Biosci. Biotechnol. Biochem., **62** (2), 407–411, 1998

JSBA

Rapid paper

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*

Tsutomu SATO,[1] Yoshinori KANAI,[2] and Tsutomu HOSHINO[1,2†]

[1]*Graduate School of Science and Technology, and*
[2]*Department of Applied Biological Chemistry, Faculty of Agriculture, Niigata University, Ikarashi, Niigata 950-21, Japan*

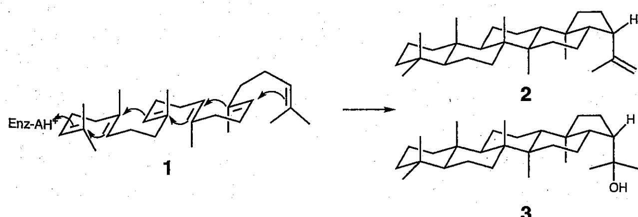
Received September 12, 1997

An overexpression system for squalene-hopene cyclase (SHC) was constructed by using the pET3a vector, which is responsible for high expression with help from the strong T7 promoter when incorporated into *E. coli* BL21(DE3). Site-directed mutagenesis experiments prove that two amino acid residues of tryptophan and aspartic acid inside the QW-motif 5 resided as active sites.

Key words: squalene; hopene; squalene cyclase; *Alicyclobacillus acidocaldarius*; site-directed mutagenesis

Squalene (**1**) is cyclized to form the pentacyclic triterpenes of hopene (**2**) and hopanol (**3**) (scheme) by the enzymic action of squalene-hopene cyclase (SHC) [EC 5.4.99.-] found in prokaryotic bacterium species. The polyolefin cyclization mechanism for squalene cyclase is apparently similar to those for oxidosqualene cyclases (OSCs) of oxidosqualene-lanosterol cyclase [EC 5.4.99.7] and cycloartenol cyclase [EC 5.4.99.8] from eucaryotes. Squalene and oxidosqualene cyclases catalyze one of the most complicated biosynthetic reactions found in nature. In the past five years, there have been remarkable advances with these enzymes; various cyclase genes have been cloned and the amino acid alignments have been deduced from bacteria,[1,2] fungi,[3,4] yeasts,[5–7] plants[8] and mammals.[9–12]

Johnson *et al.*[13] have suggested that the negative point charges of the enzyme would control the polyolefin cyclization reaction through stabilization of the carbocation generated on the cyclizing substrate. Griffin *et al.*[6] have speculated that the carbocation intermediates might be stabilized through cation-π interaction, because aromatic residues are unusually abundant in SHCs and OSCs. On the basis of the amino acid alignments involved in SHCs and OSCs, Poralla[14] has found a highly conserved repetitive motif in the two cyclases, the so-called QW motif [(K/R)(G/A)$X_{2\text{-}3}$(F/Y/W)(L/I/V)X_3Q$X_{2\text{-}5}$GXW], and assumed that the role of aromatic residues (F, Y and W) located inside the QW motif would act as the negative point charges proposed by Johnson through cation-π interaction. However, no experimental evidence for the role of the QW motifs has been presented so far. Very recently, Feil *et al.*[15] have reported that the aspartate residues in the DDTA motif of SHC, which is outside the QW motif, were essential for enzyme activity, although the function of the conserved aspartate inside the QW motif is still unknown. To date, little is known about the responsibility of the highly conserved amino acids that are found in both SHCs and OSCs, for the polyolefin cyclization mechanism.



To access the active sites, we have previously constructed the overexpression system for SHC by using the pET expression vector[16] in *E. coli*, and we report herein the first proof by the method of site-directed mutagenesis that the tryptophan and aspartate residues in the QW motif are respectively crucial for substrate recognition and stabilization of the carbocation intermediates.

Materials and Methods

Strains, plasmids and growth conditions. *Alicyclobacillus acidocaldarius* JCM 5260^T was grown at pH 3 and 60°C in a sporulation medium.[17] *E. coli* strain BL21(DE3) was used for the expression of wild-type and mutant SHC. *E. coli* strains JM109 and NM522 mutS were used for the general cloning procedure and for the first selection stage of mutagenesis, respectively. The transformants were grown at 30°C in a Luria-Bertani medium (LB) or on LB plates, both containing 50 mg/l of ampicillin. Subcloning was carried out with pUC119.

DNA sequence analysis. The DNA sequencing reaction was carried out by the dideoxy chain termination method with the AutoRead™ sequencing kit (Pharma-

---

* Part of this work was presented at 39th Symposium on the Chemistry of Natural Products (Sapporo, August 1997), Symposium Paper pp. 145-150. The nucleotide sequence data reported in this paper will appear under accession number AB007002.

† To whom correspondence should be addressed at the Faculty of Agriculture.

Abbreviations: SHC, squalene-hopene cyclase; OSC, oxidosqualene cyclase; IPTG, isopropyl-β-D-thiogalactopyranoside.

cia) and by cycle sequencing with the Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham). Automated sequencing was performed with an A.L.F.™ DNA sequencer (Pharmacia) and LIC-4000 (LI-COR).

Cloning shc and constructing the expression plasmid for SHC. Chromosomal DNA was isolated according to the procedure of Ochs *et al.*[1] Two primers were designed as follows (with the *Nde* I and *Bgl* II sites underlined): 5′-GAA TTC CAT ATG GCT GAG CAG TTG GTG GAA G-3′ and 5′-GAA TTC AGA TCT TCA CCT GCG CTC GAT GGC TTG-3′, which contained the sequences corresponding to amino acids 1–7 and 626–631 of *A. acidocaldarius* SHC, respectively. A polymerase chain reaction (PCR) was carried out by using these primers and the chromosomal DNA as a template. In a 50-μl PCR reaction, 300 ng of chromosomal DNA, 100 pmol of each primer and 1 mM of each dNTPs in a 1×*Pfu* DNA polymerase reaction buffer (Stratagene) were amplified with 2.5 U *Pfu* DNA polymerase for 30 cycles (98°C, 1 min; 68°C, 1 min; 72°C, 4 min). The resulting 1.92-kb PCR product was digested with *Nde* I and *Bgl* II, and ligated into the *Nde* I-*Bam* HI large fragment of pET3a to obtain an expression plasmid (pETSHC). The sequences of inserted DNA from the five clones were determined to inspect the *Pfu* DNA polymerase fidelity error.

Site-directed mutagenesis. The *shc* genes were mutated with oligonucleotide-directed *in vitro* mutagenesis by a Unique Site Elimination Mutagenesis kit (Pharmacia). Plasmid pKS, a pUC119 derivate containing a 1-kbp *Kpn* I–*Sac* I fragment, was used as a template for all mutagenesis reactions. The synthetic oligonucleotides used for the mutations (with changes to the wild-type sequence in bold type) were as follows:

5′-GAATATCCGT**T**GGAGAAGGTGCC**A**ATGGTC-3′ (W142L)
5′-GAATATCCGT**TC**GAGAAGGTGCC**A**ATGGTC-3′ (W142F)
5′-GAATATCCGTGG**C**AGAAGGTGCC**A**ATGGTC-3′ (E143Q)
5′-GAATATCCGTGGGA**C**AAGGTGCC**A**ATGGTC-3′ (E143D)
5′-CTCGCCGGTGT**TC**GACAC**A**GGCCTCG-3′ (W312F)
5′-CTCGCCGGTGTGG**A**ACAC**A**GGCCTCG-3′ (D313N)
5′-GGT**C**CCGGGC**A**ACTGGGCGGTGA-3′ (D350N)
5′-GGT**C**CCGGGCGA**A**TGGGCGGTGA-3′ (D350E)
5′-GGT**C**CCGGGCGACT**T**GGCGGTGA-3′ (W351L)
5′-GGT**C**CCGGGCGACT**TC**GCGGTGA-3′ (W351F)

Silent mutations were also executed to screen the desired mutants by a restriction fragment analysis (underlined in the primers described above). The *Nco* I restriction site was deleted for selection of the mutations of W142/E143, and *Stu* I and *Sma* I sites were created for the mutations of W312/D313 and D350/W351, respectively. To ascertain that the desired mutation had been carried out, the entire region of the inserted DNA was sequenced. Replacement of the *Kpn* I–*Sac* I fragment in the wild-type *shc* gene with the mutated *Kpn* I–*Sac* I fragments enabled the mutant pETSHCs to be obtained, this being confirmed by a restriction fragment analysis.

Expression of recombinant SHC. pETSHC was transformed into *E. coli* BL21(DE3). The cells were grown at 30°C to the mid-log phase (OD_{600}=0.6) in an LB medium containing 50 mg/l of ampicillin and then induced for 8 h with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG).

Protein purification. All purification steps were carried out at room temperature. The harvested cells were suspended in 50 mM Tris-HCl (pH 8.0) containing 1% Triton X-100, sonicated at 4°C for 10 min, and then subjected to centrifugation at 10,000×g for 15 min. The pH of the supernatant was adjusted to 6.0 (optimal activity), which was then heated at 60°C for 10 min to denature the *E. coli* proteins. After centrifugation, the supernatant was applied to a DEAE-Toyopearl 650 M column (Tosoh) which had been equilibrated with buffer A [10 mM sodium citrate (pH 6.0) containing 0.1% Triton X-100] and then thoroughly washed with buffer A. The wild-type and mutant SHCs were eluted with 50 mM sodium citrate (pH 6.0) containing 0.1% Triton X-100 and then dialyzed. Final purification was achieved by HPLC in a Mono Q column which had been equilibrated with 10 mM sodium citrate (pH 6.0). The adsorbed cyclase was eluted with a linear concentration gradient of sodium citrate from 0–1.0 M, enzyme activity being found at the concentration of 300 mM. All the purification steps were monitored by checking SDS-PAGE and the activities. Protein concentration was determined by the method of Lowry *et al.*[18]

Enzyme assay. The reaction mixture contained 60 mM sodium citrate (pH 6.0), 100 mM squalene, 0.2% Triton X-100, and 5 μg of the purified enzyme in a final volume of 5 ml. Incubation was conducted at 60°C for 5 min and terminated by adding 6 ml of 15% methanolic KOH. The substrate and products (hopene and hopanol) were extracted four times with 5 ml of hexane, and identified and quantified by GC in a DB-1 capillary column at 270°C.

Results and Discussion

Overexpression of A. acidocaldarius SHC in E. coli and the enzyme purification

In order to obtain a large quantity of SHC, we attempted to use the pET expression vector for high-level expression. The target gene cloned in the pET plasmid was successfully expressed by the strong T7 RNA polymerase in the host *E. coli* BL21(DE3), the expression of which is controlled by the IPTG-inducible *lac* UV5 promoter. The *shc* gene was isolated by the PCR method, which was supplemented with both an *Nde* I restriction site at the start ATG and a *Bgl* II site just after the stop codon. SHC expression plasmid pETSHC was constructed by introducing the digested *Nde* I–*Bgl* II fragment into the large *Nde* I–*Bam* HI fragment of pET3a. The sequences of inserted DNAs from the five clones were unambiguously identical, but our analyzed

sequence was very different from that reported by Ochs *et al.*[1] Very recently, the amino acid sequence has been revised in the GenBank (accession no. M73834). The alignment reported in the accession number was the same as ours, except for Ala600 and Ser601, which should be revised again to Gly and Phe, respectively, because the DNA sequence was altered from GCTTCC to GGCTTC. Thus, it can be concluded that eight QW motifs were involved in the entire SHC, differing from the seven motifs reported before.[1] SDS-PAGE of the cell-free extract showed the high expression of a 72-kDa protein at the level of 15–20% of the total *E. coli* proteins, as estimated by a densitometer. The SHC activity was also very high; the cell-free extract from 1 l of cultured cells (6.5 g) fully converted 150 mg of squalene to give hopene and hopanol in an isolation yield of 110 mg and 23 mg, respectively, while the metabolizing activity of *A. acidiocaldarius* was very little (only 4 mg of squalene was converted by using the 1 l culture).

The recombinant SHC was purified in three simple steps. Heat treatment of the cell-free extract was effective, because *E. coli* proteins could be denatured and removed without loss of activity due to the thermophilic nature. After two chromatographic procedures with DEAE-Toyopearl and Mono-Q columns, the enzyme was purified to homogeneity as shown by the SDS-PAGE data (Fig. 1). Starting from 6.5 g of cells, 10 mg of pure protein was obtained.

Construction, expression and purification of mutant SHCs

In order to search for the function of the aromatic and anionic residue, the conserved amino acids in two SHCs (W142, W312, W351, E143, D313 and D350)[1,2] were selected for the target of point mutations (Fig. 2); W351 and D350 are located inside the motif (called QW motif 5), which can only be found in SHCs and not in OSCs, whereas the residues of W142, W312, E143, D313 are highly conserved in both SHCs and OSCs but situated outside the QW motif (Fig. 2). The mutant SHCs that had been constructed as described in the Materials and Methods section were expressed at the same high level as the wild-type SHC and purified to homogeneity, giving a single band by SDS-PAGE corresponding to an apparent molecular mass of 72 kDa, identical to that of the wild type (Fig. 1).

Kinetic analyses of the wild-type and mutant SHCs

Table I shows the apparent kinetic parameters of the purified wild-type and mutant SHCs determined by Lineweaver-Burk plots. The K_m and V_{max} values for the wild-type SHC are in agreement with those reported by Feil *et al.* ($K_m = 16.7\ \mu M$, $V_{max} = 4.27$ nmol/min/μg).[15] Substitution of W351 or D350 with L or N, respectively, completely inactivated the cyclase activity, indicating that the aromatic and carboxylic moieties inside the QW motif 5 were critical to SHC catalysis. This is the first demonstration of one of the eight QW motifs (QW 5) occupying a crucial active site.

The kinetic values for all W mutants (W142, W312 and W351) revealed the following two results: (1) K_m for the W→F mutation increased (3–8 fold), whereas V_{max} remained unchanged; (2) K_m for the W→L mutants markedly increased (40 fold or unmeasurably large in the QW motif) when compared with the value for the W→F mutants. The changes in K_m value obtained from these

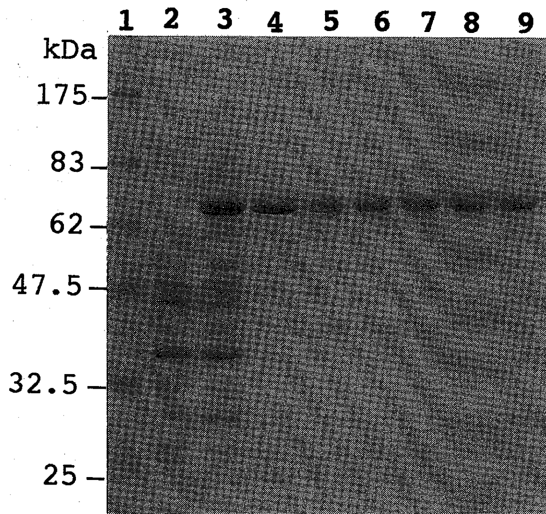


Fig. 1. Expression and Purification of Wild-type and Mutant SHCs Monitored by SDS-PAGE.

Wild-type SHC after each purification step (lanes 3–5) and mutant SHCs after Mono-Q HPLC (lanes 6–9) were electrophoresed on a 10% polyacrylamide gel and stained with Coomassie Brilliant Blue R-250. Lane 1, molecular weight standards; lane 2, cell-free extracts from *E. coli* BL21(DE3)/pET3a; lane 3, cell-free extracts from *E. coli* BL21(DE3)/pETSHC; lane 4, after heat treatment at 60°C for 10 min; lane 5, after Mono-Q HPLC; lane 6, D350N; lane 7, D350E; lane 8, W351L; lane 9, W351F. Other mutants also showed a single band by SDS-PAGE.

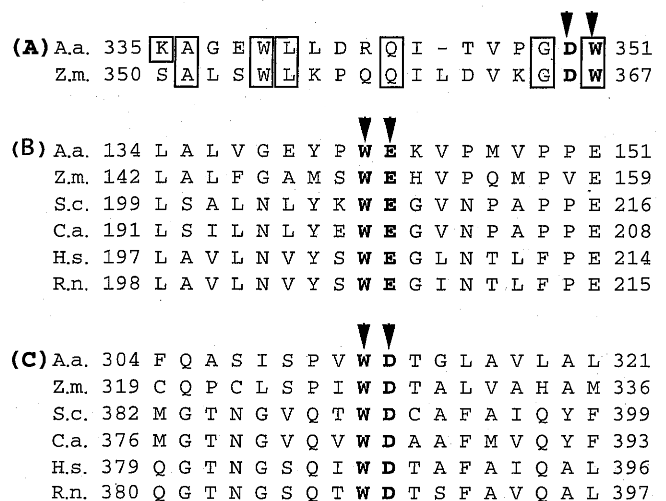


Fig. 2. Target Amino Acid Residues (W, D and E) for the Site-directed Mutagenesis. Arrows show the targeted amino acid residues.

(A) QW motif 5 involved in SHCs, eight QW motifs being present in SHC from *A. acidocaldarius*. (B) Conserved W and E located outside the QW motifs. (C) Conserved W and D outside the QW motifs. Frequently occurring residues in the QW motifs are boxed. A.a., *A. acidocaldarius* SHC; Z.m., *Zymomonas mobilis* SHC (ref. 2); S.c., *Saccharomyces cerevisiae* OSC (refs. 5 and 6); C.a., *Candida albicans* OSC (refs. 3 and 4); H.S., *Homo sapiens* OSC (refs. 7 and 8); R.n., *Rattus norvegicus* OSC (refs. 9 and 10).

Table I. Kinetic Parameters of the Wild-type and Mutant Squalene-hopene Cyclase.

SHC	K_m (μM)	V_{max} (nmol/min/μg)	V_{max}/K_m ($\times 10^2$)	Relative activity (%)
Wild-type	16.5	4.03	24.42	100
D350N	37.4	0.06	0.16	0.7
D350E	79.2	3.08	3.89	15.9
W351L*	—	—	—	0
W351F	63.5	3.85	6.06	24.8
W142L	652.7	3.87	0.59	2.4
W142F	136.8	3.65	2.67	10.9
E143Q	16.3	3.78	23.19	95.0
E143D	16.3	3.97	24.36	99.8
W312F	50.7	4.01	7.91	32.4
D313N	16.6	0.40	2.41	9.9

* No activity was detected with a large quantity of the protein (1.5 mg), comparable to a 300-fold increment with the usual assays (5 μg).

mutations suggest that strong affinity to the substrate can be induced by the aromatic indole ring of tryptophan, but not by the hydrophobicity of aliphatic side chains such as leucine. The K_m value increased in the replacement of W with F, possibly arising from the decreased π-electron density or distribution of phenylalanine than of tryptophan. Polycyclization proceeds *via* the discrete carbocation intermediates.[19,20] If the cation can be stabilized, the catalytic reaction would be accelerated. Assuming that tryptophan would act for cation stabilization as proposed by Poralla or Griffin,[6,14] V_{max} of the mutants (W→F) would be small due to the decreased π-electron density of phenylalanine compared with that of tryptophan, but neither acceleration nor reduction was observed (Table I). Therefore, tryptophan residues probably may function as the binding sites for squalene. The driving force for the molecular recognition caused by tryptophan might be CH-π interaction with methyl groups[21] and/or π-π stacking with double bonds of the squalene substrate. For CH-π interaction with methyl groups, indole ring is known to be stronger than the phenyl ring.[21,22] The C–H bond of the methyl groups would be more strongly polarized due to the adjacently located carbocations produced during the cyclization, resulting in stronger CH-π interaction with the indole rings.[21,22]

Regarding aspartic acid, the kinetic values for D313 and D350 were changed, whereas E143 remained unchanged. The E143 residue was not important for SHC activity. V_{max} for D→N was markedly decreased, whereas K_m changed little. V_{max} for the D→E mutants remained unchanged. These results strongly suggest that the carboxyl groups may serve as anionic sites to stabilize the carbocation intermediates through the ionic bonds.

This is the first report demonstrating by kinetic studies that the tryptophan and aspartic acid residues in the QW motif are critical to the cyclase activity. Further studies on other point mutations are in progress to provide better knowledge about the polycyclization mechanism.

Acknowledgments

We thank Prof. T. Ono and Dr. J. Sakakibara of School of Medicine, and Prof. T. Watanabe and Dr. N. Nikaidou, of Faculty of Agriculture at this university, for providing the technical information on molecular cloning. Part of this work was supported by a Grant-in-Aid to T. H. (No. 09660111) from the Ministry of Education, Science and Culture.

Addendum

After submitting this paper, an X-ray analysis of the hopene cyclase has been reported.[23] The X-ray data indicate that the repetitive QW motifs constitute the secondary structure and connect all helices, thus contributing to stabilization of the whole protein structure. It was reported that the side chains of the tryptophan and glutamine moieties inside the QW motifs were stacked. It is probable that the mutation of the QW motif 5 may have carried an alteration to the secondary structure; therefore, our conclusion may be erroneous that the tryptophan and aspartic acid moieties inside QW motif 5 were the active sites. However, other point mutations conducted by us have led to the similar result that the tryptophan moiety was critical to the enzyme activity; the mutations of W312L, W489L and W169V almost or completely inactivated the cyclizing function, although these tryptophan residues were located outside QW motifs. It should be also noted that W489 is near to W485 of QW motif 3, the intervening distance being only three peptidic bonds. In the case of W312L, K_m, V_{max} and V_{max}/K_m ($\times 10^2$) were estimated to be 869 μM (53-fold increase, compared to the wild type) and 0.46 nmol/min/μg (1/9-fold reduction) and 0.2 (relative activity), respectively; a very large K_m value suggests that the tryptophan moiety works as a recognition site for the squalene substrate. The X-ray analysis of the enzyme-inhibitor complex showed that W312 was situated near D374, D376 and D377, which were proposed as being responsibile for the initial protonation reaction.[23] The mutants of W489L and W169V completely lost enzyme activity, so that the kinetic values could not be measured, as was found with the mutant of W351L (Table 1). It is thus likely that these tryptophan moieties were active sites. As shown in the X-ray analysis of the enzyme-inhibitor complex,[23] these tryptophan moieties were in close contact with the inhibitor. This paper is the first to demonstrate that tryptophan residues were critical to the cyclization reaction.

References

1) D. Ochs, C. Kaletta, K-D. Entian, A. Beck-Sickinger, and K. Poralla, *J. Bacteriol.*, **174**, 298–302 (1992).
2) I. G. Reipen, K. Poralla, H. Sahm, and G. A. Sprenger, *Microbiology*, **141**, 155–161 (1995).
3) C. J. Buntel and J. H. Griffin, *J. Am. Chem. Soc.*, **114**, 9711–9713 (1992).
4) C. A. Roessner, C. Min, S. H. Hardin, L. W. Harris-Haller, J. C. McCollum, and A. I. Scott, *Gene*, **127**, 149–150 (1993).
5) E. J. Corey, S. P. T. Matsuda, and B. Bartel, *Proc. Natl. Acad. Sci. USA.*, **91**, 2211–2215 (1994).
6) Z. Shi, C. J. Buntel, and J. H. Griffin, *Proc. Natl. Acad. Sci. USA*, **91**, 7370–7374 (1994).
7) E. J. Corey, S. P. T. Matsuda, C. H. Baker, A. Y. Ting, and H. Cheng, *Biochem. Biophys. Res. Commun.*, **219**, 327–331 (1996).
8) E. J. Corey, S. P. T. Matsuda, and B. Bartel, *Proc. Natl. Acad.*

Sci. USA, **90**, 11628–11632 (1993).
9) M. Kusano, M. Shibuya, U. Sankawa, and Y. Ebizuka, *Biol. Pharm. Bull.*, **18**, 195–197 (1995).
10) I. Abe and G. D. Prestwich, *Proc. Natl. Acad. Sci. USA*, **92**, 9274–9278 (1995).
11) C. K. Sung, M. Shibuya, U. Sankawa, and Y. Ebizuka, *Biol. Pharm. Bull.*, **18**, 1459–1461 (1995).
12) C. H. Baker, S. P. T. Matsuda, D. R. Liu, and E. J. Corey, *Biochem. Biophys. Res. Commun.*, **213**, 154–160 (1995).
13) W. S. Johnson, S. D. Lindell, and J. Steele, *J. Am. Chem. Soc.*, **109**, 5852–5853 (1987).
14) K. Poralla, *Bioorg. Med. Chem. Lett.*, **4**, 285–290 (1994).
15) C. Feil, R. Sussmuth, G. Jung, and K. Poralla, *Eur. J. Biochem.*, **242**, 51–55 (1996).
16) A. H. Rosenberg, B. N. Lade, D. Chui, S-W. Lin, J. J. Dunn, and F. W. Studier, *Gene*, **56**, 125–135 (1987).
17) G. Darland and T. D. Brock, *J. Gen. Microbiol.*, **67**, 9–15 (1971).
18) O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, **193**, 265–275 (1951).
19) E. J. Corey, S. C. Virgil, H. Cheng, C. H. Baker, S. P. T. Matsuda, V. Singh, and S. Sarshar, *J. Am. Chem. Soc.*, **117**, 11819–11820 (1995).
20) T. Hoshino, E. Ishibashi, and K. Kaneko, *J. Chem. Soc. Chem. Commun.*, **1995**, 2401–2402.
21) M. Nishio and M. Hirota, *Tetrahedron*, **45**, 7201–7245 (1989).
22) M. Nishio, Y. Umezawa, and M. Hirota, *Kagaku to Seibutsu* (in Japanese), **33**, 311–318 (1995).
23) K. U. Wendt, K. Poralla, and G. E. Schulz, *Science*, **277**, 1811–1815 (1997).