Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 1846–1851, March 1999
Applied Biological Sciences

Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products

ROBERT MCDANIEL*, ARINTHIP THAMCHAIPENET†, CLAES GUSTAFSSON, HONG FU, MELANIE BETLACH, MARY BETLACH, AND GARY ASHLEY

KOSAN Biosciences, Inc., 1450 Rollins Road, Burlingame, CA 94010

Communicated by Christopher T. Walsh, Harvard Medical School, Boston, MA, December 31, 1998 (received for review December 2, 1998)

ABSTRACT **The structures of complex polyketide natural products, such as erythromycin, are programmed by multifunctional polyketide synthases (PKSs) that contain modular arrangements of functional domains. The colinearity between the activities of modular PKS domains and structure of the polyketide product portends the generation of novel organic compounds—"unnatural" natural products—by genetic manipulation. We have engineered the erythromycin polyketide synthase genes to effect combinatorial alterations of catalytic activities in the biosynthetic pathway, generating a library of >50 macrolides that would be impractical to produce by chemical methods. The library includes examples of analogs with one, two, and three altered carbon centers of the polyketide products. The manipulation of multiple biosynthetic steps in a PKS is an important milestone toward the goal of producing large libraries of unnatural natural products for biological and pharmaceutical applications.**

Few molecules have captured interest in both chemotherapy and chemistry to the extent of the polyketide erythromycin. Erythromycin and its semisynthetic derivatives are the third most widely used class of antibiotics, with current worldwide sales exceeding 3.5 billion dollars. In addition, erythromycin analogs are gaining interest for their potential use in the treatment of gastrointestinal disorders and inflammatory diseases and as next-generation antibiotics for treatment of emerging drug-resistant strains of bacteria (1, 2). The chemical challenges of erythromycin attracted the talents of R. B. Woodward and 48 colleagues who completed its synthesis in 1981 (3) and of a cadre of medicinal chemists who prepared analogs leading to the important second generation of macrolide antibiotics—clarithromycin, azithromycin, and others (4). Although such efforts effectively saturated the chemical modifications possible at the existing functional groups of the macrolide ring, most of the ring remained inaccessible to chemical modification. After the discovery of the programmed nature of modular polyketide biosynthesis (5, 6), genetic engineering strategies emerged for the production of novel polyketides (7, 8). With the work described here, it is now feasible to contemplate modifications of the chemically intractable sites of such molecules to produce hundreds or even thousands of new "unnatural" natural products by genetic engineering. Such novel macrolides could in themselves provide the basis for new pharmaceuticals or could serve as scaffolds for new semisynthetic analogs.

It is now well established that the "modular" polyketide synthases (PKSs) (*i*) each are encoded by a cluster of contiguous genes and (*ii*) have a linear, modular organization of similar catalytic domains that both build and modify the polyketide backbone. Each module contains a set of three domains—a ketosynthase, an acyltransferase (AT), and an acyl carrier protein (ACP)—that catalyze a 2-carbon extension of the growing polyketide chain (Fig. 1) (9). The choice of extender unit used by each module—acetate, propionate, or other small organic acids in the form of CoA thioesters—is determined by the specificity of the AT domain (10–12). With each 2-carbon chain extension, the final state of the β-carbonyl is embedded as a ketone, hydroxyl, methenyl, or methylene group by the presence or absence of one, two, or three additional catalytic domains in the module—a ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER). In effect, the composition of catalytic domains within a module provides a "code" for the structure of each 2-carbon unit, and the order of modules codes for the sequence of the 2-carbon units, together creating a template colinear with the polyketide product. The remarkable structural diversity of polyketides is governed by the combinatorial possibilities of catalytic domains within each module, the sequence and number of modules, and the post-polyketide synthesis cyclization and "tailoring" enzymes that accompany the PKS genes. The direct correspondence between the catalytic domains of modules in a PKS and the structure of the resulting biosynthetic product portends the possibility of modifying polyketide structure by modifying the domains of the modular PKS.

We have constructed a combinatorial library of polyketides by using 6-deoxyerythronolide B synthase (DEBS), the PKS that produces the macrolide ring of erythromycin. This was accomplished by substituting the ATs and β-carbon processing domains of DEBS with counterparts from the rapamycin PKS (RAPS) (13) that encode alternative substrate specificities and β-carbon reduction/dehydration activities. Engineered DEBS containing single, double, and triple catalytic domain substitutions catalyzed production of erythromycin macrolactones with corresponding single, double, and triple modifications. The ability to simultaneously manipulate multiple catalytic centers of the PKS demonstrates the robustness of the engineering process and the potential for creating libraries of novel polyketides that are impractical to prepare in the chemistry laboratory.

MATERIALS AND METHODS

Strains, Culture Conditions, and DNA Manipulation. DNA manipulations were performed in *Escherichia coli* XL1 Blue

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

PNAS is available online at www.pnas.org.

Abbreviations: PKS, polyketide synthase; AT, acyltransferase; KR, ketoreductase; DH, dehydratase, ER, enoylreductase; ACP, acyl carrier protein; DEBS, 6-deoxyerythronolide B synthase; RAPS, rapamycin polyketide synthase; 6dEB, 6-deoxyerythronolide B.

†Present address: Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.

*To whom reprint requests should be addressed. e-mail: mcdaniel@kosan.com.

(Stratagene) by using standard procedures (14). *Streptomyces coelicolor* CH999 (15) or *Streptomyces lividans* K4–114 (31) was used as the host for production of polyketides. *Streptomyces* spp. were cultured on R2YE agar (16) plates at 30°C. *Streptomyces* spp. protoplasts were transformed by the standard procedure (16), and transformants were selected by using 2 ml of a 500 μg/ml thiostrepton overlay.

Restriction Site Engineering. PCR mutagenesis was used to introduce restriction sites in subclones containing portions of the DEBS genes. Replacement of the DEBS domains by the RAPS cassettes were performed in the subclones before introduction into pCK7 for expression (17). The location of restriction sites at the domain boundaries was based on amino acid sequence alignment of the DEBS modules (18). The *Pst*I and *Xba*I sites in module 2 are identical to those reported (19). The remaining engineered sites generated the following sequences at the domain boundaries (restriction sites italicized): module 2 *Bam*HI, TCCGACGGT*GGATCC*GTGTTCGTC; module 5 *Bam*HI, ACTCGCCGC*GGATCC*GCGATGGTG; module 5 *Pst*I, CGGTACTGG*CTGCAG*ATCCCCACC; module 5 *Xba*I, GAGGAGGGC*TCTAGA*CTCGCCCAG; module 6 *Bam*HI, TCCGCCGGC*GGATTC*GTTTTCGTC; module 6 *Pst*I, CGGTACTGG*CTGCAG*CCGGAGGTG; and module 6 *Xba*I, GTGGGGGCC*TCTAGA*GCGGTGCAG.

Construction of Replacement Cassettes. Construction of the *rap*DH/KR4 and *rap*DH/ER/KR1 cassettes has been described (20, 21). Oligonucleotide primers used for PCR amplification of *rap*AT2 were forward, 5′-TTT*GGATCC*G-TGTTCGTCTTCCCGGGTCAGGGGTCG-3′; and reverse, 5′-TTT*CTGCAG*CCAGTACCGCTGGTGCTGGA-AGGCGTA-3′. The italicized residues indicate the *Bam*HI and *Pst*I sites used for ligation to the engineered DEBS sites. The AT/ACP linker was generated by annealing the following two oligonucleotides which create cohesive ends for ligation to the *Pst*I and *Xba*I sites in DEBS: forward, 5′-GCCGGACCGCACCACCCCTCGTGACGGAGAAC-CGGAGACGGAGAGCT-3′; reverse, 5′-CTAGAGCTCT-CCGTCTCCGGTTCTCCGTCACGAGGGGTGGTGCG-GTCCGGCTGCA-3′. This results in the amino acid sequence *PDRTTPRD*GEPETES, which contains portions of DEBS between the AT and KR italicized and the KR and ACP domains of DEBS module 2.

Production and Analysis of Polyketides. *S. coelicolor* CH999 and *S. lividans* K4–114 are genetically engineered strains containing chromosomal deletions of the entire ≈22-kilobase actinorhodin polyketide gene cluster (15). Macrolide production from DEBS is indistinguishable when expressed in either host strain. Strains expressing the mutant PKSs were grown as confluent lawns on R2YE agar medium supplemented with 5 mM sodium propionate. The Petri plates (13 × 150 mm) were fitted with sterile filter disks (Whatman no. 52, 125 mm) before filling with 50 ml of media. After 3 days of growth, the filter paper and agar were transferred to another Petri dish containing 50 ml of liquid R2YE (+ 5 mM sodium propionate), XAD-16 resin, and 6-mm glass beads for support. After 5 additional days of growth, the XAD resin was collected and extracted with 10 ml of ethanol. The ethanol extracts were dried and partitioned between ethyl acetate and saturated aqueous $NaHCO_3$. The ethyl acetate fractions were analyzed by HPLC (C-18 column, water-acetonitrile gradient) coupled to atmospheric pressure–chemical ionization. Quantitative determination of polyketides was made with evaporative light scattering detection. Compounds identified here are the major metabolites produced. Some strains also contained detectable levels of one or more minor components not reported.

Characterization of Compounds. Structure determination was based primarily on the agreement between the structure predicted for the directed mutagenesis performed and the mass spectrum. Under the chemical ionization conditions used for mass spectrometry, 6-deoxyerythronolide B (6dEB) and its analogs generate signature dehydration patterns corresponding to the ring hydroxyls and lactone group. Mass spectrometry data are provided as supplementary information. Further structural validation by NMR spectroscopy on selected compounds also was performed. Compound **11** was reported previously by Liu *et al*., who engineered a similar DEBS mutant (11).

RESULTS

Strategy for Library Construction. The DEBS multienzyme complex consists of three large subunits (>300 kDa), each containing two modules (Fig. 1). In all, there are 28 catalytic domains responsible for the priming, chain extension, β-carbon modification, and cyclization of the polyketide during biosynthesis of 6dEB (**1**) (5, 6). Thus far, individual mutagenesis strategies that have successfully altered the catalytic properties of DEBS include (*i*) deletion of modules to control chain length (22–25), (*ii*) inactivation of reduction/dehydration domains to bypass β-carbon processing steps (6, 19, 26), (*iii*) substitution of AT domains to alter starter and extender unit incorporation (10–12, 27, 28), (*iv*) addition of reduction/dehydration domains to introduce catalytic activities (20, 21), and (*v*) substitution of KR domains to control hydroxyl stereochemistry (29). Although these experiments revealed some tolerance of DEBS for alteration of individual activities, the limit of this tolerance

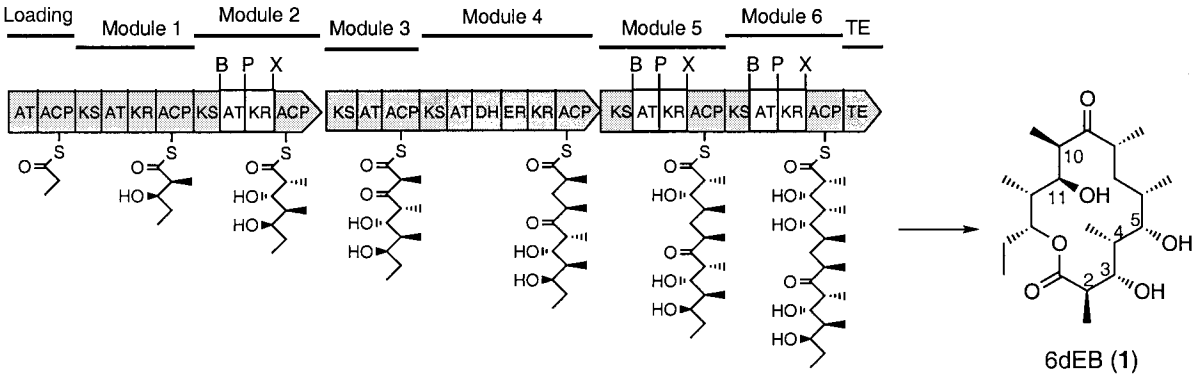

FIG. 1. Genetic architecture of DEBS. DEBS catalyzes formation of 6-deoxyerythronolide B (**1**) from decarboxylative condensations between one propionyl-CoA priming unit and six methylmalonyl-CoA extender units. For β-carbon processing, modules 1, 2, 5, and 6 contain KR domains, module 4 contains the complete KR, DH, and ER set, and module 3 lacks any functional β-carbon-modifying domains. The loading segment consists of priming AT and ACP domains, and a thioesterase (TE) catalyzes the release and cyclization of the polyketide chain. For this work, restriction endonuclease sites were engineered around AT and KR domains in modules 2, 5, and 6. (B, *Bam*HI; P, *Pst*I; X, *Xba*I; KS, ketosynthase).

was not demonstrated. The extent to which this approach is useful for producing large numbers of polyketides critically depends on the enzyme's acceptance of multiple changes in the biosynthetic pathway.

We systematically engineered single and multiple enzymatic domain substitutions in DEBS to demonstrate the broader applicability of PKS mutagenesis techniques. Modules 2, 5, and 6 of DEBS, each featuring only a KR for β-carbon processing, provided an excellent template for testing the effects of AT specificity alteration, reductive domain deletion, and reductive domain gain-of-function in three separate modules by using a series of replacement catalytic domains. For AT substitutions, the malonyl-CoA transferase from module 2 of RAPS (*rap*AT2) was used to replace methylmalonyl AT domains of DEBS. The resulting mutants were expected to incorporate acetate rather than propionate units to generate 6dEB analogs lacking a methyl substituent at the engineered positions (10, 11). Gain-of-function mutagenesis was performed by replacement of ketoreductases with cassettes containing the DH+KR domains from RAPS module 4 (*rap*DH/KR4) and the DH+ER+KR domains from RAPS module 1 (*rap*DH/ER/KR1). Successful substitution with these cassettes replaces the corresponding alcohol moieties of 6dEB with alkene and alkane carbons, respectively (20, 21). Deletion mutagenesis to convert hydroxyl groups of 6dEB to ketones was performed by substituting KR domains with a synthetic 18-aa fragment (AT/ACP linker) joining the AT and ACP domains.

Restriction sites were engineered around the boundaries of the AT and KR domains to facilitate mutagenesis (Fig. 1). The engineered sites had no effect on the level of 6dEB production. Appropriate cassettes from RAPS then were inserted into the AT or KR positions of modules 2, 5, and 6 of the full DEBS system encoded on the *Streptomyces* expression plasmid pCK7 (17). The resulting plasmids were introduced into either *S. coelicolor* CH999 or *S. lividans* K4–114, and the transformed strains were analyzed for polyketide production by HPLC/MS.

Single Catalytic Domain Substitutions. Nearly all of the strains expressing PKSs with a single mutation produced polyketides with molecular weights matching the predicted 6dEB analog and with production levels ranging from 1 to 70% of wild-type 6dEB (**1**) (Table 1). The *rap*AT2 substitutions generated functional hybrid PKSs in each of the three modules, producing 10-desmethyl (**2**), 4-desmethyl (**7**), and 2-desmethyl (**11**) 6dEB analogs as predicted. All three *rap*DH/KR4 substitutions also resulted in functional PKSs, generating 10,11-anhydro (**3**), 4,5-anhydro (**9**), and 2,3-anhydro (**13**) derivatives. The two strains carrying the AT/ACP linker substitutions in modules 5 and 6 produced 5-deoxy-5-oxo (**8**) [previously reported by Donadio *et al.* from an *ery*KR5 deletion in *Saccharopolyspora erythraea*) (6)] and 3-deoxy-3-oxo (**12**) 6dEB analogs. However, a macrolide product was not detected from the PKS with the KR deletion in module 2, suggesting that DEBS module 3 did not process the β-ketone triketide intermediate or that the product was formed at low levels. Production of 11-deoxy (**4**) and 5-deoxy (**10**) 6dEB analogs was achieved by replacing the existing KR in modules 2 or 5 with *rap*DH/ER/KR1. In addition, the C-5 ketone derivative, **8** (see above), also was produced with the *rap*DH/ER/KR1 replacement in module 5, suggesting that transfer of the unprocessed β-keto intermediate to module 6 occurs at rates competitive with ketoreduction by *rap*KR1 in module 5. The *rap*DH/ER/KR1 substitution in module 6 failed to generate a fully C-3 reduced compound, and the observed ketone (**12**) and alkene (**13**) products suggested that reductions catalyzed by the KR and ER domains are slow relative to lactone formation by the thioesterase.

An unexpected macrolide product also was observed from the PKS with the AT/ACP linker substitution in module 6, along with the expected product, **12** (Table 1). Purification and characterization by mass spectrometry and ^{1}H and ^{13}C-NMR spectroscopy revealed the structure to be 2-desmethyl-3-deoxy-3-oxo-6dEB (**14**), which arises from misincorporation of an acetate monomer in module 6. Although relaxed specificities of AT domains are known (17, 28), it is not obvious how non-AT domain replacements can affect the specificity of monomer addition.

Combined AT and KR Substitutions Within a Module. Next, substitution of both the AT and KR domains within a single module was performed in modules 2 and 6 to examine the tolerance for simultaneous alteration of extender unit and β-carbon processing within a single module. Six mutants

Table 1. Polyketides produced by AT and KR substitutions in DEBS modules 2, 5, and 6

Mutation	6dEB analog product	Compound no.	Yield*
Module 2			
*rap*AT2	10-desmethyl	**2**	0.2
AT/ACP linker	NP		
*rap*DH/KR4	10,11-anhydro	**3**	0.02
*rap*DH/ER/KR1	11-deoxy	**4**	0.2
Module 5			
*rap*AT2	4-desmethyl	**7**	0.04
AT/ACP linker	5-deoxy-5-oxo	**8**	0.1
*rap*DH/KR4	4,5-anhydro	**9**	ND
*rap*DH/ER/KR1	5-deoxy-5-oxo; 5-deoxy	**8, 10**	0.5, 0.04
Module 6			
*rap*AT2	2-desmethyl	**11**	0.7
AT/ACP linker	3-deoxy-3-oxo; 2-desmethyl-3-deoxy-3-oxo	**12, 14**	0.3, 0.4
*rap*DH/KR4	2,3-anhydro	**13**	0.4
*rap*DH/ER/KR1	3-deoxy-3-oxo; 2,3-anhydro	**12, 13**	0.3, 0.2
Module 2 (AT+KR)			
*rap*AT2+AT/ACP linker	NP		
*rap*AT2+*rap*DH/KR4	10-desmethyl-10,11-anhydro	**5**	<0.005
*rap*AT2+*rap*DH/ER/KR1	10-desmethyl-11-deoxy	**6**	<0.005
Module 6 (AT+KR)			
*rap*AT2+AT/ACP linker	2-desmethyl-3-deoxy-3-oxo	**14**	0.2
*rap*AT2+*rap*DH/KR4	2-desmethyl-3-(epi)	**15**	ND
*rap*AT2+*rap*DH/ER/KR1	2-desmethyl-3-deoxy-3-oxo	**14**	ND

*Relative to 6dEB (**1**) under similar conditions (≈20 mg/liter). ND, not determined; NP, no product.

Table 2. Combinatorial double and triple substitutions and polyketide products

Mutation			
Module 2	Module 5 or 6	6dEB analog product	Compound no.
Module 2 × module 5 double mutants			
*rap*AT2	AT/ACP linker	5-deoxy-5-oxo-10-desmethyl	**16**
*rap*AT2	*rap*DH/KR4	4,5-anhydro-10-desmethyl	**17**
*rap*AT2	*rap*DH/ER/KR1	5-deoxy-5-oxo-10-desmethyl; 5-deoxy-10-desmethyl	**16, 18**
*rap*DH/KR4	*rap*AT2	4-desmethyl-10,11-anhydro	**19**
*rap*DH/KR4	AT/ACP linker	5-deoxy-5-oxo-10,11-anhydro	**20**
*rap*DH/KR4	*rap*DH/ER/KR1	NP	
*rap*DH/ER/KR1	*rap*AT2	4-desmethyl-11-deoxy	**21**
*rap*DH/ER/KR1	AT/ACP linker	5,11-dideoxy-5-oxo	**22**
*rap*DH/ER/KR1	*rap*DH/KR4	4,5-anhydro-11-deoxy	**23**
Module 2 × module 6 double mutants			
*rao*AT2	*rap*AT2	2,10-didesmethyl	**24**
*rap*AT2	AT/ACP linker	3-deoxy-3-oxo-10-desmethyl	**25**
*rap*AT2	*rap*DH/KR4	2,3-anhydro-10-desmethyl	**26**
*rap*DH/KR4	*rap*AT2	2-desmethyl-10,11-anhydro	**27**
*rap*DH/KR4	AT/ACP linker	3-deoxy-3-oxo-10,11-anhydro	**28**
*rap*DH/ER/KR1	*rap*AT2	2-desmethyl-11-deoxy	**29**
*rap*DH/ER/KR1	AT/ACP linker	3-deoxy-3-oxo-11-deoxy	**30**
Module 2 × module 6 triple mutants			
*rap*AT14	*rap*AT2+AT/ACP linker	2,10-didesmethyl-3-deoxy-3-oxo	**31**
*rap*DH/KR4	*rap*AT2+AT/ACP linker	2-desmethyl-3-deoxy-3-oxo-10,11-anhydro	**32**
*rap*DH/ER/KR1	*rap*AT2+AT/ACP linker	2-desmethyl-3,11-dideoxy-3-oxo	**33**
Module 2 × module 5 × module 6 triple mutant			
KR2→ *rap*DH/ER/KR1	KR5→AT/ACP linker, AT6→*rap*AT2	2-desmethyl-5,11-dideoxy-5-oxo	**34**

Compound yields from all the multiple mutants fell to >0.1 mg/liter and could not be determined accurately by evaporative light scattering detection. The primary exception was compound **29**, which was producd at ≈0.2 mg/liter. NP, no product.

were constructed, with three producing the targeted doubly modified 6dEB analogs (Table 1). The absence of product from the PKS containing the *rap*AT2+AT/ACP linker double mutation in module 2 is consistent with the lack of product formation observed with the single AT/ACP linker substitution. The other two combinations in module 2, *rap*AT2+*rap*DH/KR4 and *rap*AT2+*rap*DH/ER/KR1, yielded small amounts of the expected 10-desmethyl-10,11-anhydro (**5**) and 10-desmethyl-11-deoxy (**6**) 6dEB derivatives. The PKS carrying the *rap*AT2+AT/ACP substitution in module 6 produced the anticipated 2-desmethyl-3-deoxy-3-oxo-6dEB (**14**) with identical HPLC retention time and mass fragmentation pattern as the compound unexpectedly formed by the PKS with the AT/ACP substitution alone (see above). Compound **14** was also the only product identified with the module 6 *rap*AT2+*rap*DH/ER/KR1 combination and is consistent with the slow rate of ketoreduction observed for the single *rap*DH/ER/KR1 substitution at this position. The *rap*AT2+*rap*DH/KR4 cassettes in module 6 produced a compound (**15**) with mass spectrum consistent with 2-desmethyl-6dEB (**11**), indicating that ketoreduction, but not dehydration, occurred. However, because the *rap*KR4 domain catalyzes ketoreduction with the opposite stereospecificity of *ery*KR6 (29) and because the HPLC retention time of this compound is different from **11**, we conclude that **15** is the C-3 OH epimer of **11**.

Combinatorial Double and Triple Substitutions in Different Modules. Substitutions in two separate modules next were engineered to manipulate biosynthetic steps more distant in the biosynthetic pathway. All productive single substitutions in module 2 were combined with all productive substitutions in module 5 or module 6, giving a total of 16 combinations (Table 2). Macrolide products were detected by LC/MS in the culture extracts from 15 of these mutants, although production levels decreased compared with parental single domain replacements. In each case, the mass spectrum was consistent with the compound(s) expected from the newly introduced catalytic activities (compounds **16–30**). The decline in polyketide titers by these combinatorial mutants probably reflects substrate preferences by downstream activities for distally altered regions of a biosynthetic intermediate.

Finally, triple domain substitutions were created to further test the catalytic pliancy of DEBS mutants. To optimize yields, we combined the most productive AT+KR double substitution in module 6 (*rap*AT2+AT/ACP linker) with functional AT or KR substitutions in module 2 (*rap*AT2, *rap*DH/KR, *rap*DH/ER/KR1) (Table 2). Analysis of the culture extracts indicated that these engineered DEBS produced compounds with mass spectra matching the expected 2,10-didesmethyl-3-deoxy-3-oxo (**31**), 2-desmethyl-3-deoxy-3-oxo-10,11-anhydro (**32**), and 2-desmethyl-3–11-dideoxy-3-oxo (**33**) 6dEB macrolactones. A fourth triple mutant also was engineered, this time manipulating a catalytic domain in each of three modules. The most productive single substitutions from module 2 (*rap*DH/ER/KR1), module 5 (AT/ACP linker), and module 6 (*rap*AT2) were combined in a single DEBS construct (Table 2). Again, a compound was formed with mass spectra matching the expected analog, 2-desmethyl-5,11-dideoxy-5-oxo-6dEB (**34**).

In addition to this series of combinatorial mutants, other substitutions have been used successfully to extend the number and diversity of compounds in the erythromycin library. These include replacement of the AT domains in module 1 (**35**) and module 3 (**36**) and of the DH/ER/KR domain in module 4 (**37**). Substitution of the KR in module 6 with the KR from RAPS module 2, which catalyzes reduction with opposite stereospecificity to the DEBS KR, results in the formation of a 6dEB analog with HPLC and mass properties consistent with an altered 3-hydroxyl stereochemistry (**38**). The remainder of the compounds in Fig. 2 represent combinatorial substitutions in module 2 and module 5 (**39**, **40**), module 2 and module 6 (**41–49**), module 3 and module 6 (**50**), and module 5 and module 6 (**51–61**). Finally,

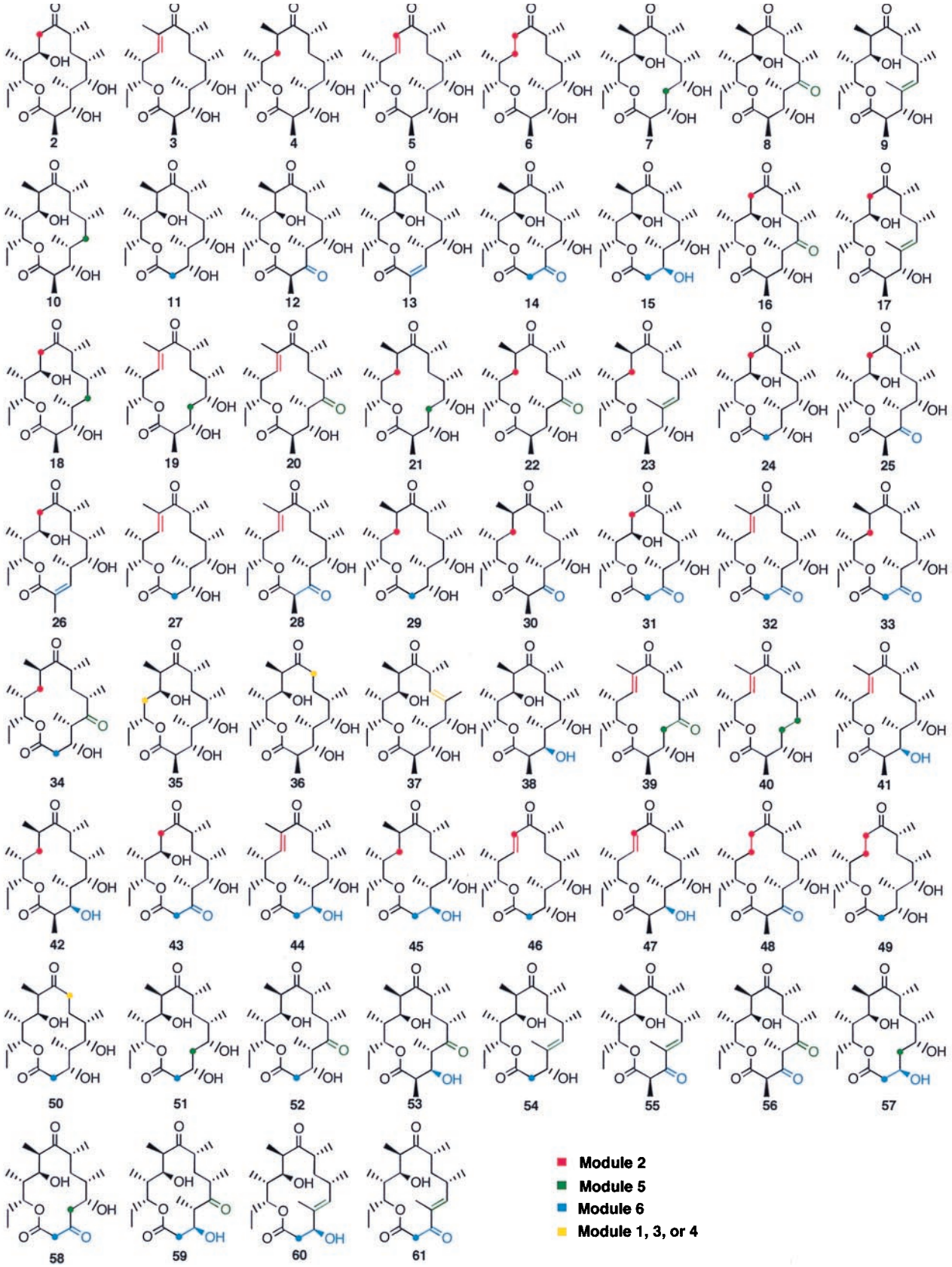

FIG. 2. DEBS combinatorial library. Colors indicate the location of the engineered carbon(s) resulting from catalytic domain substitutions in module 2 (red), module 5 (green), module 6 (blue), or modules 1, 3, or 4 (yellow).

many of the engineered DEBS reported here also produce detectable levels of one or more minor components such as acetate starter unit analogs of the major compounds (17, 24). Taking these into account, >100 macrolide products have been generated to date by using this simple combinatorial set of a 6-module scaffold and five cassettes. Many of these

6dEB analogs also have been processed further to erythromycin analogs by a DEBS-blocked mutant of *Sa. erythraea*.

DISCUSSION

Nature has exploited combinatorial biosynthesis to produce some 3,000 polyketides that are currently known‡, of which ≈150 are macrolide variants derived from ≈30 different 12-, 14- and 16-macrolactone ring structures (30). However, the natural polyketides thus far revealed represent only a small fraction of the combinatorial potential that might be realized from permutations of modules in a PKS. For example, if the two AT and five β-carbon modifier building blocks used here could be permutated into the 6-module DEBS PKS, the theoretical number of polyketides that would result is 10^7; complete permutation of the 14-module RAPS PKS with the same building blocks theoretically could yield a remarkable 10^{14} polyketides. It seems reasonable to expect that many interesting and important polyketides remain within the reservoir of yet-undiscovered molecules. The library we have created by engineering DEBS falls far short of what is theoretically possible. On the other hand, the number of polyketides we describe represents ≈3% of the total polyketides known and exceeds the total number of different macrolide ring structures yet discovered.

What is required to further approach the combinatorial potential of polyketide diversity? The experiments described demonstrate manipulation of the major combinatorial elements that can be used for engineering modular polyketide biosynthetic pathways—AT substitution, KR deletion, KR gain-of-function, and KR stereochemical alteration. Further, one or more of such modifications have been applied successfully to each of the six modules of DEBS, demonstrating a remarkable plasticity of the PKS toward foreign domains and intermediates. What is now required is to successfully apply as many of these modifications to as many modules as possible.

We have observed that, if two or more single PKS mutants are functional, it is likely that combinations of these also will produce the expected polyketide. Thus, we have invested our efforts in the stepwise approach of creating productive single mutants then combining two or more of them to prepare multiple mutants. By using the six module DEBS, the two ATs, and five β-carbon modifier components described here, there are <60 possible single mutants to be prepared. Once a modest library of productive multiple mutants has been prepared, the introduction of additional productive mutations in the library results in a multiplicative increase in the library size. For example, introduction of five new mutations into each of two virgin modules of the library of 50 mutants would produce a library of 1,150 polyketides, if all mutants were productive.

There is good reason to believe that, with appropriate efforts, many or most single PKS mutants could be prepared to produce the expected polyketides. However, as shown here, the polyketide titers of multiple mutants reflect some combination of the production losses encountered by the individual mutants. It follows that yield optimization of a relatively small number of singly altered PKSs by mutagenesis or by host alterations could substantially increase the number of polyketides produced by multiple mutants. Hence, the next achievement targeted for combinatorial biosynthesis is to increase the effectiveness of single substitutions in producing high titers of polyketides. We believe that, with achievement of this goal, generic approaches will become available for the production of numerous novel polyketides that will serve as a source of very large libraries of unnatural natural products.

‡This number was estimated by totaling the number of polyketides reported in the *Journal of Antibiotics* between 1947–1997, assuming this represents approximately half of all known polyketides.

Note Added in Proof. The construction of *S. lividans* K4–114 has now been described (31).

We thank Issac Chan-Kai and Sally Ou for technical assistance and Daniel Santi and Chaitan Khosla for helpful discussions.

1. Omura, S. (1992) *Gene* **115,** 141–149.
2. Zinner, S. H., Young, L. S., Acar, J. F. & Neu, H. C. (1997) *Expanding Indications for the New Macrolides, Azalides, and Streptogramins* (Dekker, New York).
3. Woodward, R. B., Logusch, E., Nambiar, K. P., Sakan, K., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., Chen, C. H., *et al.* (1981) *J. Am. Chem. Soc.* **103,** 3215–3217.
4. Chu, D. T. W. (1995) *Exp. Opin. Invest. Drugs* **4,** 65–94.
5. Cortés, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J. & Leadlay, P. F. (1990) *Nature (London)* **348,** 176–178.
6. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. (1991) *Science* **252,** 675–679.
7. McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1995) *Nature (London)* **375,** 549–554.
8. Katz, L. (1997) *Chem. Rev.* **97,** 2557–2576.
9. O'Hagan, D. (1991) *The Polyketide Metabolites* (Horwood, Chichester, U.K.).
10. Oliynyk, M., Brown, M. J. B., Cortes, J., Staunton, J. & Leadlay, P. F. (1996) *Chem. Biol.* **3,** 833–839.
11. Liu, L., Thamchaipenet, A., Fu, H., Betlach, M. & Ashley, G. (1997) *J. Am. Chem. Soc.* **119,** 10553–10554.
12. Ruan, X. R., Pereda, A., Stassi, D. L., Zeidner, D., Summers, R. G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M. J., Donadio, S., *et al.* (1997) *J. Bacteriol.* **179,** 6416–6425.
13. Schwecke, T., Aparicio, J. F., Molnar, I., Konig, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortes, J., Lester, J. B., *et al.* (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7839–7843.
14. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
15. McDaniel, R., Ebert-Khosla, S., Hopwood, D. & Khosla, C. (1993) *Science* **262,** 1546–1557.
16. Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M. & Schrempf, H. (1985) *Genetic Manipulation of Streptomyces: A Laboratory Manual* (The John Innes Foundation, Norwich, U.K.).
17. Kao, C. M., Katz, L. & Khosla, C. (1994) *Science* **265,** 509–512.
18. Donadio, S. & Katz, L. (1992) *Gene* **111,** 51–60.
19. Bedford, D., Jacobsen, J. R., Luo, G., Cane, D. E. & Khosla, C. (1996) *Chem. Biol.* **3,** 827–831.
20. McDaniel, R., Kao, C. M., Fu, H., Hevezi, P., Gustafsson, C., Betlach, M., Ashley, G., Cane, D. E. & Khosla, C. (1997) *J. Am. Chem. Soc.* **119,** 4309–4310.
21. Kao, C. M., McPherson, M., McDaniel, R., Fu, H., Cane, D. & Khosla, C. (1997) *J. Am. Chem. Soc.* **119,** 11339–11340.
22. Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. (1994) *J. Am. Chem. Soc.* **116,** 11612–11613.
23. Cortés, J., Wiesmann, K. E. H., Roberts, G. A., Brown, M. J. B., Staunton, J. & Leadlay, P. F. (1995) *Science* **268,** 1487–1489.
24. Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. (1995) *J. Am. Chem. Soc.* **117,** 9105–9106.
25. Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. (1996) *J. Am. Chem. Soc.* **118,** 9184–9185.
26. Donadio, S., McAlpine, J. B., Sheldon, P. J., Jackson, M. & Katz, L. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7119–7123.
27. Marsden, A. F. A., Wilkinson, B., Cortés, J., Dunster, N. J., Staunton, J. & Leadlay, P. F. (1998) *Science* **279,** 199–202.
28. Stassi, D. L., Kakavas, S. J., Reynolds, K. A., Gunawardana, G., Swanson, S., Zeidner, D., Jackson, M., Liu, H., Buko, A. & Katz, L. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 7305–7309.
29. Kao, C. M., McPherson, M., McDaniel, R., Fu, H., Cane, D. & Khosla, C. (1998) *J. Am. Chem. Soc.* **120,** 2478–2479.
30. Kirst, H. A. (1992) in *Kirk-Othmer Encyclopedia of Chemical Technology*, ed. Howe-Grant, M. (Wiley, New York), Vol. 3, pp. 169–213.
31. Zierman, R. & Betlach, M. C. (1999) *BioTechniques* **26,** 106–110.