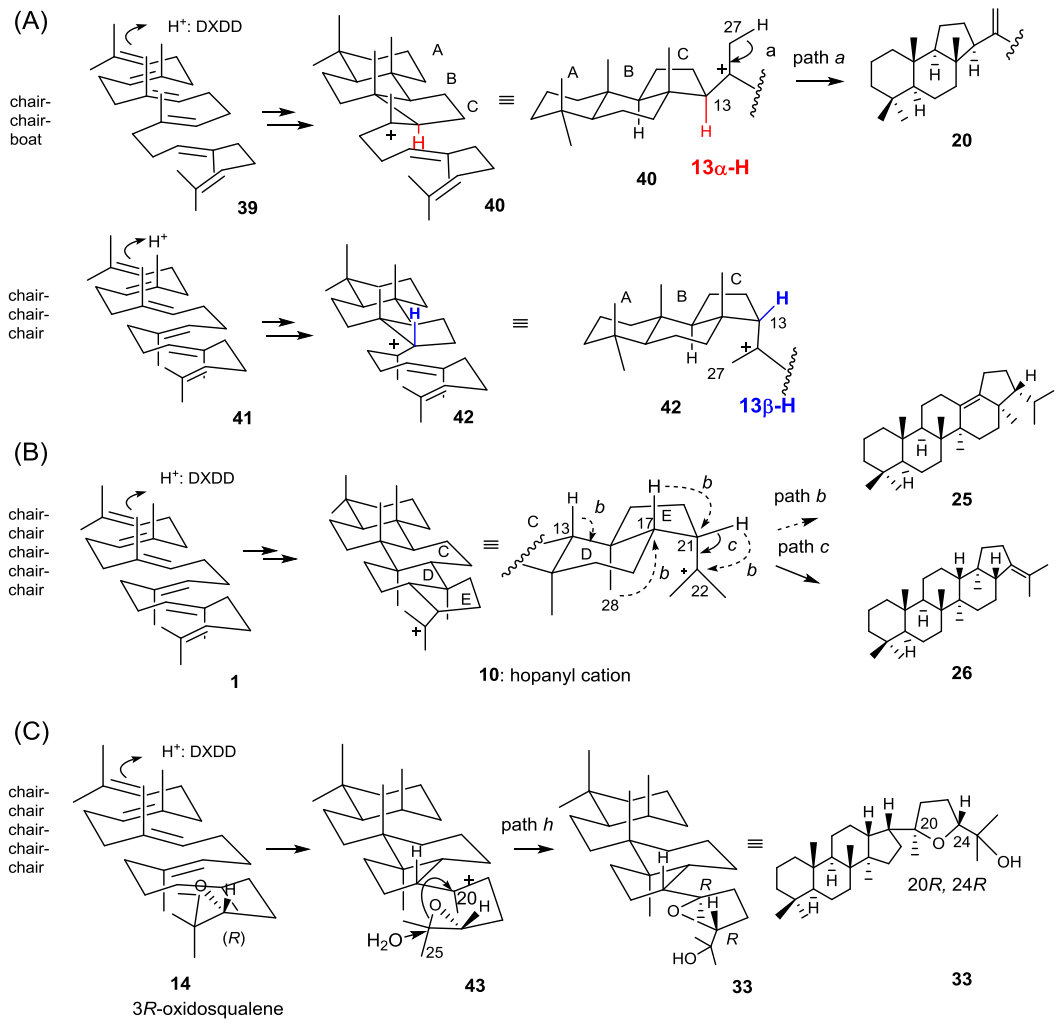
Compound names. **17**: α-polypodateraene, **18**: dammara-13(17)-24-diene, **19**: γ-polypodateraene, **20**: (17*E*)-(13αH)-malabarica-14(27),17,21-triene, **21**: 20*R*-dammara-12(13), 24-diene, **22**: 20*R*-eupha-7(8),24-diene, **23**: 20*S*-tirucalla-7(8),24-diene (a new SHC-mediated product), **24**: 17-*epi*-dammara-20(21),24-diene, **25**: neohop-13(18)-ene, **26**: hop-21(22)-ene, **27**: 20-hydroxy-17-*epi*-dammaraene, **28**: neopolypoda-5(6),13,17,21-teraene (rearranged skeleton of bicyclic polypodatetraene), **29**: 20*R*-dammara-13(17),24-dien-3α-ol, **30**: 20*R*-dammara-13(17),24-dien-3β-ol, **31**: 20*R*-dammara-12(13),24-dien-3α-ol, **32**: 20*R*-dammara-12(13),24-dien-3β-ol, **33**: (20*R*, 24*R*)-17-*epi*-epoxydammarane, **34**: polypoda-8(9),13,17,21-tetraen-3α-ol, **35**: polypoda-7(8),13,17,21-tetraen-3α-ol (3α-hydroxy-γ-polypodatetraene), **37**: polypoda-7(8),13,17,21-tetraen-3β-ol (3β-hydroxy-γ-polypodatetraene).

10.1002/cbic.201800281



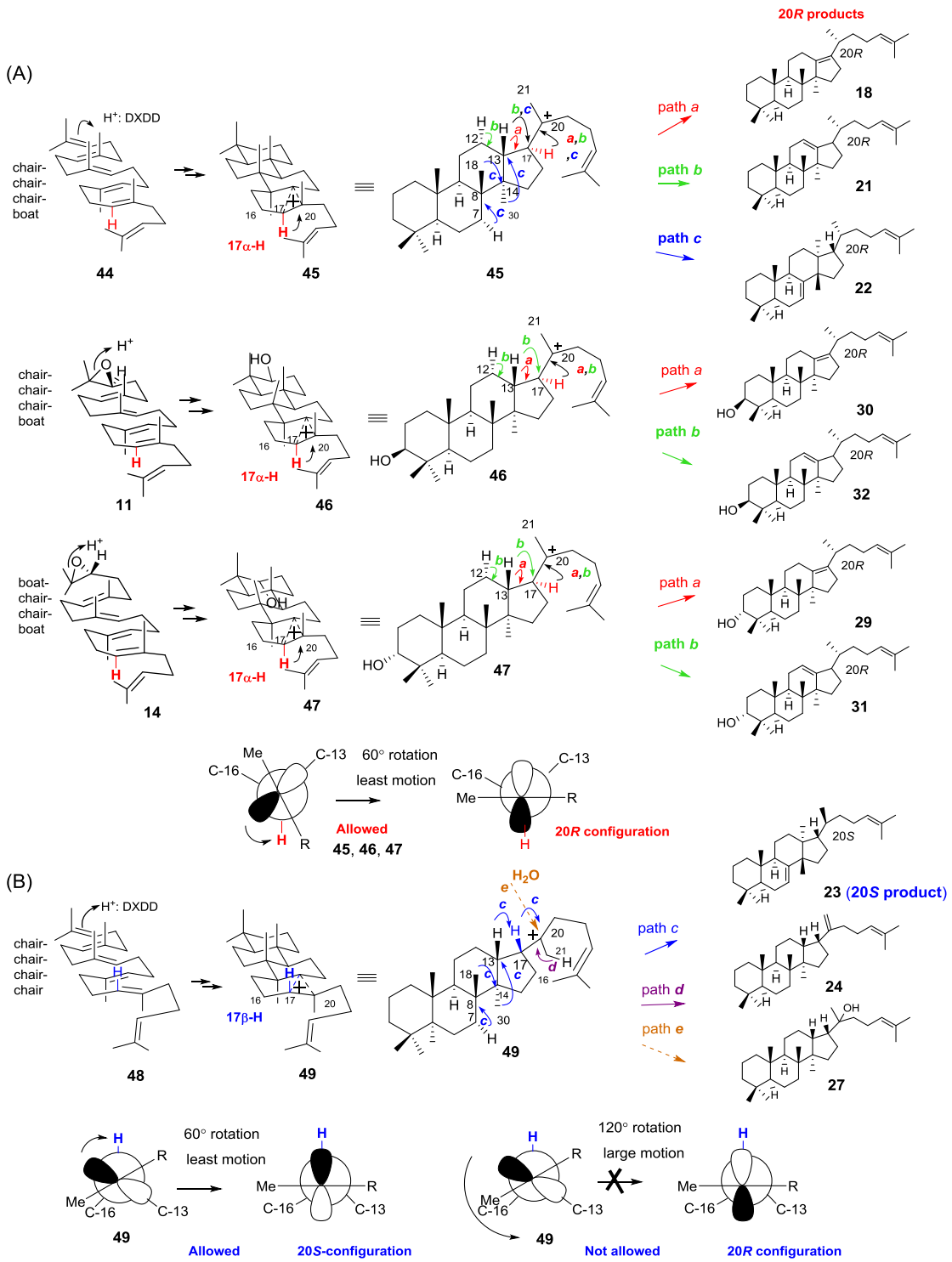
Scheme 2. Cyclization mechanisms of **1**, **11**, and **14** leading to bicyclic products **17**, **19**, **28**, and **34**–**36**. Substrates **1** and **11** fold in a chair-chair conformation, but substrate **14** folds in a boat-chair structure.

10.1002/cbic.201800281



Scheme 3. Cyclization mechanisms of **1** and **14** leading to tricyclic product **20** and pentacyclic products **25**, **26**, and **33**. (A) A chair-chair-boat conformation yielded the tricyclic **20**. (B) The all-chair conformation furnished the neohopane scaffold **25** and hopane skeleton **26**. (C) (3*R*)-oxidosqualene **14** was folded into an all-chair conformation to yield 20*R*, 24*R*-17-*epi*-epoxydammarane **33**. Notably, cyclization started from the isopropylidene moiety rather than from the epoxide ring.

10.1002/cbic.201800281



Scheme 4. Formation mechanisms of 6,6,6,5-fused tetracyclic products with 20*R* and 20*S* configurations. Cation **8** is formed by the folding of **1** into a chair-chair-chair-chair

10.1002/cbic.201800281

conformation (Scheme 1). This indicates that the native SHC furnishes the tetracyclic product(s) with the 20*S* configuration, in contrast to the A306V mutant.

Table 1. Product distribution ratio (%) obtained from the reactions of squalene (**1**) with A306X variants. The reported van der Waals volumes (nm^3) are as follows: Gly, 0.00279; Ala, 0.05702; Thr, 0.19341; Val, 0.25674; Leu, 0.37876; Phe, 0.55298, [22] thus, the steric bulk increases as follows: Gly<Ala<Thr<Val<Leu<Phe.

		Normal products		Bicycle			Tricycle	Tetracycle						Pentacycle	
Compounds	**1**	**2**	**3**	**17**	**19**	**28**	**20**	**18**	**21**	**22**	**23**	**24**	**27**	**25**	**26**
Wild type SHC	–	84	16	–	–	–	–	–	–	–	–	–	–	–	–
A306G	43	45	7	–	–	–	–	–	–	–	–	–	–	4	2
A306T	2	23	2	–	–	–	–	11	13	12	8	9	4	1	15
A306V	2	2	–	3	1	–	2	22	31	27	8	2	–	–	–
A306L	20	1	–	6	31	–	2	21	13	3	–	3	–	–	–
A306F	56	–	–	1	42	1	–	–	–	–	–	–	–	–	–

–: little or no production, below the limit of GC detection. There might be other products that have not been characterized, but the production amounts are a little, if any present.

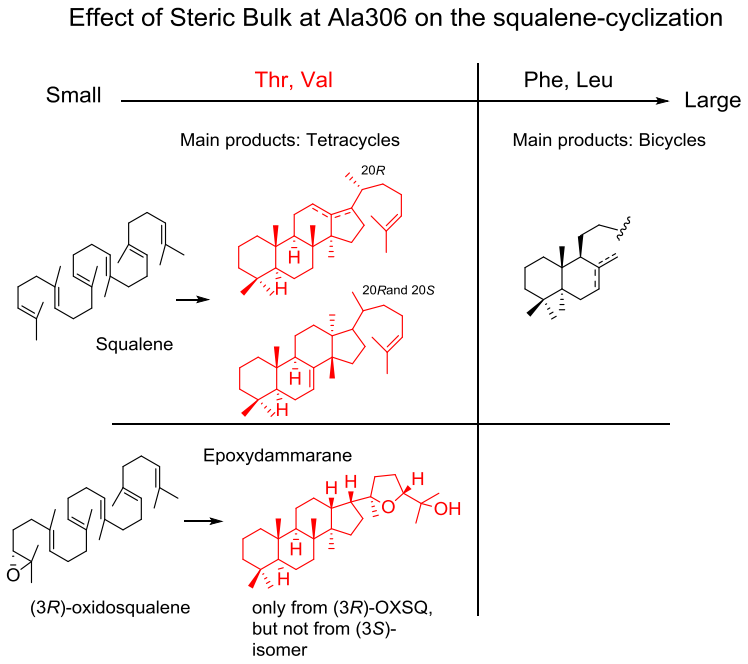
Table 2. Product distribution ratios (%) generated by incubating oxidosqualene (**11** and **14**) with A306X mutants. The steric bulk increases as follows: Ala<Thr<Val<Phe. [22]

		Normal products				Bicycle			Tetracycle				Epoxy-dammarane
Compounds	**11** & **14**	**12**	**13**	**15**	**16**	**34**	**35**	**36**	**29**	**30**	**31**	**32**	**33**
Wild type SHC	–	54.2	9.5	31.3	4.9	–	–	–	–	–	–	–	–
A306T	2	4	–	4	–	–	–	–	24	26	18	14	8
A306V	33	–	–	–	–	–	–	–	15	13	11	8	20
A306F	42	–	–	–	–	8	12	38	–	–	–	–	–

–: negligible or trace amount

10.1002/cbic.201800281

TOC graphic



The Ala306 in *A. acidocaldarius* SHC is widely conserved in known SHCs. Increases in steric bulk (A306T and A306V) led to the accumulation of 6,6,6,5-fused tetracyclic scaffolds possessing 20*R* stereochemistry in high yield (94% for A306V). The production of the 20*R* configuration indicates that squalene had been folded in a chair-chair-chair-boat conformation, indicating that the bulk at position 306 significantly affects the stereochemical fate during the polycyclization reaction. Intriguingly, epoxydammarane was generated from (3*R*)-oxidosqualene.