Alteration of the Substrate Specificity of a Modular Polyketide Synthase Acyltransferase Domain through Site-Specific Mutations†

Christopher D. Reeves, Sumati Murli, Gary W. Ashley, Misty Piagentini, C. Richard Hutchinson, and Robert McDaniel*

Kosan Biosciences, Inc., 3832 Bay Center Place, Hayward, California 94545

Received October 17, 2001; Revised Manuscript Received November 14, 2001

ABSTRACT: Cassette replacement of acyltransferase (AT) domains in 6-deoxyerythronolide B synthase (DEBS) with heterologous AT domains with different substrate specificities usually yields the predicted polyketide analogues. As reported here, however, several AT replacements in module 4 of DEBS failed to produce detectable polyketide under standard conditions, suggesting that module 4 is sensitive to perturbation of the protein structure when the AT is replaced. Alignments between different modular polyketide synthase AT domains and the *Escherichia coli* fatty acid synthase transacylase crystal structure were used to select motifs within the AT domain of module 4 to re-engineer its substrate selectivity and minimize potential alterations to protein folding. Three distinct primary regions of AT4 believed to confer specificity for methylmalonyl-CoA were mutated into the sequence seen in malonyl-CoA-specific domains. Each individual mutation as well as the three in combination resulted in functional DEBSs that produced mixtures of the natural polyketide, 6-deoxyerythronolide B, and the desired novel analogue, 6-desmethyl-6-deoxyerythronolide B. Production of the latter compound indicates that the identified sequence motifs do contribute to AT specificity and that DEBS can process a polyketide chain incorporating a malonate unit at module 4. This is the first example in which the extender unit specificity of a PKS module has been altered by site-specific mutation and provides a useful alternate method for engineering AT specificity in the combinatorial biosynthesis of polyketides.

Polyketides are a structurally diverse class of natural products that have played an important role in the development of therapeutic and agricultural chemicals (*1*, *2*). Many polyketides possess antibiotic, anticancer, immunosuppressive, or other useful biological properties. In recent years, researchers have developed a number of protein engineering strategies for manipulating polyketide synthases (PKSs),[1] the enzyme complexes responsible for biosynthesis of polyketides, to generate novel chemical structures (*3*, *4*). Modular PKSs assemble the carbon backbones of polyketides through repeated condensations between acyl-CoA thioesters, typically, malonyl, methylmalonyl, and ethylmalonyl. Thus, the diversity observed in polyketide structures is in part derived from the incorporation of different starter or extender monomers, which is controlled by the PKS.

In modular PKS systems, such as the 6-deoxyerythronolide B synthase (DEBS, Figure 1A), selection of the extender unit for each module, as well as the starter unit in many cases, is carried out by an acyltransferase (AT) domain. An AT domain catalyzes the transacylation of the monomer unit from CoA to the phosphopantetheine arm of the acyl carrier protein (ACP) in the same module. AT domains generally possess a stringent specificity for a single acyl-CoA substrate in their natural context, although some ATs can incorporate at least two different monomers with similar efficiencies. Amino acid sequence alignments between methylmalonyl-CoA (mmCoA)-specific and malonyl-CoA (mCoA)-specific AT domains always cluster into two groups according to the specificity of the domain (*5*−*7*). At least three divergent sequence motifs have been identified on the basis of such alignments and comparison to the *Escherichia coli* malonyl-CoA:ACP transacylase (FabD) crystal structure (*8*−*10*). However, the contribution of these motifs to substrate specificity has not been previously tested experimentally. In a separate study, Lau et al. (*11*) used in vivo experiments to locate a different region at the C-terminal segment of AT domains which may influence specificity. Since this segment of the AT exhibits significant variability between AT domains, it is not clear how these amino acids contribute mechanistically to substrate selectivity.

Despite the identification of regions within ATs that correlate with specificity, the most popular method of changing substrate utilization in PKSs has been via the exchange of the entire ∼300−350-amino acid AT domain of one module with a homologous AT domain encoding a different starter or extender unit, obtained from another module, usually of a heterologous PKS. A number of

† This work was supported in part by SBIR Grant 1R43DK57380-01 from the National Institutes of Health.

* To whom correspondence should be addressed. Phone: (510) 732-8400. Fax: (510) 732-8401. E-mail: mcdaniel@kosan.com.

[1] Abbreviations: 6-dEB, 6-deoxyerythronolide B; 8,8a-dOle, 8,8a-deoxyoleandolide; ACP, acyl carrier protein; AT, acyltransferase; DEBS, 6-deoxyerythronolide B synthase; DH, dehydratase; emCoA, ethylmalonyl-coenzyme A; EpoPKS, epothilone polyketide synthase; ER, enoylreductase; FabD, *E. coli* fatty acid malonyl-CoA:ACP transacylase; KR, ketoreductase; KS, ketosynthase; mCoA, malonyl-coenzyme A; mmCoA, methylmalonyl-coenzyme A; PKS, polyketide synthase; TE, thioesterase.

10.1021/bi015864r CCC: $20.00
Published on Web 12/01/2001

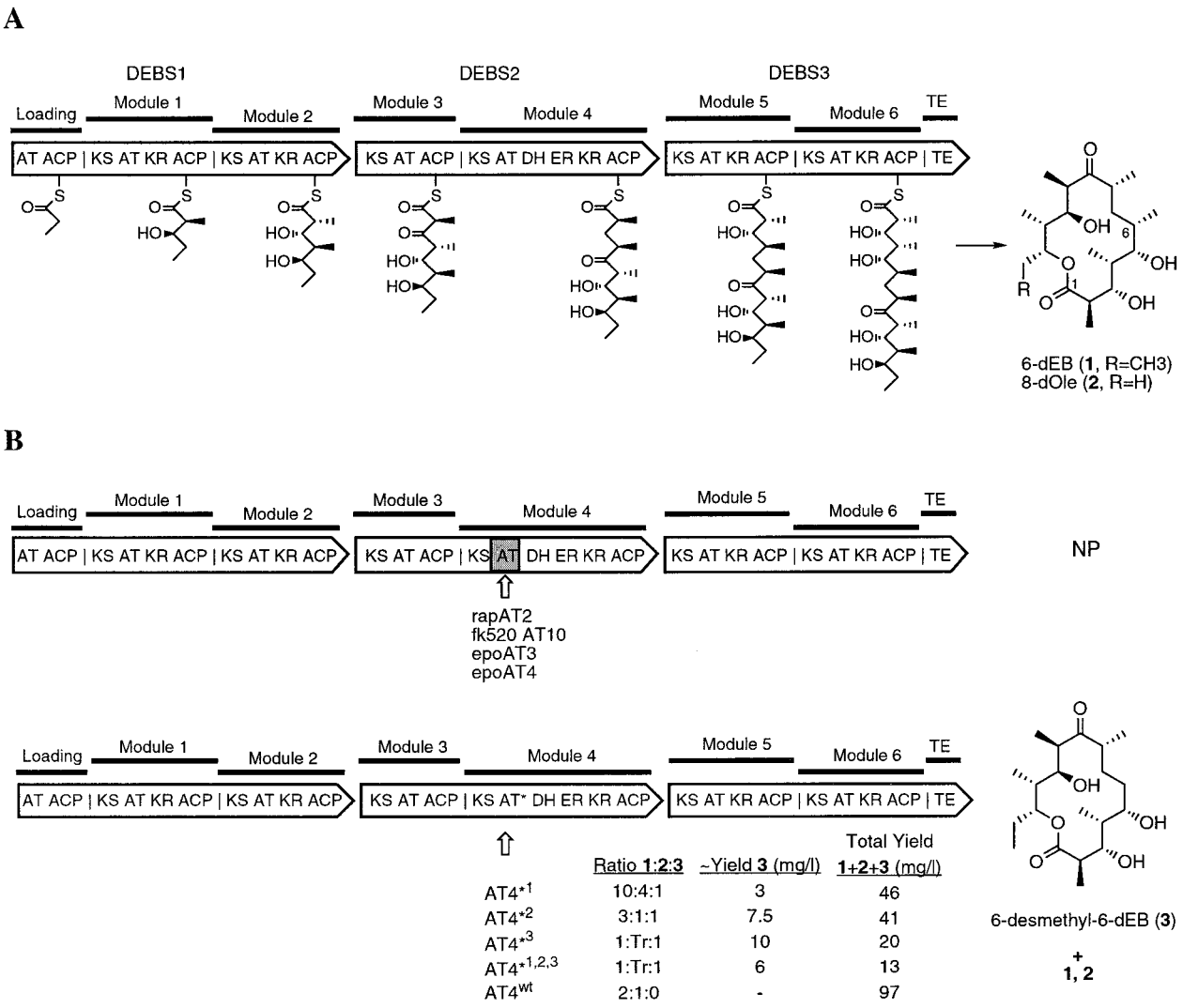


FIGURE 1: (A) Modular organization of DEBS and 6-dEB (**1**) biosynthesis. The erythromycin PKS, DEBS, uses propionyl-CoA as a loading unit and six methylmalonyl-CoA molecules for extender units. Each module in a modular PKS contains domains for one round of polyketide chain elongation (KS, AT, and ACP) and various degrees of β-keto modification (KR, DH, and ER). DEBS consists of six modules encoded on three separate polypeptide subunits. (B) DEBS module 4 mutants engineered in this study and polyketides produced. 6-Desmethyl-6-dEB (**3**) arises from incorporation of a malonate unit instead of a methylmalonate unit in module 4. See the text for a description of the mutations. Abbreviations: NP, no product observed; Tr, trace amounts; wt, wild type.

successful AT replacements with modular PKSs have been reported, in each case producing a new polyketide with the predicted structural change (*12*−*17*). However, occasionally AT replacements are only marginally successful or entirely unsuccessful, leading to only very small amounts of the desired compound or no product at all, for unknown reasons (*15*). For example, an AT cassette from module 2 of the rapamycin PKS was used successfully to alter the extender unit specificity from methylmalonate to malonate in modules 1−3, 5, and 6 of DEBS (*17*) but not module 4 (A. Thamchaipenet and R. McDaniel, unpublished observation).

If it is assumed that the genetic engineering of the PKS is properly executed, there are two possible reasons for such failures. Either the foreign AT domain causes the PKS complex to fold incorrectly and lose a necessary activity, or the replacement leads to a modified polyketide chain that is not recognized as a substrate by a subsequent ("downstream") PKS activity. In the former case, success might be achieved by changing AT domain specificity in a manner that minimized perturbation to the tertiary structure of the module, e.g., mutagenesis of a limited number of amino acids. To do so, however, requires an understanding of the rules for proper PKS folding and the structural features involved in substrate recognition.

Here we describe the alteration of the AT specificity of module 4 from DEBS by site-specific mutagenesis, utilizing the crystal structure of *E. coli* FabD (*9*) and amino acid alignments of different modular PKS AT domains to identify key residues predicted to contribute to the specificity of AT domains. All of the mutants permitted the incorporation of a malonate (as well as a methylmalonate) unit at the corresponding position of the polyketide, thereby producing 6-desmethyl-6-deoxyerythronolide B (**3**, 6-desmethyl-6-dEB). The data directly implicate specific amino acid residues as determinants of AT specificity. A series of cassette substitutions in module 4 using malonyl-CoA-specific AT domains from three different heterologous sources, however, did not produce any polyketide. Together, these results suggest that module 4 of DEBS is particularly sensitive to heterologous AT domain replacement and that the approach described here can be applied as an alternative

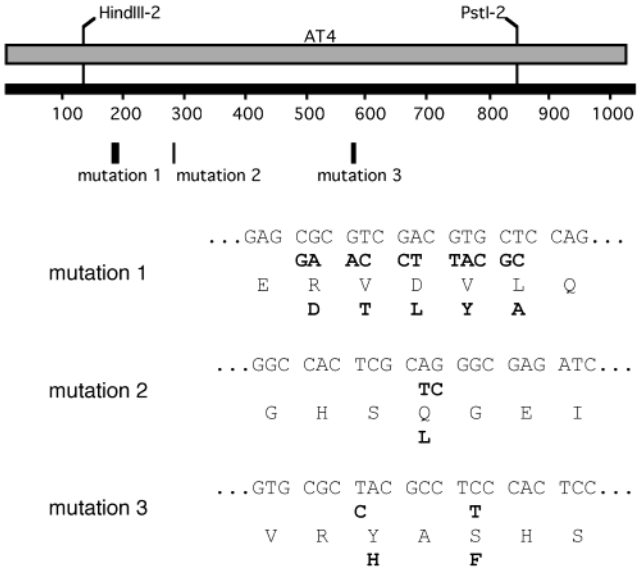


FIGURE 2: Site-directed mutations in DEBS AT4. The locations and sequences of mutations engineered in the AT domain of DEBS module 4 are shown. The scale refers to nucleotide position starting from the beginning of the conserved AT region. Also shown are the locations of the engineered restriction sites (*Hin*dIII and *Pst*I) used for substitution of the AT4 domain.

method for engineering substrate specificity of modular PKSs.

MATERIALS AND METHODS

Strains, Culture Conditions, and DNA Manipulation. All strains, culture conditions, and DNA manipulations were essentially as described previously (*18*) unless noted otherwise. The exception was that in the construction of the final *Streptomyces lividans* strains used for expression of polyketides, a second plasmid, pBOOST, was included. This plasmid, in conjunction with the pRM5-based expression plasmids used here, improves polyketide production by a mechanism that is currently under investigation (Z. Hu, personal communication).

Construction of AT Replacements in DEBS Module 4. Three different AT cassettes were used for substitution of the AT domain in module 4: fk520AT10, epoAT3, and epoAT4.[2] The cassettes were PCR amplified from DNA encoding the FK520 (*7*) and epothilone PKSs (*19*, *20*) using the following oligonucleotides (restriction sites underlined): fk520AT10, 5′-GCACAAGCTTACGGACGTCTACCACGGC (forward) and 5′-AGAACTGCAGTCTCCGCCGCGGCCGAC (reverse); epoAT3, 5′-CCGCAAGCTTCGCGAGGTGATGTGGGC (forward) and 5′-TCCGCTGCAGAAGCTGGCAACAGGCCGA (reverse); and epoAT4, 5′-CCGCAAGCTTTGCGAGGTGATGTGGGC (forward) and 5′-TCCGCTGCAGAAGCTGGCAACAGGCCGA (reverse). To replace DEBS AT4 with the three cassettes (as well as the AT4 site-specific mutants discussed below), a different set of restriction sites were engineered within the AT4 domain. A subclone was generated in which unique *Hin*dIII and *Pst*I sites were introduced into AT4 (Figure 2) using PCR amplification of segments ~1 kb upstream of the *Hin*dIII site, ~1 kb downstream of the *Pst*I site, and between the two sites. The three PCR products were assembled to make pKOS131-62A, which was verified by DNA sequencing. Both restriction sites are silent, and the altered sequences are as follows (restriction sites underlined, altered nucleotides in bold): 5′-CTCTCGCCGCACACCGACTGGAAG**C**T-**T**CTCGACGTCGTCCGCGGCGACGGC and 5′-CTGCTGATGGCGGTCGAGGAGAC**T**GC**A**GAGCACGCGGGCGCGGAAGTCACC.

The final expression plasmids were constructed as follows. The vector pKOS11-77 contains the DEBS genes with an engineered *Spe*I site downstream of the ACP of module 2 (*21*). This site appears to be functionally silent on the basis of production of 6-dEB by pKOS11-77 in *S. lividans* (*17*). The *Spe*I−*Bgl*II fragment containing the gene for DEBS2 was subcloned from pKOS11-77 into Litmus 28 (New England Biolabs), and the *Hin*dIII and *Pst*I sites were removed from the polylinker by cutting with *Bgl*II and *Afl*II and inserting a linker to make pKOS131-68. The *Bsi*WI−*Bam*HI fragment was subcloned from pKOS11-77 into Litmus 28 to make pKOS131-57. The *Apa*I−*Not*I fragment of pKOS131-62A was moved into pKOS131-57 to make pKOS131-72B, and the *Bsi*WI−*Fse*I fragment from pKOS131-72B was moved into pKOS131-68 to make pKOS131-72A. Each of the three heterologous AT domains (fk520AT10, epoAT3, and epoAT4) were moved as *Hin*dIII−*Pst*I fragments into pKOS131-72A. At this stage, the AT inserts of each subclone were confirmed by DNA sequencing. The *Spe*I−*Bgl*II fragment from each of the three subclones was moved into pKOS11-77 to obtain the final PKS expression vectors: pKOS164-35B (fk520AT10), pKOS164-35C (epoAT3), and pKOS164-35D (epoAT4).

Mutagenesis of DEBS AT4. Site-directed mutations were introduced in three regions of DEBS AT4 using the Altered-Sites mutagenesis kit (Promega). The *Hin*dIII−*Pst*I PCR product from AT4 was cloned into pALTER (Promega), and the mutations shown in Figure 2 were introduced following protocols provided by the manufacturer. All mutations were verified by DNA sequencing. Each of the mutant cassettes was introduced into the *Hin*dIII and *Pst*I sites of pKOS131-72A. Final expression plasmids were made analogous to the above AT substitutions. The resulting plasmids were pKOS215-79 (mutation 1), pKOS164-35F (mutation 2), pKOS164-35G (mutation 3), and pKOS215-35 (mutations 1−3).

Production and Analysis of Polyketides. Each of the expression plasmids was introduced into *S. lividans* K4-114 also containing pBOOST. Supernatants of strains grown in R5 liquid (with 50 mg/L thiostrepton and 200 mg/L apramycin) for 5 days were analyzed using an LC−MS system equipped with on-line extraction. The LC−MS system was comprised of a 10-port, two-position switching valve/injector, a Beckman System Gold HPLC apparatus, an Alltech ELSD detector, and a PE SCIEX API100 LC−MS-based detector equipped with an atmospheric pressure chemical ionization source. For detection of polyketides, 100−250 μL of clarified whole broth was loaded onto the guard column after a 1 min pre-equilibration with H_2O at 1 mL/min. At 30 s postinjection, a linear gradient to 15% MeCN over the course of 1 min was initiated. At 2 min, the direction of flow was reversed, and the eluent was diverted onto a Metachem Inertsil ODS-3 column (5 μm, 4.6 mm × 150 mm) pre-equilibrated with 15% MeCN. A linear gradient from

[2] The numbering used for epothilone PKS modules and domains is from the nomenclature used by Tang et al. (*19*).

15 to 100% MeCN at a rate of 1 mL/min over the course of 6 min, and then 100% MeCN for 3 min, was monitored by ELSD and MS. Under these conditions, **1** eluted at 10.0 min, **3** eluted at 9.7 min, and **2** eluted at 9.4 min. Authentic purified **1** and **2** were used as controls for compound verification. 6-dEB (**1**) was used to generate a calibration curve for polyketide titer measured by ELSD.

*Purification and Characterization of 6-Desmethyl-6-dEB (**3**). S. lividans* K4-114/pKOS164-35G with pBOOST was grown in 2 L of R5 as described above. The cells were removed by centrifugation, and the broth was filtered through a 2.5 cm × 30 cm column of XAD-16 adsorbent resin. The resin was washed with 2 volumes of water and then eluted with 4 volumes of acetone. The acetone eluate was evaporated on a rotary evaporator to a dark aqueous slurry, which was extracted three times with equal volumes of ethyl acetate. The organic extracts were combined, washed sequentially with saturated aqueous $NaHCO_3$ and brine, then dried over $MgSO_4$, filtered, and evaporated. The oily yellow residue was dissolved in 2 mL of CH_3CN, diluted with 2 mL of water, and filtered through a C_{18} Mega Bond-Elut (Varian) extraction column. The column was washed with 10 mL of a 1:1 CH_3CN/H_2O mixture, and the total eluate was evaporated. Final purification was by preparative HPLC, using a 20 mm × 50 mm C_{18} InertSil (Metachem) column and a gradient from 20 to 100% CH_3CN in water at a flow rate of 10 mL/min. Detection was by UV absorbance at 210 nm, with peak identity confirmed by mass spectrometric analysis. The appropriate HPLC fraction was evaporated, yielding 2 mg of 6-desmethyl-6-dEB (**3**): ^{1}H NMR ($CDCl_3$, 400 MHz) δ 5.154 (1H, ddd, J = 1.2, 4.0, 9.6 Hz, H-13), 4.038 (1H, br d, J = 10.4 Hz, H-5), 3.984 (1H, br d, J = 10.8 Hz, H-3), 3.779 (1H, dd, J = 0.8, 4.8 Hz, OH-11), 3.710 (1H, ddd, J = 2.4, 4.8, 10.4 Hz, H-11), 3.021 (1H, d, J = 2.4 Hz, OH-3), 2.803 (1H, dq, J = 6.8, 10.8 Hz, H-2), 2.698 (1H, br q, J = 6.8 Hz, H-10), 2.53 (1H, m, H-8), 2.186 (1H, br s, OH-5), 1.82 (unres, H-14a), 1.80 (unres, H-6a), 1.80 (unres, H-4), 1.74 (unres, H-7a), 1.73 (unres, H-12), 1.65 (1H, m, H-6b), 1.53 (1H, m, H-14b), 1.318 (3H, d, J = 6.8 Hz, Me-2), 1.260 (1H, m, H-7b), 1.051 (3H, d, J = 6.4 Hz, Me-8), 1.022 (3H, d, J = 6.8 Hz, Me-10), 1.011 (3H, d, J = 6.8 Hz, Me-4), 0.937 (3H, t, J = 7.4 Hz, H-15), 0.882 (3H, d, J = 6.8 Hz, Me-12); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 213.02 (C-9), 178.52 (C-1), 78.86 (C-3), 76.10 (C-13), 74.21 (C-5), 71.01 (C-11), 44.63 (C-2), 44.07 (C-10), 40.71 (C-12), 38.73 (C-8), 37.01 (C-4), 29.76 (C-6), 28.58 (C-7), 25.46 (C-14), 14.80 (Me-2), 12.28 (Me-8), 10.63 (C-15), 9.22 (Me-12), 6.05 (Me-10), 3.28 (Me-4). Unresolved ^{1}H resonances were assigned from two-dimensional correlation experiments.

RESULTS

Replacement of DEBS AT4 with Heterologous AT Domains Specific for Malonyl-CoA. All six of the extender AT domains in DEBS are specific for 2(*S*)-methylmalonyl-CoA (*22*). We recently described the production of five different 6-dEB analogues by replacing the AT domains from module 1, 2, 3, 5, or 6 of DEBS with the AT domain from module 2 of the rapamycin PKS (rapAT2) (*17*). Production titers of the corresponding 12-, 10-, 8-, 4-, and 2-desmethyl-6-dEB analogues ranged from ~4 to 70% of the amount of 6-dEB produced by unmodified DEBS. However, when this domain was used to replace the AT domain from module 4 of DEBS (AT4), no detectable amount of the desired polyketide, 6-desmethyl-6-dEB (**3**), was observed.

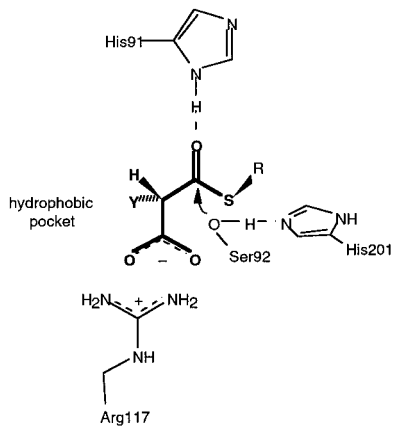
In an attempt to determine if the problem resulted from an incompatibility between the rapAT2 domain and DEBS module 4, three AT swaps were made using ATs from the FK520 PKS and the epothilone PKS (EpoPKS) (Figure 1B). Each of the chosen ATs (fk520AT10, epoAT3, and epoAT4) is more similar in sequence to DEBS AT4 than rapAT2. The first two AT domains incorporate only malonate extender units in their native PKSs, whereas the epoAT4 domain appears to be capable of incorporating either malonate or methylmalonate in epothilone biosynthesis (*19*). Therefore, a functional substitution with this domain should result in production of both **1** and **3** in the presence of both precursors. An alternative set of junctions (Figure 2) between the native and heterologous sequences was also used in these replacements to reduce the amount of heterologous sequence introduced, while still retaining those regions believed to be important for specificity. However, neither the expected 6-desmethyl compound nor 6-dEB could be detected when any of the PKSs were expressed in *S. lividans* (Figure 1B).

Expression of DEBS Containing Specific Mutations in AT4. Since the substitution of AT4 did not provide the desired 6-desmethyl analogue, we attempted more structurally conservative manipulations aimed at generating the compound. The *E. coli* FabD crystal structure (*9*) was used to locate regions in or near the active site whose amino acid sequences exhibited a strong correlation between mCoA and mmCoA specificity in AT alignments. Three such regions were chosen for mutagenesis of DEBS AT4, and in each case, the region was mutated from the canonical mmCoA-specific motif to the canonical mCoA-specific motif using site-directed mutagenesis (Figure 2). Mutation 1 consists of five amino acids immediately upstream of the highly conserved Gln residue in the active site (Gln-63 in FabD, Figure 3B). Mutation 2 is a single-amino acid substitution adjacent to the active site serine residue (Ser-92 in FabD, Figure 3B). Mutation 3 involves two amino acids adjacent to an invariant histidine residue (His-201, Figure 3B), the imidazole of which forms a hydrogen bond with the active site hydroxyl of Ser-92. All three regions have been identified as signature motifs used to discriminate between malonyl- and methylmalonyl-specific AT domains in sequence alignments (*8*, *10*), but there is no prior experimental evidence to support their role in substrate recognition.

Each of the mutations was engineered individually and in combination in DEBS, and product profiles were determined in *S. lividans* (Figure 1B). On the basis of LC−MS analysis, each PKS produced a combination of 6-dEB and two additional compounds with masses 14 amu smaller than that of 6-dEB. One of these compounds displayed the same retention and mass fragmentation as 8,8a-deoxyoleandolide (**2**, 8,8a-dOle). This compound is formed by priming of the DEBS loading domain with acetate rather than propionate and is usually observed as a minor component when DEBS is expressed in heterologous hosts (*23*). The other compound was presumed to be the desired 6-desmethyl analogue of 6-dEB and was purified for structural characterization.

^{1}H NMR and ^{13}C NMR spectroscopy established that this compound was 6-desmethyl-6-dEB (**3**). The ^{1}H NMR spectrum revealed only five methyl doublets, compared with six methyl doublets in the spectrum of 6-dEB (**1**). The ^{1}H−

A



B

E. coli FabD residue	#63	92	201
Malonyl			
avr (5)	QTPYAQ	GHSLGE	HAFH
srm (5)	RTEFAQ	GHSVGE	HGFH
nid (5)	RTEYTQ	GHSVGE	HAFH
epo (4)	QTAFTQ	GHSIGE	HAFH
sor (3)	QTAFTQ	GHSIGE	HAFH
520 (2)	DTLYAQ	GHSIGE	HAFH
rap (7)	ETGYAQ	GHSVGE	HAFH
pic (1)	ETRYTQ	GHSVGE	HAFH
epoAT4 (relaxed)	QTAFTQ	GHSAGE	HASH
Methylmalonyl			
avr (4)	RADVVQ	GHSQGE	YASH
ery (6)	RVDVVQ	GHSQGE	YASH
nid (1)	RVDVVQ	GHSQGE	YASH
epo (4)	RIDVVQ	GHSMGE	VASH
sor (2)	RVDVVQ	GHSQGE	YASH
520 (5)	RVEVVQ	GHSQGE	YASH
rap (7)	RVDVVQ	GHSQGE	YASH
pic (5)	RVDVVQ	GHSQGE	YASH
Ethylmalonyl			
srm (1)	RVDVVQ	GHSQGE	TAGH
nid (1)	RVDVVQ	GHSQGE	TAGH
tyl (1)	RVDVVQ	GHSQGE	TAGH
520 (1)	RVDVVH	GHSQGE	CPTH

FIGURE 3: (A) Hypothetical geometry of substrates in the active site of AT domains. The cleft containing the active site is oriented vertically, and the substrate bonds are bold to indicate the substrate lies in a plane slightly above the catalytic serine (Ser-92 in *E. coli* FabD). The rest of the CoA (or ACP) moiety rising up out of the active site cleft is represented by the letter R. The letter Y signifies an α-carbon substituent (i.e., hydrogen, methyl, or ethyl). (B) Alignment of divergent motifs in PKS AT domains. The divergent sequence blocks at or near the active site of PKS AT domains that correlate with malonyl, methylmalonyl, and ethylmalonyl specificity are shown. These correspond to the regions of DEBS AT4 mutated in this study (see Figure 2). Each sequence is a consensus sequence compiled by one or more AT domains from that particular PKS cluster, and the number in parentheses next to the PKS name indicates the number of AT domains that are represented. For reference, the number of the FabD residue is shown at the top of the alignment. The residue immediately following the invariant GHS motif protrudes into the hydrophobic pocket and is a branched-chain amino acid in all known malonyl-specific AT domains or a glutamine in most methylmalonyl-specific AT domains. The amino acids of region 1 lie on the left side of the catalytic cleft and may affect the size of the hydrophobic pocket. The amino acids of region 3 lie on the right side of the catalytic cleft and could affect the size and shape of the substrate binding pocket. Abbreviations: 520, FK520; avr, avermectin; epo, epothilone; nid, niddamycin; pic, picromycin; rap, rapamycin; sor, soraphen; srm, spiramycin; tyl, tylosin.

^{1}H COSY spectrum clearly revealed coupling between the H-5 resonance at δ 4.038 and the H-7b resonance at δ 1.28 to resonances at δ 1.66 and 1.80. ^{1}H−^{13}C HSQC confirmed that the δ 1.66 and 1.80 resonances arise from the assigned C-6 methylene at δ_C 29.76. ^{1}H−^{13}C HMBC data revealed long-range ^{1}H−^{13}C correlations completely consistent with the structure of **3**. A comparison of the ^{13}C NMR shifts between **1** and **3** reveals only four carbons with $\Delta\delta$ values of >1: C-5 ($\Delta\delta = -2.1$), C-6 ($\Delta\delta = -5.7$), C-7 ($\Delta\delta = -9.0$), and Me-4 ($\Delta\delta = -3.6$). The upfield shifts for C-5, C-6, and C-7 are again consistent with the assigned structure; the large upfield shifts at C-6 and C-7 upon loss of Me-6 reflect the difference in shifts between the corresponding carbons in methylcyclohexane and cyclohexane ($\Delta\delta$ −5.5 and −8.2 for C-1 and C-2, respectively) (*24*). The large upfield shift for Me-4 may reflect conformational changes in this region of the molecule arising from loss of interactions with Me-6.

Thus, in each of the PKSs, the selectivity of the AT domain has been altered to incorporate a malonate extender unit, although not exclusively. The average yield of 6-desmethyl-6-dEB (**3**) ranged from ∼3 to 10 mg/L among the mutants (Figure 1). Each single mutation resulted in a different product ratio, with AT4*^{3} resulting in the highest mCoA to mmCoA selectivity and AT4*^{1} resulting in the lowest (Figure 1). The triple mutant appeared to yield no discernible benefit over the AT4*^{3} mutation alone. The 6-desmethyl analogue of **2** could not be detected by LC−MS in any of the strains.

DISCUSSION

Effect of AT Substitutions in Module 4 of DEBS. Several examples of mmCoA to mCoA AT replacements have been previously reported which collectively include every module of DEBS except module 4 (*13*−*15*, *17*). Stassi et al. were able to make a functional AT swap in module 4 with an ethylmalonyl-CoA (emCoA)-specific AT which was capable of incorporating either a butyryl unit in the presence of emCoA or a propionate unit if no emCoA was available (*16*). However, a substantial decrease in polyketide titer accompanied the substitution (L. Katz, personal communication). Here, we have reported the substitution of DEBS AT4 with three different heterologous mCoA AT cassettes that did not produce any measurable polyketide. The reason for the apparent low-level production or nonproduction from these substitutions is not known. However, the subsequent synthesis of 6-desmethyl-6-dEB by the site-specific AT4 mutants in reasonable yield demonstrates that DEBS can process the unnatural pentaketide chain lacking an α-methyl group. These results together suggest that substitution of AT domains in module 4 easily alters the conformation of the module to disrupt efficient transfer of the acyl-CoA substrate to the ACP domain. This is evident with the epoAT4 cassette, which should not pose any substrate barriers to DEBS because of its presumed relaxed mCoA and mmCoA specificity. Furthermore, some previously reported AT substitutions have resulted in a "relaxation" of specificity which suggests that the rest of the module can affect the conformation of a heterologous AT domain (*16*, *25*). It is also consistent with the observation that alterations to a region at the C-terminus of AT domains, and located a considerable distance from the active site, can still affect

substrate specificity (*11*), possibly through steric forces that distort the AT domain.

Hypothetical Model for Acyl-CoA Binding in AT Domains. The *E. coli* FabD crystal structure (*9*) shows five water molecules in a plane just above and surrounding the active site serine (Ser-92). Five atoms of an acyl-CoA substrate (the thioester carbonyl oxygen and sulfur atoms, the two carboxylate oxygen atoms, and the α-carbon substituent atom) can be positioned in the active site such that they occupy nearly the same positions as these water molecules (Figure 3A). This configuration places the carboxyl group of the substrate in the proximity of a conserved arginine residue (Arg-117 in FabD) and is therefore consistent with biochemical experiments that have directly implicated this residue in carboxylate binding (*26*) (J. Kennedy, personal communication). Figure 3B also indicates the relative positions of the conserved glutamine (Gln-63), serine (Ser-92), and histidine (His-201) residues adjacent to the divergent motifs that were engineered in this study. In this model, the substituent on the α-carbon of an acyl-CoA substrate would point toward a hydrophobic pocket with a conserved Gln residue (equivalent to Gln-11 of *E. coli* FabD) lying immediately above it.

Identification of Amino Acid Residues That Alter AT Specificity. We have shown that three different motifs within the DEBS AT4 domain can be engineered to change the substrate selectivity from mCoA exclusively to allow both mCoA and mmCoA to be utilized. Since the intracellular concentrations of these two substrates are not known in *S. lividans*, it is not possible to calculate the relative specificities of the mutant AT domains based on the polyketide titers. We assume that the decline in overall polyketide titer relative to that of the wild-type enzyme results primarily from a reduction in the level of mmCoA incorporation at module 4 and a slower turnover for the intermediate which contains malonate. However, we cannot rule out changes in the level of enzyme expression as a result of the introduced mutations. These issues could be better addressed with in vitro systems such as those recently developed for individual modules of DEBS (*27*, *28*).

On the basis of the *E. coli* FabD structure, all three of these regions lie within the active site cleft of the AT domain. Prior to this study, these divergent motifs could be used to predict the specificity of a mCoA or mmCoA AT domain, but it was not known if these motifs were directly involved in substrate selectivity or were only an indication of specificity resulting from evolutionary constraints. Our results suggest these motifs are directly involved in substrate selectivity and that those amino acids in region 3 (Figure 3B) are the most critical residues. This is consistent with sequence alignments comparing mCoA-, mmCoA-, and emCoA-specific AT domains. For example, the consensus motifs for regions 1 and 2 are the same for both mmCoA- and emCoA-specific AT domains (Figure 3B), and therefore, how these amino acids contribute directly to binding of the two different substrates is not apparent. However, the consensus motifs for region 3 are different for mCoA (His-X-Phe-His), mmCoA ATs (Tyr-X-Ser-His), and emCoA ATs (Thr-Ala-Gly-His for several ATs) (Figure 3B). The size differences in side chains between the Phe, Ser, or Gly residues immediately adjacent to the conserved His residue (His-201) could possibly account for the differences in the binding of the corresponding acyl side chains of mCoA, mmCoA, and emCoA substrates. Furthermore, the amino acids of regions 1 and 2 of the epoAT4 domain are consistent with the mCoA motif, whereas the residues of region 3 suggest the origin of its relaxed mCoA and mmCoA specificity (Figure 3B).

It is surprising that all three regions could be altered individually to change the specificity of the domain, indicating that many different amino acids can contribute to substrate binding, perhaps either directly or indirectly. Therefore, it is likely that AT4 mutants with properties even better than those described here can be obtained. Furthermore, the identification of these critical residues for substrate utilization could facilitate the tailoring of AT specificity for non-natural extender units and might be used in conjunction with recent precursor engineering techniques (*29*) to greatly enhance the diversity of ketide units that can be used in combinatorial biosynthesis technologies.

ACKNOWLEDGMENT

We thank John Carney and Nina Viswanathan for assistance with LC−MS analysis and Ralph Reid for helpful scientific discussions.

REFERENCES

1. O'Hagan, D. (1991) *The polyketide metabolites*, Ellis Horwood, Chichester, U.K.
2. Hopwood, D. (1997) *Chem. Rev. 97*, 2465−2497.
3. McDaniel, R., and Khosla, C. (2001) in *Enzyme technologies for pharmaceutical and biotechnological applications* (Kirst, H. A., Yeh, W.-K., and Zmijewski, M. J., Jr., Eds.) pp 397−426, Marcel Dekker, Inc., New York.
4. Weissman, K. J., and Staunton, J. (2001) in *Enzyme technologies for pharmaceutical and biotechnological applications* (Kirst, H. A., Yeh, W.-K., and Zmijewski, M. J., Jr., Eds.) pp 427−470, Marcel Dekker, Inc., New York.
5. Schwecke, T., Aparicio, J. F., Molnar, I., Konig, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortes, J., Lester, J. B., Böhm, G. A., Staunton, J., and Leadlay, P. F. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7839−7843.
6. Kakavas, S. J., Katz, L., and Stassi, D. (1997) *J. Bacteriol. 179*, 7515−7522.
7. Wu, K., Chung, L., Revill, P. W., Katz, L., and Reeves, C. D. (2000) *Gene 251*, 81−90.
8. Haydock, S. F., Aparicio, J. F., Molnár, I., Schwecke, T., Khaw, L. E., König, A., Marsden, A. F. A., Galloway, I. S., Stuanton, J., and Leadlay, P. F. (1995) *FEBS Lett. 374*, 246−248.
9. Serre, L., Verbree, E. C., Dauter, Z., Stuitje, A. R., and Derewenda, Z. S. (1995) *J. Biol. Chem. 270*, 12961−12964.
10. Ikeda, H., Nonomiya, T., Usami, M., Ohta, T., and Omura, S. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 9509−9514.
11. Lau, J., Fu, H., Cane, D. E., and Khosla, C. (1999) *Biochemistry 38*, 1643−1651.
12. Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W., and Rao, R. N. (1996) *Gene 183*, 231−236.
13. Oliynyk, M., Brown, M. J. B., Cortes, J., Staunton, J., and Leadlay, P. F. (1996) *Chem. Biol. 3*, 833−839.
14. Liu, L., Thamchaipenet, A., Fu, H., Betlach, M., and Ashley, G. (1997) *J. Am. Chem. Soc. 119*, 10553−10554.
15. Ruan, X. R., Pereda, A., Stassi, D. L., Zeidner, D., Summers, R. G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M. J., Donadio, S., and Katz, L. (1997) *J. Bacteriol. 179*, 6416−6425.
16. Stassi, D. L., Kakavas, S. J., Reynolds, K. A., Gunawardana, G., Swanson, S., Zeidner, D., Jackson, M., Liu, H., Buko, A., and Katz, L. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 7305−7309.

17. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., Betlach, M., and Ashley, G. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 1846−1851.
18. Tang, L., Fu, H., and McDaniel, R. (2000) *Chem. Biol. 7*, 77−84.
19. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B. (2000) *Science 287*, 640−642.
20. Molnár, I., Schupp, T., Ono, M., Zirkle, R. E., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratman, A., Cyr, D. D., Gorlach, J., Mayo, J. M., Hu, A., Goff, S., Schmid, J., and Ligon, J. M. (2000) *Chem. Biol. 7*, 97−109.
21. Kao, C. M., Luo, G., Katz, L., Cane, D. E., and Khosla, C. (1995) *J. Am. Chem. Soc. 117*, 9105−9106.
22. Marsden, A. F. A., Caffrey, P., Aparicio, J. F., Loughran, M. S., Staunton, J., and Leadlay, P. F. (1994) *Science 263*, 378−380.
23. Kao, C. M., Katz, L., and Khosla, C. (1994) *Science 265*, 509−512.
24. Breitmaier, E., and Voelter, W. (1987) in *Carbon-13 NMR Spectroscopy*, pp 186−187, VCH, New York.
25. Xue, Q., Ashley, G., Hutchinson, C. R., and Santi, D. V. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 11740−11745.
26. Rangan, V. S., and Smith, S. (1997) *J. Biol. Chem. 272*, 11975−11978.
27. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) *Science 284*, 482−485.
28. Wu, N., Kudo, F., Cane, D. E., and Khosla, C. (2000) *J. Am. Chem. Soc. 122*, 4847−4852.
29. Pohl, N. L., Hans, M., Lee, H. Y., Kim, Y. S., Cane, D. E., and Khosla, C. (2001) *J. Am. Chem. Soc. 123*, 5822−5823.

BI015864R