ORIGINAL ARTICLE

The Complete Biosynthetic Gene Cluster of the 28-Membered Polyketide Macrolactones, Halstoctacosanolides, from *Streptomyces halstedii* HC34

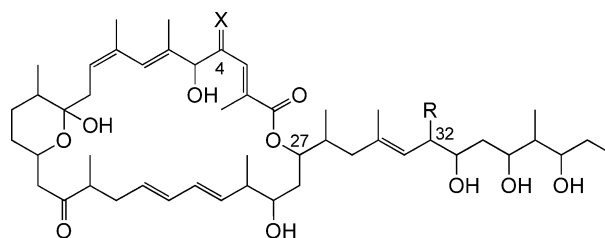**Shigehiro Tohyama, Katsumi Kakinuma, Tadashi Eguchi**

Abstract Halstoctacoanolides A and B are 28-membered polyketide macrolactones and were isolated from *Streptomyces halstedii* HC34. The biosynthetic gene cluster (*hls* cluster) of halstoctacosanolides was completely identified from the genome library of *Streptomyces halstedii* HC34. DNA sequence analysis of *ca.* 100 kb region revealed that there were seven type I polyketide synthases (PKSs) and two cytochrome P450 monooxygenases in this cluster. Involvement of the gene cluster in the halstoctacosanolide biosynthesis was demonstrated by the gene disruption of P450 monooxygenase genes. The mutants produced a new deoxygenated halstoctacosanolide derivative, halstoctacosanolide C, which confirmed that the *hls* gene cluster was essential for the biosynthesis of halstoctacosanolides.

Keywords halstoctacosanolides, biosynthetic gene cluster, polyketide synthase, P450 monooxygenase, gene disruption

Introduction

Polyketide macrolides are one of the largest bioactive natural product families. They have various complex structures and a huge variety of bioactivities have been reported for them. Three examples are the antibiotic erythromycin, immunosuppressant rapamycin and the antitumor agent epothilone, and there are many other polyketide macrolides that are clinically important [1]. Polyketide chains are generally biosynthesized in a similar manner by the multi-modular enzymes, type I polyketide synthases (PKSs) [1]. Each module is responsible for one elongation of an extender unit and modification of the β-carbon by coordinated domains. In generally, the polyketide structures well reflect the domain structure of its own biosynthetic PKSs. Because these domain structures are easily assigned from the gene information, it is now possible to predict the polyketide structure biosynthesized by PKSs when once the DNA sequences of the genes encoding PKSs are determined. Based on this context, the predictions of the polyketide productivity and the searches of new polyketide compounds from organisms having the PKS genes, especially from the major secondary metabolite producer actinomycetes, have been accomplished [2~6]. Molecular genetic studies of the actinomycetes, thus, present promising opportunities for the discovery of a new polyketide compound potentially useful for drug development.

T. Eguchi (Corresponding author), **S. Tohyama, K. Kakinuma:** Department of Chemistry & Materials Science, and Department of Chemistry, Tokyo Institute of Technology, 2-12-1, O-okayama, Meguro-ku, Tokyo152-8551, Japan.
Email: eguchi@cms.titech.ac.jp

Halstoctacosanolide A; X = O, R = OH
Halstoctacosanolide B; X = H$_2$, R = OH
Halstoctacosanolide C; X = H$_2$, R = H

Fig. 1 Structures of halstoctacosanolides.

Halstocacosanides A and B are polyketides which have a 28-membered macrolactone structure and were discovered in the genome of *Streptomyces halstedii* HC34 using the strategy outlined in the previous paragraph. Subsequent isolation and evaluation of these compounds demonstrated that they possess moderate antimicrobial activity against several microorganisms [3]. However, the complete gene cluster had not been cloned. To confirm the complete biosynthetic gene cluster and in an effort to generate halstoctacosanolide analogues through genetic engineering, *ca.* 100 kb segment of DNA containing the putative halstoctacosanolide (*hls*) biosynthetic gene cluster from *Streptomyces halstedii* HC34 was cloned and sequenced. In this paper, we report the cloning and characterization of the complete gene cluster for halstoctacoanolide biosynthesis. Also described are isolation and structure determination of a new halstoctacosanolide derivative produced by gene disruption of cytochrome P450 monooxygenases in the *hls* cluster.

Results and Discussion

Cloning and Sequencing of Complete Halstoctacosanolide Biosynthetic Gene Cluster

Previously, three cosmids (7-4F, 3-8A and 10-10H) including total *ca.* 45 kb DNA sequence had been cloned for the halstoctacosanolide biosynthetic gene and the DNA sequence contain three open reading flames (ORFs) encoding PKS [3]. To identify the complete gene cluster, genomic walking using a *S. halstedii* HC34 genome library was curried out by standard hybridization method. Based on the known DNA sequences, DIG labeled probes were prepared for the hybridization. Using the sequence of most downstream region as a probe, the next overlapped cosmid was cloned. By repeating this procedure for several times, three cosmids (18-9Q, 22-5B and 22-7D) were newly identified, and *ca.* 60 kb of contiguous DNA was sequenced. There are nine ORFs in the whole *hls* cluster, which encode seven PKS genes and two cytochrome P450 monooxygenases as shown in Figure 2. The DNA sequence has been deposited in the DNA Data Base of Japan (DDBJ) under the accession number AB241068.

Table 1 Summary of the *hls* genes and comparison with database

| Gene	Size (aa)	Proposed function	Closest homologue (gene and origin)		Identity/similarity	Accession No.
hlsA	4,194	PKS (module L, 1 and 2)	*oleAI*	*Streptomyces antibioticus*	48/60%	AAF82408
hlsB	7,111	PKS (module 3~6)	ORF13	*Streptomyces aizunensis*	52/65%	AAX98188
hlsC	5,220	PKS (module 7~9)	ORF13	*Streptomyces aizunensis*	52/64%	AAX98188
hlsD	5,322	PKS (module 10~12)	ORF9	*Streptomyces aizunensis*	56/68%	AAX98184
hlsE	3,809	PKS (module 13, 14)	*fscE*	*Streptomyces* sp. FR-008	50/61%	AAQ82567
hlsF	5,232	PKS (module 15~17)	ORF14	*Streptomyces aizunensis*	53/65%	AAX98189
hlsG	2,099	PKS (module 18)	*conD*	*Streptomyces neyagawaensis*	53/64%	AAZ94389
hlsH	404	Cytochrome P450 monooxygenase	*cyp881*	*Streptomyces peucetius*	42/58%	CAE53712
hlsI	414	Cytochrome P450 monooxygenase	*nocL*	*Nocardia* sp.	44/58%	AAT09797
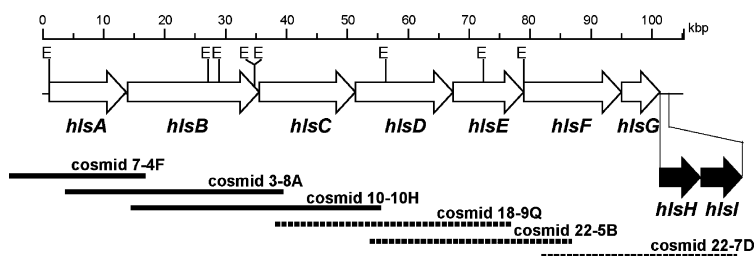
Fig. 2 Organization of the *hls* cluster in *S. halstedii* HC34.

Each arrow represents ORF, open arrow; PKS, closed arrow; cytochrome P450 monooxygenase. The cosmid clones are denoted as overlapping lines and the cosmids cloned in the present study are shown in dashed line. E, *Eco*RI restriction site.
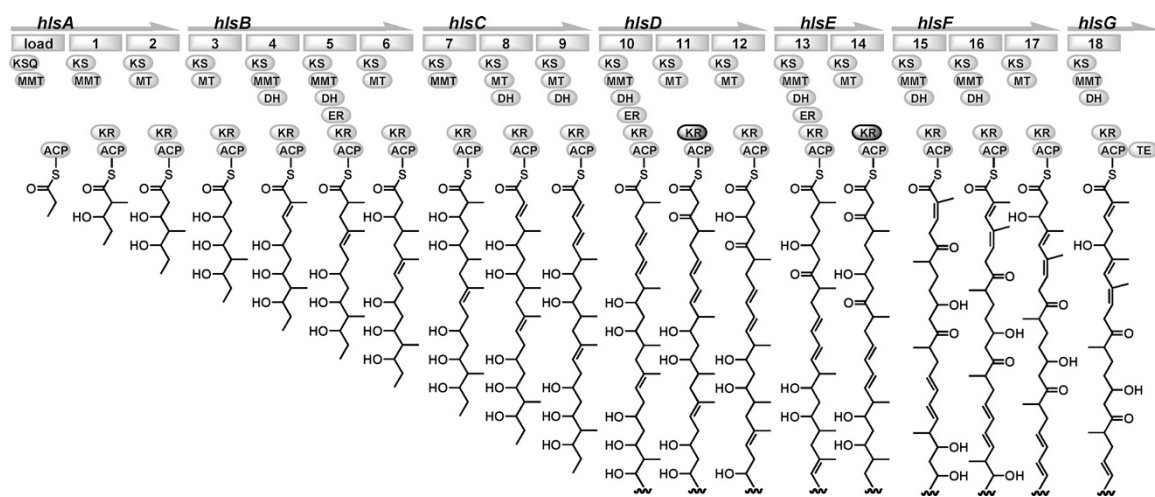
Fig. 3 Modular structure of the halstoctacosanolide PKSs and intermediates in polyketide chain biosynthesis.

Each circle represents an enzymatic domain in the PKS multifunctional proteins. MT, malonyl transferase; MMT, methylmalonyl transferase; ACP, acyl carrier protein; KS, ketoacyl-ACP synthase; KSQ, inactive KS; KR, ketoacyl ACP reductase; DH, hydroxylthioester dehydratase; ER, enoyl reductase. Inactive KR domains are shown in dark gray at module 11 and 14.

PKS Genes

There are seven PKS genes, *hlsA* to *hlsG*, containing eighty-eight catalytic domains organized into one loading and eighteen extender modules in a *ca.* 100 kb region of the *hls* gene cluster. The number of modules is same as the number of the extender units, which is required for the construction of halstoctacoanolides. These genes are arranged straight forwardly head to tail, and there is no gene encoding non-PKS as seen in the rifamycin biosynthetic gene cluster [7]. When the chain elongation reactions sequentially proceed from upstream to downstream, the PKS gene structures well reflect the structure of polyketide chain of halstoctacosanolides except for a ketonic moiety at C-17 and an acetal function at C-11 as shown in Figure 2. These functional groups are presumably unreduced by the KR modules and remain in the carbonyl functions during the elongation steps by PKSs. However, since all modules contain KR domain, certain contradiction appeared between the gene structure and the chemical structure. Some inactive KR domains have been reported to lack the active site motif, GXGXXGXXXA for NADP-binding or the catalytic triad, Lys, Ser and Tyr in the active site [8]. However, in the case of *hls* KR, all the important amino acid residues appeared to be conserved in all KR. Although the exact reason is not clear at the moment, a similar case was reported in avermectin PKSs, module 10 [9].

Halstoctacoanolides A and B have one trisubstituted *Z* olefin at C-8 position. Usually, most of the double bond geometry found in polyketide chains is *E*-configuration. It has been suggested that the stereochemistry of β-hydroxyl group derived from KR domain might control the double bond geometry in DH domain [8]. An Asp residue of KR domain was proposed to be an important residue for determining the stereospecificity. However, in our case, the Asp residue is found in all KR domains by multiple alignments. It is known that there are a few exceptions to this hypothesis [10], and in the *hls* gene cluster this is also not the case.

Two Genes Encoding Cytochrome P450

Halstoctacosanolide A has two additional oxygen functional groups (C-4 and C-32) rather than those derived from the usual PKSs chain elongation pathway, which are considered to be introduced after the polyketide chain elongation. Accordingly, it had been expected that there are several genes encoding the oxidation enzymes. In fact, two genes, *hlsH* and *hlsI*, were identified to encode cytochrome P450 monooxygenases in the downstream region of the PKS genes. In the rapamycin biosynthesis, the methylene group at C-9 of rapamycin polyketide chain was reported to be oxidized to the carbonyl group by a single cytochrome P450 enzyme, RapJ [11]. Therefore, it seemed likely that a single enzyme, either HlsH or HlsI, catalyzes the oxidation of methylene group at C-4 to carbonyl group, and another may be responsible for the hydroxylation at C-32. However, since the amino acid sequences of HlsH and HlsI are quite similar (identity 42%) and belong to the same CYP family CYP107 found in secondary metabolism [12, 13], it was not easily predictable which enzyme is responsible to either of oxidations. Therefore, to determine the function of each gene, the gene disruption experiments were performed.
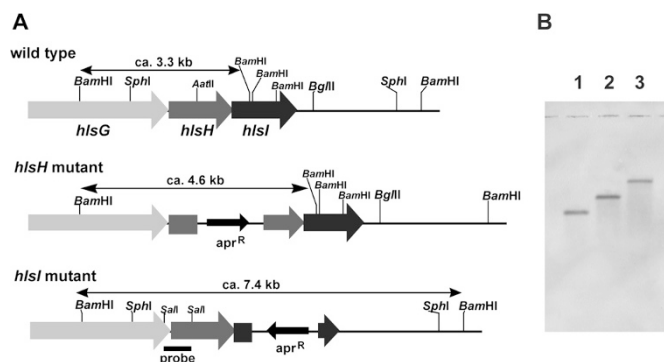
Fig. 4 Targeted disruption of the *hlsH* and *hlsI* genes.

A: Genome appearances of wild type, *hlsH* and *hlsI* mutants and the expected *Bam*HI fragment of each genome DNA. B: Analysis of chromosomal DNA by Southern hybridization with *ca*. 500 bp *Sal*I fragment probe. Genome DNA fragments digested by *Bam*HI from lane 1; wild type, lane 2; *hlsH* mutant, and lane 3; *hlsI* mutant.

Furthermore, the gene disruption experiments for these cytochrome P450 monooxygenases, *hlsH* and *hlsI*, could be expected to generate a new halstoctacosanolide analogue.

Disruptions of the Genes Encoding Cytochrome P450

The gene disruption experiments were carried out by gene insertional inactivation according to the previously reported method for the disruption of the *vin* genes in the same strain [14]. The apramycin resistant gene cassette was thus inserted into the *Aat*II site that existed in *hlsH* and *Bam*HI sites found in *hlsI*. Each fragment was subcloned into *E. coli - Streptomyces* shuttle vector pWHM3. The resulting plasmids were separately transformed into protoplasts of *S. halstedii* HC34 and the clones showing proper phenotype were selected. The gene disruptions of the resulting mutants, which showed apramycin resistance, were confirmed using Southern hybridization against *Bam*HI fragments of *S. halstedii* genomic DNA, and the expected sizes of *Bam*HI fragments (wild type; *ca*. 3.3 kb, *hls*H mutant; *ca*. 4.6 kb, *hls*I mutant; *ca*. 7.4 kb) appeared as shown in Figure 4.

The fermentation broths of the wild type and two mutants were separately extracted with ethyl acetate, and the TLC analysis of the resulting extracts was performed as shown in Figure 5. The *hlsH* mutant was found to lose the productivity of halstoctacosanolide A and to produce only halstoctacosanolide B. In the *hlsI* mutant, neither halstoctacosanolide A or B was produced at all. It should be noted that both mutants appeared to produce a new less polar compound. Thus, the isolation study of this compound was pursued.
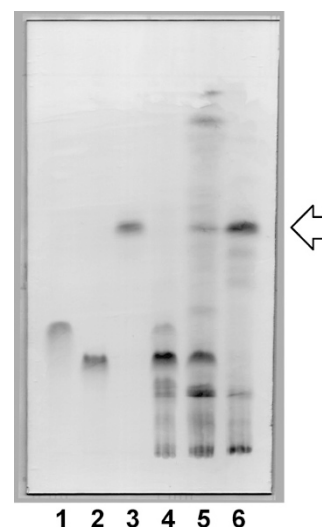

Fig. 5 TLC analysis of halstoctacosanolides and the fermentation broths of wild type, *hlsH* and *hlsI* mutants.

Lane 1~3: sample of halstoctacosanolides A, B, and C, respectively; lane 4~6: EtOAc extracts from the fermentation broths of wild type, *hlsH* and *hlsI* mutants, respectively. TLC was developed three times with ethyl acetate. Arrow indicates the position of halstoctacosanolide C.

Isolation and Structure Determination of a New Halstoctacosanolide Analogue

In the TLC analysis, the *hlsI* mutant seemed to produce a relatively large amount of the new compound. Thus, isolation study was carried out from the fermentation broth of the *hlsI* mutant according to the previous method for halstoctacosanolides [3]. Forty-two mg of the new compound, named halstoctacosanolide C, was successfully isolated from 1.5 liters culture. The physico-chemical properties of halstoctacosanolide C are summarized in

Table 2. The molecular formula was established as $C_{48}H_{78}O_{10}$ on the basis of HRFAB-MS, which corresponds to a deoxygenated derivative of halstoctacosanolide B. In UV spectrum, the maximum absorption wavelength and the molar absorption coefficient were also quite similar to those of halstoctacosanolide B. However, the structure determination of halstoctacosanolide C was not straightforward because halstoctacosanolide C was obtained as an inseparable mixture of two components. The same phenomenon was observed in the case of halstoctacosanolide B, which suggested that halstoctacosanolide C also exists in an equilibrium of the hemiacetal and seco-ketone forms. Therefore, methyl

Table 2 Physico-chemical properties of halstoctacosanolide C

Halstoctacosanolide C		
Appearance		White powder
MP		63~65°C
$[\alpha]_D^{27}$		−6.5 (c 1.36, $CHCl_3$)
Molecular formula		$C_{48}H_{78}O_{10}$
HRFAB-MS	Calcd:	837.5493 [M+Na]$^+$
	Found:	837.5521
UV λ_{max} (ε) (in MeOH)		226 (48,000)
IR ν (neat) cm^{-1}		3410 (br), 2960, 2930, 1700

Table 3 NMR data of halstoctacosanolide C methyl acetal, halstoctacosanolide A, and halstoctacosanolide B methyl acetal in $CDCl_3$
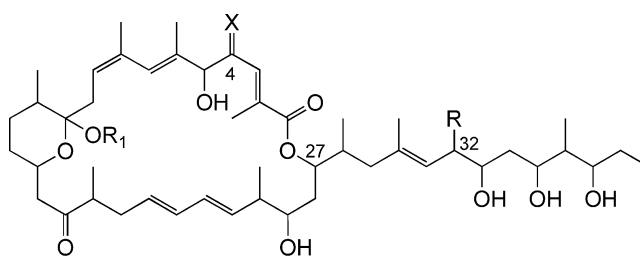
X = O, R = OH, R$_1$=H; Halstoctacosanolide A
X = H$_2$, R = OH, R$_1$=Me; Halstoctacosanolide B methyl actal
X = H$_2$, R = H, R$_1$=Me; Halstoctacosanolide C methyl actal

Position	Halstoctacosanolide C methyl acetal		Halstoctacosanolide A	Halstoctacosanolide B methyl acetal
	δ_C	δ_H (multiplicity, J in Hz)	δ_C	δ_C
1	167.98 (s)		165.79 (s)	167.74 (s)
2	129.54 (s)		145.50 (s)	129.38 (s)
3	137.12 (d)	6.70 (ddd, 1.2, 6.4,7.9)	126.70 (d)	136.93 (d)
4	35.44 (t)	2.54/2.56 (m)	198.87 (s)	34.80 (t)
5	75.28 (d)	4.29 (t, 6.1)	82.38 (d)	74.81 (d)
6	138.75 (s)		133.71 (s)	138.49 (s)
7	125.29 (d)	5.94 (br s)	129.40 (d)	125.23 (d)
8	133.06 (s)		134.92 (s)	133.25 (s)
9	123.28 (d)	5.32 (dd, 6.2, 7.3)	122.56 (d)	122.66 (d)
10	34.06 (t)	2.04/2.25 (m)	39.91 (t)	33.58 (t)
11	101.30 (s)		98.92 (s)	101.21 (s)
12	36.37 (d)	1.61 (m)	35.71 (d)	35.93 (d)
13	27.36 (t)	1.40/1.55 (m)	27.42 (t)	27.16 (t)
14	31.91 (t)	1.26/1.60 (m)	32.34 (t)	31.74 (t)
15	67.75 (d)	4.03 (ddt, 9.5, 11.8, 2.6)	67.51 (d)	67.69 (d)
16	47.96 (t)	2.23 (m)/2.72 (dd, 9.9, 14.1)	48.15 (t)	47.80 (t)
17	212.55 (s)		215.38 (s)	213.19 (s)
18	47.74 (d)	2.68 (sextet, 6.8)	48.05 (d)	47.80 (d)
19	35.04 (t)	2.05/2.52 (m)	34.75 (t)	34.98 (t)
20	129.67 (d)	5.55 (dt, 14.2, 7.2)	130.68 (d)	129.68 (d)

Table 3 (Continued)

| Position | Halstoctacosanolide C methyl acetal | | Halstoctacosanolide A | Halstoctacosanolide B methyl acetal |
	δ_C	δ_H (multiplicity, J in Hz)	δ_C	δ_C
21	132.40 (d)	5.99 (m)	131.47 (d)	132.55 (d)
22	129.96 (d)	5.95 (m)	129.36 (d)	130.10 (d)
23	135.74 (d)	5.43 (dd, 7.8, 14.2)	136.22 (d)	135.63 (d)
24	42.90 (d)	2.06 (m)	41.14 (d)	43.60 (d)
25	71.47 (d)	3.20 (ddd, 1.6, 6.4, 9.9)	68.72 (d)	72.31 (d)
26	34.99 (t)	1.52/1.65 (m)	31.98 (t)	33.74 (t)
27	73.93 (d)	5.11 (ddd, 1.4, 3.8, 10.9)	75.73 (d)	73.94 (d)
28	34.03 (d)	2.05 (m)	31.68 (d)	33.00 (d)
29	44.11 (t)	2.04/2.25 (m)	44.93 (t)	44.66 (t)
30	135.66 (s)		136.15 (s)	137.69 (s)
31	122.92 (d)	5.29 (t, 7.5)	128.13 (d)	126.99 (d)
32	36.35 (t)	2.19/2.26 (m)	72.25 (d)	72.31 (d)
33	69.38 (d)	3.90 (m)	71.44 (d)	71.61 (d)
34	41.32 (t)	1.44/1.83 (m)	38.78 (t)	38.13 (t)
35	73.56 (d)	4.17 (dt, 9.9, 2.4)	73.77 (d)	73.29 (d)
36	41.32 (d)	1.43 (m)	40.63 (d)	41.19 (d)
37	78.25 (d)	3.78 (ddd, 1.8, 5.5, 8.1)	79.08 (d)	77.83 (d)
38	28.07 (t)	1.46/1.53 (m)	28.08 (t)	27.88 (t)
39	10.40 (q)	0.92 (t, 7.6)	10.44 (q)	10.48 (q)
2-CH$_3$	12.72 (q)	1.84 (d, 1.2)	14.95 (q)	12.75 (q)
6-CH$_3$	14.66 (q)	1.56 (d, 1.1)	14.74 (q)	14.60 (q)
8-CH$_3$	24.11 (q)	1.75 (s)	24.31 (q)	24.22 (q)
12-CH$_3$	16.12 (q)	0.79 (d, 6.2)	16.13 (q)	16.26 (q)
18-CH$_3$	15.08 (q)	1.05 (d, 7.0)	16.46 (q)	14.62 (q)
24-CH$_3$	15.14 (q)	1.02 (d, 6.8)	11.20 (q)	16.15 (q)
28-CH$_3$	14.90 (q)	0.87 (d, 6.5)	14.23 (q)	14.62 (q)
30-CH$_3$	15.99 (q)	1.66 (s)	16.28 (q)	16.57 (q)
36-CH$_3$	4.94 (q)	0.91 (d, 7.1)	4.23 (q)	5.00 (q)
11-OCH$_3$	47.51 (q)	3.06 (s)		47.42 (q)

acetal derivatization was performed as in the case of halstoctacosanolide B. Thus, halstoctacosanolide C was treated with methanol under acidic condition and the methyl acetal derivative of halstoctacosanolide C was smoothly obtained. Since NMR spectra of methyl acetal derivative of halstoctacosanolide C at room temperature caused line broadening, all measurements were carried out at 50°C. ^{1}H and ^{13}C NMR spectra of methyl acetal derivative of halstoctacosanolide C were quite similar to those of methyl acetal derivative of halstoctacosanolide B except for disappearance of one oxymethine and appearance of one methylene (Table 3). Further analysis including ^{1}H-^{1}H COSY, HMQC and HMBC experiments established that halstoctacosanolide C is the C-32 deoxygenated derivative of halstoctacosanolide B.

These results clearly revealed the halstoctacosanolide biosynthetic pathway as follows. Firstly, seven PKSs (HlsA to G) generate halstoctacosanolide C from nine malonyl CoAs and ten methylmalonyl CoAs, and then the methylene group at C-32 is hydroxylated by HlsI to produce halstoctacosanolide B. Finally, halstoctacosanolide A is produced through oxidation of methylene group at C-4 of halstoctacosanolide B into carbonyl by HlsH as shown in Figure 6. By the TLC experiments, the *hlsH* mutant appeared to produce halstoctacosanolide B with a small amount of halstoctacosanolide C. This may be due to a downstream polar effect caused by the gene disruption of *hlsH*, and the transcriptional disorder of the *hlsI* gene may be occurred.

In the present study, we successfully cloned and sequenced the complete gene cluster for halstoctacosanolide biosynthesis, and revealed that seven ORFs encoded type I
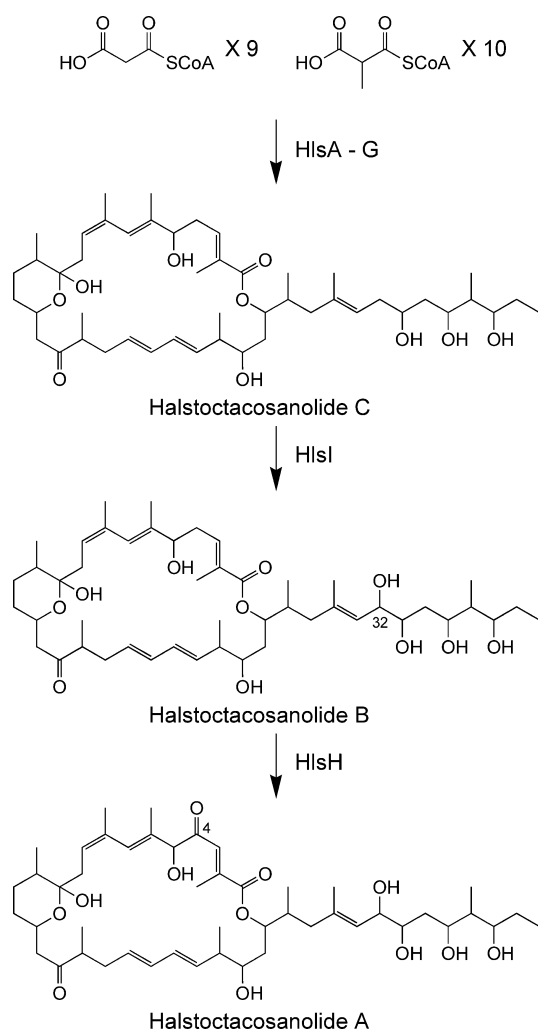
Fig. 6 Biosynthetic pathway of halstoctacosanolides in *S. halstedii* HC34.

PKS and two ORFs encoded P450 monooxygenase in this *hls* cluster. Involvement of the *hls* gene cluster in halstoctacosanolide biosynthesis was confirmed by the gene disruption of *hlsH* and *I*. Furthermore, we successfully generated a new halstoctacosanolide analogue through gene manipulations.

Experimental

General

Optical rotation was measured with a JASCO DIP-360 spectrometer. Mass spectra were obtained on a JEOL JMS-700 in the positive FAB mode using 3-nitrobenzyl alcohol as a matrix. UV and IR spectra were recorded on a Shimadzu UV-160A spectrophotometer and Horiba FT-710 spectrometer, respectively. NMR spectra were obtained by a

JEOL LA-400 and/or a Bruker DRX-500 spectrometers. Chemical shifts are reported in δ value relative to internal tetramethylsilane (δ 0.00). Silica gel column chromatography was carried out with a Kieselgel 60 (70~230 or 230~400 mesh, Merck) and preparative thin-layer silica gel chromatography was performed on PLC plates (silica gel 60 F254, Merck).

Bacterial Strains, Plasmids, and Culture Conditions

Halstoctacosanolide producer *S. halstedii* HC34 was used as the source of DNA of the genome library and disruption experiments. For halstoctacosanolide production, wild type and mutants were cultivated in vicenistatin production medium [3] (3% starch, 1.5% soya flake, 0.2% yeast extract, 0.5% corn steep liquor, 0.3% NaCl, 0.05% $MgSO_4 \cdot 7H_2O$, 0.0005% $CoCl_2 \cdot 6H_2O$ and 0.3% $CaCO_3$, at pH 7.1) for 3 days at 28°C with shaking. *S. lividans* TK24 [15] and *E. coli - Streptomyces* shuttle vector pWHM3 [16] were used to prepare plasmids for gene disruption and *S. lividans* was cultivated in SK No. 2 medium (2% starch, 0.5% glucose, 0.5% yeast extract, 0.3% peptone, 0.3% beef extract, 0.02% KH_2PO_4 and 0.06% $MgSO_4 \cdot 7H_2O$) at 28°C. Transformants were selected in R5 agar plate [15] with appropriate antibiotic. *E. coli* DH5a (TaKaRa) and plasmid pUC119 (TaKaRa), pBluescript II SK(+) (Stratagene) and LITMUS28 (New England Biolabs) was used for routine subcloning of DNA fragments. *E. coli* was grown at 37°C in Luria-Bertani (LB) broth or on LB agar supplemented with the appropriate antibiotics for selection.

DNA Manipulations and Southern Hybridization

Isolation of *S. halstedii* genomic DNA and construction of its genome library were described previously [17]. Purification of plasmids from *E. coli* was carried out by GFX micro Plasmid Prep Kit (Amersham Biosciences). General DNA manipulations in *E. coli* were performed according to standard protocols [18]. To proceed gene walking, Southern hybridization was carried out using a probe derived from the digested DNA fragment existing most downstream of the isolated cosmid. Labeling of DNA with dioxigenin-dUDP and detection of the hybridized DNA were carried out using DIG DNA Labeling Kit and Detection Kit (Roche Diagnostics).

Sequencing and Sequence Analysis

DNA sequencing was performed on double-stranded DNA templates by dideoxynucleotide chain-termination method with a Model 4000-1 (Li-COR) and ThermoSequencase Cycle Sequencing Kit, and a part of these was ordered to Hitachi High-Technologies. DNA sequences were analyzed with the Align-IR (LI-COR) and the DNASIS program

(Hitachi Software Engineering). Protein sequence homology searches of databases were carried out using the BLAST and the FASTA programs.

Gene Disruption

To disrupt the *hlsH* gene, the plasmid, named pWHM-dH, was prepared as follows. Firstly the *Bam*HI-*Bgl*II fragment (*ca.* 4.5 kb) around the *hlsH* gene was subcloned into the *Bam*HI site of pBluescript II SK(+), and then the apramycin resistant cassette, *aac3*(*VI*) [19], was inserted into the *Aat*II site of the *hlsH* gene of the resulting plasmid. Finally the *Hin*dIII-*Xba*I fragment (*ca.* 6 kb) containing the resulting disrupted *hlsH* gene was subcloned into *Hin*dIII-*Xba*I site of pWHM3. In the case of disruption of the *hlsI* gene, the plasmid, named pWHM-dI, was prepared as follows. The *Sph*I fragment (*ca.* 5 kb) involving the *hlsI* gene was subcloned into *Sph*I site of PUC119, whose *Bam*HI site had been removed, and then the apramycin resistant cassette was inserted into the *Bam*HI site of the *hlsI* gene of the resulting plasmid. Finally the *Sph*I fragment (*ca.* 6 kb) containing the resulting disrupted *hlsI* gene was subcloned into *Sph*I site of pWHM3. The resulting pWHM-dH and -dI were separately introduced into protoplasts of *S. lividans* TK24, and the plasmids were isolated from *S. lividans* TK24. The plasmids were separately purified by a standard CsCl-ethidium bromide gradient centrifugation method [15]. Each pure plasmid was introduced into protoplasts of *S. halstedii*. For homologous recombination, the appropriate transformants were sporulated on R5 agar plates for 20 days. After singlation, the colonies showing the proper phenotype, thiostrepton-sensitive and apramycin-resistant, were selected.

Isolation and Purification of Halstoctacosanolide C

The fermentation broth (1.5 liters) of the *hlsI* mutant was centrifuged to obtain a mycelium cake and supernatant. The mycelium cake was extracted with acetone (500 ml). The extract was filtered through a pad of Celite and evaporated to give an aqueous suspension. The suspension, after being adjusted to pH 10, was extracted three times with ethyl acetate. The combined organic layer was washed with brine, dried over MgSO$_4$, and concentrated *in vacuo*. The supernatant was also extracted three times with ethyl acetate. The combined organic layer was washed with brine, dried over MgSO$_4$, and concentrated *in vacuo*. Then, the crude residues thus obtained were combined and subjected to a flash silica gel column chromatography, which was eluted with 80% ethyl acetate in hexane to afford halstoctacosanolide C (42 mg).

Preparation of Halstoctacosanolide C Methyl Acetal

A solution of halstoctacosanolide C (13.2 mg, 0.016 mmol) in methanol (3.0 ml) was treated with pyridinium *p*-toluenesulfonate (3 mg) for 3 hours at room temperature. The reaction mixture was diluted with Tris-HCl buffer (1 M, pH 8) and the resulting mixture was evaporated to remove methanol, whereupon the rest was extracted three times with ethyl acetate. The combined extract was washed with brine, dried over MgSO$_4$, filtered, and concentrated *in vacuo*. The residue was purified by preparative TLC (5% methanol in CHCl$_3$) to afford methyl acetal derivative of halstoctacosanolide C (11.3 mg, 0.013 mmol, 83%) as white powder. mp 75~77°C; $[\alpha]_D^{28}=-19.7$ (*c* 1.05, CHCl$_3$); IR (neat): 3410 (br), 2960, 2930, 1700 cm^{-1}; HRFAB-MS (*m/z*); Calcd for C$_{49}$H$_{80}$O$_{10}$Na: 851.5649 [M+Na]$^+$, Found: 851.5630.

References

1. Staunton J, Weissman KJ. Polyketide biosynthesis: A millennium review. Nat Prod Rep 18: 380–416 (2001)
2. Ōmura S, Ikeda H, Ishikawa J, Hanamoto A, Takahashi C, Shinose M, Takahashi Y, Horikawa H, Nakazawa H, Osonoe T, Kikuchi H, Shiba T, Sakaki Y, Hattori M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. Proc Natl Acad Sci USA 98: 12215–12220 (2001)
3. Tohyama S, Eguchi T, Dhakal RP, Akashi T, Otsuka M, Kakinuma K. Genome-inspired search for new antibiotics. Isolation and structure determination of new 28-membered polyketide macrolactones, halstoctacosanolides A and B, from *Streptomyces halstedii* HC34. Tetrahedron 60: 3999–4005 (2004)
4. Boakes S, Oliynyk M, Cortes J, Bohm I, Rudd BAM, Revill WP, Staunton J, Leadlay PF, A new modular polyketide synthase in the erythromycin producer *Saccharopolyspora erythraea*. J Mol Microbiol Biotech 8: 73–80 (2004)
5. McAlpine JB, Bachmann BO, Piraee M, Tremblay S, Alarco AM, Zazopoulos E, Farnet CM. Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example. J Nat Prod 68: 493–496 (2005)
6. Bachmann BO. Decoding chemical structures from genomes. Nat Chem Biol 1: 244–245 (2005)
7. August PR, Tang L, Yoon YJ, Ning S, Muller R, Yu TW, Taylor M, Hoffmann D, Kim CG, Zhang X, Hutchinson CR, Floss HG. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. Chem Biol 5: 69–79 (1998)
8. Reid R, Piagentini M, Rodriguez E, Ashley G, Viswanathan N, Carney J, Santi DV, Hutchinson CR, McDaniel R. A

model of structure and catalysis for ketoreductase domains in modular polyketide synthases. Biochemistry 42: 72–79 (2003)

9. Ikeda H, Nonomiya T, Usami M, Ohta T, Ōmura S. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. Proc Natl Acad Sci USA 96: 9509–9514 (1999)

10. Palaniappan N, Kim BS, Sekiyama Y, Osada H, Reynolds KA. Enhancement and selective production of phoslactomycin B, a protein phosphatase IIa inhibitor, through identification and engineering of the corresponding biosynthetic gene cluster. J Biol Chem 278: 35552–35557 (2003)

11. Gregory MA, Petkovic H, Lill RE, Moss SJ, Wilkinson B, Gaisser S, Leadlay PF, Sheridan RM. Mutasynthesis of rapamycin analogues through the manipulation of a gene governing starter unit biosynthesis. Angew Chem Int Ed 44: 4757–4760 (2005)

12. Lamb DC, Ikeda H, Nelson DR, Ishikawa J, Skaug T, Jackson C, Ōmura S, Waterman MR, Kelly SL. Cytochrome P450 complement (CYPome) of the avermectin-producer *Streptomyces avermitilis* and comparison to that of *Streptomyces coelicolor* A3(2). Biochem Biophys Res Commun 307: 610–619 (2003)

13. Parajuli N, Basnet DB, Chan Lee H, Sohng JK, Liou K. Genome analyses of *Streptomyces peucetius* ATCC 27952 for the identification and comparison of cytochrome P450 complement with other *Streptomyces*. Arch Biochem Biophys 425: 233–241 (2004)

14. Ogasawara Y, Kakinuma K, Eguchi T. Involvement of glutamate mutase in the biosynthesis of the unique starter unit of the macrolactam polyketide antibiotic vicenistatin. J Antibiot 58: 468–472 (2005)

15. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. Practical *Streptomyces* genetics. John Innes Foundation, Norwich (2000)

16. Vara J, Lewandowska-Skarbek M, Wang YG, Donadio S, Hutchinson CR. Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in *Saccharopolyspora erythraea* (*Streptomyces erythreus*). J Bacteriol 171: 5872–5881 (1989)

17. Ogasawara Y, Katayama K, Minami A, Otsuka M, Eguchi T, Kakinuma K. Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in *Streptomyces halstedii*. Chem Biol 11: 79–86 (2004)

18. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratry manual, 2nd Edition. Cold Spring Harbor, NY: Cold Spring Harbor laboratory, New York (1989)

19. Blondelet-Rouault MH, Weiser J, Lebrihi A, Branny P, Pernodet JL. Antibiotic resistance gene cassettes derived from the W interposon for use in *E. coli* and *Streptomyces*. Gene 190: 315–317 (1997)