Substrate Specificity of the Loading Didomain of the Erythromycin Polyketide Synthase†

Janice Lau,‡ David E. Cane,§ and Chaitan Khosla*,‡,||

Departments of Chemical Engineering, Chemistry, and Biochemistry, Stanford University, Stanford, California 94305-5025, and Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912

Received March 16, 2000; Revised Manuscript Received June 16, 2000

ABSTRACT: The priming of many modular polyketide synthases is catalyzed by a loading acyltransferase−acyl carrier protein (AT_L−ACP_L) didomain which initiates polyketide biosynthesis by transferring a primer unit to the ketosynthase domain of the first module. Because the AT_L domain influences the choice of the starter unit incorporated into the polyketide backbone, its specificity is of considerable interest. The AT_L−ACP_L didomain of the 6-deoxyerythronolide B synthase (DEBS) was functionally expressed in *Escherichia coli*. Coexpression of the Sfp phosphopantetheinyl transferase from *Bacillus subtilis* in *E. coli* leads to efficient posttranslational modification of the ACP_L domain with a phosphopantetheine moiety. Competition experiments were performed with the holo-protein to determine the relative rates of incorporation of a variety of unnatural substrates in the presence of comparable concentrations of labeled acetyl-CoA. Our results showed that the loading didomain of DEBS can accept a surprisingly broad range of substrates, although it exhibits a preference for unbranched alkyl chain substrates over branched alkyl chain, polar, aromatic, and charged substrates. In particular, its tolerance toward acetyl- and butyryl-CoA is unexpectedly strong. The studies described here present an attractive prototype for the expression, analysis, and engineering of acyltransferase domains in modular polyketide synthases.

Modular polyketide synthases (PKSs) are large multifunctional enzyme complexes that are responsible for catalyzing the biosynthesis of structurally diverse polyketide natural products (*1*, *2*). These compounds possess a broad range of biological activities and are formed by the successive condensations of carboxylic acid monomers in a process that is mechanistically related to the biosynthesis of fatty acids. Individual modules of active sites in modular PKSs are sequentially organized, and are responsible for catalyzing all the reactions associated with one chain extension cycle. The ability of these enzymes to generate such an enormous variety of products is partly attributed to their use of a wide array of both starter and extender units to synthesize the polyketide backbones. Previous studies by several laboratories have shown that the choice of these substrates is primarily controlled by individual acyltransferase (AT) domains within each PKS. In particular, the primer units for initiating polyketide chain assembly are selected by loading acyltransferases (AT_L) (*3*−*6*), whereas the extender units incorporated into the polyketide backbones during chain elongation are chosen by the extender acyltransferases (*7*−*10*).

6-Deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea* is a modular PKS that catalyzes the formation of 6-deoxyerythronolide B (6-dEB) (**1**), the macrocyclic core of the antibiotic erythromycin (*11*, *12*). The assembly of 6-dEB involves six decarboxylative condensation steps between a propionyl-CoA derived primer unit and six methylmalonyl-CoA derived extender units (Figure 1). Each elongation cycle is catalyzed by one of the six modules of DEBS, which are arranged into three large multidomain proteins, DEBS1, DEBS2, and DEBS3. Each polypeptide is comprised of two modules, with the N-terminus of DEBS1 containing two additional catalytic domains—a loading acyltransferase (AT_L) and a loading acyl-carrier protein (ACP_L)—for delivering the propionyl primer to the first module for polyketide biosynthesis.

Although extender AT domains have been shown to exhibit a high degree of substrate specificity, the AT_L domains of modular PKSs are known to accept unnatural primer units (*13*−*16*). For example, the AT_L domain of the avermectin polyketide synthase can accept more than 40 different carboxylic acids as alternative starter units (*13*). Likewise, the AT_L domain of DEBS can accept and transfer acetyl, butyryl, and isobutyryl groups in addition to its natural propionyl substrate (*14*−*17*). Since these unnatural primers are processed intact through the downstream modules of the corresponding PKSs, quantitative analysis of the specificity of the entire PKS has been possible. However, since isolated AT domains have not been prepared thus far, the tolerance and specificity of individual AT domains remains unknown.

To study the intrinsic primer unit specificity of DEBS, we expressed and purified from *Escherichia coli* the loading

† This research was supported by grants from the National Institutes of Health (CA-66736 to C.K. and GM-22172 to D.E.C.) and by a grant from the Strategic Research Fund of Zeneca Co. to C.K. J.L. is a recipient of an NSF Predoctoral Fellowship and a Stanford-NIH Predoctoral Training Fellowship.

* To whom correspondence should be addressed. Phone/Fax: (650) 723-6538. E-mail: ck@chemeng.stanford.edu.

‡ Department of Chemical Engineering, Stanford University.

§ Department of Chemistry, Brown University.

|| Departments of Chemistry and Biochemistry, Stanford University.

10.1021/bi000602v CCC: $19.00
Published on Web 08/03/2000

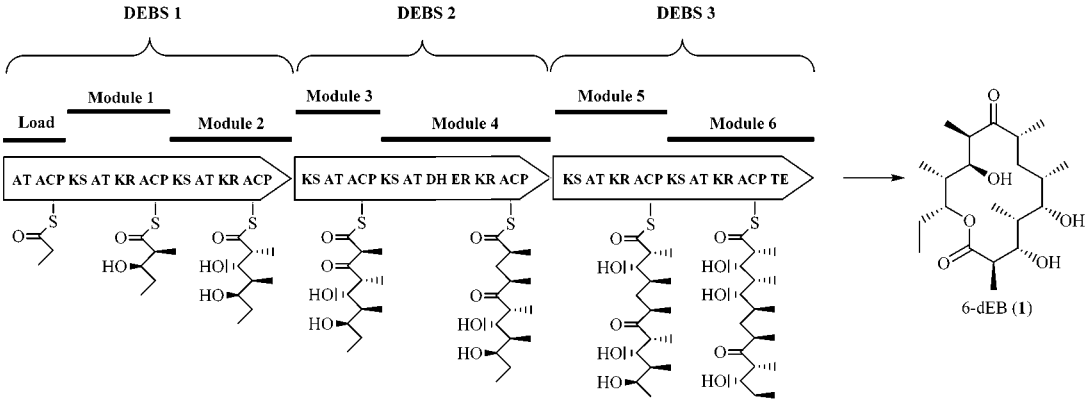

FIGURE 1: Modular organization of the 6-deoxyerythronolide B synthase (DEBS), which catalyzes the biosynthesis of 6-deoxyerythronolide B (6-dEB) (**1**). DEBS consists of three large multifunctional proteins, DEBS1, DEBS2, and DEBS3, with each containing two modules. Each module catalyzes one cycle of 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).

acyltransferase domain of DEBS as part of a bifunctional AT_L-ACP_L protein. (Our attempts thus far to express the isolated AT domain in the absence of the ACP domain have not been successful, yielding only inactive protein as insoluble inclusion bodies.) Because it has recently been shown that the appropriate engineering of N- and C-terminal ends plays an important role in the assembly of functional modules or domains of PKSs (*18*, *19*), three variants of the AT_L-ACP_L protein, differing only in their C-terminal linker downstream of the ACP_L domain, were initially constructed. All three variants were successfully isolated from *E. coli*, and they exhibited comparable catalytic activities. Although posttranslational modification of the loading didomain is not required for the biosynthetic competence of the AT_L domain, phosphopantetheinylation is necessary for investigating the transfer of the starter acyl intermediate to the loading acyl-carrier protein (ACP_L) (*20*). The Sfp phosphopantetheinyl transferase from *Bacillus subtilis* (*21*) is known to be able to modify a wide variety of heterologous carrier domains (*20*, *22*). Here, we report that Sfp can efficiently modify the ACP_L domain of the AT_L-ACP_L protein with phosphopantetheine in vivo. The resulting *holo*-enzyme was assayed against a number of substrates to probe its ability to discriminate between structurally different primer units. The results of this study are described here.

EXPERIMENTAL PROCEDURES

Chemicals. [1-^{14}C]Acetyl-CoA (50 mCi/mmol) and [1-^{14}C] propionyl-CoA (54 mCi/mmol) were obtained from Moravek Biochemicals. All other unlabeled starter and extender coenzyme A derivatives—acetyl-, propionyl-, *n*-butyryl-, isobutyryl-, benzoyl-, phenylacetyl-, β-hydroxybutyryl-, crotonoyl-, methylmalonyl-, and malonyl-CoA—were purchased from Sigma Chemical Company and used without further purification.

Manipulation of DNA and Strains. DNA manipulations were performed in *E. coli* XL1 Blue (Stratagene) using standard culture conditions (*23*). Polymerase chain reactions (PCR) were carried out using *Pfu* polymerase (Stratagene) as recommended by the manufacturer.

Construction of AT_L-ACP_L Didomain. In creating a functional AT_L-ACP_L didomain, three different linkers were designed and independently fused to the C-terminus of the loading ACP domain. The genes corresponding to these three variants are cloned into pET21c (Novagen) to form the expression plasmids, pJL636, pJL637, and pJL638. Previously, Aparicio and co-workers demonstrated that limited proteolysis of DEBS1 with trypsin leads to the generation of a 60-kDa fragment, which houses the AT_L-ACP_L didomain (*6*). The resulting fragment was found to be specifically labeled after incubation with [^{14}C]propionyl-CoA. Plasmid pJL636 encodes for an AT_L-ACP_L didomain that was intended to mimic this proteolytic fragment. It has a C-terminus at the tryptic site normally located 12 amino acids downstream of the loading ACP. The construction of this plasmid was accomplished as follows. An *Nde*I−*Sph*I restriction fragment (nucleotides 578−1807; GenBank accession no. M63676) corresponding to the N-terminal region of the didomain was excised from the plasmid, pKOS001−11, which harbors the "DEBS1+TE" gene in a derivative of the pUC19 vector. A *Sph*I−*Eco*RI fragment (nucleotides 1807−2228) corresponding to the C-terminal region was amplified by PCR using the primers 5′-GCA TGC **GGC GGT TCC TGC TGG CCG CGG CCC**-3′ and 5′-GAA TTC **CGG GTT TCC CGT TGT GC**-3′ (bases complementary to the DEBS gene are shown in bold). The isolated *Nde*I−*Sph*I and *Sph*I−*Eco*RI fragments were subsequently ligated together and introduced into *NdeI/Eco*RI-digested pET21c to produce pJL636.

Plasmid pJL637 encodes an AT_L-ACP_L didomain whose C-terminus is fused to the C-terminal amino acids of DEBS1. This plasmid was constructed as follows. A PCR-amplified *Sph*I−*Spe*I fragment (nucleotides 1807−2193) corresponding to the C-terminal region of the didomain was inserted into a subclone containing a PCR-amplified *Spe*I−*Eco*RI fragment (nucleotide 10953−11216) corresponding to the interpolypeptide linker at the C-terminus of DEBS1. The primers for the amplification of the *Sph*I−*Spe*I fragment are 5′-GCA TGC **GGC GGT TCC TGC TGG CCG CGG CCC**-3′ and 5′-ACT AGT **GAG TGC TTC CGC GAG CGC**-3′. The primers for the amplification of the *Spe*I−*Eco*RI fragment are 5′-ACT AGT GAG CTC GGC ACC GAG GTC-3′ and 5′-GAA TTC **TGC CCG TCG AGC TCC CGG**-3′. The resulting subclone was digested with *Sph*I and *Eco*RI, and the 649-bp *Sph*I−*Eco*RI fragment was isolated and ligated to the *Nde*I−*Sph*I restriction fragment described above. The

final 1.9-kb *Nde*I−*Eco*RI fragment was introduced into *Nde*I−*Eco*RI-digested pET21c to produce pJL637.

Analogous to plasmid pJL637, plasmid pJL638 encodes an AT_L−ACP_L didomain that was capped by the C-terminal end of DEBS2. This plasmid was constructed as follows. An *Spe*I−*Xho*I fragment (nucleotide 10471−10719; GenBank accession no. 63677) corresponding to the interpolypeptide linker at the C-terminus of DEBS2 was amplified by PCR using the primers 5′-ACT AGT **ACG CTG TTC GCG GCC TGA**-3′ and 5′-CTC GAG **GTC CTC TCC CCC GCC**-3′. The resulting fragment was introduced into *Spe*I−*Xho*I-digested pJL637 to produce pJL638.

Expression and Purification of AT_L-ACP_L Didomains. Plasmid pJL636 was introduced via transformation into *E. coli* BL21-CodonPlus-RIL (Stratagene) for expression of the *apo* AT_L−ACP_L didomain. The strain BL21-CodonPlus-RIL carries a chloramphenicol resistance gene and contains extra copies of the genes that encode the tRNAs of rare *E. coli* codons. One liter cultures of BL21-CodonPlus-RIL/pJL636 were grown at 37 °C in 2 L flasks containing LB medium supplemented with 100 μg/mL carbenecillin and 12 μg/mL chloramphenicol. Expression of the loading didomain was induced with 1 mM IPTG at an OD_{600} of 0.6. After induction, incubation was continued for 5 h at 30 °C. The cells were then harvested by centrifugation at 2500*g* and resuspended in 25 mL of 50 mM Tris, pH 8.0.

All purification procedures were performed at 4 °C. The resuspended cells were disrupted by two passages through a French press at 13 000 psi. The supernatant was collected by centrifugation at 10 000*g* and loaded onto a previously equilibrated Ni-NTA column (4 cm, Qiagen). The column was washed with 10 mM imidazole in 50 mM Tris, pH 8.0, and the *apo* AT_L−ACP_L didomain eluted at 50 mM imidazole. Pooled fractions containing the loading didomain were concentrated to a volume of 1 mL with a Centriprep-30 concentrator (Amicon) and subsequently diluted to 20 mL in 100 mM NaH_2PO_4, pH 7.2, 2.5 mM DDT, 1 mM EDTA, and 20% glycerol (buffer A). The protein was applied to a previously equilibrated Resource Q column (6 mL, Pharmacia) at 2 mL/min. An initial gradient from 0 to 130 mM NaCl in buffer A was run at 1 mL/min for 13 min. This was followed by a shallower gradient from 130 to 200 mM NaCl at 1 mL/min for 28 min. Peak fractions that eluted from the column at 150−180 mM NaCl were pooled and concentrated with a Centriprep-30 concentrator (Amicon). The purified protein was flash-frozen in liquid nitrogen and stored at −80 °C. Protein concentrations were estimated by densitometric scanning of Coomassie blue-stained protein. A typical 1 L culture produced about 1 mg of purified enyzme.

For expression of the *holo* AT_L−ACP_L didomain, plasmids pJL636, pJL637, and pJL638 were independently cotransformed into BL21-CodonPlus-RIL containing the plasmid pRSG56 (*18*), which carries a kanamycin resistance gene and the *sfp* gene. One liter cultures of these recombinant *E. coli* strains were grown at 37 °C in 2 L flasks containing LB medium supplemented with 100 μg/mL carbenicillin, 50 μg/mL kanamycin, and 12 μg/mL chloramphenicol. The expression and purification steps for the three AT_L−ACP_L variants were performed in the same manner as described above.

Radioactive Labeling of Apo AT_L−ACP_L Didomain. To examine the acyltransferase activity of the *apo* AT_L−ACP_L didomain, reaction mixtures were set up in a total volume of 25 μL containing 2.2 μM *apo* AT_L−ACP_L (encoded by pJL636), 1 mM propionyl-CoA (222 μM [1-^{14}C]propionyl-CoA + 778 μM unlabeled propionyl-CoA), 100 mM NaH_2PO_4, pH 7.2, 2.5 mM DTT, 1 mM EDTA, and 20% glycerol. The reactions were incubated at room temperature and quenched at specific times (1, 5, 15, and 30 min) with 50 μL of an ice-cold 10% trichloroacetic acid (TCA) solution. Precipitation was completed by incubation on ice for 10 min. The precipitate was pelleted by centrifugation, washed three times with 100 μL of 10% TCA, and solubilized in 200 μL of a 20 mM NaOH, 2% SDS solution. The suspension was combined with 5 mL of liquid scintillation fluid, and the incorporated ^{14}C label was quantified by liquid scintillation counting (LSC).

Radioactive Labeling of Holo AT_L−ACP_L. To examine the extent of in vivo phosphopantetheinylation of the AT_L−ACP_L didomain as a result of the coexpression of Sfp, reaction mixtures were set up in a total volume of 25 μL containing 2 μM *holo* AT_L−ACP_L (encoded by pJL636), 1 mM propionyl-CoA (222 μM [1-^{14}C]propionyl-CoA + 778 μM unlabeled propionyl-CoA), 100 mM NaH_2PO_4, pH 7.2, 2.5 mM DTT, 1 mM EDTA, and 20% glycerol. The reaction mixtures were incubated at room temperature and quenched at specific times (1, 5, 15, and 30 min) with 50 μL of ice-cold 10% TCA. Precipitation was completed by incubation on ice for 10 min. The precipitate was pelleted by centrifugation, washed three times with 100 μL of 10% TCA, and solubilized in 200 μL of a 20 mM NaOH, 2% SDS solution. The suspension was combined with 5 mL of liquid scintillation fluid, and the incorporated ^{14}C label was quantified by LSC.

Incorporation of [^{14}C]Propionyl-CoA and [^{14}C]Acetyl-CoA into AT_L−ACP_L Didomain as a Function of Substrate Concentration. To assess the saturating substrate concentration for the loading didomain, reaction mixtures were set up in a total volume of 25 μL containing 3.6 μM *holo* AT_L−ACP_L (encoded by pJL636), 300 μM [1-^{14}C]propionyl-CoA or [1-^{14}C]acetyl-CoA, 100 mM NaH_2PO_4, pH 7.2, 2.5 mM DTT, 1 mM EDTA, and 20% glycerol. The concentration of unlabeled propionyl- or acetyl-CoA in the mixtures ranged from 700 μM to 3.7 mM. The reactions were incubated at room temperature and quenched at specific times (5, 10, and 15 min) with 50 μL of ice-cold 10% TCA. Precipitation of the loading didomain and quantification of the incorporated ^{14}C label were performed as described above.

Analysis of the Specificity of the AT_L−ACP_L Didomain for Different Starter Units. Since the AT-catalyzed acylation of the loading ACP is a very rapid event, and since the enzyme and substrate are necessarily present in stoichiometric ratios, measurement of the absolute kinetic parameters for this reaction in the presence of different CoA thioesters was not possible. However, the relative specificity of the loading didomain for different substrates could be measured. To this end, reaction mixtures were set up in a total volume of 25 μL containing 3.6 μM *holo* AT_L−ACP_L (encoded by pJL636), 2 mM acetyl-CoA (300 μM [1-^{14}C]acetyl-CoA + 1.7 mM unlabeled acetyl-CoA), 2 mM unlabeled competing coenzyme A derivative (propionyl-, *n*-butyryl-, isobutyryl-, benzoyl-, phenylacetyl-, β-hydroxybutyryl-, crotonoyl-, methylmalonyl-, or malonyl-CoA), 100 mM NaH_2PO_4, pH 7.2, 2.5 mM DTT, 1 mM EDTA, and 20% glycerol. After

incubation at room temperature for 10 min, the reactions were quenched with 50 μL of ice-cold 10% TCA. Precipitation of the loading didomain and quantification of the incorporated ^{14}C label were performed as described above.

RESULTS

Construction and Purification of AT_L-ACP_L Didomain. Previous studies have demonstrated the importance of the presence of appropriate linkers for the expression of functional modules or domains of PKSs as individual proteins (*18*, *19*). To assess the significance of engineering suitable linkers for the expression of a catalytically active AT_L-ACP_L didomain, three variants of this protein were constructed. It has been observed by Pereda and co-workers that the 107 amino acids, which precede the loading acyltransferase domain of DEBS (as defined by homology), are essential in creating an active AT_L protein that can complement a DEBS mutant lacking this domain for the production of erythromycin in *S. erythraea* (*19*). On the basis of this observation, all three variants of the AT_L-ACP_L didomain were designed to retain this 107 amino acid segment, which extends from the natural start codon of DEBS1 to the onset of the AT_L domain as defined by sequence alignment of various AT domains, at their N-termini.

Although the three variants possess the same N-terminus, they differ in the construction of their C-termini. The variant encoded by plasmid pJL636 was modeled after a proteolytic fragment generated by limited trypsinization of DEBS1 that was specifically labeled with [^{14}C]propionyl-CoA (*6*). Because the tryptic site corresponding to this fragment is normally located 12 amino acids downstream of the loading acyl-carrier protein domain, a short linker consisting of these 12 amino acids was fused to the C-terminus of the AT_L-ACP_L didomain. The other two variants, encoded by plasmids pJL637 and pJL638, possessed longer C-terminal tails whose sequences were based on those of the C-termini of DEBS1 and DEBS2, respectively.

The *apo* and *holo* forms of the three AT_L-ACP_L variants were produced in *E. coli* as C-terminal hexahistidine $(His)_6$-tagged fusion proteins. They were purified by nickel affinity chromatography followed by anion-exchange chromatography to >95% homogeneity (Figure 2). All three variants can be acylated with [^{14}C]propionyl- and [^{14}C]acetyl-CoA and, as determined by SDS−PAGE/autoradiography, they exhibited comparable catalytic properties (data not shown). Therefore, in the remainder of our studies, we focused on the protein derived from pJL636.

Posttranslational Phosphopantetheinylation of AT_L-ACP_L Didomain. All polyketide synthases require posttranslational modification of their ACP domains to become catalytically active (*20*). Covalent attachment of the 4′-phosphopantetheine moiety of coenzyme A to the conserved serine residue of the ACP_L domain is therefore essential for transfer of the primer unit from the AT_L domain to the ACP_L domain. However, a key barrier to this process is the inability of the inherent phosphopantetheinyl transferases of *E. coli* to posttranslationally modify the AT_L-ACP_L didomain. To overcome this obstacle, the Sfp phosphopantetheinyl transferase from *B. subtilis*, which is capable of modifying the *apo* forms of many heterologous recombinant proteins (*21*, *22*), was coexpressed with the loading didomain.

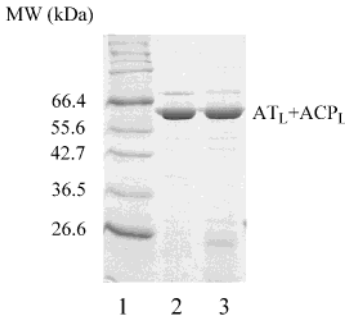

FIGURE 2: Purity of recombinant *apo* and *holo* AT_L+ACP_L didomain (encoded by plasmid pJL636) overproduced in and purified from *E. coli*. Protein samples were resolved by SDS−PAGE (10%) and stained with Coomassie blue. Lanes 1−3 correspond to the molecular weight markers, the *apo* AT_L+ACP_L protein, and the *holo* AT_L+ACP_L protein, respectively. Quantitative densitomer scanning revealed that the proteins were >95% pure. The contaminant protein that appears as a minor band around 70 kDa was uncharacterized, but does not label in the presence of any radiolabeled CoA thioesters, and is unlikely to have acyltransferase activity.

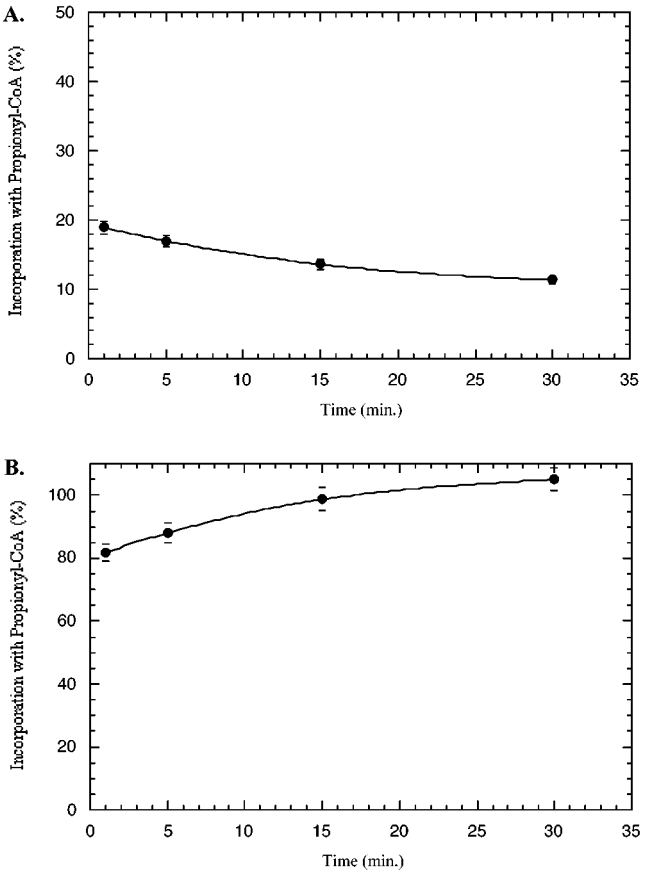

FIGURE 3: Incorporation of propionyl-CoA into the (A) *apo* AT_L+ACP_L protein, and (B) *holo* AT_L+ACP_L protein. The initial substrate concentration in each reaction mixture was 1 mM. Labeling is plotted as the extent of incorporation of [^{14}C]propionyl-CoA into 2 μM of AT_L+ACP_L protein, as monitored by radioassay. Complete labeling of both the AT_L and ACP_L domains of the protein is designated as 100%. Thus, for the apo-protein the theoretical maximum level of labeling would be 50%. See experimental procedures for details.

In the absence of posttranslational modification, only the acyltransferase domain of the AT_L-ACP_L protein is capable of accepting a substrate. As shown in Figure 3A, incubation of the *apo*-protein with [^{14}C]propionyl-CoA resulted in an initial burst of ^{14}C-labeling followed by a gradual decrease,

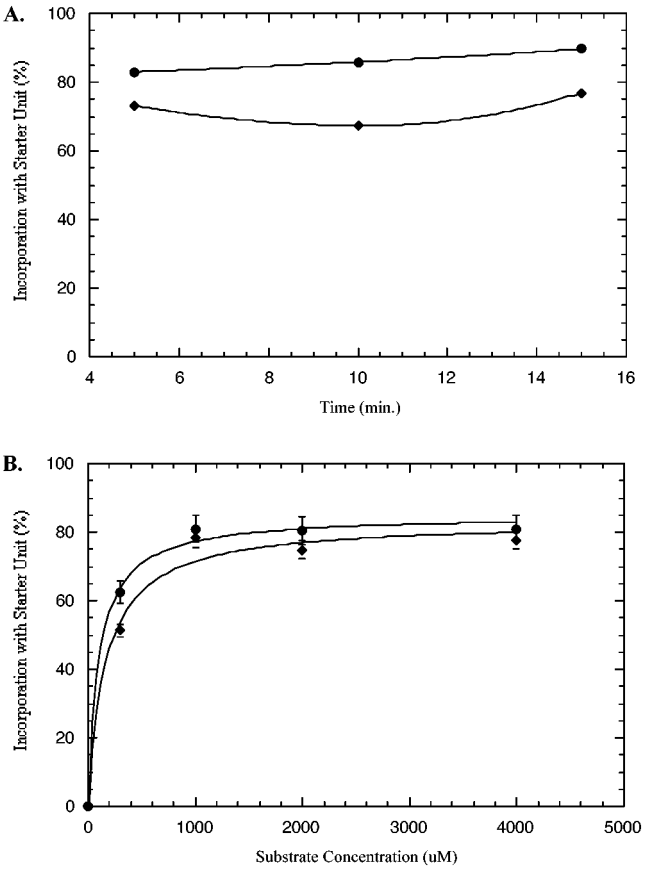

FIGURE 4: (A) Incorporation of propionyl- (◆) and acetyl-CoA (●) into 3.6 μM of *holo* AT_L+ACP_L protein. The initial substrate concentration in each reaction mixture was 2 mM. (B) Incorporation of propionyl- (◆) and acetyl-CoA (●) into 3.6 μM of *holo* AT_L+ACP_L protein as a function of substrate concentration after 10 min. Labeling activity was measured as the extent of incorporation of [^{14}C]propionyl- or [^{14}C]acetyl-CoA into the AT_L+ACP_L protein as monitored by radioassay. Complete labeling of both the AT_L and ACP_L domains of the protein is designated as 100%. See experimental procedures for details.

suggesting progressive hydrolysis of the labeled substrates from the AT_L domain to a steady-state level of occupancy corresponding to ca. 20% of the available AT active sites. In contrast, when posttranslationally modified AT_L−ACP_L didomain was incubated with [^{14}C]propionyl-CoA under similar conditions, nearly 100% of the estimated AT and ACP active sites in the protein were labeled (Figure 3B). Addition of separately purified Sfp to the reaction mixture did not result in an increase in the labeling intensity after 30 min incubation (data not shown). Thus, Sfp can efficiently convert the *apo* form of the loading didomain into the *holo* form in vivo. Moreover, the transacylation reaction is relatively fast, as judged by its ability to reach 80% completion within the first minute of incubation with propionyl-CoA (Figure 3B). This is consistent with earlier studies on homologous acyltransferases from Type II fatty acid synthases, where k_{cat} values as high as 1.6×10^3 s^{-1} have been reported (*24*).

Substrate Specificity of AT_L−ACP_L Domain. Earlier studies have demonstrated that DEBS can accept and process unnatural primer units (*14*−*17*). We were therefore interested in quantifying the specificity of the loading didomain for alternative substrates. Owing to the rapid rate of the transacylation reaction catalyzed by AT_L, it was not possible to measure the absolute kinetic parameters for the loading didomain in the presence of natural and unnatural substrates. Therefore, the substrate specificity of the loading didomain was measured via competition assays in which labeling of the *holo*-protein by acetyl-CoA was quantified in the presence of comparable concentrations of a variety of unlabeled coenzyme A thioesters. Since these experiments were performed under conditions that were expected to yield stoichiometric acylation of the loading didomain, the percent acylation by the labeled substrate is proportional to the ratio of the "apparent" V/K_M for the labeled substrate versus the unlabeled competing cosubstrate. To validate this assumption, a control experiment was performed in which equimolar concentrations (2 mM) of radiolabeled propionyl-CoA and unlabeled acetyl-CoA were used. The ratio of incorporation of these two substrates calculated from this experiment was indistinguishable from that obtained by using equimolar concentrations (2 mM) of radiolabeled acetyl-CoA and unlabeled propionyl-CoA (data not shown).

Table 1: Substrate Specificity of the Loading Didomain of DEBS

source of primer unit	labeling by acetyl-CoA (dpm)	incorporation rates relative to propionyl-CoA[a]
acetyl-CoA	2008	50 ± 2
propionyl-CoA	671	100
n-butyryl-CoA	857	60 ± 1
isobutyryl-CoA	1399	38 ± 4
crotonoyl-CoA	1469	30 ± 1
benzoyl-CoA	1914	0 + 2
phenylacetyl-CoA	1765	9 ± 2
β-hydroxybutyryl-CoA	1871	2 ± 9
(*RS*)-methylmalonyl-CoA	1736	11 ± 1
malonyl-CoA	1840	4 ± 1

[a] The amount of the competing (unlabeled) substrate that was incorporated into the protein in each experiment was determined by subtracting the amount of labeled acetyl groups in the presence of the competing substrate from that obtained in its absence. Although acetyl-CoA was used as the labeled substrate in these experiments to maximize signal-to-noise, the relative incorporation rates reported here are normalized to propionyl-CoA, which is the natural substrate of the loading didomain of DEBS. The error bars reported are based on two independent experiments. Background labeling in the absence of the loading didomain or in the presence of control proteins was under 50 dpm; this background value was subtracted from the counts reported here.

To determine the optimal quench time and concentration of the labeled substrate to be used in these competition experiments, *holo* AT_L−ACP_L was independently incubated with varying concentrations of propionyl-CoA and acetyl-CoA, and the reactions were quenched at different times. As illustrated in Figure 4, maximal incorporation of both substrates was achieved after 5 min at a substrate concentration of 2 mM. Therefore, measurements of the relative rates of incorporation of acyl groups derived from different coenzyme A thioesters were obtained by incubating 2 μM of *holo* AT_L−ACP_L with equimolar ratios (2 mM) of the two competing substrates for 5 min.

As shown in Table 1, the loading didomain of DEBS exhibits a 2-fold preference toward its natural substrate, propionyl-CoA, over acetyl- and *n*-butyryl-CoA, and a 3-fold preference over isobutyryl- and crotonoyl-CoA. The incorporation of aromatic substrates into the loading didomain was significantly reduced; indeed, benzoyl-CoA is not a substrate at all within detectable limits, whereas phenylacetyl

primer units are transacylated at 10-fold lower levels compared to propionyl primer units. The loading didomain also strongly discriminates against polar and charged primer units, such as β-hydroxybutyryl-, methylmalonyl-, and malonyl-CoA derived acyl groups. Notably, however, methylmalonyl-CoA appears to be tolerated as a substrate by the enzyme and is in fact a better substrate than malonyl-CoA. To confirm that methylmalonyl-CoA and malonyl-CoA were indeed capable of acylating the loading didomain, the protein was incubated with radiolabeled substrates, and the products of the reaction were analyzed by SDS−PAGE/autoradiography. In both cases, labeled protein could be detected after 5 min (data not shown). Moreover, the extent of labeling observed on the apo-protein was significantly less than that observed on the holo-protein, suggesting that the pathway of acyl transfer in the presence of these substrates was similar to that in the cases of acetyl-, propionyl-, and butyryl-CoA.

DISCUSSION

Although several previous reports have measured the substrate specificity of AT domains in the context of multimodular PKSs (*16*−*18*, *25*, *26*), this is the first study to quantify the substrate specificity of an isolated AT derived from a modular PKS. The availability of the AT_L−ACP_L didomain of DEBS in a recombinant form from *E. coli* therefore affords detailed analysis of the structure and specificity of an acyltransferase. In contrast to the expression of complete PKS modules, which require careful engineering of N- and C-terminal sequences, expression of a functional loading didomain appears to be more permissive. The three variants of the AT_L−ACP_L protein described here, which possess the same natural N-terminus derived from DEBS but differ in their C-termini, exhibit comparable catalytic capabilities. It should be noted, however, that although a suitable C-terminal linker may not be essential for self-acylation of the loading didomain by a primer unit, its presence may be necessary for efficient transfer of the substrate to the downstream modules of DEBS. The recombinant proteins described here could therefore provide interesting models for future tests of the role of linkers in chain transfer.

Although the *apo* form of the loading didomain is successfully expressed in *E. coli*, the native phosphopantetheinyl transferases of *E. coli* are not able to posttranslationally prime the ACP_L domain with phosphopantetheine. In the absence of such posttranslational modification, the substrates utilized by the AT_L domain would not be transferred to the ACP_L domain, and would eventually be hydrolyzed. This is illustrated by the gradual decrease in labeling of the *apo* AT_L+ACP_L protein with [^{14}C]propionyl-CoA over time (Figure 3A). Our results confirm that coexpression of Sfp in *E. coli* can lead to complete conversion of the *apo* AT_L+ACP_L into the *holo* form (Figure 3B). Moreover, both catalytic sites of the *holo* AT_L+ACP_L protein can be fully occupied simultaneously, suggesting that phosphopantetheinylation of the ACP_L domain may also induce a change in the didomain protein so as to stabilize the otherwise labile acyl-AT_L intermediate.

Earlier analysis of the primer unit specificity of the bimodular DEBS1+TE, which includes the loading didomain, module 1, module 2, and the thioesterase domain of DEBS, revealed that this multifunctional enzyme has a 32-fold preference for propionyl units over acetyl units and an 8-fold preference for propionyl units over butyryl units (*25*). More recently, a similar study has been performed with a mutant of DEBS1+TE, which has an inactivated ketosynthase domain in module 1 (KS1°). Here, *N*-acetylcysteamine analogues of the natural diketide intermediate were used to prime module 2, which converted these substrates into the corresponding triketide lactone products (*26*). In this system, the natural C_5 diketide (corresponding to a propionyl primer unit at the loading domain and a propionyl extender) was only a 12-fold better substrate than the C_4 diketide (which corresponds to an acetyl primer unit at the loading domain), whereas a longer chain C_6 diketide (corresponding to a butyryl primer unit at the loading domain) was 1.6-fold superior to the natural C_5 diketide. In this report, we quantify the specificity of the loading didomain of DEBS alone for its natural substrate. Our measurements show that the loading didomain is suprisingly tolerant toward longer and shorter substrates with less than a 2-fold specificity for natural propionyl units over unnatural acetyl and butyryl units (Table 1). Taken together, these results imply that the principal determinants of primer unit specificity of DEBS lie not in the loading didomain but in the active sites of the downstream modules. Thus, module 1 appears to be the principal point of discrimination between propionyl and butyryl primer units, whereas module 2 appears to be the principal point of discrimination between growing chains derived from propionyl and acetyl primer units.

Table 1 also provides additional new insights into the molecular recognition features of the active site of the AT_L domain. Although the substrate binding pocket appears to have the greatest preference for unbranched alkyl chains, it is also tolerant toward branched alkyl chains and olefinic functional groups. The ability to incorporate bulkier, polar, and charged substrates is further reduced but still measurable in most cases. The preference for methylmalonyl-CoA over malonyl-CoA may reflect the fact that methylmalonyl-CoA more closely resembles propionyl-CoA than malonyl-CoA. Indeed, Smith and co-workers have shown that a single residue within the acyltransferase domain of the vertebrate fatty acid synthase plays an important role in discriminating between (neutral) primer units and (α-carboxylated) extender units (*27*). Thus, it may be feasible to further broaden or alter the substrate specificity of the loading didomain of DEBS via protein engineering. In this regard, the heterologous expression of the loading didomain in *E. coli* offers a convenient system for constructing and analyzing directed or random mutants.

REFERENCES

1. O'Hagan, D. (1991) *The Polyketide Metabolites*, E. Norwood, New York.
2. Katz, L., and Donadio, S. (1993) *Annu. Rev. Microbiol. 47*, 875−912.
3. Marsden, A. F. A., Wilkinson, B., Cortés, J., Dunster, N. J., Staunton, J., and Leadlay, P. F. (1998) *Science 279*, 199−202.
4. Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W., and Rao, R. N. (1996) *Gene 183*, 231−236.
5. Hunziker, D., Yu, T.-W, Hutchinson, C. R., Floss, H. G., and Khosla, C. (1998*) J. Am. Chem. Soc. 120*, 1092−1093.
6. Aparicio, J. F., Caffery P., Marsden, A. F. A, Staunton, J., and Leadlay, P. F. (1994) *J. Biol. Chem. 269*, 8524−8528.

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. Oliynyk, M., Brown, M. J. B., Cortés, J., Staunton, J., and Leadlay, P. F. (1996) *Chem. Biol. 3*, 833−839.
9. 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.
10. 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.
11. Cortés, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) *Nature 348*, 176−178.
12. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) *Science 252*, 675−679.
13. Dutton, C. J., Gibson, S. P., Goudie, A. C., Holdom, K. S., Pacey, M. S., Ruddock, J. L., Bu'Lock, J. D., and Richards, M. K. (1991) *J. Antibiot. 44*, 357−365.
14. Kao, C. M., Katz, L., and Khosla, C. (1994) *Science 265*, 509−512.
15. Kao, C. M., Luo, G., Katz, L., Cane, D. E., and Khosla, C. (1995) *J. Am. Chem. Soc. 117*, 9105−9106.
16. Wiesmann, K. E. H., Cortés, J., Brown, M. J. B., Cutter, A. L., Staunton, J., and Leadlay, P. F. (1995) *Chem. Biol. 2*, 583−589.
17. Pieper, R., Luo, G., Cane, D. E., and Khosla, C. (1995) *J. Am. Chem. Soc. 117*, 11373−11374.
18. Gokhale, R. S, Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) *Science 284*, 482−485.
19. Pereda, A., Summers, R. G., Stassi, D. L., Ruan, X., and Katz, L. (1998) *Microbiology 144*, 543−553.
20. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. N., and Flugel, R. S. (1997) *Curr. Opin. Chem. Biol. 1*, 309−315.
21. Nakano, M. M., Corbell, N., Besson, J., and Zuber, P. (1992) *Mol. Gen. Genet. 232*, 313−321.
22. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., Lacelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989*) Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
24. Joshi, V. C., and Wakil, S. J. (1971) *Arch. Biochem. Biophys. 143*, 493−505.
25. Pieper, R., Ebert-Khosla, S., Cane, E., and Khosla, C. (1996) *Biochemistry 35*, 2054−2060.
26. Chuck, J., McPherson, M., Huang, H., Jacobsen, J. R., Khosla, C., and Cane, D. E. (1997) *Chem. Biol. 4*, 757−766.
27. Rangan, V. S., and Smith, S. (1997) *J. Biol. Chem. 272*, 11975−11978.

BI000602V