

Site-directed Mutagenesis Experiments on the Putative Deprotonation Site of Squalene-hopene Cyclase from *Alicyclobacillus acidocaldarius*

Tsutomu SATO, Masanori KOUDA, and Tsutomu HOSHINO[†]

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

Received November 18, 2003; Accepted December 4, 2003

To provide insight into the catalytic mechanism for the final deprotonation reaction of squalene-hopene cyclase (SHC) from *Alicyclobacillus acidocaldarius*, mutagenesis experiments were conducted for the following ten residues: Thr41, Glu45, Glu93, Arg127, Trp133, Gln262, Pro263, Tyr267, Phe434 and Phe437. An X-ray analysis of SHC has revealed that two types of water molecules (“front water” and “back waters”) were involved around the deprotonation site. The results of these mutagenesis experiments allow us to propose the functions of these residues. The two residues of Gln262 and Pro263 probably work to keep away the isopropyl group of the hopanyl cation intermediate from the “front water molecule,” that is, to place the “front water” in a favorable position, leading to the minimal production of by-products, *i.e.*, hopanol and hop-21(22)-ene. The five residues of Thr41, Glu45, Glu93, Arg127 and Trp133, by which the hydrogen-bonded network incorporating the “back waters” is constructed, increase the polarization of the “front water” to facilitate proton elimination from the isopropyl moiety of the hopanyl cation, leading to the normal product, hop-22(29)-ene. The three aromatic residues of Tyr267, Phe434 and Phe437 are likely to play an important role in guiding squalene from the enzyme surface to the reaction cavity (substrate channeling) by the strong affinity of their aromatic residues to the squalene substrate.

Key words: squalene; hopene; terpene cyclase; triterpene

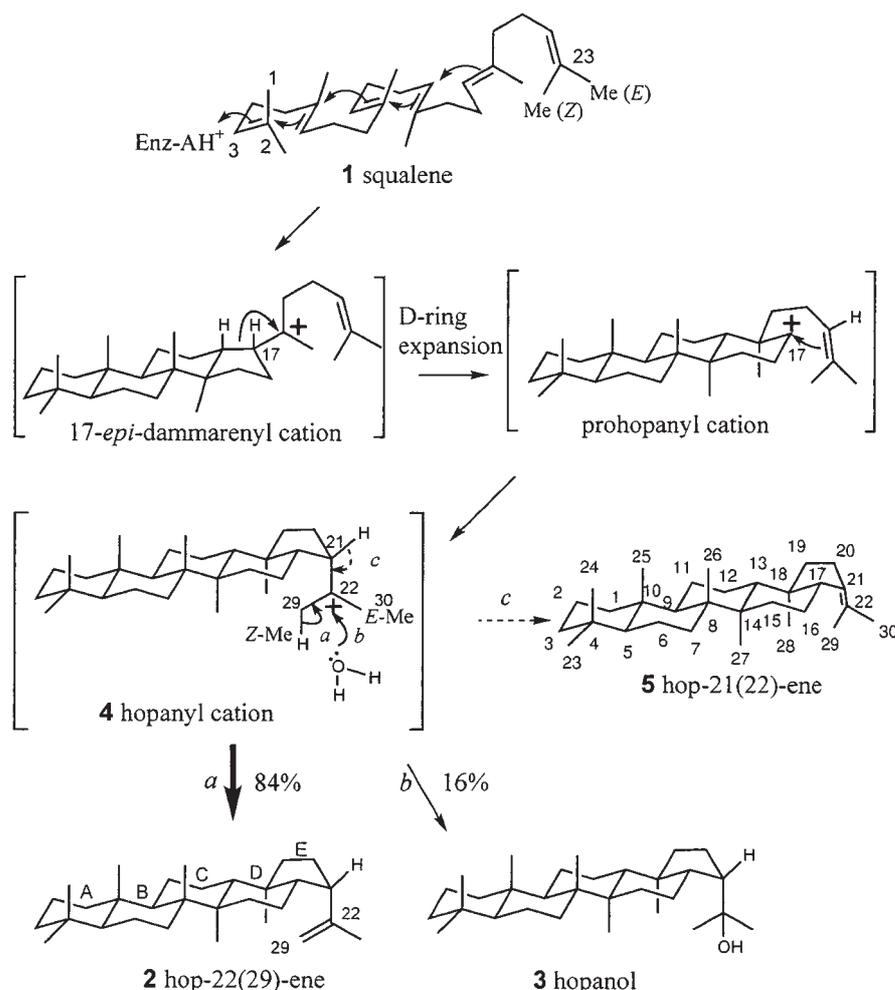
The polycyclization of squalene **1** into pentacyclic hopene **2** (hop-22(29)-ene) and hopanol **3** is one of the most complicated biochemical reactions (Scheme 1)^{1–3} and is catalyzed by squalene-hopene cyclases (SHCs) [EC 5.4.99-] from prokaryotic species. The polycyclization reaction proceeds under precise enzymatic control to form five rings and nine chiral centers. The formation of **2** occurs by a proton elimination reaction from the methyl group at C-29 of hopanyl cation **4** (path *a*), while

that of **3** occurs by the addition of a water molecule to final cation C-22 of **4** (path *b*), as shown in Scheme 1. The ratio of product **2** to **3** is 84:16. Numerous triterpene skeletons are formed by the biocyclization of common substrate **1** or (3*S*)-2,3-oxidosqualene. The diversity of triterpene skeletons found in nature originates from many types of different termination reactions by triterpene cyclases.^{1–5} Tetrahymanol from protozoa is biosynthesized by water attack on the cation at C-21 to the final cation having the 6/6/6/6/6-fused pentacyclic skeleton. As for lanosterol from vertebrates and fungi, the final protosterol cation (a 6/6/6/5-fused tetracyclic ring system) undergoes backbone rearrangement by 1,2-shifts of the hydride and methyl group and a subsequent deprotonation reaction at C-9, leading to the introduction of a double bond at C8–C9 (lanostane numbering). Cycloartenol from plants is formed *via* the protosterol cation in a similar manner to that of lanosterol, but the deprotonation reaction occurs from the methyl group of C-19 after the backbone rearrangement (lanostane numbering), resulting in the formation of a cyclopropane ring.

In the past half a decade, there have been remarkable advances in understanding the catalytic mechanisms by *A. acidocaldarius* SHC.^{1,2} Site-directed mutants and substrate analogs^{1,2} have enabled us to establish that this polycyclization reaction consisted of eight reaction steps, including two ring-expansion processes for the formation of 6-membered C- and D-rings.^{6,7} The polycyclization reaction is triggered by proton attack on the terminal double bond of **1**, the acidic proton being donated by the DXDDTA motif.⁸ The transient carbocations that are generated during the polycyclization cascade are stabilized through cation- π interaction by such aromatic amino acid residues as F365 and F605.^{9,10} The steric bulk size of the amino acid residues affects the stereochemical destiny during the polycyclization process, which has been demonstrated by mutation experiments on I261 and L607.^{11,12} However, it remains uncertain as to which amino acids are responsible for the final deprotonation reaction of **4** to

[†] To whom correspondence should be addressed. Fax: +81-25-262-6854; E-mail: hoshitsu@agr.niigata-u.ac.jp

Abbreviations: SHC, squalene-hopene cyclase; GC, gas chromatography



Scheme 1. Polycyclization Reaction of Squalene **1** into Hopene **2** and Hopanol **3** Catalyzed by Squalene-hopene Cyclase.

The production ratio of **2:3** is *ca.* 84:16. The acyclic and flexible squalene molecule is cyclized through a discrete carbocation to form final cation **4** (the hopanyl cation). Products **2** and **3** are formed by regiospecific deprotonation (path *a*) from the methyl group of C-29 of **4**, derived from Z-Me at C23 of **1**, and hydroxylation (path *b*) at the C-22 cation of **4**, respectively.¹⁴ The structure of hop-21(22)-ene **5** is also shown, which occurs from the erroneous deprotonation reaction of H-21 of intermediate **4** (path *c*) by the mutated cyclases. Compound **5** is not produced by the native cyclase, but is formed by the mutated cyclases.

form **2**.

We have previously established, by using norsqualene lacking methyl group at C-23, that the terminal isopropylidene moiety of **1** was indispensable for initiating the polycyclization reaction and for constructing the five-membered E-ring.¹³ Incubation of deuterated **1** having an *E*- or *Z*-trideuteriomethyl group at C-23 has revealed that proton elimination occurred exclusively from the *Z*-methyl group at the terminal position of **1**, but not from the *E*-methyl group (Scheme 1).^{1,14} In addition, we have suggested that the binding force for the *Z*-methyl group to SHC was stronger than that for the corresponding *E*-methyl group.¹³ Based on the X-ray three-dimensional crystal structures of SHC, Wend *et al.* have suggested the putative structure for the deprotonation site involving some amino acids and water molecules.^{2,15,16} According to their suggestion from the results of an X-ray structural analysis, we have illustrated the putative structure around the deprotonation site as shown in Fig. 1. The catalytic base

responsible for the deprotonation reaction is suggested to be a water molecule (named “front water”) whose polarization can be enhanced by other types of water (named “back waters”) that construct the hydrogen-bonding network associated with the seven residues of T41, E45, E93, R127, W133, Q262 and Y267. The Q262 residue bears a further hydrogen bond with the “front water”. The strongly polarized “front water” can therefore store the proton generated from Me-29 to form **2**, while **3** is produced if the “front water” adds as hydroxyl to the C-22 cation of **4**, instead of accepting the proton (Fig. 1).^{15,16} In addition, the P263 and F437 residues are close to the “front water”. To date, the functions of the three residues of E45, F129 and F605, which may be responsible for the deprotonation site, have been evaluated.^{10,17,18} The π -electrons of F605 work to stabilize the transient carbocations at C-17 and C-22 that are formed during the polycyclization process through cation- π interaction.¹⁰ Residues E45¹⁸ and F129¹⁷ play a crucial role in the catalysis. However, no

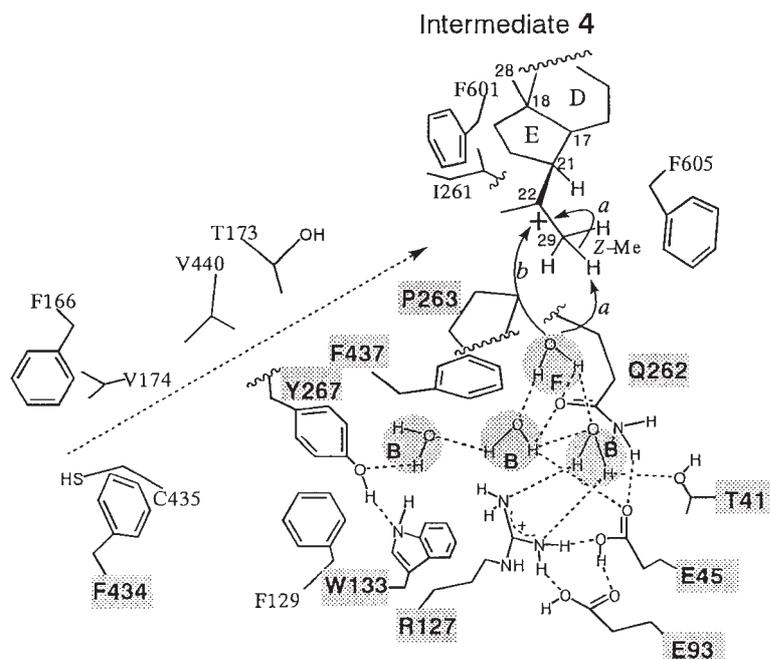


Fig. 1. Putative Structure for the Deprotonation Site of *A. acidocaldarius* SHC.

The locations of amino acid residues and water molecules are illustrated according to the X-ray crystal structure of SHC.¹⁶⁾ Two types of water molecules are depicted: the “front water” and “back waters” are represented by F and B, respectively. The “front water” and “back waters” are associated with each other by hydrogen bonds, and the latter ones further construct the hydrogen-bonded network involving several amino acids. The increased polarity of the “front water” can effectively behave as the catalytic base to form **2**. The “front water” further forms the hydrogen bond with Q262. The amino acid residues responsible for the formation of the hydrogen-bonded network including the “back waters” are as follows: T41, E45, E93, R127, W133 and Y267. The P263 and F437 residues are close to the “front water”. The F434 and C435 residues, situated on the mobile loop, may behave as a gate that permits substrate passage. Y267 and F437 are aligned close to the substrate passage to guide the substrate into the reaction cavity from the gate. This channel is shown with a dotted arrow.

results of mutagenesis experiments targeted on any other of these described residues have been reported before.

Figure 1 also shows some amino acids responsible for substrate entrance from the enzyme surface into the active site cavity through a channel which contains a narrow constriction formed by the four F166, V174, F434 and C435 residues.^{15,16)} It has been suggested that the F434 and C435 residues on the mobile loop may also behave as a gate that permits substrate passage.^{15,16)} Milla *et al.* have recently demonstrated that C435 actually has such a gate function by using a thiol-modifying inhibitor.¹⁹⁾ The function of F434 located at the gate was also examined and is described here. The locations of Y267 and F437 are also near the interface between the putative deprotonation site and the substrate gate (Fig. 1).

In order to understand in detail the catalytic mechanism for the final deprotonation reaction, we carried out mutagenesis experiments targeted on the ten residues, T41, E45, E93, R127, W133, Q262, P263, Y267, F434 and F437, and functional analyses of the amino acid residues are reported here.

Materials and Methods

Analytical methods. ¹H-NMR and ¹³C-NMR spectra were measured in C₆D₆ solutions by a Bruker DPX 400

instrument. The reported chemical shifts of ¹H- and ¹³C-NMR (ppm) are relative to 7.28 and 128.0 ppm of the solvent peak. MS spectra were recorded with a Jeol SX 102 mass spectrometer. GC analyses were carried out by a Shimadzu GC-8A instrument with a DB-1 capillary column (0.53 mm × 30 m, injection temperature of 290°C, column temperature of 200–270°C at 1°C min⁻¹, and flow rate of carrier N₂ of 0.37 kg/cm²). Specific rotation values were measured at 25°C by a Horiba SEPA-300 instrument.

General methods. The detailed experimental protocol has been reported in the previous papers^{20,21)} for the overexpression systems with the pET vector in *E. coli* BL21 (DE3), DNA sequence analyses, preparation of cell-free homogenates and enzyme purification methods.

Site-directed mutagenesis. All the site-directed mutagenesis experiments were performed as described in the previous papers.^{20,21)} The following primers were used: T41A, 5'-pd[CTCCGCTTCCATGGCGACGTTGCTC-AG]-3' (+Nco I); E45A, 5'-pd[GAGGACGTACGCCGCTTCCATGGTGACGTTGCTCAG]-3' (+Nco I); E93A, 5'-pd[CACGACCATCGCGGCCTACGTCGC-GC]-3' (-BsiW I); R127Q, 5'-pd[GAGTCGTCCGAA-GTGTTACGCGTATGTGGCTGG]-3' (+Mlu I); W133A, 5'-pd[GTGTTACGCGTATGGCGCTGGCG-

CTGG]-3' (+*Mlu* I); Q262G, 5'-pd[CTGGGGCGG-GATTGGGCCCTTGGTTTTACG]-3' (+*Apa* I); Q262A, 5'-pd[GGAGACGGCAGTTGGGGCGGGAT-TGCGCCGCCTTGG]-3' (-*Pvu* II); P263G, 5'-pd[GC-TGGGGCGGAATTCAGGGCCCTTGGTTTTACG]-3' (+*EcoR* I); P263A, 5'-pd[GCTGGGGCGGAATTCAGGGCCCTTGGTTTTAC]-3' (+*EcoR* I); Y267A, 5'-pd[GCCTTGGTTTTGCCGCGCTCATAGCGCTCAAG-ATTC]-3' (+*Aor51* HI); F434A, 5'-pd[CTTCGCCGA-AGTCCGATGCCGGGATGTGGTTCGG]-3' (+*Sph* I); and F437A, 5'-pd[GGTCACTTCGCCGGCGTCCGAG-AACGG]-3' (+*Nae* I). The bold letters designate the altered bases, and the italic letters show the target mutations. The underlined letters show the silent mutations for easy screening of the desired mutants by a restriction fragment analysis. The created or deleted restriction sites are shown in parentheses. To ascertain that the desired mutation had been carried out, the entire region of all the inserted DNAs was sequenced.

Structural determination of product 5. Compound **5** had previously been isolated from the reaction mixture of the mutated SHC of I261G (24% of all products).¹¹ However, we did not previously report the NMR data. Complete NMR assignments of **5** were done by 2D-NMR methods including COSY 45, HOHAHA, NOESY, HMQC and HMBC. The spectroscopic data for compound **5** (solid, hop-21(22)-ene) are as follows: ¹H-NMR (400.13 MHz, C₆D₆) δ_{H} (ppm): 2.40 (1H, m, H-20), 2.30 (1H, m, H-16), 2.25 (1H, m, H-20), 1.95 (3H, bs, H-29), 1.84 (1H, bd, *J* = 12.4 Hz, H-17), 1.78 (3H, bs, H-30), 1.78 (2H, m, H-1 and H-16), 1.69 (1H, m, H-19), 1.66 (1H, m, H-11), 1.63 (2H, m, H-2 and H-6), 1.62 (1H, m, H-7), 1.59 (2H, m, H-12), 1.52 (1H, m, H-13), 1.51 (1H, m, H-3), 1.47 (1H, m, H-15), 1.42 (2H, m, H-9 and H-11), 1.41 (2H, m, H-2 and H-6), 1.37 (1H, m, H-7), 1.35 (1H, m, H-15), 1.28 (1H, m, H-3), 1.13 (1H, m, H-19), 1.125 (3H, s, H-27), 1.10 (3H, s, H-26), 1.04 (3H, s, H-23), 0.994 (3H, s, H-24), 0.988 (3H, s, H-25), 0.89 (1H, m, H-1), 0.88 (1H, m, H-5), 0.84 (3H, s, H-28). ¹³C-NMR (100.6 MHz, C₆D₆) δ_{C} (ppm): 135.8 (C-21), 120.6 (C-22), 56.5 (C-5), 56.3 (C-17), 50.8 (C-9), 48.4 (C-13), 44.6 (C-18), 42.4 (C-3), 42.2 (C-14), 41.7 (C-8), 40.6 (C-1), 39.5 (C-19), 37.7 (C-10), 33.62 (C-23), 33.60 (C-7), 33.4 (C-4), 33.1 (C-15), 28.8 (C-20), 24.0 (C-12), 23.7 (C-16), 23.0 (C-30), 21.8 (C-24), 21.3 (C-11), 19.7 (C-29), 19.09 (C-6), 19.07 (C-2), 16.87 (C-27), 16.84 (C-26), 16.1 (C-25). EIMS fragments *m/z* (%): 410 (45) [M⁺], 395 (14), 367 (38), 341 (54), 231 (26), 218 (10), 205 (16), 203 (18), 191 (95), 189 (100), 161 (68). HREIMS for C₃₀H₅₀: calcd., 410.3913; found, 410.3900. [α]_D²⁵ +29.1 (*c* = 0.68, CHCl₃), *cf.* +29.8 (*c* = 0.3–0.6, CHCl₃) in ref. 22.

Product distribution pattern. The *E. coli* transformant, in which the mutated SHC was expressed, was cultured at 30°C for 20 h in a Luria-Bertani medium (100 ml) containing 50 mg/l of ampicillin. The cells

were collected by centrifugation and washed twice with a Tris-HCl buffer (pH 8.0, 50 mM). To the collected pellets, 10 ml of a Tris buffer supplemented with 1% Triton X-100 (w/v) was added. The bacterial cells were subjected to ultrasonication at 4°C and then centrifuged to remove the cell debris to prepare a cell-free homogenate. The pH value of the supernatant was adjusted to 6.0 by adding 0.5 M citric acid. Five ml of the cell-free extract thus prepared, which corresponds to ca. 0.5 mg of the homogeneously purified SHC, was used for evaluating the product distribution. One mg of **1** was separately incubated at 45°C for 16 h. A hexane extract was obtained from the reaction mixture. The detergent was removed by passing the extract through a short SiO₂ column [*n*-hexane:EtOAc = 100:20], and then the extract subjected to GC analyses to examine the product distribution pattern and to quantify the product amounts. The enzymatic products were identified by comparing their GC-MS fragment patterns with those of authentic samples that had previously been isolated by us.

Enzyme assay. All the mutated SHCs were homogeneously purified according to the protocol described in the previous paper.^{20,21} A mixed solution, which was composed of 0.5 mM squalene, 0.2% Triton X-100 and 5 μ g of the homogeneously purified enzyme in a sodium citrate buffer (60 mM, pH 6.0), was prepared in a final volume of 5 ml for the enzyme reactions. Incubation was done for 60 min at different temperatures (30, 35, 40, 45, 50, 55, 60, 65 or 70°C) to examine the thermal stability of each cyclase. To evaluate the kinetic parameters, incubation was conducted for 60 min at 45°C before adding 15% methanolic KOH (6 ml) to terminate the enzyme reaction. The lipophilic products (**2**, **3** and, in some cases, **5**) and unreacted **1** were extracted with *n*-hexane (5 ml \times 4) from each reaction mixture, and the final products were quantified by GC analyses with a DB-1 capillary column (30 m in length). The kinetic values of *K_m* and *k_{cat}* were estimated from Lineweaver-Burk plots.

Results and Discussion

To understand the catalytic mechanism for the deprotonation reaction, the amino acid residues shown in Fig. 1 were mutated. The following twelve mutants were constructed: T41A, E45A, E93A, R127Q, W133A, Q262G, Q262A, P263G, P263A, Y267A, F434A and F437A. According to the location sites shown in Fig. 1, these mutants are classified into three categories. Category A: Q262G, Q262A, P263G and P263A, which were located between **4** and the “front water”. Category B: T41A, E45A, E93A, R127Q and W133A, which were positioned around the “back waters”. Category C: Y267A, F434A and F437A, which were situated near the substrate channel.

Product distribution pattern of mutated SHCs

We have previously shown that point mutants, which had been altered at active sites, afforded prematurely cyclized products. Based upon the product distribution patterns, the positions of numerous active residues inside the central active cavity have been identified.¹⁾ With a large amount of the cell-free homogenate from *E. coli* encoding the mutated SHCs, **1** was incubated at 45 °C for 16 h, and the enzymic products were analyzed by GC. As a typical example, the product distribution patterns for the wild-type and P263G mutant are shown in Fig. 2. The production of **3** by the P263G mutant was significantly higher than that of the wild-type. Together with **2** and **3**, compound **5** was additionally produced by the P263G mutant. No other aberrant cyclization product was detected. Product **5** was identified to be hop-21(22)-

ene (Scheme 1) by co-injection and comparing the GC-MS fragment pattern with that of an authentic sample which had previously been isolated by us from the I261G mutant¹¹⁾ (see the Materials and Methods section).

The product distribution patterns of all the mutated SHCs are summarized in Table 1. All the mutants produced hydroxylated **3** in a higher yield than that of the wild-type (Table 1), indicating that the “front water” was used to add to the C-22 carbocation of **4** instead of accepting the proton of Me-29 as the catalytic base (Fig. 1). It is of particular interest that the Q262G, G262A, P263G and P263A mutants (category A) produced **3** as the main product (60–73%) instead of **2**. These mutants (category A) also produced **5**, although in a small quantity, which can be produced by the proton

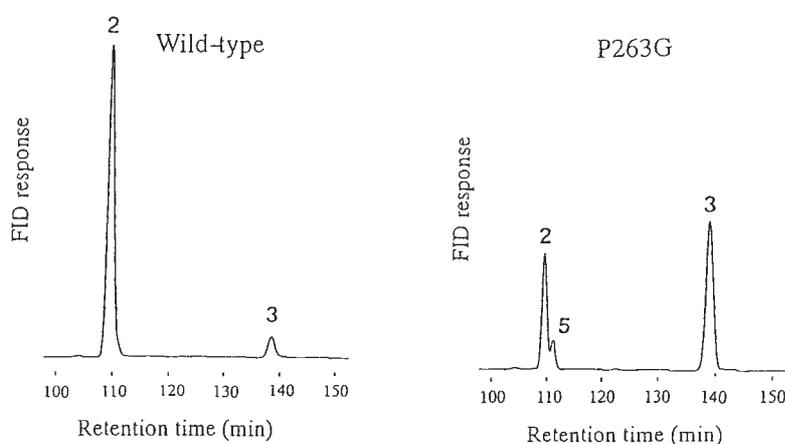


Fig. 2. Gas Chromatograms Showing the Product Distribution of the Wild-type and Mutated P263G SHCs. A marked increase in hopanol **3** and new product **5** were found for the P263G mutant.

Table 1. Product Ratios (%) of the Wild-type and Mutant SHCs and Their Relative Activities (%)

Squalene **1** (1.00 mg) was incubated with 5 ml of a cell-free extract, the amount of which corresponds to a half mg of each purified enzyme, at pH 6.0, 45 °C, and for 16 h (see the Materials and Methods section). Product ratios were estimated by GC analyses. Figures in parentheses show the relative activity (%) of each mutated SHC, compared to that of the wild-type, which were determined by incubating **1** (1.00 mg) with 5 μg of the homogeneously purified SHC for 60 min. The incubation was done for a short time by using a small amount of the SHC, because a large amount of the SHC and a prolonged incubation fully converted **1** into **2** and **3**.

SHC	Category*	Hopene (2)	Hopanol (3)	5
Wild-type		84.1 (100)	15.9 (100)	—
Q262G	(A)	22.4 (96)	71.8 (1753)	5.8
Q262A	(A)	35.9 (113)	60.1 (1071)	4.0
P263G	(A)	20.1 (87)	73.4 (1502)	6.5
P263A	(A)	21.1 (76)	72.9 (1540)	6.0
T41A	(B)	82.0 (92)	18.0 (97)	—
E45A	(B)	61.3 (41)	38.7 (133)	—
E93A	(B)	67.4 (36)	32.6 (98)	—
R127Q	(B)	68.2 (54)	31.8 (98)	—
W133A	(B)	76.4 (74)	23.6 (112)	—
Y267A	(C)	75.3 (88)	24.7 (129)	—
F434A	(C)	77.9 (58)	22.1 (85)	—
F437A	(C)	67.5 (130)	32.5 (331)	—

* Category A–C was classified according to the location site (see Fig. 1).

elimination reaction of H-21 of intermediate **4** (path *c* in Scheme 1), but **5** was not produced by the mutants of categories B and C (Table 1).

Enzyme activities of mutated SHCs

The specific enzyme activities of mutant categories A–C against the incubation temperature are illustrated in Figs. 3–5, respectively. The specific activities for the formation of **2**, **3** or all the products (**2** + **3** + **5**) are shown in I–III, respectively. The nine E45A, E93A, R127Q, W133A, Q262G, Q262A, P263G, P263A and F437A mutants had decreased activities in the high-temperature range (Figs. 3–5). However, the optimal temperature of the three T41A, Y267A and F434A mutants remained unchanged. The kinetic data for mutant categories A–C are shown in Tables 2–4, respectively.

Category A

Figures 3 I and II show that the production of **2** was

decreased, while that of **3** was markedly increased at lower temperatures by these mutations. The data in Table 1 further support this result. The markedly increased specific activities for all the products (**2** + **3** + **5**) at low temperatures (Fig. 3 III) were thus due to the significantly increased production of **3**. We have previously reported that the replacement of Phe365 and Phe605 by Tyr or Trp gave a marked increase in k_{cat} at lower temperatures. However, the mutants replaced by aliphatic amino acids lacking π -electrons gave premature cyclization products, which were generated from the corresponding cationic intermediate(s), without completion of the polycyclization reaction.^{9,10} Thus, aromatic π -electrons play a crucial role in the acceleration of reaction velocity, resulting from the stabilization of transient carbocations through a cation- π interaction.^{9,10,23} However, the methyl group and hydrogen atom of the Ala and Gly mutants (Q262G, Q262A, P263G and P263A) cannot stabilize a transient carbocation(s), due to the lack of π -electrons. The residues of

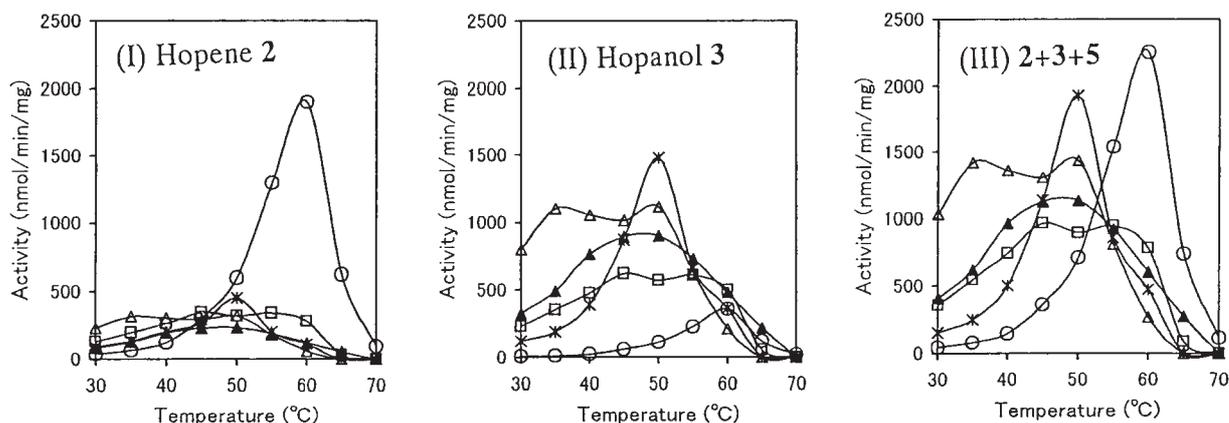


Fig. 3. Specific Enzyme Activities of the Wild-type and Mutated SHCs in Category A against Incubation Temperature.

One mg of **1** was incubated with 5 μ g of each homogeneously purified SHC for 60 min at pH 6.0. The enzyme activities estimated by the produced amounts of **2**, **3**, and all the products (**2** + **3** + **5**) are illustrated in panels I, II and III, respectively. The definition of symbols is as follows: \circ wild-type, \triangle Q262G, \square Q262A, * P263G, \blacktriangle P263A.

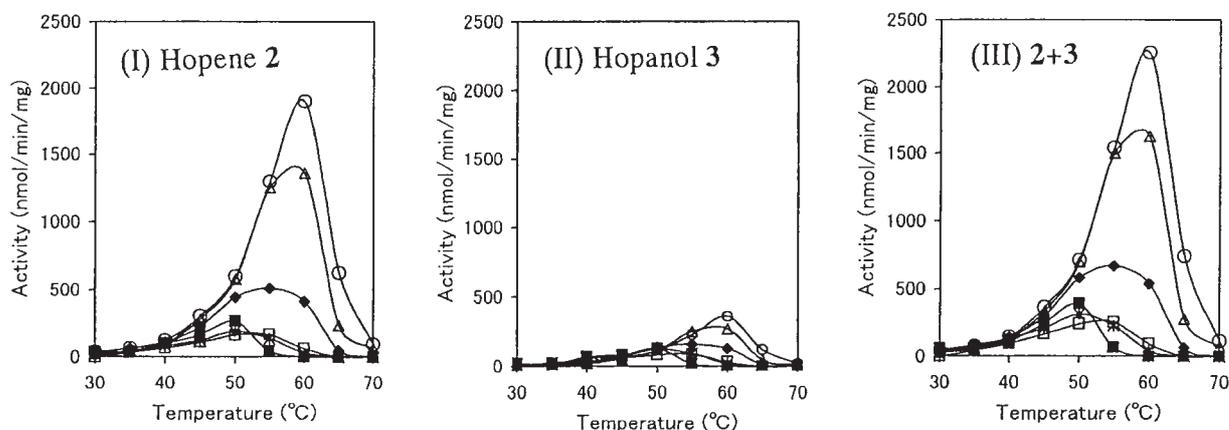


Fig. 4. Specific Enzyme Activities of the Wild-type and Mutated SHCs in Category B against Incubation Temperature.

The incubation conditions were the same as those described for Fig. 3. The enzyme activities estimated by the amounts of **2**, **3**, and **2** + **3** are depicted in panels I, II and III, respectively. \circ wild-type, \triangle T41A, * E45A, \square E93A, \blacksquare R127Q, \blacklozenge W133A.

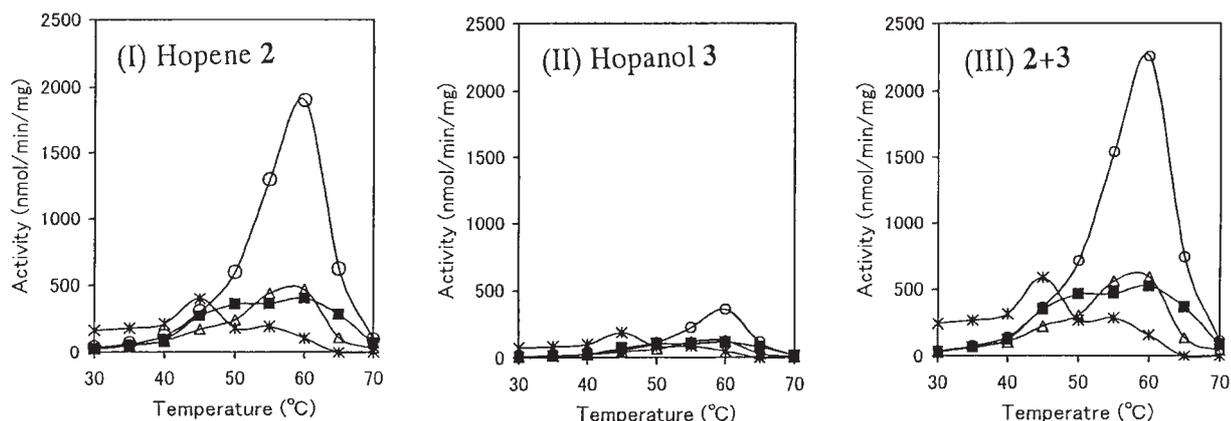


Fig. 5. Specific Enzyme Activities of the Wild-type and Mutated SHCs in Category C against Incubation Temperature.

The incubation conditions were the same as those described for Figs. 3 and 4. The enzyme activities estimated by the amounts of **2**, **3**, and **2 + 3** are depicted in panels I, II and III, respectively. ○ wild-type, ■ Y267A, △ F434A, * F437A.

Table 2. Kinetic Data and Optimal Temperatures for the Wild-type and Mutated SHCs in Category A

The kinetic values of K_m and k_{cat} were determined by incubating at 45°C, pH 6.0 and for 60 min, during which, the purified proteins were not denatured. These kinetic data were each estimated for the formation of **2**, **3** and all the products (**2 + 3 + 5**).

SHC	Estimated product	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m	Relative activity (%)	Optimal temperature for 60 min ($^{\circ}C$)
Wild-type	2	1.62×10	51.8	3.19	100.0	60
	3	1.62×10	9.8	0.60	100.0	60
	2 + 3	1.62×10	61.6	3.79	100.0	60
Q262G	2	1.56×10^2	82.1	5.26×10^{-1}	16.5	35–50
	3	1.56×10^2	263.2	1.69	281	35–50
	2 + 3 + 5	1.56×10^2	366.5	2.35	61.8	35–50
Q262A	2	1.02×10^2	72.8	7.14	22.4	45–55
	3	1.02×10^2	121.9	1.20	200	45–55
	2 + 3 + 5	1.02×10^2	202.8	1.99	52.4	45–55
P263G	2	2.37×10^2	63.4	2.68×10^{-1}	8.4	50
	3	2.37×10^2	231.5	9.77×10^{-1}	162.8	50
	2 + 3 + 5	2.37×10^2	315.4	1.33	35.1	50
P263A	2	1.85×10^2	65.6	3.55×10^{-1}	11.1	50
	3	1.85×10^2	226.6	1.22	203.3	50
	2 + 3 + 5	1.85×10^2	310.9	1.68	44.3	50

Q262 and P263 are located between **4** and the “front water”, as shown in Fig. 1.^{15,16} Table 2 shows that the values of k_{cat} and K_m for the formation of all the products (**2 + 3 + 5**) were higher than those of the wild-type (k_{cat} : 1.2–1.6-fold for **2**, 12.4–26.9-fold for **3**, 3.3–6.0-fold for all the products; K_m : 5.9–14.6-fold). This finding indicates that the affinity of the cyclase to **1** was less than the wild-type, despite the reaction velocity of these mutants having been faster for the formation of all the products. The decreased binding of **1** would have occurred due to the decreased steric bulk size of the Ala or Gly mutants, indicating the requirement of an appropriate steric bulk size at positions 262 and 263. Perturbation around these residues may have resulted in closer contact of the “front water” to cationic intermediate **4** and thus led to the significantly higher production of **3**. The slightly accelerated reaction

velocity for the formation of **2** at lower temperatures, and the unusual formation of **5**, which was produced by the deprotonation reaction from H-21 (path *c*), may also be explained in terms of the closer orientation of the “front water” to **4**. As an alternative explanation, perturbation around the “front water” may have promoted the formation of **5**, because the deprotonation reaction from H-21 for the formation of **5** would be more preferable to that from H-29 for **2** (a Zaitsev rule). We thus propose that residues Q262 and P263 worked to situate the catalytic base of the “front water” at the appropriate position. The favorable placement of the “front water” would be effective for suppressing the production of by-products **3** and **5**. Indeed, the Q262G and P263G mutants having the smaller side chain produced **3** and **5** in higher yields than the Q262A and P263A mutants (Table 1). We have previously reported

Table 3. Kinetic Data and Optimal Temperatures for the Wild-type and Mutated SHCs in Category B

The incubation conditions were the same as those described for Table 2. These kinetic data were each estimated for the formation of **2**, **3** and all the products (**2 + 3**).

SHC	Estimated product	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m	Relative activity (%)	Optimal temperature for 60 min ($^{\circ}\text{C}$)
Wild-type	2	1.62×10	51.8	3.19	100.0	60
	3	1.62×10	9.8	6.05×10^{-1}	100.0	60
	2 + 3	1.62×10	61.6	3.79	100.0	60
T41A	2	1.67×10	44.1	2.64	82.8	60
	3	1.67×10	9.8	5.87×10^{-1}	97.0	60
	2 + 3	1.67×10	53.9	3.23	85.2	60
E45A	2	2.13×10	20.7	9.72×10^{-1}	30.5	50
	3	2.13×10	13.3	6.24×10^{-1}	103	50
	2 + 3	2.13×10	34.0	1.60	42.2	50
E93A	2	1.89×10	21.2	1.12	35.1	55
	3	1.89×10	10.4	5.50×10^{-1}	90.9	55
	2 + 3	1.89×10	31.6	1.67	44.1	55
R127Q	2	2.55×10	26.9	1.05	32.9	50
	3	2.55×10	12.8	5.02×10^{-1}	83.0	50
	2 + 3	2.55×10	39.7	1.56	41.2	50
W133A	2	1.69×10	37.3	2.21	69.3	55
	3	1.69×10	11.5	6.80×10^{-1}	112.4	55
	2 + 3	1.69×10	48.8	2.89	76.3	55

Table 4. Kinetic Data and Optimal Temperatures for the Wild-type and Mutant SHCs in Category C

The incubation conditions were the same as those described for Table 2 and 3. These kinetic data were each estimated for the formation of **2**, **3** and all the products (**2 + 3**).

SHC	Estimated product	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m	Relative activity (%)	Optimal temperature for 60 min ($^{\circ}\text{C}$)
Wild-type	2	1.62×10	51.8	3.19	100.0	60
	3	1.62×10	9.8	6.05×10^{-1}	100.0	60
	2 + 3	1.62×10	61.6	3.79	100.0	60
Y267A	2	1.97×10^2	38.5	1.95×10^{-1}	6.1	50–60
	3	1.97×10^2	12.6	6.40×10^{-2}	10.6	50–60
	2 + 3	1.97×10^2	51.1	2.59×10^{-1}	6.8	50–60
F434A	2	8.16×10	32.3	3.96×10^{-1}	12.4	60
	3	8.16×10	9.2	1.13×10^{-1}	18.7	60
	2 + 3	8.16×10	41.5	5.09×10^{-1}	13.4	60
F437A	2	9.55×10	61.5	6.44×10^{-1}	20.2	45
	3	9.55×10	29.6	3.10×10^{-1}	51.7	45
	2 + 3	9.55×10	91.1	9.54×10^{-1}	25.1	45

the enzymic products from the mutants of Phe605A,¹⁰⁾ but **5** could not be isolated, possibly due to its low production. We reexamined the enzymic products from the mutants of Phe605A; GC-MS analyses clearly showed the involvement of **5** in the reaction mixtures, although in a small amount (3.8%). As found with the P263A and P263G mutants, the less bulky size at the 605 position would also have induced perturbation around the “front water”, resulting in the high production of **3**¹⁰⁾ and **5**. As shown in Fig. 1, the terminal amide group of Q262 forms a hydrogen bond with the “front water”. The substitution of Gln with Ala or Gly could also disrupt the hydrogen bond, thus leading to higher

production of **3** and **5** due to perturbation around the “front water”.

Category B

As shown in Table 1, **2** and **3** were produced, although there was no detectable amount of **5**, by the category B mutants (T41A, E45A, E93A, R127Q and W133A). The production of **2** was less throughout the entire temperature range examined (Fig. 4 I) than that by the wild-type, but that of **3** was almost the same as that of the wild-type, being slightly higher in the low-temperature ranges (Fig. 4 II). The specific activities for the formation of (**2 + 3**) were smaller than those of the

wild-type throughout the entire temperature range examined (Fig. 4 III). The kinetic analyses revealed that k_{cat} for the production of **2** was lower (40–85%), but that for the production of **3** was slightly higher (106–135%) than by the wild-type (Table 3). These mutations would have disrupted the hydrogen-bonding network constructed for the “back waters”, thus lowering the polarity of the “front water” and leading to a slower reaction for the deprotonation of H-29 (path *a*) to form **2**. The nucleophilic property of the “front water” to the C-22 cation (path *b*) for the formation of **3** was slightly increased. This may suggest that the correct position of the “front water” could have been slightly disordered due to breakage of the hydrogen bond between these amino acid residues and the “back waters” by a single amino acid substitution. The difference in K_m for the T41A, E45A, E93A, R127Q and W133A mutants was minor (1.0–1.5-fold) in comparison with that of the wild-type. This finding suggests that the affinity of **1** to the cyclase was not affected by these mutations (Table 3); that is, the interactive contact of substrate **1** with these amino acid residues was little. The relative activity, k_{cat}/K_m , for the formation of **2** was significantly decreased by the mutations. This finding further supports the idea that the hydrogen-bonding network, which had been constructed by assembling the residues of E45, E93, R127 and W133 including the “back waters”, strongly worked to polarize the “front water”, leading to the higher production of **2** than of **3**. There was little difference in the relative activity (k_{cat}/K_m) for the formation of **3**, suggesting that the position of the “front water” remained unchanged or only slightly altered by these site-directed mutations. There was also little difference in the kinetic values and specific activities with incubation temperature (Table 3 and Fig. 4) for the T41A and W133A mutants when compared to the figures for other mutants E45A, E93A and R127Q. This finding indicates that polarization of the “front water” could have been archived mainly by the three E45, E93 and R127 residues.

Category C

The F434 residue is located at the substrate gate (Fig. 1). Compared to the wild-type, the Ala mutant had significantly lower affinity (5-fold in K_m), although difference in reaction velocity was small (1/1.6 fold in k_{cat}), as shown in Table 4. This may suggest an important role of the aromatic residue for capturing substrate **1** and for guiding **1** into the reaction cavity. The K_m value of Y267A for the formation of **2**, **3** or (**2** + **3**) was also each significantly decreased by 12-fold. The Y267 residue participates in the hydrogen-bonding network with the “back water” (Fig. 1); thus, Y267 may be classified into category B. The lower of k_{cat} value for the formation of **2** is characteristic of category B. Indeed, k_{cat} decreased by 1/1.3 fold with a small change. The difference in K_m for category B was little by the point mutations, as shown in Table 3. However, mutated

Y267A had an exceptionally large K_m value, probably due to the location of Y267 in the proximity of the substrate passage. A higher K_m value was also observed with the F434 mutant situated at the substrate entrance. In respect of the specific activity against the incubation temperature, mutated F437A showed a different profile from that of the F434A or Y267A mutant; at lower temperatures, the production of **2** and **3** was each higher than that of the wild-type (Fig. 5 I–III). As shown in Table 4, the increase in k_{cat} for the production of **3** (3-fold) was higher than that of **2** (1.2-fold). This profile is characteristic of category A (Fig. 3 I–III). The higher values of K_m and k_{cat} with F437A at lower temperatures also supports the classification of F437 into category A, but not into category B, suggesting that the F437 residue is likely to have been situated near the “front water”, and not near the “back waters”. However, the increase in the formation of **3** (3-fold) was significantly smaller than that of the category A mutants (12–27-fold) (compare Tables 2 and 4). In addition, the formation of product **5**, characteristic of category A, was not detected by the F437A mutant. These findings suggest that F437 was located at three interfaces among the “front water”, “back waters” and substrate channel, which is in a good agreement with the inference from the X-ray analysis (Fig. 1).

We have previously reported the kinetic data for the F129A mutant:¹⁷⁾ K_m $1.85 \times 10 \mu\text{M}$ and k_{cat} 13.7 min^{-1} for the formation of **2**; K_m $1.85 \times 10 \mu\text{M}$ and k_{cat} 2.5 min^{-1} for that of **3**. This K_m value is nearly equal to that for the wild-type, while k_{cat} is significantly lower. F129 is not involved in the hydrogen-bonded network with the “back waters” and is located around residues Y267 and W133 as shown in Fig. 1. The hydrogen bond could be disrupted around Y267 and W133 due to the decreased bulk size by mutating Phe into Ala, leading to a slower reaction rate as shown for category B. Thus, no detectable amount of **5** and no increased amount of **3** was observed for the mutant F129A.

In conclusion, the results of the site-directed mutagenesis experiments described here allowed us to identify the amino acid residues that play a crucial role in the final deprotonation reaction leading to **2**. Substitution of Gln with Ala at the 262 position disrupted a hydrogen bond of the side chain of Gln with the “front water”, thus enabling the disordered “front water” to approach final cation **4**. This could lead to the formation of a larger quantity of **3** than **2**. Pro263 did not form a hydrogen bond with the “front water”, but the Ala and Gly mutants produced **3** in a significantly large amount. This would have been due to perturbation of the correct position of the “front water” induced by the decreased bulk size of Ala and Gly. The F437A mutant also showed a high production of **3**, but the increase was less than that by P263, suggesting that Pro263 was closer to the “front water” than in F437. Several amino acids are responsible for the hydrogen-bonded network with the

“back waters”, which form further hydrogen bonds with the “front water”, thus increasing the polarity of the “front water” to give rise to a strong catalytic base for the “front water”. The disruption of the hydrogen-bonded network responsible for the “back waters” led to reduced polarity of the “front water” (a weakened catalytic base), thus giving less activity for the formation of **2**. The residues located at the substrate entrance or passage were also examined. The F434A mutant resulted in significantly decreased binding to the substrate. The Y267 and F437 residues were aligned near the passage. The mutation of these residues into Ala also decreased the binding of the substrate with the cyclase. The Phe166 aromatic residue was also located at the substrate gate (Fig. 1). Aromatic residues may play an important role in capturing a hydrophobic substrate and guiding it into the reaction cavity. To validate this assumption, further mutation experiments are necessary with the residues F166, V174, V440 and so on (see Fig. 1).

A variety of quenching sites for the last cation generated during the polycyclization cascade leads to structural diversity of the triterpene skeleton. At the present time, the amino acid residue(s) responsible for final reaction step (quenching of the final carbocation, *e.g.*, the deprotonation site or hydroxylation site) of the triterpene cyclase family is still unknown, except for the SHC. If detailed information can be collected for the final reaction step, it will be possible to create many “unnatural natural” triterpenes in future. We have demonstrated that the steric bulk size of the crucial residues, which are in close contact with the folded and/or extended substrate, significantly affected the folding conformation of squalene, leading to novel triterpenes.^{11,12} Very interesting reports have recently appeared that squalene analogs (truncated C₂₉ analogs and an elongated C₃₅ analog) were successfully cyclized to give novel carbocyclic skeletons by using *A. acidocaldarius* squalene-hopene cyclase^{24,25} and that the natural indole diterpene, petromindole, was synthesized in a single step from the indole derivative of the geranylgeranyl isoprenoid, 3-(ω -oxidogeranylgeranyl)-indole, by using *Arabidopsis thaliana* lupeol synthase (a plant oxidosqualene cyclase).²⁶ The combined use of altered triterpene cyclases and substrate analogs will also be promising for creating further unnatural natural compounds and for one-pot syntheses of natural terpenoids.

Acknowledgments

This work was supported by grant-aid to T. H. (Nos. 13306009 and 15380081) and to T. S. (No. 15780083).

References

- 1) Hoshino, T., and Sato, T., Squalene-hopene cyclase: catalytic mechanism and substrate recognition. *J. Chem.*

- Soc. Chem. Commun.*, 291–301 (2002).
- 2) Wendt, K. U., Schulz, G. E., Corey, E. J., and Liu, D. R., Enzyme mechanisms for polycyclic triterpene formation. *Angew. Chem. Int. Ed.*, **39**, 2812–2833 (2000).
- 3) Abe, I., Rohmer, M., and Prestwich, G. D., Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. *Chem. Rev.*, **93**, 2189–2206 (1993).
- 4) Segura, M. J. R., Jackson, B. E., and Matsuda, S. P. T., Mutagenesis approaches deduce structure-function relationship in terpene synthases. *Nat. Prod. Rep.*, **20**, 304–317 (2003).
- 5) Kushiro, T., Shibuya, M., and Ebizuka, Y., Cryptic regioselectivity in deprotonation step of triterpene biosynthesis catalyzed by new members of lupeol synthase. *Tetrahedron Lett.*, **40**, 5553–5556 (1999).
- 6) Hoshino, T., Kouda, M., Abe, T., and Ohashi, S., New cyclization mechanism for squalene: a ring expansion step for the five-membered C-ring intermediate in hopene biosynthesis. *Biosci. Biotechnol. Biochem.*, **63**, 2038–2041 (1999).
- 7) Sato, T., Abe, T., and Hoshino, T., On the cyclization mechanism of squalene: a ring expansion process of the five-membered D-ring intermediate. *J. Chem. Soc. Chem. Commun.*, 2617–2618 (1998).
- 8) Sato, T., and Hoshino, T., Functional analysis of the DXDDTA motif in squalene-hopene cyclase by site-directed mutagenesis experiments: initiation site of the polycyclization reaction and stabilization site of the carbocation intermediate of the initially cyclized A-ring. *Biosci. Biotechnol. Biochem.*, **63**, 2189–2198 (1999).
- 9) Hoshino, T., and Sato, T., Functional analysis of phenylalanine 365 in hopene synthase, a conserved amino acid in the families of squalene and oxidosqualene cyclases. *J. Chem. Soc. Chem. Commun.*, 2005–2006 (1999).
- 10) Hoshino, T., Kouda, M., Abe, T., and Sato, T., Functional analysis of Phe605, a conserved aromatic amino acid in squalene-hopene cyclase. *J. Chem. Soc. Chem. Commun.*, 1485–1486 (2000).
- 11) Hoshino, T., Abe, T., and Kouda, M., Unnatural natural triterpenes produced by altering isoleucine into alanine at position 261 in hopene synthase and the importance of having the appropriate bulk size at this position for directing the stereochemical destiny during the polycyclization cascade. *J. Chem. Soc. Chem. Commun.*, 441–442 (2000).
- 12) Sato, T., Sasahara, S., Yamakami, T., and Hoshino, T., Functional analyses of Tyr420 and Leu607 of *Alicyclobacillus acidocaldarius* squalene-hopene cyclase. Neoa-chillapentaene, a novel triterpene with the 1,5,6-trimethylcyclohexene moiety produced through a folding of the constrained boat structure. *Biosci. Biotechnol. Biochem.*, **66**, 1660–1670 (2002).
- 13) Hoshino, T., and Kondo, T., The cyclization mechanism of squalene in hopene biosynthesis: the terminal methyl groups are critical to the correct folding of this substrate both for the formation of the five-membered E-ring and for the initiation of the polycyclization reaction. *J. Chem. Soc. Chem. Commun.*, 731–732 (1999).
- 14) Hoshino, T., Nakano, S., Kondo, T., and Miyoshi, A., Report in preparation. Preliminary results have been reported at Annual Meeting Japan Society for Biosci. Biotechnol. and Agrochem., April 1998, Nagoya, Ab-

- stract p. 256.
- 15) Wendt, K. U., Poralla, K., and Schulz, G. E., Structure and function of a squalene cyclase. *Science*, **277**, 1811–1815 (1997).
 - 16) Wendt, K. U., Lenhart, A., and Schulz, G. E., The structure of the membrane protein squalene-hopene cyclase at 2.0 Å resolution. *J. Mol. Biol.*, **286**, 175–187 (1999).
 - 17) Sato, T., and Hoshino, T., Catalytic function of the residues of phenylalanine and tyrosine conserved in squalene-hopene cyclases. *Biosci. Biotechnol. Biochem.*, **65**, 2233–2242 (2001).
 - 18) Dang, T., and Prestwich, G. D., Site-directed mutagenesis of squalene-hopene cyclase: altered substrate specificity and product distribution. *Chem. Biol.*, **7**, 643–649 (2000).
 - 19) Milla, P., Lenhart, A., Grosa, G., Viola, F., Weihofen, W. A., Schulz, G. E., and Balliano, G., Thiol-modifying inhibitor for understanding squalene cyclase function. *Eur. J. Biochem.*, **269**, 2108–2116 (2002).
 - 20) Sato, T., Kanai, Y., and Hoshino, T., Overexpression of squalene-hopene cyclase by the pET vector in *Escherichia coli* and first identification of tryptophan and aspartic acid residues inside the QW motif as active sites. *Biosci. Biotechnol. Biochem.*, **62**, 407–411 (1998).
 - 21) Sato, T., and Hoshino, T., Kinetic studies on the function of all the conserved tryptophans involved inside and outside the QW motifs of squalene-hopene cyclase: stabilizing effect of the protein structure against thermal denaturation. *Biosci. Biotechnol. Biochem.*, **63**, 1171–1180 (1999).
 - 22) Ageta, H., Shiojima, K., and Arai, Y., Acid-induced rearrangement of triterpenoid hydrocarbons belonging to the hopane and migrated hopane series. *Chem. Pharm. Bull.*, **35**, 2705–2716 (1987).
 - 23) Dougherty, D.A., Cation- π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science*, **271**, 163–168 (1996).
 - 24) Hoshino, T., and Ohashi, S., Importance of the methyl group at C(10) of squalene for hopene biosynthesis and novel carbocyclic skeletons with 6/5 + 5/5 + (6) ring system(s). *Org. Lett.*, **4**, 2553–2556 (2002).
 - 25) Abe, I., Tanaka, H., and Noguchi, H., Enzymatic formation of an unnatural hexacyclic C35 polyprenoid by bacterial squalene cyclase. *J. Am. Chem. Soc.*, **124**, 14514–14515 (2002).
 - 26) Xiong, Q., Zhu, X., Wilson, W. K., Ganesan, A., and Matsuda, S. P. T., Enzymatic synthesis of an indole diterpene by an oxidosqualene cyclase: mechanistic, biosynthetic, and phylogenetic implications. *J. Am. Chem. Soc.*, **125**, 9002–9003 (2003).