

## Nucleotide Sequence Analysis of a Second Set of the Polyketide Synthase $\beta$ -Ketoacyl synthase and Chain Length Factor Genes from the Salinomycin-producing *Streptomyces albus*

Chang-Gu Hyun, Kwan-Hyung Park,  
C. Richard Hutchinson<sup>1</sup> and Joo-Won Suh\*

Department of Biological Science, Biotechnology Research Institute,  
Myong Ji University, Yongin 449-728, Korea

<sup>1</sup>School of Pharmacy & Department of Bacteriology,  
University of Wisconsin, Madison, WI 53706

(Received November 22, 1996 / Accepted December 20, 1996)

The pWHM220 cosmid with a 24-kb insert cloned from *Streptomyces albus* ATCC 21838 induces the biosynthesis of a polyether antibiotic similar to salinomycin in *Streptomyces lividans*. We have analyzed this region by DNA sequencing as well as Southern blot hybridization with type I and type II polyketide synthase (PKS) probes. Surprisingly, we found another set of type II PKS genes only 10-kb from the original PKS genes, *salABCDE*. The DNA sequence revealed two complete open reading frames (ORFs) named *salB2* and *salC2*, and one partial ORF that does not resemble any known DNA or deduced protein sequence. The *salC2* should code for chain length determining factor while the deduced amino acid sequence encoded by *salB2* exhibits high similarity to  $\beta$ -ketoacyl synthase from different PKS gene clusters. The highest identity was found for  $\beta$ -ketoacyl synthases from *S. argillaceus* (MtmP, 59.1% identity), the mithramycin producer and from *S. venezuelae* ISP5230 (JadA, 52.3% identity), the jadomycin producer. The *SalC2* protein clearly resembles its counterparts in other aromatic PKS gene clusters that are believed to influence the length of the polyketide chain. The highest identities observed were to that of *S. argillaceus* (MtmK, 62.3%) and *S. venezuelae* ISP5230 (JadB, 55.1%) proteins. Moreover, the deduced amino acid sequences of the *salB2* and *salC2* products were 29.0% identical.

**Key words:** *Streptomyces albus*,  $\beta$ -ketoacyl synthase, chain length factor

Polyketides are a large group of structurally diverse secondary metabolites, several of which have applications as antibiotics, immunosuppressants, anticancer agents, and veterinary products. In recent years genetic studies have yielded detailed information on the organization and function of genes involved in the biosynthesis of polyketides in microorganism (16, 17).

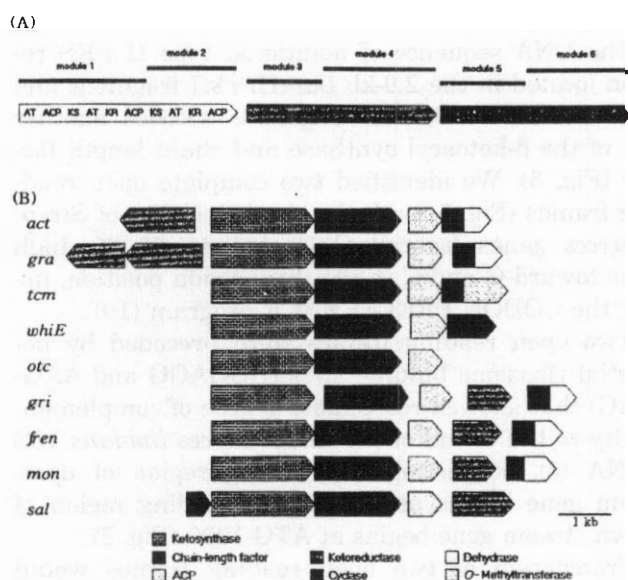
Cloning, sequencing, and functional analysis of these genes have improved our understanding of the genetic programming of polyketide synthases (16, 24, 25, 26). Two classes of PKSs have been studied from microorganisms (type I and type II (Fig. 1)). Various examples of the type II PKS have been identified (3, 5, 7, 8, 9, 12, 14, 15, 21, 22, 32, 35), which consists of a multienzyme complex formed by several polypeptides with individual enzymatic activities. In contrast, there is a paucity of information about type I PKS where the enzymatic activities are present as

separate domains in one multifunctional polypeptide. This is the case for the PKS involved in biosynthesis of the macrolide antibiotics, such as erythromycin (11) and rapamycin (30).

All of these metabolites share a common mechanism of biosynthesis that has largely resisted in vivo analysis. The carbon skeleton of a polyketide results from the sequential condensation of fatty acids like acetate, propionate, and butyrate. This process is catalyzed by polyketide synthases in a manner that is conceptually similar to the biosynthesis of long-chain fatty acids catalyzed by the fatty acids synthases found in all organisms (17).

We reported recently the type II PKS gene cluster from the salinomycin producer (8). Five of the genes cloned in pWHM220 were characterized by sequence analysis to confirm that the 3.9-kb DNA fragment contained type II PKS homologs (Fig. 2). The five open reading frames were identified by sequence analysis and named *salA-salE*. Since *salA*, *salB*, *salC*, *salD*, and *salE* genes encode proteins

\* To whom correspondence should be addressed



**Fig. 1.** Comparison of the organization between type I and type II PKSs. (a) Type I PKS for erythromycin biosynthesis. (b) The known gene clusters encoding type II PKSs. References: *act* (12), *fren* (5), *gra* (33), *gri* (35), *mon* (3), *otc* (21), *sal* (8), *tcm* (7), *whiE* (9).

highly similar to the following enzymes: cyclase,  $\beta$ -ketoacyl synthase, chain length determining factor, acyl carrier protein, and ketoreductase; it is very likely that these five *S. albus* genes govern the synthesis of some type of aromatic polyketide.

In this paper we report the additional PKS genes which was found by further sequencing of pWHM220. These new genes, named *salB2* and *salC2* showed homology with *tcmKL*(7). The two genes seems to encode  $\beta$ -ketoacyl synthase and chain length factor homologs, compose of minimal PKS together with acyl carrier protein.

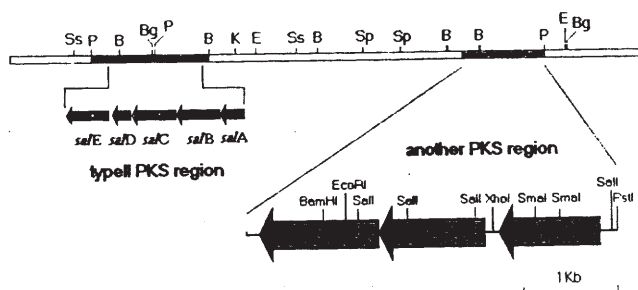
## Materials and Methods

### Bacterial strains, bacteriophage, and plasmids

*S. albus* ATCC21838 was obtained from the American Type Culture Collection (Rockville, Md). *Escherichia coli* DH5 $\alpha$  was purchased from Bethesda Research Laboratories, Inc. (Gaithersburg, Md) and was used for subcloning and sequencing. *E. coli* vector used for subcloning was pUC19 (GIBCO-BRL, Gaithersburg Md) and for sequencing, M13mp18 and M13mp19 were used (33). The cosmid clone pWHM220 has been described elsewhere (8). The cosmid vector pKC505 (27) was obtained from Richard Baltz, Eli Lilly Co. Ltd. (Indianapolis, Ind.).

### Media, enzymes, and chemicals

Bacterial strains were grown on LB medium where 50  $\mu$ g/ml ampicillin was added, when neces-



**Fig. 2.** Restriction map of 24-kb region of *S. albus* ATCC 21838 genomic DNA cloned in pWHM220.

sary. Restriction enzymes, ligase, and klenow enzyme were purchased from Takara Co. All enzyme reactions were carried out according to the recommendations of manufacturers, unless indicated otherwise. All chemicals were purchased from Sigma Chemical Co. (U.S.A.).

### DNA manipulation

The general DNA manipulation was carried out following Sambrook *et al* (28). Plasmid DNA and bacteriophage RF DNA were isolated from *E. coli* by using alkaline SDS extraction of cell lysate. A large-scale DNA purification was performed by banding in CsCl-ethidium bromide density gradients. Single-stranded DNA was isolated by precipitating phage particles. Agarose gel electrophoresis was performed in Tris-acetate or Tris-borate buffer.

### Nucleotide sequencing and analysis of sequence

Nested deletions were constructed with the Erase a Base system (Promega Biotech) according to the manufacturer's instruction. Derivatives of pUC19 and M13mp19 containing *S. albus* DNA were digested with *XbaI-PstI* and *BamHI-KpnI*, respectively before exonuclease III treatment.

Double-stranded templates were subcloned into M13mp18 and single-stranded templates were sequenced by the dideoxy chain-termination method using ( $\alpha$ - $^{32}$ P)dATP and Sequenase version 2.0 (United States Biochemicals) according to the manufacturer's instructions. To avoid compressions, 7-deaza-dGTP was used instead of dGTP. Labeled DNA was separated on 6% polyacrylamide wedge gels.

Primary DNA sequence data were analyzed and assembled by using software from DNASIS and BLAST programs (1,13). DNA and protein sequence were analyzed with the Genetics Computer Group software package (10).

## Results and Discussion

### Southern analysis of pWHM220 and isolation

The *actI* gene encoding a type II PKS involved in the biosynthesis of actinorhodin (12), an aromatic polyketide which has played a key role in the elucidation of polyketide synthesis mechanism and  $\beta$ -ketoacyl synthase domain of the *eryA* type I PKS gene(11) were used for Southern analysis of cosmid pWHM220 (Fig. 2). This established the presence of 1.4-kb, 2.4-kb, and 2.9-kb *Bam*HI-*Pst*I fragments that hybridized to *actI* probe, but very faintly to *eryKS*. Because previously described (8) type II polyketide synthase genes from *S. albus* were found in the 3.8-kb *Bam*HI fragment that contains one *Pst*I site and yields 1.4-kb and 2.4-kb *Bam*HI-*Pst*I fragments, we concluded that the 2.9-kb *Bam*HI-*Pst*I fragment should be located in different place from the 3.8-kb *Bam*HI fragment.

1	G	AGC	CTG	GCC	CAG	CAC	GTC	GGC	CGG	GCC	TGC	CGC	GGC	CAC	GTG	GAC	46
47	CGC	GCT	GCC	CAC	GGG	AGG	CGG	CCG	GTV	AAC	RCR	GTC	GTG	ATC	ACC	GGG	94
1																	8
95	ATC	GGC	GTC	GTC	GCC	CCC	GGC	GCG	GTV	GGC	ACC	GCC	RCR	TTC	TGG	GAC	142
1																	24
143	CTG	L	CTC	ACC	T	GTC	GGC	RCR	ACC	T	GCC	ATC	CGC	RCR	GTC	ACC	190
25	L																40
191	ACC	T	CGC	RCR	CTA	CGC	TCC	CGC	GTC	GCC	ACC	GAG	GTC	GAC	TTC	ACC	238
41	T																56
239	GCC	ACC	CAC	HCA	TGG	GAC	TCC	CCG	ACA	CCG	AAC	GCC	TGG	ACC	T	CGC	286
57																	72
287	GCA	CAG	TTC	GCA	CTG	L	GTC	V	GCC	ACA	RCR	GAA	GCC	ATC	GCC	AGC	334
73																	88
335	GTC	ACC	GCG	GAC	CGC	ATC	GGC	G	RCR	AAC	CCC	CTG	RCR	ACC	GGC	GTC	382
89	V																104
383	GCG	AGC	GCC	ACA	ATC	GGC	TGC	ACC	ACG	AGC	CTG	GCC	ACC	CAG	TAC	GCC	430
105																	120
431	CTC	L	AGC	GAC	TGC	GGC	T	ACC	T	TGG	ACC	CTC	GAC	CAC	ACC	T	478
121																	136
479	GAA	E	TCC	CTC	L	TAC	Y	GAC	D	TAC	TTC	GTC	ACC	CTC	CCT	GCC	526
137																	152
527	GCC	ACC	CGC	GAC	RCR	GGC	GCA	CAG	GGC	CCC	GTC	GCT	CTC	CTC	GTC	TCC	574
153																	168
575	TGC	C	ACC	TCC	GGC	CTG	L	GAC	D	GCC	ATC	GTC	GGC	CAC	GGC	GCC	622
169																	184
623	GAA	E	GGC	AGC	GCC	ACA	GAC	D	ATC	L	GTC	V	GTC	GCC	GGC	GGA	670
185																	200
671	GTG	V	CCC	ATC	GCC	ACA	ATG	M	GCC	TGC	TTC	GAC	CGC	CTG	RCR	CTC	718
201																	216
719	CGC	AAC	GAC	GAC	D	CCC	ACC	ACA	ATC	GCC	AGC	RCR	CCC	TTC	GAC	D	766
217																	232
767	GAC	D	GGA	TGC	TGC	TGC	S	CGC	ACA	AGG	RCR	GCG	CCG	G	CCG	TCC	814
233																	248
815	TGC	S	AAC	N	ACG	T	CCC	P	GCC	ACA	GCC	ACG	RCR	CCC	TAC	GGC	862
249																	264
863	GCG	GTC	ACC	GCC	ACA	CAC	AGC	AGC	GCC	CAC	CAC	ATG	ACG	GGC	GGC	CTG	910
265																	280
911	GGG	GCA	CTG	L	GAG	ATG	M	GCC	GAC	D	GCC	ATC	RCR	GCC	GCC	CTC	958
281																	296
959	CGG	RCR	CTG	L	AAC	CCC	P	GAC	GCC	D	GTC	V	GAC</				

The DNA sequence of additional type II PKS region located in the 2.9-kb *Bam*HI-*Pst*I fragment and adjacent 1.3-kb *Bam*HI fragment, encodes another set of the  $\beta$ -ketoacyl synthase and chain length factor (Fig. 3). We identified two complete open reading frames (Fig. 5) with the characteristics of *Streptomyces* genes (overall G+C content, 75.2%; high bias toward G and C in the third codon position, using the CODON PREFERENCE program (10)).

Translation of two open reading frames would result in a polypeptide with 411 amino acids and

1393	CCC AAC GGC CTG GGC ACC AGA GCC TGG TGG GAC GCC GTG CTG TGC GGG	1440
32	P N G L G T R A W W D A V L C G	47
1441 48	CGC R ACC T GGA C CTG L GGC CCC P ATC T ACC CGC TTC GAC CGT CGG R CTA C CC P CGP F	1488 63
1489 64	GTA R CGC R ATC I GCC GGC GAG E ATC CCC GGC TTC GTC GAC DAG GAC DAC HAC ATC I	1536 79
1537 80	CCC AGC S AGA R CTG L CTG CCC TCC ATC ACC TACC GAC CGC GGC C ACC TC GCG ATC I GCC ATG L	1584 95
1585 96	GTC GCG GCC GAA E GAA GCA CTG CGC GAC GCT AAC GTG AGC CGC P ACC GCC GAC D	1632 111
1633 112	CTC CCG GCA Y GGC GGC GGC GTG ATC ACC GCC AGS TCC SCG GA GGC GGC G	1680 127
1681 128	GCA GAA E TTC GGC GAA CGG GGA CTG GCC GCA CTG TWG AGC AAA GGC GCC A	1728 143
1729 144	CAG CAG H GTC S AGC GCC ATY CAG TCC TFC GCG TCC TFC CAC GCG GCA GCC A	1776 159
1777 160	CGC GCA CAG ATC TCC ATC CGG CAC CGG CTG CGC GGC CAC GGC TCG ACC T	1824 175
1825 176	GTC GTC S AGC S GAA E CAG A GCC GGC GGC ATC GAC GCA CTC ACC CGC GCC CGR R	1872 191
1873 192	CGG CGG ATC R CGC GAC GGG GCA CTG CTC ATG GTC ACC GGC GGC ATC GAC D	1920 207
1921 208	TCC ATA CTG L TGC GCA TGG GGC TGG GCC GCG CAC CTG GCG GAC GGC CGG R	1968 223
1969 224	TTC AGC S CCC GGC ACC GAA CCC CGC CGG GCC TAC CGG CCC TTC GCG GCC A	2016 239
2017 240	ACG GCC GAC GGC CAC GCG GTC GGC GAG GGC GGC ACC CTG LTC GTC CTG L	2064 255
2065 256	GAG GAC D GCC CGG GCC ACC GCC CGC CGG GGC GCC ATT GGC TAY GGC GTC V	2112 271
2113 272	ATC GGC GGC TGC GCC ACC ACC TTC GAC GGC CCC GAC DCG CCP ACA CTG L	2160 287
2161 288	CGC CAG A GCC GCG GAA LE CTC GGC ATG GCC AAG GCC GGC CTG GCC CCC GAA E	2208 303
2209 304	CAC GTG GAT D GTG GTT TTC CCC CGA R ACC GGC GCC CCG GAG R CGC R GCC GAC D	2256 319
2257 320	CTC L GTC V GAG S CAG A CCG CTG L TGC CCG A CTG TTC GGA P CCC TAY GGA GTA V	2304 335
2305 336	CGC GTC V ACC T GTG CCG AAG ACC T ATG ACC GGC CGG CTG GGC GAG A GGC GGC G	2352 351
2353 352	TGS GCC A CTG GAC GTG GCA ACC GCA CTG CTC GCC ATC CTG CGC GAG E AAG GTC V	2400 367
2401 368	GTA CCC CCG ACC ATT GGA ACC GGA CGC GTC GCC GAC GAG TGC CGC CTG L	2448 383
2449 384	GAC CTG GTC ACC T GGA GCC CCA P CGG GAC TGC CCG GGC TGC GGG TG CGC R	2496 399
2497 400	TGG GTA CTG L GCC CGC GGA R CGG GGC GGC TTC AAC TSC GCC A GCG ATG CTC L	2544 415
2545 416	CGG GGC CCT C P CAG ACQ CAG TGA CGG CCC GCC CCT AAC GAG ACG GAG CGG G	2592 422
2593	AAC ACA CGC GCG GCC AGG GAC GCC GGC CCC GCG GGC CGC GCA AAG GGA	2640
2641	AAC ACC CCG GCG SCC GAC TCC CCG CGC CTC CCG GCG CTC CTC GGC CGT	2688
2689	TCC GCG TTC GGC CTT TCG CCT TCC GGC CTC GCG CTC CTT TGC CTG C	2732

**Fig. 3.** Nucleotide sequence of a 2,732-bp region encoding *salB2* and *salC2*. The predicted peptide sequence is shown below the nucleotide sequence. Putative ribosome binding sites are underlined.



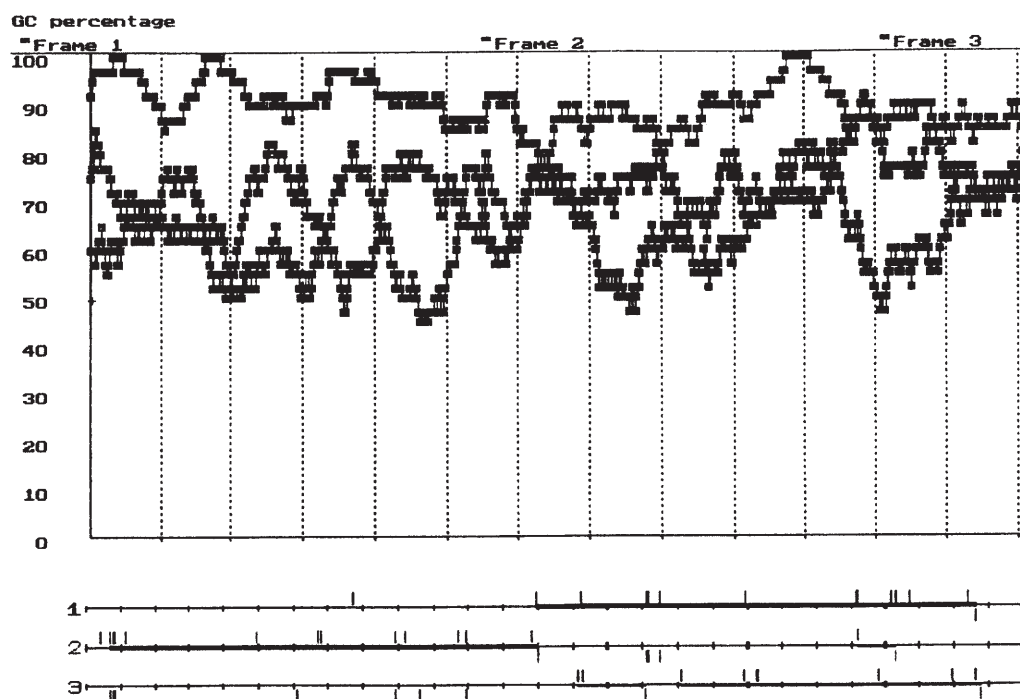


Fig. 4. Analysis of guanine and cytosine content at third positions, for 2,732-bp with the *salB2* and *salC2* genes of pWHM220. ATG and GTG codons are indicated by vertical bars, stop codons are designated by crosses.

422 amino acids, respectively, and a molecular weight of 43,621 and 43,738, and a calculated isoelectric point of 7.72 and 9.60, respectively. The gene organization for the two genes seemed to be translationally coupled.

### Deduced function of the proteins

The two new genes cloned in pWHM220 were proved to have homology with *tcmKL* by sequence analysis. TcmK and TcmL compose minimal PKS for tetracenomycin biosynthesis (7). The separate components of the minimal PKS are ketoacyl synthase (KS) which catalyzes the condensation reac-

tion and may also carry an acyltransferase (AT) domain for loading the starter unit to the KS, a second protein with a high degree of homology to the KS which is involved, at least in part, in determination of the polyketide chain length, and an acyl carrier protein (ACP).

The two open reading frames were identified by CODONPREFERENCE and named *salB2*, *salC2*. Sequence comparisons with GAP program (10) indicated that SalB2 polypeptide bears strong similarity to many other  $\beta$ -ketoacyl synthases. The similarity is especially strong in the region surrounding 169-Cys and 347-Ser, which corresponds to the

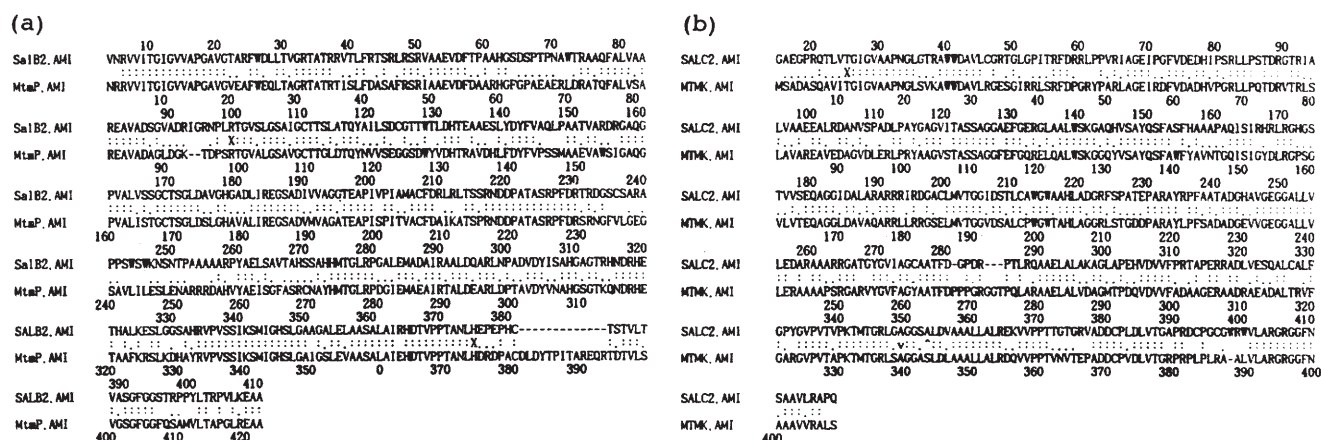


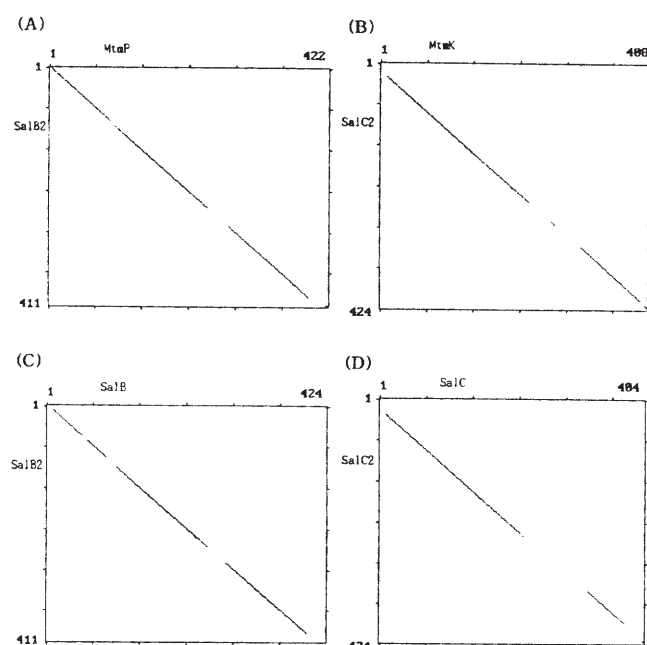
Fig. 5. Amino acid sequence comparisons using Genetic Computer Group program. (a) *SalB2* and *MtmP*,  $\beta$ -ketoacyl synthase for mithramycin biosynthesis (59.1% identity); (b) *SalC2* and *MtmK*, chain length factor from *S. argillaceus* (62.3% identity).

A. $\beta$ -ketoacyl synthase active site		Ref
<i>S. albus</i> SalB2	RGAGPVALVSSGTSGLDAVGHGA	this study
<i>S. argillaceus</i> MtmP	IGAAGPVALISTGTSGLDSLGHAV	22
<i>S. venezuelae</i> JadA	VGAEGPNTVVSTGTSGLDSVGYAR	15
<i>S. glaucescens</i> TcmK	AGAEGPVTIVSTGTSGLDAVGYGT	7
<i>S. griseus</i> Gri	VGAEGPATVVSTGTSGLDAVGHAV	35
<i>S. rimosus</i> Otc	AEAEPAGVVSAGTSGLDVLTHAA	21
<i>S. violaceoruber</i> Gra	AGAEGPVTIVSDGTSGLDSVGYAV	32
<i>S. peucetius</i> Dau	AGAEGPVTIVSAGTSGLDSIGYAC	14
<i>S. coelicolor</i> Act	VGAEGPVTIVSTGTSGLDSVGNVA	12
<i>S. albus</i> SalB	VGAEGPSTVVSTGTSGLDSVGYAV	8
Active site system	*A*GP *S G G T*G ** *	

B. Acyltransferase active site		Ref
<i>S. albus</i> SalB2	PVSSIKSMIGHSLGAAGALELAASAL	this study
<i>S. argillaceus</i> MtmP	PVSSIKSMIGHSLGAIGSLEVAASAL	22
<i>S. venezuelae</i> JadB	PVSSIKSMVGHSLGAIGSIEIAASAL	15
<i>S. glaucescens</i> TcmK	PVSSIKSMIGHSLGAIGSLELAACAL	7
<i>S. griseus</i> Gri	PVSAIKSMVGHSLGAIGSIEIAACAL	35
<i>S. rimosus</i> Otc	PISSIKSMIGHSLGAIGSLEVAASAL	21
<i>S. violaceoruber</i> Gra	PVSSIKSMIGHSLGAIGSIEIAASVL	32
<i>S. peucetius</i> Dau	PISSIKSMIGHSLGAVGSLEVAATAL	14
<i>S. coelicolor</i> Act	PVSSIKSMVGHSLGAIGSLEIAACVL	12
<i>S. albus</i> SalB	PVSSIKSMVGHSLGAIGSIEIAASAL	8
Active site system	P*S**KSM GHSLGA **E*AA* L	

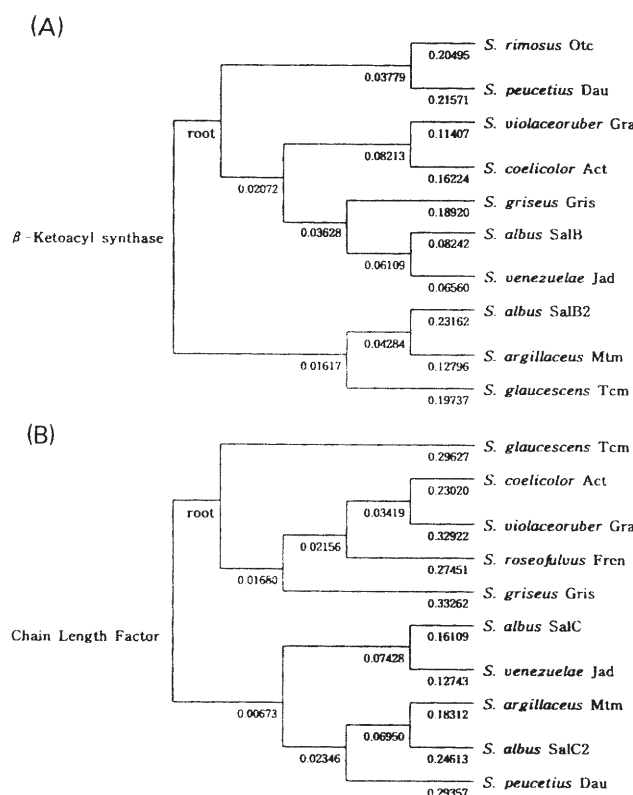
**Fig. 6.** Sequence alignment of amino acids of  $\beta$ -ketoacyl synthase (KS) and acyltransferase (AT) with conserved active sites of other KS and AT, respectively. Conserved amino acids for all peptides are shown under each alignment. Amino acids at active sites are marked in box. References for each peptide are shown.



**Fig. 7.** COMPARE-DOTPLOT analysis of the deduced protein products of (A) SalB2 versus MtmP, 59.1% identity; (B) SalC2 versus MtmK, 62.3% identity; (C) SalB2 versus SalB, 54.9% identity; (D) SalC2 versus SalC, 54.8% identity. A window size of 30 was used at a stringency of 15.

active site in  $\beta$ -ketoacyl synthase enzyme (Fig. 6).

The deduced *salC2* product resembles several



**Fig. 8.** Dendrogram showing relationships of type II PKS components,  $\beta$ -ketoacyl synthase and chain length determining factor. PILEUP program of GCG (10) was used to analyze the data.

chain length factor with the following percentages of identical amino acids (Fig. 4, 7): MtmK for mithramycin biosynthesis, 62.3% (22), JadB for jadomycin biosynthesis, 55.1% (15), and SalC for polyketide biosynthesis in *S. albus*, 54.8% (8).

All of the genes are transcribed in the same direction. A phylogenetic tree, involving a large set of sequence including those of *salB2* and *salC2*, is shown in Fig 8. Recent studies have shown that combination of genes from different PKSs can be expressed to produce functional hybrid PKS. The PKS gene set identified in this study is an important addition to the range of available PKSs and will enable further analysis of the molecular basis for PKS programming to be made.

### Relationship between the two polyketide gene clusters

We had previously shown that *S. albus* has a gene cluster (*salA-E*) encoding type II PKS. And this region could complement mutations in the *tcmKL* region (8). In this study, we showed that pWHM220 cosmid clone had another copy of genes for  $\beta$ -ketosynthase and chain length factor. The presence of multiple copies of the same or similar genes in

*Streptomyces* is not usual but not odd (11, 23).

Here we report a second set of PKS genes (*salB2*, *salC2*) in *S. albus* that are similar to another set, the *salBC* genes from another PKS gene cluster (8); *salB2* encodes the  $\beta$ -ketoacyl synthase and *salC2* encodes chain length factor in the same orientation. Since the original PKS gene cluster and the additional PKS gene cluster are close each other (only 10-kb apart) and have the same orientation, we speculate that these two gene clusters might have arisen by tandem genetic duplication (2).

However, we do not exclude that the two PKS gene clusters have different functions. Since *act*- and *whiE*-PKS genes, the two sets of related genes, are expressed at different stages in the life-cycle of the organism, *S. coelicolor*, to produce antibiotic and spore pigments, respectively. It was found that each of the three subunits of the *whiE* minimal PKS could complement lesions in the *act*-PKS to produce actinorhodin. Conversely, the corresponding *act*-PKS subunits would complement mutations in the *whiE* locus to restore spore pigmentation (34).

## Acknowledgments

This work was supported by a grant of Genetic Engineering Research Program of the Ministry of Education, Korea, in 1995.

## References

1. Altschul S.F., W. Gish, W. Miller, E.W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
2. Anderson, R.P. and J.R. Roth. 1977. Tandem genetic duplication in phage and bacteria. *Annu. Rev. Microbiol.* **31**, 473-506.
3. Arrowsmith, T.J., F. Malpartida, D.H. Sherman, A. Birch, D.A. Hopwood, and J. A. Robinson. 1992. Characterization of actI-homologous DNA encoding polyketide synthase genes from the monensin producer *Streptomyces cinnamonensis*. *Mol. Gen. Genet.* **234**, 254-64.
4. Ben Shen and C. Richard Hutchinson. 1993. Enzymatic synthesis of a bacterial polyketide from acetyl and malonyl coenzyme A. *Science* **262**, 1535-1540.
5. Bibb, M.J., D.H. Sherman, S. Omura, and D.A. Hopwood. 1994. Cloning, sequencing and deduced functions of a cluster of *Streptomyces* genes probably encoding biosynthesis of the polyketide antibiotic frenolicin. *Gene* **142**, 31-39.
6. Bibb, M.J., P.R. Findlay, and M.W. Johnson. 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* **30**, 157-66.
7. Bibb, M.J., S. Biro, H. Motamedi, J.F. Collins, and C.R. Hutchinson. 1989. Analysis of the nucleotide sequence of the *Streptomyces glaucescens tcmI* genes provides key information about the enzymology of polyketide antibiotic biosynthesis. *EMBO J.* **8**, 2727-2736.
8. Borell, Chris W., Y.G. He, H.J. Kwon, C.G. Hyun, K. Fumio, J.W. Suh, H.J. Jin, H.G. Floss and C.R. Hutchinson. 1996. Production of antibiotics salinomycin by introduction of a gene cluster containing type II polyketide synthase genes from *Streptomyces albus* ATCC 21838 into *Streptomyces lividans*. *Microbiology* (UK) in submitted.
9. Davis, N.K. and K.F. Chater. 1990. Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Mol. Microbiol.* **4**, 1679-1691.
10. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.
11. Donadio, S., M.J. Staver, J.B. McAlpine, S.J. Swanson and L. Katz. 1991. Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.
12. Fernandez-Moreno, M.A., E. Martinez, L. Boto, D.A. Hopwood, and F. Malpartida. 1992. Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J. Biol. Chem.* **267**, 19278-19290.
13. Gish, W. and D.J. States. 1993. Identification of protein coding regions by database similarity search. *Nature Genet.* **3**, 266-272.
14. Grimm, A., Madduri, K., Ali, A. and Hutchinson, C. R. 1994. Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase. *Gene* **151**, 1-10.
15. Han, L., Yang, K., Ramalingan, E., Mosher, R.H. & L.C. Vining. 1994. Cloning and characterization of polyketide synthase genes for Jadomycin B biosynthesis in *Streptomyces venezuelae* ISP5230. *Microbiology* **140**, 3379-3389.
16. Hopwood, D.A., C. Khosla, D.H. Sherman, M.J. Bibb, S. Ebert-Khosla, E.-S. Kim, R. McDaniel, W.P. Revell, R. Torres, T.-W. Yu. 1993. Toward an understanding of the programming of aromatic polyketide synthases: a genetics-driven approach, p. 267-275. In G. D. Hegeman, R.H. Baltz, and P.L. Skatrud (ed.), *Genetics and molecular biology of industrial microorganisms*. American Society for Microbiology, Washington, D. C.
17. Hopwood, D.A., and D.H. Sherman. 1990. Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* **24**, 37-66.
18. Hutchinson, C.R. and I. Ikeda, I. 1995. Polyketide synthase gene manipulation; A structure-function approach in engineering novel antibiotics. *Annu. Rev. Microbiol.* **49**, 201-238.
19. Khosla, C., R. McDaniel, S. Ebert-Khosla, R. Torres, D.H. Sherman, M.J. Bibb, and D.A. Hopwood. 1993. Genetic construction and functional analysis of hybrid polyketide synthases containing heterologous acyl carrier proteins. *J. bacteriol.* **175**, 2197-2204.
20. Kim, E.S., D.A. Hopwood, and D.H. Sherman. 1994.



- Analysis of type II  $\beta$ -ketoacyl synthase specificity in *Streptomyces coelicolor* A3(2) by *trans* complementation of actinorhodin synthase mutants. *J. Bacteriol.* **176**, 1801-1804.
21. Kim, E.S., M.J. Bibb, M.J. Butler, D.A. Hopwood, and D.H. Sherman. 1994. Sequences of the oxytetracycline polyketide synthase-encoding *otc* genes from *Streptomyces rimosus*. *Gene* **141**, 141-142.
  22. Lombo, F., G. Blanco, E. Fernandez, C. Mendez, and J.A. Salas. 1996. Characterization of *Streptomyces argillaceus* genes encoding a polyketide synthase involved in the biosynthesis of the antitumor mithramycin. *Gene* **172**, 87-91.
  23. Magdalena, J., M. Forsman, M.V. Lenzini, A. Brans, and J. Dusart. 1992. Two different  $\beta$ -lactamase genes are present in *Streptomyces cacaoi*. *FEMS Microbiol. Lett.* **99**, 101-106.
  24. McDaniel, R., S. Ebert-Kosla, D.A. Hopwood, and C. Khosla. 1993. Engineered biosynthesis of novel polyketides. *Science* **262**, 1546-50.
  25. McDaniel, R., S. Ebert-Kosla, D.A. Hopwood, and C. Khosla. 1993. Engineered biosynthesis of novel polyketides: manipulation and analysis of an aromatic polyketide synthase with unproven catalytic specificities. *J. Am. Chem. Soc.* **115**, 11671-11675.
  26. McDaniel, R., S. Ebert-Kosla, D.A. Hopwood, and C. Khosla. 1995. Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* **375**, 549-54.
  27. Richardson, M. A., S. Kuhstoss, P. Solenberg, N. Schaus, and R.N. Rao. 1987. A new shuttle vector, pKC505, for streptomycetes: its use in the cloning of three different spiramycin-resistance genes from a *Streptomyces ambofaciens* library. *Gene* **61**, 231-241.
  28. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2nd ed., Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.
  29. Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
  30. Schweche, T., J.F. Aparicio, I. Molnar, A. Konig, L. E. Khaw, S.F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J.J.B. Lester, G.A. Bohm, J. Staunton and P. F. Leadley. 1995. The biosynthesis gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. USA* **93**, 7839-7843.
  31. Sherman, D.H., E.-S. Kim, M.J. Bibb, and D.A. Hopwood. 1992. Functional replacement of genes for individual polyketide synthase components in *Streptomyces coelicolor* A3(2) by heterologous genes from a different polyketide pathway. *J. bacteriol.* **174**, 6184-6190.
  32. Sherman, D.H., F. Malpartida, M.J. Bibb, H.M. Kieser, and D.A. Hopwood. 1989. Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tu22. *EMBO J.* **8**, 2717-2725.
  33. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors *Gene* **33**, 103-119.
  34. Yu, T.W. and D.A. Hopwood. 1995. Ectopic expression of the *Streptomyces coelicolor* *whiE* genes for polyketide spore pigment synthesis and their interaction with the *act* genes for actinomorphin biosynthesis. *Microbiology* **141**, 2779-2791.
  35. Yu, T.W., M.J. Bibb, W.P. Reville, and D.A. Hopwood. 1994. Cloning, sequencing, and analysis of the griseusin polyketide synthase gene cluster from *Streptomyces griseus*. *J. Bacteriol.* **176**, 2627-34.