Journal of Industrial Microbiology & Biotechnology (2001) 27, 170–176
© 2001 Nature Publishing Group 1367-5435/01 $17.00
www.nature.com/jim

Organization of biosynthetic gene cluster for avermectin in *Streptomyces avermitilis*: analysis of enzymatic domains in four polyketide synthases

H Ikeda[1,2], T Nonomiya[2] and S Ōmura[1,2]

[1]*School of Pharmaceutical Sciences, Kitasato University, Tokyo 108-8641, Japan;* [2]*Research Center for Biological Function, The Kitasato Institute, Tokyo 108-8642, Japan*

The analysis of the incorporation of ^{13}C-labeled precursors into avermectins indicates that the avermectin aglycons are synthesized by head-to-tail condensation of various acyl groups, which is similar to the biosynthesis of other polyketides. Polyketide synthases (PKS) use the appropriate CoA ester as a primer and add acetate units from malonyl-CoA and propionate units from methylmalonyl-CoA to assemble the polyketides. Avermectin aglycons are formed by addition to the starter unit (2-methylbutyrate or isobutyrate) of 12 acyl condensations in the order P–A–A–A–A–P–P–A–P–A–P–A (P, propionyl; A, acetyl). Within the 90-kb gene cluster for avermectin biosynthesis, the central 65-kb segment was found to be required for aglycon biosynthesis by phenotypic analysis of strains containing deletion or insertion mutations in this region. A complete sequence analysis of the 65-kb segment indicated that this segment encodes avermectin PKS. The avermectin PKS genes are organized into two converging blocks of ORFs. From the results of sequencing analysis, a feature of the two regions, *aveA1/aveA2* and *avea3/aveA4*, is that they encode four kinds of large multifunctional polypeptides containing 55 domains which possess putative fatty acid synthase-like activities. The avermectin PKS (AVES 1–4) appear to contain two, three, or four modules. AVES 1 and 2 contain two and four modules, respectively, whereas AVES 3 and AVES 4 each contains three modules. The 12 modules correspond to the 12 cycles required for synthesis of the avermectin aglycon. *Journal of Industrial Microbiology & Biotechnology* (2001) **27,** 170–176.

Keywords: avermectin; anthelmintic; *Streptomyces avermitilis*; polyketide synthase; biosynthetic gene cluster

Introduction

Avermecins [5,12] and the closely related milbemycins [25] are 16-membered macrocyclic polyketides that are currently attracting great interest for veterinary health because the compounds have potent antiparasitic and broad-spectrum activity against nematode and arthropod parasites. A semisynthetic derivative of avermectins, ivermectin (C22, C23 dihydroavermectin B1), has been used as an antiparasitic agent since 1981, and as an agricultural pesticide and an antiparasitic agent since 1985. Ivermectin has become a major animal health agent and is also being used as a human health agent in the program to eradicate onchocerciasis [3,7].

Avermectins are produced by *Streptomyces avermitilis* which was isolated from a soil sample collected in Shizuoka Prefecture, Japan. They are pentacyclic polyketide-derived compounds linked to a disaccharide of the methylated deoxysugar oleandrose. Eight major compounds result from structural differences at three positions — C5, C22–C23, and C26 (Figure 1) [5]. Among these components, the B1 component has the most effective antiparasitic activity. Avermectin biosynthesis can be classified into three-stage formation of the polyketide-derived initial aglycons (6,8*a*-seco-6,8*a*-deoxy-5-oxoavermectin aglycons), modification of the initial aglycons to generate avermectin aglycons, and glycosylation of avermectin aglycons with a methylated deoxysugar nucleotide, TDP-L-oleandrose, to generate avermectins [8,13,14]. After the presumptive aliphatic polyketide-derived precursors are lactonized to generate initial aglycons, post-polyketide modifications including oxidative cyclization, reduction, and/or methylation occur to form the avermectin aglycons. In the terminal biosynthetic steps, *O*-glycosylation at C13 and C4′ involving dTDP-L-oleandrose yields avermectins [15].

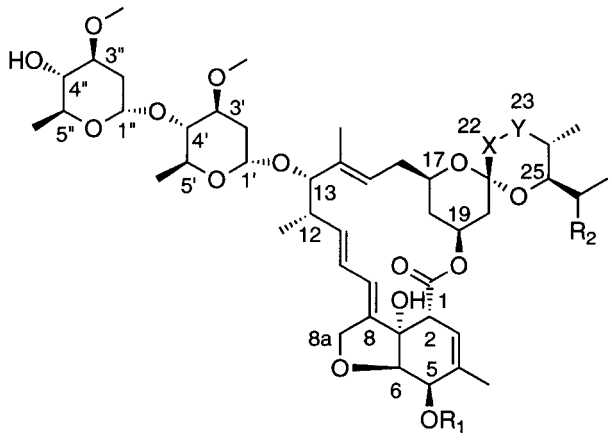
Labeling studies have shown that avermectin aglycons are formed by extension of the starter unit (2-methylbutyrate or isobutyrate) with seven acetate and five propionate units [6]. The biosynthesis of initial aglycons requires a complex polyketide synthase (PKS) that carries out a total of 12 successive cycles of elongation that resemble the steps in long-chain fatty acid biosynthesis. Following each condensation, the β-keto group is reduced to a hydroxyl group by a specific ketoreductase (KR) and the hydroxyl group is dehydrated to generate a double bond by a specific dehydrase (DH). Enoylreductase (ER), which reduces a double bond, is not required during polyketide chain elongation to form the avermectin aglycons because no fully saturated β-carbon chain is found in the avermectin aglycons [20].

We report here a detailed analysis of the modular and domain organization of the avermectin PKS, which differs

Abbreviations: ORF, open reading frame; PKS, polyketide synthase; FAS, fatty acid synthase; AVES, 6,8a-seco-6,8a-deoxy-5-oxoavermectin aglycon synthase; DEBS, 6-deoxyerythronolide B synthase; RAPS, rapamycin polyketide synthase; ACP, acyl carrier protein; AT, acyltransferase; DH, dehydratase; KR, β-ketoacyl-ACP reductase; KS, β-ketoacyl-ACP synthase; TE, thioesterase; kb, kilobase pairs(s); aa, amino acid.

Correspondence: Dr S Ōmura, Research Center for Biological Function, The Kitasato Institute, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8642, Japan

Received 21 September 1999; accepted 14 September 2000



		R_1	R_2	X-Y
Avermectin	A1a	CH_3	C_2H_5	CH=CH
	A1b	CH_3	CH_3	CH=CH
	A2a	CH_3	C_2H_5	CH_2-CH(OH)
	A2b	CH_3	CH_3	CH_2-CH(OH)
	B1a	H	C_2H_5	CH=CH
	B1b	H	CH_3	CH=CH
	B2a	H	C_2H_5	CH_2-CH(OH)
	B2b	H	CH_3	CH_2-CH(OH)
Ivermectin	B1a	H	C_2H_5	CH_2-CH_2
	B1b	H	CH_3	CH_2-CH_2

Figure 1 Structural formulae for avermectins and ivermectins. Both sugars are α-L-oleandrose. Ivermectin is chemically synthesized from avermectin B1 components (B1a and B1b) by selective hydrogenation, and commercial product of ivermectin consists of more than 80% of ivermectin B1a and less than 20% of ivermectin B1b.

significantly from that of other type I PKS of macrolide-producing streptomycetes, both in the constitution of modules and in the mode of polyketide chain initiation, elongation, and termination.

Materials and methods

Bacterial strains and media

S. avermitilis K139 was used as the source of DNA for construction of the genomic libraries. *Escherichia coli* strain NM522 F^- Δ(*srlA–recA*)::Tn*10* was used as a host for cosmid recombinant derivatives and for plasmid subcloning. *E. coli* strain HB101 F^+ and strain JM108 were used as transposon ($\gamma\delta$) donor and recipient, respectively. Luria–Bertani medium was used in *E. coli* propagation and mating.

Vectors, DNA manipulation

Cosmid and plasmid preparations, DNA restriction digestion, size fractionation, DNA fragment isolation, ligation reactions, lambda packaging, and gel electrophoresis were performed by standard procedures. pOCUS-2, a derivative of pUC19, or pUC19 was the routine cloning vector, and pKU402 [18] was the cosmid vector used for genomic DNA library construction. The cosmid library was made using the sonic extract and freeze–thaw lysate of *E. coli* strain SMR10. The primary clones containing *aveD* [17] were extended by genomic walking using the cosmid library to obtain the entire gene cluster for avermectin biosynthesis.

DNA sequencing and analysis

Transposon ($\gamma\delta$)-facilitated DNA sequencing [18] was employed to sequence the avermectin biosynthetic gene cluster. Subclones containing a 1- to 3-kb segment were introduced into *E. coli* HB101 F^+ and their transformants were mated with *E. coli* JM108 to transpose $\gamma\delta$ into target recombinant plasmids. After selection of transconjugants containing the transposition, each transconjugant was screened by polymerase chain reaction using vector and transposon primers to confirm the transposition and to determine the position of the priming site caused by the transposition. DNA segments amplified using vector and transposon primers were used as double-stranded DNA templates. DNA sequencing reactions were performed using a Thermo Sequenase™ kit (Amersham) and a vector or transposon primer labeled with the fluorescence dye IRD40 at the 5′ terminus of each oligonucleotide (Aloka, Japan). The nucleotide sequences were read from a LI-COR Model 4000L sequencer on both DNA strands. DNA and deduced protein sequence homology searches of databases were performed using the BLAST [1], FASTA [22], and FramePlot [19] programs. Multiple alignment and phylogenetic analyses were performed using ClustalW version 1.7 [26].

Results and discussion

The region encoding avermectin PKS in the gene cluster for avermectin biosynthesis

The entire avermectin biosynthetic gene cluster was identified by genomic walking using a cosmid library and a DNA segment containing *aveD* [16,17,21]. A total of 82.0 kb of contiguous DNA was cloned and sequenced, and the deduced gene organization within this region was determined [18]. The nucleotide sequence of the avermectin biosynthetic gene cluster was shown to contain 18 ORFs spanning a distance of 82 kb [18]. The organization of the *ave* modules encoding the multifunctional PKS was complex, with the 12 modules encoding avermectin PKS organized as two groups, *aveA1/aveA2* and *aveA3/aveA4*, six of which are convergently transcribed (Figure 2). The region involving biosynthesis of the initial aglycons is located in the central 65-kb segment of the gene cluster for avermectin biosynthesis. The sequence analysis suggested that *aveC* gene is translationally coupled with *aveA2*. On the other hand, *aveE* is transcribed in the same direction as *aveA3–aveA4*. These two converging transcription units share one transcriptional terminator ($\Delta G = -38.9$ kcal/mol) which is in between *aveC* and *aveE* (Figure 2).

KS domain

The 12 modules of avermectin PKS are organized within four large polypeptides (Figure 2): AVES 1 (414 kDa) contains modules 1–2, AVES 2 (666 kDa) modules 3–6, AVES 3 (575 kDa) modules 7–9, and AVES 4 modules 10–12 [18]. Each module contains one

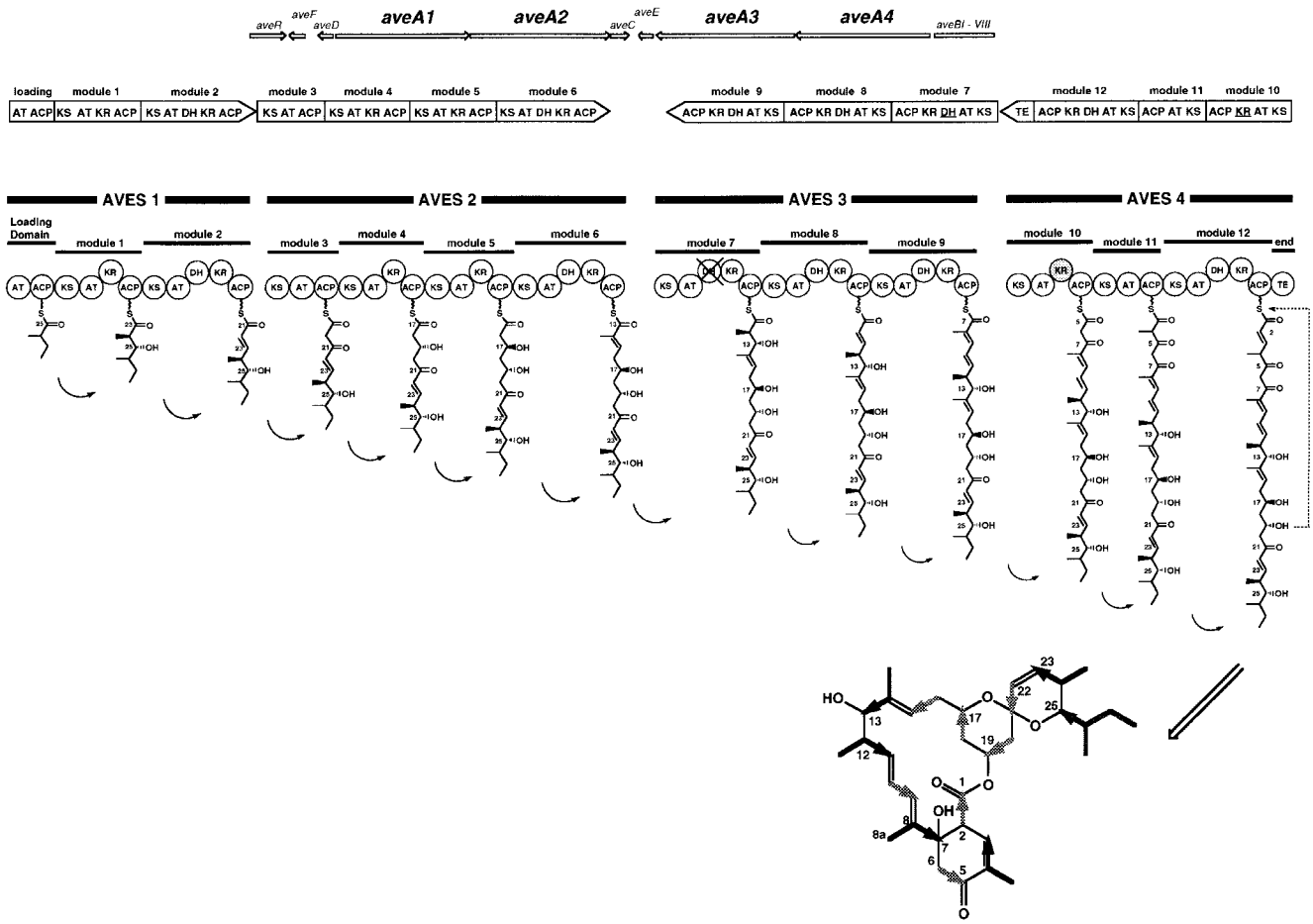


Figure 2 Model for 6,8*a*-seco-6,8*a*-deoxy-5-oxoavermectin aglycon formation and predicted domain structure of the avermectin PKS. The central four genes, *aveA1*, *aveA2*, *aveA3*, and *aveA4*, encode multifunctional PKS, AVES 1, AVES 2, AVES 3, and AVES 4, respectively. Each circle represents an enzymatic domain in the PKS multifunctional polypeptide. Abbreviations: ACP, acyl carrier protein; AT, acyltransferase; DH, dehydratase; KR, β-ketoacyl-ACP reductase; KS, β-ketoacyl-ACP synthase; TE, thioesterase. The reaction order from modules 7 to 9 and 10 to 12 in *aveA3* and *aveA4*, respectively, is drawn in the direction opposite to the gene order on the genome. The crossed-out domain in module 7 is nonfunctional. The shaded domain in module 10 does not function in polyketide chain elongation.

of the following domain organizations: KS, AT, and ACP; KS, AT, KR, and ACP; KS, AT, DH, KR, and ACP. The first module in AVES 1 encodes the initial AT and ACP domains which function to load the starting acyl group onto the AVES 1 and last module in AVES 4 encodes a final domain (TE) which releases the completed aliphatic polyketide chain to form a lactone. In each case, the order of the modules along the polypeptides reflects the order in which each module is used for polyketide chain extension. The order of the domains within each module is KS, AT (DH, KR where applicable), ACP, exactly as for other type I PKS and FAS.

The KS domains of type I PKS and FAS are the most highly conserved (about 30–50% sequence identity over the whole domain) of all the constituent domains in these multifunctional enzymes [4,9,10]. Among the KS domains in avermectin PKS, there is high (71–81%) deduced amino acid sequence similarity over the whole domain and the nucleotide sequence of each domain also has high similarity (~85%). Although the guanosine plus cytosine (GC) contents of the *S. avermitilis* genome are about 73 mol%, which is similar to that of other streptomycetes, the region encoding all KS domains has a relatively high GC content of about 80%. Analysis of multiple sequence alignments of KS domains in all modules revealed that two histidine residues, respectively, 135- and 175-aa C-terminal of the active site cysteine, are invariant in all these domains (Figure 3A). The positions of these two histidines are very similar to those in the other type I PKS modules [2,10]. It has been speculated that one of these histidine residues acts as a general base to increase the nucleophilicity of the active site cysteine [2].

AT domain

All the AT domains in avermectin PKS, as with other type I PKS and FAS, contain the conserved motif sequence GHSXG, in which the serine residue is the active site (Figure 3B). In addition, four other recently proposed invariant residues are conserved in all AT domains [2] except the AT domain (ATload) in the N-terminus of AVES 1. Alanine and histidine residues in the sequence motif yAsHs about 105-aa C-terminal of the active site, a glutamine residue in the N-terminus, and an arginine residue 25-aa C-terminal of the active site serine are invariant in AT1 to 12 domains. Polyketide-derived compounds are composed of a variety of acyl building blocks. The AT domains of avermectin PKS appear to have diverged among themselves more than the other domains. The AT1 shows

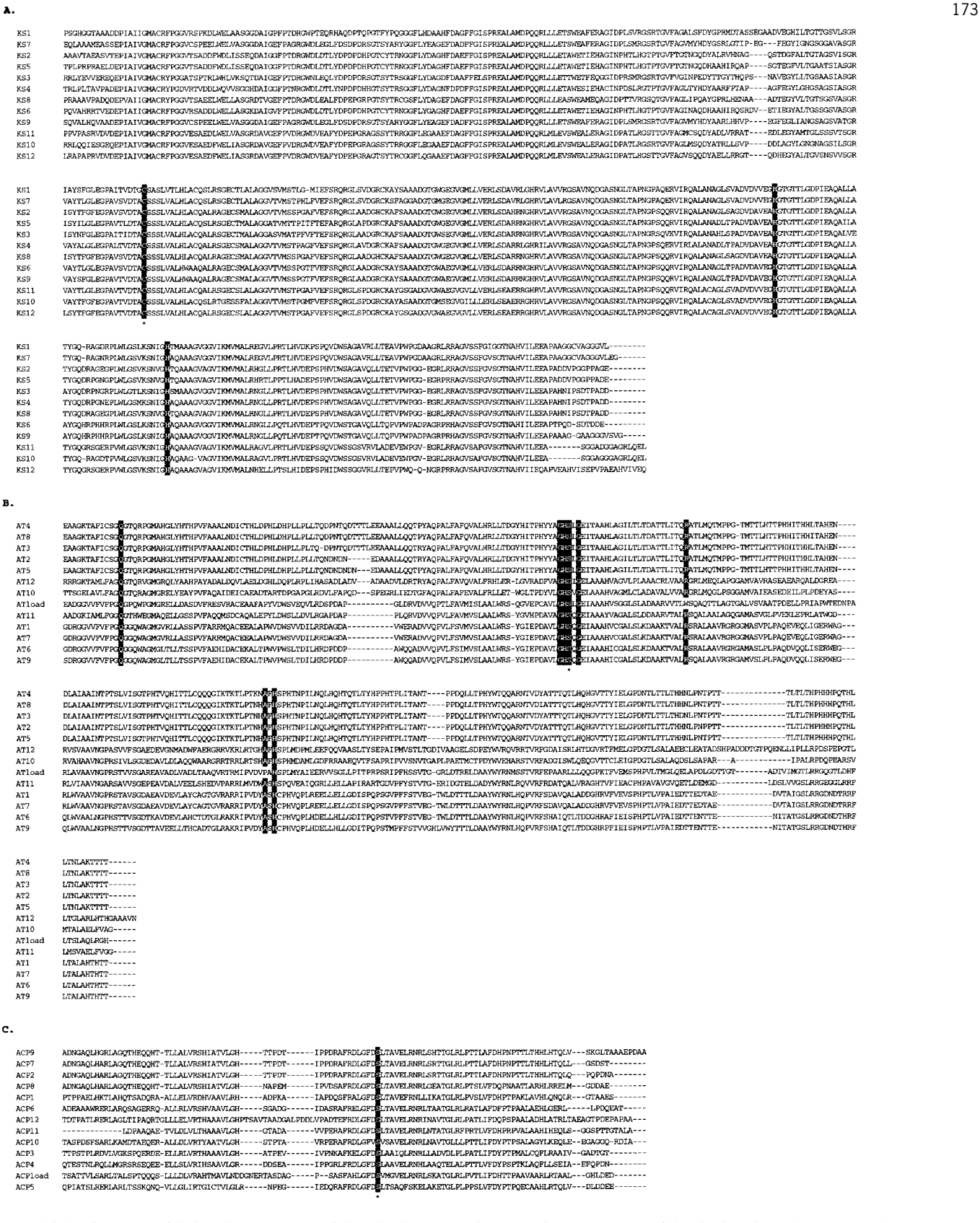


Figure 3 Multiple alignments of deduced aa sequences of domains in avermectin PKS. The aa sequences of the six domains, (A) KS, (B) AT, (C) ACP, (D) DH, (E) KR, and (F) TE were aligned using ClustalW program. The reversed letters indicate conserved invariant residue(s) or motifs for the function of the domains. The active site residues are marked with an asterisk. For comparative purposes, oleandomycin (OLE), erythromycin (ERY), tylosin (TYL), and niddamycin (NID) TE domains have been included in the TE alignment.

Figure 3 (continued).

26.1%, 26.1%, 26.0%, 26.2%, 26.3%, 34.0%, and 33.9% aa identity to AT2, AT3, AT4, AT5, AT8, AT10, and AT12, respectively, while the AT1 shows 82.7%, 100%, 79.1%, 53.9%, and 50.8% aa identity to AT6, AT7, and AT9, and AT11, and ATload, respectively. Sequence alignment generated by a number of programs that perform pairwise comparisons showed that several AT domains of type I PKS are clustered into at least three groups [18,23]. The first group contains only the proposed malonate loading functions, the second group contains methylmalonyl loading functions, and the third group contains an acyl residue derived from the monocarboxylic acid loading function. The sequence alignment of AT domains in avermectin PKS coincides in this manner (AT1, 6, 7, 9, 11 and AT2, 3, 4, 5, 8, 10, 12 possess methylmalonyl and malonyl loading functions, respectively). The ATload is somewhat different from the other 12 AT domains. Two of four invariant residues in regular AT domains, which are 25- and 105-aa C-terminal of the active site serine, respectively, were not conserved in the ATload. Interestingly, ATload clustering the third group is found in the N-terminal of the first module of DEBS and AVES only [11,18]. The actual assignment of the substrate specificity for each domain will have to await identification of the polyketide encoded by this pathway.

ACP domain

There is an ACP domain associated with each of the 12 modules in the avermectin PKS, and an additional ACP (ACPload) is located immediately upstream of the first KS domain (Figure 2). Similarities between all ACP domains range from 41.2% to 59.2% aa identity. Although the sequence homology among the all ACP domains is not high compared to that of KS domains, the sequence around the phosphopantheteine binding site, the serine residue in the LGFDS motif, is relatively conserved (Figure 3C). Remarkably, ACP10 has the active site sequence LGFVS where all previously identified ACPs in type I PKS possess aspartate adjacent to the serine. A similar observation was also found in ACP11 of RAPS, in which the active site sequence was LGINS [2].

DH domains

In avermectin PKS, all modules encoded putative KS, AT, and ACP domains responsible for all of the acyl condensation processes, but this was not the case for functions involved in the processing of β-carbons. Modules 1, 4, 5, and 10 encoded a KR domain only. Modules 3 and 11 lakced a KR domain. Modules 2, 6, 7, 8, 9, and 12 carried DH and KR domains (Figure 2). In modules, 2, 6, 8, 9, and 12, where DH activity is predicted to be required for initial

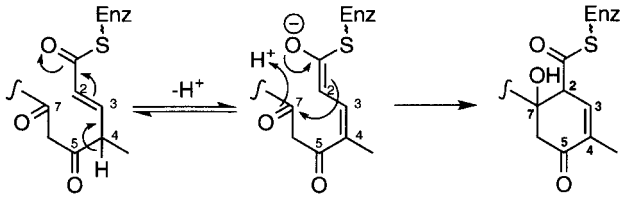


Figure 4 The formation of the cyclohexene ring at C2–C7. Ring formation is performed by aldol condensation between the C2 enoyl and C7 carbonyl.

aglycon biosynthesis, there is good alignment with the DH domains from DEBS and RAPS, and the active site motif HXXXGXXXXP [4,11] is present (Figure 3D). The dysfunctional DH domain in module 7 is consistent with the retention of a hydroxyl group at C13. The DH domain of module 7 contains a partially conserved dehydratase consensus sequence with two mismatched amino acids in the dehydratase motif. The DH domain in module 7 contains YXXXGXXXXS in the corresponding region, which would readily account for its inactivity. Interestingly, the first amino acid replacement of histidine (H) to tyrosine (Y) in the corresponding dehydratase consensus sequence is caused by a one-base change in the corresponding nucleotide sequence, in which CAC (H) has been converted to TAC (Y). The DH2 contains a partially conserved dehydratase consensus sequence with one mismatched amino acid in the active motif HXXXGXXXXS. The DH, which corresponds to C22–C23 dehydration, seems to have partial dehydratase activity because two intermediates containing β-hydroxyl or enoyl carbons at C22–C23 were processed in subsequent acyl condensation reactions. Although the ordinary active site motif sequence, HXXXGXXXXP, was introduced into the corresponding region of the DH domain in module 2 on the chromosome by gene replacement, the resulting recombinant strains still produced components 1 and 2 containing β-hydroxyl or enoyl carbons at C22–C23, respectively [18]. The active site sequence, HXXXGXXXXS, in the module 2 does not interfere with the dehydratase activity and the partial activity may be due to another reason.

KR domain

Similarities between all KR domains range from 39.7% to 89.2% aa identity. Although the conserved motif, GXGXXGXXXA, proposed to be the NADP(H) binding site [24] in KR domains of other type I PKS, was found in modules 2, 6, 7, 8, 9, 10, and 12, KR domains of modules 1, 4, and 5 contain GXGXXAXXXT in the corresponding region (Figure 3E), where KR activity is predicted to be required for initial aglycon biosynthesis (Figure 2). Module 10 encodes KS, AT, KR, and ACP domains, but the KR domain would be nonfunctional because C7 must be a carbonyl residue in order to form the cyclohexene ring at C2–C7 by aldol condensation between the C2 enoyl and the C7 carbonyl (Figure 4). Thus, the structure of the initial aglycon would not require the activity of KR10, but there is no convincing evidence from the sequence alignments that the KR of either module is inactive.

TE domain

A TE domain was identified at the C-terminus of AVES 4, adjacent to the ACP domain of module 12 (Figure 2). The avermectin PKS TE shows 37.7%, 39.2%, 34.8%, and 35.3% aa identity to the erythromycin, niddamycin, oleandomycin, and tylosin PKS and TEs, respectively. Analysis of alignment of these sequences revealed the invariant GXHXG and GXH motifs, which are considered to be essential to be essential for TE activity (Figure 3F) [10].

Conclusions

(1) A type I PKS cluster for avermectin aglycon biosynthesis was isolated and sequenced from an avermectin-producing strain of *S. avermitilis*. It comprises four ORFs, AVES 1, AVES 2, AVES 3, and AVES 4 which encode two, four, three, and three modules, respectively. Analysis of deduced aa sequences of these four AVES containing 55 domains indicates that almost all domains in the 12 modules contain the invariant active sites conserved among type I PKS. Five AT domains, AT1, AT6, AT7, AT9, and AT11, are homologous to ATs specific for malonyl CoA, while the other seven, AT2, AT3, AT4, AT5, AT8, AT10, and AT12, are homologous to those ATs specific for methylmalonyl CoA.

(2) The NADP(H) binding motif of three KR domains, KR1, 4, and 5, are different from the motif conserved in type I PKS. These KR domains contain the motif sequence, GXGXXAXXXT, in the NADP(H) binding site. The dehydratase active site motif of DH2, containing the sequence of HXXXGXXXXS, is also different from that of the motif conserved in type I PKS.

(3) In module 10, there is a predicted active site for reduction (KR) which is not reflected in the ultimate structure of avermectin and which may be nonfunctional. A nonfunctional DH domain is found in module 7, and contains YXXXGXXXXS in the corresponding dehydratase-conserved motif. This replacement of histidine to tyrosine in the motif sequence is caused by a one-base change in the corresponding nucleotide sequence, in which CAC (H) is replaced by TAC (Y).

Acknowledgements

This work was supported by Grants-in Aid of the Ministry of Education, Science and Culture Japan no. 10680573 (to H.I.) and "Research for the Future" program of the Japan Society for the Promotion of Science, JSPS-RFTF96I00304 (to S.O. and H.I.).

References

1 Altschul SF, W Gish, W Miller, EW Myers and DJ Lipman. 1990. Basic local alignment search tool. *J Mol Biol* 215: 403–410.
2 Aparicio JF, I Molnar, T Schwecke, A Konig, SF Haydock, LE Khaw, J Staunton and PF Leadlay. 1996. Organization of biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. *Gene* 169: 9–16.
3 Aziz MA, S Diallo, IM Diop, M Lariviere and M Parta. 1982. Efficacy and tolerance of ivermectin in human onchocerciasis. *Lancet* 2: 171–173.
4 Bevitt DJ, J Corts, SF Haydock and PF Leadlay. 1992. 6-Deoxyerythronolide B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur J Biochem* 204: 39–49.
5 Burg RW, BM Miller, EE Baker, J Birnbaum, SA Currie, R Hartman, Y-L Kong, RL Monaghan, G Olson, I Putter, JB Tunac, H Wallick, EO Stapley, R Oiwa and S Ōmura. 1979. Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. *Antimicrob Agents Chemother* 15: 361–367.

6 Cane DE, T-C Liang, LK Kaplan, MK Nallin, MD Schulman, OD Hensens, AW Douglas and G Albers-Schönberg. 1983. Biosynthetic origin of the carbon skeleton and oxygen atoms of the avermectins. *J Am Chem Soc* 105: 4110–4112.
7 Chabala JC, H Mrozik, RL Tolman, P Eskola, LH Peterson, MF Woods, MH Fisher and WC Campbell. 1980. Ivermectin, a new broad-spectrum antiparasitic agent. *J Med Chem* 23: 1134–1136.
8 Chen TS, OD Hensens and MD Schulman. In: Campbell WC (Ed.), Ivermectin and Avermectin. Springer-Verlag, New York, 1989, p 55.
9 Cortés J, SF Haydock, GA Roberts, DJ Bevitt and PF Leadlay. 1990. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* 348: 176–178.
10 Donadio S and L Katz. 1992. Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*. *Gene* 111: 51–60.
11 Donadio S, MJ Staver, JB McAlpine, SJ Swanson and L Katz. 1991. Modular organization of the genes required for complex polyketide biosynthesis. *Science* 252: 675–679.
12 Egerton JR, DA Ostling, LS Blair, CH Eary, D Suhayda, S Cifelli, RF Riek and WC Campbell. 1979. Avermectins, new family of potent anthelmintic agents: efficacy of the B1a component. *Antimicrob Agents Chemother* 15: 372–378.
13 Ikeda H and S Ōmura. 1993. Genetic aspects of the selective production of useful components in the avermectin producer *Streptomyces avermitilis*. *Actinomycetology* 7: 133–144.
14 Ikeda H and S Ōmura. 1995. Control of avermectin biosynthesis in *Streptomyces avermitilis* for the selective production of a useful component. *J Antibiot* 48: 669–682.
15 Ikeda H and S Ōmura. 1997. Avermectin biosynthesis. *Chem Rev* 97: 2591–2609.
16 Ikeda H, Y Takada, C-H Pang, H Tanaka and S Ōmura. 1993. Transposon mutagenesis by Tn*4560* and its applications in avermectin-producing *Streptomyces avermitilis*. *J Bacteriol* 175: 2077–2082.
17 Ikeda H, R-L Wang, T Ohta and S Ōmura. 1998. Cloning of the gene encoding avermectin B 5-*O*-methyltransferase in avermectin-producing *Streptomyces avermitilis*. *Gene* 206: 175–180.
18 Ikeda H, T Nonomiya, M Usami, T Ohta and S Ōmura. 1999. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc Natl Acad Sci USA* 96: 9509–9514.
19 Ishikawa J and Hotta K. 1999. FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. *FEMS Microbiol Lett* 174: 251–253.
20 MacNeil DJ, JL Occi, KM Gewain, T MacNeil, PH Gibbons, CL Ruby and SJ Danis. 1992. Complex organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase. *Gene* 115: 119–125.
21 MacNeil T, KM Gewain and DJ MacNeil. 1993. Deletion analysis of the avermectin biosynthesis genes of *Streptomyces avermitilis* by gene cluster displacement. *J Bacteriol* 175: 2552–2563.
22 Pearson WR. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* 183: 63–98.
23 Ruan X, D Stassi, SA Lax and L Katz. 1997. A second type-I PKS gene cluster isolated from *Streptomyces hygroscopicus* ATCC29253, a rapamycin-producing strain. *Gene* 203: 1–9.
24 Scrutton NS, A Berry and RN Perham. 1990. Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* 343: 38–43.
25 Takiguchi Y, H Mishima, M Okuda and M Terao. 1980. Milbemycins, a new family of macrolide antibiotics. Fermentation, isolation and physico-chemical properties. *J Antibiot* 33: 1120–1127.
26 Thompson JD, Higgins DG, Gibson TJ. 1994. Improved sensitivity of profile searches through the use of sequence weights and gap excision. *Comput Appl Biosci* 10: 19–29.