Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: Y. Fukuda, T. Watanabe and T. Hoshino, *Org. Biomol. Chem.*, 2018, DOI: 10.1039/C8OB02009D.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/obc

View Article Online
DOI: 10.1039/C8OB02009D

Mutated variants of squalene-hopene cyclase: Enzymatic syntheses of triterpenes bearing oxygen-bridged monocycles and a new 6,6,6,6,6-fusded pentacyclic scaffold, named neogammacerane, from 2,3-oxidosqualene

Yoriyuki Fukuda, Takashi Watanabe, 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

Electronic Supporting Information (ESI) available: Primers used for the mutagenesis, GC profiles of the enzymatic reactions, EIMS and NMR spectra of the enzymatic products, and the NMR analyses for proposing structures.

View Article Online
DOI: 10.1039/C8OB02009D

Abstract

Squalene-hopene cyclase (SHC) catalyzes the conversion of acyclic squalene molecule into a 6,6,6,6,5-fused pentacyclic hopene and hopanol. SHC is also able to convert (3*S*)-2,3-oxidosqualene into 3β-hydroxyhopene and 3β-hydroxyhopanol and can generate 3α-hydroxyhopene and 3α-hydroxyhopanol from (3*R*)-2,3-oxidosqualene. Functional analyses of active site residues toward the squalene cyclization reaction have extensively been reported, but investigations of the cyclization reactions of (3*R*,*S*)-oxidosqualene by SHC have rarely been reported. The cyclization reactions of oxidosqualene with W169X, G600F/F601G and F601G/P602F were examined. The variants of the W169L generated new triterpene skeletons possessing a 7-oxabicyclo[2.2.1]heptane moiety (oxygen-bridged monocycle) with (1*S*,2*S*,4*R*)- and (1*R*,2*S*,4*S*)-stereochemistry, which were produced from (3*R*)- and (3*S*)-oxidosqualenes, respectively. The F601G/P602F double mutant also furnished a novel triterpene, named neogammacer-21(22)-en-3β-ol, consisting of a 6,6,6,6,6-fused pentacyclic system, in which Me-29 at C-22 of the gammacerane skeleton migrated to C-21. We propose to name this novel scaffold neogammacerane. The formation mechanisms of the enzymatic products from 2,3-oxidosqualene are discussed.

Introduction

The polycyclization cascades of squalene 1 and (3*S*)-2,3-oxidosqualene **4** have been attractive to organic chemists and biochemists for more than 70 years since Ruzicka and coworkers proposed the "biogenetic isoprene rule". [1] The reactions proceed with complete regiospecificity and stereospecificity to yield sterols and triterpenes[2] with remarkable structural diversity; more than 100 different carbon frameworks that exhibit important biological activities can be produced.[3] Squalene-hopene cyclase (SHC) converts squalene **1** into a 6,6,6,6,5-fused pentacyclic hopene **2** and hopanol **3** at a ratio of approximately 5:1. Substrate **1** folds into an all-chair conformation inside the enzyme cavity, as shown in Scheme 1A. SHC can also tolerate both (3*S*)-2,3-oxidosqualene **4** and (3*R*)-2,3-oxidosqualene **7**, yielding 3β-hydroxyhopene **5** and 3β-hydroxyhopanol **6** from **4** and 3α–hydroxyhopene **8** and 3α–hydroxyhopanol **9** from **7** (Scheme 1B), in which **4** folds into an all-chair structure, while **7** folds into a boat-chair-chair-chair-chair conformation in the reaction cavity.[4] These polycyclization cascades are attained by a single enzyme to yield five C-C bonds and nine stereocenters. Recent studies on the SHC from the bacterium sp. *Alicyclobacillus acidocaldarius* (*Aa*SHC) have revealed that *Aa*SHC has high potential for developing new compounds, including terpene and non-terpenoid skeletons, which result from SHC-mediated enzymatic reactions.[4d,5, 6] We

View Article Online
DOI: 10.1039/C8OB02009D

have reported functional analyses of active site residues for the cyclization reactions of squalene molecule **1** by *Aa*SHC.[2b, 2j] Intriguingly, we recently discovered that the site-specific mutant A306V of *Aa*SHC generated an epoxydammarane skeleton from (3*R*)-oxidosqualene **7**.[7] This result indicated that the proton from the DXDD sequence[8] attacked the terminal double bond, but not on the epoxide ring, despite the cyclization reaction usually occurring from the epoxide ring, as seen in Scheme 1B. This unusual cyclization outcome prompted us to reinvestigate the cyclization reactions of **4** and **7** with mutated SHCs.

We previously reported the cyclization reactions of **1** with W169X variants (X=Val, Phe and Tyr) and that π-electrons are necessary at this position for complete polycyclization,[9] but we have not reported on the reactions of **4** and **7**. Prokaryotic SHCs accept **1**, **4** and **7** as substrates to yield **2**, **3**, **5**, **6**, **8** and **9**, as shown in Schemes 1A and 1B, but eukaryotic oxidosqualene cyclases (OSCs) only accept (3*S*)-**4** to generate cyclic triterpenes possessing a 3β-hydroxyl group.[2] Comparison of the amino acid alignment of prokaryotic SHCs to eukaryotic OSCs suggests that OSCs lack Gly600, which is conserved among prokaryotic SHCs. Therefore, we created the ΔG600 variant and incubated **1**, **4** and **7** with the ΔG600-*Aa*SHC variant.[10] No reaction occurred for substrates **1** and **7**, but (3*S*)-**4** underwent the polycyclization reaction to yield mono- and tricyclic products with the 3β-hydroxyl group.[10] This is the first report of the successful transformation of the prokaryotic *Aa*SHC into a eukaryotic OSCs. Here, we describe the incubation experiments of **1**, **4** and **7** with two double mutants (G600F/F601G and F601G/P602F) to generate detailed information as to why the ΔG600 variant only accepts (3*S*)-**4**. The ΔG600 variant generated mono- and tricyclic products, but did not furnish further cyclized products, such as tetra- and pentacyclic scaffolds. We attempted to generate tetra- and pentacyclic products with the 3β-hydroxyl group and/or develop the potential of the *Aa*SHC to create unnatural compounds. As shown in Fig. S1, the Trp169 residue is in proximity to the C/D-ring formation sites, while the Gly600 and Phe601 residues are located primarily around the D/E-ring sites in the reaction cavity.

Here, we report that rare and novel triterpene scaffolds, which possess the 7-oxabicyclo[2.2.1]heptane ring and the novel 6,6,6,6,6-fused pentacyclic scaffold (named neogammacerane scaffold), are generated by the enzymatic reactions of **4** and **7** with mutated SHCs (W169L and F601G/P602F) and discuss the formation mechanisms of these new products from **4** and **7**.

View Article Online
DOI: 10.1039/C8OB02009D

Results

We constructed 6 site-specific mutants: W169V, W169L, W169F, W169Y, G600F/F601G and F601G/P602F. Primers used for the preparation of these mutants are shown in Table S1.The preparation of cell-free extracts and incubation conditions are described in the Experimental section. Incubation experiments for **1** with the W169V and W169L variants gave no enzymatic product (Fig. S2), although Phe and Tyr mutants with aromatic π-electrons afforded 17-*epi*-dammara-20(21),24-diene.[9] The two double mutants, G600F/F601G and F601G/P602F, also generated no enzymatic product in an incubation reaction with **1** (Fig. S12.2), as we reported in the enzymatic reaction of **1** with the ΔG600 mutant. [10] However, the enzymatic reactions of oxidosqualene (**4** and **7**) with all variants tested here yielded products in good conversion, as seen in the Figs. S3 and S4 and Figs. S12.2 and S12.3.

Isolation of products 10─16 from W169X variants.

Fig. S3 shows the product distributions obtained from the incubation of the racemic mixture of **4** and **7** with native SHC and the W169X variants. Peaks, **5**, **6**, **8** and **9**, were confirmed to be 3β-hydroxyhopene, 3β-hydroxyhopanol, 3α-hydroxyhopene and 3α-hydroxyhopanol, respectively, as isolated in our previous paper.[11] The new prominent peaks **10─16** were isolated to determine the structures. A racemic mixture of **4** and **7** (100 mg) was incubated with cell-free homogenates (200 mL) prepared from a 2.3 L-culture of the W169L variant at 55°C (pH 6.0) for 20 h. After heating the reaction mixture at 70-80°C for 20 min after adding 15% KOH/MeOH (100 mL), lipophilic materials were extracted with hexane. The hexane extract was subjected to SiO_2 column chromatography and eluted with hexane/EtOAc (100:0─100:2) to afford approximately 8 fractions, one of which contained almost pure product **12** (18.2 mg). Separation of other products was unsuccessful. The fractions containing product **13** were again applied to SiO_2 column chromatography (hexane/EtOAc=100:0─100:5), followed by normal phase HPLC (hexane/2-PrOH=100:0.05─100:0.15), yielding pure **13** (1.2 mg). The fractions including products **10** and **11** were combined and subjected to normal phase HPLC (hexane/2-PrOH=100:0.05) to afford **10** (1.1 mg) and **11** (4.2 mg) in a pure state. The isolation yields did not correspond to the production yields (see the GC profile in Fig. S3), as there were still many fractions that required further purification. In a similar way, incubation of the racemic mixture of **4** and **7** (100 mg) with cell-free homogenates (200 mL) from the W169V mutants was performed. SiO_2 column chromatography (hexane/EtOAc=100:0.1─100:6) afforded two fractions: one contained a mixture of **14**, **16** and **5** and the other fraction was a mixture of **15** and **8**. Compounds

View Article Online
DOI: 10.1039/C8OB02009D

5 and **8** were insoluble in MeOH. By washing the two fractions with MeOH, **5** and **8** were removed by filtration. Products **14** and **16** were purified via normal-phase HPLC (hexane/2-PrOH=100:0.1) and pure **15** was also isolated using the same HPLC conditions.

Structures of products 10—16 from W169X variants

Substrates **4** and **7** possess 8 vinylic Me groups, which usually appear at δ_H 1.8—1.5. The ^{1}H-NMR spectrum of product **10** (600 MHz, $CDCl_3$、Fig. S5.2) indicated the presence of 5 vinylic Me groups, which appeared at δ_H 1.68 (s, 3H) and 1.60 (s, 12H) and the presence of 3 aliphatic Me groups (δ_H 1.06, s, 3H, Me-24; 0.90, s, 3H, Me-25; and 1.37 (s, 3H, Me-26). These chemical shifts imply that **10** is a monocyclic product, as shown in the monocyclic structure of product **12** (Fig. 1). However, two oxygenated carbon signals were observed at δ_C 86.26 (d, C-3) and 88.08 (s, C-6) in the ^{13}C-NMR spectrum (150 MHz, $CDCl_3$), suggesting that **10** has an oxygen-bridged monocyclic structure that differs from the typical monocyclic skeleton (like **12**), which possesses one oxygenated carbon. H-3 (δ_H 3.77, d, J=5.0 Hz, 1H) was correlated with C-3 in the HMQC spectrum (Fig. S5.7), and the following HMBC cross peaks were clearly found: H-3/C-6, H-3/C-4 (δ_C 41.76, s), H-3/C-5 (δ_C 56.27, d), Me-26/C-6, Me-26/C-1 (δ_C 28.75, t), Me-26/C-5, Me-24/C-5 and Me-25/C-5. Furthermore, the COSY and TOCSY spectra showed that the ^{1}H-^{1}H spin coupling network was observed between the following three protons: H-3/H-2/H-1. These NMR data unambiguously indicate that **10** has a 7-oxabicyclo[2.2.1]heptane moiety, as shown in Fig. S5.9. The ^{1}H-NMR spectrum of product **11** (600 MHz, $CDCl_3$、Fig. S6.2) also indicated the presence of 5 olefinic Me groups (δ_H 1.68, s, 3H) and 1.60 (s, 12H), the δ_H of which were identical to those of product **10**, but the δ_H of three aliphatic Me groups (δ_H 1.04, s, 3H, Me-24; δ_H 1.01, s, 3H, Me-25; and δ_H 1.32, s, 3H, Me-26) were somewhat different from those of **10**. Furthermore, two oxygenated carbons (δ_C 86.06, d, C-3; δ_C 86.71, s, C-6) were also found in the ^{13}C-NMR spectrum of **11** (150 MHz, $CDCl_3$), and the δ_C of Me-26 was different between **11** and **10**. These NMR data indicate that the side chains of **11** and **10** were identical, but the stereochemistry of monocyclic ring are different, thus **11** and **10** are diastereomers. Detailed NMR analyses, including 2D NMR (Fig. S6.9) for product **11** demonstrated that product **11** also had a 7-oxabicyclo[2.2.1]heptane moiety, which is further supported by the same EIMS spectra between **10** and **11** (compare Fig. S5.1 with S6.1).

Products **10** and **11** are diastereomers because the complete separation of the two compounds was successfully achieved. The *R*- and *S*-stereochemistry at the three

View Article Online
DOI: 10.1039/C8OB02009D

stereocenters were determined as follows. Authentic (3*S*)-**4** was supplied from cultures of *Saccharomyces cerevisiae*, which produces only (3*S*)-**4**, but not (3*R*)-**7**,[2f, 2j] and **4** was isolated via SiO_2 column chromatography using hexane/EtOAc (100:0.5─100:20). The racemic mixture of **4** and **7** was synthesized by treatment of **1** with *m*PCBA, [6i,j] and (3*R*)-**7** was successfully obtained by separating the racemic mixture of **4** and **7** with a CHIRAL PACK IA (Daicel Industries LTD) using hexane/THF (100:0.05). Fig. S4 shows gas chromatograms (GC) of hexane extracts obtained by incubating the racemic mixture of **4** and **7** (A), pure **7** (B) and pure **4** (C) with the W169L variant. Fig. S4C indicates that products **11**, **12** and **13** were produced from **4**, but no detectable amount of substrate **4** was found in the GC trace, indicating that **4** was almost fully converted into products. By contrast, only product **10** was generated from **7** (Fig. S4B), but the yield was significantly low. Therefore, the stereocenters at C-3 of products **10** and **11** must have *R*- and *S*-configurations, respectively. Careful comparison of the NOESY spectra (600 MHz, $CDCl_3$) of **10** and **11** indicated that **10** had a distinct NOE between Me-25 and H-2 (Fig. S5.6.2), but **11** did not exhibit a clear NOE between Me-25 and H-2 (Fig. S6.6.2), even though a strong NOE was observed between Me-24 and H-2 (Fig. S5.6.2). This clear difference in NOE between **10** and **11** coincides with the conformational structures illustrated in Fig. S5.10 and Fig. S6.10. Consequently, the complete structures of products **10** and **11**, including the stereochemistry, are depicted in Fig. 1.

This bridged bicyclic ether (7-oxabicyclo[2.2.1]heptane) moiety of **11** was reported to be generated by the enzymatic reactions of (3*S*)-11-fluoro- and (3*S*)-14-fluorooxidosqualenes with native SHC (Fig. S6.11).[12] Robustell et al. reported ^{1}H-NMR chemical shifts for fluorinated compounds **A** and **B** (Fig. 1C), but did not report all of the ^{1}H NMR data (especially with respect to the δ_H of H-5). Furthermore, they did not assign the ^{1}H- and ^{13}C-signals of the oxygen-bridged monocycle and the configuration at C-5. [12] Through detailed NOESY analyses of **10** and **11**, we unambiguously propose the *S*-stereochemistry at C-5; strong NOEs of H-5/ Me-24 and H-5/Me-26 for **10** (Fig. S5.6.2) and distinct NOEs of H-5/H-1 and H-5/H-2 for **11** (Fig. S6.6.2). Notably, compounds **A** and **B**, which have the same oxygen-bridged monocycle as that of **11**, are produced from fluoro-derivatives of **4** (not genuine **4**). Fluorine is a strong electron-withdrawing atom, and thus, the π-electron density on the fluorine-substituted double bond significantly decreased, leading to the termination of the polycyclization reaction at the mono- (**A**, **B**) and bicyclic stages (**C**, **D**) without completion of the polycyclization reaction. We further generated a new product, **10**, in addition to **11**, although the production yield of **10** was relatively lower than that of **11** (see Fig. S3 and

View Article Online
DOI: 10.1039/C8OB02009D

S4). Moreover, compound **11** from (3*S*)-**4** has not been found in nature. This is the first report of conversion of a genuine oxidosqualene into the oxygen-bridged monocyclic triterpenes **10** and **11** by an enzymatic reaction of triterpene cyclases (including prokaryotic SHCs and eukaryotic OSCs). The variant W169L also generated **12**, which was first isolated from *Achillea odorata* and named achilleol A [13] and was prepared from site-directed variants from OSCs[14a, b] and SHCs.[10, 14c] The MS and NMR data are given in Fig. S7. We also isolated product **13** from the enzymatic reaction of **4** with the A306F variant.[7] The MS and NMR data are shown in Fig. S8.

The ^{1}H-NMR spectrum of product **14** (600 MH, $CDCl_3$, Fig. S9.2) indicated four aliphatic Me groups, δ_H 0.956 (s, 3H, Me-23); 0.779 (s, 3H, Me-24); 0.852 (s, 3H, Me-25); 0.975 (s, 3H, Me-26), and further showed three olefinic Me groups, δ_H 1.595 (s, 3H, Me-28); 1.681 (s, 3H, Me-29); and 1.600 (s, 3H, Me-30). One of the 8 Me groups involved in **4** and **7** was missing in the ^{1}H-NMR spectrum; however, in turn, methylidene protons (δ_H 4.59, s, 1H, H-27; and 4.87, s, 1H, H-27) were observed, and the methylidene group was further confirmed by the DEPT pulse sequences (δ_C 108.8, t, C-14; 154.5, s, C-14). Detailed HMBC analyses (Fig. S9.4) indicated that **14** was composed of a 6,6,5-fused tricyclic system, as shown in Fig. 1. A strong NOE was observed between Me-26 and H-13 (δ_H 2.12, m, 1H), indicating the β-orientation for H-13. H-3 (δ_H 3.20, 1H) exhibited a double-doublet splitting pattern (*J*=11.0 and 5.2 Hz), and thus, the OH function at C-3 is β-oriented. Consequently, the whole structure is as shown in Fig. 1.

The ^{1}H-NMR spectrum of product **15** (600 MHz, $CDCl_3$, Fig. S10.2) showed two olefinic Me groups (δ_H 1.68, s, 3H, Me-26; 1.60, s, 3H, Me-27), which were correlated with H-24 (δ_H 5.11, bt, *J*=6.8 Hz) in the COSY spectrum, indicating that the isopropylidene group is present in **15**. Methylidene protons (H-21) were also found at the following chemical shifts: δ_H 4.88, s, 1H and δ_H 4.91, s, 1H. Five aliphatic Me groups were observed as follows: δ_H 0.935 (s, 3H, Me-28); δ_H 0.834 (s, 3H, Me-29); δ_H 0.845 (s, 3H, Me-19); δ_H 0.942 (s, 3H, Me-18); and δ_H 0.901 (s, 3H, Me-30). The detailed HMBC analyses (Fig. S10.4) demonstrated that **15** consisted of a 6,6,6,5-fused tetracyclic skeleton. Definitive NOEs were found between H-13 (δ_H 2.02, m, 1H) and H-17 (δ_H 2.63, m, 1H), between Me-30 and H-9 (δ_H 1.43, m, 1H) and between Me-18 and H-13, demonstrating the stereochemistry of 13β-H, 17β-H and 30α-Me. The H-3 proton (δ_H 3.41, 1H) gave a broad singlet, indicating that the OH group is arranged in α-disposition. The whole structure of **15** is as shown in Fig. 1.

The structural determination of **16** was conducted in a similar manner to the method for **15**. Detailed 2D analyses were as shown in Fig. S11.4. The stereochemistry in the

View Article Online
DOI: 10.1039/C8OB02009D

tetracyclic skeleton was identical to that of **15**, but, contrary to **15**, the OH group at C-3 was determined to be β-oriented (δ_H 3.13, dd, *J*=10.9, 5.2 Hz, 1H), and thus, the structure of **16** is as shown in Fig. 1.

Isolation of products 16, 20 and 22─24 from the F601G/P602F mutant.

Cell-free homogenates (750 mL) were prepared from a 30 L culture of the double mutant F601G/P602F. A mixture of **4** and **7** (300 mg) was incubated with the homogenates, prepared as above, at 45°C (pH 6.0) for 20 h. After terminating the enzymatic reaction with 15% KOH/MeOH, lipophilic materials were extracted with hexane. Triton X-100 was removed using a short SiO_2 column and eluted with hexane/EtOAc (100:30). SiO_2 column chromatography (hexane/EtOAc=100:0─100:20) afforded **20** (7.8 mg) in an almost pure state. A mixture of **22**, **23**, **16** and **24** was loaded on a SiO_2 column impregnated with 5% $AgNO_3$, followed by elution with hexane/EtOAc=100:0─100:5 to afford **16** (8.0 mg) and **22** (5.6 mg) in an almost pure state. A fraction containing a mixture of **22**, **23** and **24** was acetylated with Ac_2O/Py, and the acetate mixture was subjected to 5% $AgNO_3$-SiO_2 column chromatography and eluted with hexane/EtOAc (100:0─100:5) to yield the acetates of **23** (1.8 mg) and **24** (0.7 mg) that contained some purities, but the purities were good enough for structural determination (see Figs. S14.9 and S15.9). Compounds **12** and **17─21** were produced from the ΔG600 and double mutant G600F/F601G variants. Previously, we reported that compounds **12** and **17─21** were produced by incubating the racemic mixture of **4** and **7** with the ΔG600 variant.[10] Product **22** was also isolated from the enzymatic reaction of a mixture of **4** and **7** with the A306T[7] and A306V variants.[7] The NMR data are shown in Fig. S13. Products **23** and **24** are new metabolites that have not been reported before.

Structures of products 16, 20 and 22─24 from the F601G/P602F mutant.

In the ^{1}H NMR spectrum of **23**-acetate (600 MHz, C_6D_6, Fig. S14.2), no olefinic Me protons were found, suggesting that **23** is a fully cyclized product. Further, no olefinic proton was found. Detailed HMBC analyses (Fig. S14.9) indicated that **23** consists of a 6,6,6,6,5-fused pentacyclic skeleton, similar to **2**, but the double bond position was different from that of hop-22(29)-ene **2**. The presence of an isopropyl residue was demonstrated by distinct COSY correlations; the two aliphatic methyl protons of Me-29 (δ_H 1.14, *d*, *J*=6.8 Hz, 3H) and Me-30 (δ_H 1.18, d, *J*=6.8 Hz, 3H) were correlated with allylic proton H-22 (δ_H 2.87, m, 1H). Me-29 and Me-30 both had clear HMBC cross peaks with C-21 (δ_C 136.5, s), and Me-28 (δ_H 1.13, s, 3H) showed a

View Article Online
DOI: 10.1039/C8OB02009D

definitive HMBC correlation with C-17 (δ_C 140.4, s). Thus, the tetra-substituted double bonds are positioned at C-17 and C-21. Further HMBC analyses are shown in Fig. S14.9. The splitting pattern of H-3 (δ_H 4.83, dd, J=11.6 and 4.4 Hz, 1H) and a distinct NOE between H-5 (δ_H 0.811, bd, J=11.6 Hz, 1H) and H-3 (Fig. S14.6) indicates that the 3-OH is β-oriented. Therefore, product **23** was determined to be 3β-hydroxy-neohop-17(21)-ene, as shown in Fig. 2 and Fig. S14.9.

Product **24** had two vinylic Me groups (δ_H 1.79, s, 3H for Me-29; δ_H 1.97, s, 3H for Me-30), suggesting that an isopropylidene residue remains in **24** without participating in the polycyclization reaction. However, this possibility was denied because the DEPT data clearly indicated the presence of a tetra-substituted double bond (δ_C 120.7, s and δ_C 135.7, s; C-21 and C-22). The distinct HMBC correlations between the two vinylic Me signals (Me-29 and Me-30) and the two sp^2 carbons (C-21 and C-22) were observed (Fig. S15.8), and thus, both Me-29 and Me-30 are located at both C-21 and C-22 (Fig. S15.9). Distinct NOEs were observed between Me-27 (δ_H 1.10, s, 3H) and Me-28 (δ_H 0.855, s, 3H) and between H-13 (δ_H 1.54, m, 1H) and H-17 (δ_H 1.84, m, 1H), indicating that Me-27 and Me-28 are α-oriented, whereas H-13 and H-17 are placed in the β-orientation in the D-ring of **24**. A distinct HMBC cross peak was found for Me-28/C-19 (δ_C 39.48, t)/C-17 (δ_C 56.23, d). H-19 (δ_H 1.13, m, 1H and δ_H 1.70, m, 1H), which was assigned by the HMQC spectrum, had a COSY correlation with H-20 (δ_H 2.23, m, 1H and δ_H 2.40, m, 1H). The chemical shifts of H-20 indicate that H-20 is in the allylic position. A clear spin-coupling between Me-30 and H-17 was observed through homoallylic spin coupling (see the COSY and TOCSY spectra, Fig. S15.4 and Fig. S15.5), and a distinct NOE for Me-30/H-16 allowed the assignment of Me-30. In addition, Me-30 showed a HMBC cross peak for C-17. All of the NMR data described above suggests that the E-ring of **24** is composed of a 6-membered ring, but not the 5-membered ring from the hopane skeleton. Product **24** is different from the gammacerane scaffold **E** and tetrahymanol **F** consisting of the 6,6,6,6,6-fused pentacyclic skeleton (see Fig. 2). Thus, to our knowledge, **24** is a novel triterpene (i.e., in a ACS SciFinder search). We propose the name neogammacerane (compound **G**, Fig. 2) for the fundamental carbon framework of **24**, in which the two Me groups are positioned at C-21 and C-22.

Discussion

Formation mechanisms and distribution of products 10─16 by the W169X variants

View Article Online
DOI: 10.1039/C8OB02009D

Scheme 2 shows the formation mechanisms of products **10**—**16** by the W169X mutants. Scheme 2A shows the cyclization mechanisms for the formation of products **10**—**13** by the W169L variant; (3*S*)-**4** folds into a chair conformation to give monocyclic **25** bearing a C-6 cation. Proton elimination from Me-25 of **25** affords **12**. The chair structure of **25** underwent a ring flip to give a boat structure **26**, in which the OH group is rearranged into a flagpole position of **26** from the equatorial position of **25**. The flagpole OH was approximately located at the C-6 cation of **26**, leading to a nucleophilic attack of the OH on the C-6 cation to afford product **11** with (1*R*,2*S*,4*S*)-configurations (IUPAC numbering). (3*R*)-**7** was also converted into the oxygen-bridged monocycle **10**. As described in Scheme 1, (3*R*)-**7** folds into a boat structure to give **27**, where the OH group is arranged at the bowsprit position, and therefore, the OH group of **27** is very far from the C-6 cation. Cation **27** underwent a ring flip to afford the chair structure **28** with α-OH. The highly nucleophilic OH group attacked the C-6 cation, yielding **10** with (1*S*,2*S*,4*R*)-stereochemistry (IUPAC numbering). However, the distance between the OH and C-6 cation of **28** is longer than that in **26**, which may have led to the lower yield of **10** than **11** of approximately 3-fold (see Table 1). The W169L variant also folds **4** into a chair-chair conformation to give bicyclic **29** with the C-8 cation. The deprotonation of H_{ax}-7 generated **13** with a C-7—C-8 double bond.

Scheme 2B depicts the cyclization mechanism of the conversion of **4** and **7** into **14**—**16** by the W169V variant. (3*S*)-**4** is folded in a chair-chair-chair conformation in the reaction cavity of this variant to generate a 6,6,5-fused tricyclic cation **30**, which undergoes proton elimination from Me-27, yielding **14**. The folding of **4** and **7** into a chair-chair-chair-chair conformation and a boat-chair-chair-chair conformation affords the 17-*epi*-damamrenyl cations **31** with 3β-OH and **32** with 3α-OH, respectively. The deprotonation reactions from the Me-21 of **31** and **32** affords **16** and **15**, respectively.

In our preceding paper,[9] we reported that the W169F and W169Y variants afforded tetracyclic 17-*epi*-dammara-20(21),24-diene (3-deoxy-**15** and **16**) from the enzymatic reactions of **1** in addition to **2** and **3**, but the W169V variant was completely inactive to **1**, and thus, we propose that W169 functions to stabilize the tetracyclic intermediary cation via cation-π interactions.[9] By contrast, the reactions of oxidosqualenes (**4** and **7**) with the Phe and Tyr variants only afforded the normal products **5**, **6**, **8** and **9**, but no detectable amount of the premature products with the tetracyclic dammarenyl skeleton was found (Fig. S3). The π-electron densities increase as follows: Phe<Tyr<Trp. Thus, the richest π-electron densities (Trp) are necessary for the full cyclization reaction of **1**,[9] but in the case of **4** and **7**, the highest π-electron densities, such as those for Trp, are not necessary for complete polycyclization. Moreover, squalene **1** underwent no reaction

View Article Online
DOI: 10.1039/C8OB02009D

with the two aliphatic Val and Leu mutants (Fig. S2), but **4** and **7** underwent polycyclization (Fig. S3). The epoxide rings of **4** and **7** have significantly higher nucleophilicity toward the DXDD motif (proton donor) than the double bond of **1**, easily triggering polycyclization of **4** and **7**. Otherwise, the orientation of substrate **4** or **7** inside the reaction cavity may be somewhat different from that of **1**, in which the substrate heads of **4** or **7** may be located near the DXDD motif, but the head of **1** may be not close to the acidic motif.

<Scheme 2 and Table 1, here>

Table 1 shows that the Val mutant generated the prematurely cyclized tri- **14** and tetracycles **15** and **16**. Production of the abortive tetracyclic **14**—**16** indicated that π-electrons are required for complete cyclization to generate pentacyclic products. However, the Leu variant afforded a significantly large amount of monocyclic products (approximately 72 %) and a small amount of bicyclic product (approximately 7%) but did not produce tri- and tetracyclic **14**—**16**. In other words, the product distribution is distinct for the Val and Leu variants, even though there is only a single CH_2 unit difference between the two amino acid residues. For the Val variant, **4** and **7** could be folded into somewhat extended or stretched conformations (chair-chair-chair, chair-chair-chair-chair or boat-chair-chair-chair structures) in the reaction cavity, and thus, the polycyclization reaction terminated at the later reaction stages of tri- and tetracyclic **30**—**32**. As shown in Fig. S3, the Val mutant also produced pentacyclic **5** and **6**. The somewhat large reaction cavity of the Val mutant may have led to the final products **5** and **8**. However, the Leu mutant constrained **4** and **7** in an unextended or curled conformation inside the reaction cavity, in which only mono- and bicyclic conformations (ca 10:1) are organized, and thus, the polycyclization cascade mainly stopped at the earlier reaction stage **25**—**28** (monocyclic intermediates). We surmise that the reaction cavity of the Leu variant, which is responsible for the cyclization reaction of **4** and **7**, is fairly smaller than that of the Val variant. Therefore, the organization of chair or boat structures may be mainly limited to the monocyclic stages. To ascertain this theory, the X-ray crystal structures, which are complexed with **4** and **7**, should be resolved.

Formation mechanisms and distribution of products 22—24 by the G600F/F601G and F601G/P602F double mutants

Fig. S12.3A shows the GC profile, which was obtained from the incubation experiment, of racemic **4** and **7** with the ΔG600 variant. As described previously,

View Article Online
DOI: 10.1039/C8OB02009D

products **12** and **17—21** were found.[10] The G600F/F601G mutant yielded the same products, **12** and **17—21**, as the ΔG600 variant, despite the somewhat different yields of each product reported between the two mutants (Table 2 and compare Figs. S12.3A with Fig. S12.3B). Fig. S12.1 shows the amino acid alignments of the mutated and wild-type *Aa*SHCs. The G600F/F601G variant was constructed by exchanging Gly at 600 for Phe and Phe at 601 for Gly (see Fig. S12.1). The ΔG600 and G600F/F601G mutants both have the same Phe residue at position 600, which may have led to their identical production of mono- **12** and tricycles **17—21**; the large steric bulk of the Phe residue[15] at position 600 may have prevented full cyclization. This theory prompted us to create the double mutant F601G/P602F, which may enable the generation of tetra- and pentacyclic products because the Gly residue, which has the smallest steric size,[15] is substituted at position 601; two Gly residues are aligned at the positions of both 600 and 601of the F601G/P602F mutant (see Fig. S12.1). As shown in Table 2 and Fig. S12.3, the F601G/P602F variant successfully produced tetra- (**16** and **22**, yield: 25%) and pentacyclic products (**23** and **24**, yield 6.5%) in addition to tricycle (**20**, yield:13%). It is to be noted that all of the products possessed a 3β-hydroxyl group, indicating that (3*S*)-**4** only reacted with three mutants, but **1** and **7** were inactive (see Fig. S12.2 and S12.3); thus, the three mutated SHCs are categorized as OSCs. Currently, it is unknown why the three mutants can only accept **4** as the substrate. One possible answer is that these mutations may bring about either looser or tighter packing at the local sites of the reaction cavity, and thus, substrates **1**, **4** and **7** could be arranged in different geometries in the reaction cavities of the three mutants. The epoxide head of **4** may be arranged in proximity to the DXDD motif, but the heads of **1** and **7** may have failed to access the motif. Further studies are necessary, including X-ray crystallographic analyses, to determine the exact altered substrate specificities and product distribution differences between the G600F/F601G and F601G/P602F variants.

Scheme 3 illustrates the formation mechanisms of **5** and **6** from **4** by the native SHC; **1** and **4** fold into a chair-chair-chair conformation to provide the 6,6,5-fused tricyclic (13β-H)-malabaricanyl cation **30**. Further cyclization of **30** into a chair-chair-chair-chair conformation affords the 6,6,6,5-fused tetracyclic 17-*epi*-dammarenyl cation **31**. Subsequently, ring expansion and further cyclization of **31** yields the 6,6,6,6,5-fused pentacyclic hopanyl cation **33**. In our preceding paper,[6j] we revealed that a robust hydrophobic interaction occurs between Me-24 (*E*-Me at C-23 of **4**) and the D/E ring formation sites (marked with a pink crescent shape, the *E*-Me binding site), and this strong interaction enforces the correct positioning of Me-30 (*Z*-Me at C-23 of **4**) at its binding site, where a highly polarized water molecule is located in proximity to the C-

View Article Online
DOI: 10.1039/C8OB02009D

22 cation of **33** (marked with a blue crescent). The polarized water molecule (a strong base) abstracts the proton from the *Z*-Me (Me-29 of **33**) to yield **5** (path *a*). Product **6** is produced via the nucleophilic attack of the polarized water molecule on the C-22 cation of **33** (path *b*). This mechanism was proposed by Wendt et al.[16] The Phe601 and the Gly600 residues are located primarily at the D/E-ring formation site in the reaction cavity (Fig. S1). Consequently, the strong hydrophobic interaction between the *E*-Me (Me-24) and D/E-ring sites is critical to the formation of cation **33** with the 5-membered E-ring, as well as for smooth deprotonation. Perturbation near the C/D/E-ring formation site may prevent the construction of an organized chair conformation at each local site(s) in the reaction cavity, leading to an altered product distribution.

<Scheme 3 here>

Scheme 4 depicts the mechanisms of formation of products **12** and **17—21** by the ΔG600 and the G600F/F601G variants. As described above, the two variants have Phe residues at position 600 (see Fig. S12.1). The Phe residue has significantly larger steric bulk than the Gly residue (van der Waal's volumes (nm^3): 0.00279 for Gly; 0.55298 for Phe).[15] The amino acid residue positioned at 600 is located in close proximity to the C/D-ring formation sites (Fig. S1). The embedded Phe, with a large steric volume at position 600, collides with substrate **4** due to tight packing, as represented by the grey crescent shape. This perturbation at the local site(s) could prevent the organization of the normal all-chair folding conformation. Only the initially cyclized A-ring (cation **34**) could be organized to give **12**, which resulted from deprotonation of Me-25. A chair-chair-chair fold could afford the tricyclic cation **30**, followed by the nucleophilic attack of a water molecule on the C-14 cation of **30** to give the 6/6/5-fused tricycle **20** with 13β-H. The tight packing at the local site(s) could exert a further influence on the folding conformation to afford the chair-chair-boat folding conformation **35**, leading to the 6/6/5-fused tricyclic cation **36** with 13α-H. Deprotonation from Me-27 could afford **19**, and a water attack of the C-14 cation could give **21**. A series of 1,2-shift reactions—H-13 to the C-14 cation, Me-26 to C-13, and then deprotonation of H-9—could give **17**. Consecutive rearrangement reactions—H-13 to C-14 cation, Me-26 to C-13, and H-9 to C-8, followed by the proton elimination of H-11—furnish **18**.

<Scheme 4, here>

Scheme 5 shows the formation mechanisms of products **16** and **22—24** by the F601G/P602F variant. Contrary to variants ΔG600 and G600F/F601G, the local domain at positions 600 and 601($Gly^{600}Gly^{601}$) of the F601G/P602F variant is more loosely

View Article Online
DOI: 10.1039/C8OB02009D

packed than that of wild type because Gly residues are present at positions 600 and 601. As shown in Fig. S1, the domain is located near the D/E-ring site (represented by a purple crescent). The loosely packed domain could afford an unconstrained conformation **37**, near the D/E-ring formation sites (yield 30%) because no or little hydrophobic interaction occurs between the purple crescent and **37**. The dotted black line, depicted in **37**, indicates the long axis along the all-chair folding conformation for the formation of A─E rings inside the reaction cavity of the wild-type SHC (see also Scheme 3). The long axis is far from the purple crescent near the D/E-ring sites. This perturbation halted polycyclization at the tricyclic **30** (13% yield, from a chair-chair-chair conformation) and tetracyclic stages **31** (17% yield, from a chair-chair-chair-chair structure) to yield **20** and **16**, respectively. As described above, **20** was also produced by the variants ΔG600 and G600F/F601G (Scheme 4A), and **16** was generated from the W169V variants (Scheme 2). Furthermore, the folding conformation of **4** was altered into a chair-chair-chair-boat structure (8% yield), yielding the tetracyclic cation **39** with 17α-H, which underwent the hydride shift of H-17 to the C-20 cation, and the deprotonation of H-13 gave product **22**. This conformational change at the D-ring site of **4** would have resulted from the geometry of **38**, which is distinct from that of **37**. The long axis (dotted line) shown in **37** was largely rotated to the right (red line in **38**), where the D-ring site of **4** was near the purple crescent, possibly leading to a collision between them, which would cause a boat structure for the D-ring formation site of **39** (like **30** in Scheme 4). However, **4** was, in part, folded into an all-chair conformation via the small rotation of the dotted line (yield: 6.4%), albeit in a somewhat distorted conformation, to yield the pentacyclic intermediate **40**. Intermediate **40** could be generated from the geometry shown by the red line in **40**, in which somewhat appropriate, but not perfect, distances between the D/E-ring site and the purple crescent could be set up for the hydrophobic interaction to give a chair-chair-chair-chair-chair conformation to form pentacyclic **40**, albeit at a low yield (6.4%). Then, **40** was converted into two different intermediates: **33** (hopanyl cation), consisting of a 6,6,6,6,5-fused pentacyclic skeleton, and **41** consisting of a 6,6,6,6,6-fused pentacyclic skeleton (gammacerane skeleton). The imperfect hydrophobic interaction of the D/E-ring formation site with the *E*-Me (Me-24) led to failure to place the polarized water at the accurate position, and hence, the distance between the Z-Me (Me-30 of **4**, i.e., Me-29 of **33**) and the polarized water molecule is large. Thus, the deprotonation reaction from the Me-29 of **33** and a water attack on the C-22 cation did not occur; instead, the hydride shift of H-21 to the C-22 cation, followed by the proton elimination of H-17, afforded the neohopane scaffold **23**. This imperfect hydrophobic interaction could

View Article Online
DOI: 10.1039/C8OB02009D

further afford **41** with 6-membered E-ring, because of less steric strain of 6-membered chair structure than that of 5-membered structure **33**. A polarized water molecule is also far from the C-21 cation of **41**, and thus, the nucleophilic attack of a water molecule at the C-21 cation did not occur. Therefore, tetrahymanol scaffold (compound **H**: 3β-hydroxytetrahymanol) was not produced. Instead, the migration of Me-29 of cation **41** to the C-21 cation, followed by the elimination of H-21, resulted in the production of **24**. An alternative biosynthetic pathway to form **41** may be possible (green arrow): a ring expansion of **33** to **41** through stabilization of the secondary C-21 cation of **41** via cation-π interactions with aromatic amino acids;[17a] for example, Phe602 from the F601G/P602F variant or the Phe605 residue of the native SHC,[17a,b] which may possibly be near the C-21 cation of **41**. To confirm the cyclization mechanisms described above, the X-ray crystallographic analyses of the ΔG600, G600F/F601G and F601G/P602G mutants complexed with **4** will be necessary.

<Scheme 5 and Table 2, here>

In conclusion, we present the first enzymatic syntheses of triterpenes **10** and **11** bearing the 7-oxabicyclo[2.2.1]heptane moiety and novel neogammacerane triterpene scaffold **24** by employing mutated SHCs. The oxygen-bridged monocyclic skeleton has been reported via chemical syntheses, such as those found in the cyclization of 14,15-oxidogeranylgeranyl acetate with $SnCl_4$[18.] and in the cyclization of 2,3-oxidosqualene with Ph_4BF_4/HEIP.[19.] However, these fundamental scaffolds, **10**, **11** and **24**, have not been identified in nature. These new triterpenes were only produced from the enzymatic cyclization reactions of 2,3-oxidosqualene, but not from squalene cyclization. The Val mutant accepts almost equal amounts of **4** and **7** to afford 3β- and 3α-hydroxy-cyclic triterpenes, but the W169L mutant has significantly high selectivity for (3*S*)-**4** (73 %), while (3*R*)-**7** was used at a very small amount (6 %). The double mutants G600F/F601G and F601G/P602F also showed high substrate selectivity; only (3*S*)-**4** was recognized, yielding 3β-hydroxy-derivatives of cyclic triterpene scaffolds. SHC has excellent potential for creating new terpenoid compounds[2b, 2i, 4,5] and is also a promising tool for generating non-terpenoid scaffolds by selecting substrates and employing further mutated SHCs.

Experimental

General analytical methods: NMR spectra were recorded in $CDCl_3$ on a Bruker DMX 600 or DPX 400 spectrometer, and the chemical shifts (δ) are given in ppm relative to

View Article Online
DOI: 10.1039/C8OB02009D

the residual solvent peak δ_H 7.26 and δ_C 77.0 as the internal reference for ^{1}H- and ^{13}C-NMR spectra, respectively. In the case of C_6D_6 solutions, the chemical shifts are given in ppm relative to the solvent peak (δ_H 7.28; δ_C 128.0). The coupling constants *J* are given in Hz. GC analyses were performed on a Shimadzu GC-8A or GC2014 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 details are shown in Table S1; primers used for mutagenesis and protocols are described.

Incubation conditions and GC analysis

The SHC described here is from *A. acidocaldarius*.[11, 20] The standard culture of the cloned *Escherichia coli* and incubation conditions were performed according to our published protocols.[11, 20] Cell-free homogenates were used as the enzyme sources. The cell-free extracts were prepared as follows. The *E. coli* BL 21 strains encoding the mutated SHCs, W169X, G600F/F601G and F601G/P602F, were cultured in 100 mL of a LB medium. To the cultured cells, a buffer solution (Tris, pH 8.0, 10 mL) containing 1% Triton X-100 was added, and suspensions were subjected to ultrasonication and centrifugation. The homogenates were used as the enzymatic sources. To 1.0 mg of **1** or a racemic mixture of **4** and **7** emulsified in a distilled water containing Triton X-100 (20 mg), 2 mL of the cell-free homogenates (equivalent to 40 μg of the pure SHC) and 0.5 M citrate buffer (pH 6.0, 3 mL) were added and incubated at 55°C for 20 h. Then, deactivation of the enzyme activities was achieved by adding 15% KOH/MeOH and heating at 70°C for 45 min. The lipophilic materials from the reaction mixture were extracted with hexane. The Triton X-100 included in the hexane extract was removed by a short SiO_2 column by eluting with hexane/EtOAc (100:20), followed by GC analyses, shown in Figs. S2─S4 for W169X variants and in Figs. S12.2─S12.3 for ΔG600, G600F/F9601G and F601G/P602F variants. The GC conditions are also described in the supporting information.

Spectroscopic data of enzymatic products

We have isolated products **12**, **13**, **17-21**and **22** from some mutants and reported the NMR data. Compound **12** and **13** were isolated from ΔG600 variant,[10] and from A306F variant,

View Article Online
DOI: 10.1039/C8OB02009D

[6] respectively. These products were identified by the NMR data (see Figs. S7 for **12**, S8 for **13** and S13 for **22**). Products **17-21** were isolated from ΔG600 mutant and these NMR data were reported in the ref. 10. Product **22** was isolated from A306T variant,[7] but we have not isolated products **10**, **11**, **14**, **15**, **16**, **23** and **24** before. The NMR and EIMS data of these products are given below.

Products 10. $[\alpha]_D^{25}$= -5.5 (*c*=0.096, EtOH); ^{1}H NMR (600 MHz,$CDCl_3$): δ=0.901 (s, 3H, Me-25), 1.06 (s, 3H, Me-24), 1.16 (m, 1H, H-1), 1.22 (m, 1H, H-5), 1.24 (m, 1H, H-7), 1.37 (m, 1H, H-7), 1.38 (s, 3H, Me-26), 1.58 (m, 1H, H-2), 1.60 (s, 12H, Me-27, Me-28, Me-29 and Me-30, total 4 x Me), 1.62 (m, 1H, H-1), 1.68 (s, 3H, Me-23), 1.80 (m, 1H, H-2), 1.89 (m, 1H, H-8), 1.98 (m, 1H, H-8; m, 2H, H-15; m, 2H, H-19, total 5H), 2.03 (m, 2H, H-11; m, 2H, H-12, total 4H), 2.06 (m, 2H, H-16; m, 2H, H-20, total 4H), 3.77 (d, *J*=5.0 Hz, 1H, H-3), 5.11 (m, 1H, H-10; m, 1H, H-13; m, m, 1H, H-17; m, 1H, H-21, total 4H); ^{13}C NMR (150 MHz, $CDCl_3$): δ=16.18 (q, C-29), 16.21 (q, C-28), 16.27 (q, C-27), 17.67 (q, C-30), 19.87 (q, C-25), 21.41 (q, C-26), 25.68 (q, C-23; t, C-7), 26.66 (t, C-20), 26.73(t, C-16), 26.77 (t, C-2), 28.24 (t, C-11; t, C-12), 28.75 (t, C-1), 32.44 (q, C-24), 39.62 (t, C-19), 39.72 (t, C-15), 39.75 (t, C-8), 41.76 (s, C-4), 56.27 (d, C-5), 86.26 (d, C-3), 88.08 (s, C-6), 124.2 (d, C-21), 124.3 (d, C-17), 124.4 (d, C-13), 124.7 (d, C-10), 131.2 (s, C-22), 134.9 (s, C-18), 135.2 (s, 2C, C-9 and C-14). Since the chemical sifts are close to each other, assignments of the following carbons may be exchangeable: between C-2, C-16 and C-20, between C-8, C-15 and C-19, between C-10, C-17 and C-21, and between C-9, C-14, C-18 and C-21. MS (EI) see Fig. S15.1. HRMS (EI): *m*/*z* 426 (M^+): calcd. for $C_{30}H_{50}O$: 426.3862; found: 426.3888.

Product 11. $[\alpha]_D^{25}$= -8.2 (*c*=0.445, EtOH); ^{1}H NMR (600 MHz,$CDCl_3$): δ=1.01 (s, 3H, Me-25), 1.04 (s, 3H, Me-24), 1.17 (dd, *J*=8.6,5.8 Hz, 1H, H-5), 1.32 (s, 3H, Me-26), 1.37 (m, 1H, H-7), 1.41 (m, 1H, H-7), 1.43 (m, 1H, H-1), 1.50 (m, 1H, H-1), 1.60 (s, 12H, Me-27, Me-28, Me-29, Me-30, total 4 x Me), 1.68 (m, 1H, H-2; s, 3H, Me-23, total 4H), 1.89 (m, 1H, H-8), 1.91 (m, 1H, H-2), 1.97 (m, 1H, H-8; m, 2H, H-15; m, 2H, H-19, total 5H), 2.02 (m, 2H, H-11; m, 2H, H-12, total 4H), 2.06 (m, 2H, H-16; m, 2H, H-20, total 4H), 3.71 (d, *J*=5.4, 1H, H-3), 5.09 (t, *J*=7.0 Hz, 1H, H-21), 5.12 (m, 3H, H-10, H-13 and H-17); ^{13}C NMR (150 MHz, $CDCl_3$): δ=15.95 (q, C-29), 15.99 (q, C-28), 16.05 (q, C-27), 17.66 (q, C-30), 18.88 (q, C-26), 23.38 (q, C-25), 25.67 (q, C-23), 25.75 (t, C-2), 26.08 (q, C-24), 26.16 (t, C-7), 26.67 (t, C-20), 26.77 (t, C-16), 28.21 (t, C-12), 28.25 (t, C-11), 39.01 (t, C-1), 39.72 (t, C-19), 39.75 (t, C-15), 39.82 (t, C-8), 45.23 (s, C-4), 55.19 (d, C-5), 86.06 (d, C-3), 86.71(s, C-6), 124.2 (d, C-21; d, C-17,

View Article Online
DOI: 10.1039/C8OB02009D

total 2C), 124.4 (d, C-13), 124.5 (s, C-10), 131.2 (s, C-22), 134.9 (s, C-18), 135.2 (s, C-14), 135.4 (s, C-9). Since the chemical sifts are close to each other, assignments of the following carbons may be exchangeable: between C-8, C-15 and C-19, between C-11 and C-12, between C-16 and C-20, between C-27, C-28 and C-29, between C-10, C-13, C-17 and C-21, and between C-14 and C-18. MS (EI) see Fig. S6.1. HRMS (EI): *m*/*z* 426 (M^+): calcd. for $C_{30}H_{50}O$: 426.3862; found: 426.3866.

Product 14. $[\alpha]_D^{25}$= -5.6 (*c*=0.151, EtOH); ^{1}H NMR (600 MHz,$CDCl_3$): δ=0.67 (dd, *J*=12.0, 2.1 Hz, 1H, H-5), 0.779 (s, 3H, Me-24), 0.852 (s, 3H, Me-25), 0.956 (s, 3H, Me-23), 0.975 (s, 3H, Me-26), 1.08 (td, *J*=12.4, 4.1 Hz, 1H, H-1; td, *J*=12.4, 4.1 Hz, 1H, H-7, total 2H), 1.30 (dd, *J*=12.5, 7.7Hz, 1H, H-9), 1.41 (m, 1H, H-6), 1.44 (m, 1H, H-11), 1.51 (m, 1H, H-1), 1.52 (m, 1H, H-11), 1.54 (m, 1H, H-6), 1.55 (m, 1H, H-7), 1.56 (m, 1H, H-12), 1.57 (m, 2H, H-2), 1.59 (s, 3H, Me-28), 1.60 (s, 3H, Me-30), 1.68 (s, 3H, Me-29), 1.87 (m, 1H, H-15), 1.98 (m, 1H, H-12), 2.00 (m, 1H, H-16), 2.02 (m, 1H, H-15), 2.05 (m, 2H, H-20), 2.12 (m, 1H, H-13), 2.15 (m, 1H, H-16), 2.18 (m, 2H, H-19), 3.20 (dd, *J*=11.0, 5.2 Hz, H-3), 4.59 (s, 1H, H-27), 4.87 (s, 1H, H-27), 5.09 (br t, *J*=6.7 Hz, H-21), 5.12 (br t, *J*=6.7 Hz, 1H, H-17); ^{13}C NMR (150 MHz, $CDCl_3$): δ=15.30 (q, C-24), 15.60 (q, C-25), 16.02 (q, C-28), 17.69 (q, C-30), 18.98 (t, C-6), 20.68 (t, C-11), 24.68 (q, C-26), 25.70 (q, C-29), 26.74 (t, C-20), 26.87 (t, C-16), 27.41 (t, C-2), 27.62 (t, C-12), 28.03 (q, C-23), 36.44 (s, C-10), 36.47 (t, C-7), 38.68 (t, C-1), 38.71 (s, C-4), 39.24 (t, C-15), 39.70 (t, C-19), 45.23 (s, C-8), 55.37 (d, C-9), 55.68 (d, C-5), 56.28 (d, C-13), 79.17 (d, C-3), 108.8 (t, C-27), 124.2 (d, C-17), 124.4 (d, C-21), 131.3 (s, C-22), 135.1 (s, C-18), 154.5 (s, C-14). The assignments of C-2 and C-12 and those of C-16 and C-20 may be exchangeable. MS (EI) see Fig. S9.1. HRMS (EI): *m*/*z* 426 (M^+): calcd. for $C_{30}H_{50}O$: 426.3862; found: 426.3830.

Product 15. $[\alpha]_D^{25}$= -10.3 (*c*=0.165, EtOH); ^{1}H NMR (600 MHz,$CDCl_3$): δ=0.834 (s, 3H, Me-29), 0.845 (s, 3H, Me-19), 0.901 (s, 3H, Me-30), 0.935 (s, 3H, Me-28), 0.942 (s, 3H, Me-18), 1.22 (m, 2H, H-12 and H-15), 1.24 (m, 2H, H-5 and H-11), 1.28 (m, 1H, H-1), 1.30 (m, 1H, H-7), 1.41 (m, 2H, H-6), 1.42 (m, 1H, H-1), 1.43 (m, 1H, H-9), 1.55 (m, 1H, H-2), 1.57 (m, 1H, H-15), 1.59 (m, 1H, H-11), 1.60 (br s, 3H, Me-27), 1.63 (m, 1H, H-7), 1.67 (m, 1H, H-12), 1.68 (br s, 3H, Me-26), 1.79 (m, 2H, H-16), 1.96 (m, 1H, H-2), 1.98 (m, 1H, H-22), 2.02 (m, 1H, H-13), 2.08 (m, 1H, H-23), 2.13 (m, 1H, H-22), 2.16 (m, 1H, H-23), 2.63 (m, 1H, H-17), 3.41 (br s, 1H, H-3), 4.88 (s, 1H, H-21), 4.91 (s, 1H, H-21), 5.11 (br t, *J*=6.8 Hz, H-24); ^{13}C NMR (150 MHz, $CDCl_3$): δ=15.82 (q, C-18), 16.04 (q, C-19), 16.92 (q, C-30), 17.70 (q, C-27), 18.27 (t,

View Article Online
DOI: 10.1039/C8OB02009D

C-6), 21.84 (t, C-11), 22.11 (q, C-29), 24.94 (t, C-12), 25.39 (t, C-2), 25.70 (q, C-26), 27.32 (t, C-23), 28.26 (t, C-16), 28.27 (q, C-28), 33.05 (t, C-15), 33.62 (t, C-1), 34.91 (t, C-7), 37.27 (s, C-10), 37.63 (s, C-4), 38.59 (t, C-22), 41.03 (s, C-8), 44.02 (d, C-17), 44.70 (d, C-13), 49.53 (d, C-5), 49.98 (s, C-14), 50.66 (d, C-9), 79.29 (d, C-3), 108.8 (t, C-21), 124.4 (d, C-24), 131.4 (s, C-25), 152.3 (s, C-20). MS (EI) see Fig. S10.1. HRMS (EI): *m*/*z* 426 (M^+): calcd. for $C_{30}H_{50}O$: 426.3862; found: 426.3846.

Product 16. $[\alpha]_D^{25}$= + 14.2 (*c*=0.359, EtOH); ^{1}H NMR (600 MHz,C_6D_6): δ=0.764 (dd, *J*=11.3, 2.0 Hz, 1H, H-5), 0.89 (m, 1H, H-1), 0.921 (s, 3H, Me-19), 0.931 (s, 3H, Me-29), 1.027 (s, 3H, Me-18), 1.09 (s, 3H, Me-30), 1.13 (s, 3H, Me-28), 1.23 (m, 2H, H-1 and H-6), 1.33 (m, 1H, H-15), 1.36 (m, 1H, H-9), 1.37 (m, 1H, H-7), 1.48 (m, 1H, H-12), 1.50 (m, 1H, H-6), 1.58 (m, 1H, H-11), 1.59 (m, 1H, H-1), 1.63 (m, 1H, H-7), 1.65 (m, 1H, H-1), 1.66 (m, 1H, H-15), 1.76 (s, 3H, Me-27), 1.82 (s, 3H, Me-26), 1.88 (m, 1H, H-12),, 1.92 (m, 1H, H-16), 1.98 (m, 1H, H-11), 2.01(m, 1H, H-16), 2.02 (m, 1H, H-13), 2.19 (m, 1H, H-22), 2.35 (m, 1H, H-23), 2.36 (m, 1H, H-22), 2.43 (m, 1H, H-23), 2.75 (m, 1H, H-17), 3.13 (dd, *J*=10.9, 5.2 Hz, 1H, H-3), 5.19 (s, 1H, H-21), 5.20 (s, 1H, H-21), 5.40 (br t, 1H, H-24); ^{13}C NMR (150 MHz, C_6D_6): δ=15.66 (q, C-29), 16.02 (q, C-18), 16.43 (q, C-19), 17.13 (q, C-30), 17.74 (q, C-27), 18.66 (t, C-6), 22.26 (t, C-2), 25.37 (t, C-12), 25.82 (q, C-26), 27.83 (t, C-11), 27.91 (t, C-23), 28.20 (q, C-28), 28.69 (t, C-16), 33.43 (t, C-15), 35.39 (t, C-7), 37.33 (s, C-10), 39.09 (t, C-22), 39.18 (s, C-4), 39.30 (t, C-1), 41.05 (s, C-8), 44.39 (d, C-17), 45.02 (d, C-13), 50.11 (s, C-14), 51.22 (d, C-9), 56.27 (t, C-5), 78.75 (d, C-3), 109.4 (t, C-21), 124.9 (d, C-24), 131.3 (s, C-25), 152.2 (s, C-20). The assignments of C-11 and C-23 may be exchangeable. MS (EI) see Fig. S11.1. HRMS (EI): *m*/*z* 426 (M^+): calcd. for $C_{30}H_{50}O$: 426.3862; found: 426.3845.

Product 23 acetate. $[\alpha]_D^{25}$= + 132 (*c*=0.27, $CHCl_3$); ^{1}H NMR (400 MHz,C_6D_6): δ=0.811 (br d, *J*=11.6 Hz, 1H, H-5), 0.868 (s, 3H, Me-25), 0.88 (m, 1H, H-1), 1.00 (s, 3H, Me-26), 1.03(s, 6H, Me-23 and Me-24), 1.13 (s, 3H, Me-28), 1.14 (d, *J*=6.8 Hz, 3H, Me-29), 1.18 (d, *J*=6.8 Hz, 3H, Me-30), 1.20 (s, 3H, Me-27), 1.27 (m, 1H, H-11), 1.30 (m, 1H, H-9), 1.37 (m, 1H, H-6), 1.40 (m, 1H, H-7), 1.46 (m, 2H, H-15), 1.52 (m, 1H, H-6), 1.54 (m, 1H, H-11), 1.55 (m, 2H, H-7 and H-12), 1.60 (m, 1H, H-19), 1.64 (m, 1H, H-1), 1.65 (m, 1H, H-13), 1.68 (m, 1H, H-12), 1.72 (m, 1H, H-2), 1.86 (m, 1H, H-2), 1.88 (s, 3H, acetyl C*H*$_3$), 1.93 (m, 1H, H-19), 2.18 (m, 1H, H-16), 2.36 (m, 1H, H-20), 2.45 (m, 1H, H-20), 2.52 (m, 1H, H-16), 2.87 (m, 1H, H-22), 4.83 (dd, *J*=11.6, 4.4 Hz, 1H, H-3); ^{13}C NMR (100 MHz, C_6D_6): δ=15.23 (q, C-27), 16.40 (q, C-26), 16.43 (q,

View Article Online
DOI: 10.1039/C8OB02009D

C-25), 16.83 (q, C-24), 18.55 (t, C-6), 19.42 (q, C-28), 20.24 (t, C-16), 20.86 (q, acetyl $\underline{C}H_3$), 21.50 (q, C-29), 21.63 (t, C-11), 22.10 (q, C-30), 24.11 (t, C-12), 24.33 (t, C-2), 26.81 (d, C-22), 27.94 (t, C-20), 28.10 (q, C-23), 32.27 (t, C-15), 33.73 (t, C-7), 37.21 (s, C-10), 37.98 (s, C-4), 38.56 (t, C-1), 41.92 (s, C-14), 42.09 (t, C-19), 42.29 (s, C-8), 49.78 (d, C-13), 50.19 (s, C-18), 50.92 (d, C-9), 55.55 (d, C-5), 80.53 (d, C-3), 136.5 (s, C-21), 140.4 (s, C-17), 169.9 (s, acetyl $\underline{C}O$). The assignments of proton signals Me-29 and Me-30 may be exchangeable. Since the chemical sifts are close to each other, assignments of the following carbons may be exchangeable: between C-2 and C-12, between C-8 and C-14 and between C-25 and C-26. MS (EI) see Fig. S14.1. HRMS (EI): *m*/*z* 468 (M^+): calcd. for $C_{32}H_{52}O_2$: 468.3967; found: 468.3932.

Product 24 acetate. $[\alpha]_D^{25}$= + 24.1 (*c*=0.23, $CHCl_3$) ; ^{1}H NMR (600 MHz,$CDCl_3$): δ=0.81 (br d, *J*=11.6 Hz, H-5), 0.855 (s, 3H, Me-28), 0.88 (m, 1H, H-1), 0.897 (s, 3H, Me-25), 1.037 (s, 3H, Me-23), 1.037 (s, 3H, Me-26), 1.049 (s, 3H, Me-24), 1.104 (s, 3H, Me-27), 1.13 (m, 1H, H-19), 1.27 (m, 1H, H-9), 1.29 (m, 1H, H-11), 1.34 (m, 2H, H-7 and H-15), 1.44 (m, 1H, H-6), 1.48 (m, 1H, H-15), 1.53 (m, 1H, H-11), 1.54 (m, 1H, H-13), 1.55 (m, 2H, H-2 and H-12), 1.57 (m, 1H, H-6), 1.60 (m, 1H, H-1), 1.70 (m, 2H, H-7 and H-19), 1.74 (m, 1H, H-2), 1.78 (m, 1H, H-16), 1.793 (s, 3H, Me-29), 1.84 (m, 1H, H-17), 1.885 (s, 3H, acetyl $C\underline{H_3}$), 1.965 (s, 3H, Me-30), 2.23 (m, 1H, H-20), 2.33 (m, 1H, H-16), 2.40 (m, 1H, H-20), 4.83(dd, *J*=11.8, 4.6 Hz, 1H, H-3); ^{13}C NMR (150 MHz, $CDCl_3$): δ=14.99 (q, C-28), 16.10 (q, C-25), 16.70 (q, C-26), 16.84 (q, C-24), 16.87 (q, C-27), 18.58 (t, C-6), 19.65 (q, C-30), 20.85 (q, acetyl $\underline{C}H_3$), 21.29 (t, C-11), 23.00 (q, C-29), 23.69 (t, C-12), 23.94 (t, C-16), 24.10 (t, C-2), 28.14 (q, C-23), 28.77 (t, C-20), 33.12 (t, C-15), 33.50 (t, C-7), 37.16 (s, C-10), 37.99 (s, C-4), 38.45 (t, C-1), 39.48 (t, C-19), 41.69 (s, C-8), 41.87 (s, C-14), 44.55 (s, C-18), 48.43 (d, C-13), 50.44 (d, C-9), 55.45 (d, C-5), 56.23 (d, C-17), 80.56 (d, C-3), 120.7 (s, C-21), 135.7 (s, C-22), 169.9 (s, acetyl $\underline{C}O$). The assignments of the following carbons may be exchangeable between C-8 and C-14 and between C-2, C-12 and C-16, due to the very close values. The discrimination of C-21 and C-22 are impossible by the various NMR techniques including 2D NMR analyses. MS (EI) see Fig. S15.1. HRMS (EI): *m*/*z* 468 (M^+): calcd. for $C_{32}H_{52}O_2$: 468.3967; found: 468.3932

Conflicts of Interest

There are no conflicts of interest to declare.

Acknowledgements

View Article Online
DOI: 10.1039/C8OB02009D

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: oxidosqualene; enzyme catalysis; hopene; neogammacerane; oxygen-bridged monocycle.

References

1. A. Eschenmoser, L. Ruzicka, O. Jeger, D. Arigoni, *Helv. Chim. Acta* 1955, **38**, 1890–1904.
2. For a review of squalene, oxidosqualene cyclases, see: a) K. U. Wendt, G. E. Schulz, E. J. Corey and 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 and J. N. Johnston, *Chem. Rev.* 2005, **105**, 4730–4756; d) M. J. R. Segura, B. E. Jackson and 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 and T. T. Wang, *Chem. Rec.* 2008, **8**, 302–325; g) W. D. Nes, *Chem. Rev.* 2011, **111**, 6423–6451; h) G. Siedenburg and D. Jendrossek, *Appl. Environ. Microbiol.*, 2011, **77**, 3905-3915; i) P.-O. Syrén, S. Henche, A. Eichler, B. M. Nestl and B. Hauer, *Curr. Opin. Struct. Biol.* 2016, **41**, 73-82; j) T. Hoshino, *Org. Biomol. Chem*., 2017, **15**, 2869-2891.
3. R. Xu, G. C. Fazio, and S. P. T. Matsuda, *Phytochemistry* 2004, **65**, 261–291.
4. a) M. Rohmer, C. Anding, G. Ourisson, *Eur. J. Biochem*., 1980, **112**, 541-547; b) P. Bouvier, Y. Berger, M. Rohmer, G. Ourisson, *Eur. J. Biochem*., 1980, **112**, 549-556; c) R. Rohmer, P. Bouvier, G. Ourisson, *Eur. J. Biochem*., 1980, **112**, 557-560; d) T. Hoshino, S. Nakano, T. Kondo, T. Sato and A. Miyoshi, *Org. Biomol. Chem.*, 2004, **2**, 1456-1470.
5. For example, the Hauer group reported the following papers: a) M. Seitz, P. O. Syrén, M. Seitz, B. M. Nestle, B. Hauer, *ChemBioChem*, 2013, **14**, 436-439; b) S. C. Hammer, A. Marjanovic, J. M. Dominicus, B. M. Nestle, B. Hauer, *Nat. Chem. Biol*., 2015, **11**, 121-126; c) L. C. Kühnel, B. M. Nestl and B. Hauer, *ChemBioChem*, 2017, **18**, 2222-2225; d) S. A. Bastian, S. C. Hammer, N. Kreß, B. M. Nestle, B. Hauer, *ChemCatChem.*, 2017, **9**, 4364-4368.
6. a) T. Hoshino and S. Ohashi, *Org. Lett.*, 2002, **4**, 2553-2556; b) S. Nakano, S. Ohhashi and T. Hoshino, *Org. Biomol. Chem.*, 2004, **2**,2012-2022; c) T. Hoshino, I. Kudo, S. Nakano and S. Ohashi, *Org. Biomol. Chem.*, 2004, **2**, 2650-2657; d) T. Abe and T. Hoshino, *Org. Biomol. Chem.*, 2005, **3**, 3127-3139; e) T. Hoshino, Y.

View Article Online
DOI: 10.1039/C8OB02009D

Yonemura, T. Abe and Y. Sugino, *Org. Biomol. Chem.*, 2007, **5**, 792-801; f) T. Hoshino, Y. Kumai and T. Sato, *Chem. Eur. J.*, 2009, **15**, 2091-2100; g) J. Cheng and T. Hoshino, *Org. Biomol. Chem.*, 2009, **7**, 1689-1699; h) Y. Yonemura, Y. Ohyama and T. Hoshino, *Org. Biomol. Chem*., 2012, **10**, 440-446; i) K. Takahashi, Y. Sasaki and T. Hoshino, *Eur. J. Org. Chem.*,2018, 1477-1490; j) I. Kaneko, Y. Terasawa and T. Hoshino, *Chem. Eur. J*., 2018, **24**, 11139-11157; k) T. Abe and T. Hoshino, *Org. Biomol. Chem*., 2005, **3**, 3127-3139; (l) H. Tanaka, H. Noguchi and I. Abe, *Org. Lett.*, 2004, **6**, 803-806. (m) H. Tanaka, H. Noguchi and I. Abe, *Org. Lett.*, 2005, **7**, 5873-5876. (n) H. Tanaka, N. Noguchi and I. Abe, *Tetrahedron Lett*., 2004, **45**, 3093-3096.

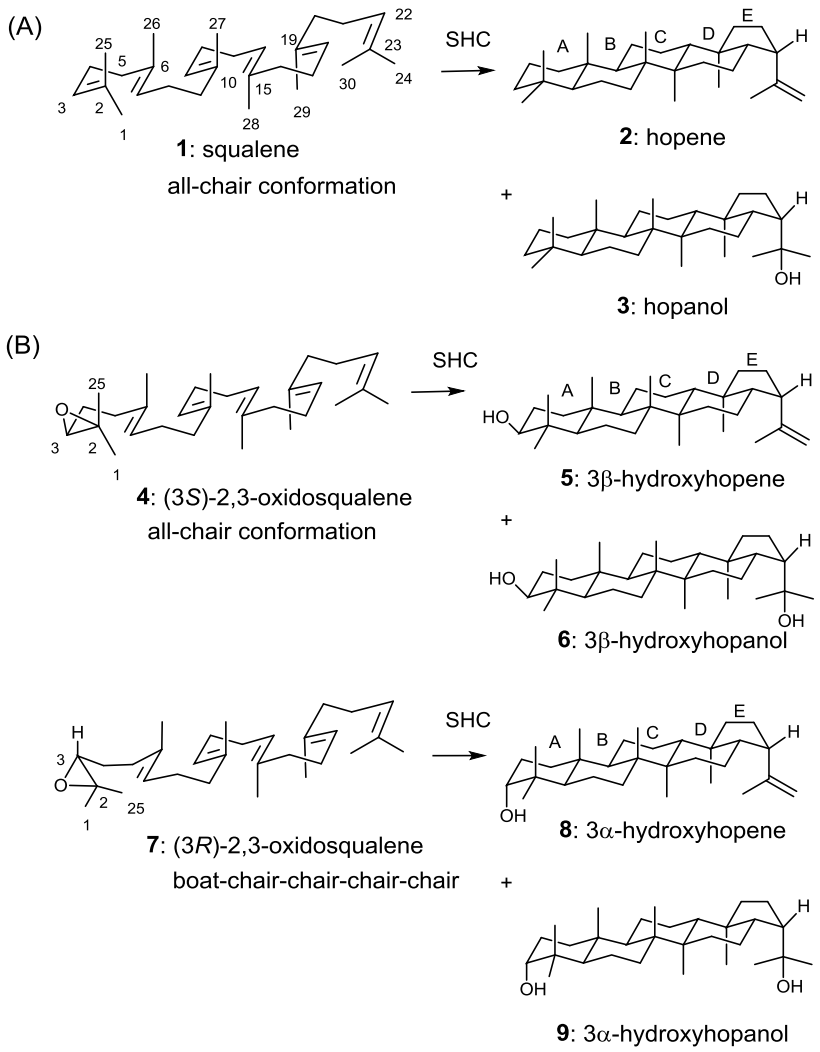
7. N. Ideno, S. Umeyama, T. Watanabe, M. Nakajima, T. Sato and T. Hoshino, *ChemBioChem*, 2018, in press, DOI: 10.1002/cbic.201800281.
8. a) T. Sato and T. Hoshino, *Biosci, Biotechnol. Biochem*., 1999, **63**, 2189-2198; (b) T. Dang, G. D. Prestwich, *Chem. Biol*., 2000, **7**, 643–649.
9. T. Sato, T. Abe and T. Hoshino, *Chem. Commun.* 1998, 2617-2618.
10. T. Hoshino, K. Shimizu and T. Sato, *Angew. Chem. Int. Ed*., 2004, **43**, 6700-6703.
11. a) T. Sato and T. Hoshino, *Biosci. Biotechnol. Biochem*, 1999, **63**, 1171-1180; b) T. Sato and T. Hoshino, *Biosci. Biotechnol. Biochem.*, 1999, **63**, 2189-2198.
12. a) B. Robustell, I. Abe and G. D. Prestwich, *Tetrahedron L*ett., 1998, **39**, 957-960; b) B. Robustell, I. Abe and G. D. Prestwich, *Tetrahedron Lett.*, 1998, **39**, 9385-9388.
13. A. F. Barrero, E. J. Alvarez-Manzaneda and R. Alvarez-Manzaneda, *Tetrahedron Lett.*, 1989, **30**, 3351-3352.
14. a) C-H. Chang, H-Y. Wen, W-S. Shie, C-T. Lu, M-E. Li, Y-T. Liu, W-H. Li and T-K. Wu, *Org. Biomol. Chem*., 2013, **11**, 4214-4219; b) S. P. T. Matsuda, L. B. Darr, E. A. Hart, J. B.R. Herrera, K. E. McCann, M. M. Meyer, J. Pang, H. G. Schepmann, *Org. Lett.*, 2000, **2**, 2261-2263; c) T. Sato and T. Hoshino, *Biosci. Biotechnol. Biochem*., 1999, **63**, 2189-2198.
15. Z-h Lin, H-x Long, Z. Bo, Y-q Wang, , Y-z Wu, *Peptides*, 2008, **29**, 1798-1805.
16. a) K. U. Wendt, K. Poralla, G. E. Schulz, *Science*, 1997, **277**, 1811-1815; b) K. U. Wendt, A. Lenhart, G. E. Schulz, *J. Mol. Biol.*, 1999, **286**, 175-187; c) D. J. Reinert, G. Balliano, G. E. Schulz, *Chem. Biol*., 2004, **6**, 1717-1720.
17. a) 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; b) T. Hoshino, M. Kouda, T. Abe and T. Sato, *Chem. Commun.*, 2000, 1485-1486.
18. E.E. van Tamlen and R. G. Nadeau, *Biorg. Chem*., 1982, **11**, 197-218.

View Article Online
DOI: 10.1039/C8OB02009D

19. Y. Tian, X. Xin, L. Zhang and J. Qu, *Org. Lett.*, 2016, **18**, 268-271.
20. a) T. Sato, T. Hoshino, *Biosci, Biotechnol. Biochem.*, 1999, **63**, 2189-2198; b) T. Sato, S. Sasahara, T. Yamakami, T. Hoshino, *Biosci, Biotechnol. Biochem.*, 2002, **66**, 1660-1670.

View Article Online
DOI: 10.1039/C8OB02009D

Figure and Scheme 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 reaction cavity. (B) Polycyclization reactions of (3*S*)-2,3-oxidosqualene **4** and (3*R*)-2,3-oxidosqualene **7**. Here, **4** folds into an all-prechair structure, similar to the polycyclization reaction of **1**, leading to 3β-hydroxyhopene **5** and 3β-hydroxyhopanol **6**. However, **7** folds into a boat-chair-chair-chair-chair conformation, resulting in the production of 3α-hydroxyhopene **8** and 3α-hydroxyhopanol **9**.

View Article Online
DOI: 10.1039/C8OB02009D

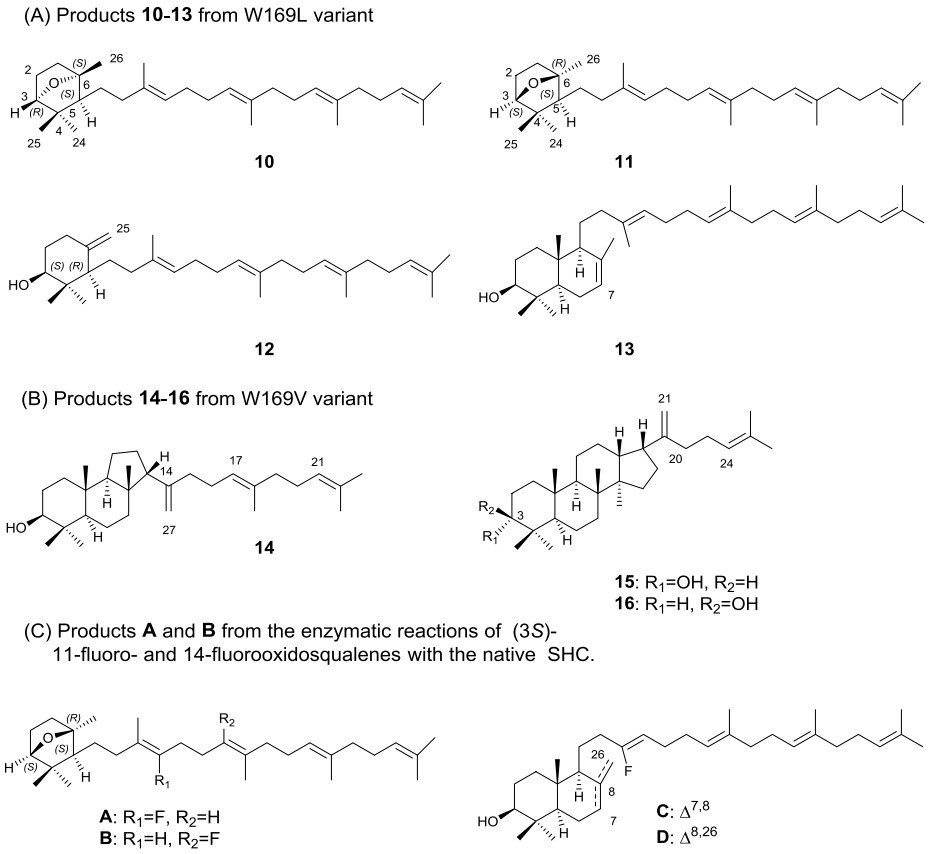


Figure 1. Product structures of **10**—**16**, generated by incubating a mixture of **4** and **7** with W169X variants. The W169L variant afforded **10**—**13** (A), and the W169V variant gave **14**—**16** (B). Compounds **A**—**D** are enzymatic products obtained by incubation of (3*S*)-11-fluoro- and (3*S*)-14-fluorooxidosqualenes with wild-type SHC (C). The structure of the 7-oxabicyclo[2.2.1]heptane moiety of **11** is identical to those of **A** and **B**, including configurations at C-3 and C-6, but their C-5 stereochemistry was not previously determined.[12] Compounds **C** and **D** were reported to be produced only from (3*S*)-14-fluorooxidosqualene.

Compound names:**10**: (1*S*,2*S*,4*R*)-1,3,3-trimethyl-2-((3*E*,7*E*,11*E*)-3,8,12,16-tetramethylheptadeca-3,7,11,15-tetraen-1-yl)-7-oxabicyclo[2.2.1]heptane; **11**: (1*R*,2*S*,4*S*)-1,3,3-trimethyl-2-((3*E*,7*E*,11*E*)-3,8,12,16-tetramethylheptadeca-3,7,11,15-tetraen-1-yl)-7-oxabicyclo[2.2.1]heptane; **12**: achilleol A; **13**: polypoda-7(8), 13,17,21-tetraen-3β-ol (3β-hydroxy-γ-polypodatetraene); **14**: (13β-*H*)-malabarica-14(27), 17, 21-trien-3β-ol; **15**: 17-*epi*-dammara-20(21), 24-dien-3α-ol; **16**: 17-*epi*-dammara-20(21), 24-dien-3β-ol.

View Article Online
DOI: 10.1039/C8OB02009D

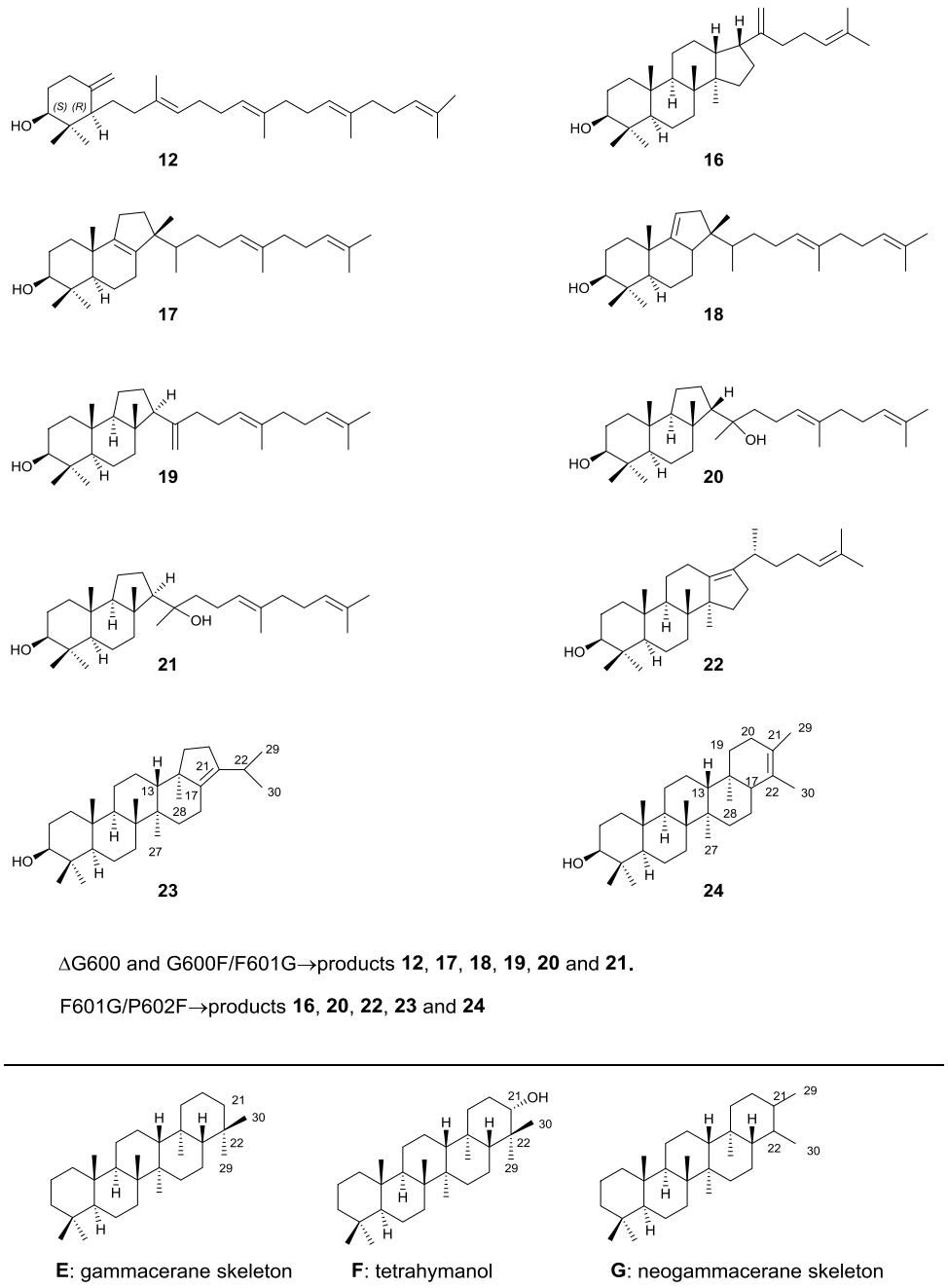


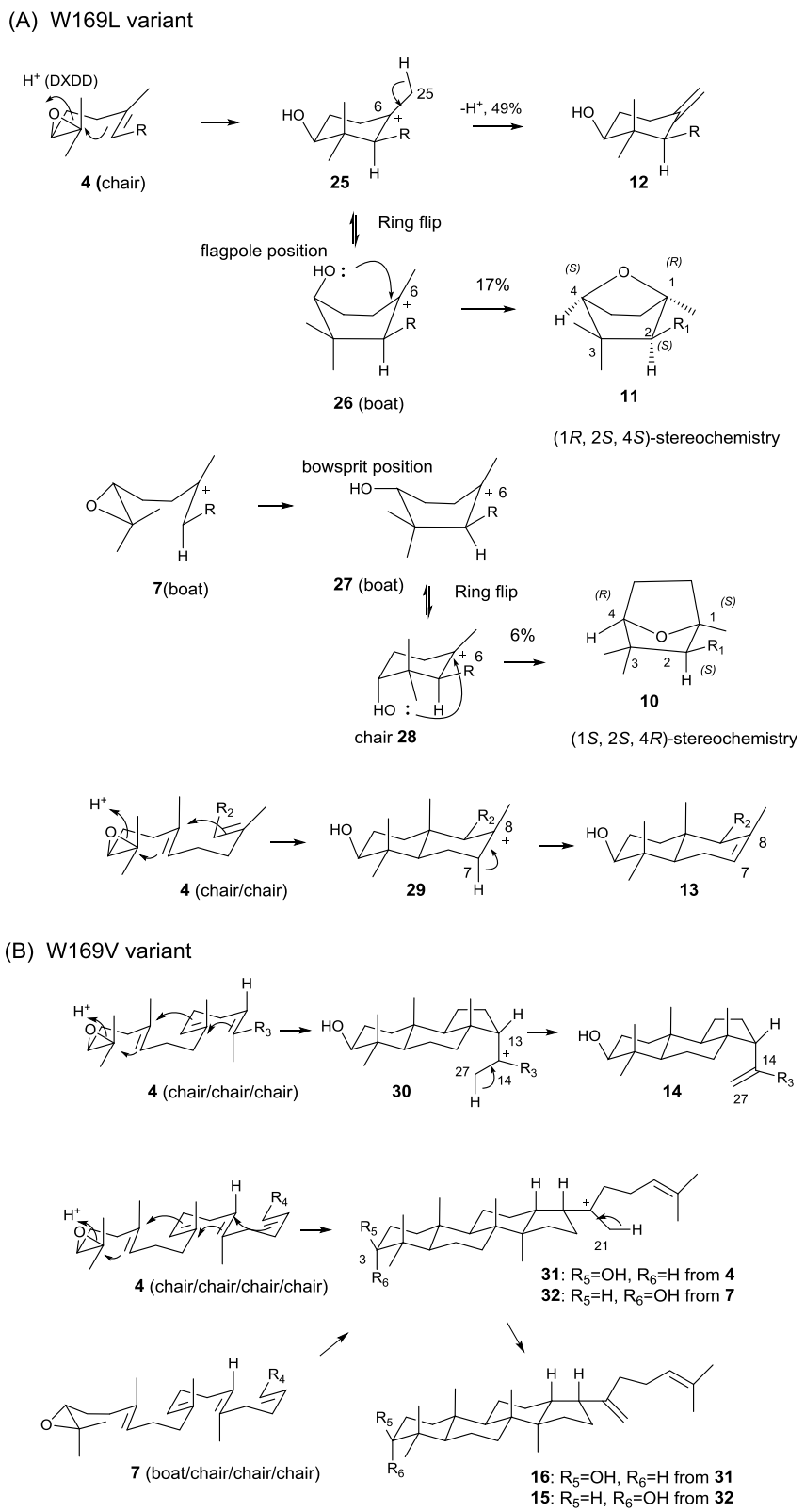
Figure 2. Product structures of **17**—**24**, which are produced from enzymatic reactions of a racemic mixture of **4** and **7** with SHC mutants of ΔG600 and double mutants of G600F/F601G and F601G/P602F. Products **12** and **16**, generated by W169X mutants (see Fig. 1 and Scheme 2), were also furnished from these site-specific mutants. ΔG600

View Article Online
DOI: 10.1039/C8OB02009D

and G600F/F601F variants afforded products **12**, **17**, **18**, **19**, **20** and **21**. F601G/P602Fgenerated products **16**, **20**, **22**, **23** and **24**.

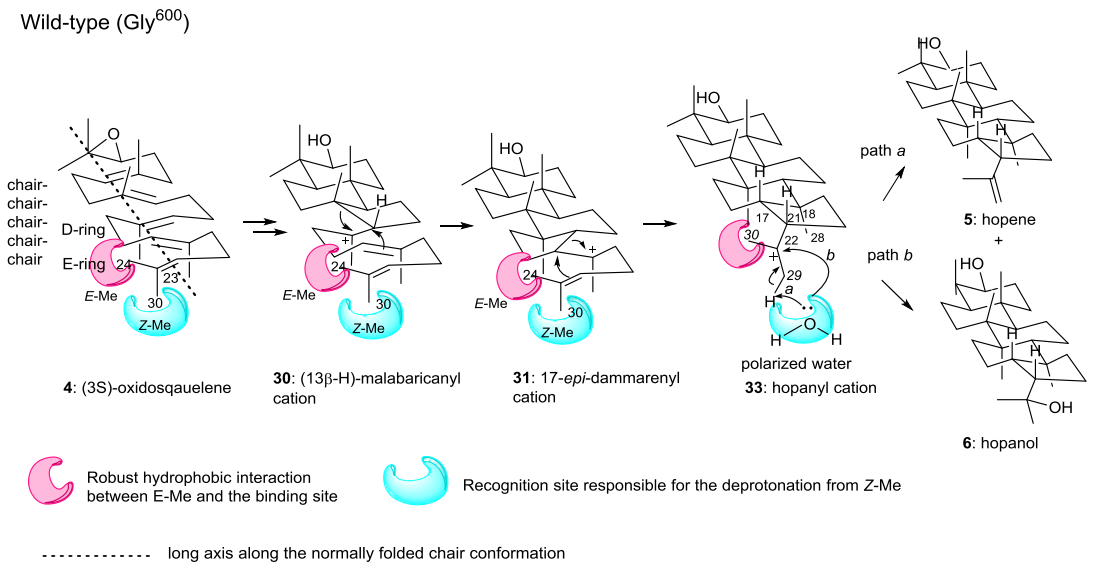
Compound names. **17**: podioda-8(9),17,21-trien-3β-ol; **18**: podioda-9(11),17,21-trien-3β-ol; **19**: (13α-*H*)-malabarica-14(27), 17, 21-trien-3β-ol; **20**: (13β-*H*)-malabarica-17, 21-dien-3β,14-diol; **21**: (13α-*H*)-malabarica-17, 21-dien-3β,14-diol; **22**: (20*R*)-dammara-13(17),24-dien-3β-ol; **23**: hop-17(21)-en-3β-ol; **24**: novel carbocyclic triterpene. Gammacerane skeleton consists of a 6.6.6.6.6-fused pentacycle, as exemplified by compounds **C** and **D**. We name neogammacerane skeleton for **24**, in which Me-29 at C-22 in gammacerane skeleton is migrated to C-21, and thus neogammacera-21(22)-en-3β-ol is proposed for **24**.

View Article Online
DOI: 10.1039/C8OB02009D



Scheme 2. Cyclization mechanisms of **4** and **7** into monocyclic products **10**—**12** and **13** (A) by the W169L variant and into **14**—**16** (B) by the W169V variant.

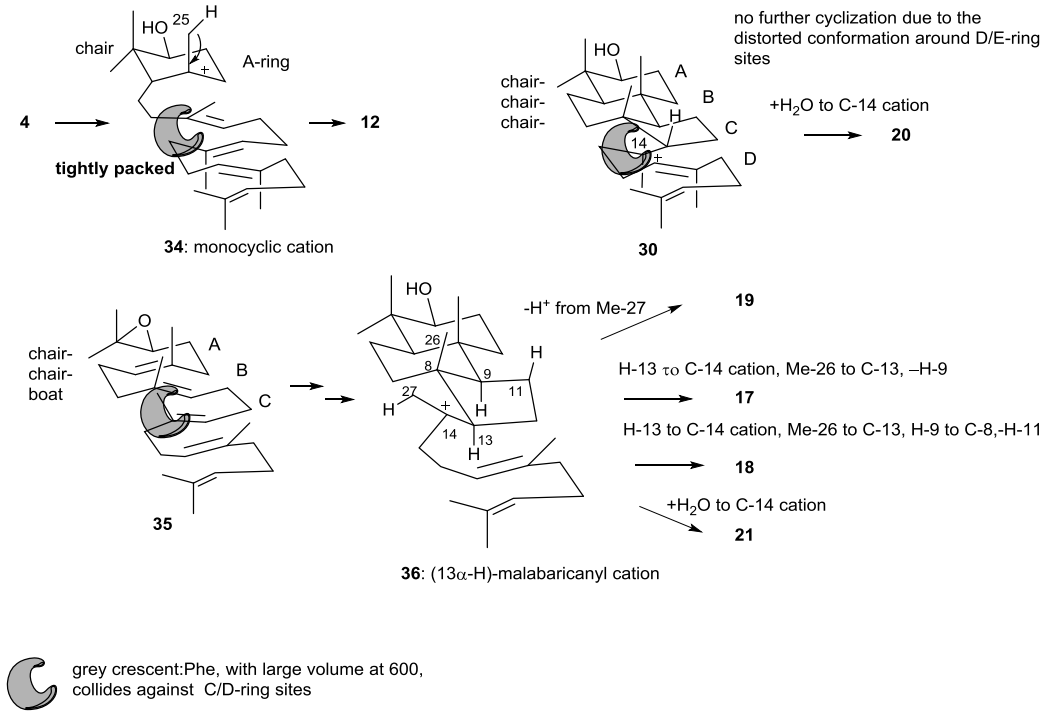
View Article Online
DOI: 10.1039/C8OB02009D



Scheme 3. Cyclization mechanism of (3*S*)-oxidosqualene **4** to form 3β-hydroxyhopene **5** and 3-hydroxyhopanol **6** by native SHC.[6j]

A pink crescent represents a robust hydrophobic interaction between the *E*-Me on the terminal double bond of **4** and the binding site of the reaction cavity. This strong interaction leads to the 5-membered E-ring of hopanyl cation **33** and contributes to the placement of Me-30 (*Z*-Me) of **4** at the correct position of the recognition site (a blue crescent), which plays a crucial role for the deprotonation reaction from *Z*-Me to generate **5** (path *a*) and for the nucleophilic attack of a water molecule to afford **6** (path *b*). The highly ordered hydrogen-bonding networks generate a strongly polarized water molecule, which is in proximity to Z-Me and abstracts the proton of *Z*-Me.[6j,16] This scheme indicates that the *Z*-Me (Me-30) of **4** is led to the Me-29 of cation **33**. The dotted line indicates the long axis along the normally folded conformation (all chair structures).

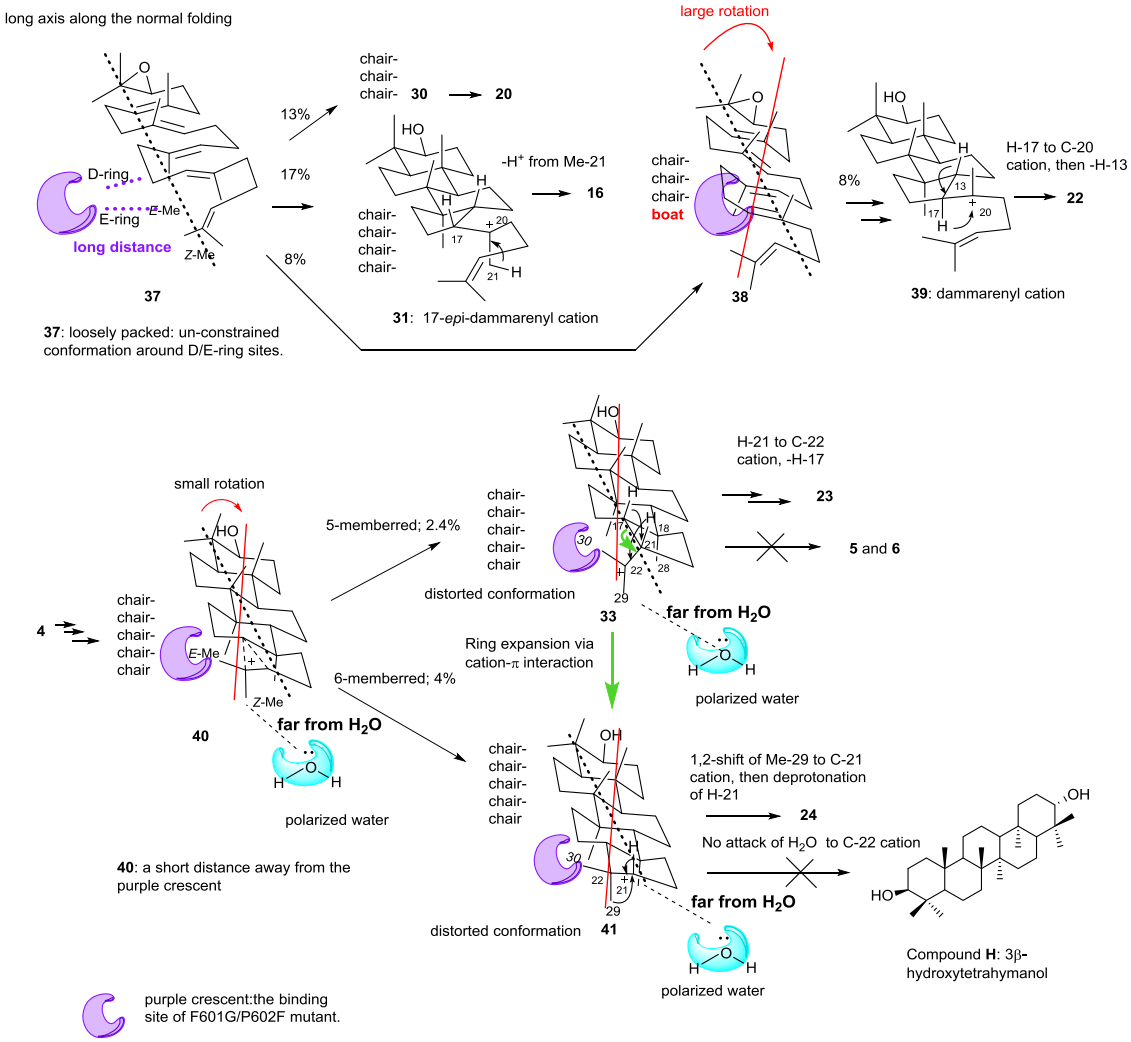
View Article Online
DOI: 10.1039/C8OB02009D



Scheme 4. Formation mechanisms of products **12** and **17**—**21** from (3*S*)-**4**, generated by mutants ΔG600 and G600F/F601G

The two mutants generated similar enzymatic products. G600F/F601G and ΔG600 variants have Phe residues at the same position 600. The Phe residue has a significantly larger steric bulk than Gly residue.[15] The amino acid residue positioned at 600 is located in close proximity to C/D-ring formation sites (Fig. S1), marked with a grey crescent shape. The substituted Phe at position 600 could collide with the C/D-ring sites, leading to the abortive cyclization reactions terminated at the monocyclic cation **34** and tricyclic cation **30**, produced from a chair-chair-chair conformation, and tricyclic cation **36**, generated from chair-chair-boat structure **35**. Further reactions, such as deprotonation, rearrangement reactions and water attack of cations, could afford **12** and **17**—**21**.

View Article Online
DOI: 10.1039/C8OB02009D



Scheme 5. Mechanisms for production of **16**, **20**, and **22**—**24** from (3*S*)-**4** by F601G/P602F variant.

This mutant has Gly residues, which has the smallest steric size, at both 600 and 601 residues, indicating that the local sites at positions 600-601 (marked with a purple crescent) are loosely packed. Thus, little or no hydrophobic interaction exists near D/E-ring formation sites, leading to the formation of unconstrained conformation **37** near the D/E-ring formation sites (total yield: 30%). The purple crescent is away from the long axis (the black dotted line), which is formed by the normal all-chair conformation in the reaction cavity (see also the dotted line in Scheme 3). The wide interspace between the purple crescent and long axis (dotted line) exerted further influence on the folding conformation of **4**; the dotted line shown in **37** was largely rotated into the red line of **38**, leading to boat structure **38** in the D-ring site (8%), which resulted from the close

View Article Online
DOI: 10.1039/C8OB02009D

collision with the purple crescent (like **30** in Scheme 4). A small rotation of the dotted line could afford the all-chair conformation **40**, albeit somewhat distorted. This distorted folding conformation could afford the incorrect positioning of polarized water (blue crescent).

Table 1. Product distribution ratios (%) obtained from reactions of a racemic mixture **4** and **7** in presence of W169X variants.

		Monocycle			Bicycle	Tricycle	Tetracycle		Normal products (pentacycle)			
Compounds	**4** & **7**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**5**	**6**	**8**	**9**
Wild SHC	5.0	—	—	—	—	—	—	—	39.2	6.7	42.4	6.7
W169V	4.1	—	—	—	—	11.0	14.9	21.1	19.2	0.6	27.4	1.7
W169L	20.0	5.7	17.2	48.7	7.2	—	—	—	0.6	—	0.6	—
W169F	5.7	—	—	—	—	—	—	—	41.8	4.3	42.9	5.1
W169Y	4.3	—	—	—	—	—	—	—	44.4	3.4	44.4	3.5

—: Little production or no production, below limit of GC detection. Other products may exist that were not characterized, but amounts were small, if present. (3*S*)-**4** afforded products **11**, **12** and **13** in high yields, but (3*R*)-**7** was converted into only **10** in low yield.

Table 2. Product distribution ratios (%) generated by reactions of oxidosqualenes (**4** and **7**) with variants of ΔG600, G600F/F601G and F601G/P602F. Reported van der Waal's volumes (nm^3) are as follows: 0.00279 for Gly; 0.55298 for Phe; 0.2279 for Pro.[15]

		Monocycle	Tricycle					Tetracycle		Pentacycle	
Compounds	**4** & **7**	**12**	**17**	**18**	**19**	**20**	**21**	**16**	**22**	**23**	**24**
ΔG600- SHC	54.4	16.7	9.9	6.4	4.7	6.1	1.8	—	—	—	—
G600F/F601G	56.0	10.3	7.9	6.7	6.0	13.1	*	—	—	—	—
F601G/P602F	55.3	—	—	—	—	13.2	*	17.0	8.0	2.4	4.1

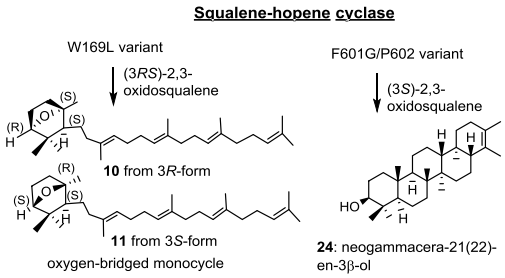
—: Little production or no production, below limit of GC detection. There might be other products that have not been characterized, but production amounts are small, if present.

*: production yield was nearly zero.

Total yield of tetra- and pentacycles generated from F601G/P602F variant was up to 31.5%; that of the tricycle was smaller (ca 13%)

View Article Online
DOI: 10.1039/C8OB02009D

Table of Contents



First enzymatic syntheses of triterpenes bearing a 7-oxabicyclo[2.2.1]heptane moiety and a novel scaffold, named neogammacerane, by the mutated squalene-hopene cyclases.