Manipulation of Modular Polyketide Synthases

Leonard Katz

Abbott Laboratories, Abbott Park, Illinois 60064-3500

Received May 22, 1997 (Revised Manuscript Received August 28, 1997)

Contents

Leonard Katz was born in Montreal, Canada, in 1943. He received a B.Sc. from McGill University in 1965 and a Ph.D. from the University of California at Santa Barbara in 1970. He was a Postdoctoral Fellow at the University of California, San Diego, from 1970 to 1974 and Assistant Professor of Biology at New York University from 1974 to 1977. After a brief tenure at Schering-Plough, Inc., he joined Abbott Laboratories in Illinois in 1979 where he is currently Senior Project Leader in Antibacterial Discovery Research. He also holds the position of Research Fellow in the Volwiler Research Society at Abbott. He has had a long-term interest in the genetics and biochemistry of the biosynthesis of erythromycin and in the use of genetics to change the structure of the antibiotic. Current interests also include the use of bacterial genomics in the discovery of novel antibacterial agents.

I. Introduction

Polyketides are a class of natural products synthesized by bacteria, fungi, and plants through the successive condensation of simple carboxylic acids[1] that have widespread use in human and veterinary medicine, agriculture, and animal nutrition. The group includes plant flavanoids, fungal aflatoxins, as well as hundreds of compounds that can inhibit the growth of bacteria, viruses, fungi, parasites, or human tumor cells. A number of antifungal polyketides also have immunosuppressive activity. Polyketides can vary widely in structure, from the eight-carbon compound 6-methylsalicylic acid to brevitoxin B, which contains 50 carbon atoms in its chain.

Although diverse in structure and properties, polyketides can be grouped into two overall classes: the aromatic and the complex polyketides. Aromatic polyketides are produced through the condensation of acetate groups (with the exception of the starter unit) which generate (β-) keto groups that remain largely unreduced during and after growth of the acyl chain. The acyl chains, either during or after completion of their synthesis, undergo enzymatic programmed or spontaneous folding to allow aldol condensations to take place resulting in the formation of 6-membered rings. The aromatic rings are subsequently reduced through enzymatically controlled dehydrations. Examples of aromatic polyketides include actinorhodin, daunorubicin, tetracenomycin, and the tetracyclines.

The complex polyketides are structurally more diverse than the aromatic family and, unlike the mainly acetate-derived aromatics, are composed of acetates, propionates and butyrates in varying ratios. Furthermore, as will be discussed in section II, because of a fundamental difference from the aromatic group in the chemistry of their synthesis, processive β-carbonyl reduction, as well as the structural constraints imposed by the presence of methyl side chains, these molecules cannot undergo folding and aromatization. Many of the compounds cyclize through lactonization; others remain as long acyl chains. Examples of complex polyketides, shown in Figure 1, are the macrolide antibiotics erythromycins A, B, C, and D (**1**, **2**, **3**, and **4**, respectively) and spiramycin I (**5**), the antifungal compound rapamycin (**6**) (which also exhibits immunosuppresssive activity), the anthelmintic compound avermectin $A1_a$ (**7**), the antifungal polyene amphotericin B (**8**), and the

S0009-2665(96)00025-8 CCC: $28.00

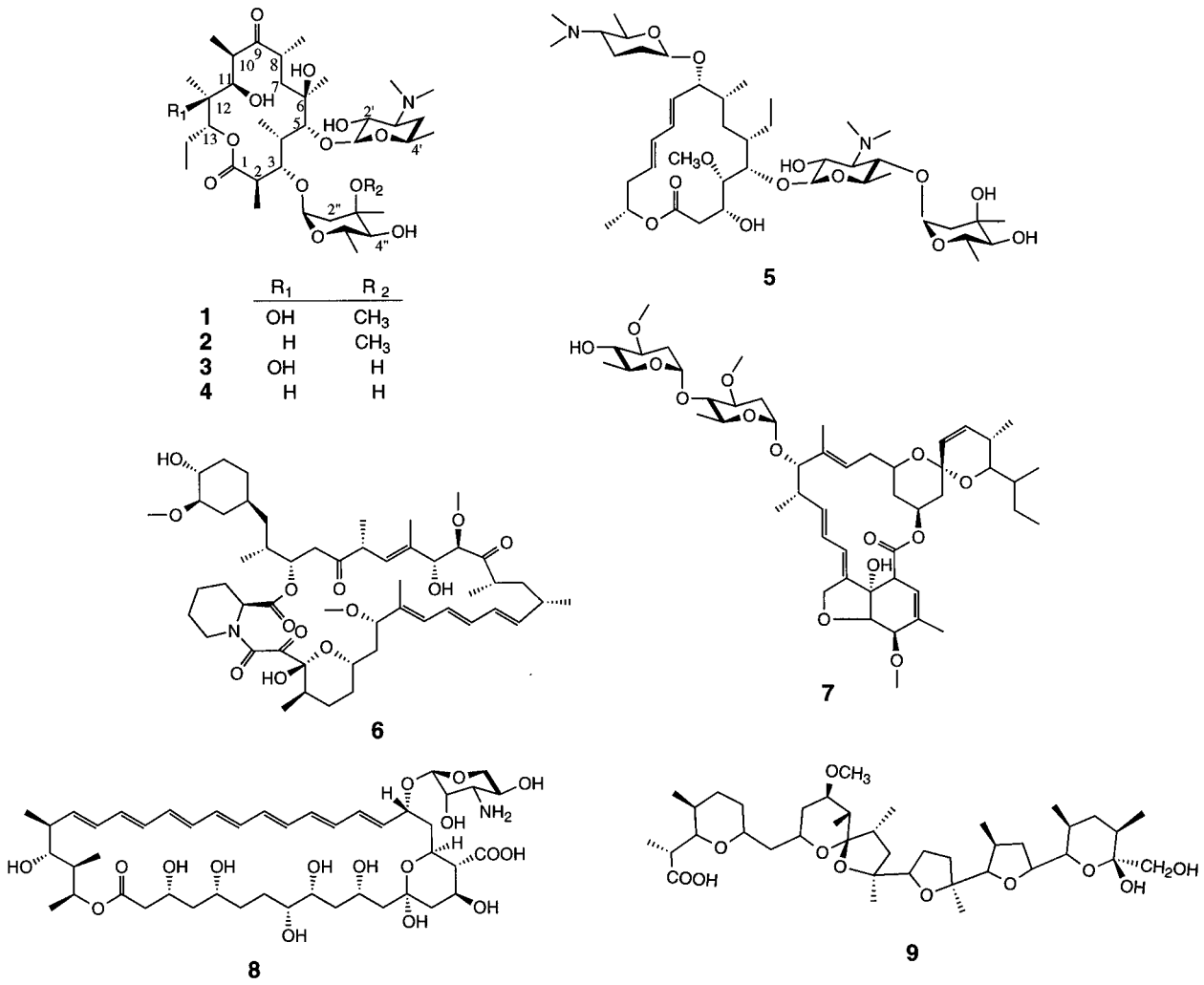


Figure 1. Structures of complex polyketides.

polyether compound nigericin (**9**). Although the chemistry of complex polyketide synthesis has been understood for some time, in the last five years significant leaps in our understanding of the genetics and biochemistry of complex and aromatic polyketide biosynthesis have taken place. The processes of complex polyketide biosynthesis will be described in detail in section II with emphasis placed on the synthesis of polyketide component of erythromycin to provide a basis for the discussion, in section IV, of the rationally designed changes in polyketide structure that have been made by manipulation of the polyketide biosynthetic genes.

It can be seen that several of the compounds are highly modified after acyl chain synthesis and may be linked to sugars. In general, synthesis of the acyl chain is only one component in the synthesis a very complex molecule. The best-studied example is erythromycin, produced by the Gram positive mycelia-forming bacterium *Saccharopolyspora erythraea*. Erythromycin is composed of a 14-membered macrolactone ring to which are attached the 6-deoxysugars D-desosamine at C-5 and L-cladinose at C-3. In addition to the two glycosidations, the polyketide ring also undergoes P450-mediated hydroxylation at C-6 and C-12. As will be described in section III, synthesis of the deoxysugars and modification of the polyketide backbone are highly controlled, but some of the steps in the biosynthesis of the final product of the pathway may be bypassed. The impact of changing the polyketide structure through genetic manipulation on subsequent post polyketide modification is discussed.

This review focuses exclusively on the complex polyketides and does not provide any discussion of the elegant work conducted in many laboratories on the genetics and biochemistry of aromatic polyketide biosynthesis. Recent reviews on these subjects have appeared.[2−4]

II. Complex Polyketide Biosynthesis

A. Chemistry

Polyketide synthesis chemically and biochemically resembles the synthesis of fatty acids. Synthesis of both are initiated by a Claisen condensation between a starter carboxylic acid and a dicarboxylic acid. The example shown in Figure 2 uses propionate (**10**) and methylmalonate (**11**) as enzyme-linked thioesters as starter and extender, respectively. As will be discussed in greater detail below, the starter unit is linked via a thioester to the cysteine residue in the active site of the enzymatic unit that carries out the condensation reaction, abbreviated as the ketosynthase. The extender unit is linked via a thioester to the pantotheine side chain which itself is attached as a phosphoester to a serine residue in an acyl carrier (ACP) unit. In polyketide synthesis, the KS

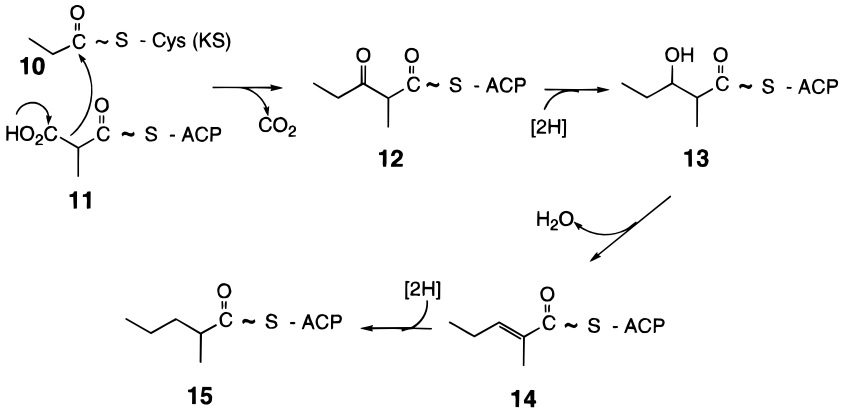


Figure 2. Reduction cycle of β-carbonyl formed from Claisen condensation of carboxylic acids. Key: ACP, acyl carrier protein domain of polyketide synthase; KS, β-ketoacyl ACP synthase domain of polyketide synthase.

and ACP are components of the large polyketide synthase (PKS), discussed in more detail below; in fatty acid synthesis they are components of the fatty acid synthase (FAS). The resulting diketide (**12**) is two carbons longer than the starter and carries an α methyl side chain from the methylmalonyl moiety. The chain grows by two carbons, therefore, at each condensation cycle. As is shown in Figure 2, the diketide is linked to the ACP. The energy supplied by the exergonic decarboxylation of the extender permits breakage of the thioester bond in the starter and formation of the diketide. In both polyketide and fatty acid synthesis, the next step is reduction of the β-carbonyl which takes place after every condensation step in fatty acid synthesis but is not obligatory in polyketide biosynthesis. The scheme for full β-carbonyl reduction for the diketide **12** is therefore shown only to illustrate the chemistry; it does not indicate the necessity for the process at any given step. Full reduction is a three-step process requiring three distinct enzymatic functions: first, ketoreduction to produce the secondary alcohol (**13**) where an electron is supplied by NADPH, followed by dehydration leaving the α,β unsaturation in the polyketide chain (**14**). The final step is enoylreduction, again employing NADPH as electron donor, to result in the production of a methylene function at the β-carbon (**15**). Once the reduction cycle is completed, the nascent chain enters the next cycle in the position on the enzyme equivalent to that of a starter to condense with the next extender unit.

Two significant differences exist between fatty acid and complex polyketide synthesis. First, fatty acid synthesis only uses malonyl moieties as extender units to build an acyl chain and usually acetate as starter; the fatty acid synthase may also use a branched chain carboxylate as the starter unit. In contrast, polyketide synthesis in bacteria can use malonyl, methylmalonyl, ethylmalonyl, and, for a single condensation in the synthesis of ascomycin, propylmalonyl units as extenders. Thus, not only must the polyketide synthase choose the correct extender to use for incorporation at every step of the synthesis, the correct sterochemistry must also be preserved. In fungi, malonyl units are used as extenders and side chains are added by C-methylation using S-adenosylmethionine to donate the methyl groups. Secondly, in polyketide synthesis, a full cycle of β-carbonyl reduction (ketoreduction, dehydration, and enoylreduction) does not necessarily take place after every condensation as in fatty acid synthesis. This is best illustrated in reviewing the structure of erythromycin (**1**). The keto group at C-9 results from the lack of ketoreduction after the third condensation cycle in the building of the polyketide backbone. The hydroxyls that are present at C-13 (the lactone oxygen), C-11, C-5, and C-3 result from ketoreduction after the first, second, fifth, and sixth condensations and the failure to undergo the subsequent dehydration steps. Finally, the methylene function at C-7 results from the full cycle of reduction after the fourth condensation step.

Although it was thought that β-carbonyl processing takes place after each step in polyketide synthesis (as has been known for many years for fatty acid synthesis), only recently has definitive evidence for processive synthesis emerged. Cane and Yang showed that *N*-acylcysteamine thioesters, five carbons in length, carrying the correct degree of reduction of the carbonyl functions and correct stereochemistry (Figure 3, **16**) were incorporated intact into erythromycin when fed *to Sac. erythraea*[5] and Yue *et al.* showed that a seven-carbon thioester could be incorporated intact into tylosin.[6] Subsequent work showed intact incorporation of *N*-acylcysteamine thioesters of short to intermediate length acyl chains of correct proposed structure (**18** and **20**) into nargenicin (**19**),[7] methymycin (**21**),[8] nonactin,[9] aspyrone,[10] or dehydrocurvularin.[11,12] Further evidence was obtained when acyl chain intermediates with the correct stereochemistry of the side chains and the correct degree of processing and of sizes corresponding to two through six condensation cycles were recovered from fermentation broths of mutants blocked in the synthesis of the polyketide components of mycinamycin,[13,14] tylosin,[15] or erythromycin.[16] The compound produced in the case of erythromycin (**17**) is consistent with the proposed product of the first two condensation steps and correct degree of reduction in the synthesis of the erythromycin polyketide backbone, followed by release of the nascent chain from the polyketide synthase and spontaneous lactonization. This compound was also made by feeding **16** to a host carrying an altered polyketide synthase that could only permit two condensation steps indicating that reduction at the second cycle must have taken place immediately after condensation.[17]

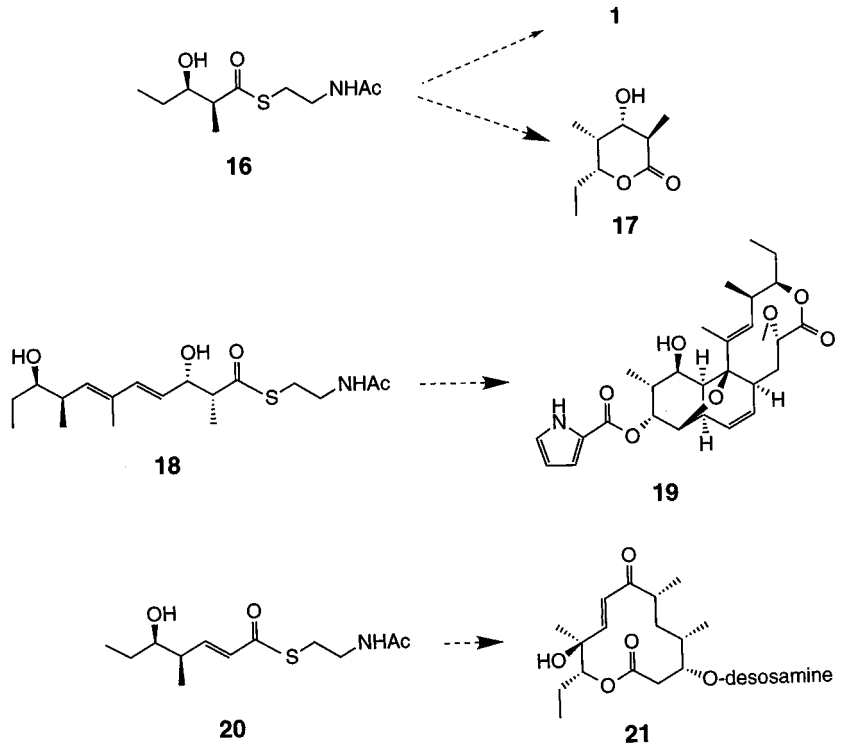


Figure 3. Examples of incorporation intact of *N*-acylcysteamine thioesters of acyl chains into polyketides.

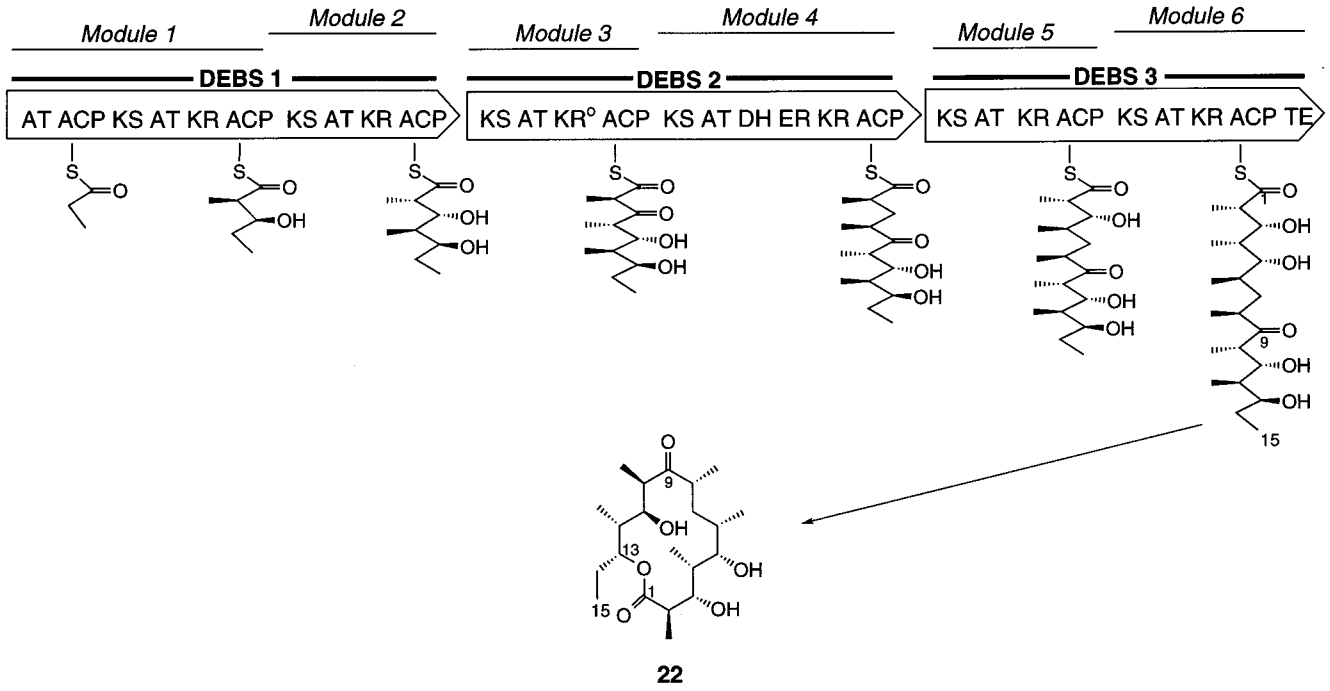
B. The Polyketide Synthase

The complexity of the synthesis of polyketides underscores the programming required for the enzymes involved to produce the correct structure. For example, to make the macrolactone ring of erythromycin, the polyketide synthase must (1) select the correct units for the condensation, (2) determine the correct stereochemistry for the side chains, (3) determine the correct degree of reduction at each cycle, and (4) permit only six condensation cycles before terminating the synthesis. For the more complicated molecule rapamycin, the polyketide synthase must initiate the synthesis with an unusual cyclohexyl starter, permit 14 condensations, each with a specified degree of reduction, and then pass the acyl chain to an enzyme that catalyzes an amide linkage to a derivatized amino acid. Until recently, no single polyketide synthase had been isolated intact, most likely because the sizes of the enzymes had been grossly underestimated by researchers. Our current understanding of the nature of the enzymes and the way they function has come from the isolation and characterization of their corresponding genes.

The first complex polyketide synthase (PKS) to be sequenced was the deoxyerythronolide B synthase (DEBS), involved in the synthesis of the polyketide backbone of erythromycin. One of the original mutants that was blocked in the synthesis of the aglycon but which could convert exogenously fed post polyketide intermediates into erythromycin[18] was used in a complementation test to identify a chromosomal segment of the *Sac. erythraea* chromosome adjacent to the gene *ermE* (that determined erythromycin resistance) as containing some or all of the PKS genes.[19] Subsequent directed insertions into the DNA segment indicated that the PKS gene(s) spanned a region of approximately 35 kb.[20,21] DNA sequencing of the segment not only revealed the overall size and organization of the PKS, but also enabled a proposed understanding of the step by step biosynthesis of the polyketide backbone.[21,22]

The schematic organization of the Ery PKS at the DNA and polypeptide level is shown in Figure 4. The PKS is encoded by three genes, designated *eryAI*, *eryAII*, and *eryAIII*, each approximately 10 kb in length. The *eryAI* and *eryAII* genes are separated by a 1.46 kb segment with characteristics of an insertion (IS) element.[23] The *eryAII* and *eryAIII* genes overlap: the ATGA sequence contains the TGA stop codon for *eryAII* and the ATG start codon of *eryAIII*. The three genes encode the polypeptides, each over 3300 amino acids in length named DEBS (deoxyerythronolide B synthase) 1, DEBS2, and DEBS3, respectively.[24] On the basis of comparing the putative amino acid sequences to known sequences of fatty acid synthases, it was determined that DEBS was composed of enzymatic domains resembling β-ketoacyl ACP synthases (KS, the condensing enzymes), acyltransferases (AT), β-ketoreductases (KR), dehydratases (DH), transenoylreductases (ER), and acyl carrier proteins (ACP), as well as a single domain near the carboxy terminus of DEBS3 resembling thioesterases.[20−22,24−26] Thus, the DEBSs are multifunctional proteins resembling the type I fatty acid synthases found in vertebrates which carry a single copy of each domain. In addition, as can be seen in Figure 4, each DEBS protein carries two sets of the domains KS-AT-ACP, six in all, and a varying combination of the DH, ER, and KR domains within each set. DEBS1 also carries the additional AT and ACP near the NH_2 end of the polypeptide. The KS-AT-(DH-ER-KR)-ACP organization, termed modules, resembles the organization of the domains in vertebrate fatty acid synthases.

Once the presence of modules and the extent of β-carbonyl reduction that each could permit on the basis of the presence of functional KR, DH, and ER



22

Figure 4. Scheme of DEBS showing functional domains in the three-component polypeptides and scheme of acyl chain growth showing structure as ACP thioester at the end of each condensation/reduction cycle. Modules are shown to indicate component functions. Key: ACP, acyl carrier protein; AT, acyltransferase; DH, dehydratase; ER, enoylreductase, KS, β-ketoacyl ACP synthase; KR, β-ketoreductase; TE, KR°, nonfunctional KR domain.

domains were determined, it immediately became apparent that each module controlled one condensation and reduction cycle and that the biochemical order of reactions to produce **22** followed the linear order of modules in DEBS.[21] It was proposed that the AT domain at the beginning of the DEBS1 binds propionyl CoA and transfers it to the pantotheine arm of the adjacent ACP domain and then to the active-site cysteine residue of the KS domain of module 1. The AT domain in the first module binds methylmalonyl CoA and transfers it to the ACP domain in module 1. The diketide is formed through transfer of the starter to the extender and then reduced to the β-hydroxy form by the KR1 domain in module 1. Since no other functions are present, the diketide is transferred from the ACP to the KS domain in module 2 where it undergoes condensation with the methylmalonyl moiety sitting on the cognate ACP domain. Again, only β-ketoreduction takes place due to the absence of functions other than the KR in module 2. After completion of the cycle, the triketide is transferred to the KS domain of module 3. Sequence analysis indicated that the KR domain in the third module would be dysfunctional due to the absence of a functional NADPH binding site, indicated in Figure 4 as KR°. Thus after the third condensation, the β-keto group formed would not be reduced but would persist through the succeeding biosynthesis steps and appear at C-9 in **22**. In the fourth cycle, ketoreduction would be followed by dehydration and enoyl reduction to produce the methylene function which would be present at C-7 of **22** due to the presence of the DH and ER domains in module 4. The reduction cycles in the fifth and six steps would be similar to those described for modules 1 and 2. After the sixth cycle, the 15-carbon chain is released from the ACP domain of DEBS3 through the action of the TE domain in module 6 and forms the lactone **22** by attack of the C-1 carboxylate on the C-13 hydroxyl.

Two of the many aspects relating to the specificity of the Ery PKS implied from the proposed mechanism of acyl chain synthesis require further discussion. The first involves the ACP domain. In fatty acid synthases, which use one set of active sites to conduct seven or more condensations and reductions, the ACP domain interacts twice with the *same* KS domain in each cycle, first to receive the growing chain through the condensation reaction and then to pass the reduced chain back to initiate the next cycle. In type I vertebrate systems, the FAS is believed to function as a head to tail homodimer and the KS of one chain interacts with the ACP of the other.[27] In synthesis of the polyketide **22**, the ACP of each module is proposed to interact with two *different* KS domains: the KS of its cognate module to engage in the condensation reaction and the KS of the next module used in acyl chain elongation for acyl chain thioester exchange. How an ACP domain determines the correct KS domain with which to interact is not yet understood.

The second aspect of specificity deals with the AT domains. The modular hypothesis proposed that the specificity for the extender domains resides in the AT domains. Thus in DEBS, the AT domains of the six modules all select methylmalonyl CoA for binding and subsequent acyl transfer over malonyl CoA or other dicarboxylic CoA esters present in the host. The basis for the specificities for malonate and methylmalonate are discussed below. In addition, the model proposed that the stereochemistry of the side chains in **22** is determined by the methylmalonyl enantiomer chosen for incorporation by the particular AT domain: (2*S*)-mmCoA for the ATs in modules 2, 5, and 6 and (2*R*)-mmCoA for the ATs in modules 1, 3, and 4. It has since been shown, in elegant work done

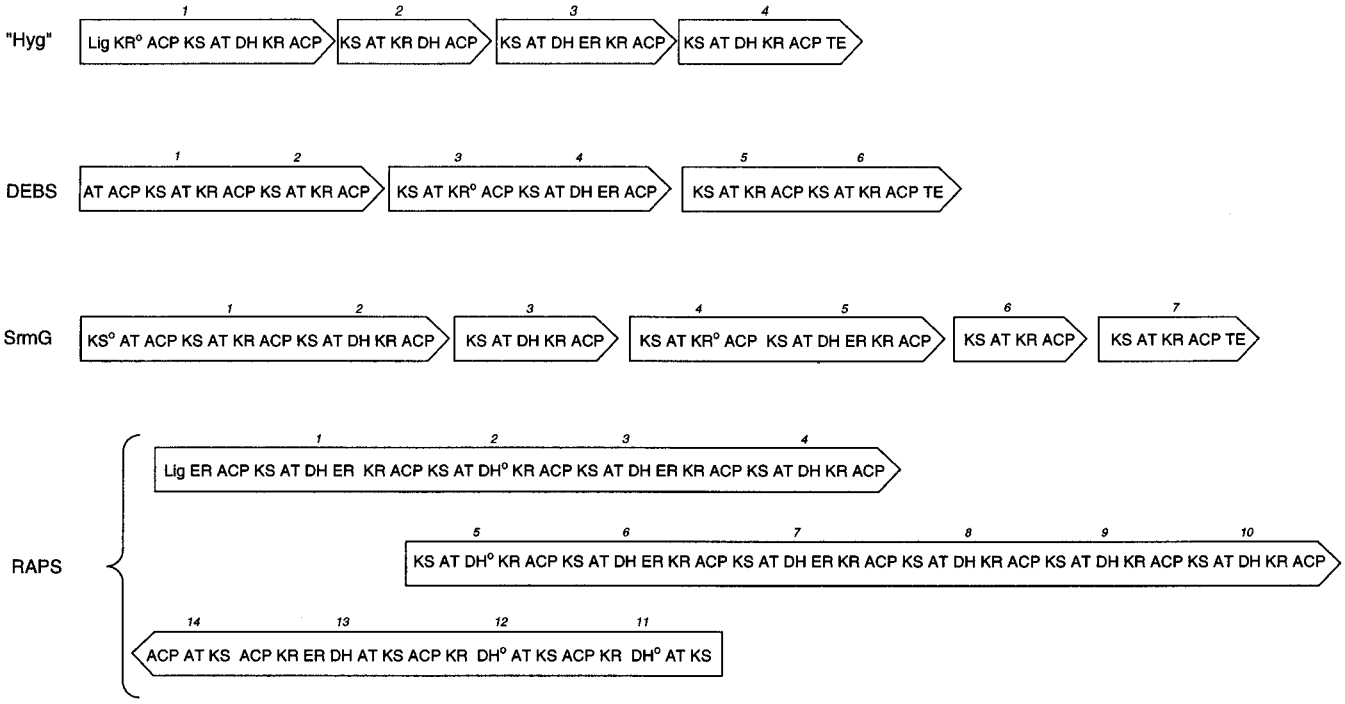


Figure 5. Modular organization of various polyketide synthases. Abbreviations are as in Figure 4, except for the addition of Lig for CoA ligase. Modules are numbered to indicate their position in synthesis of the cognate polyketide. Nonfunctional domains are indicated by the superscript "o". Relative directions of modules indicates relative direction at the gene level. Relative positions of the polypeptides are indicated only for "Hyg", DEBS and SrmG. Order and relative orientation of modules for the RAPS are shown. References: DEBS, ref 21; "HyG", ref 36; RAPS, ref 29; SrmG, ref 31.

in P. F. Leadlay and J. Staunton's laboratories, that all six of the AT domains in purified DEBS proteins bind and catalyze the acyl transfer (most likely to their cognate ACPs) of only the (2*S*)-mmCoA enantiomer.[28] Thus, in three of the six condensations in the synthesis of **22**, epimerization must take place, either of the 2-methyl group in mmCoA before the condensation, or of the α side chain in the nascent polyketide chain after the condensation. Extensive examinations of the DEBS polypeptides have not revealed sequences which could serve as epimerase domains.

In the six years since the determination of the nucleotide sequence of the *eryA* genes, several other complex PKS genes have also been sequenced, in whole or or part, including rapamycin,[29,30] spiramycin,[31] oleandomycin,[32] and avermectin.[33] In each case, the PKS is organized in modules similar to that in DEBS, with one module servicing a single condensation and reduction cycle. Three such organizations, along with that for the Ery PKS are shown in Figure 5. The largest set sequenced are the *rapA*−*C* genes involved in the synthesis of rapamycin, which span more than 70 kb and encode 14 modules. The Rap PKS is designated rapamycin synthase (RAPS). The first gene, *rapA*, encodes an unusual set of loading functions consisting of domains resembling a CoA ligase, an enoylreductase and an ACP which function to activate, reduce, and move the dihydroxycyclohexyl starter to the KS domain of the first module. There is no TE domain at the end of the last gene, *rapC*, presumably since the acyl chain must be moved to another enzyme (probably though thioester exchange) which catalyzes amide bond formation between the polyketide and pipecolic acid. The *rap* PKS genes encode only three polypeptides, the largest of which contains six modules.

Other examples of PKS modules have been found. PKS genes for the synthesis of a heptaene macolide identical in backbone structure to candicidin have been cloned; hybridization with various domains of the *eryA* genes as probes indicated the PKS genes are organized into 21 contiguous modules, spanning a region of greater than 100 kb.[34] Several PKS-like genes have also been uncovered in a number of bacterial chromosomes through genome sequencing initiatives. Examples include *Bacillus subtilis*, *Mycobacterium leprae*, and *Mycobacterium tuberculosis*: the annotated sequences are accessible in the GenBank and EMBL public databases.[35] To date the identities of the compounds encoded by these genes are not known. In addition, PKS-like sequences have been uncovered through the inadvertant cloning of genes that were not originally targeted. For example, we have discovered one PKS-like cluster in the strain of *Streptomyces venezuelae* that produces methymycin and two new PKS-like clusters in the strain of *Streptomyces hygroscopicus* that produces rapamycin. One of the clusters, named "*hyg*" was sequenced completely.[36] The organization of the PKS is shown in Figure 5. The cluster is composed of four modules, each contained within an individual polypeptide. In addition, the first module contains a loading domain very similar to that found in RAPS 1.

Since rapamycin is composed of both malonate and methylmalonate building blocks, after the *rap* PKS genes were sequenced and the correspondence of modules to condensation cycles determined, the Leadlay−Staunton group compared the sequences of the 14 AT domains in RAPS with the seven AT domains in DEBS.[29,37] Using sequence alignments generated by a number of programs that perform pairwise comparisons, they showed that the various

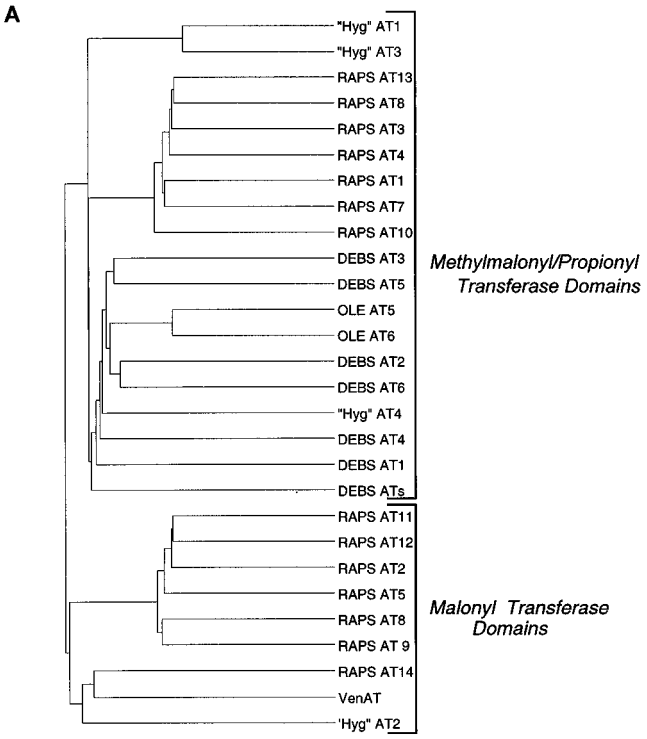


Figure 6. Alignment of acyltransferase domains from various polyketide synthases. A. Dendrogram from PILEUP analysis (GCG, Madison WI) showing clustering of malonyl or methylmalonyl/propionyl loading domain sequences. B. Consensus sequences of malonyl and methylmalonyl loading domains. All letters shown represent invariant amino acids for a given given (malonyl or methylmalonyl) family. Bolded letters indicate differences between the families. Gaps indicate no consensus at the given position. The asterisk (*) signifies the serine residue which is linked to the acyl CoA in the acyl-AT complex.

AT domains cluster into two groups, one that contains only the proposed malonate loading functions and the other that contains all the methylmalonyl and propionyl loading functions. In addition, they showed that within the methylmalonyl cluster, the RAPS AT domains subclustered from the DEBS AT domains, suggesting that the RAPS modules all had a common origin or that one was formed from another through gene duplication and subsequent divergence. With the sequencing of additional PKS genes, including a set from the same host that carries the *rap* PKS genes, it was possible to ask how two PKSs from the same host are related along sequence lines. A dendrogram from a PILEUP analysis from the GCG software package[38] of the various AT domains is shown in Figure 6A. Again, the known or suspected malonyl and methylmalonyl/propionyl loading domains cluster into cognate groups. Furthermore, the RAPS methylmalonyl AT domains sub-cluster from the others in the group but the DEBS AT domains do not themselves form a separable group. Three of the four AT domains, from modules 1, 3, and 4 of the "Hyg" PKS cluster with the methylmalonyl loading domains; the VenAT and "Hyg" AT2 domains fall into the malonyl loading group. The latter two domains were used in AT replacement experiments descibed in section VII.

Using the sequences from the 28 AT domains shown in Figure 6A, as well as from a number of additional AT domains from other PKS-like sequences obtained in our laboratories, signature sequences for malonyl and methylmalonyl loading domains could be identified and are shown in Figure 6B. The active-site serine residue (identified by the asterisk (*)) is preceded 30 residues upstream by the amino acid threonine and followed immediately by the residue valine in the more than 12 streptomycete PKS malonyl AT domains examined. Similarly, in the more than two dozen methylmalonyl loading domains examined, the amino acid 30 residues upstream of the serine is valine and the residue immediately following the serine is glutamine. Additional differences between the two types of loading domains in the 40 residue sequence shown in Figure 6B are also present. As more sequences become available, it will be of interest to determine whether these differences are maintained and whether signature sequences for ethylmalonyl and other extender loading domains can also be uncovered.

III. Post Polyketide Modification in Erythromycin Biosynthesis

Since the biochemical fates of the altered polyketides produced by manipulation of the DEBS-encoding genes in *Sac. erythraea* (see section VII) may be impacted by the presence of enzymes that normally process the erythromycin lactone, features of the pathway of erythromycin biosynthesis will be reviewed briefly here. The genes, enzymes and the pathway of erythromycin biosynthesis are shown in Figure 7. After synthesis of **22** by DEBS, the compound is hydroxylated at C-6 by the P450 hydroxylase EryF, encoded by the *eryF* gene, to produce the compound erythronolide B (EB; **23**).[39,40] Although this step always takes place in wild-type strains of *Sac. erythraea*, it is bypassable: mutants disrupted in *eryF* produce the bioactive compound 6-deoxyerythromycin.[39] The next steps involve addition of the two deoxysugars, first L-mycarose to yield 3-α-mycarosylerythronolide B (**24**), and then desosamine to produce erythromycin D (**4**), the first bioactive macrolide of the pathway. The genes involved in the biosynthesis of L-mycarose are designated *eryBI*−*eryBVII* and the genes involved in the synthesis of desosamine are designated *eryCI*−*eryCVI*. The pathways of synthesis of L-mycarose and D-desosamine from TDP 4,6-deoxyglucose have been proposed from the identities of the putative gene products.[41] The glycosyltransferases that attach mycarose and desosamine to the lactone ring are *eryBIV* and *eryCII*, respectively.

Erythromycin D (**4**) is converted to erythromycin A (**1**) through the action of two enzymes, a P450 hydroxylase, the product of the *eryK* gene, acting at C-12[42] and an *O*-methyltransferase, the product of the *eryG* gene, that methylates the hydroxyl at C″-3 on the mycarose moiety.[43,44] (Methylated L-mycarose is designated L-cladinose, but the compound has not been isolated as a free sugar except by removal from **1**.) Since both erythromycins B (**2**) and C (**3**) could be found, in varying proportion, in fermentation broths of *Sac. erythraea*, it was originally believed that the pathway of **4** to **1** could proceed either by

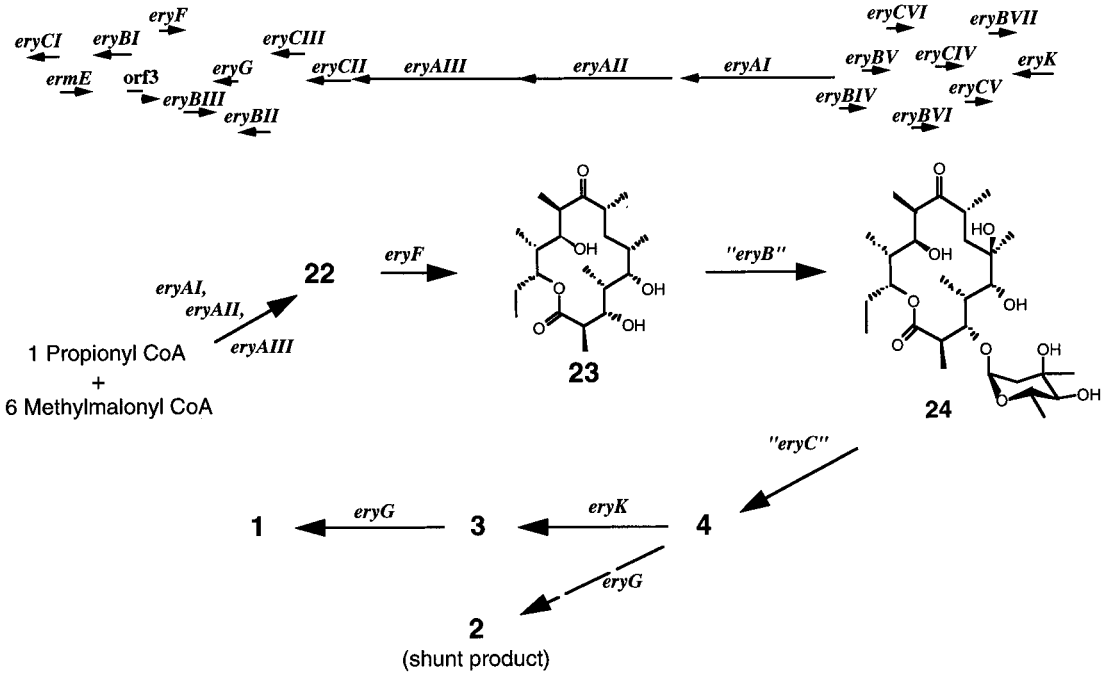


Figure 7. The erythromycin biosynthesis gene cluster in *Saccharopolyspora erythraea* and their role in erythromycin biosynthesis. The genes involved in the various steps of the pathway are shown. "*eryB*" and "eryC" indicate all the *eryB* and *C* genes for the synthesis of the two sugars from TDP-4,6-deoxyglucose through attachment to form the intermediates of the pathway shown.

Table 1. Summary of Known or Proposed Functions of Erythromycin Biosynthesis Genes

locus	role in Er biosynthesis	function	ref(s)
eryAI	macrolactone synthesis	polyketide synthase	21
eryAII	macrolactone synthesis	polyketide synthase	21
eryAIII	macrolactone synthesis	polyketide synthase	21
eryBI	L-mycarose synthesis	not known	41
eryBII	L-mycarose synthesis	3-ketoreductase	41
eryBIII	L-mycarose synthesis	not known	41
eryBIV	L-mycarose synthesis	4-ketoreductase	41
eryBV	L-mycarose attachment	mycarosyltransferase	41
eryBVI	L-mycarose synthesis	not known	41
eryBVII	L-mycarose synthesis	5-epimerase	41
eryCI	D-desosamine synthesis	3-aminotransferase	41, 49
eryCII	D-desosamine synthesis	not known	41
eryCIII	D-desosamine attachment	desosaminyltransferase	41
eryCIV	D-desosamine synthesis	3,4-dehydratase	41
eryCV	D-desosamine synthesis	not known	41
eryCVI	D-desosamine synthesis	3-aminodimethyltransferase	41
eryF	C-6 hydroxylation	P450 monooxygenase	39, 50
eryG	C″-3 *O*-methylation	*O*-methyltransferase	43, 44
eryI	not known	thioesterase	39, 50
eryK	C-12 hydroxylation	P450 monooxygenase	42, 45
ermE	resistance	*N*-methyltransferase	46

first C-12 hydroxylation followed by C″-3 *O*-methylation, or by *O*-methylation followed by hydroxylation. With the cloning and expression of *eryK* and subsequent purification of the protein, it was found that enzyme preferred **4** as a substrate over **2** by a factor of 1900-fold.[45] Similarly, EryG shows a 4-fold preference for 3 over 4 as a substrate.[44] Thus, the pathway is believed to proceed **4** → **3** → **1** and, if **2** is seen, it is as a shunt product.

All the genes involved in the biosynthesis of erythromycin have been found to be clustered in a segment of the chromosome surrounding the gene *ermE*, that encodes a 23S ribosomal RNA *N*-methyltransferase that confers self-resistance to erythromycin.[46] The *ermE* gene was cloned from *Sac. erythraea* by selection for erythromycin resistance in the heterologous host *Streptomyces lividans*.[47] Cosmids carrying *ermE* and up to 40 kb of flanking DNA sequences were used in a series of complementation experiments with mutants blocked in erythromycin production or in disruption experiments to identify erythromycin biosynthesis genes.[19,21,22,39,41−44,48] The >65 kb segment of DNA shown in Figure 7, from *eryCI* through 7 kb beyond the last known erythromycin biosynthesis gene *eryK* has been sequenced.[20,21,41,42,49−51] The entire biochemical pathway from the synthesis of the first diketide through the synthesis of **1**, including all of the reactions to produce the two deoxysugars from TDP-4,6-deoxyglucose, is encoded by the genes in this segment. A summary of the genes and their corresponding known or proposed role in erythromycin synthesis is shown in Table 1. Interestingly, no regulatory genes for the pathway have been uncovered. Since the synthesis of erythromycin occurs at the maximum rate during the late logarithmic and stationary growth phases in

Sac. erythraea, it is not known, therefore, how the biosynthesis of the antibiotic is regulated.

IV. Overview of Genetic Changes in PKS Genes

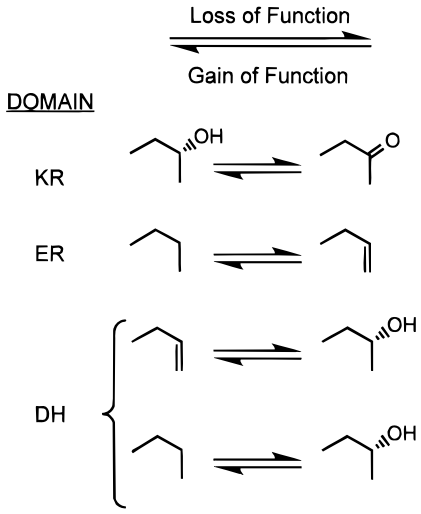
Current technology is available to introduce a wide variety of specified genetic changes into PKS genes, from a change of a single nucleotide pair to the deletion of large segments of DNA. Changes, therefore, can be directed at a single functional domain or at an entire module (~5 kb) and can be categorized as loss, gain or change of function(s). Changes affecting the condensation reactions are also distinguished from changes affecting the processing of the β-carbonyl groups generated after the condensation.

A. Changes in β-Carbonyl Processing

1. Loss of Function

The three domains involved in β-carbonyl processing function in hierarchical order (KR-DH-ER). Thus, if all three are present in a given domain, loss of KR function obviates the requirement for the DH and ER functions and loss of DH domain will render the otherwise functional ER domain unnecessary. The structure of resultant polyketide at the site determined by the corresponding β-carbonyl-processing domain is shown in Scheme 1. (Attachment of the

Scheme 1



polyketide backbone to the ACP as a thioester is implied: see Figure 2 for details.) Loss of the KR results in the appearance of a keto group in place of the original hydroxyl. Loss of the ER leaves the ene in the polyketide. Loss of the DH domain leaves the hydroxyl side chain in place of the methylene function, if a functional ER domain is otherwise present within the affected module; if an ER domain is not present, the ene formerly present at the site in the polyketide is replaced by an hydroxyl group. Examples of loss of KR and ER functions and their effects on polyketide structure are described in section VII.

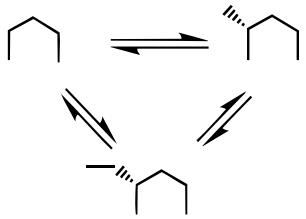
2. Gain of Function

Structural changes resulting from gain of β-carbonyl-processing functions in PKS modules are also shown in Scheme 1. These changes are the reverse of those shown for corresponding loss of function. Replacement of an hydroxyl side chain with a methylene group in the polyketide chain through gain of DH function cannot occur without the presence of the ER function within the affected module, however. One example of gain of function will be described.

B. Changes in Polyketide Side Chains

Since the structure of the side chain at a given site in the polyketide chain is determined by the structure of the extender unit incorporated at the corresponding site in the nascent chain, changes in the side chains which appear in the final product will result from replacement of one extender unit for another. Free interchange of the three common substrates used as extender units, malonyl CoA, methylmalonyl CoA, and ethylmalonyl CoA will allow side-chain replacements shown in Scheme 2. Since the AT domains determine the nature of the cognate extender unit incorporated, AT swapping results in corresponding side-chain changes.

Scheme 2



C. Changes in the Starter Side Chain

The starter unit, propionyl CoA, acetyl CoA, dihydroxycyclohexenyl CoA (or dihydroxycyclohexenecarboxylic acid), etc., is also determined by either an AT or ligase function present as the first functional domain at the 5′ end of the PKS gene(s). Domain swaps at these sites can result in changes in the starting side chain. An example is described in section VII.

D. Changes in Acyl Chain Length

Since the length of the acyl chain is determined by the number of condensations, each of which is determined by a single module within a PKS, chain length can be changed by removing or adding modules. To date, only removal of modules, resulting in the shortening of the PKS chain, has been reported. A brief summary will be provided in section VII.

V. Methodology

Genetic manipulations of the DEBS and the modular polyketide synthase encoded by *srmG*, involved in the synthesis of platenolide in *Streptomyces ambofaciens*, have been reported and will be described in detail in section VII. For both PKSs, changes have been made in the polyketide-producing host, *Sac. erythraea or S. ambofaciens*, where the PKS-encoding genes are manipulated in their naturally occurring location, the host chromosome. While using these hosts to make gene replacements are fairly time consuming and can take up to several months to complete, they offer the potential to produce all potential altered polyketide compounds that emanate

from the complete polyketide biosynthesis pathway in a single host. This is best evidenced in the case of the *Sac. erythraea* strain in which the AT domain of module 1 in *eryAI*, which normally specifies a methylmalonyl extender, was replaced with an AT domain that specifies a malonyl extender in the first condensation cycle catalyzed by DEBS, as shall be described in some detail in section VI. The product produced by the host that carried the altered PKS was determined to be 12-desmethylerythromycin B, indicating that not only had the PKS recognized and condensed an unusual extender unit, but that all of the remaining steps of polyketide assembly and processing could take place on the altered nascent acyl chain. In addition, all the biochemical steps that normally followed the completion of polyketide synthesis, C-6 hydroxylation, attachment of the deoxysugars at C-3 and C-5, and *O*-methylation at the C″-3 position of the neutral sugar L-mycarose after attachment to the macrolactone ring, but not the normal C-12 hydroxylation step, could take place. Thus, using the natural host to make the change in the PKS not only enabled the synthesis of a novel natural product, but also provided information on the specificities of a number of enzymes in the pathways that operate after the synthesis of the polyketide.

The other system to make changes in a modular polyketide synthase, that will be described in greater detail below, employs a plasmid carrying the PKS genes that can replicate in streptomycetes and *E. coli*. The *E. coli* host is used perform the gene replacement experiments and the host *Streptomyces coelicolor* is used produce the polyketide product. Since the actual gene replacement in the streptomycete is the most time-consuming aspect of the manipulation, employing an *E. coli* host to perform the genetic exchange both facilitates and speeds up the process considerably. However, since only the PKS genes are introduced into the heterologous host, only the polyketide lactone or intermediates in its synthesis can be produced. Conversion to more advanced erythromycin intermediates would require the construction of host strains carrying additional erythromycin biosynthesis genes. In addition, drawbacks relating to precursor supply, which will be further addressed in section VII, also limit this system.

A. The *Sac. erythraea* System

The system employed at Abbott Laboratories to effect genetic changes in the *eryA* genes is described here. We have used *Sac. erythraea* ER720[52] as the host since we have found it amenable to the introduction of exogenous DNA either via transformation of protoplasts or by electroporation. This strain is a fairly low level producer of erythromycin and has given rise to some difficulty in the recovery of novel erythromycins produced through genetically altering the PKS genes. Although *Sac. erythraea* strains that produce higher levels of erythromycin would most likely be better choices as starting hosts, it has been our experience that such strains are more refractory to genetic manipulation, at least under the conditions that we have attempted.

Gene replacement is a two-step process that requires two independent genetic exchanges between the chromosome and a DNA vector carrying the desired DNA segment that will eventually be placed in the PKS-encoding gene. In *Sac. erythraea*, plasmids comprise the DNA vectors employed for gene replacements. In the first step, a plasmid defective in replication in *Sac. erythraea* carrying the desired sequence to eventually be placed in an *ery A* gene is introduced into the host. Although *E. coli* plasmids, pBR322, pUC18, etc. have been used for gene replacements in other systems, we have found pIJ702 and its derivatives to be the most useful plasmids to perform gene replacement experiments in the erythromycin biosynthesis genes *in Sac. erythraea*. pIJ702 is derived from pIJ350, a plasmid itself derived from the multicopy plasmid pIJ101 isolated from *S. lividans*, into which was inserted a DNA segment conferring thiostrepton resistance from *Streptomyces azureus* and the gene *melC* from *Streptomyces antibioticus*.[53] A particularly useful derivative of pIJ702 is the vector pWHM3, also called pCS5, developed by C. R. Hutchinson and colleagues that is basically a fusion of pUC18 and pIJ702: pCS5 can be propaged in *E. coli*.[54] A similar plasmid used by P. F. Leadlay and associates for gene replacement in *Sac. erythraea* is called pTED2.[55] For reasons that are unclear, pIJ702 does not stably replicate in *Sac. erythraea* and is thought to be lost at the time when cells carrying the plasmid under selection for thiostrepton resistance begin to sporulate on solid medium.[56] If the plasmid carries a segment of the *Sac. erythraea* chromosome, homologous recombination between plasmid and chromosome can take place, resulting in the integration of the plasmid sequence and maintenance of the plasmid in the chromosome through succeeding generations in the presence of the drug. Throughout the growth cycle a low frequency of excision of the plasmid through a reversal of the integration process can also occur. Removal of the drug results in the loss of the plasmid from cells in which excision has taken place. When two segments of the *Sac. erythraea* chromosome are placed in pCS5 to flank a desired segment from a heterologous host, an independent recombination event at each of the *Sac. erythraea* sequences can take place resulting in the replacement of a chromosomal segment with a desired sequence introduced on the plasmid. An example of a gene replacement experiment in *Sac. erythraea* is described here and illustrated in Figure 8.

Two noncontiguous segments of approximately 1.1 kb, from the regions of *eryAI* immediately upstream and downstream, respectively, of the DNA segment encoding the AT domain of module 1 were inserted into pCS5, to create the vector termed pCS5/*ery*AT1-flanks, as shown in Figure 8A. (The coordinates of the fragments with respect to the *eryA* sequence are shown: GenBank accession no. M63676.) The segments were introduced as *Eco*RI-*Bam*HI and *Bam*HI-*Hin*dIII fragments and were engineered to contain the cloning sites *Avr*II and *Nsi*I, respectively in the two segments, as shown. A segment from an uncharacterized PKS-encoding region from a strain of *Streptomyces venezuelae* believed to encode an AT domain was cloned into pCS5/*ery*AT1-flanks using PCR to introduce *Avr*II and *Nsi*I sites upstream and downstream of the segment, respectively and to allow the coding sequence of the entire ~3 kb segment to

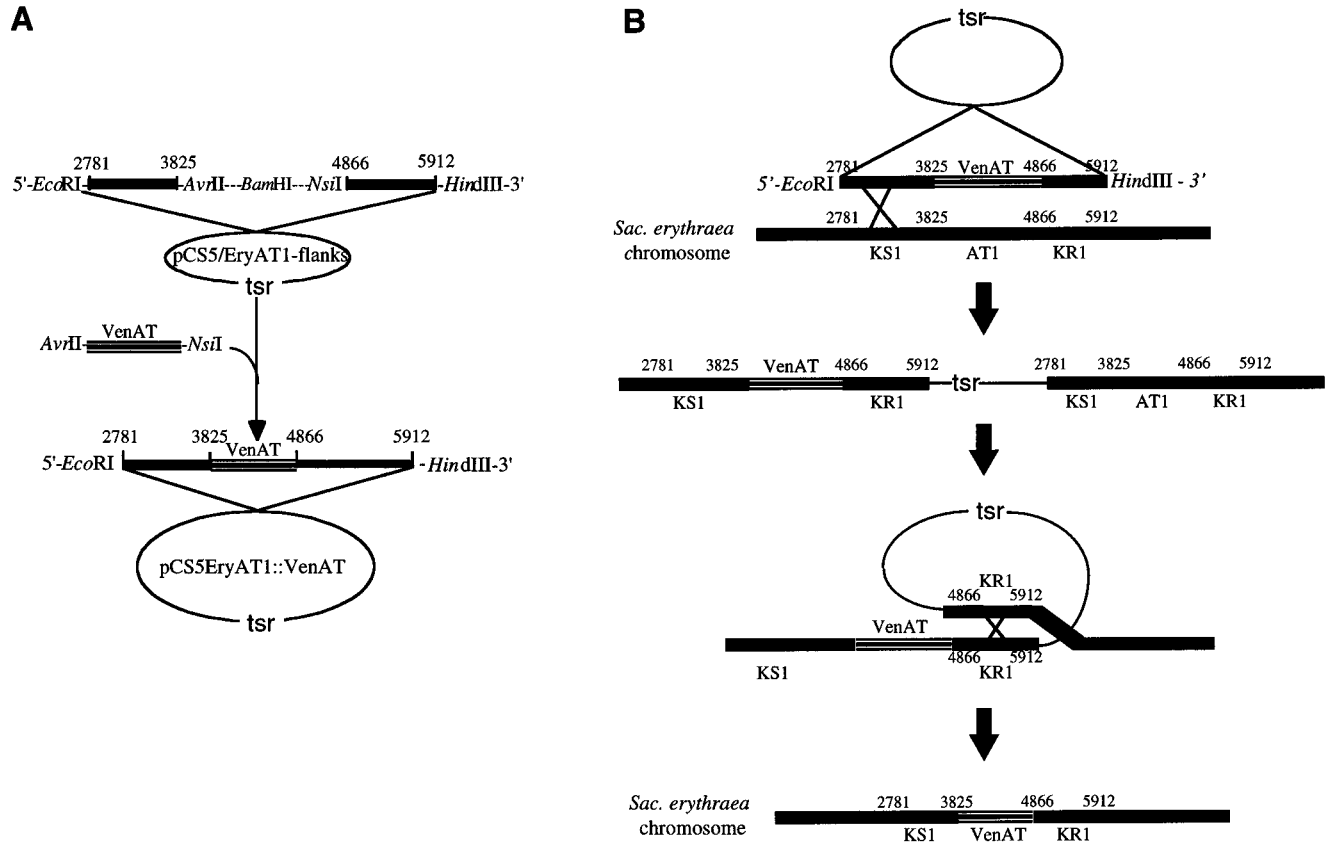


Figure 8. Scheme of a gene replacement experiment. A. Illustration of strategy used to clone the VenAT heterologous AT domain into vector pCS5 carrying two segments of the *eryAI* gene from *Sac. erythraea*. The two segments in bold represent the upstream and downstream sequences adjacent to the AT domain in the ery PKS that will be replaced; each is cloned independently in pCS5 by PCR as an *Eco*RI-*Bam*HI or *Bam*HI-*Hind*III fragment to produce pCS5/EryAT1 flanks containing the *Avr*II-*Bam*HI-*Nsi*I "stuffer" segment. The VenAT DNA segment is engineered to contain *Eco*RI and *Nsi*I ends and cloned into pCS5/EryAT1 flanks using these two enzymes to release the "stuffer" segment and result in a plasmid containg the VenAT encoding sequence adjoining the ery PKS upstream and downstream sequences. B. Scheme showing two-step homologous recombination to place VenAT domain in chromosome between KS1 and KR1 domains in *eryAI*. Genetic exchange is indicated by a large X. Description of the selection and screening for the correct genetic construct is described in the text.

lie within a single open reading frame. The plasmid, called pEryAT1::venAT, was introduced by transformation into protoplasts of *Sac. erythraea* ER720 employing selection for thiostrepton resistance. The colonies that arose carried the plasmid in the integrated state: the plasmid had undergone a single recombination event with the chromosome in either the *Eco*RI-*Avr*II (upstream) or the *Nsi*I-*Hind*III (downstream) region of homology. The location of the recombination event was determined by examination of the chromosomal structure of the recombinant by Southern hybridization or by PCR analysis. (The illustration shown in Figure 8B depicts recombination in the upstream sequence.) Integration of the plasmid sequence into *eryAI* destroys the integrity of the gene and results in the loss of erythromycin production, which was monitored by TLC analysis or by the bioassay of products produced from the host. Once the strain was determined to contain the correct integration, it was plated in the absence of thiostrepton for a number of passages and then up to several hundred colonies were tested for the inability to grow in the presence of the drug. Thiostrepton-sensitive colonies were generated from the excision of the integrated vector from the chromosome and loss of the excised plasmid from the cell through failure to replicate. Excision took place through homologous recombination between sequences in the segments either upstream or downstream of the *ven*AT sequence. If the exchange occurred in the upstream segment, the plasmid would have excised almost precisely as it had integrated, leaving behind the unaltered chromosome. Such strains would carry the *ery*AT1 sequence and would be expected to produce erythromycin. If the exchange took place in the downstream sequence, the circular DNA element removed from the chromosome would have taken with it the segment of the chromosome encoding *ery*AT1 and have left behind the segment of the plasmid encoding *ven*AT (Figure 8B). Since genetic exchange takes place between identical copies of the sequence, the segment left behind on the chromosome would, therefore, resemble the segment constructed in pEryAT1::venAT, that is the *eryAI* gene with the AT in module I replaced with that from an *S. venezuelae* PKS gene in frame with the upstream and downstream segments encoding DEBS1. Again, Southern hybridization or PCR analysis is used to demonstrate that the desired gene replacement has occurred. Once the correct strain has been made it could be examined for the polyketide products it produced.

It is not understood why pBR322 or other *E. coli* plasmids that do not replicate in *Sac. erythraea* cannot be used to perform replacement experiments in this host. It is possible that a small amount of

replication is required to establish the plasmid in the host before it undergoes recombination. Perhaps pIJ702 undergoes abortive replication to generate single-stranded intermediates which are the DNA elements that undergo genetic exchange. Single-stranded DNA has been used to perform gene replacements in streptomycetes.[57]

The single biggest drawback to the gene replacement system described above is the necessity to follow without selection the randomness of the second recombination step that results in excision of the integrated plasmid from the chromosome. The frequency of this event cannot be predicted nor is it possible to predict, for any given colony that arises that has lost the plasmid, at which site the recombination event has occurred or whether it is favored at a given sequence. Consequently, following the excision requires the blind testing of many colonies, isolation of their DNA, Southern hybridizations, etc. This is time consuming and labor intensive. Improvements to the system are sorely needed. One such improvement would be to add a marker to the plasmid that can be used to select for cells wherein the plasmid has been *lost*. A recent example is the *rpsL* gene which was cloned from *Streptomyces roseosporus* in a plasmid to confer the dominant phenotype of streptomycin sensitivity in a streptomycin resistant background in that host.[58] If *rpsL* functions in *Sac. erythraea*, selection for streptomycin-resistance would yield colonies in which pCS5 (carrying *rpsL*) has been excised from the chromosome and lost from the cell.

B. The *S. coelicolor*/pCK7 System

The system developed by C. Khosla and associates is described in detail in a review by C. Khosla this issue and will be described only briefly here. Plasmid pCK7 contains the origin of replication from SCP2*, a stable, low copy plasmid from *S. coelicolor*, the *tsr* gene for thiostrepton resitance, a ColEI origin of replication and an *amp* gene for replication and selection of ampicillin resistance in *E. coli*, the gene *actII* ORF4, a positive regulator of the actinorhodin PKS genes in *S. coelicolor*,[59] and the three *eryA* genes cloned from the chromosome *of Sac. erythraea* downstream from P*actI*, a promoter under the control of the *actII* ORF4 product.[60] The host employed for production of the polyketide is *Streptomyces coelicolor* CH999, which carries a deletion of the entire *act* cluster (>20 kb). *S. coelicolor* CH999/pCK7 produces the macrolactone of erythromycin, 6-dEB (**22**), under the control of *actII ORF4*, which normally turns on actinorhodin production at the time when cells are in late log and stationary phase.

Derivatives of pCK7, which contain segments of the *eryA* genes that have been genetically modified have also been made. Changes to the *ery* PKS genes are made by cloning the desired segment, along with regions of *eryA* required for the two-step crossovers, into a temperature-sensitive *E. coli* plasmid that carries a chloramphenicol resistance marker and is compatible with pCK7. When the two plasmids are present in the same cell, survival of the host in the presence of chloramphenicol and carbenicillin at high temperature results from homologous recombination between the two plasmids to form a cointegrate whose replication is driven from the ColEI origin in pCK7. Second step recombination can be monitored through the loss of chloramphenicol resistance. Refinements to the system have been made to screen directly for the strain that contains the desired replacement. Once the final construct is made and checked, it can be moved into the *S. coelicolor* host by standard transformation procedures.

VI. Examples of Genetic Manipulation of PKS Genes

A. Loss of β-Carbonyl-Processing Functions

1. Ketoreduction

Introduction of the first change into a complex PKS was reported for the KR domain of the fifth module of DEBS.[21] Conveniently, two *Nco*I restriction sites separated by a distance of 813 bp were found in the *eryAIII* gene in the region encoding the KR5 domain. Deletion of the 813 bp segment would remove 271 amino acids from DEBS3, which comprised the first two-thirds of the KR5 domain as well as a segment from the AT5 to KR5 interdomain region, but which would leave the reading frame intact. A 9.5 kb segment of the *eryAIII* gene, which contained the KR5 domain as well as at least 4 kb of upstream and downstream flanking sequences, was cloned in an *E. coli* vector and the 813 bp segment was removed by digestion with *Nco*I followed by religation. The segment was then cloned in pCS5 for use in the subsequent gene replacement procedure in *Sac. erythraea* as described in section IV. As shown in Figure 9, the resulting strain produced the macrolactone 5,6-dideoxy-5-oxoerythronolide B (**25**) as well as the further biosynthesis products 3-α-mycarosyl-5,6-dideoxy-5-oxoerythronilde B (**26**) and 3-α-mycarosyl-5-deoxy-5-oxoerythronilde B (**27**). The 5-oxo derivatives were expected to be made if synthesis of the polyketide could bypass the step of ketoreduction after the fifth condensation and continue through the end of the sixth cycle. That this occurred indicated that the 271 amino acid deletion did not adversely affect the functions of the remaining domains in DEBS3 or the necessary interactions between DEBS3 and DEBS2 for interprotein nascent chain passage required for the start of the fifth condensation cycle. Compound **27** was produced by EryF-mediated hydroxylation at C-6 of **26**, followed by addition of L-mycarose at C-3. The next step, addition of D-desosamine, was blocked by the presence of the keto group in place of the normal hydroxyl function at C-5. The appearance of **26**, albeit in relatively minor amounts, indicated that the C-6 hydroxylation step had been bypassed. As described in section II, loss of C-6 hydroxylation of the aglycon does not shut down the remaining steps of the erythromycin biosynthesis pathway, but until the 5-oxo derivatives were uncovered, 6-deoxy derivatives carrying the deoxysugars had not been observed previously in fermentations with strains carrying a functional *eryF* gene. It was subsequently shown, using purified EryF protein, that **25** could be hydroxylated at C-6 but that it was a much poorer substrate than the normal substrate **22**.[61]

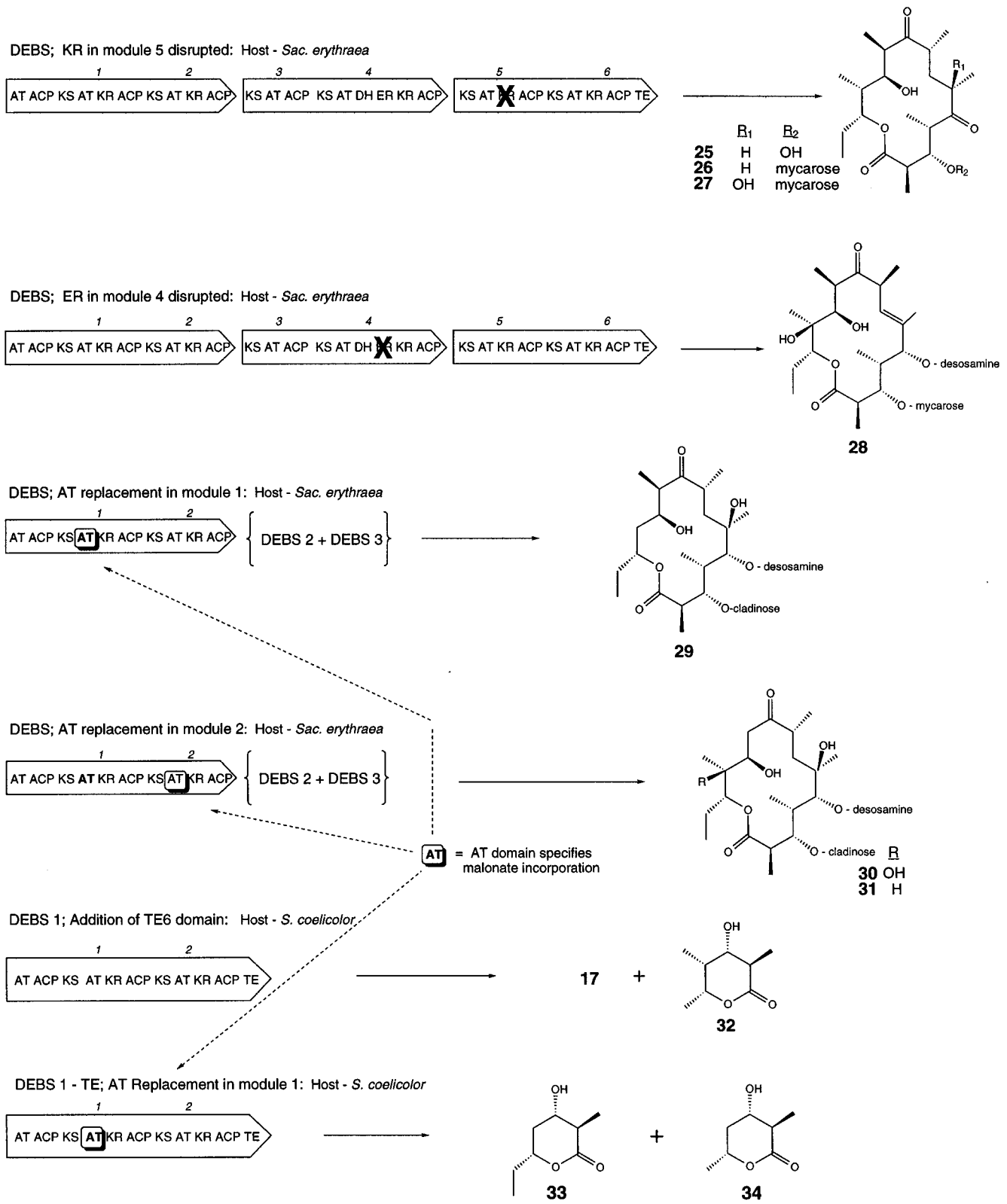


Figure 9. Summary of genetic manipulations of PKS genes and corresponding products produced by host strains carrying altered PKS genes. See Figures 4 and 5 for abbreviations.

The 5-oxo- derivatives were the first rationally designed complex polyketides made by reprogramming a modular PKS and raised the prospects that PKSs could tolerate significant changes not only in their own architecture but also in the structure of the nascent chains they could ultimately synthesize. In addition, the KR5 deletion experiment also provided the first direct piece of evidence that DEBS was indeed a modular PKS and that, at least for the synthesis of 6-dEB, colinearity does exist between the genetic order of the modules and the chemical order of reactions in the synthesis of 6DEB.

2. Enoylreduction

The ER domain in the fourth module of DEBS was changed by the introduction of changes in four base pairs in the DNA to result in the substitution of two adjacent glycine residues in the putative NADPH

binding site to the sequence serine-proline.[62] The ser-pro sequence was found in the NADPH binding site in the segment of DEBS2 (module 3) corresponding to the location of a KR domain which was required to be dysfunctional (see section II), hence the strategy employed to introduce a minimum alteration of the ER domain and retain the same overall number of residues and yet ensure loss of function. The changes introduced an *Nhe*I site in the DNA which could be used to follow the progress of the subsequent gene replacement employed to exchange the wild type ER domain for the mutant one in the *eryAII* gene. The resulting strain produced the compound $\Delta^{6,7}$-anhydroerythromycin C (**28**) in very small amounts (Figure 9). It is of interest to point out that many changes were introduced into the ER domain, from deletions of large segments down to one or two amino acid replacements in different locations of the domain. The one reported is the only change that allowed the mutant strain to produce any identifiable polyketide.

The 6,7 ene function in the resulting compound was predicted on the basis of the predicted biochemistry for 6-dEB synthesis. It was not known, however, whether the compound $\Delta^{6,7}$-anhydroerythronolide B produced by the mutant PKS would be further processed. Apparently, all succeeding reactions that normally take place starting with 6-dEB, proceeded with $\Delta^{6,7}$-anhydroerythronolide B with the exception of the C-6 hydroxylation and *O*-methylation at C″-3.

Attempts to produce an altered polyketide through inactivation of the dehydratase function of the DH4 domain of DEBS have not been successful either in P. F. Leadlay's laboratory[63] or at Abbott Laboratories. Gene replacements to change the single histidine residue in the active site of the DH4 domain resulted in the failure of the host to produce a detectable compound. It is not currently known whether the single amino acid substitution resulted in an intolerable change in the structure of DEBS 2 or whether the nacent chain carrying an hydroxyl group at what would ultimately become C-7 of the lactone ring could not be further processed by the DEBS.

Gain of a DH function in DEBS has recently been reported using the *S. coelicolor*/pCK7. Since this will be discussed in detail in the review by C. Khosla, it will be described briefly at the end of section VI.

B. Changes in the Extender Unit: AT Replacements

Three streptomycete PKS AT domains, designated RAPS AT14, VenAT2, and "Hyg"AT2, have been used to replace the AT domains in modules 1 and 2 of DEBS in *Sac. erythraea*.[64] The RAPS AT14 domain was predicted to specify a malonyl extender unit based on the absence of a methyl (or ethyl) side chain at C-9 of rapamycin. The carbons at C-9 and C-8 are introduced into the nacent rapamycin acyl chain at the 14th condensation step, presumably controlled by the module 14 of RAPS.[29] The VenAT2 and "Hyg"AT2 domains were also predicted to specify malonyl extenders on the basis of the PILEUP analysis and the resulting dendrogram displayed in Figure 6. Each of the three AT domains, encoding ~330 amino acids,[25] were independently introduced into modules 1 or modules 2 of the *eryAI* gene in the chromosome of *Sac. erythraea* through the gene replacement procedures outlined above and illustrated in Figure 8. Since the number of amino acids introduced in each case was roughly equivalent to the number replaced, the "hybrid" DEBS molecules were all roughly of the same size.

Three independent isolates of each strain, carefully analyzed to ensure that the desired replacement had been made and that the reading frame was not disrupted, were examined for the production of polyketides under the standard conditions used for production of erythromycin. For the replacements at module 1, all strains produced bioactive compounds that migrated on TLC plates slightly faster than erythromycin A. Mass spectrometry and NMR analysis confirmed that the bioactive compound produced in each case was 12-desmethylerythromycin B (**29**; Figure 9) which, in addition to lacking the methyl side chain at C-12 of the macrolactone ring, is also lacking the hydroxyl group normally present at C-12. Interestingly, the three strains differed in the amount of polyketide produced, although for any given construct, the replicates all produced roughly the same level of product. The strain carrying the "Hyg"AT2 domain in module 1 of DEBS was the best polyketide producer of the three, producing approximately one-half the amount of **29** as the level of **1** made by the parent strain. The strains carrying the RAPS AT14 or VenAT domains in module 1 produced about one-fourth the level of **29** as the strain carrying the "Hyg"AT2 domain.

Careful examination of the fermentation broths indicated that the other congeners of 10- or 12-desmethylerythromycin, the C and D derivatives, were not produced to any appreciable levels in the strains carrying an AT replacement. In addition, the aglycon or 3-α-mycarosylated intermediates of **29** were also not detected. Taken together, these findings indicated that most of the enzymes that normally act upon the completed polyketide compound in the erythromycin pathway, 6-dEB, also act efficiently on 6-deoxy-12-desmethylerythronolide B. Thus, after its synthesis, 6-deoxy-12-desmethylerythronolide B is, in linear sequence, efficiently hydroxylated at C-6 by EryF, mycarosylated at the C-3 hydroxyl by EryBV, desosaminylated at the C-5 hydroxyl by EryCIII and *O*-methylated at the C″-3 hydroxyl by EryG. Hydroxylation at C-12, however, does not take place to any measurable extent. This is not surprising since P450-mediated hydroxylation (which follows a free-radical mechanism) at the secondary carbon (C-12) in 12-desmethylerythromycin would be expected to be more than 500-fold less favored energetically than the hydroxylation at the corresponding tertiary carbon in erythromycin. What is somewhat surprising is the finding that the 12-desmethyl compound was fully converted to the B form, indicating that efficient methylation of the D congener had taken place. Biochemical experiments have demonstrated that erythromycin D is a rather poor substrate for EryG-mediated *O*-methylation and, in mutants lacking EryK, small amounts of **2** relative to **4** are produced.[42] High-level production of **29**, therefore, suggests that removal of the methyl side chain at C-12 improves the overall kinetics of

the methylation, perhaps by enhancing the binding of the D compound to EryG.

The basis for the differences in rates of production of **29** in the four strains (wild-type and three "hybrids") is presently not understood but it is not likely that it is due to different rates of synthesis of the altered DEBS molecules. Since the three DNA segments inserted *into Sac. erythraea* came from streptomycete hosts which, together with *Sac. erythraea*, share high G + C bias in third codon positions, it is not expected that the rates of translation through the AT1 domains would vary by 4-fold in the four different constructs. It is more likely that the difference in levels of polyketides produced is reflective of the relative difference in activity of the hybrid PKSs. Since, in the course of polyketide chain growth, the AT domain passes the extender to the cognate ACP, it is reasonable to expect that the efficiency of the interaction between ACP1 and the AT domain positioned in module 1 can vary greatly among the four DEBS1 molecules in question. To determine the basis for the differences in levels of polyketide produced, it will be necessary to isolate the various hybrid PKSs and examine them biochemically.

The expected 10-desmethylerythromycin derivatives (**30** and **31**) were found when the "Hyg"AT2 and RAPS AT14 domains were used to replace the AT2 domain in DEBS (Figure 9).[64] Strains carrying the VenAT domain at module 2 did not produce detectable levels of polyketide product. Again, the strain carrying the "Hyg"AT2 domain produced approximately four times as much product as the strain carrying the RAPS AT14 domain, but the total amount of polyketide produced in the best strain was at least 5-fold lower than the level of **29** produced when the same exogenous AT domain was introduced at module 1. In the two cases of exogenous AT replacement at module 2, relatively equal amounts of the A (**29**) and B forms (**30**) were produced. Failure to hydroxylate all the molecules is probably due to reduced binding of **30** to EryK. In addition, primary TLC analysis indicated the presence of several putative intermediates in the pathway to **30** and, potentially, new products that resulted from side reactions. Overall, in contrast to the wild-type strain and the strain carrying the AT replacement in module 1, the polyketide and post-polyketide biosynthesis pathways did not appear to operate as efficiently when the AT replacement was located in module 2.

P. Leadlay and J. Staunton and associates have used an *S. coelicolor*/pCK-like system to generate an AT replacement in module 1 of DEBS.[65] The starting system employed the *eryAI* gene to which the sequence encoding the TE domain from module 6 was placed at the 3′ end and which normally makes triketide lactones **17** and **32** (Figure 9) produced from the two condensations directed by DEBS 1 (described in more detail in section VII.D). Replacement of the AT1 domain of DEBS 1with the AT2 domain from the second module of the rapamycin PKS (again expected to direct incorporation of a malonyl extender unit into the nascent polyketide chain) led to the production of **33** and **34** in which the methyl side chain had been replaced by a proton, indicating incorporation of malonate in place of methylmalonate (Figure 9). The basis for the generation of two structurally different polyketides is discussed below.

C. Changes in Starter Units

The 16-membered macrocylic lactones produced in the biosynthetic pathways of the macrolide antibiotics spiramycin and tylosin are platenolide I (**35**) and tylactone (**36**), respectively (Figure 10). Among the structural differences between the two polyketides are the side chains: methyl in **35** and ethyl in **36**. Platenolide I biosynthesis begins with the condensation of actetate and malonate; tylactone biosynthesis starts with the condensation of propionate and methylmalonate. At the genetic level, the two PKSs have very similar organizations. As shown in Figure 10, the loading module of the platenolide I PKS gene (*srmG*) in the spiramycin producer, *S. ambofaciens*, encodes the KS^o-ATs-ACPs functions. Immediately upstream of the KS^o segment is the *srmG* promoter.[66] As discussed in section II, KS^o is a nonfunctional domain that resembles KS domains but carries the amino acid glutamine in place of cysteine at the active site and hence cannot form the thioester linkage with the starter compound required for the condensation reaction. Platenolide I synthesis is expected follow the following pathway at the biochemical level: (i) binding of acetyl CoA to the AT loading domain; (ii) transfer of the acetyl moiety to the ACP loading domain; and (iii) transfer of the acetyl moiety to the active-site cysteinyl residue of the KS1 domain where it participates in the first condensation reaction with a malonyl moiety at the ACP1 domain. Interestingly, the loading module of the tylactone PKS gene (*tylG*) in *S. fradiae*, the tylosin producer, also encodes KS^o-ATs-ACPs functions but in this case, since tylactone synthesis starts with propionate, the ATs is believed to specify binding and transfer of a propionyl residue to ACPs. At the genetic level, the loading modules from the two PKS systems are similar in length and codon bias.

In an elegant experiment (Figure 10), the segment encoding the KS^o-ATs-ACPs functions from *tylG* was introduced into the *S. ambofaciens* chromosome to replace the KS^o-ATs-ACPs segment of *srmG* precisely, resulting in the creation of a hybrid PKS under the control of the *srmG* promoter.[31] The methodology to perform the gene replacement experiment was similar to that described in section V. In addition to the gene replacement in *srmG*, the strain constructed also carried a deletion of a segment of the chromosome upstream of the *srmG* promoter involved in the synthesis and attachment of the sugars to the macrolactone. Mass spectrometric and NMR analysis demonstrated that the genetic construct produced the single compound 16-methylplatenolide I (**37**), indicating that the loading module replacement directed the hybrid platenolide PKS to start the synthesis with propionate in place of acetate. Of interest was the finding that the strain carrying the hybrid PKS produced the hybrid polyketide at levels 3−4-fold higher than the level of platenolide I synthesis in the control strain. The basis for the rise in the level of polyketide synthesis is not currently understood. If the overall rate of polyketide production is determined by the level of precursor supply, it is curious

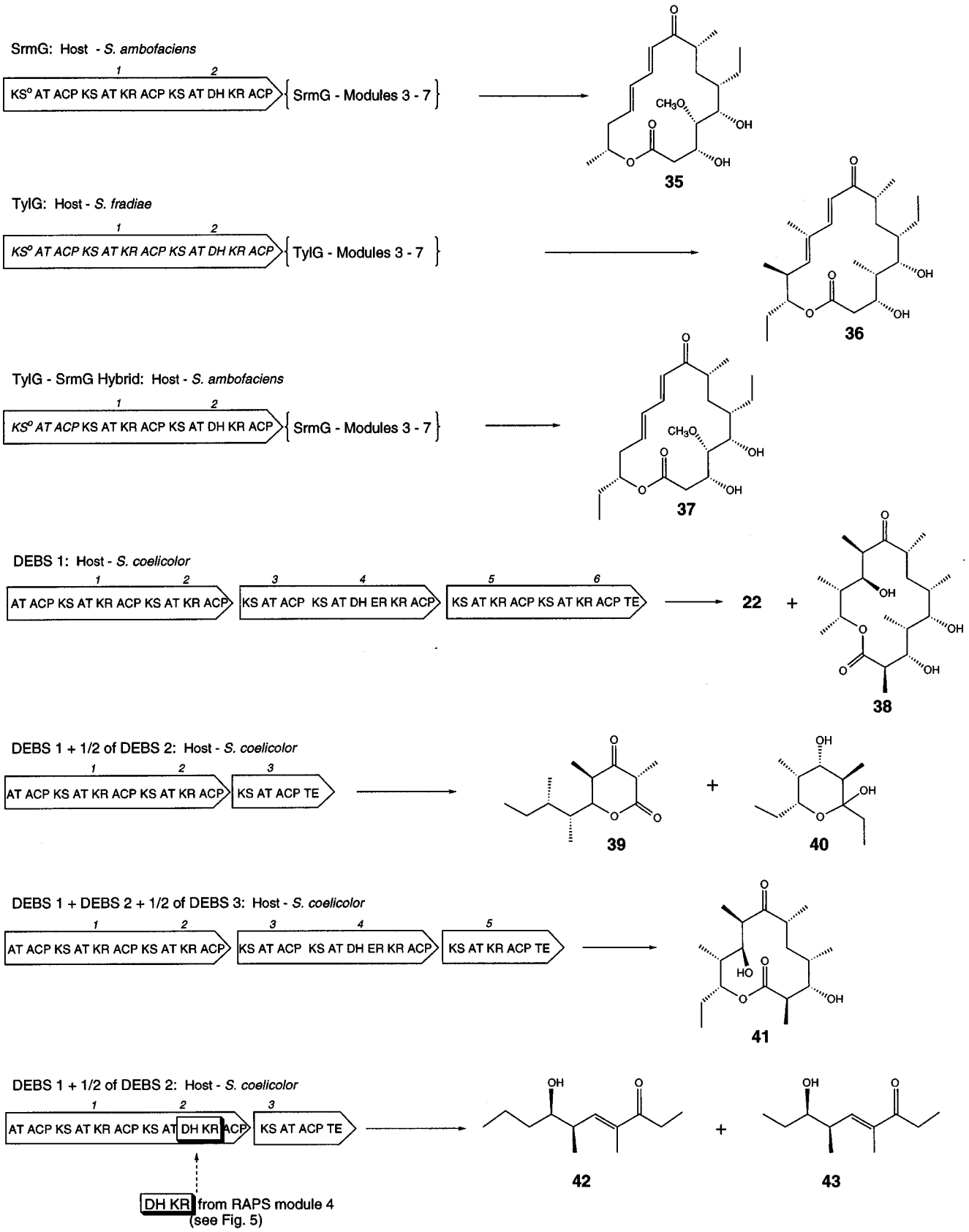


Figure 10. Summary of genetic manipulations of PKS genes and corresponding products produced by host strains carrying altered PKS genes. See Figures 4 and 5 for abbreviations.

that the availability of free acetyl CoA would limit the rate of polyketide synthesis. This point will be readdressed below.

Although the replacement experiment in *srmG* somewhat resembles the AT replacements in *eryA*, it is important to note the differences. In the cases where AT replacements were made in *eryA*, only the AT domains were swapped; in *srmG*, both the AT and its cognate ACP domains were exchanged with an exogenous AT−ACP cognate set. Although the AT exchanges in both genes resulted in a similar types of changes in the structure of the resulting polyketides,

at the biochemical level the experiments indicate substantially different effects. In the *eryA* case, five different AT domains in total were placed in module 1: the resident DEBS AT1 domain, RAPS AT2 and RAPS AT14 domains, the "Hyg"AT2 domain from another PKS gene set present in the rapamycin producer *S. hygroscopicus*, and the VenAT domain from a partially characterized PKS from *S. venzuelae*. In each of these cases, during polyketide synthesis, the AT domain sitting in module 1 had to interact with the same ACP1 domain to pass the extender unit to the pantotheine side chain of ACP1 to form a stable thioester bond. Similarly, in the AT replacements constructed in module 2, the AT sitting in module 2 had to interact with the same ACP2. In seven of the nine cases compared, the AT−ACP interaction was between a noncognate pair, of which six were observed to be functional. In the *srmG* case, one AT−ACP cognate pair was exchanged for another. Thus the non-cognate interaction was between the ACPs domain from *tylG* and the KS1 domain from *srmG*. ACPn−KSn+1 (where n = module number) interactions involve thioacyl transfer of the nascent polyketide from the ACP domain to the next KS in sequence for subsequent condensation. As described in section II, the ACP domain plays a central role in polyketide synthesis, interacting with every domain within its resident module as well as the KS domain from the subsequent one used in polyketide synthesis. Taken together, the AT replacement studies demonstrate quite convincingly that ACP domains in PKSs can interact with noncognate AT and KS domains. If their role is to provide a structural support for the panthotheine side chain that participates in the chemistry of polyketide synthesis, it is highly likely that the ACPs in PKSs would also be interchangeable.

D. Changes in Acyl Chain Length

This section is treated in extensive detail in the article by C. Khosla in this issue and will be only briefly reviewed here. With a single exception, the *S. coelicolor* CH999/pCK7 system (see section V) and its derivatives were used for the constructions described in this section. Results are summarized in Figure 10.

Starting with the three *eryA* genes, it was observed that the *S. coelicolor* strain produced two polyketides: the normal 6-dEB compound (**22**) as well as 8,8a-deoxyoleandolide (**38**) in a 4:1 ratio.[59] The **38** compound is produced by the DEBS using acetyl CoA as the starter in place of propionyl CoA and has been found in extremely small amounts in fermenters producing high levels of erythromycin.[67] Since the amount of polyketide produced by *S. coelicolor* is low by comparison to that produced by *Sac. erythraea*, the substantial synthesis of **38** indicates a shortage of supply of propionyl CoA in that host. Addition of propionate to the growth media resulted in exclusive production of **17**, no doubt through conversion to propionyl CoA by an acyl CoA carboxylase. In *Sac. erythraea*, efficient synthesis of erythromycin can occur in a mutant lacking propionyl carboxylase activity[68] but does not take place at all in a mutant disrupted in the gene *eryM*, which encodes a malonyl decarboxylase, unless propionate is added to the growth medium.[69] These findings suggest, therefore, that in *Sac. erythraea*, propionyl CoA is derived from the EryM-catalyzed decarboxylation of methylmalonyl CoA (itself derived from catabolism of isoleucine and other amino acids) and that EryM or its functional equivalent is not present in *S. coelicolor*.

A significant advance in PKS biochemistry was made when it was observed that a pCK7 derivative carrying only the *eryAI* gene (encoding DEBS1) could produce small quantities of the triketide lactone **17** that is formed through the action of modules 1 and 2 and which undergoes spontaneous lactonization after release from ACP2.[70] Placement of the TE domain from module 6 at the end of module 2 increased the level of production of the triketide lactone and also resulted in production of **38**, a triketide lactone that starts with acetate (again, probably due to the depletion of propionyl CoA). A similar construct made in *Sac. erythraea* also resulted in the production of **17** but the nor derivative was not observed, likely since there was no shortage of propionyl CoA.[55] Although it is thought that the TE domain serves the role of both thioesterase and cyclase in 6-dEB formation (by directing the released carboxylate function at the end of the acyl chain to the C-13 hydroxyl group in the chain, rather than to water), it is likely that placement of the TE domain at the end of module 2 enhances the efficiency of release of the triketide from DEBS1 thereby increasing the overall rate of triketide lactone production.

These findings indicated that DEBS1 can function in the absence of DEBS2 and DEBS3. However, since, with a single exception, **17** was never seen in fermentation beers of wild-type *Sac. erythraea* or in mutants producing erythromycin pathway intermediates, it is safe to conclude that premature release of the nascent chain from ACP2 does not occur to any appreciable extent. The exception is the strain of *Sac. erythraea* that carried the KR5 deletion which produced exceedingly small amounts of **17**. At the time of its identification, it was proposed that the compound was produced by premature release from ACP2 during acyl chain synthesis.[71]

A construct containing modules 1, 2 (DEBS1), and 3 with the TE6 domain placed after ACP3 (half of DEBS2) was constructed in the pCK system and found to make a nine carbon tetraketide with the correct levels of reduction at the three carbonyls but which had undergone two different spontaneous lactonizations to produce compounds **39** and **40**.[72] Placement of the TE6 after ACP5 in module 5 and removal of module 6 resulted in the production of the 12-membered lactone **41**.[73] That the predicted products were made from DEBS complexes in which one-half of either DEBS2 or DEBS3 was missing indicated that preservation of the orginal architecture of these proteins is not required for their function.

E. Gain of Function

A construct containing modules 1, 2 (DEBS1), and 3 with the TE6 domain placed after ACP3 (half of DEBS2) was engineered wherein the KR domain of DEBS module 2 was replaced with the DH and KR domains of the fourth module of the rapamycin PKS employing the *S. coelicolor*/pCK7 system. The resulting strain produced the acyclic tetraketides **42**

and **43** (which had undergone terminal decarboxylation) that contained the ene function at the predicted position. These findings indicated that the DH domain functioned in processing the carbonyl generated after the second condensation.[74] The appearance of compound **43** indicated that the PKS employed acetyl CoA rather propionyl CoA as the starter as had been seen in other polyketides made in the *S. coelicolor* system. Hence, the slighly larger DEBS 1 component of the eryPKS, produced by the addition of a DH domain to module 2, not only functioned as expected during the second condensation cycle but also could interact properly with DEBS 2 to pass the growing polyketide chain.

VII. Conclusions

The examples described here made it clear that it is becoming increasingly possible to change the structure of complex polyketides in a predicted manner by introducing directed changes in modular polyketide synthases. The AT replacements, first with domains taken from well-characterized PKSs which catalyzed the synthesis of known products, and then with domains taken from partially characterized PKS-like sequences from which the compounds they produced (if any) were not known, dramatizes the concept that the large diversity of polyketide structures found in nature is built on enzymes with common structural and biochemical bases with slightly different (and interchangeable?) units. The ability to use uncharacterized AT domains to make predicted side-chain changes also underscores the significance of the clustering of the domains by function, first determined by Leadlay and colleagues. As new AT domains become sequenced, will it be possible to replace methyl groups with ethyl or propyl groups? Replacements at the starter unit are even more interesting since there are a great variety of structures that are known or presumed to start polyketide biosynthesis. We have seen that propionate can be substituted for acetate, but will it be possible to replace the erythromycin starter AT with heterologous loading domains that will yield the erythromycin polyketide backbone with, for example, the branched chain starter from avermectin, or with other, more exotic compounds? At the end of the cycle, can we attach pipecolate or another amino acid to the erythromycin polyketide backbone? Finally, which of the wide variety of changes that can be made to the polyketide structure will be tolerated by the enzymes that process the polyketide into the bioactive macrolide compound?

To date we have learned that not only are the modular PKSs somewhat flexible for the structures of the substrates they process, but also that the individual components of complex PKS assemblies can function on their own. Splitting a bimodular polypeptide in half does not seem to affect the function of at least the N-terminal module contained within. The assembling of heterologous components at the modular level or larger into functional novel PKS complexes is still some time away but as we learn more about the structure of modular PKSs, the prospects of creating novel polyketides by genetic manipulation of PKS genes, perhaps even combinatorially,[75] will steadily improve.

VIII. Acknowledgments

I am grateful to my many excellent colleagues at Abbott Laboratories for their dedication and fortitude over the many years we have engaged in this research. Finally, I thank Rich Summers for reviewing the manuscript.

IX. References

(1) (1) Birch, A. J. *Science* **1967**, *156*, 202.
(2) Hutchinson, C. R.; Fujii, I. *Annu. Rev. Microbiol.* **1995**, *49*, 201.
(3) Hopwood, D. A.; Sherman, D. H. *Annu. Rev. Genet.* **1990**, *24*, 37.
(4) Katz, L.; Donadio, S. *Annu. Rev. Microbiol.* **1993**, *47*, 875.
(5) Cane, D. E.; Yang, C. *J. Am. Chem. Soc.* **1987**, *109*, 1255.
(6) Yue, S.; Motamedi, H.; Yamamoto, Y.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1987**, *109*, 1253.
(7) Cane, D. E.; Ott, W. R. *J. Am. Chem. Soc.* **1988**, *110*, 4840.
(8) Cane, D. E.; Lambalot, R. H.; Prabhakaren, P. C.; Ott, W. R. *J. Am. Chem. Soc.* **1993**, *115*, 522.
(9) Spavold, Z. M.; Robinson, J. A. *J. Chem Soc., Chem. Commun.* **1988**, 4.
(10) Staunton, J.; Sutkowski, A. C. *J. Chem Soc., Chem. Commun.* **1991**, 1110.
(11) Yoshizawa, Y.; Li, Z.; Reese, P. B.; Vederas, J. C. *J. Am. Chem. Soc.* **1990**, *112*, 3212.
(12) Li, Z.; Martin, F. M.; Vederas, J. C. *J. Am. Chem. Soc.* **1992**, *114*, 1531.
(13) Kinoshita, K.; Takenaka, S.; Hayashi, M. *J. Chem. Soc., Chem. Commun.* **1988**, 943.
(14) Kinoshita, K.; Takenaka, S.; Hayashi, M. *J. Chem. Soc., Perkin Trans.* **1991**, *1*, 2247.
(15) Huber, M. L.; Paschal, J. W.; Leeds, J. P.; Kirst, H. A.; Wind, J.; Miller, F. D.; Turner, J. A. *Antimicrob. Agents Chemother.* **1990**, *34*, 1535.
(16) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Gene* **1992**, *115*, 97.
(17) Pieper, R.; Kao, C. M.; Khosla, C.; Luo, G.; Cane, D. E. *Chem. Soc. Rev.* **1996**, *25*, 297.
(18) Martin, J. R.; Perun, T. J. *Biochemistry* **1967**, *6*, 1728.
(19) Tuan, J. S.; Weber, J. M.; Staver, M. J.; Leung, J. O.; Donadio, S.; Katz, L. *Gene* **1990**, *90*, 21.
(20) Weber, J. M.; Leung, J. O.; Maine, G. T.; Potenz, R. H.; Paulus, T. J.; Dewitt, J. P. *J. Bacteriol.* **1990**, *172*, 2372.
(21) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Science* **1991**, *252*, 675.
(22) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. *Nature* **1990**, *346*, 176.
(23) Donadio, S.; Staver, M. J. *Gene* **1993**, *126*, 147.
(24) Bevitt, D. J.; Cortes, J.; Haydock, S. F.; Leadlay, P. F. *Eur. J. Biochem.* **1992**, *204*, 39.
(25) Donadio, S.; Katz, L. *Gene* **1992**, *111*, 51.
(26) Caffrey, P.; Bevitt, D. J.; Staunton, J.; Leadlay, P. F. *FEBS Lett.* **1992**, *304*, 225.
(27) Smith, S. *FASEB J.* **1994**, *8*, 1248.
(28) Marsden, A. F.; Caffrey, P.; Aparracio, J. F.; Loughran, M. S.; Staunton, J.; Leadlay, P. F. *Science* **1994**, *263*, 378.
(29) Schweke, T.; Aparacio, J. F.; Molnar, I.; Konig, A.; Khaw, L. E.; Haydock, S. F.; Oliynyk, M.; Caffrey, P.; Cortes, J.; Lester, J. B.; Bohm, G. A.; Staunton, J.; Leadlay, P. F. *Proc. Nat. Acad. Sci. U.S.A.* **1995**, *92*, 7839.
(30) Aparacio, J. J.; Molnar, I.; Schweke, T.; Konig, A.; Haydock, S. F.; Khaw, L. E.; Staunton, J.; Leadlay, P. F. *Gene* **1996**, *169*, 9.
(31) Kuhstoss, S.; Huber, M.; Turner, J. R.; Paschal, J. W.; Rao, N. *Gene* **1996**, *183*, 231.
(32) Swan, D. G.; Rodriguez, A. M.; Vilches, C.; Mendez, C.; Salas, J. A. *Mol. Gen. Genet.* **1994**, *242*, 358.
(33) MacNeil, D. J.; Occi, J. L.; Gewain, K. M.; MacNeil, T.; Gibbons, P. H.; Ruby, C. L.; Danis, S. J. *Ann. N.Y. Acad. Sci.* **1994**, *721*, 123.
(34) Hu, Z.; Bao, K.; Zhou, X.; Zhou, Q.; Hopwood, D. A.; Kieser, T.; Deng, Z. *Mol. Microbiol.* **1994**, *14*, 163.
(35) GenBank accession numbers: U11039; U00024; AD000014; U00023; U00010. EMBL accession numbers: Z83858; Z85982; Z83857Z274697.
(36) Ruan, X.; Stassi, D. L.; Lax, S.; Katz, L. *Gene* **1997**, in press.
(37) Haydock, S. F.; Aparacio, J. F.; Molnar, I.; Schweke, T.; Khaw, L. E.; Konig, A.; Marsden, A. F.; Galloway, I. S.; Staunton, J.; Leadlay, P. F. *FEBS Lett.* **1995**, *374*, 246.
(38) Devereux, J.; Haeberli, P.; Smithies, O. *Nucleic Acids Res.* **1984**, *12*, 387.
(39) Weber, J. M.; Leung, J. O.; Swanson, S. J.; Idler, K. B.; McAlpine, J. B. *Science* **1991**, *252*, 114.
(40) Andersen, J. F.; Hutchinson, C. R. *J. Bacteriol.* **1992**, *174*, 725.
(41) Summers, R. G.; Donadio, S.; Staver, M. J.; Wendt-Pienkowski, E.; Hutchinson, C. R.; Katz, L. *Microbiology*, in press.
(42) Stassi, D. L.; Donadio, S.; Staver, M. J.; Katz, L. *J. Bacteriol.* **1993**, *175*, 182.

(43) Weber, J. M.; Schoner, B.; Losick, R. *Gene* **1989**, *75*, 235.
(44) Paulus, T. J.; Tuan, J. S.; Luebke, V. E.; Maine, G. T.; DeWitt, J. P.; Katz, L. *J. Bacteriol.* **1990**, *172*, 2541.
(45) Lambalot, R. H.; Cane, D. E.; Aparacio, J. J.; Katz, L. *Biochemistry* **1995**, *34*, 1858.
(46) Uchiyama, H.; Weisblum, B. *Gene* **1985**, *38*, 103.
(47) Thompson, C. J.; Kieser, T.; Ward, J. M.; Hopwood, D. A. *Gene* **1982**, *20*, 51.
(48) Stanzak, R.; Matsushima, P.; Baltz, R. H.; Rao, N. *Bio/Technology* **1986**, *4*, 229.
(49) Dhillon, N.; Hale, R. S.; Cortes, J.; Leadlay, P. F. *Mol. Microbiol.* **1989**, *3*, 1405.
(50) Haydock, S. F.; Dowson, J. A.; Dhillon, N.; Roberts, G. A.; Cortes, J.; Leadlay, P. F. *Mol. Gen. Genet.* **1991**, *230*, 120.
(51) Pereda, A.; Summers, R. G.; Katz, L. *Gene*, in press.
(52) Dewitt, J. P. *J. Bacteriol.* **1985**, *164*, 969.
(53) Katz, E.; Thompson, C. J.; Hopwood, D. A. *J. Gen. Microbiol.* **1983**, *129*, 2703.
(54) Vara, J. A.; Lewandoska-Skarbek, M.; Wang, Y.-G.; Donadio, S.; Hutchinson, C. R. *J. Bacteriol.* **1989**, *170*, 5872.
(55) Cortes, J.; Wiesmann, K. E.; Roberts, G. A.; Brown, M. J.; Staunton, J.; Leadlay, P. F. *Science* **1995**, *268*, 1487.
(56) Weber, J. M.; Losick, R. *Gene* **1988**, *68*, 173.
(57) Hilleman, D.; Puehler, A.; Wohlleben, W. *Nucleic Acids Res.* **1991**, *19*, 727.
(58) Hosted, T. J.; Baltz, R. H. *J. Bacteriol.* **1997**, *179*, 180.
(59) Fernandez-Moreno, M. A.; Cabello, J. L.; Hopwood, D. A.; Malpartida, F. *Cell* **1991**, *66*, 769.
(60) Kao, C. M.; Katz, L.; Khosla, C. *Science* **1994**, *265*, 509.
(61) Andersen, J. F.; Tatsuta, K.; Gunji, H.; Ishiyama, T.; Hutchinson, C. R. *Biochemistry* **1992**, *32*, 1905.
(62) Donadio, S.; McAlpine, J. B.; Sheldon, P. J.; Jackson, M.; Katz, L. *Proc. Nat. Acad. Sci. U.S.A.* **1993**, *90*, 7119.
(63) Leadlay, P. F.; Staunton, J.; Aparacio, J. F.; Bevitt, D. J.; Caffrey, P.; Cortes, J.; Marsden, A.; Roberts, G. A. *Biochem. Soc. Trans.* **1993**, *21*, 218.
(64) Ruan, X.; Stassi, D. L.; Pereda, A.; Zeidner, D.; Shivakumar, A.; Kakavas, S.; Jackson, M.; Summers, R. G.; Katz, L. *J. Bacteriol.*, in press.
(65) Oliynik, M.; Brown, M. J.; Cortes, J.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **1996**, *3*, 833.
(66) Geistlich, M.; Losick, R.; Turner, J. R.; Rao, N. *Mol. Microbiol.* **1992**, *6*, 2019.
(67) Kibawage, I. O.; Janssen, G.; Busson, R.; Hoogmartens, J.; Vanderhaeghe, H.; Verbist, L. *J. Antibiot.* **1987**, *40*, 1.
(68) Donadio, S.; Staver, M. J.; Katz, L. *Mol. Microbiol.* **1996**, *19*, 977.
(69) Hsieh, Y.-J.; Kolattukudy, P. E. *J. Bacteriol.* **1994**, *176*, 714.
(70) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 11612.
(71) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Gene* **1992**, *115*, 97.
(72) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1996**, *118*, 9184.
(73) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1995**, *117*, 9105.
(74) McDaniel, R.; Kao, C. M.; Fu, H.; Hezevi, P.; Gustafson, C.; Betlach, M.; Ashley, G.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1997**, *119*, 4309.
(75) Khosla, C.; Zawada, R. J. *Trends Biotechnol.* **1996**, *14*, 335.

CR960025+