Dissecting the Role of Acyltransferase Domains of Modular Polyketide Synthases in the Choice and Stereochemical Fate of Extender Units†

Janice Lau,‡ Hong Fu,§ David E. Cane,*,|| and Chaitan Khosla*,‡,⊥,#

Departments of Chemical Engineering, Chemistry, and Biochemistry, Stanford University, Stanford, California 94305-5025, KOSAN Biosciences, Inc., 1450 Rollins Road, Burlingame, California 94110, and Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912

Received August 21, 1998; Revised Manuscript Received October 22, 1998

ABSTRACT: Modular polyketide synthases (PKSs), such as the 6-deoxyerythronolide B synthase (DEBS), are large multifunctional enzyme complexes that are organized into modules, where each module carries the domains needed to catalyze the condensation of an extender unit onto a growing polyketide chain. Each module also dictates the stereochemistry of the chiral centers introduced into the backbone during the chain elongation process. Here we used domain mutagenesis to investigate the role of the acyl transferase (AT) domains of individual modules in the choice and stereochemical fate of extender units. Our results indicate that the AT domains of DEBS do not influence epimerization of the (2*S*)-methylmalonyl-CoA extender units. Hence, stereochemical control of the methyl-branched centers generated by DEBS most likely resides in the ketosynthase (KS) domains of the individual modules. In contrast, several recent studies have demonstrated that extender unit specificity can be altered by AT domain substitution. In some of these examples, the resulting polyketide was produced at considerably lower titers than the corresponding natural product. We analyzed one such attenuated mutant of DEBS, in which the methylmalonyl transferase domain of module 2 was replaced with a malonyl transferase domain. As reported earlier, the resulting PKS produced only small quantities of the expected desmethyl analogue of 6-deoxyerythronolide B. However, when the same hybrid module was placed as the terminal module in a truncated 2-module PKS, it produced nearly normal quantities of the expected desmethyl triketide lactone. These results illustrate the limits to modularity of these multifunctional enzymes. To dissect the role of specific amino acids in controlling AT substrate specificity, we exchanged several segments of amino acids between selected malonyl and methylmalonyl transferases, and found that a short (23−35 amino acid) C-terminal segment present in all AT domains is the principal determinant of their substrate specificity. Interestingly, its length and amino acid sequence vary considerably among the known AT domains. We therefore suggest that the choice of extender units by the PKS modules is influenced by a "hypervariable region", which could be manipulated via combinatorial mutagenesis to generate novel AT domains possessing relaxed or altered substrate specificity.

The modular polyketide synthase 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea* is an assembly of three large multifunctional proteins—DEBS1, DEBS2, and DEBS3—that catalyzes the biosynthesis of the erythromycin macrolactone 6-deoxyerythronolide B (6-dEB) (**1**) (*1*, *2*). Each of the three proteins is organized into two modules, and each module carries the catalytic domains needed to incorporate one of six methylmalonyl-CoA extender units onto a growing polypropionate chain (Figure 1) (*1*, *2*). The stereochemistry of the methyl-branched centers arising from the individual condensation reactions is also determined by the polyketide synthase as each successive C_3 unit is added. In 6-dEB, the three chiral carbons at the C-2, C-4, and C-10 positions all have the D-methyl configuration, whereas those at the C-6, C-8, and C-12 positions have the L-methyl configuration. Given the considerable interest in the genetic and chemical reprogramming of modular polyketide synthases (PKSs) (*3*−*5*), a better understanding of the molecular basis for the choice and stereochemical fate of the extender units may lead to novel ways of generating structurally diverse polyketides.

Incorporation of deuterated propionate into erythromycin A demonstrated that formation of the D-methyl centers at C-2, C-4, and C-10 of the polyketide chain results from decarboxylative inversion of the (2*S*)-methylmalonyl-CoA substrates (*6*). More recently, based on stereospecific acylation both of intact DEBS proteins and of selected partial proteolytic fragments, Leadlay, Staunton, and their collaborators showed that (2*S*)-methylmalonyl-CoA, and not its (2*R*)-epimer, is the sole substrate of all six AT domains of DEBS (*7*). By following the formation of triketide lactone (**4**), catalyzed by DEBS1+TE [a truncated PKS containing the

† This research was supported by grants from the National Institutes of Health (CA-66736 to C.K. and GM-22172 to D.E.C.), by a National Science Foundation Young Investigator Award (to C.K.), and by a David and Lucile Packard Fellowship for Science and Engineering (to C.K.). J.L. is a recipient of an NSF Predoctoral Fellowship.

* Corresponding authors.
‡ Department of Chemical Engineering, Stanford University.
§ KOSAN Biosciences, Inc.
|| Department of Chemistry, Brown University.
⊥ Department of Chemistry, Stanford University.
Department of Biochemistry, Stanford University.

10.1021/bi9820311 CCC: $18.00
Published on Web 01/13/1999

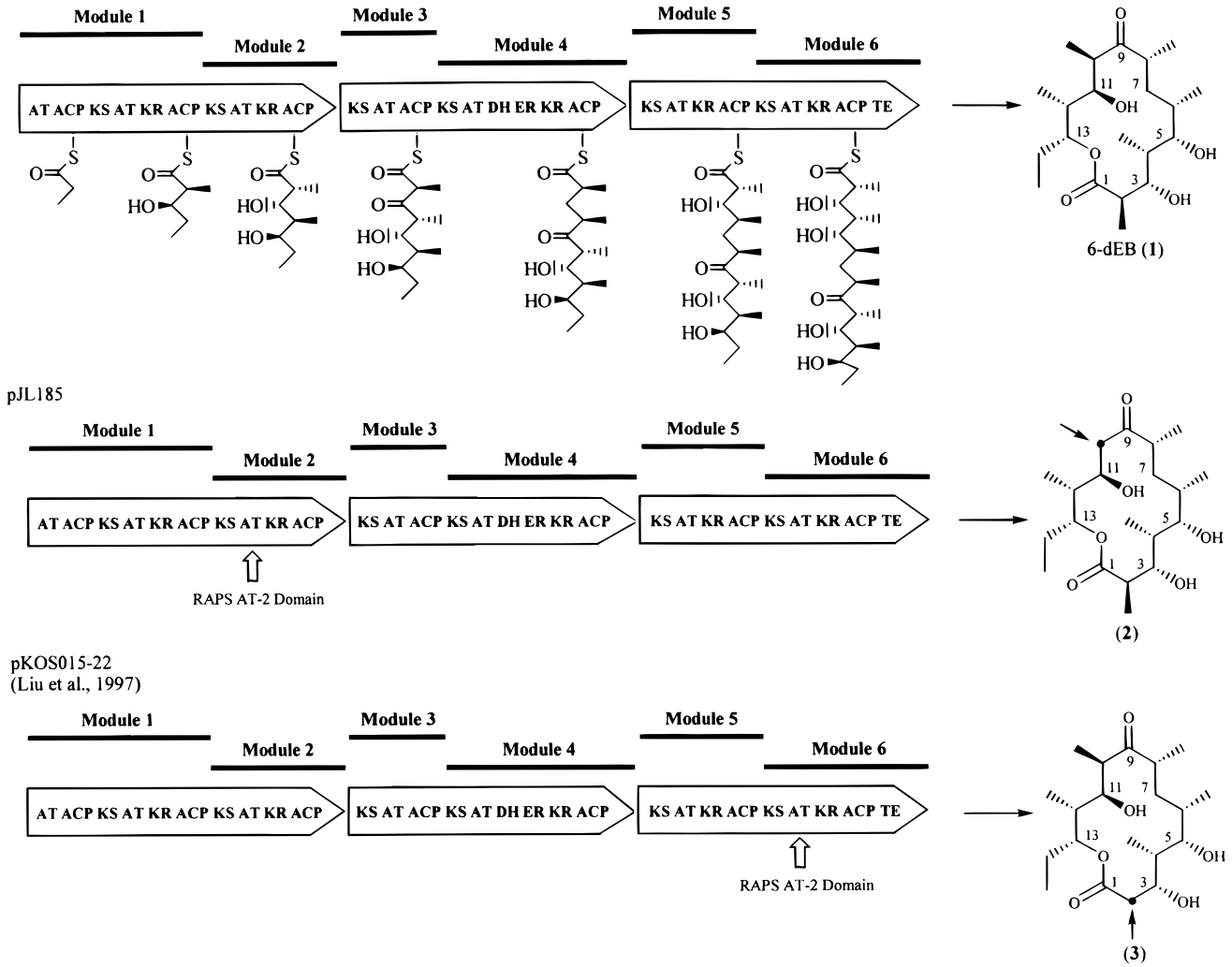

FIGURE 1: Modular organization of the 6-deoxyerythronolide B synthase (DEBS), which catalyzes the biosynthesis of 6-deoxyerythronolide B (6-dEB) (**1**). DEBS consists of three large multifunctional proteins (DEBS1, DEBS2, and DEBS3) that each possess two modules. Each module catalyzes one cycle of chain extension and associated β-ketoreduction for the biosynthesis of 6-dEB. The active sites denote acyltransferases (AT), acyl carrier proteins (ACP), β-ketoacyl-ACP transferases, β-ketoreductases, a dehydratase (DH), an enoylreductase (ER), and a thioesterase (TE). Substitution of AT2 with the RAPS AT2 domain (pJL185) leads to the production of 10-desmethyl-6-dEB (**2**). Substitution of AT6 with the RAPS AT2 domain (pKOS015-52) leads to the production of 2-desmethyl-6-dEB (**3**).

first two modules of DEBS fused to the thioesterase domain (TE) derived from module 6], in the presence of (2*S*)-[2-^{2}H]-methylmalonyl-CoA, they also observed the loss of deuterium from the (2*S*)-methylmalonyl-CoA substrate at the C-4 position of the triketide lactone, which carries the L-methyl center generated by module 1 (corresponding to C-12 in 6-dEB) (*8*). Together, these findings suggest that generation of the L-methyl centers may involve an epimerization step that is catalyzed by DEBS itself. Because construction of the polyketide chain is processive, this epimerization step would have to take place either in the methylmalonyl moiety prior to the condensation step or in the polyketide product immediately following the ketosynthase reaction. Additional evidence bearing on the timing of the epimerization reaction comes from the recent finding that the stereochemistry of β-ketoreduction is an intrinsic property of the relevant ketoreductase (KR) domains and is independent of both the stereochemistry and the degree of substitution adjacent to the β-keto group that undergoes reduction (*9*).

If epimerization were indeed to occur before condensation, it is conceivable that the relevant acyltransferase domains (AT) of DEBS might epimerize the enzyme-bound (2*S*)-methylmalonyl group to the requisite (2*R*)-methylmalonyl diastereomer prior to the ketosynthase-catalyzed condensations that generate the β-ketoacyl polyketide intermediates with L-methyl groups in the α-position. To test this hypothesis, we engineered two chimeric derivatives of DEBS1+TE, in which the AT domains of module 1 (AT1) and of module 2 (AT2) were independently replaced by the AT domain of module 4 (AT4) (Figure 2). Heterologous expression of DEBS1+TE in *Streptomyces coelicolor* CH999 normally results in formation of the triketide (2*R*,3*S*,4*S*,5*R*)-2,4-dimethyl-3,5-dihydroxy-*n*-heptanoic acid δ-lactone (**4**) (*10*, *11*). We now report that each of the chimeric proteins, in which either AT1 or AT2 was replaced by AT4, produces the triketide lactone (**4**) with unaltered stereochemistry.

Although the mechanism for controlling the stereochemistry of the methyl-branched centers remains unclear, it is well-established that the extender-unit specificity toward malonyl- or methylmalonyl-CoA is exclusively controlled by the AT domains of the individual modules (*12*−*14*). In strains harboring some of these recombinant PKSs, however, modified polyketides were produced at considerably lower titers than erythromycin itself. We explore some of the basis

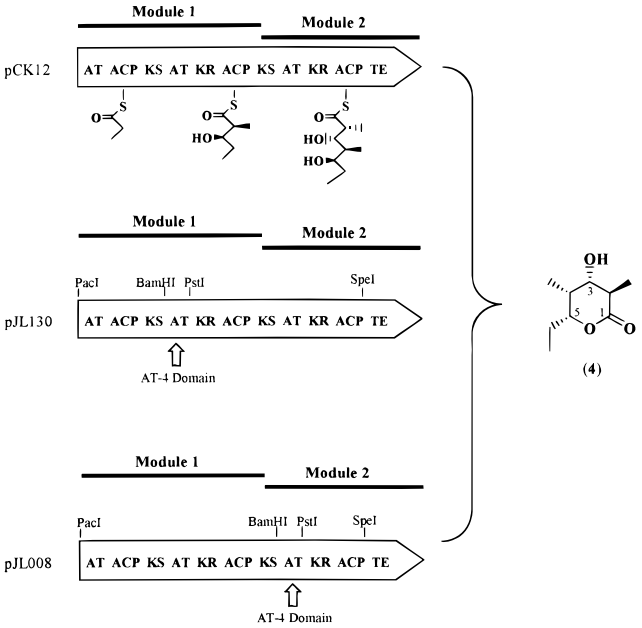

FIGURE 2: Modular organization of DEBS1+TE, which catalyzes the synthesis of triketide lactone (**4**). DEBS1+TE consists of DEBS1 fused to the thioesterase (TE) domain from DEBS3. Substitution of AT1 or AT2 with the AT4 domain (pJL130 and pJL008) leads to the same triketide product generated by the parent PKS construct.

for these differences in production levels. Moreover, to identify the structural determinants of substrate specificity in the AT domains, hybrid AT domains were constructed by replacing fragments of a malonyl-CoA-specific AT domain with homologous segments from a methylmalonyl-CoA-specific AT domain, and vice versa. The results of these studies are also described here.

MATERIALS AND METHODS

Strains and Culture Conditions. *Streptomyces coelicolor* CH999 (*15*) was used as the host for production of polyketides from the engineered expression plasmids. *S. coelicolor* strains were grown on R2YE agar plates (*16*). DNA manipulations were performed in *Escherichia coli* XL1 Blue (Stratagene) using standard culture conditions (*17*).

Manipulation of DNA and Strains. Manipulation and transformation of DNA in *E. coli* were performed by standard procedures (*17*). Polymerase chain reaction (PCR) was performed with Pfu polymerase (Stratagene) under conditions recommended by the manufacturer. *S. coelicolor* protoplasts were transformed by standard procedures (*16*), and transformants were selected using 1.5 mL of a 500 μg/mL thiostepton (Sigma) overlay.

Construction of Plasmids. Plasmids pJL130 and pJL008 are derivatives of pCK12, which harbors the DEBS1+TE gene in the *Streptomyces* expression vector (*11*). In plasmids pJL130 and pJL008, a 1.0-kb *Bam*HI−*Pst*I fragment corresponding to either the AT1 domain or the AT2 domain was replaced by a 1.0-kb *Bam*HI−*Pst*I fragment corresponding to the AT4 domain of DEBS, respectively (Figure 2). These plasmids were constructed as follows. The AT4 cassette was amplified by PCR with flanking *Bam*HI and *Pst*I restriction sites (nucleotides 6052−7097; GenBank accession no. M63677). This cassette was introduced independently into two different plasmids, pKOS011-16 and pKOS001-11,

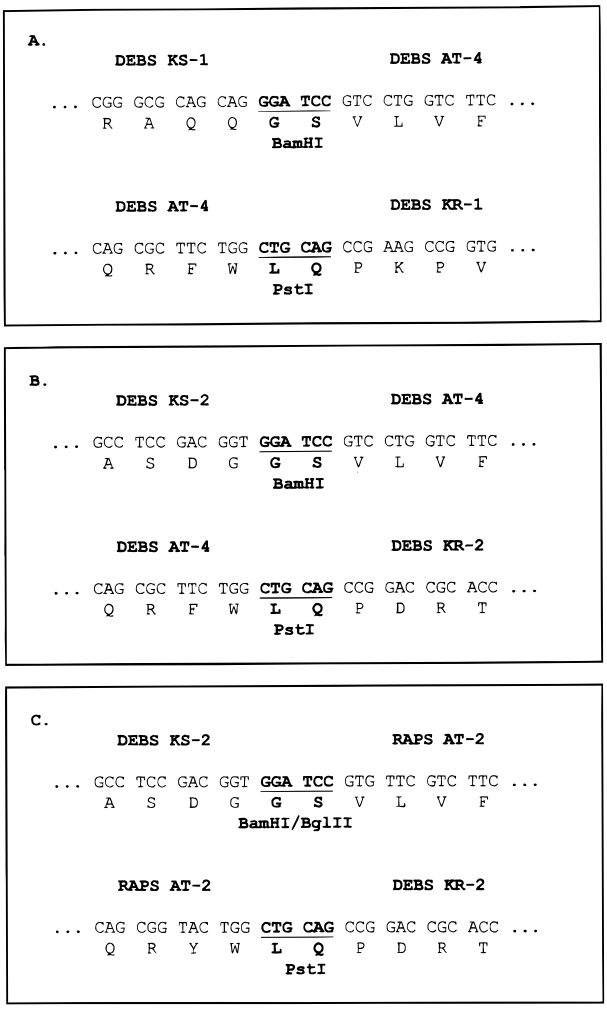

FIGURE 3: Nucleotide sequence at the boundaries of the AT substitutions. (A) Substitution of AT1 with AT4 (pJL130). (B) Substitution of AT2 with AT4 (pJL008). (C) Substitution of AT2 with RAPS AT2 (pKOS008-51 or pJL185).

which both contain the DEBS1+TE gene (*11*) in a derivative of the pUC19 vector. Plasmid pKOS011-27 was engineered by PCR to have *Bam*HI and *Pst*I restriction sites flanking the AT1 domain (nucleotides 3828−4891; GenBank accession no. M63676), while plasmid pKOS001-11 was engineered to have the same restriction sites flanking the AT2 domain (nucleotides 8259−9307; GenBank accession no. M63676). The sequences of the engineered boundaries are shown in Figure 3. The resulting 10-kb fragments encoding the modified DEBS1 genes were subsequently digested with *Pac*I and *Spe*I and ligated into *Pac*I/*Spe*I-digested pCK12 (*11*) to produce the expression plasmids pJL130 and pJL008.

Plasmids pKOS008-51 and pJL185 are derivatives of pCK12 and pCK7 (*11*, *18*), respectively, in which the 1.0-kb *Bam*HI−*Pst*I fragment corresponding to the entire AT2 domain of DEBS is replaced with a 0.9-kb *Bgl*II−*Pst*I fragment corresponding to the intact AT2 domain of the rapamycin polyketide synthase (RAPS) (Figures 1 and 4). These plasmids were constructed as follows. The RAPS AT2 cassette was PCR-amplified with flanking *Bgl*II and *Pst*I restriction sites (complement of nucleotides 54631−55556; GenBank accession no. X86780) and subcloned into pCR-Script (SK+) (Stratagene) to generate pKOS008-50. Because of the presence of a second natural *Pst*I restriction site

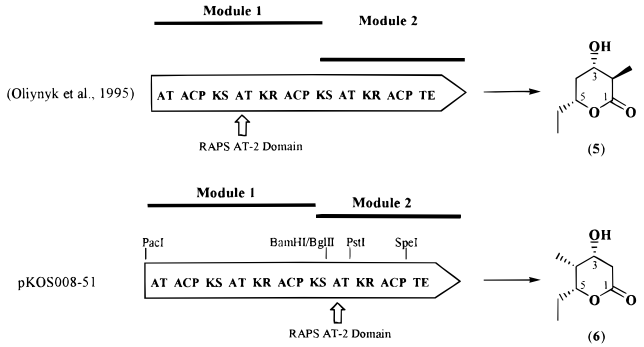

FIGURE 4: Bimodular DEBS/RAPS hybrid fusions and triketide products. Substitution of AT1 with the RAPS AT2 domain leads to the production of (**5**), while substitution of AT2 with the RAPS AT2 domain (pKOS008-51) leads to the production of (**6**).

(nucleotide 55084) in the middle of the RAPS AT2 domain, pKOS008-50 was digested completely with *Bgl*II but only partially with *Pst*I. The 925-bp fragment corresponding to the intact RAPS AT2 domain was then isolated and ligated into *Bam*HI/*Pst*I-digested pKOS001-11. The sequences of the engineered boundaries are shown in Figure 3. The resulting 10-kb fragment encoding the chimeric DEBS1 gene with the heterologous AT2 domain was digested with *Pac*I and *Spe*I and ligated into *Pac*I/*Spe*I-digested pCK12 (*11*) and pKOS011-77 to produce the expression plasmids pKOS008-51 and pJL185, respectively. Plasmid pKOS011-77 is a modified version of pCK7 (*18*), in which a *Spe*I restriction site was engineered by PCR at the same nucleotide position as pCK12 (*11*) in DEBS1 (nucleotide 10953; GenBank accession no. 63676).

Plasmid pJL286 is a derivative of pJL185, in which a 124-bp *Xho*I−*Sst*I fragment of the RAPS AT2 domain encoding the sequence motif that is suggested to be responsible for malonyl-CoA specificity is excised and replaced with an analogous fragment from the DEBS AT2 domain encoding the corresponding sequence motif suggested for methylmalonyl-CoA specificity (*19*). This plasmid is constructed as follows. A PCR-amplified *Bgl*II−*Xho*I fragment from the RAPS AT2 domain (complement of nucleotides 55430−55556) was ligated to a PCR-amplified *Xho*I−*Sst*I fragment extending from nucleotides 8403−8548 of the DEBS AT2 domain. This 271-bp *Bgl*II−*Sst*I fragment was used to replace the native *Bgl*II−*Sst*I fragment of the RAPS AT2 domain (complement of nucleotides 43929−55556) in pKOS008-50. The engineered boundaries for this hybrid AT domain are shown in Figures 5 and 6. The resulting plasmid was completely digested with *Bgl*II and partially digested with *Pst*I. The 946-bp fragment corresponding to the hybrid AT domain was isolated and ligated into *Bam*HI/*Pst*I-digested pKOS001-11. The 10-kb fragment encoding the chimeric DEBS1 gene was digested with *Pac*I and *Spe*I and ligated into *Pac*I/*Spe*I-digested pKOS011-77 to produce the expression plasmid pJL286.

Plasmid pJL259 is a derivative of pJL185, in which a 198-bp *Pin*AI−*Pst*I fragment encoding the C-terminal region of the RAPS AT2 domain is excised and replaced with its homologous counterpart from the DEBS AT2 domain. This plasmid is constructed as follows. Plasmid pKOS008-50 was digested with *Bgl*II and *Pin*AI. The 727-bp *Bgl*II−*Pin*AI fragment from the RAPS AT2 domain (complement of nucleotides 54829−55556) was isolated and ligated to a PCR-amplified *Pin*AI−*Pst*I fragment extending from nucleotides 9082 to 9307 of the DEBS AT2 domain. The engineered boundaries for this hybrid AT domain are shown in Figures 5 and 6. The resulting hybrid AT cassette was subcloned into a derivative of pCR-Script (SK+). This plasmid was completely digested with *Bgl*II and partially digested with *Pst*I. The 952-bp fragment corresponding to the hybrid AT domain was isolated and ligated into *Bam*HI/*Pst*I-digested pKOS001-11. The 10-kb fragment encoding the chimeric DEBS1 gene was digested with *Pac*I and *Spe*I and ligated into *Pac*I/*Spe*I-digested pKOS011-77 to produce the expression plasmid pJL259.

Plasmid pJL285 is a derivative of pJL185, in which a 68-bp *Nhe*I−*Sty*I fragment corresponding to a subregion of the RAPS *Pin*AI−*Pst*I fragment is replaced with its counterpart from the DEBS AT2 domain. This plasmid is constructed as follows. A PCR-amplified *Nhe*I−*Sty*I fragment extending from nucleotides 9089−9193 of the DEBS AT2 domain was ligated to a PCR-amplified *Sty*I−*Pst*I fragment from the RAPS AT2 domain (complement of nucleotides 54631−54754). This 227-bp *Nhe*I−*Pst*I fragment was introduced into a subclone containing a PCR-amplified *Bgl*II−*Nhe*I fragment from the RAPS AT2 domain (complement of nucleotides 54822−55556). The engineered boundaries for this hybrid AT domain are shown in Figures 5 and 6. The resulting plasmid was completely digested with *Bgl*II and partially digested with *Pst*I. The 961-bp fragment corresponding to the hybrid AT domain was isolated and ligated into *Bam*HI/*Pst*I-digested pKOS001-11. The 10-kb fragment encoding the chimeric DEBS1 gene was digested with *Pac*I and *Spe*I and ligated into *Pac*I/*Spe*I-digested pKOS011-77 to produce the expression plasmid pJL285.

Plasmid pJL287 is a derivative of pJL185, in which a 123-bp *Sty*I−*Pst*I fragment corresponding to a second subregion of the RAPS *Pin*AI−*Pst*I fragment is replaced with its counterpart from the DEBS AT2 domain. This plasmid is constructed as follows. A PCR-amplified *Sty*I−*Pst*I fragment extending from nucleotides 9193−9307 of the DEBS AT2 domain was introduced into a subclone containing a PCR-amplified *Bgl*II−*Sty*I fragment from the RAPS AT2 domain (complement of nucleotides 54754−55556). The engineered boundaries for this hybrid AT domain are shown in Figures 5 and 6. The resulting plasmid was completely digested with *Bgl*II and partially digested with *Pst*I. The 916-bp fragment corresponding to the hybrid AT domain was isolated and ligated into *Bam*HI/*Pst*I-digested pKOS001-11. The 10-kb fragment encoding the chimeric DEBS1 gene was digested with *Pac*I and *Spe*I and ligated into *Pac*I/*Spe*I-digested pKOS011-77 to produce the expression plasmid pJL287.

Plasmid pJL305 is a derivative of pCK7 (*18*), in which an 89-bp *Nhe*I−*Sty*I fragment encoding a subregion of the DEBS AT6 domain that is homologous to the one encoded by the *Nhe*I−*Sty*I fragment of the DEBS AT2 domain is replaced with the 68-bp *Nhe*I−*Sty*I fragment of the RAPS AT2 domain. This plasmid is constructed as follows. A PCR-amplified *Nhe*I−*Sty*I fragment from the RAPS AT2 domain (complement of nucleotides 54754−54822) was ligated to a PCR-amplified *Sty*I−*Pst*I fragment extending from nucleotides 17678−17792 of the DEBS AT6 domain (GenBank accession no. M63677). This 182-bp *Nhe*I−*Pst*I fragment was introduced into a subclone containing a PCR-amplified *Bam*HI−*Nhe*I fragment extending from nucleotides 16765−

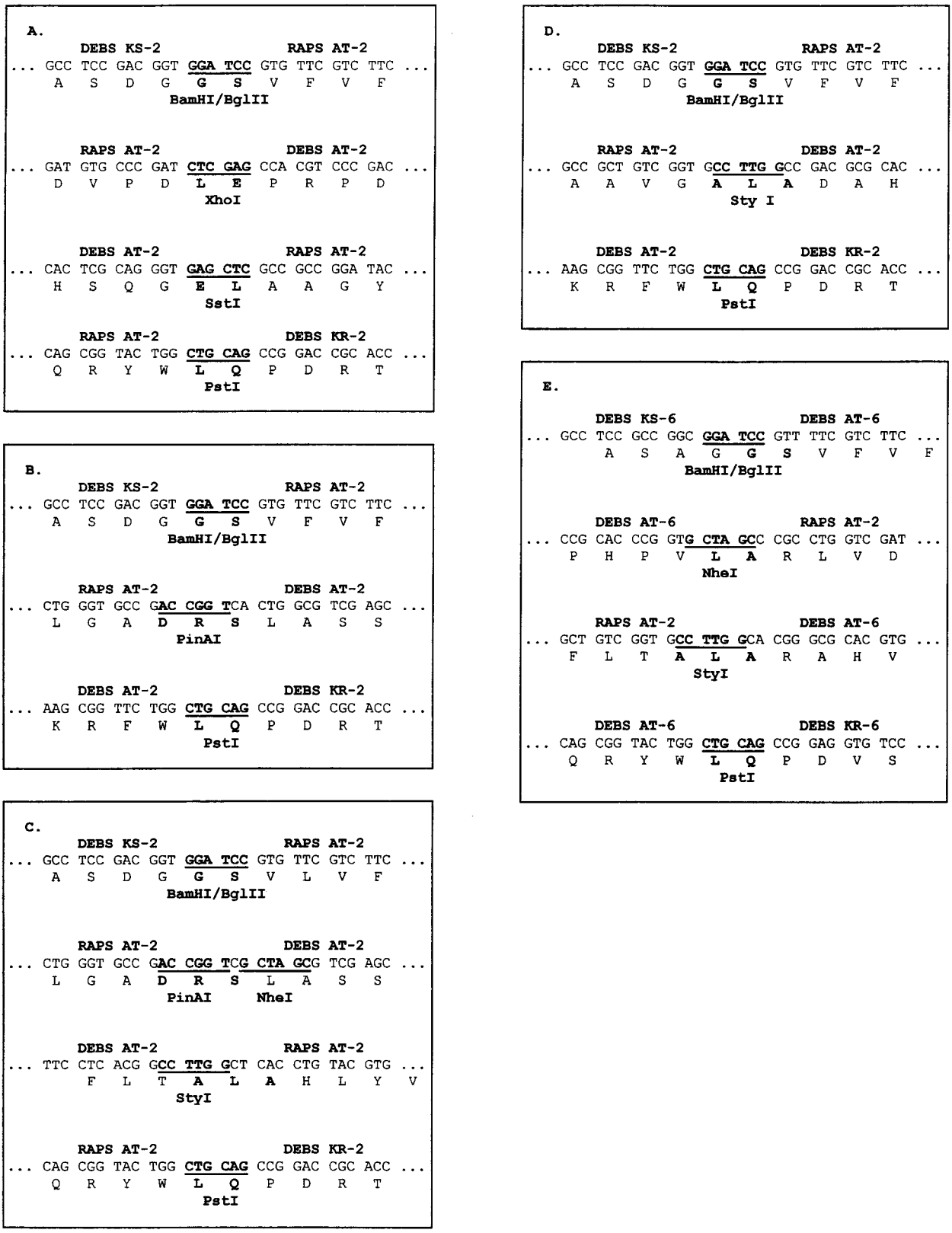

FIGURE 5: Nucleotide sequence at the boundaries of the DEBS/RAPS hybrid AT2 and AT6 domains. (A) Hybrid AT2 domain used in the construction of pJL286. (B) Hybrid AT2 domain used in the construction of pJL259. (C) Hybrid AT2 domain used in the construction of pJL285. (D) Hybrid AT2 domain used in the construction of pJL287. (E) Hybrid AT6 domain used in the construction of pJL305.

17589 of the DEBS AT6 domain. The engineered boundaries for this hybrid AT domain are shown in Figures 5 and 6. The resulting plasmid was digested with *Bam*HI and *Pst*I, and the 1.0-kb fragment corresponding to the hybrid AT domain was isolated and ligated into *Bam*HI/*Pst*I-digested pKOS015-52. Plasmid pKOS015-52 is a derivative of pUC19 containing the DEBS3 gene. It was engineered by PCR to have *Bam*HI and *Pst*I restriction sites flanking the AT6 domain (nucleotides 16765−17792; GenBank accession no. M63677). The 9-kb fragment encoding the chimeric DEBS3 gene was digested with *Bgl*II (nucleotide 11087) and *Eco*RI (nucleotide 20226) and ligated into *Bgl*II/*Eco*RI-digested

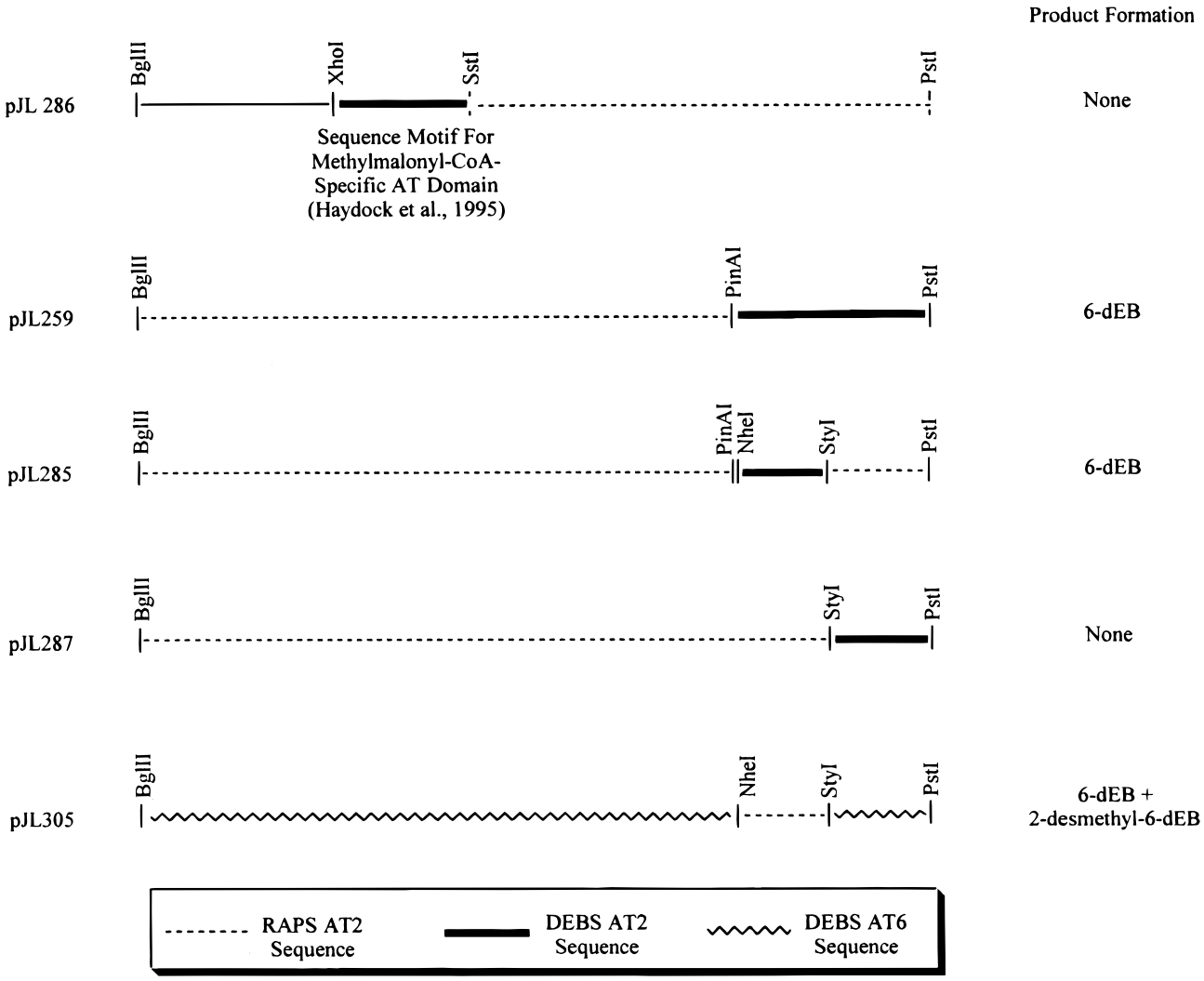

FIGURE 6: Diagrams of the DEBS/RAPS hybrid AT2 and AT6 domains in pJL286, pJL259, pJL285, pJL287, and pJL305.

pKOS011-77 to produce the expression plasmid pJL305.

Purification and Characterization of Polyketides. Each of the recombinant strains was grown on 200 mL of R2YE agar containing 50 μg/mL thiostrepton and 600 mg/L propionic acid for 6 days. The medium was extracted twice with 2 volumes of ethyl acetate. The organic phase, which contained the desired polyketides, was dried in a SpeedVac. AP−CI mass spectroscopy was performed on the crude extracts.

To isolate compounds **1**, **3**, **4**, and **6** for nuclear magnetic resonance (NMR) structural studies, the crude extracts were redissolved in ethyl acetate for spotting on thin-layer chromatography (TLC) plates (Baker Si250F Silica Gel). Migration was allowed for approximately 1.5 h, using ethyl acetate−hexane (60:40). The polyketides were visualized with a vanillan−sulfuric acid−ethanol (3:1:96) spray. After TLC separation, the regions of interest were scraped from the plate and extracted twice with an equal volume of ethyl acetate. The organic phase was dried in a SpeedVac, and the isolated polyketides were redissolved in deuterated chloroform for NMR analysis.

RESULTS

Replacement of AT1 or AT2 of DEBS1+TE with AT4. To investigate the role of the AT domains in epimerizing the extender units prior to incorporation into the polyketide backbone, the recombinant strains CH999/pJL130 and CH999/pJL008 were generated by transforming *S. coelicolor* CH999 with the appropriate plasmids in which either AT1 or AT2 of the DEBS1+TE gene was replaced with the AT4 domain of DEBS, respectively. These resultant strains both synthesized a triketide (**4**) identical to that produced by the parent DEBS1+TE strain (*11*) (Figure 2). The stereochemistry of the products was established by ^{1}H NMR spectroscopy and comparison of the data with those for authentic **4**.

Replacement of DEBS AT2 with RAPS AT2 in a 2-Module or 6-Module Scaffold. To test the ability of the heterologous AT2 domain of RAPS to recruit a malonyl-CoA extender unit for incorporation into the polyketide backbone in module 2 of DEBS, the recombinant strain CH999/pKOS008-51 was transformed with a plasmid in which the methylmalonyl-CoA-specific AT2 domain of the bimodular DEBS1+TE gene was replaced with the malonyl-CoA-specific RAPS AT2 domain. This mutant led to the production of 2-desmethyl triketide lactone (**6**) (Figure 4). This novel metabolite was produced at ∼10 mg/L, which is half that of the production levels of the parent triketide lactone (**4**) (∼20 mg/L). The structure of this product was confirmed by ^{1}H and ^{13}C NMR spectroscopy (data not shown).

Similarly, the recombinant strain CH999/pJL185 harbors a plasmid in which the AT2 domain of the 6-module DEBS

system was replaced with the RAPS AT2 domain. The DNA sequence encoding module 2 in pJL185 is identical to that present in pKOS008-51. This mutant led to the production of a metabolite thought to be 10-desmethyl 6-dEB (**2**) (Figure 1). However, this putative polyketide was generated in amounts insufficient for exhaustive spectroscopic characterization (<1 mg/L); its existence in the fermentation broth was suggested by AP−CI mass spectroscopy ($M + H^+ = 373$, $M - H_2O + H^+ = 355$, $M - 2H_2O + H^+ = 337$). Nevertheless, our result is consistent with that of Katz and co-workers (*13*), who produced the corresponding 10-desmethyl analogue of erythromycin by similar AT swaps into DEBS in the native erythromycin producer *S. erythraea*. Although the exact titers were not reported, these scientists reported the isolation of only a few milligrams of the product from a 27-L fermentation, suggesting that the titer of this novel erythromycin in *S. erythraea* was considerably lower than that of the parent natural product in the same host.

Replacement of DEBS AT2 or DEBS AT6 in the 6-Module System with a DEBS/RAPS Hybrid AT Domain. To determine the regions of the AT domains that dictate their substrate specificity, the recombinant strains CH999/pJL286, CH999/pJL259, CH999/pJL285, and CH999/pJL287 were each transformed with a plasmid in which the AT2 domain of the full DEBS system was replaced with a hybrid DEBS/RAPS AT domain as depicted in Figure 6. Of the four mutants, CH999/pJL286 and CH999/pJL287 did not lead to the production of any detectable polyketides. However, CH999/pJL259 and CH999/pJL285 both produced approximately 10 mg/L of the 6-dEB macrolactone (**1**). The structure of this product was again confirmed by ^{1}H NMR spectroscopy. No 10-desmethyl-6-dEB (**2**) could be detected in the extracts of these recombinant strains.

Likewise, the recombinant strain CH999/pJL305 carries a plasmid in which the AT6 domain of the 6-module DEBS system was replaced with a chimeric DEBS/RAPS domain as shown in Figure 6. This mutant led to the biosynthesis of both 6-dEB (**1**) and the 2-desmethyl 6-dEB compound (**3**) in comparable amounts (∼5 mg/L). The structure of **2** was also confirmed by ^{1}H NMR spectroscopy.

DISCUSSION

Role of Acyltransferase Domains in Controlling the Stereochemistry of Methyl-Branched Centers in Polyketides. Although (2*S*)-methylmalonyl-CoA serves as the sole substrate for the AT domains of DEBS, the expected D-methyl center, generated by decarboxylative inversion of the extender units, is found only in the growing polyketide chains produced by modules 2, 5, and 6. By contrast, the epimeric L-methyl configuration is observed at C-12 and C-8, positions in the erythromycin polyketide generated by modules 1 and 3, respectively. Since the L-methyl configuration at the α-position of the acyl chain generated by module 4 is certainly due to the reduction of its corresponding α-methyl-α,*β*-unsaturated ester intermediate, it is unknown whether module 4 contains epimerase activity or not. Notwithstanding this ambiguity, the possible ability of the AT4 domain to epimerize its bound (2*S*)-methylmalonyl substrate can still be evaluated by alternately replacing the AT1 or AT2 domain of the DEBS1+TE mutant with the AT4 domain. If the AT4 domain were to have no intrinsic epimerase activity, then both of the derived hybrid PKSs would produce the natural (2*R*,3*S*,4*S*,5*R*)-2,4-dimethyl-3,5-dihydroxy-*n*-heptanoic acid *δ*-lactone (**4**). On the other hand, if the donor AT4 domain or one or the other of the replaced AT1 and AT2 domains were able to mediate epimerization, then one of the two hybrid PKSs would be expected to produce a diastereomer of **4**.

From our studies, the observed interchangeability of DEBS AT4 with either the AT1 or the AT2 domain in formation of **4** suggests not only that the donor AT domain retains its catalytic activity upon introduction into another module, but also that there is no instrinsic epimerase activity associated with any of these three AT domains. We have also recently reported that replacement of the DEBS KR2 domain with either the KR2 or the KR4 domain of the rapamycin PKS results in formation of the 3-*epi*-triketide lactone, thereby establishing that the configuration of the polyketide hydroxyl groups is controlled directly by the operative KR domain and is not correlated with the stereochemistry or degree of substitution at the adjacent α-position (*9*). We have also shown that replacement of the DEBS ACP3 domain with the DEBS ACP6 domain in a truncated trimodular PKS does not result in alteration of stereochemistry of the relevant methyl-branched centers (*20*). Thus, it can be inferred that neither the AT, the KR, nor the ACP domain is responsible for the epimerization that results in generation of the L-methyl-branched centers derived from (2*S*)-methylmalonyl-CoA. By a process of elimination, the critical epimerase activity is therefore likely to be associated directly with the individual KS domains, a hypothesis that is amenable to direct experimental verification through construction of the relevant hybrid PKSs.

Processing of Modified Polyketide Intermediates in Modular PKSs. Although the AT domains of DEBS do not appear to epimerize their (2*S*)-methylmalonyl-CoA substrate, these domains exercise a strict specificity for methylmalonyl-CoA over malonyl-CoA, with a strict specificity for the (2*S*)-enantiomer of the former substrate. Significantly, Leadlay, Staunton, and co-workers have shown that replacement of the AT1 domain of DEBS1+TE with the RAPS AT2 domain results in biosynthesis of the predicted 4-desmethyl triketide lactone (**5**) (*12*). Similarly, replacement of the individual AT1, AT2, or AT6 domain of DEBS with heterologous AT domains that are specific for malonyl-CoA leads to production of the predicted desmethyl 6-dEB or erythromycin derivative (*13*, *14*).

However, the yields of the products generated by these chimeras vary greatly depending on the position of the polyketide chain at which the modified extender unit is incorporated. Katz and co-workers have demonstrated that production of the 10-desmethylerythromycin derivatives, resulting from replacement of the DEBS AT2 domain with a heterologous AT domain specific for malonyl-CoA, is significantly lower than that of the 12-desmethylerythromycin derivatives, which results from replacement of the DEBS AT1 domain with a malonyl-CoA-specific AT domain (*13*). This suggests that intermediates, which are altered at different positions of the polyketide backbone, may be recognized or processed to varying degrees by the modules downstream of the one responsible for the modification. Recently, Liu and co-workers have also observed that replacement of the final AT domain of DEBS (AT6) with the malonyl-CoA-specific RAPS AT2 domain leads to sizable production of

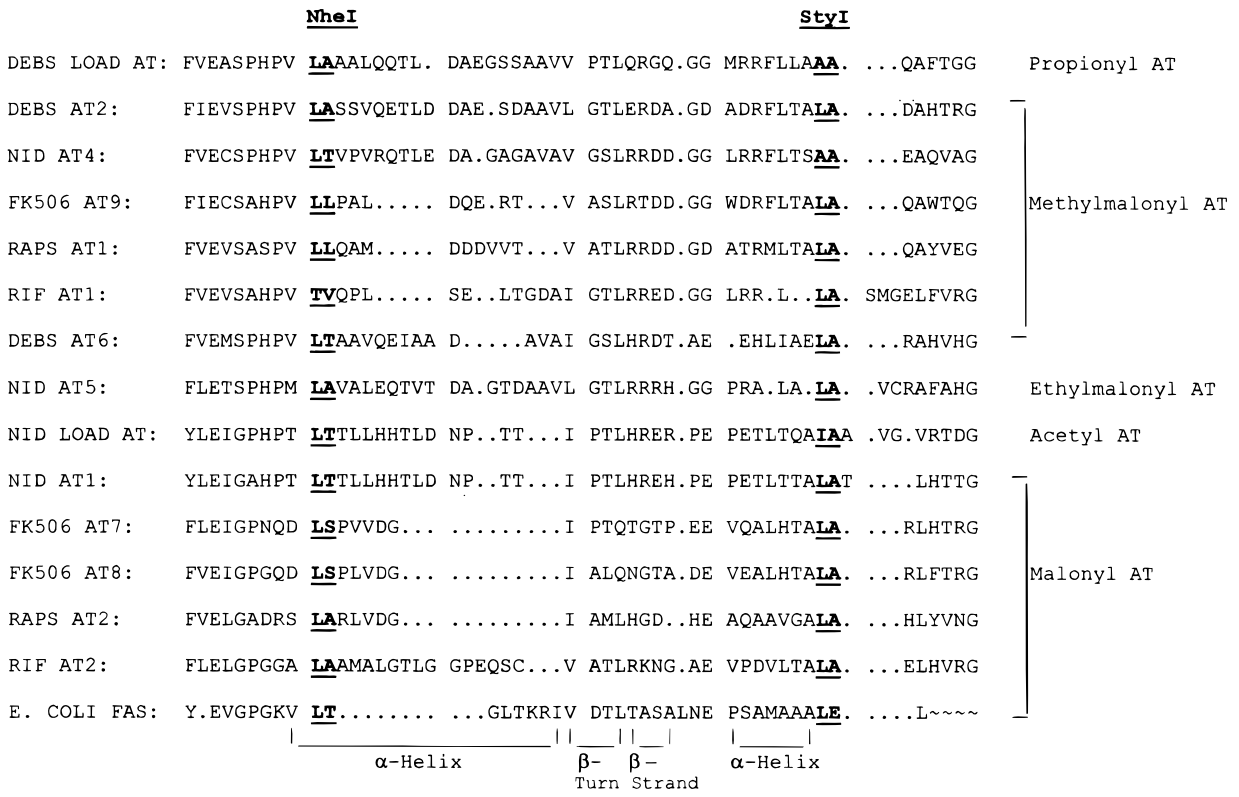

FIGURE 7: Alignment of amino acid sequences corresponding to the *Nhe*I−*Sty*I fragments of the malonyl transferase of the *E. coli* fatty acid synthase, and selected AT domains from the erythromycin (DEBS), rapamycin (RAPS), niddamycin (NID), FK506, and rifamycin (RIF) PKS clusters. The secondary structure corresponding to this region of the malonyl transferase of the *E. coli* fatty acid synthase is designated.

the 2-desmethyl-6-dEB compound (**3**) (Figure 1) (14). In the latter case, the resulting modified intermediate generated by module 6 does not have to undergo further downstream processing and can simply be directly cyclized by the thioesterase domain.

Consistent with these independent results, we note in our studies that replacement of the DEBS AT2 domain with the RAPS AT2 domain in DEBS1+TE leads to the production of approximately 10 mg/mL of the expected 2-desmethyl triketide lactone (**6**). By contrast, the same domain swap in the full (6-module) DEBS system results in a considerably lower yield of the corresponding 10-desmethyl-6-dEB (**2**) (<1 mg/L). This finding indicates that the heterologous RAPS AT2 domain is fully capable of recruiting a malonyl-CoA extender unit for incorporation onto the growing polyketide chain, and that subsequent processing of the altered intermediate by downstream modules acts as the main obstacle to efficient product formation. Whether this reduced yield is due to an intrinsic barrier due to discrimination against the altered triketide intermediate by one or more downstream modules, or is a consequence of aberrant intermodular interactions between the mutant and native proteins, cannot be distinguished at this time.

Methylmalonyl- versus Malonyl-CoA Substrate Specificity of Acyltransferase Domains in Modular PKSs. To delineate the region of the DEBS AT2 domain that controls substrate specificity, a series of chimeric AT domains were constructed by replacing fragments of the RAPS AT2 domain (which utilizes malonyl-CoA as substrate) with homologous regions from the DEBS AT2 domain (which uses methylmalonyl-CoA). The resulting hybrid AT domains were used to replace the entire AT2 domain of the full (6-module) DEBS system. Most chimeras thus generated were apparently inactive. In particular, when a 124-bp *Xho*I−*Sst*I fragment encoding the divergent sequence motif identified by Leadlay and co-workers for the malonyl-CoA-specific RAPS AT2 domain (*19*) was replaced with a 145-bp fragment encoding the sequence motif for the methylmalonyl-CoA-specific DEBS AT2 domain, the resulting recombinant strain, CH999/pJL286, did not lead to any detectable product formation. However, when a 198-bp *Pin*AI−*Pst*I fragment at the 3′-end of the RAPS AT2 domain was replaced with a homologous 225-bp fragment from the 3′-end of the DEBS AT2 domain, the resulting recombinant strain, CH999/pJL259, produced approximately 10 mg/L of the 6-dEB macrolactone (**1**) (Figure 6).

In an effort to refine the structural basis for extender-unit specificity, the 198-bp *Pin*AI−*Pst*I fragment at the 3′-end of the RAPS AT2 domain was further subdivided into a 68-bp *Nhe*I−*Sty*I fragment and a 123-bp *Sty*I−*Pst*I fragment. These two smaller regions were independently swapped with their DEBS AT2 counterparts, and the resultant hybrid AT domains were used to replace the AT2 domain of the 6-module DEBS. The recombinant strain CH999/pJL287, in which the 123-bp RAPS *Sty*I−*Pst*I was swapped with the 114-bp DEBS *Sty*I−*Pst*I fragment in the full DEBS, did not generate any detectable polyketides. On the other hand, the recombinant strain CH999/pJL285, in which the 68-bp RAPS *Nhe*I−*Sty*I fragment was swapped with the 104-bp DEBS *Nhe*I−*Sty*I fragment in the full DEBS, produced about 10 mg/L of 6-dEB (**1**) (Figure 6).

The above results suggest that the *Nhe*I−*Sty*I fragment plays an important role in controlling the substrate specificity of at least one AT domain. To test the generality of this

conclusion, we wished to conduct a reverse experiment whereby substitution of this segment in a methylmalonyl-AT domain with its counterpart from a malonyl-AT domain would lead to the formation of the expected desmethyl analogue of the natural product. Given the low yield of 10-desmethyl-6-dEB (**2**) as described above, we targeted instead the AT domain of module 6 for this experiment, since Liu and co-workers have previously shown that substitution of DEBS AT6 by RAPS AT2 leads to the biosynthesis of 2-desmethyl-6-dEB (**3**) in good yield (10 mg/L) (*14*). A hybrid AT domain was constructed by swapping a 89-bp *Nhe*I−*Sty*I fragment of the DEBS AT6 domain, which is homologous to the 104-bp *Nhe*I−*Sty*I fragment of the DEBS AT2 domain, with the 68-bp *Nhe*I−*Sty*I fragment of the RAPS AT2 domain. Again, the design of this hybrid was based on well-established sequence alignments of AT domains from modular PKSs (*19*, *21*). The recombinant strain CH999/pJL305 produced the predicted 2-desmethyl-6-dEB (**3**) (∼5 mg/L) as well as comparable amounts of 6-dEB (**1**) (Figure 6). Although the exact reason for the production of both compounds remains to be established, a mixture of **1** (<1 mg/L) and **3** (10 mg/L) is also known to be present in extracts from the recombinant strain previously described by Liu and co-workers (*14*, *22*).

Taken together, our findings suggest that the 35 amino acid segment encoded by the *Nhe*I−*Sty*I fragment of the DEBS AT2 domain or its counterparts in other AT domains, such as DEBS AT6 and RAPS AT2, are important determinants of specificity for methylmalonyl- or malonyl-CoA substrates. Sequence comparison among analogous segments from these and other known AT domains reveals a surprisingly high number of charged residues (15−30%) within this region (Figure 7). Moreover, there is also a striking diversity with respect to both the length (23−35 residues) and sequence of these amino acid segments. Most remarkably, this region of the AT domains also varies considerably from its counterpart in the malonyl transferase of the *E. coli* fatty acid synthase, whose crystal structure has recently been determined (*23*). The secondary structure of this region of the *E. coli* malonyl transferase is shown in Figure 7. The greatest variation in the length of the amino acid sequences encoded by the *Nhe*I−*Sty*I fragments of these AT domains is most notably in the region of first α-helix of the *E. coli* malonyl transferase. A definitive consensus sequence or sequences that correlate the primary protein structure of the AT domains to substrate specificity cannot as yet be deduced. But based on an analogy with the immunoglobulins, we propose that substrate specificity of AT domains in modular PKSs is controlled by a "hypervariable" region located at the C-terminal end of the domain, whose boundaries can be defined from sequence comparison, but whose selectivity cannot be explained based on primary structure alone.

Sequence analysis of 20 AT domains from DEBS and RAPS by Leadlay and co-workers led to the identification of signature sequences that could be used to discriminate between malonyl transferases and methylmalonyl transferases (*19*). Subsequent studies on the FK506 (*24*), rifamycin (*25*), and niddamycin (*26*) PKSs have confirmed the utility of these signature sequences in analyzing naturally occurring PKS modules. Our experimental results suggest that the predictive power of these signature sequences is probably due to evolutionary constraints rather than mechanistic ones. Thus, whereas the a priori prediction of substrate specificity of AT domains found in nature could benefit from the above-mentioned signature sequences, the design of AT domains with altered or relaxed substrate specificity will probably require structure-based and/or combinatorial genetic manipulation of the hypervariable region defined in this study.

REFERENCES

1. Cortés, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) *Nature 348*, 176−178.
2. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) *Science 252*, 675−679.
3. Katz, L. (1997) *Chem. Rev. 97*, 2557−2575.
4. Khosla, C. (1997) *Chem. Rev. 97*, 2577−2590.
5. Staunton, J., and Wilkinson, B. (1997) *Chem. Rev.* 2611−2629.
6. Cane, D. E., Liang, T., Taylor, P. B., Chang, C., and Yang C. (1986) *J. Am. Chem. Soc. 108*, 4957−4964.
7. Marsden, A. F. A., Caffrey, P., Aparicio, J. F., Loughran, M. S., Staunton, J., and Leadlay, P. F. (1994) *Science 263*, 378−380.
8. Weissman, K. J., Bycroft, M., Grice, P., Hanefeld, U., Staunton, J., and Leadlay, P. F. (1997) *Biochemistry 36*, 13849−13855.
9. Kao, C. M., McPherson, M., McDaniel, R. N., Fu, H., Cane, D. E., and Khosla, C. (1998) *J. Am. Chem. Soc. 120*, 2478−2479.
10. Cortés, J., Wiesmann, K. E. H., Roberts, G. A., Brown, M. J. B., Staunton, J., and Leadlay, P. F. (1995) *Science 268*, 1487−1489.
11. Kao, C. M., Luo, G., Katz, L., Cane, D. E., and Khosla, C. (1995) *J. Am. Chem. Soc. 11*, 9105−9106.
12. Oliynyk, M., Brown, M. J. B., Cortés, J., Staunton, J., and Leadlay, P. F. (1996) *Chem. Bio. 3*, 833−839.
13. Ruan, X., Pereda, A., Stassi, D. L., Zeidner, D., Summers, R. G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M. J., Donadio, S., and Katz, K. (1997) *J. Bacteriol. 170*, 6416−6425.
14. Liu, L., Thamchaipenet, A., Fu, H., Betlach, M., and Ashley G. (1997) *J. Am. Chem. Soc. 119*, 10553−10554.
15. McDaniel, R., Ebert-Khosla, S., Hopwood, D., and Khosla, C. (1993) *Science 262*, 1546−1557.
16. Hopwood, D. A., et al., and Schrempf, H. (1985) *Genetic Manipulation of Streptomyces: A Laboratory Manual*, The John Innes Foundation, Norwich, U.K.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Kao, C. M., Luo, G., Katz, L., Cane, D. E., and Khosla, C. (1994) *J. Am. Chem. Soc. 116*, 11612−11613.
19. Haydock, S. F., Aparicio, J. F., Molnar, I., Schwecke, T., Khaw, L. E., König, A., Marsden, A. F. A., Galloway, I. S., Staunton, J., and Leadlay, P. F. (1997) *FEBS Lett. 374*, 246−248.
20. McDaniel, R., Kao, C. M., Hwang, S. J., and Khosla, C. (1997) *Chem. Biol. 4*, 667−674.
21. Donadio, S., and Katz, L. (1992) *Gene 111*, 51−60.
22. Fu, H., unpublished results.
23. Serre, L., Verbree, E. C., Dauter, Z., Stuitje, A. R., and Derewenda, Z. S. (1995) *J. Biol. Chem. 270*, 12961−12964.
24. Motamedi, H., Cai, S. J., Shafiee, A., and Elliston, K. O. (1997) *Eur. J. Biochem. 224*, 74−80.
25. August, P. R., Tang, L., Yoon, Y. J., Ning, S., Müller, R., Yu, T. W., Taylor, M., Hoffmann, D., Kim, C. G., Zhang, X., Hutchinson, C. R., and Floss, H. G. (1998) *Chem. Biol. 5*, 69−79.
26. Kakavas, S. J., Katz, L., and Stassi, D. (1997) *J. Bacteriol. 179*, 7515−7572.

BI9820311