Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 9509–9514, August 1999
Biochemistry

Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*

HARUO IKEDA*, TOMOKO NONOMIYA†, MASAYO USAMI* TOSHIO OHTA†, AND SATOSHI ŌMURA*†‡

*School of Pharmaceutical Sciences, Kitasato University, Tokyo 108-8641, Japan; and †Research Center for Biological Function, The Kitasato Institute, Tokyo 108-8642, Japan

Contributed by Satoshi Ōmura, June 23, 1999

ABSTRACT **Analysis of the gene cluster from *Streptomyces avermitilis* that governs the biosynthesis of the polyketide anthelmintic avermectin revealed that it contains four large ORFs encoding giant multifunctional polypeptides of the avermectin polyketide synthase (AVES 1, AVES 2, AVES 3, and AVES 4). These clustered polyketide synthase genes responsible for avermectin biosynthesis together encode 12 homologous sets of enzyme activities (modules), each catalyzing a specific round of polyketide chain elongation. The clustered genes encoding polyketide synthase are organized as two sets of six modular repeats, *aveA1-aveA2* and *aveA3-aveA4*, which are convergently transcribed. The total of 55 constituent active sites makes this the most complex multifunctional enzyme system identified to date. The sequenced DNA region contains 14 additional ORFs, some of which encode polypeptides governing other key steps in avermectin biosynthesis. Between the two sets of polyketide synthase genes lie two genes involved in postpolyketide modification, one of which encodes cynthochrome P450 hydroxylase that probably catalyzes furan ring formation at C6 to C8a. Immediately right of the large polyketide synthase genes is a set of genes involved in oleandrose biosynthesis and its transglycosylation to polyketide-derived aglycons. This cluster includes nine genes, but one is not functional in the biosynthesis of avermectin. On the left side of polyketide synthase genes, two ORFs encoding methyltransferase and nonpolyketide synthase ketoreductase involved in postpolyketide modification are located to the left of the polyketide synthase genes, and an adjacent gene encodes a regulatory function that may be involved in activation of the transcription of avermectin biosynthetic genes.**

Streptomyces avermitilis produces avermectin, a series of eight related pentacyclic lactones that contain a disaccharide of the methylated deoxysugar oleandrose (1). Avermectin and the related compounds milbemycin and nemadectin are potent anthelmintic compounds, the former two of which are used commercially in animal healthcare and agriculture. The semisynthetic derivatives of avermectin C22, C23 dihydroavermectin B1, ivermectin, are widely used for the treatment of diseases caused by nematodes and arthropods in veterinary and agricultural fields, respectively. Ivermectin has been used for livestock farming and health care of companion animals. The efficacy of ivermectin in human onchocerciasis has made it a promising candidate for the control of this insidious and intractable tropical disease (2). Ivermectin also has been found to be effective against human strongyloidiasis in Okinawa, Japan (3) as well as Asian countries.

The genes encoding functions for the biosynthesis of many antibiotics are clustered. A natural fertility system has been used to examine linkage between mutations in several strains defective in avermectin production (4), and the avermectin biosynthetic gene cluster has been cloned and characterized (5, 6). Avermectin biosynthesis can be classified into three stages as follows: formation of the polyketide-derived initial aglycon (6, 8a-seco-6, 8a-deoxy-5-oxoavermectin aglycons), modification of the initial aglycon to generate avermectin aglycons, and glycosylation of avermectin aglycons to generate avermectins. A region of 90 kilobases (kb) is required for the biosynthesis of avermectin, and the central 65-kb segment was found to be required for the formation of the aglycon by phenotypic analysis of strains containing deletion or insertion mutations in this region. The avermectin polyketide synthase (PKS) uses a variety of acyl units. Labeling studies have shown that the starting acyl group for avermectin is derived from valine (isobutyryl-CoA) or isoleucine (2-methylbutyryl-CoA) to yield the a and b components, respectively (7). The labeling studies also have revealed that avermectin aglycon is formed by extension of the starter unit with seven acetate and five propionate units (7). After the presumptive aliphatic polyketide-derived precursor is lactonized to generate an initial aglycon, a series of modifications, including furan ring closure, ketoreduction, and/or methylation occur to form avermectin aglycons. In the terminal biosynthetic steps, *O*-glycosylation at C13 and C4′ is performed by using deoxythymidine diphosphate (dTDP)–oleandrose to form avermectins.

The engineered biosynthesis of altered avermectin derivatives is of great interest for enhancement of the potency and spectrum of activity of this drug and reduction of its toxicity (2, 8, 9). Therefore, we performed a detailed study of the organization of the avermectin biosynthetic genes in *S. avermitilis*.

MATERIALS AND METHODS

Bacterial Strains and Media. *S. avermitilis* K139 (4) was used as the source of DNA for the construction of the genomic libraries. Media for vegetative growth, sporulation and avermectin production were used as described (4). *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 (γδ) donor and recipient, respectively. *E. coli* strain SM10 was used as a source of λ packaging extract. Luria–Bertani medium was used in *E. coli* propagation and mating.

Vectors and DNA Manipulation. Cosmid and plasmid preparations, DNA restriction digestion, size fractionation, DNA fragment isolation, ligation reactions, λ packaging, and gel electrophoresis were performed by standard procedures (10). *E. coli* transformation was performed according to the procedure of Hanahan (11). POCUS, a derivative of pUC19, or pUC19 was the routine cloning vector, and pKU402 (12) was the cosmid vector used for genomic DNA library construction. The cosmid library was made by using the sonic extract and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: PKS, polyketide synthase; kb, kilobase; ACP, acyl carrier protein; dTDP, deoxythymidine diphosphate.

‡To whom reprint requests should be addressed. E-mail: omura-s@kitasato.or.jp.

freeze-thaw lysate of *E. coli* strain SM10. The primary clones containing *aveD* (13) were extended by genomic walking by using the cosmid library to obtain the entire gene cluster for avermectin biosynthesis.

DNA Sequencing and Analysis. Transposon (γδ)-facilitated DNA sequencing (14) was used to sequence the avermectin biosynthetic gene cluster. About 1- to 3-kb restriction fragments that were isolated by *Bam*HI, *Pst*I, *Sst*I, or *Kpn*I digestion were subcloned into plasmid vectors. These subclones were introduced into *E. coli* HB101 F^+, and their transformants were mated with *E. coli* JM108 to transpose γδ into target recombinant plasmid. After selection of transconjugants containing the transposition, each transconjugant was screened by PCR by 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 by using vector and transposon primers were used as double-stranded DNA templates. DNA sequencing reactions were performed by using a Thermo Sequenase kit (Amersham Pharmacia) and a vector or transposon primer labeled with the fluorescence dye IRD40 at the 5′ terminus of each oligonucleotide (Aloka, Tokyo). The nucleotide sequences were read from a Li-Cor (Lincoln, NE) Model 4000L sequencer on both DNA strands. DNA and deduced protein sequence homology searches of databases were performed by using the BLAST (15) and FASTA (16) programs. Sequences were analyzed using the FramePlot (17) developed by J. Ishikawa of The National Institute of Infectious Diseases (Tokyo). Multiple alignment and phylogenetic analyses were performed by using CLUSTALW 1.7 (18).

Avermectin Extraction and Complementation Analysis. Avermectins and related compounds were extracted according to the published procedures (4). Silica gel thin layer chromatography was routinely performed to detect avermectins and related compounds. Complementation analysis was performed as described (13).

RESULTS AND DISCUSSION

The entire avermectin biosynthetic gene cluster was identified by genomic walking using a cosmid library and a DNA segments containing *aveD* (13). A total of 82.0 kb of contiguous DNA was cloned and sequenced, and the deduced gene organization within this region is shown in Fig. 1. The nucleotide sequence of the avermectin biosynthetic gene cluster was completely determined and was shown to contain 18 ORFs spanning a distance of 82 kb (Table 1). The organization of the *ave* modules encoding the multifunctional PKS was more complex than the organization of the modules encoding erythromycin (19, 20), pikromycin (21), or rifamycin PKS (22). The 6 modules encoding erythromycin or pikromycin PKS, or 10 modules encoding rifamycin PKS, are transcribed in one direction whereas the 12 modules encoding avermectin PKS are organized as two groups, *aveA1-aveA2* and *aveA3-aveA4*, six of which are convergently transcribed (Fig. 1). The organization of the modules encoding rapamycin PKS was also more complex, but these modules are organized as two groups (10 and 4 modules, respectively) that are transcribed in two directions (23). The most striking feature was the presence of four extraordinarily large ORFs, *aveA1*, *aveA2*, *aveA3*, and *aveA4* (Fig. 1). The nomenclature for the genes, proteins, and modules in avermectin PKS is very confused. To prevent confusing, we have changed to rationalize the nomenclature in consideration of polyketide-elongation reaction order. These ORFs encoded the multifunctional PKSs AVES 1 (6, 8a-6, 8a-deoxy-5-oxoavermectin aglycon synthase; 414 kDa) containing the first two modules required for polyketide chain extension, AVES 2 (666 kDa) containing the next four modules with the activities required to continue chain elongation up to C13 (Fig. 2), AVES 3 (575 kDa) containing the next three modules with activities required to continue chain elongation up to C7 (Fig. 2), and AVES 4 (510 kDa) containing three modules required to complete the polyketide-derived initial aglycon, 6,8a-seco-6,8a-deoxy-5-oxoavermectin aglycon, respectively (Fig. 2). In total, the four large proteins contribute 55 catalytic functions. The first and last modules encode unique domains similar to the erythromycin PKS modules. Chain initiation on the avermectin PKS may require initial acyltransferase and acyl carrier protein domains to load the starting acyl group onto the PKS. Such domains were identified at the N terminus of AVES 1, adjacent to a β-ketoacyl–acyl carrier protein (ACP) synthase domain of module 1 in Fig. 2. On the other hand, chain termination on the avermectin PKS may require a thioesterase domain that releases the completed polyketide from PKS to form a lactone. Such a thioesterase domain was identified at the C terminus of AVES 4, adjacent to the acyl carrier protein domain of module 12 (Fig. 2). Each module showed one of the following domain organizations: β-ketoacyl-ACP synthase, acyltransferase, and acyl carrier protein; β-ketoacyl-ACP synthase, acyltransferase, β-ketoacyl-ACP reductase, and acyl carrier protein; or β-ketoacyl-ACP synthase, acyltransferase, dehydratase, β-ketoacyl-ACP reductase, and acyl carrier protein. In avermectin PKS, all modules encoded putative β-ketoacyl-ACP synthase, acyltransferase, and acyl carrier protein domains responsible for all of the acyl condensation processes but for a variable number of functions involved in the processing of β-carbons. As shown in Fig. 2, modules 1, 4, 5, and 10 encoded a

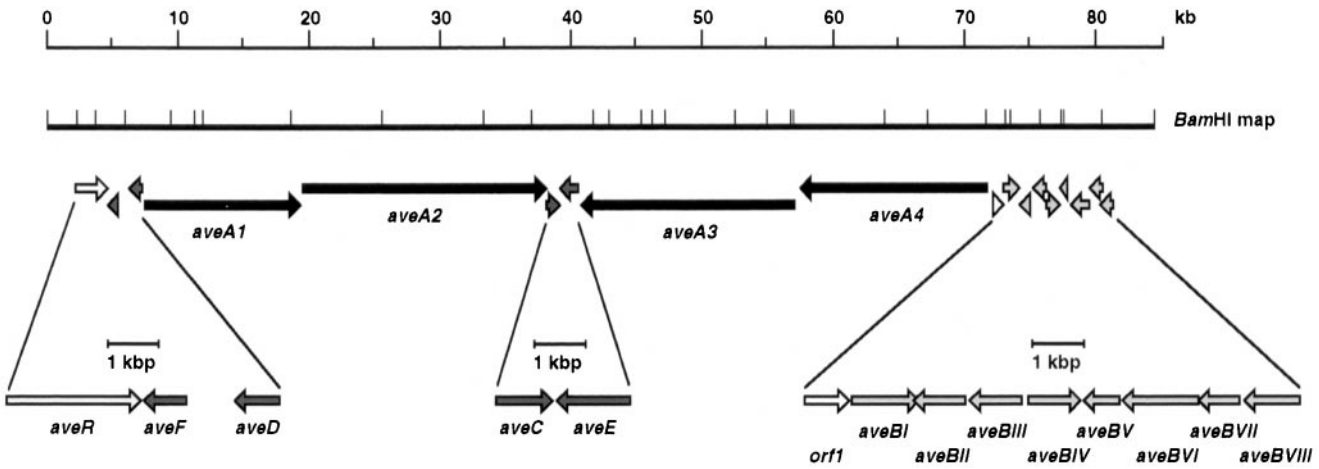

FIG. 1. Organization of the gene cluster for avermectin biosynthesis. The direction of transcription and relative sizes of the ORFs deduced from analysis of the nucleotide sequence are indicated.

Table 1. Deduced functions of ORFs in the avermection biosynthetic gene cluster

Polypeptide	Amino acids, no.	Proposed function or sequence similarities detected					
AVES 1	3,953	PKS					
Loading domain			AT(MB)			ACP	
Module 1		KS	AT(P)		KR	ACP	
Module 2		KS	AT(A)	DH	KR	ACP	
AVES 2	6,239	PKS					
Module 3		KS	AT(A)			ACP	
Module 4		KS	AT(A)		KR	ACP	
Module 5		KS	AT(A)		KR	ACP	
Module 6		KS	AT(P)	DH	KR	ACP	
AVES 3	5,532	PKS					
Module 7		KS	AT(P)	DH*	KR	ACP	
Module 8		KS	AT(A)	DH	KR	ACP	
Module 9		KS	AT(P)	DH	KR	ACP	
AVES 4	4,881	PKS					
Module 10		KS	AT(A)		KR*	ACP	
Module 11		KS	AT(P)			ACP	
Module 12		KS	AT(A)	DH	KR	ACP	TE
ORF 1	238	Reductase ?†					
AveBI	412	Glycosyl transferase					
AveBII	355	dTDP-glucose 4,6-dehydratase					
AveBIII	299	α-D-Glucose-1-phosphate thymidylyltransferase					
AveBIV	343	dTDP-4-keto-6-deoxy-L-hexose 4-reductase					
AveBV	225	dTDP-4-keto-6-deoxyhexose 3,5-epimerase					
AveBVI	469	dTDP-4-keto-6-deoxy-L-hexose 2,3-dehydratase					
AveBVII	257	dTDP-6-deoxy-L-hexose 3-*O*-methyltransferase					
AveBVIII	347	dTDP-4-keto-6-deoxy-L-hexose 2,3-reductase					
AveC	347	?					
AveD	283	C5-*O*-methyltransferase					
AveE	456	Cytochrome P450 hydroxylase					
AveF	302	C5-ketoreductase					
AveR	949	Positive regulator					

Abbreviations are the same as in Fig. 2. AT(MB), acyltransferase incorporating a 2-methylbutyrate or isobutyrate extender unit; AT(A), acyltransferase incorporating an acetate extender unit; AT(P), acyltransferase incorporating a propionate extender unit. Partial or low-level sequence similarities are indicated by question marks.
*Enzyme activity is possible nonfunctional (see text).
†The gene product is not involved in avermectin biosynthesis.

β-ketoacyl-ACP reductase domain only. Modules 3 and 11 lacked a β-ketoacyl-ACP reductase domain. Modules 2, 6, 7, 8, 9, and 12 carried β-ketoacyl-ACP reductase and dehydratase domains. C21 is likely to be a carbonyl group because the spiroketal is probably formed by ketalization of the carbonyl group at C21 and two hydroxyl groups at C17 and C25. The β-hydroxy carbons at C17 and C25 are generated by β-ketoacyl-ACP reductase domains in module 1 or 5, respectively, during chain elongation by acyl condensation. The PKS functions predicted to be required to synthesize the initial aglycon were compared with the domains identified in the avermectin PKS, and one β-ketoacyl-ACP reductase domain in module 10 and one dehydratase domain in module 7 appeared to be nonfunctional. Module 10 encodes β-ketoacyl-ACP synthase, acyltransferase, β-ketoacyl-ACP reductase, and acyl carrier protein domains, but the β-ketoacyl-ACP reductase domain is nonfunctional because C7 must be a carbonyl residue to form the cyclohexene ring at C2 to C7 by aldol condensation between the C2 enoyl and the C7 carbonyl. Thus, the structure of the initial aglycon would not require the activity of a β-ketoacyl-ACP reductase in module 10, but there is no convincing evidence from the sequence alignments that the β-ketoacyl-ACP reductase of either module is inactive. The dysfunctional dehydratase domain in module 7 is consistent with the retention of a hydroxyl group at C13. The dehydratase domain of module 7 contains a partially conserved dehydratase consensus sequence with two mismatched amino acids in the dehydratase motif. The active site motif HXXXGXXXXP (24) is present in functional dehydratase domains in modules 6, 8, 9, and 12. In contrast, the dehydratase 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) is replaced to TAC (Y). Dehydratase domain 2 contains a partially conserved dehydratase consensus sequence with one mismatched amino acid in the active motif HXXXGXXXXS. The dehydratase that corresponds to C22–23 dehydration seemed to have partial dehydratase activity because two intermediates containing β-hydroxyl or enoyl carbons at C22–23 were processed in subsequent acyl condensation reactions. Although the ordinary active site motif sequence, HXXXGXXXXP, was introduced into the corresponding region of the dehydratase domain in module 2 on the chromosome by gene replacement, the resultant recombinant strains produced both components 1 and 2 containing β-hydroxyl or enoyl carbons at C22–23, respectively (data not shown). The active site sequence HXXXGXXXXS in the module 2 would not interfere with the dehydratase activity, and the partial activity may be attributable to other regions of the domain or "downstream" domain acting prematurely on the substrate.

Polyketide-derived compounds are composed of a variety of acyl building blocks. Sequence alignment generated by a number of programs that perform pairwise comparisons showed that 57 acyltransferase domains of erythromycin, pikromycin, tylosin, niddamycin, rapamycin, and avermectin PKS are clustered into at least three groups. The first group contains only the proposed malonate loading functions, the

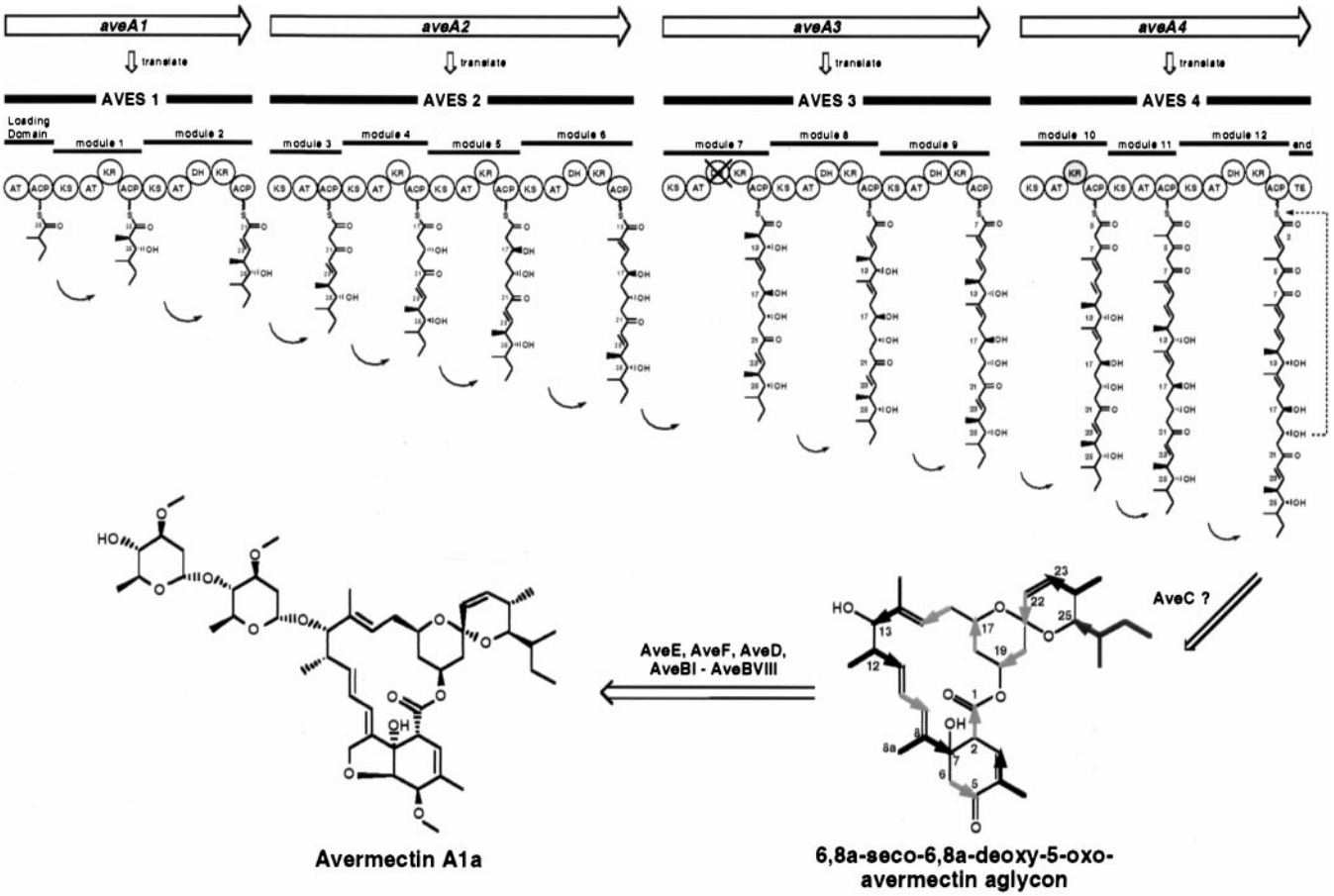

FIG. 2. Model for 6,8a-seco-6,8a-deoxy-5-oxoavermectin aglycon formation and predicted domain structure of the avermectin PKS. Each circle represents an enzymatic domain in the PKS multifunctional polypeptide. AT, acyltransferase; DH, dehydratase; KR, β-ketoacyl-ACP reductase; KS, β-ketoacyl-ACP synthase; TE, thioesterase. The reaction order from module 7 to 9 and from 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.

second group contains methylmalonyl loading functions, and the third group contains an acyl residues derived from monocarboxylic acid loading functions (Fig. 3*A*). As shown in Fig. 3*A*, the second group (methylmalonyl loading functions) contains ethylmalonyl (extending butyrate) and hydroxymalonyl (extending glycolate) loading functions in tylosin and niddamycin PKS. Furthermore, the signature sequences for malonyl and methylmalonyl (without ethymalony and hydroxymalonyl) loading domains were identified and are shown in Fig. 3*B*. The alignment of the domain sequence of the third group is similar to that of methylmalonyl loading domains, but the consensus sequence was not found around the active serine residue. The actual assignment of the substrate specificity for each domain will have to await identification of the polyketide encoded by this pathway.

Two ORFs, *aveC* and *aveE*, lie between the two sets of PKS genes shown in Fig. 1. These ORFs are adjacent to *aveA2* and *aveA3*, respectively. Nucleotide sequence analysis of this region suggested that *aveC* is translationally coupled to *aveA2*. Although mutation of *aveC* led to the production of component 2 containing a β-hydroxy carbon at C22–23 (25), there was no homology between the deduced amino acid sequence of the gene product of *aveC* and the active site motif of a putative dehydratase. Therefore, the function of AveC is not obvious. It may be involved in postpolyketide modification. The sequencing results implied that *aveE* may be on a single transcription unit with *aveA4* and *aveA3*. The deduced amino acid sequence in the C-terminal region of AveE resembles the consensus sequence of cytochrome P450 hydroxylase. This result and the previous complementation profiles (26) suggested that furan ring closure is initiated by introduction of an oxygen atom into the allylic methyl residue at C8a, which would be catalyzed by cytochrome P450 hydroxylase.

Nine other ORFs were detected to the right of the PKS genes, some of which showed significant sequence similarity to genes involved in deoxysugar biosynthesis and transglycosylation of dTDP-sugar in the gene cluster for macrolide antibiotic biosynthesis (Fig. 1). The gene product of the first ORF, orf1, which is adjacent to *aveA4*, resembles reductases, but the gene product of this ORF may not function in avermectin biosynthesis. The genetic loci for oleandrose biosynthesis and transglycosylation are immediately downstream of orf1. Seven genes, *aveBII*, *aveBIII*, *aveBIV*, *aveBV*, *aveBVI*, *aveBVII*, and *aveBVIII*, are responsible for deoxysugar biosynthesis, and the eighth gene, *aveBI*, encodes a glycosyltransferase that apparently catalyzes transfer of oleandrose onto polyketide-derived aglycons (Table 1). Beyond *aveBVIII* lie genes that resemble a homolog of thioesterase and a large subunit of glutamate synthase, and the mutants deleted in these regions by transreplacement with recombinant plasmid still produced avermectins, indicating the limit of the avermectin biosynthetic cluster (data not shown).

Three ORFs were detected to the left of the PKS genes (Fig. 1). Two these ORFs were found upstream of *aveA1* and were shown to be involved in the modification of the initial aglycons. *aveD*, immediately upstream of *aveA1*, encodes a C5 *O*-methyltransferase requiring *S*-adenosyl-L-methionine (13). Just downstream of the *aveD* locus is a gene (*aveF*) encoding

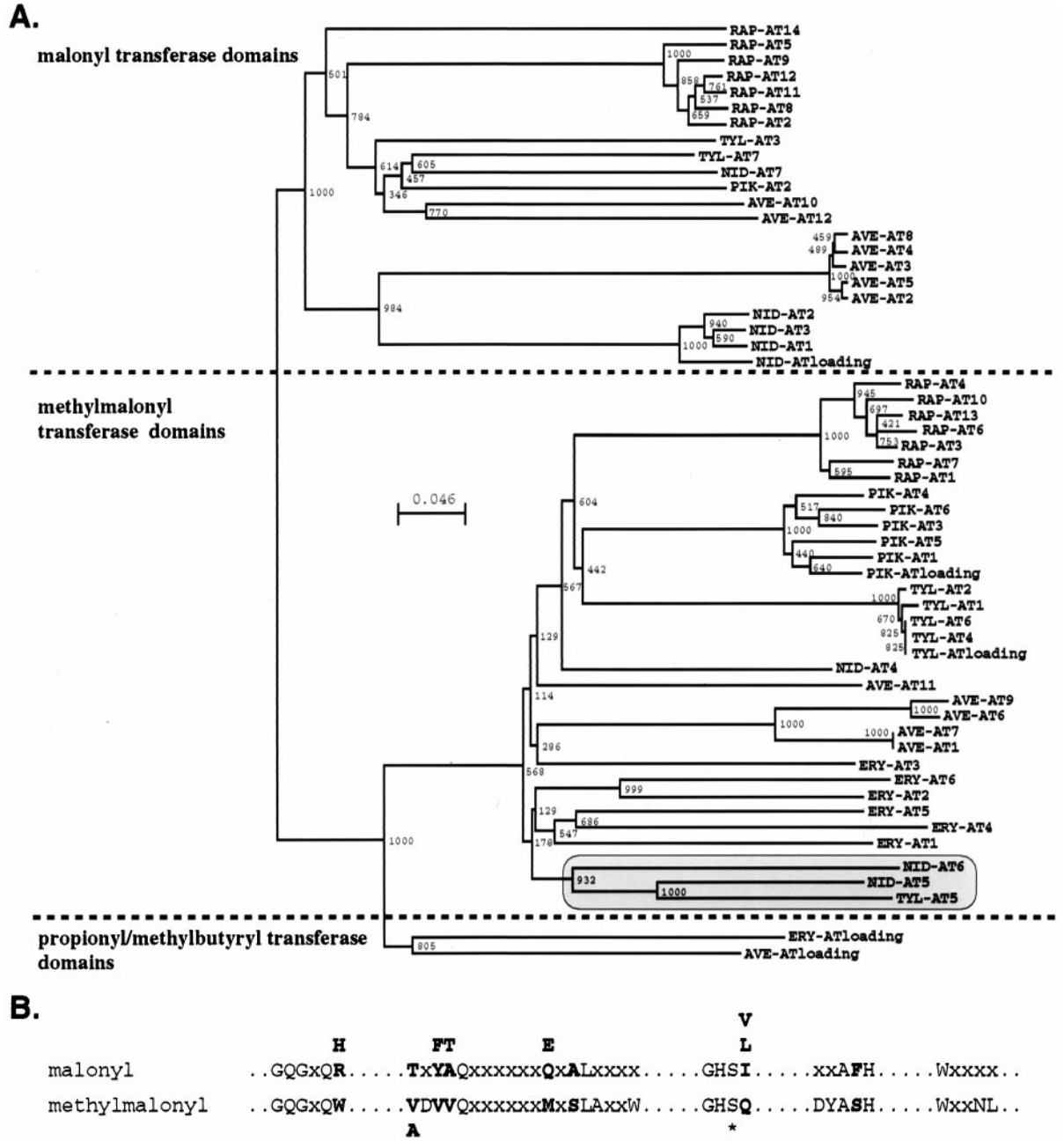

FIG. 3. Phylogenetic analysis of acyltransferases. (*A*) Phylogenetic tree of amino acid sequences of acyltransferase domains from actinomycete type I PKSs showing clustering of malonyl, methylmalony, or propionyl/methylbutyryl loading domain sequence. Domains in the shared box are ethylmalonate (TYL-AT5 and NID-AT5) or hydroxymalonate loading function (NID-AT6). Multiple alignment and phylogenetic analysis using the bootstrapping method were performed by using CLUSTALW. The number of amino acid substitutions is proportional to the length of the horizontal lines. Bootstrap tree is 1,000. AVE, avermectin PKS module; ERY, erythromycin PKS module; NID, niddamycin PKS module; PIK, pikromycin PKS module; RAP, rapamycin PKS module; TYL, tylosin PKS module. (*B*) Putative consensus sequences of malonyl and methylmalonyl loading domains. All letters shown represent invariant amino acids for the malonyl or methylmalonyl loading domain. Bold letters indicate significant differences between malonyl and methylmalonyl loading domains. The asterisk indicates the serine residue that is linked to the acyl-CoA in the acyl: acyltransferase complex.

a C5-ketoreductase similar to members of the family of short-chain alcohol dehydrogenases. The highest resemblance was found to 3-ketoacyl-ACP/CoA reductase (GenBank accession no. CAA07629) of *Streptomyces coelicolor* A3(2) (AveF: 47% identity), which is involved in the γ-butyrolactone synthesis. AveF also contained the NAD(P)H-binding consensus motif GXXXGXGXXXAXXXA (27) in its N terminus. Unlike erythromycin and pikromycin biosynthesis, avermectin biosynthesis requires the catalytic function of not only multifunctional PKS but also a non-PKS ketoreductase. Interestingly, disruption of the promoter region of *aveD* caused a lack of not only C5 *O*-methylation activity (*aveD*) but also C5 ketoreduction (*aveF*) (28), suggesting that both genes occupy a single transcription unit. The last ORF, *aveR*, lies immediately to the left of *aveF*. Mutants carrying transpositions of Tn*4560* were previously reported to be defective in avermectin biosynthesis and bioconversion of avermectin aglycons, and the transpositions in these mutants occurred to the left of the avermectin biosynthetic gene cluster and were accomplished by deletions in the region adjacent to the transposition (29). Nucleotide sequence analysis indicated that *aveR* is transcribed in the opposite direction to the *aveD-aveF* operon. The product

of *aveR* showed significant sequence identity with RapH in the rapamycin biosynthetic gene cluster (23), ORF6 in a second type-I PKS gene cluster of *Streptomyces hygroscopicus* ATCC29253 (30), and *cho* ORF3 of *Streptomyces* sp. (31). These gene products contain a sequence that is predicted to constitute a helix-turn-helix motif for DNA binding. The phenotype of *aveR* mutants, its complementation profile by introducing DNA segments, and sequence similarity to proteins carrying the helix-turn-helix motif indicated that AveR has a potential regulatory function. Beyond *aveR* lie genes that resemble an integral membrane and an ABC transport protein, and the mutants lacking these regions in the chromosome still produced avermectins, indicating the other probable limit of the avermectin biosynthetic cluster (data not shown).

With 18 genes spanning 82 kb of DNA capable of producing polyketide-derived avermectin (Table 1), in which one gene is not function in avermectin biosynthesis, the avermectin biosynthetic cluster represents the least complex yet most versatile modular PKS system reported to date. Information regarding the avermectin gene cluster is also useful for the elucidation of the other related anthelmintic compounds such as milbemycin and nemadectin. The comparison of avermectin PKS with other PKS in gene clusters for these related compounds should allow us to understand the natural evolution of biological catalysis. Such studies should facilitate the genetic manipulation of multistep biosynthetic pathways to generate molecular diversity in natural products.

We thank D. A. Hopwood for critical reading of the manuscript and valuable comments. We also thank S. Ouchi and M. Ukegawa for the excellent technical assistance. This work was supported by Grant-in-Aid of the Ministry of Education, Science and Culture Japan 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.).

1. Burg, R. W., Miller, B. M., Baker, E. E., Birnbaum, J., Currie, S. A., Hartman, R., Kong, Y.-L., Monaghan, R. L., Olson, G., Putter, I., *et al.* (1979) *Antimicrob. Agents Chemother.* **15,** 361–367.
2. Aziz, M. A., Diallo, S., Diop, I. M., Lariviere, M. & Parta, M. (1982) *Lancet* **ii,** 171–173.
3. Shikiya, K., Kinjo, N., Uehara, T., Uechi, H., Ohshiro, J., Arakaki, T., Kinjo, F., Saito, A., Iju, M. & Kobari, M. (1992) *Intern. Med.* **31,** 310–312.
4. Ikeda, H., Kotaki, H. & Ōmura, S. (1987) *J. Bacteriol.* **169,** 5615–5621.
5. MacNeil, T., Gewain, K. M. & MacNeil, D. J. (1993) *J. Bacteriol.* **175,** 2552–2563.
6. Ikeda, H. & Ōmura, S. (1993) *Actinomycetologica* **7,** 133–144.
7. Cane, D. E., Liang, T.-C., Kaplan, L. K., Nallin, M. K., Schulman, M. D., Hensens, O. D., Douglas, A. W. & Albers-Schönberg, G. (1983) *J. Am. Chem. Soc.* **105,** 4110–4112.
8. Shoop, W. L., Egerton, J. R., Eary, C. H., Haines, H. W., Michael, B. F., Mrozik, H., Eskola, P., Fisher, M. H., Slayton, L., Ostlind, D. A., *et al.* (1996) *Int. J. Parasitol.* **26,** 1237–1242.
9. Dutton, C. J., Gibson, S. P., Goudie, A. C., Holdom, K. S., Pacey, M. S., Ruddock, J. C., Bu'Lock, J. D. & Richards, M. K. (1991) *J. Antibiot.* **44,** 357–365.
10. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1999) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
11. Hanahan, D. (1985) in *DNA Cloning: A Practical Approach*, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 109–132.
12. Pang, C.-H., Shiiyama, M., Ikeda, H., Tanaka, H. & Ōmura, S. (1994) *Actinomycetologica* **8,** 21–25.
13. Ikeda, H., Wang, R.-L., Ohta, T. & Ōmura, S. (1998) *Gene* **206,** 175–180.
14. Strathmann, M., Hamilton, B. A., Mayeda, C. A., Simon, M. I., Meyerowitz, E. M & Palazzolo, J. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1247–1250.
15. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
16. Pearson, W. R. (1990) *Methods Enzymol.* **183,** 63–98.
17. Ishikawa, J. & Hotta, K. (1999) *FEMS Microbiol. Lett.* **174,** 251–253.
18. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Comput. Appl. Biosci.* **10,** 19–29.
19. Cortes, J., Haydock, S. F., Robert, G. A., Bevitt, D. J. & Leadlay, P. F. (1990) *Nature (London)* **348,** 176–178.
20. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. (1991) *Science* **252,** 675–679.
21. Xue, Y., Zhao, L., Liu, H.-W. & Sherman, D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 12111–12116.
22. August, P. R., Tang, L., Yoon, Y. J., Ning, S., Muller, R., Yu, T. W., Taylor, M., Hoffmann, D., Kim, C. G., Zhang, X., *et al.* (1998) *Chem. Biol.* **5,** 69–79.
23. Molnár, I., Aparicio, J. F., Haydock, S. F., Khaw, L. E., Schwecke, T., König, A., Stauton, J. & Leadlay, P. F. (1996) *Gene* **169,** 1–7.
24. Bevitt, D. J., Cortés, Haydock, S. F. & Leadlay, P. F. (1992) *Eur. J. Biochem.* **204,** 39–49.
25. Ikeda, H., Pang, C.-H., Endo, H., Ohta, T., Tanaka, H. & Ōmura, S. (1995) *J. Antibiot.* **48,** 773–775.
26. Pang, C.-H., Matsuzaki, K., Ikeda, H., Tanaka, H. & Ōmura, S. (1995) *J. Antibiot.* **48,** 59–66.
27. Persson, B., Krook, M. & Jórnvall, H. (1991) *Eur. J. Biochem.* **200,** 537–543.
28. Ikeda, H., Takada, Y., Pang, C.-H., Matsuzaki, K., Tanaka, H. & Ōmura, S. (1995) *J. Antibiot.* **48,** 95–97, 1995.
29. Ikeda, H., Takada, Y., Pang, C.-H., Tanaka, H. & Ōmura, S. (1993) *J. Bacteriol.* **175,** 2077–2082.
30. Ruan, X., Stassi, D., Lax, S. A. & Katz, L. (1997) *Gene* **203,** 1–9.
31. Horii, M., Ishizaki, T., Paik, S. Y., Manome, T. & Murooka, Y. (1990) *J. Bacteriol.* **172,** 3644–3653.