BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 95–102 95
DOI: 10.1021/bi901826g

Protein−Protein Recognition between Acyltransferases and Acyl Carrier Proteins in Multimodular Polyketide Synthases[†]

Fong T. Wong,[‡] Alice Y. Chen,[‡] David E. Cane,[§] and Chaitan Khosla*[,‡,∥]

[‡]*Departments of Chemical Engineering, and* [∥]*Chemistry, and Biochemistry, Stanford University, Stanford, California 94305, and* [§]*Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912-9108*

Received October 26, 2009; Revised Manuscript Received November 15, 2009

ABSTRACT: Acyltransferase (AT) domains of multimodular polyketide synthases are the primary gatekeepers for stepwise incorporation of building blocks into a growing polyketide chain. Each AT domain has two substrates, an α-carboxylated CoA thioester (e.g., malonyl-CoA or methylmalonyl-CoA) and an acyl carrier protein (ACP). Whereas the acyl-CoA specificity of AT domains has been extensively investigated, little is known about their ACP specificity. Guided by recent high-resolution structural insights, we have systematically probed the protein−protein interactions between AT domains, ACP domains, and the linkers that flank AT domains. Representative AT domains of the 6-deoxyerythronolide B synthase (DEBS) have greater than 10-fold specificity for their cognate ACP substrates as compared to other ACP domains from the same synthase. Both of the flanking (N- and C-terminal) linkers of an AT domain contributed to the efficiency and specificity of transacylation. As a frame of reference, the activity and specificity of a stand-alone AT domain from the "AT-less" disorazole synthase (DSZS) were also quantified. The activity (k_{cat}/K_M) of this AT was >250-fold higher than the corresponding values for DEBS AT domains. Although the AT from DSZS discriminated modestly against ACP domains from DEBS, it exhibited >40-fold higher activity *in trans* in the presence of these heterologous substrates than their natural AT domains. Our results highlight the opportunity for regioselective modification of a polyketide backbone by *in trans* complementation of inactivated AT domains. They also reinforce the need for more careful consideration of protein−protein interactions in the engineering of these assembly line enzymes.

Polyketide synthases (PKSs)[1] are multifunctional enzymes responsible for the production of polyketide natural products, many of which have important medicinal applications (*1*−*6*). Modular polyketide synthases, such as the 6-deoxyerythronolide B synthase (DEBS), are organized into modules comprised of multiple catalytic domains. To assemble short acyl-CoA precursors into complex polyketide products, each module contains minimally three functional domains: a β-ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP). The KS domain receives the growing polyketide chain from the upstream module and subsequently catalyzes formation of the C−C bond between this substrate and an ACP-bound extender unit that is selected by the AT domain.

Many biosynthetic engineering strategies for the modification of polyketides are predicated on the apparent modularity of these multimodular assembly lines. For example, two different modular approaches have been proposed for the incorporation of unnatural extender units into the polyketide backbone. In one approach, the AT domain of a target module is replaced by a naturally occurring homologue with a distinct extender unit specificity (*7*−*9*). An alternative approach involves complementation of a module harboring a site-specifically inactivated AT with a stand-alone AT protein having the desired extender unit specificity (*10*). In practice, the efficiency of both approaches rests upon the extent to which the protein−protein interactions necessary for structural integrity and catalytic activity of the engineered module are preserved as a result of introducing a heterologous AT. Here we attempt to pinpoint these protein−protein interactions and to quantify their contributions to acyl transfer kinetics.

Our investigations into specific protein−protein interactions involving the AT domain took advantage of the development of a multidomain [KS][AT] construct harboring catalytically active KS and AT domains as well as three flanking "linker" sequences (*11*, *12*), a catalytically active AT construct that includes flanking linkers at both the N- and C-termini (*13*), and an active ACP construct without any linkers (*12*). The X-ray structures of prototypical [KS][AT] constructs (*14*, *15*) and a solution NMR structure of a prototypical ACP from DEBS (*16*) have been solved (Figure 1). For comparison purposes, we have also studied a representative stand-alone AT protein from an "AT-less" multimodular PKS, the disorazole synthase (DSZS) (*17*, *18*). The DSZS and homologous synthases are comprised of modules lacking dedicated AT domains; instead, extender units are transacylated onto all of the ACP domains of these synthases

[†]This research was supported by grants from the NIH (GM087934 to C.K. and GM022172 to D.E.C.), by a Stanford Graduate Fellowship to A.Y.C., and by a National Science Scholarship from the Agency of Science, Technology, and Research (A*STAR), Singapore, to F.T.W.

*Corresponding author. E-mail: Khosla@stanford.edu. Phone: (650) 723-6538. Fax: (650) 725-7294.

[1]Abbreviations: AT, acyltransferase; CoA, coenzyme A; ACP, acyl carrier protein; DEBS, 6-deoxyerythronolide B synthase; DSZS, disorazole synthase; KS, β-ketosynthase; PKS, polyketide synthase; TE, thioesterase; DTT, dithiothreitol; LC-MS/MS, liquid chromatography tandem mass spectrometry; TCEP, tris(2-carboxyethyl)phosphine; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

Published on Web 11/18/2009

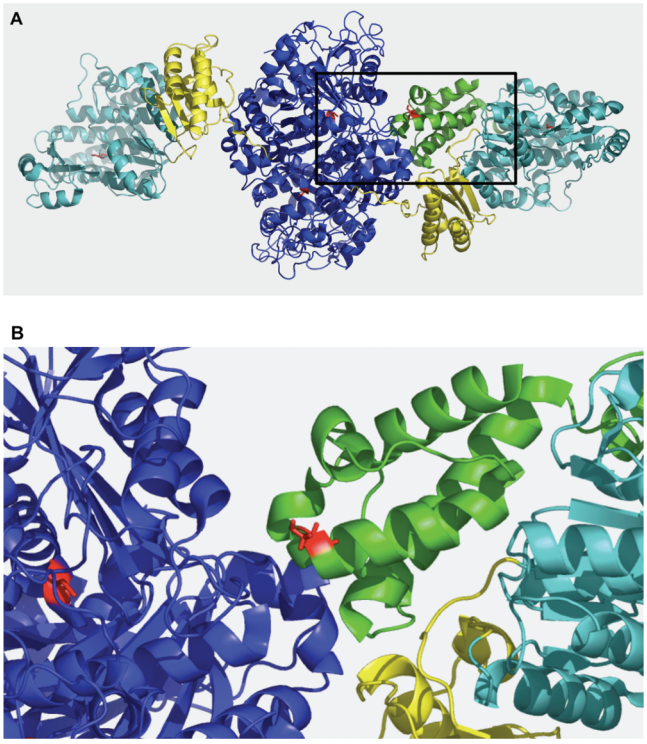

FIGURE 1: PatchDock model for the NMR solution structure of the ACP2 domain of DEBS (PDB ID 2JU2 (*16*)) docked onto the X-ray crystal structure of the [KS3][AT3] fragment of DEBS (PDB ID 2QO3 (*15*)). The [KS][AT] protein forms a homodimer. The KS domain, AT domain, and linkers are shown in blue, cyan, and yellow, respectively. The ACP domain is shown in green. The active site residues of the KS, AT, and ACP domains are highlighted in red. (B) is the enlarged view of boxed area in (A).

by a single stand-alone AT protein (*19*, *20*). Our results reveal that AT−ACP interactions are both specific and variable and therefore warrant more careful consideration in designing efficient strategies for engineering polyketide biosynthesis.

EXPERIMENTAL PROCEDURES

Reagents and Chemicals. DL-[2-*methyl*-^{14}C]Methylmalonyl-CoA and [^{14}C]malonyl-CoA were from American Radiolabeled Chemicals. All other chemicals were from Sigma. SDS−PAGE gradient gels were from Bio-Rad and Invitrogen. Ni-NTA affinity resin was from Qiagen. The HiTrap-Q anion-exchange column was from Amersham Pharmacia.

Construction and Expression of Discrete AT Domains from DEBS and DSZS. All AT domains from DEBS were expressed as stand-alone proteins with their KS-to-AT and post-AT linkers (*13*). The DNA sequence encoding DEBS AT1 was amplified by PCR as an *Nde*I−*Eco*RI fragment using primers 5′-AAAAAACATATGCAGGTCGTCGAAGGCGAGCGG-3′ and 5′-TTTTGAATTCACGCGGTAGCGCAGCGCGGAAACC-3′. This *Nde*I−*Eco*RI fragment was cloned into the pET28 expression vector to yield plasmid pAYC129. The DNA sequence encoding DEBS AT2 was constructed similarly with primers 5′-AAAACATATGGAACCGGAGCCCGTGCCGCAACC-3′ and 5′-TTTTGAATTCACCCGGTAGAACCAGCCGTCGAGC-3′ to yield plasmid pAYC130. The DNA sequence encoding DEBS AT4 was amplified similarly with primers 5′-AAAAAACATATGGCAGAGCAGGAGGCCGCCCGC-3′ and 5′-TTTTGAATTCACCTCCGCGCCGCGCACGCCCAGC-3′ to yield plasmid pAYC131. The DNA sequence encoding DEBS AT5 was amplified similarly with primers 5′-AAAAACATATGGAGGCCGACGAGCCCGAGCCGG-3′ and 5′-TTTTGAATTCACCTGGTAGCGCCAGTCGTCGTCC-3′ to yield plasmid pAYC132. The DNA sequence encoding DEBS AT6 was constructed similarly with primers 5′-AAAACATATGGAGCCCGAGCCGCTGCCGGAACC-3′ and 5′-TTTTGAATTCACGCGGTAGCGGCTGTCGGCGAGC-3′ to yield plasmid pAYC133. All five plasmids were introduced into *Escherichia coli* BL21(DE3) via transformation. Proteins were expressed and purified using the previously described protocol (*11*). The expression and purification of DEBS AT3 with the post-AT linker (pAYC23) have been previously described (*13*).

Two versions of the AT gene from DSZS were constructed. The DNA sequence encoding DSZS AT without the post-AT linker (denoted DSZS AT_S) was amplified by PCR as an *Nde*I−*Eco*RI fragment using primers 5′-AAAAACATATGAAAGCATACATGTTTCCCGGGC-3′ and 5′-TTTTTGAATTCGCGACGACGAGGGGCTGGGCG-3′. Cosmid pKOS254-190.1 was used as the template (*17*). This *Nde*I−*Eco*RI fragment was cloned into the pET21 expression vector to yield plasmid pFW3. The DNA sequence encoding DSZS AT with the post-AT linker (denoted DSZS AT_L) was constructed similarly with primers 5′-AAAAACATATGAAAGCATACATGTTTCCCGGGC-3′ and 5′-TTTTTGAATTCGCCGCGGCGCTGCCCAGCACC-3′ to yield plasmid pFW5. The plasmids were transformed into *E. coli* BL21(DE3) for protein expression. Proteins were expressed and purified using the previously described protocol (*11*).

DNA sequences encoding various AT6 constructs ($AT6_{AT}$, $AT6_{PAT}$, $AT6_{KSAT}$) were amplified from pAYC133 as *Nde*I−*Eco*RI fragments and cloned into the pET21 expression vector. The DNA sequence encoding the core AT6 domain, excluding the C-terminal helix from the AT domain and flanking linkers (denoted $AT6_{AT}$), was constructed with primers 5′-AAAAACATATGGGCGTGGTTTTCGTCTTCCCAGGTC-3′ and 5′-TTTTTGAATTCGCCCCGATGGCCACGGCGTCCGC-3′ to yield pFW59. The DNA sequence encoding the core AT6 domain with the C-terminal flanking (post-AT) linker (denoted $AT6_{PAT}$) was constructed similarly with primers 5′-AAAAACATATGGGCGTGGTTTTCGTCTTCCCAGGTC-3′ and 5′-TTTTTGAATTCGCGCGGTAGCGGCTGTCGGCGAG-3′ to yield pFW60. The DNA sequence encoding the core AT6 domain with the N-terminal flanking KS-to-AT linker, excluding the C-terminal helix and post-AT linker (denoted $AT6_{KSAT}$), was constructed with primers 5′-AAAAACATATGGAGCCCGAGCCGCTGCCGGAA-3′ and 5′-TTTTTGAATTCGCCCCGATGGCCACGGCGTCCGC-3′ to yield plasmid pFW61.

DNA sequences encoding the rest of the AT6 constructs ($AT6_{small}$ and $AT6_{small,PAT}$) were amplified from pAYC133 as *Nde*I−*Eco*RI fragments and cloned into the pET28 expression vector. The DNA sequence encoding the core AT6 domain with the proximal 15 residues from the KS-to-AT linker and less the post-AT linker and its C-terminal helix (denoted as AT_{small}) was constructed with primers 5′-AAAAACATATGGGCGTTGCGGCTCCCGGTGCCA-3′ and 5′-TTTTTGAATTCGCCCCGATGGCCACGGCGTCCGC-3′ to yield plasmid pFW75. The DNA sequence encoding AT6 along with the 15 residues from the KS-to-AT linker and the post-AT linker (denoted as $AT_{small,PAT}$) was constructed with 5′-AAAAACATATGGGCGTTGCGGCTCCCGGTGCCA-3′ and 5′-TTTTTGAATTCGCGCGGTAGCGGCTGTCGGCGAG-3′ to yield plasmid pFW76. The plasmids were transformed into *E. coli* BL21(DE3) for protein

expression. Proteins were expressed and purified using the previously described protocol (*11*).

Construction and Expression of DEBS [KS][AT] Didomains. The expression and purification of [KS3][AT3] (pAYC02) and [KS6][AT6] (pAYC11) didomains from DEBS have been previously described (*11*).

Construction and Expression of AT^o Mutants of DEBS Module 3+TE and Module 6+TE. The construction of plasmids encoding module 3 (pRSG34) and module 6 (pRSG54) with the thioesterase (TE) domain covalently attached has been previously described (*21*). Module 3/AT^o+TE: To inactivate the acyltransferase domain of module 3+TE (pRSG34), the AT active site Ser was mutated into Ala using primers 5′-GCGGTCGTGGGGCACGCGCAGGGCGAGATCG-3′ and 5′-CGATCTCGCCCTGCGCGTGCCCCACGACCGC-3′ to yield plasmid pAYC136. Module 6/AT^o+TE: Similarly, primers 5′-GCCGTCATCGGCCATGCGCAGGGCGAGATCG-3′ and 5′-CGATCTCGCCCTGCGCATGGCCGATGACGGC-3′ were used to inactivate the AT domain of module 6+TE (pRSG54) to yield plasmid pAYC138. The plasmids, along with pRSG34 and pRSG54, were transformed into *E. coli* BAP1 (*22*) to produce the desired holo proteins with pantetheinylated ACP domains (*23*). Proteins were expressed and purified using the previously described protocol (*11*).

Construction and Expression of Discrete DEBS and DSZS Acyl Carrier Proteins. A N-terminal hexahistidine-tagged DEBS ACP6 was constructed for ACP6 protein expression. The ACP6 sequence was amplified from pRSG54 with primers 5′-AAAAACATATGGCGGCCCCGGCGCGGGAGATG-3′ and 5′- TTTTTGAATTCTCAGAGCTGCTGTCCTATGTGGTC-3′. The resulting *Nde*I−*Eco*RI fragment was then inserted into a pET28 expression vector to yield pFW55.

Expression plasmid for DEBS ACP3 was constructed from a previously described ACP3 expressing construct, pVYA5 (*12*). An *Xba*I−*Nde*I cassette, containing a hexahistidine tag upstream of a flag tag, was prepared by thermally annealing primer 5′-AAAAAATCTAGAAATAATTTTGTTTAACTTTAACTTTAAGAAGGAGATATACCATGCCGCATCATCATCATCATCACAGCAGCGACTACAAAGACGATGACGACAAGCTGCATATGAAAAAA-3′ and its complementary primer for 5 min at 95 °C. This cassette was then digested and inserted into pVYA5, replacing the corresponding *Xba*I−*Nde*I fragment to yield pSHIV9. The original *Xba*I−*Nde*I fragment in pVYA5 contains only an N-terminal hexahistidine tag.

The sequence encoding DSZS ACP1 was amplified from cosmid pKOS254-190.4 (*17*) using primers 5′-AAAAACATATGGCGCCTGCAGGGGCAGGACAG-3′ and 5′-TTTTTGAATTCGCTGCCGACCTCGCGGGGACGCG-3′. The *Nde*I−*Eco*RI fragment was then inserted into the pET21 expression vector to yield pFW50.

The plasmids were transformed into *E. coli* BAP1 (*22*) to produce the desired holo proteins with pantetheinylated ACP domains (*23*). Proteins were expressed and purified using the previously described protocol (*11*).

Limited Proteolysis of Stand-Alone DEBS AT6. Stand-alone DEBS AT6 with both linkers (30 μM, in 100 mM phosphate at pH 7.2) was incubated with varying concentrations of trypsin with 2.5 mM TCEP at room temperature for 30 min. Trypsin (4.2−42 μM in ddH_2O) was added to make up final molar trypsin:AT protein ratios of 1:5, 1:15, 1:50, and 1:100. The total reaction volume is 10 μL. The reaction was quenched in 5 μL of SDS−PAGE sample buffer. The proteins were separated on a SDS−PAGE gel, and the protein of interest was eluted from the gel and subjected to in-gel trypsin digestion and LC-MS/MS analysis.

In-Gel Trypsin Digestion and LC-MS/MS Analysis. The excised gel band from above was cut into 1 × 1 mm pieces before being reduced with 50 mM DTT and alkylated with 100 mM acrylamide. The gel was then destained with 50 mM 1:1 acetonitrile:ammonium bicarbonate solution. This was dried on a Speedvac and reconstituted in 12 ng/μL of trypsin, in the presence of protease Max. The mixture was incubated at 50 °C for 1 h before being spun down at 12*g* for 30 s. Peptides were extracted and dried using a Speedvac. This was reconstituted in 10 μL of 2% acetonitrile/water with 0.1% formic acid.

The sample was then injected onto Eksigent NanoLC-2D capillary HPLC with a self-packed C18 reverse-phase column. MS spectra were collected with an LCQ Deca XP Plus (Thermo Scientific) that was set to the data-dependent acquisition mode to analyze the top three abundant peaks. Data were analyzed using the SORCERER processor (SageN Research) with the Swiss-PROT database. An eight-node cluster for accelerated Sequest database searching was used.

Acylation of Trypsin-Digested Fragments of Stand-Alone AT6. Stand-alone DEBS AT6 (30 μM, in 100 mM phosphate at pH 7.2) was digested, as described above, at molar trypsin:AT protein ratios of 1:5, 1:15, 1:50, and 1:100 with 2.5 mM TCEP at room temperature. No trypsin was added to the control reaction. The mixture was incubated with 200 μM [^{14}C]methylmalonyl-CoA after 30 min of trypsin digestion. The total volume of the reaction mixture is 10 μL. The reaction was quenched after 1 min of [^{14}C]methylmalonyl CoA addition with 5 μL of SDS−PAGE sample buffer. The samples were loaded on a SDS−PAGE gel. The gel was dried using a Bio-Rad gel-drying system and analyzed using a Packard phosphorimager.

Acylation and Transacylation of Module 3+TE, Module 6+TE Mutants, and Stand-Alone ACPs in the Presence of Discrete AT Proteins. For kinetic assays, the reaction volume corresponding to each time point was 10 μL. All ACPs used were in their holo form.

Transacylation of DEBS ACP6 in the Presence of AT6 Constructs. AT6 constructs (20 μM, in 100 mM phosphate at pH 7.2) were incubated at room temperature with 2.5 mM TCEP and 200 μM [^{14}C]methylmalonyl-CoA with discrete DEBS ACP6 protein (80 μM). Nine microliters of the mixture was removed at 1 and 5 min time points and quenched with 5 μL of SDS−PAGE loading buffer and loaded on a SDS−PAGE gel. The gel was dried using a Bio-Rad gel-drying system and analyzed using a phosphorimager.

Acylation and Transacylation of Module 3+TE and Module 6+TE Mutants in the Presence of Discrete AT Proteins. Wild-type or mutant modules (30 μM, in 100 mM phosphate at pH 7.2) were incubated for 10 min on ice with 2.5 mM TCEP and 30 μM [^{14}C]methylmalonyl-CoA (for DEBS ATs) or [^{14}C]malonyl-CoA (for DSZS ATs), alone or in conjunction with a discrete AT protein (30 μM). Samples were quenched with 5 μL of SDS−PAGE loading buffer and loaded on a SDS−PAGE gel. Samples were processed as described above.

Transacylation Rates of Modular ACPs by Discrete AT Proteins. Module 3/AT^o+TE (20−100 μM, in 100 mM phosphate at pH 7.2) were incubated at room temperature with 0.03 μM discrete DSZS AT_S, 2.5 mM TCEP, and 200 μM [^{14}C]malonyl-CoA. Nine microliters of the mixture was removed at intervals of 1 min and quenched with 5 μL of SDS−PAGE

loading buffer. The samples were loaded on a SDS−PAGE gel and processed as described above.

Transacylation Rates of Discrete ACPs by DSZS AT$_S$. Discrete ACPs (20−150 μM, in 100 mM phosphate, pH 7.2) were incubated at room temperature with DSZS AT$_S$ (0.003 μM for cognate DSZS ACP1 transacylation, 0.03 μM for transacylation of DEBS ACPs), 2.5 mM TCEP, and 200 μM [^{14}C]malonyl-CoA. Nine microliters of the mixture was removed at 1 min intervals and quenched with 5 μL of SDS−PAGE loading buffer and loaded on a SDS−PAGE gel. Samples were processed as described above.

Transacylation Rates of Discrete ACPs by DEBS ATs and [KS][AT]s. Discrete ACPs (20−100 μM, in 100 mM phosphate, pH 7.2) were incubated at room temperature with discrete DEBS ATs (and [KS][AT]s), 2.5 mM TCEP, and 200 μM [^{14}C]methylmalonyl-CoA. Depending on the ACP−AT combination used, 0.5−10 μM DEBS AT (or [KS][AT]) was used for each series of tranacylation assays. Nine microliters of the mixture was removed at intervals of 2 min and quenched with 5 μL of SDS−PAGE loading buffer and loaded on a SDS−PAGE gel. Samples were processed as described above.

Transacylation assays for k_{cat}/K_M calculations (Table 1) were repeated at least three independent times.

Docking Models of DEBS AT and ACP. Homology models for ACP3, ACP6, and AT6 from DEBS were derived using the I-TASSER server (*24*−*26*).

In silico docking experiments involving AT (or [KS][AT]) and ACP proteins were performed using PatchDock (*27*, *28*) and then further refined and reranked with FireDock (*29*, *30*). For PatchDock simulations, ATs and ACPs were set as receptor and ligand, respectively, under default complex-type settings.To simulate ACP docking on AT, a maximum distance constraint of 25 Å was placed between AT and ACP active site Ser residues. This was to ensure that the catalytic AT Ser residue was not beyond the reach of the approximately 18 Å long pantetheinylated arm of the holo-ACP.

RESULTS

Design and Production of Stand-Alone AT Constructs from DEBS. Analysis of the X-ray structures of [KS3][AT3] and [KS5][AT5] from DEBS led to the design of a prototypical stand-alone AT from module 3 of DEBS (*13*). From an initial analysis of two AT3 constructs, both of which contain the N-terminal KS-to-AT linker with or without the C-terminal post-AT linker, it was observed that the absence of the post-AT linker negatively affects KS activity but not AT activity (*13*). In this study we constructed and analyzed a panel of stand-alone AT constructs from module 3 as well as other DEBS modules. As shown in Figure 2, these constructs included the KS-to-AT linker, the post-AT linker, neither linker, or both linkers.

AT constructs harboring both the KS-to-AT linker and the post-AT linker from DEBS modules 3, 4, 5, and 6 (denoted as AT3, AT4, AT5, and AT6, respectively) were expressed in *E. coli* as soluble proteins at approximately 3, 10, 15, and 50 mg of purified protein per liter of culture, respectively. No soluble protein was obtained for the AT domains from modules 1 or 2. To understand the contributions of the N- and C-terminal linkers to AT activity and specificity, we also attempted to produce truncated derivatives of AT6 that lacked either or both linkers (Figure 2). Our initial attempt at determining linker boundaries for these truncated constructs entailed rational design based on our knowledge of previously solved crystal structures of [KS][AT] (*14*, *15*). Because the AT domain harbors a conserved C-terminal helix that interacts with the KS-to-AT linker, some of our constructs also lacked this helix. AT6 constructs without the KS-to-AT linker (denoted AT6$_{PAT}$) and without both linkers (denoted as AT6$_{AT}$) were produced with a yield of 30 and 7 mg per liter of culture, respectively. Soluble proteins were not obtained for the AT6 construct without the post-AT linker and C-terminal helix (denoted as AT6$_{KSAT}$).

A Truncated Acyltransferase Construct Developed through Limited Proteolysis. Analysis of the above constructs (detailed below) revealed that some proteins had comparable activity to the reference stand-alone AT constructs, whereas the activity of others was substantially attenuated. Together, these results illustrated the limits of structure- and sequence-based engineering of stand-alone enzymes that are ordinarily part of larger multidomain proteins. We therefore turned to limited proteolysis as an alternate strategy for identifying an active truncated derivative of AT6, which included the KS-to-AT linker and post-AT linker (Figure 2). After treatment of the AT6 protein with varying amounts of trypsin, [^{14}C]methylmalonyl-CoA was added to the partially digested mixture in order to identify fragments competent for self-acylation. The corresponding

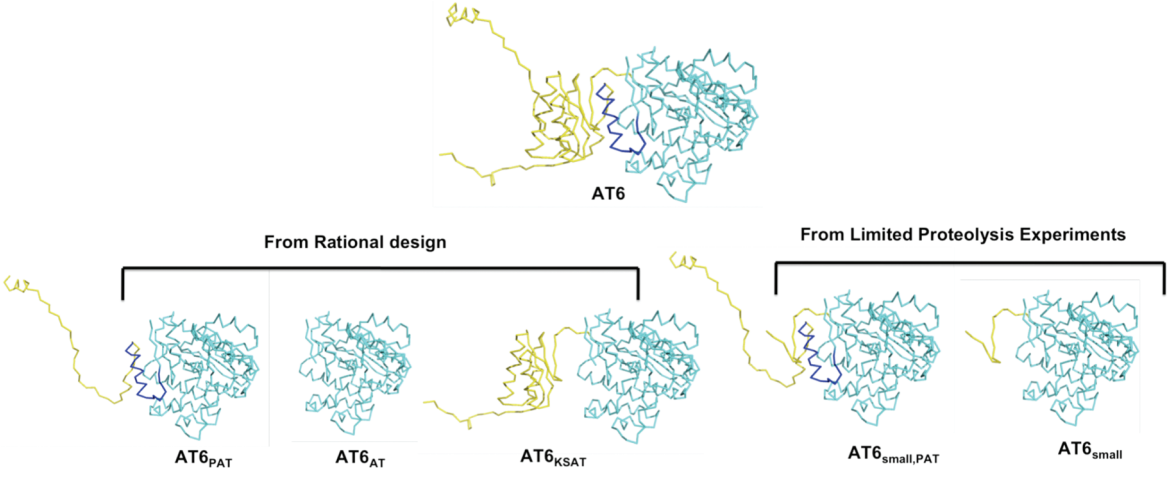

FIGURE 2: Ribbon models of truncated AT6 constructs from rational design and limited proteolysis experiments. The model for AT6 was generated using the I-TASSER server (*24*−*26*). The AT domain, C-terminal helix, and both linkers are shown in cyan, dark blue, and yellow, respectively, and are all present in construct "AT6".

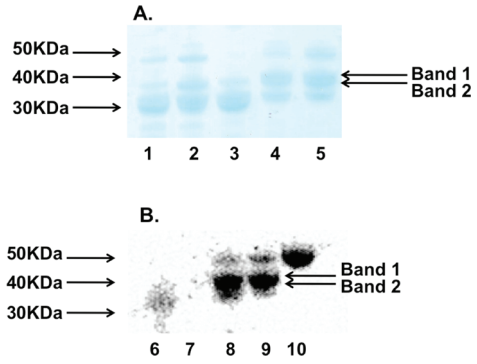

FIGURE 3: Limited proteolysis of AT6 and identification of active truncated variants. (Panel A) Partial trypsin digestion of DEBS AT6 (30 min at room temperature). AT6 was incubated with molar ratios of trypsin:AT6 ranging from 1:5 (lane 1), 1:15 (lane 2), 1:15 (lane 3, incubated for 1 h), 1:50 (lane 4), to 1:100 (lane 5). (Panel B) Radio SDS−PAGE analysis of trypsin-digested AT6 fragments incubated with [^{14}C]methylmalonyl-CoA. AT6 was incubated with molar ratios of trypsin:AT6 of 1:5 (lane 6), 1:15 (lane 7), 1:50 (lane 8), and 1:100 (lane 9). AT6 with no trypsin added was used as a control (lane 10).

protein bands from an SDS−PAGE gel were extracted and subjected to LC-MS/MS analysis in order to identify the relevant N- and C-termini (Figure 3). In particular, the sequence of the 40 kDa proteolytic fragment was analyzed. LC-MS/MS generated peptide sequences (> 95% confidence) were used to determine its boundaries. Overall, 67% coverage of the predicted amino acid sequence of this fragment (predicted MW 39422 Da) was obtained. More importantly, continuous coverage of the N- and C-termini was obtained over 37 and 59 residues, respectively. Whereas unambiguous border identification is not possible by LC-MS/MS, these results motivated us to test the properties of the predicted construct, which starts at the amino acid sequence GVAAPGATTGTA (denoted as $AT6_{small,PAT}$, Figure 2). The construct was expressed as a soluble protein and purified in a yield of 6 mg per liter of culture. As seen in Figure 4, robust self-acylation of this protein was observed in the presence of [^{14}C]methylmalonyl-CoA. The corresponding protein with the same N-terminus but lacking the C-terminal helix was obtained as a soluble protein ($AT6_{small}$) but was only weakly acylated by [^{14}C]methylmalonyl-CoA.

Acylation and Transacylation Activities of Alternative Stand-Alone AT Constructs. A functional AT construct from DEBS should be able to self-acylate as well as to transacylate its cognate ACP domain, while also retaining specificity for methylmalonyl-CoA. To investigate these properties of the alternative forms of AT6 as well as the truncated AT6 derivatives that were obtained via limited proteolysis (Figure 2), assays with [^{14}C]methylmalonyl-CoA and ACP6 were performed. Radio SDS−PAGE (Figure 4) revealed that $AT6_{AT}$ and $AT6_{small}$ self-labeled more weakly than AT6. Similarly, $AT6_{AT}$ and $AT6_{small}$ were unable to transfer the extender unit onto the ACP (Figure 4). $AT6_{PAT}$, $AT6_{small,PAT}$, and AT6 showed a comparable ability to acylate the ACP.

Expression and Purification of Recombinant DSZS AT. The disorazole synthase (DSZS) is a member of an "AT-less" type I polyketide synthase with a stand-alone AT that transacylates extender units onto all ACP domains of this PKS. This AT is part of a larger (91 kDa) protein that also includes a C-terminal putative oxidoreductase domain separated by a linker of unknown function (*17*). Based on sequence alignments with known acyltransferases from structurally characterized bacterial fatty

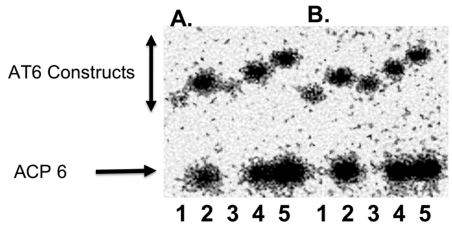

FIGURE 4: Radio SDS−PAGE analysis of transacylation and acylation activities of truncated DEBS AT6 constructs with DEBS ACP6. [^{14}C]Methylmalonyl-CoA was incubated for 1 min (panel A) and 5 min (panel B) with $AT6_{AT}$ (lane 1), $AT6_{PAT}$ (lane 2), $AT6_{small}$ (lane 3), $AT6_{small,PAT}$ (lane 4), and AT6 (lane 5) with DEBS ACP6.

acid synthases (*31*), two versions of the disorazole AT were engineered, one with the ca. 100 residue linker (DSZS AT_L) and the other without this linker (DSZS AT_S). Both constructs were expressed as soluble proteins at yields of 80−110 mg of purified protein per liter of culture.

Expression, Purification, and Characterization of ACP Proteins. To measure the transacylation rate of an ACP by an AT, the kinetics of radiolabel transfer from [^{14}C]malonyl-CoA (for DSZS AT) or [^{14}C]methylmalonyl-CoA onto each ACP was quantified. The ACP3 and ACP6 domains from DEBS were expressed using similar constructs as described before (*12*). DSZS is predicted to have eight distinct ACP domains, one in each module except for module 7, which has two ACPs (*17*). Based on sequence alignments with type I and II PKS acyl carrier proteins, a stand-alone acyl carrier protein (DSZS ACP1) was constructed from module 1 of DSZS. This construct was expressed in *E. coli* BAP1 (*22*) and purified as soluble holo-ACP at a yield of 1 mg per liter of culture.

Expression and Purification of AT^o Mutants of DEBS Module 3+TE and Module 6+TE. To attempt complementation of intact modules by stand-alone AT domains, mutants of DEBS modules 3 and 6 with inactive AT domains were engineered (denoted module 3/AT^o+TE and module 6/AT^o+TE, respectively). In each case AT inactivation was achieved via a Ser → Ala mutation at the AT active site. Both mutants were produced as soluble proteins at approximately 50 mg of purified protein per liter of culture.

ACP Specificity of Acyltransferases. In an intact DEBS module, an AT domain catalyzes acyl transfer from methylmalonyl-CoA to the ACP domain within the same module via a covalent acyl-AT intermediate. As an initial assessment of whether DEBS AT domains are specific for their natural ACP partners, the AT domains from modules 3−6 of DEBS (with intact KS-to-AT and post-AT linkers) were individually incubated with DEBS module 3/AT^o+TE or module 6/AT^o+TE derivatives of DEBS in the presence of [^{14}C]methylmalonyl-CoA. As seen in Figure 5, although the AT proteins were comparably labeled by the methylmalonyl extender unit, different combinations of modules and stand-alone AT proteins showed different degrees of AT-to-ACP transacylation capability. Moreover, the transacylation preferences for the two modules were distinct from one another.

We therefore quantified the specificity of different AT constructs from DEBS modules 3 and 6 for ACP3 and ACP6. A representative transacylation rate versus [ACP] data set is plotted in Figure 6, the slope of which yielded an apparent k_{cat}/K_M value. The results of all such experiments are summarized in Table 1. For both the [KS][AT] and the AT6 constructs, a 10−20-fold preference is observed for the cognate ACP. Moreover, the rate

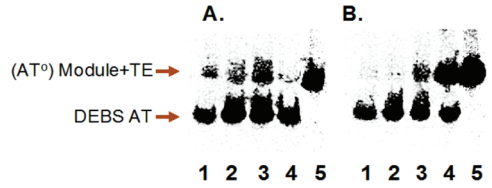

FIGURE 5: Radio SDS−PAGE analysis of acylation and transacylation activities of stand-alone AT proteins from DEBS using AT° mutants of modules 3 and 6 as ACP substrates. [^{14}C]Methylmalonyl-CoA was incubated with module 3/AT°+TE (panel A, lanes 1−4) or module 6/AT°+TE (panel B, lanes 1−4) in the presence of alternative acyltransferases. Lanes in each panel: 1, DEBS AT3; 2, DEBS AT4; 3, DEBS AT5; 4, DEBS AT6; 5, the corresponding wild-type module with an active AT (control).

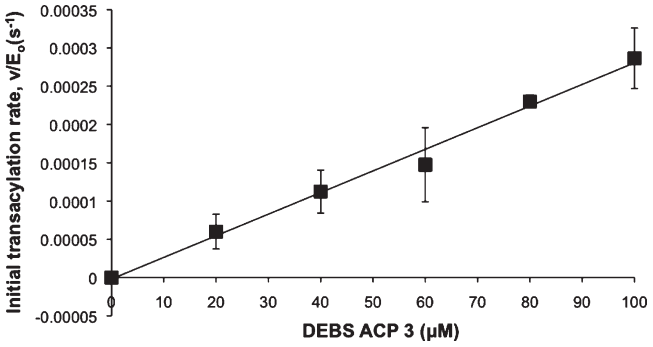

FIGURE 6: A typical transacylation rate (s^{-1}) vs DEBS [ACP] (μM) plot. 20−100 μM DEBS ACP3 was incubated with 10 μM DEBS AT6. [^{14}C]Methylmalonyl-CoA was held constant at 200 μM. Initial rates were determined at room temperature at least three independent times. k_{cat}/K_M [(μM s)$^{-1}$] was obtained through linear fitting of the data.

Table 1: Apparent k_{cat}/K_M Values for Transacylation of ACPs by Alternative AT Constructs[a]

AT	ACP	k_{cat}/K_M [(M s)$^{-1}$]
[KS3][AT3]	ACP3	390 ± 50
	ACP6	21 ± 8
[KS6][AT6]	ACP3	6 ± 1
	ACP6	86 ± 14
AT6	ACP3	3 ± 0.3
	ACP6	69 ± 7
AT6$_{small,PAT}$	ACP3	4 ± 1
	ACP6	19 ± 3
DSZS AT$_S$	DSZS ACP1	(87 ± 13) × 10^3
	ACP3	(14 ± 3) × 10^3
	Module 3/AT°+TE	(9 ± 1) × 10^3

[a]ATs and ACPs are from DEBS, unless otherwise indicated.

constants are comparable in magnitude for the [KS6][AT6] and the AT6 constructs. In contrast, both the activity and specificity of the smaller AT construct AT6$_{small,PAT}$ were modestly reduced. The latter construct was chosen for kinetic analysis, because it was the most active of the truncated AT6 derivatives analyzed in this study.

We also performed similar measurements with the AT protein from the DSZS. To test the acylation and transacylation activities of DSZS AT$_S$ and DSZS AT$_L$, each protein was incubated with [^{14}C]malonyl-CoA in the presence of the AT° mutants of DEBS module 3+TE or module 6+TE. As seen in Figure 7, both enzymes catalyzed self-acylation and transacylation to the ACP domain with comparable efficiency. Thus, the 100-residue post-AT linker did not appear to have a significant effect on either

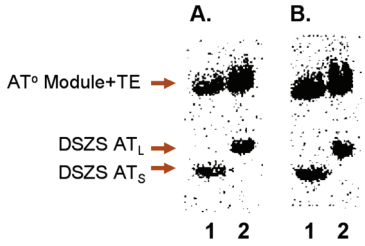

FIGURE 7: Radio SDS−PAGE analysis of DSZS AT activity. [^{14}C]Malonyl-CoA was incubated with DEBS module 3/AT°+TE (panel A) or module 6/AT°+TE (panel B) in the presence of recombinant DSZS AT constructs. The discrete AT labeled the ACP domain of each module. Lanes: 1, DSZS AT$_S$ (DSZS AT without post-AT linker); 2, DSZS AT$_L$ (DSZS AT with post-AT linker).

component of DSZS AT activity. When [^{14}C]methylmalonyl-CoA was used instead of [^{14}C]malonyl-CoA, no acyl transfer to either of the two DEBS modules was detected, although some labeling of the DSZS AT proteins could be observed (data not shown). Most remarkably, the activity of DSZS AT$_S$ was considerably higher against all ACPs evaluated than even that of the cognate AT (Table 1). Thus, although DSZS AT has higher specificity for its own substrates, it is still able to efficiently charge heterologous ACP domains.

It should be noted that, in all of the above experiments, AT-ACP specificity was quantified by measuring and comparing the overall specificity constant (k_{cat}/K_M) of the transacylation reaction. The observed differences in this parameter should not be mistaken for inherent AT-ACP affinity, which would be more appropriately investigated via protein binding experiments such as isothermal titration calorimetry.

DISCUSSION

In a multimodular PKS, the starter unit is elongated by decarboxylative condensation with extender units as it progresses through the enzymatic assembly line. These extender units are selected by acyltransferase (AT) domains that act as gatekeepers in each PKS module. The importance of the AT to the stringent incorporation of a specific extender unit in the synthesis of polyketide building blocks makes it vital that the mechanism and structure of these domains be well elucidated in order to develop efficient strategies for the regiospecific engineering of extender unit incorporation in polyketide biosynthesis. Here we have quantified the ACP specificity of representative AT domains derived from multimodular PKSs and have examined the role of flanking linkers on both AT activity and specificity.

A key finding of our study is that, whereas stand-alone AT constructs harboring intact N- and C-terminal linkers have properties comparable to the corresponding AT domains that are part of larger multidomain systems, deletion of either linker has a deleterious effect on both the activity and specificity of the mutant AT. The most active of the truncated constructs consisted of the core AT domain flanked by the proximal 15 C-terminal residues of the KS-to-AT linker and the post-AT linker (AT6$_{small,PAT}$). Nonetheless, the activity of even that truncated AT was measurably attenuated. It thus appears that the intact linkers play an important role in ACP recognition and/or catalysis of acyl transfer. The mechanistic basis for such an effect remains to be elucidated.

Our investigations into the ACP specificity of AT domains focused on DEBS ACP3 and ACP6 as substrates, because these

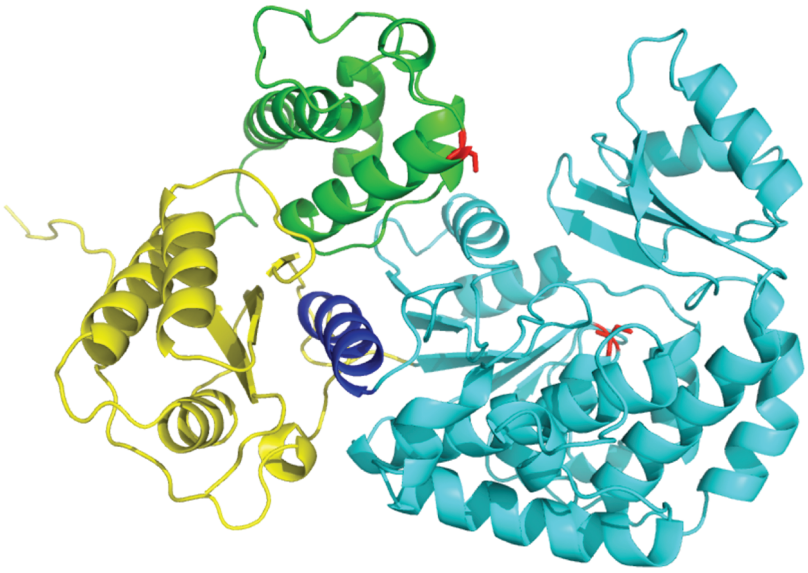

FIGURE 8: PatchDock docking model for AT3 with linkers (truncated from PDB ID 2QO3 (*15*)) and DEBS ACP3 (homology model generated by I-TASSER). The AT domain, linkers, C-terminal helix, and ACP domain are shown in cyan, yellow, dark blue, and green, respectively. Active site Ser residues are shown in red.

two ACPs are anticipated to have the most different electrostatic surfaces among the six DEBS ACPs (*16*). Similarities in k_{cat}/K_M values between the [KS][AT] didomain and stand-alone AT constructs (Table 1) suggested that the presence of the KS domain does not significantly influence AT-ACP recognition. This situation is radically different from KS-ACP recognition, which is influenced by the presence of both the AT domain and the post-AT linker (*13*). We note that the above analysis of AT-ACP specificity is based on *in trans* analysis of enzyme activity that ordinarily functions *in cis*. Earlier studies on nonribosomal peptide synthetases (*32*) and polyketide synthases (*7*) have observed minimal effects on the catalytic activities of heterologous adenylation and AT domains, respectively, when introduced in an *in cis* configuration. Thus, the actual consequence of suboptimal AT−ACP interactions on the turnover number of the overall assembly line remains to be elucidated.

Protein−protein docking simulations were performed with PatchDock (*27*, *28*) using the known structures of DEBS [KS3][AT3] (PDB ID 2QO3) and ACP2 (PDB ID 2JU2) as well as a homology model of ACP3 (generated via the I-TASSER server (*24*−*26*)). Several of the lowest energy docking models emerging from this exercise placed the ACP on the side of the AT protein that opens to the gorge leading to the AT active site serine. A representative AT3-ACP3 docking model is shown in Figure 8. In this model, helix I, helix II, and loop II regions of the ACP are docked onto the β strands of the KS-to-AT linker, the smaller subdomain of AT3, and the C-terminal helix of the AT. All other docking models generated by PatchDock simulations exhibited similar orientations between the AT and the ACP. The distance between the active sites was 23 Å, consistent with the interaction of the 18 Å phosphopantetheinyl arms with the AT active site Ser.

This model for AT-ACP interaction may explain the observed specificity of AT3 and AT6 for their cognate ACPs (Table 1). The proposed AT−ACP interactions involve docking of the ACP onto the KS-to-AT linker and C-terminal helix of the AT. The sequences of both regions are highly variable among the DEBS modules and highlight the importance of considering this interdomain interface in future biosynthetic engineering strategies for altering extender unit specificity of PKSs. Indeed, earlier work has already demonstrated that mutations in the C-terminal helix can influence AT substrate specificity (*33*).

Perhaps not surprisingly, the DSZS AT has a markedly higher (250−1000-fold) transacylation efficiency for its cognate DSZS ACP1 in comparison to the maximum activity of stand-alone AT domains from DEBS. The evolution of this high *in trans* AT activity is presumably facilitated by the intermodular nature of DSZS AT's mode of action on its cognate ACP substrates. It is nonetheless impressive that the rate at which DSZS AT catalyzes acyl transfer onto heterologous ACP domains is only modestly lower (10-fold) than acyl transfer to its cognate ACP domains. This observation is in stark contrast to an estimated ~5000-fold difference in the rate at which the malonyl-CoA:acyl carrier protein transacylase from the fatty acid synthase acylates its own ACP as compared to DEBS ACP2 and suggests that the DSZS acyltransferase may be a superior candidate for *in trans* complementation of engineered PKSs (*10*). This hypothesis is currently being tested in our laboratories.

In conclusion, this study highlights the importance of AT−ACP specificity in multimodular PKSs. Such interactions, which were overlooked in virtually all earlier attempts to engineer PKSs via domain substitution, must clearly be carefully considered in the design of structurally stable and catalytically active hybrid PKSs. Our results also suggest improved boundaries for AT domain substitution that include the flanking linker regions.

ACKNOWLEDGMENT

We thank the Vincent Coates Foundation Mass Spectrometry Laboratory at Stanford University (SUMS). We are also grateful to Chris Adams at SUMS for help with LC-MS/MS analysis. We thank Shiven Kapur for plasmid pSHIV9.

REFERENCES

1. Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., and Cane, D. E. (2007) Structure and Mechanism of the 6-Deoxyerythronolide B Synthase. *Annu. Rev. Biochem. 76*, 195–221.
2. Fischbach, M. A., and Walsh, C. T. (2006) Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chem. Rev. 106*, 3468–3496.

3. Hill, A. M. (2006) The Biosynthesis, Molecular Genetics and Enzymology of the Polyketide-Derived Metabolites. *Nat. Prod. Rep. 23*, 256–320.
4. Weissman, K. J., and Leadlay, P. F. (2005) Combinatorial Biosynthesis of Reduced Polyketides. *Nature Rev. Microbiol. 3*, 925.
5. Wenzel, S. C., and Müller, R. (2005) Formation of Novel Secondary Metabolites by Bacterial Multimodular Assembly Lines: Deviations from Textbook Biosynthetic Logic. *Curr. Opin. Chem. Biol. 9*, 447.
6. McDaniel, R., Welch, M., and Hutchinson, C. R. (2005) Genetic Approaches to Polyketide Antibiotics. 1. *Chem. Rev. 105*, 543.
7. Hans, M., Hornung, A., Dziarnowski, A., Cane, D. E., and Khosla, C. (2003) Mechanistic Analysis of Acyl Transferase Domain Exchange in Polyketide Synthase Modules. *J. Am. Chem. Soc. 125*, 5366–5374.
8. Ruan, X., Pereda, A., Stassi, D. L., Zeidner, D., Summers, R. G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M., Donadio, S., and Katz, L. (1997) Acyltransferase Domain Substitutions in Erythromycin Polyketide Synthase Yield Novel Erythromycin Derivatives. *J. Bacteriol. 179*, 6416–6425.
9. Oliynyk, M., Brown, M. J. B., Cortés, J., Staunton, J., and Leadlay, P. F. (1996) A Hybrid Modular Polyketide Synthase obtained by Domain Swapping. *Chem. Biol. 3*, 833–839.
10. Kumar, P., Koppisch, A. T., Cane, D. E., and Khosla, C. (2003) Enhancing the Modularity of the Modular Polyketide Synthases: Transacylation in Modular Polyketide Synthases Catalyzed by Malonyl-CoA:ACP Transacylase. *J. Am. Chem. Soc. 125*, 14307–14312.
11. Chen, A. Y., Schnarr, N. A., Kim, C. Y., Cane, D. E., and Khosla, C. (2006) Extender Unit and Acyl Carrier Protein Specificity of Ketosynthase Domains of the 6-Deoxyerythronolide B Synthase. *J. Am. Chem. Soc. 128*, 3067–3074.
12. Kim, C. Y., Alekseyev, V. Y., Chen, A. Y., Tang, Y., Cane, D. E., and Khosla, C. (2004) Reconstituting Modular Activity from Separated Domains of 6-Deoxyerythronolide B Synthase. *Biochemistry. 43*, 13892–13898.
13. Chen, A. Y., Cane, D. E., and Khosla, C. (2007) Structure-Based Dissociation of a Type I Polyketide Synthase Module. *Chem. Biol. 14*, 784–792.
14. Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7-Å Crystal Structure of a 194-kDa Homodimeric Fragment of the 6-Deoxyerythronolide B Synthase. *Proc. Natl. Acad. Sci. U.S.A. 103*, 11124–11129.
15. Tang, Y., Chen, A. Y., Kim, C. Y., Cane, D. E., and Khosla, C. (2007) Structural and Mechanistic Analysis of Protein Interactions in Module 3 of the 6-Deoxyerythronolide B Synthase. *Chem. Biol. 14*, 931–943.
16. Alekseyev, V. Y., Liu, C. W., Cane, D. E., Puglisi, J. D., and Khosla, C. (2007) Solution Structure and Proposed Domain Domain Recognition Interface of an Acyl Carrier Protein Domain from a Modular Polyketide Synthase. *Protein Sci. 16*, 2093–2107.
17. Carvalho, R., Reid, R., Viswanathan, N., Gramajo, H., and Julien, B. (2005) The Biosynthetic Genes for Disorazoles, Potent Cytotoxic Compounds that Disrupt Microtubule Formation. *Gene 359*, 91–98.
18. Kopp, M., Irschik, H., Pradella, S., and Rolf Müller, R. (2005) Production of the Tubulin Destabilizer Disorazol in Sorangium Cellulosum: Biosynthetic Machinery and Regulatory Genes. *ChemBioChem 6*, 1277–1286.
19. Cheng, Y., Tang, G., and Shen, B. (2003) Type I Polyketide Synthase Requiring a Discrete Acyltransferase for Polyketide Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A. 100*, 3149–3154.
20. Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., and Piel, J. (2008) Exploiting the Mosaic Structure of Trans-Acyltransferase Polyketide Synthases for Natural Product Discovery and Pathway Dissection. *Nat. Biotechnol. 26*, 225–233.
21. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) Dissecting and Exploiting Intermodular Communication in Polyketide Synthases. *Science 284*, 482–485.
22. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. Coli. *Science 291*, 1790–1792.
23. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E., and Flugel, R. S. (1997) Post-Translational Modification of Polyketide and Nonribosomal Peptide Synthases. *Curr. Opin. Chem. Biol. 1*, 309–315.
24. Zhang, Y. (2008) I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinf. 9*, 40.
25. Zhang, Y. (2007) Template-Based Modeling and Free Modeling by I-TASSER in CASP7. *Proteins 69*, 108–117.
26. Wu, S. T., Skolnick, J., and Zhang, Y. (2007) Ab Initio Modeling of Small Proteins by Iterative TASSER Simulations. *BMC Biol. 5*, 17.
27. Duhovny, D., Nussinov, R., and Wolfson, H. J. (2002) Efficient Unbound Docking of Rigid Molecules. *Lect. Notes Comput. Sci. 2452*, 185–200.
28. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H. J. (2005) PatchDock and SymmDock: Servers for Rigid and Symmetric Docking. *Nucleic Acids Res. 33*, W363–W367.
29. Andrusier, N., Nussinov, R., and Wolfson, H. J. (2007) FireDock: Fast Interaction Refinement in Molecular Docking. *Proteins: Struct., Funct., Bioinf. 69*, 139–159.
30. Mashiach, E., Schneidman-Duhovny, D., Andrusier, N., Nussinov, R., and Wolfson, H. J. (2008) FireDock: A Web Server for Fast Interaction Refinement in Molecular Docking. *Nucleic Acids Res. 36*, W229–232.
31. Jenke-Kodama, H., Sandmann, A., Muller, R., and Dittmann, E. (2005) Evolutionary Implications of Bacterial Polyketide Synthases. *Mol. Biol. Evol. 22*, 2027–2039.
32. Linne, U., Stein, D. B., Mootz, H. D., and Marahiel, M. A. (2003) Systematic and Quantitative Analysis of Protein-Protein Recognition between Nonribosomal Peptide Synthetases Investigated in the Tyrocidine Biosynthetic Template. *Biochemistry 42*, 5114–5124.
33. Lau, J., Fu, H., Cane, D. E., and Khosla, C. (1999) Dissecting the Role of Acyltransferase Domains of Modular Polyketide Synthases in the Choice and Stereochemical Fate of Extender Units. *Biochemistry 38*, 1643–1651.