BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

pubs.acs.org/biochemistry

Structure and Mechanism of the *trans*-Acting Acyltransferase from the Disorazole Synthase

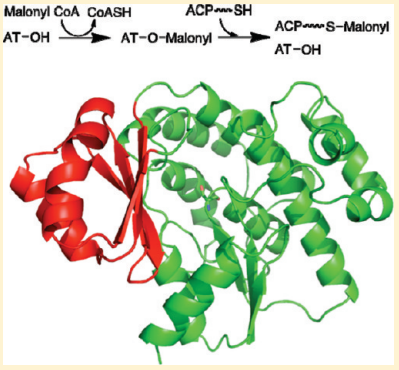
Fong T. Wong,[†] Xi Jin,[‡] Irimpan I. Mathews,[∥] David E. Cane,[⊥] and Chaitan Khosla*[,†,‡,§]

[†]Department of Chemical Engineering, [‡]Department of Chemistry, and [§]Department of Biochemistry, Stanford University, Stanford, California 94305, United States

[∥]Stanford Synchrotron Radiation Lightsource, Stanford University, Menlo Park, California 94025, United States

[⊥]Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912-9108, United States

ABSTRACT: The 1.51 Å resolution X-ray crystal structure of the *trans*-acyltransferase (AT) from the "AT-less" disorazole synthase (DSZS) and that of its acetate complex at 1.35 Å resolution are reported. Separately, comprehensive alanine-scanning mutagenesis of one of its acyl carrier protein substrates (ACP1 from DSZS) led to the identification of a conserved Asp45 residue on the ACP, which contributes to the substrate specificity of this unusual enzyme. Together, these experimental findings were used to derive a model for the selective association of the DSZS AT and its ACP substrate. With a goal of structurally characterizing the AT−ACP interface, a strategy was developed for covalently cross-linking the active site Ser → Cys mutant of the DSZS AT to its ACP substrate and for purifying the resulting AT−ACP complex to homogeneity. The S86C DSZS AT mutant was found to be functional, albeit with a transacylation efficiency 200-fold lower than that of its wild-type counterpart. Our findings provide new insights as well as new opportunities for high-resolution analysis of an important protein−protein interface in polyketide synthases.



Polyketides are medicinally important natural products with a variety of pharmacological properties.[1] Their structural features can be mapped onto individual enzymes on multifunctional assembly lines, called polyketide synthases (PKSs), in a highly modular fashion. A typical PKS consists of 2−20 enzyme modules, where each module contains minimally a β-keto-synthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP) domain. The KS domain receives the growing polyketide chain from the upstream module and subsequently catalyzes chain elongation with an ACP-bound extender unit as the cosubstrate. The extender unit is transacylated onto the pantetheinyl arm of the ACP by the AT.

Within the modular PKS family, some PKSs are comprised of enzyme modules that lack dedicated AT domains. Examples of these "AT-less" PKSs include the leinamycin, migrastatin, and disorazole synthases.[2−8] In these cases, acyltransferase activity comes from a discrete protein that transacylates *in trans* the same extender unit onto each ACP domain of that PKS. We recently reported the reconstitution and biochemical properties of the malonyl-CoA-specific *trans*-acting AT from the disorazole synthase (DSZS).[9] Notably, this AT has considerably higher activity toward heterologous ACP domains than that of normally *cis*-acting AT domains derived from canonical PKS assembly lines. Therefore, to rationally exploit its promise for biosynthetic engineering of novel polyketide antibiotics, we sought to determine the atomic structure of the DSZS AT protein.

The X-ray crystal structures of a number of ATs from polyketide synthases and the related fatty acid synthases have been determined. These include the multipurpose ATs from *Streptomyces coelicolor* and *Escherichia coli* in complex with acetate or malonyl-CoA [Protein Data Bank (PDB) entries 1NM2[10] and 2G2Z[11]] and the AT domains of modules 3 and 5 of the 6-deoxyerythronolide B synthase (PDB entries 2HG4[12] and 2QO3[13]). In this study, we report two X-ray crystal structures of the *trans*-AT from DSZS. We have also developed a novel approach toward the eventual crystallographic mapping of the AT−ACP protein−protein interface.

EXPERIMENTAL PROCEDURES

Reagents and Chemicals. [^{14}C]Malonyl-CoA was from American Radiolabeled Chemicals. 1,3-Dibromopropanone was obtained from Alfa Aesar. SDS−PAGE gradient gels were from Invitrogen. Ni-NTA agarose was from Qiagen. The Hi-Trap-Q anion exchange column was from GE Healthcare. The thrombin cleavage capture kit was obtained from EMD4Biosciences (Novagen, 69022). Anti-FLAG M2 affinity gel was from Sigma-Aldrich (A2220). BMH and BMOE were obtained from Thermo Scientific Pierce Protein Research Products. All other chemicals were from Sigma.

Protein Constructs. DSZS AT, DEBS ACP3, and DEBS ACP6 were expressed from plasmids pFW3, pSHIV9, and

Received: April 25, 2011
Revised: June 26, 2011
Published: June 27, 2011

Table 1. Crystallographic Parameters and Data Collection and Refinement Statistics

	native	acetate complex
Crystallographic Parameters		
space group	$P2_12_12_1$	$P2_12_12_1$
unit cell dimensions (Å)	43.43 Å, 54.14 Å, 121.41 Å	43.03 Å, 54.01 Å, 120.71 Å
	90.0°, 90.0°, 90.0°	90.0°, 90.0°, 90.0°
Data Collection Statistics		
resolution limit (Å)	32.6−1.51	35.0−1.35
no. of observed reflections	424975	326939
no. of unique reflections	44785	61823
completeness (overall/outer shell)	97.7/86.9	98.6/96.9
redundancy (overall/outer shell)	9.5/4.4	5.3/3.5
R_{sym} [a] (%) (overall/outer shell)	5.1/44.3	6.7/48.6
$R_{mrgd\text{-}F}$ [b] (%) (overall/outer shell)	6.8/48.7	8.5/57.0
I/σ (overall/outer shell)	26.7/3.4	15.7/2.8
Refinement Statistics		
resolution limit (Å)	32.6−1.51	35.0−1.35
no. of reflections/%	42520/97.7	58636/98.6
no. of reflections used for R_{free}	2258	3118
$R_{(working)}$ [c] (%)	16.2	16.1
R_{free} [d] (%)	18.2	17.4
model contents/average B (Å^2)		
protein atoms	2156/19.8	2167/15.1
ligand atoms	0	0
ions/buffer atoms	0	31/21.3
water molecules	299/30.4	276/25.5
Ramachandran favored/outliers	99.6/0.0	99.6/0.0
rmsd		
bond lengths (Å)	0.016	0.016
bond angles (deg)	1.511	1.613

[a] $R_{sym} = \Sigma|I_{avg} - I_i|/\Sigma I_i$. [b] For $R_{mrgd\text{-}F}$, see ref 32. [c] $R_{(working)} = \Sigma|F_p - F_{pcalc}|/\Sigma F_p$, where F_p and F_{pcalc} are observed and calculated structure factors, respectively. [d] R_{free} is calculated with 5% of the data.

pFW55, respectively.[9] His-tagged only DSZS ACP1 was expressed from plasmid pFW69. The DSZS ACP1 construct containing a His tag, thrombin cleavage site, and FLAG tag in sequence upstream of the ACP was amplified using primers 5′-AAAAAACATATGGACTACAAAGACGATGACGACAAGCTGGCGCCTGCAGGGGCAGGACAG-3′ and 5′-TTTTTGAATTCTCATGCCGACCTCGCGGGGACGCG-3′ with cosmid pKOS254-190.4[8] as the template. The amplified DNA fragment was digested with *Nde*I and *Eco*RI and cloned into a pET28 vector to yield pFW96. The S86C mutant of the DSZS AT (pFW88) was obtained using the Quikchange Site-directed Mutagenesis Kit (Stratagene) with primer 5′-TTCCTGGCCGGCCACTGCCTGGGCGAGTTCAGC-3′ and its antiparallel primer with pFW3 as the template. Alanine mutants of DSZS ACP1 were created with the Quikchange kit using pFW69 as the template.

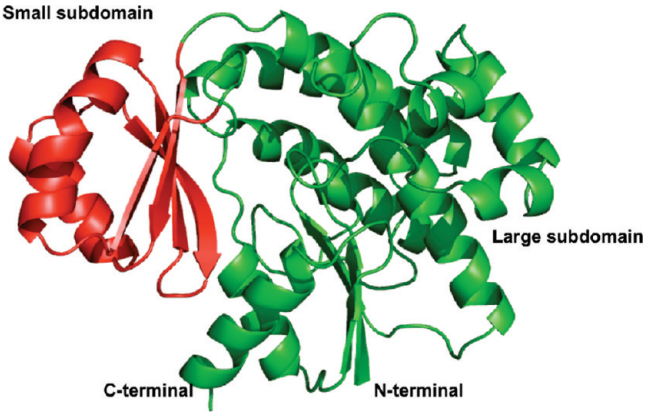


Figure 1. X-ray structure of the DSZS AT. The small subdomain is colored red.

Table 2. Alignment of AT Structures with DSZS AT

aligned structure (PDB entry)	backbone atom rmsd (Å) (no. of atoms aligned)	backbone atom rmsd (Å) within a 10 Å radius of the catalytic serine (no. of atoms aligned)
S. coelicolor MAT (1NM2)	1.0 (732)	0.5 (166)
E. coli MAT (2G2Z)	0.9 (881)	0.5 (194)
DEBS AT3 (2QO3)	2.4 (785)	0.9 (190)
DEBS AT5 (2HG4)	1.5 (833)	0.7 (122)

Protein Expression and Purification. Plasmid pFW3 (DSZS AT) was introduced into *E. coli* BL21(DE3) by electroporation. The resulting transformant was grown in LB medium at 37 °C until the culture optical density reached 0.6. The culture was cooled to 18 °C, then induced with 0.2 mM isopropyl β-D-thiogalactopyranoside, and grown for an additional 15 h. Cells were harvested by centrifugation (4420*g* for 15 min). The cell pellet was resuspended in lysis/wash buffer [50 mM phosphate, 10 mM imidazole, and 500 mM NaCl (pH 7.6)] and lysed by sonication (6 × 30 s, on ice). After centrifugation at 42700*g* for 45 min, the supernatant was incubated with Ni-NTA agarose for 1 h. The resin was washed with 10 column volumes of lysis/wash buffer, and the bound protein was eluted with 4 column volumes of elution buffer [50 mM phosphate, 150 mM imidazole, and 300 mM NaCl (pH 7.6)]. The eluant was diluted to 50 mL with water and applied to a HiTrap-Q anion exchange column at a rate of 2 mL/min. DSZS AT was eluted in the flow-through. The protein was exchanged with 20 mM Tris-HCl buffer (pH 7.2) for use in crystallization. A typical yield of DSZS AT was 50 mg/L of culture volume. Holo-ACPs were expressed using BAP1 cells, which have the *sfp* phosphopantetheinyl transferase gene from *Bacillus subtilis* incorporated into their chromosome.[14] Proteins used for cross-linking reactions and assays were expressed and purified as described previously.[9]

Crystallization. Two types of crystals of DSZS AT were grown at 22 °C using the hanging drop vapor diffusion technique. Both crystal types were grown from a 9 mg/mL protein stock in 20 mM Tris-HCl (pH 7.2). The first type of crystal was grown in hanging drops comprised of 1.2 μL of protein solution and 1.2 μL

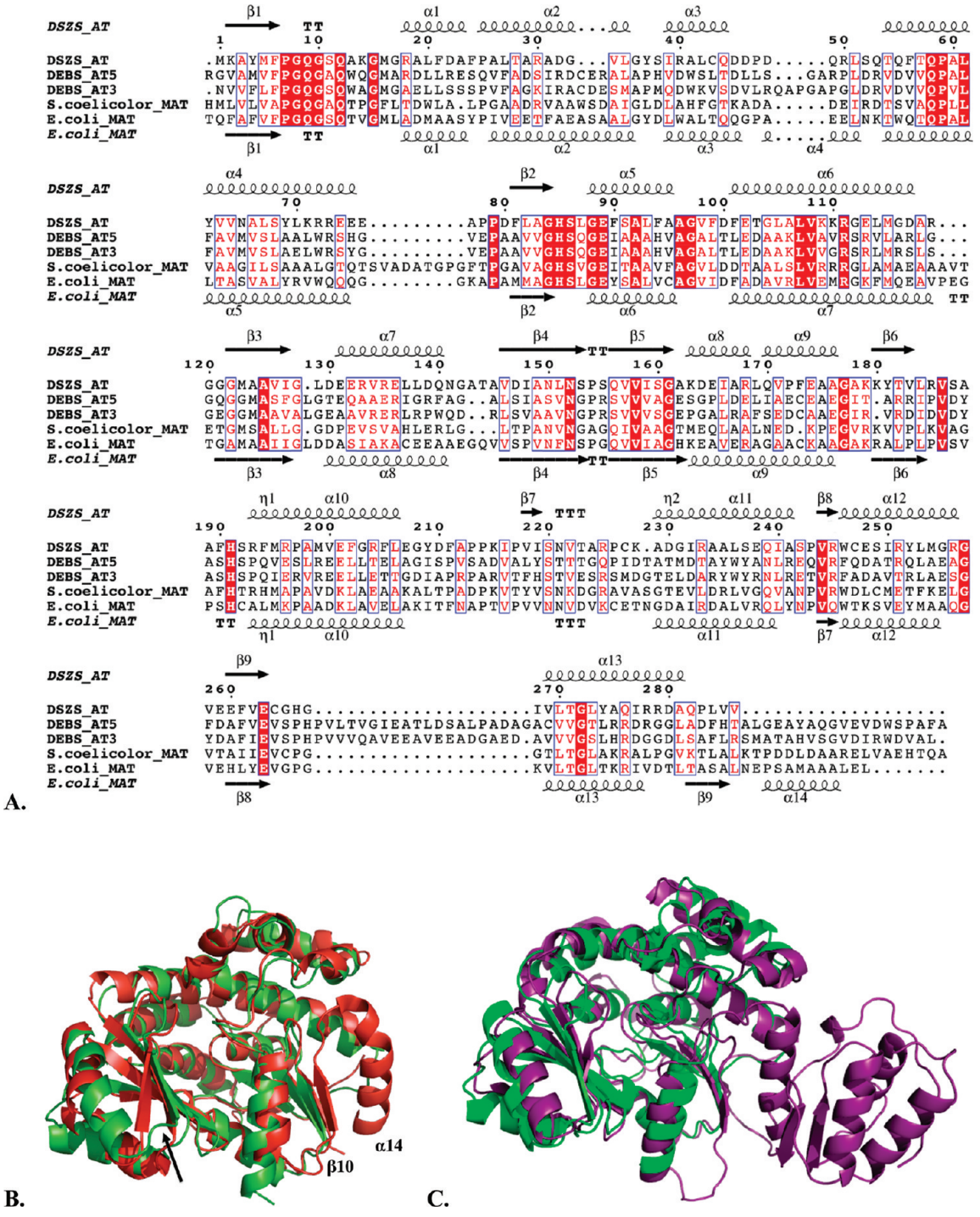


Figure 2. (A) Sequence alignment of structurally characterized acyltransferases (AT). DSZS AT, *S. coelicolor*, and *E. coli* acyltransferases (MAT) are malonyl-specific enzymes, whereas the two ATs from the 6-deoxyerythronolide B synthase (DEBS) are methylmalonyl-specific. Sequences were aligned with ClustalW2.[33] The image of aligned sequences was generated using ESPript.[34] Secondary structures of DSZS AT (top) and *E. coli* MAT (bottom) are annotated. Similar and strictly conserved residues are indicated by red letters and red boxes, respectively. The GHSXG catalytic motif is present in all sequences. The missing gap in sequence alignment is due to an extra C-terminal β sheet in the DEBS ATs. (B) Structural alignment of DSZS AT (green) and *S. coelicolor* MAT (PDB entry 1NM2,[10] red). C-Terminal secondary structural elements of *S. coelicolor* MAT, which are absent in DSZS AT, are labeled. Differences in β-sheet arrangements can also be observed (arrow). (C) Alignment of DSZS AT (green) and the DEBS AT5 domain (PDB entry 2HG4,[12] purple).

of reservoir solution containing 30% PEG 3350, 40 mM ammonium acetate, and 0.1 M Tris-HCl (pH 7.5). A solution containing 32% PEG 3350 was used as the cryoprotectant for this crystal, whose low-temperature data yielded the structure of the native DSZS AT. A second type of crystal, which was used to obtain the acetate-bound complex, was grown in an analogous fashion. The

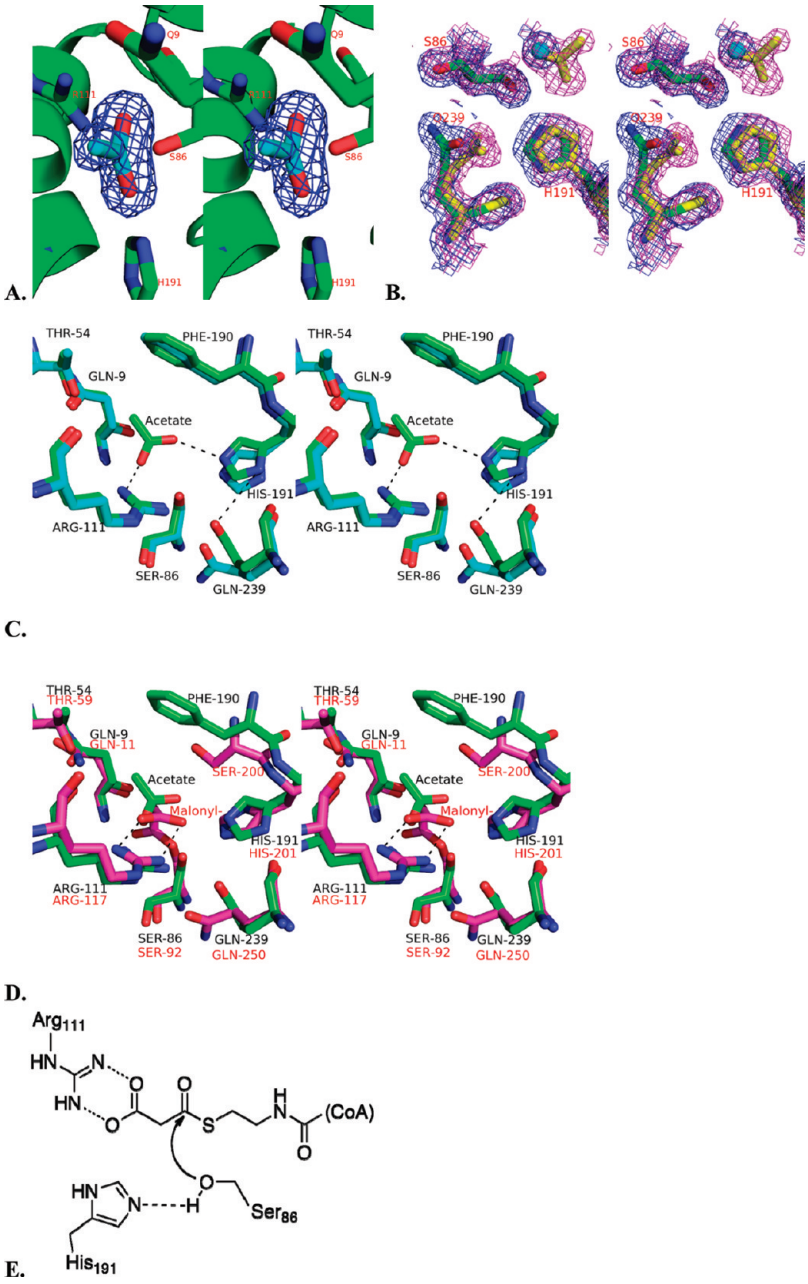


Figure 3. (A) Stereoview of an OMIT $F_o - F_c$ electron density map of the active site, contoured at 3σ. The refined structure of the acetate, which was removed during the derivation of the omit map, is colored cyan. (B) Stereoview of the $2F_o - F_c$ electron density map (contoured at 1σ) of the active sites of both the free (blue) and acetate-bound DSZS AT (magenta). (C) Stereoview of the aligned active sites of free (cyan) and acetate-bound DSZS AT (green). (D) Active site alignment of the acetate-bound DSZS AT (green, black residue labels) and malonyl-AT from *E. coli* (PDB entry 2G2Z,[11] magenta, red residue labels). Note that the imidazole side chain of DSZS AT His191 is rotated relative to the corresponding His201 of the *S. coelicolor* MAT. Accordingly, Gln239 of DSZS AT can interact with N_δ of His191, whereas the corresponding Gln250 of the MAT lacks this interaction and is oriented differently. Hydrogen bonding is represented by dashed lines. (E) Proposed reaction mechanism for DSZS AT.

reservoir solution contained 30% PEG 3350, 200 mM ammonium acetate, and 0.1 M Tris-HCl (pH 7.5); 32% PEG 3350 was also used as the cryoprotectant for this crystal.

Data Collection. Diffraction data for the native DSZS AT structure were collected at Stanford Synchrotron Radiation Lightsource (SSRL) beamline 11-1. The data set was collected at 100 K using a Quantum 315 CCD detector. Diffraction data for the acetate-bound complex were collected at SSRL beamline 12-2 at 100 K using a Dectris Pilatus 6M detector. Data were processed using XDS.[15] Data collection statistics are listed in Table 1.

Structure Solution and Refinement. The structure of the DSZS AT was determined by molecular replacement, using PDB entry 2G1H[11] as the search model. After several cycles of manual model building using Arp/Warp[16] and refinement using REFMAC,[17] the final $R_{(working)}$ and R_{free} values for the native enzyme structure were 16.2 and 18.2%, respectively. The structure of the acetate-bound DSZS AT was determined by

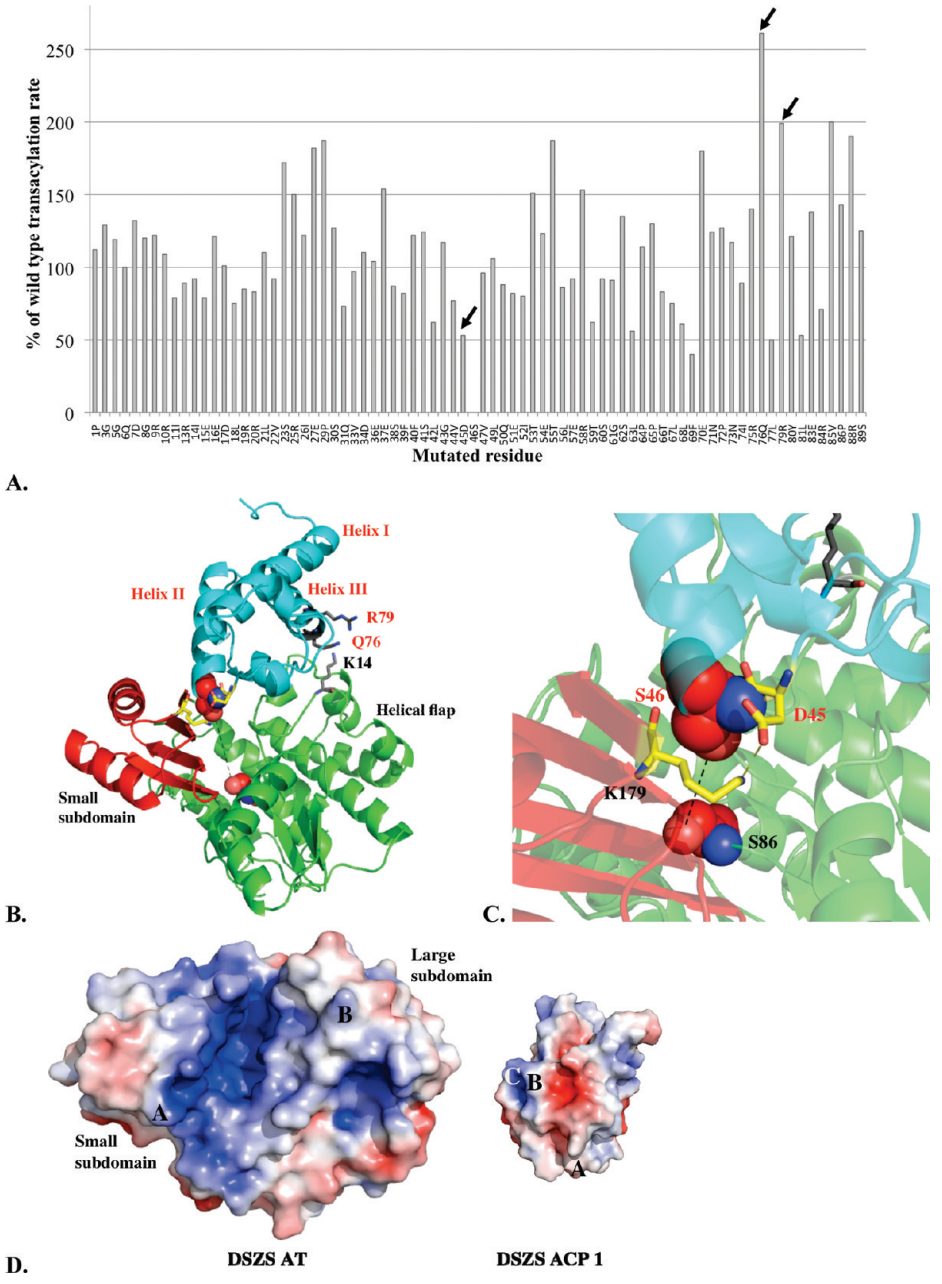


Figure 4. (A) Quantification of the transacylation rate of alanine-scanning DSZS ACP1 mutants by DSZS AT. The transacylation rates are expressed as a percentage of the wild-type DSZS ACP1 transacylation rate. Mutation in the phosphopantetheinylation site (Ser46) resulted in an inactive ACP. Wild-type alanine residues are not represented. (Arrows denote surface residues with a ≥2-fold increase or decrease in transacylation rate.) (B−D) Docking model of DSZS AT and DSZS ACP1. (B) The DSZS ACP1 homology model (cyan, red labels) was used as a ligand protein to dock with the DSZS AT structure (green with the small subdomain colored red, black labels). The active site serine residues of the AT (Ser86) and the pantetheinate attachment site of the ACP (Ser46) are shown as spheres. The Asp45-Lys179 pair is colored yellow. Surface residues Gln76 and Arg79 (ACP), along with Lys14 (AT), are colored gray. (C) Interaction between helix II (Asp45) of the ACP and Lys179 on the small subdomain of the AT (red). (D) Electrostatic surface maps of ACP and AT docking interfaces, generated with the adaptive Poisson−Boltzmann solver (APBS) in PyMOL.[35,36] Colors range from blue (positive) to white to red (negative). Docked ACP is rotated 180° from the AT−ACP interface such that annotations of A−C on AT and ACP interfaces should match up. On the ACP, A is Asp45, B is Gln76, and C is Arg79. On the AT, A is Lys179 and B is Lys14.

molecular replacement, using the coordinates of the native protein as the search model. Manual model building and refinement using Arp/Warp[16] and REFMAC[17] resulted in final $R_{(working)}$ and R_{free} values of 16.1 and 17.4%, respectively. Water molecules were included in the refinement via examination of the hydrogen bonds in the $2F_o - F_c$ and $F_o - F_c$ maps. The refinement statistics are listed in Table 1.

Alanine-Scanning Analysis of DSZS ACP1. To examine transacylation of DSZS ACP1 mutants by DSZS AT, we incubated DSZS ACP1 mutants (120 μM) with 0.003 μM DSZS AT in 100 mM sodium phosphate buffer (pH 7.2) and 2.5 mM TCEP at room temperature; 200 μM [2-^{14}C]malonyl-CoA was added to start the reaction. The total reaction volume was 30 μL. Nine microliters of the mixture was removed at 1, 2, and 3 min

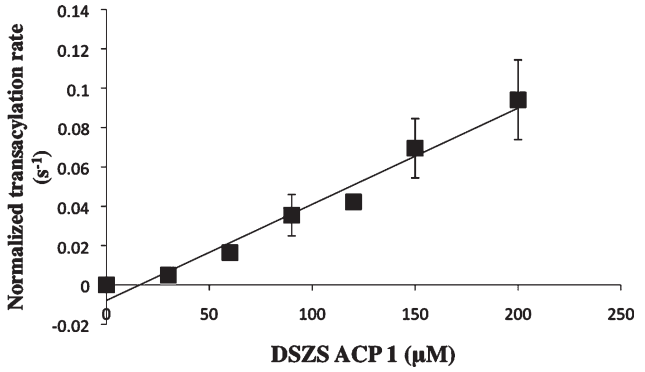


Figure 5. Kinetic analysis of transacylation of holo-ACP1 from DSZS by the S86C mutant of the DSZS AT. Each data point was obtained in duplicate.

and the reaction quenched with 5 μL of SDS−PAGE Laemmli buffer. The samples were loaded onto a SDS−PAGE gel. The gel was dried using a Bio-Rad gel-drying system and analyzed using a phosphorimager. On the basis of preliminary activity data, D45A and F69A mutants of the ACP were analyzed using liquid chromatography and mass spectrometry to quantify the extent of post-translational phosphopantetheinylation. Whereas the D45A mutant was exclusively in its holo form, the F69A mutant was a mixture of its holo and apo forms. Therefore, the results for the F69A ACP1 mutant were not carried forward into docking model calculations.

Docking Model of DSZS AT and DSZS ACP1. A homology model of DSZS ACP1 was derived using the iterative threading assembly refinement (I-TASSER) server.[18,19] The top threading templates used were 2AFDA, 2AFEA,[20] and 2JU1A.[21] Macromolecular docking calculations for DSZS AT and DSZS ACP1 were performed with PatchDock[22,23] and refined by FireDock.[24,25] On the basis of the mutagenesis results and predicted positions of the residues on the DSZS ACP1 homology model, a set of ACP surface residues with the most significant perturbations to transacylation activity were identified as potential interacting residues. These ACP surface residues, Asp45, Gln76, and Arg79, were then specified as constraints to narrow the search for a docking model. The top 100 solutions from PatchDock were manually inspected to select models with accessibility between the active serine on the AT (Ser86) and the pantetheinylated site on the ACP (Ser46). This property is required for transacylation. These models were further refined and reranked with FireDock. The most energetically favored model was chosen as the final model.

Rates of Transacylation of DSZS ACP1 by the S86C Mutant of DSZS AT. Holo-DSZS ACP1 [30−200 μM, in 100 mM phosphate (pH 7.2)] was incubated at room temperature with 0.3 μM S86C mutant DSZS AT, 2.5 mM TCEP, and 200 μM [^{14}C]malonyl-CoA. A 9 μL portion of the mixture was removed at intervals of 1 min, and the reaction was quenched with 5 μL of SDS−PAGE loading 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 phosphorimager.

Cross-Linking Reactions. For the cross-linking assays, 5 μM AT was incubated with 40 μM holo-ACP in 100 mM sodium phosphate (pH 7.2) with 120 μM 1,3-dibromopropanone at room temperature. At time points of 10, 30, 60, 120, 300, and 600 s, 5 μL of the reaction mixture was removed and the reaction quenched with an equal volume of Laemmli loading buffer. The AT used in these assays was either the wild-type DSZS AT or its S86C mutant. ACPs used included DSZS ACP1, DEBS ACP3, and DEBS ACP6. Samples were loaded on 4 to 12% Bis-Tris SDS−PAGE gels.

For comparison between cross-linkers, 120 μM BMH, BMOE, or DBP was used with 5 μM S86C DSZS AT and 40 μM DSZS ACP1. A 6 μL sample of the reaction mixture was removed at time points of 30, 60, and 120 s and the reaction quenched. Samples were loaded on 10% Bis-Tris SDS−PAGE gels.

In a larger-scale reaction, FLAG-tagged DSZS ACP1 (40 μM) and the S86C mutant of DSZS AT (40 μM) were co-incubated in 100 mM sodium phosphate buffer (pH 7.2) along with 300 μM 1,3-dibromopropanone at 4 °C overnight.

Purification of FLAG-Tagged ACP. The FLAG-tagged DSZS ACP1 construct (pFW96) contains (in sequence, from the N-terminus) a His tag, a thrombin cleavage site, and a FLAG tag upstream of the DSZS ACP1 gene. Its yield was comparable to that of His-tagged DSZS ACP1. The His tag was cleaved using the thrombin cleavage capture kit, and the His tag was removed by Ni-NTA purification. The desired protein eluted in the flow-through.

Purification of the Cross-Linked AT−ACP Complex. The cross-linked reaction mixture was buffer exchanged (via a 3 kDa molecular mass cutoff filter) into lysis/wash buffer. The mixture then underwent Ni-NTA purification, as described previously.[9] The eluant was buffer exchanged into TBS [50 mM Tris-HCl (pH 7.4) and 150 mM NaCl]. The resulting mixture was then subjected to immunoprecipitation with the anti-FLAG antibody resin and eluted using 6 × 1 mL of 0.1 M glycine hydrochloride (pH 3.5). The eluant was buffer exchanged into FPLC buffer [50 mM sodium phosphate (pH 7.2)] and then further applied on a HiTrap-Q anion exchange column at a rate of 2 mL/min. The AT−ACP adduct was eluted at 170 mM NaCl [50 mM sodium phosphate (pH 7.2)]. Starting from 7 mg of AT, we produced 1 mg of the AT−ACP adduct.

RESULTS

Crystallization. Heterologous expression and purification of the 33 kDa *trans*-AT from DSZS have been described previously.[9] The protein yielded crystals that revealed its structure at 1.51 Å resolution. Notwithstanding our best efforts to soak these crystals with malonate or malonyl-CoA, the ligand-bound AT complex was not observed. However, crystals of the acetate-bound form of the *trans*-AT were obtained when the acetate concentration was increased in the crystallization buffer. The latter crystals provided the structure of the acetate-bound DSZS AT at 1.35 Å resolution.

Overall Structure. The DSZS AT is a $\alpha\beta$-hydrolase protein consisting of two subdomains (Figure 1). The smaller subdomain (residues 122−183) is a four-stranded antiparallel β sheet capped with two α helices. The larger subdomain consists of 10 α helices ranging from 6 to 21 residues in length and a short three-stranded parallel β sheet. The buried active site is accessible via a gorge between the two subdomains.

Overall, the architecture of the DSZS AT is similar to that of other structurally characterized ATs, including the *S. coelicolor* and *E. coli* malonyl CoA:ACP transacylases (MATs) as well as the methylmalonyl-specific AT domains of the 6-deoxyerythronolide B synthase.[10−13] On the basis of tertiary structure alignments, the DSZS AT is more similar to the discrete type II ATs

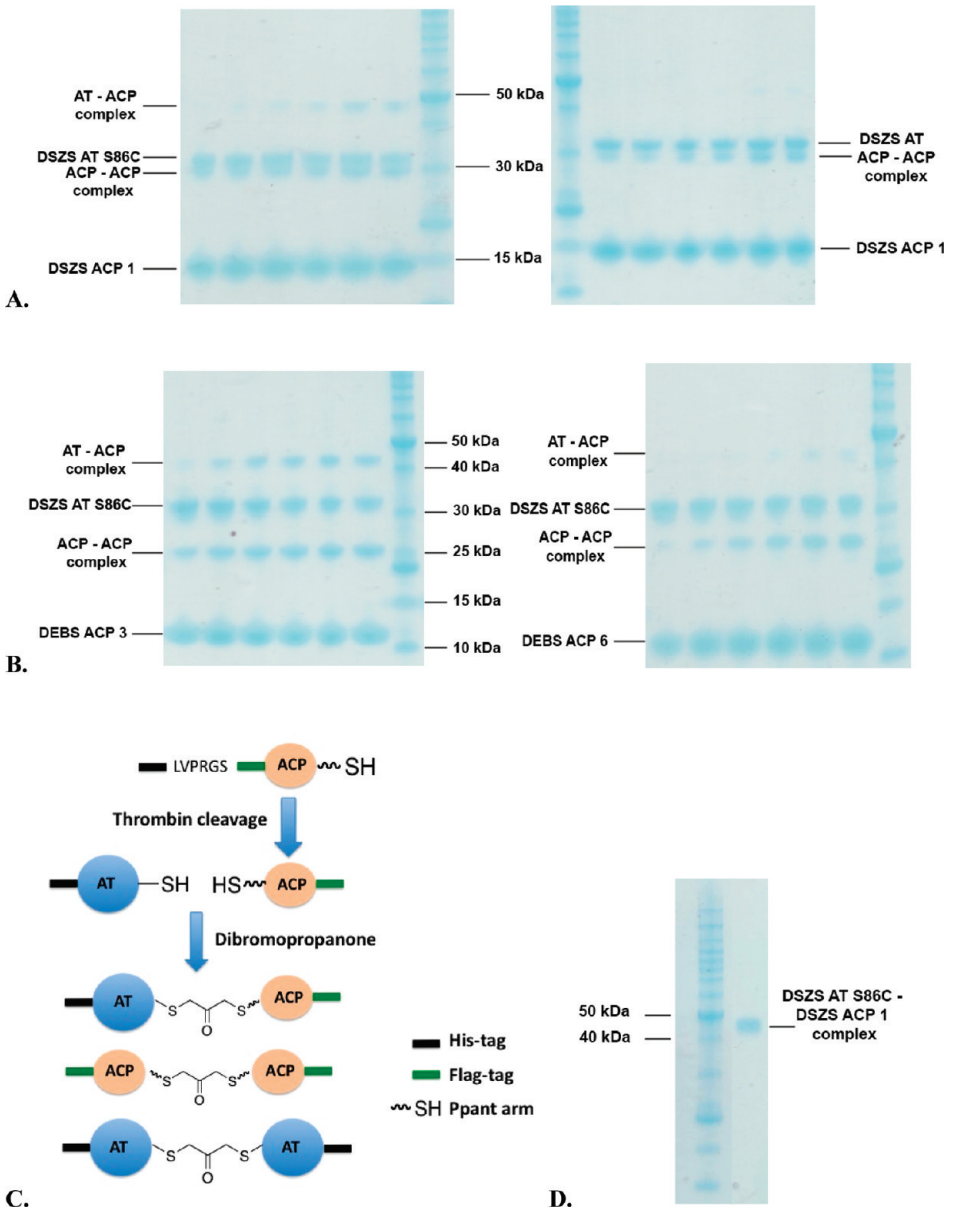


Figure 6. (A) Cross-linking of the S86C mutant (left) and the wild type (right) of the 33 kDa DSZS AT with the 14 kDa DSZS holo-ACP1. (B) Cross-linking of the S86C mutant AT with the 11 kDa DEBS holo-ACP3 (left) or DEBS ACP6 (right). In each panel, the lanes from left to right are samples taken at 10, 30, 60, 120, 300, and 600 s, respectively, at room temperature. (C) Cross-linking reaction of AT and ACP with orthogonal tags. The AT−ACP heterodimer complex can be separated from the two homodimers by a two-step purification procedure. (D) FPLC-purified, cross-linked DSZS AT S86C−DSZS ACP1 complex (45 kDa) The total yield was ≤10% with respect to the AT. Approximately 10 μg of Benchmark protein marker (Invitrogen) was loaded in all gels.

than to the type I AT domains (Table 2). Overall, the structures of ATs are more conserved in the vicinity of their active sites than over the rest of the proteins (Table 2). The other ATs also contain an extra C-terminal helix (α14) and β strand (β10) compared to the DSZS AT (Figure 2B,C).

Catalysis. The active sites of all ATs possess a Ser-His catalytic dyad.[10,11] Ser86 in the DSZS AT, which is part of a conserved GHSXG motif, is deprotonated by $N_{\varepsilon 2}$ of His191, located 2.7 Å from the Ser hydroxyl O_{δ}. In turn, $N_{\delta 1}$ of His191 in the DSZS AT is stabilized and oriented by hydrogen bonding to the carbonyl oxygen of Gln239. All of these interactions are also evident in the acetate-bound DSZS AT structure. In the latter, His191 has moved slightly to accommodate the acetate anion in the active site, and the side chain of Gln239 is positioned 0.9 Å closer to His191 (Figure 3B,C). The high resolution (1.5 and 1.35 Å) of acetate-free and acetate-bound structures provides good confidence for the modeling of both ligand and active site side chains into the electron density contours.

In the structure of the DSZS AT in a complex with acetate, only one acetate ion is located near the catalytic Ser-His dyad (Figure 3A). The plane of its carboxylate is at a 40° incline relative to that of the Arg111 guanidine. Instead of forming a bidentate salt bridge with Arg111, the carboxylate appears to be stabilized through hydrogen bonds with both Arg111 and His191 (Figure 3C). The orientation of the bound acetate is entirely analogous to that observed earlier in the *S. coelicolor*

dx.doi.org/10.1021/bi200632j |*Biochemistry* 2011, 50, 6539–6548

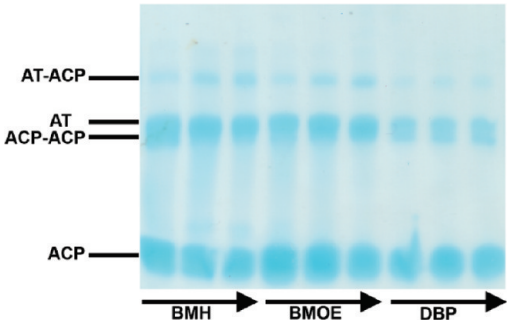


Figure 7. Cross-linking of the 33 kDa DSZS AT S86C mutant with 14 kDa DSZS holo-ACP1 using BMH, BMOE, and DBP cross-linkers. Samples were taken at 30, 60, and 120 s at room temperature.

MAT−acetate complex and is predicted to mimic the free carboxylate moiety of a bound malonyl substrate, as exemplified by structural superposition of the DSZS AT active site with the known structure of the malonyl-AT from *E. coli* (Figure 3D).[10,11] Together, these structural insights support the widely accepted mechanism of AT catalysis, outlined in Figure 3E.

In addition to Ser86, His191, and Arg111, the substrate-binding pocket of the DSZS AT also contains other conserved amino acids found in homologous AT domains, including Gln9 and Phe190. These residues may play a role in both catalysis and substrate specificity.[10,11] α-Substituted substrates, such as methylmalonyl-CoA, would presumably be prevented from binding in the active site by unfavorable steric interactions with the side chains of both Phe190 and Gln9. In turn, Gln9 is positioned by hydrogen bonding with Thr54 (Figure 3C).

AT−ACP Docking Interface. Although the specificity of DSZS AT is influenced by the identity of the ACP substrate used in the reaction,[9] the association of these two proteins is relatively weak. To derive a model for this protein−protein interaction, we therefore first sought to obtain experimental constraints through alanine-scanning mutagenesis of DSZS ACP1, a natural substrate of DSZS AT. The transacylation activity of DSZS AT was quantified with each member of the DSZS ACP1 mutant library as a substrate (Figure 4A). Interestingly, DSZS AT exhibited a higher specificity for several ACP mutants versus wild-type ACP, including 2- and 2.5-fold higher for the R79A and Q76A surface mutants, respectively. In contrast, DSZS ACP1 with its surface residue Asp45 mutated to Ala exhibited a 50% decrease in transacylation activity.

To derive a docking model of DSZS AT and DSZS ACP1, we first constructed a homology model of the latter protein, as described in Experimental Procedures. The top DSZS ACP1 homology model has an rmsd of 0.97 Å with DEBS ACP2 (PDB entry 2JU2[21]).

The docking orientation of the modeled ACP1 and DSZS AT was calculated using PatchDock and FireDock.[22−25] Among the docking solutions from PatchDock, we selected two properties. First, the models should predict surface interactions between the AT and the experimentally identified surface residue Asp45 of the ACP. Second, they should allow for unhindered access of the ACP phosphopantetheinyl arm to the active site Ser86 of the AT. The latter ensures that transacylation of the ACP is possible. Through further refinement of these models in FireDock, the most energetically favored model was chosen as our final model (Figure 4B−D). In this model, the distance between the active site Ser86 of the AT and the Ser46 cofactor attachment site on ACP1 was constrained to 20 Å, consistent with the length of a fully extended phosphopantetheine arm. The model also accounts for our experimental observations; Asp45 of DSZS ACP1 was predicted to form a salt bridge with Lys179 on the AT surface (Figure 4C). Given that these residues are highly conserved among naturally occurring ACP and AT sequences, the validity of this docking model is further supported by the observed promiscuity of DSZS AT for heterologous ACP substrates.[9]

The DSZS AT−ACP1 docking model is also supported by the observed increase in the transacylation activity of the ACP1 Q76A and R79A mutants (Figure 4A). Specifically, the model predicts that these changes should each remove unfavorable interactions with Lys14 on the AT. Perhaps not surprisingly, these residues are unique to ACP1 among different ACP domains of DSZS. Other DSZS ACPs harbor Glu or Ala in place of Gln76, and Asp, Glu, Gly, or Ala in place of Arg79.

The proposed DSZS AT−ACP1 docking model generally resembles a previously reported model for docking between the *S. coelicolor* MAT and the actinorhodin ACP.[10] In both these cases, the residues in and around the conserved Asp-Ser motif of the ACP were deduced to engage with the AT. The main difference is that, in the model for the *S. coelicolor* proteins, the conserved helical flap of the large AT subdomain acted as the primary ACP docking site,[10] whereas our model proposes that the most significant AT−ACP interactions should involve the smaller ferredoxin-like subdomain. As summarized in Figure 2B, significant differences were observed in the β-sheet arrangement of the small subdomain of the two AT proteins. Presumably, these differences play an important role in influencing the outcome of these *in silico* docking studies. Alternatively, the absence of a C-terminal helix and a β-strand in the DSZS AT could also allow ACP substrates to preferentially dock at the small subdomain of this AT.

A Covalently Linked AT−ACP Complex. Notwithstanding the alanine-scanning mutagenesis and modeling efforts summarized above, a more detailed understanding of AT−ACP interactions in assembly line PKSs will require cocrystallization of these two proteins. Their relatively weak mutual binding affinity, however, makes this an elusive goal. In our own laboratory, repeated attempts to cocrystallize the DSZS AT with DSZS ACP1 substrate yielded only crystals of the AT protein. We therefore have sought to develop a fundamentally different experimental approach based on covalent cross-linking of the ACP to the S86C mutant of the DSZS AT. To the best of our knowledge, a Ser → Cys mutant of a polyketide synthase AT has not previously been characterized.

Qualitative and quantitative analysis of the S86C mutant of the DSZS AT revealed that the protein was able to catalyze self-acylation with malonyl-CoA and transacylation of the malonyl intermediate onto a holo-ACP cosubstrate (data not shown). In the presence of DSZS ACP1, the mutant showed a k_{cat}/K_M of 500 ± 40 M^{-1} s^{-1} (Figure 5), corresponding to a 200-fold decrease compared to the value of 87000 M^{-1} s^{-1} observed for the wild-type DSZS AT.[9] Thus, it appeared that the S86C mutant was a suitable partner for cross-linking experiments with bifunctional electrophilic reagents such as 1,3-dibromopropanone (DBP).

Whereas no cross-linking was observed between the wild-type AT and a holo-ACP, the S86C mutant could be cross-linked to DSZS ACP1, DEBS ACP3, and DEBS ACP6 in the presence of dibromopropanone (Figure 6A,B). Via attachment of a FLAG affinity tag to the AT and a His_6 tag to the ACP, the cross-linked

protein complex could be purified from the reaction mixture via sequential affinity purification on a Ni-NTA column followed by an anti-FLAG column (Figure 6C,D). To further optimize the yield of cross-linked DSZS ACP1 and the S86C mutant, we evaluated alternative homobifunctional cross-linkers with longer spacer lengths, including bis(maleimido)ethane (BMOE) and bis(maleimido)hexane (BMH). The thiol-to-thiol lengths of BMH, BMOE, and DBP are 16, 11, and 6 Å, respectively. As shown in Figure 7, both BMH and BMOE yielded significantly larger quantities of the AT−ACP complex compared to DBP. Thus, it is possible to generate stable complexes of this AT−ACP pair of proteins.

DISCUSSION

Although the vast majority of known PKS assembly lines contain dedicated acyltransferase (AT) domains in each module, some systems, such as the disorazole synthase (DSZS), lack dedicated AT domains within individual modules. These AT-less PKSs have a single, discrete AT that catalyzes *in trans* acylation of modular ACP domains. This atypical mechanism for extender unit incorporation presents intriguing opportunities for engineering polyketide biosynthesis.[26] A critical prerequisite for the exploitation of this potential, however, is a detailed atomic-level understanding of the mechanism of ACP recognition by a *trans*-acting AT. As a step in this direction, we have now determined the first X-ray crystal structure of a *trans*-acting AT, the DSZS AT.

Overall, DSZS AT is similar to structurally characterized ATs from both type I and type II PKSs, with greater similarity to malonyltransferases from type II PKS systems. The major difference between DSZS AT and previously characterized ATs is expected to reside in the mode and specificity of AT−ACP interactions. In particular, DSZS AT is a slightly smaller protein because of the absence of the usual C-terminal helix and β-strand. The DSZS AT also has notable differences in the β-sheet arrangement of its small subdomain. Either or both of these differences could result in changes in predicted AT−ACP docking interactions that are centered on the small subdomain instead of the large subdomain of the AT. Our mutagenesis and modeling efforts have also suggested that the DSZS AT interacts with its ACP substrates at a conserved Asp45 on the ACP, a feature that is presumably responsible for its promiscuity toward heterologous ACP domains.

Consistent with the conservation of the catalytic mechanism of all ATs, the active site of the DSZS AT is closely related in sequence and active site structure to other crystallographically characterized ATs.[10,11,27] The importance of a nucleophilic serine residue in the active site is supported by the behavior of the S86C mutant whose transacylation efficiency was reduced by 200-fold compared to that of wild-type DSZS AT. This effect is a factor of 100-fold stronger than the effect of the analogous S101C mutation of the recombinant thioesterase of chicken fatty acid synthase[28] but is still modest compared to the 6×10^5-fold decrease in activity in the S192C mutant of the serine protease trypsin.[29] The introduction of a cysteine residue into the active site of the AT has allowed preparation of functionally relevant AT−ACP adducts using thiol-specific cross-linking reagents. Although only a modest fraction of the AT was conjugated to ACP partners (Figure 6), the attachment of distinct affinity tags on the two proteins allowed purification of this cross-linked adduct. Further improvements in the yield of such adducts might be achieved with suitable mechanism-based cross-linking methods and reagents[30,31] and may eventually provide thoroughly pursued direct structural insights into the AT−ACP protein−protein interface.

AUTHOR INFORMATION

Corresponding Author

*E-mail: khosla@stanford.edu. Telephone: (650) 723-6538. Fax: (650) 725-7294.

Funding Sources

This research was supported by grants from the National Institutes of Health (GM087934 to C.K. and GM022172 to D. E.C.) and by a National Science Scholarship from the Agency of Science, Technology and Research (A*STAR), Singapore, to F. T.W.

ACKNOWLEDGMENT

Portions of this research were conducted at the Stanford Synchrotron Radiation Lightsource (SSRL), a division of the SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Basic Energy Sciences, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

ABBREVIATIONS

AT, acyltransferase; CoA, coenzyme A; ACP, acyl carrier protein; DEBS, 6-deoxyerythronolide B synthase; DSZS, disorazole synthase; KS, β-ketosynthase; PKS, polyketide synthase; PDB, Protein Data Bank; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; FPLC, fast performance liquid chromatography; SSRL, Stanford Synchrotron Radiation Lightsource; TCEP, tris(2-carboxyethyl)phosphine; rmsd, root-mean-square deviation; MAT, malonyl-CoA:ACP transacylase; DBP, 1,3-dibromopropanone; BMOE, bis(maleimido)ethane; BMH, bis(maleimido)hexane.

REFERENCES

(1) Khosla, C. (2009) Structures and Mechanisms of Polyketide Synthases. *J. Org. Chem. 74*, 6416–6420.

(2) 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.

(3) Cheng, Y., Tang, G., and Shen, B. (2003) From the Cover: Type I Polyketide Synthase Requiring a Discrete Acyltransferase for Polyketide Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A. 100*, 3149–3154.

(4) Piel, J., Hui, D., Fusetani, N., and Matsunaga, S. (2004) Targeting Modular Polyketide Synthases with Iteratively Acting Acyltransferases from Metagenomes of Uncultured Bacterial Consortia. *Environ. Microbiol. 6*, 921–927.

(5) Tang, G., Cheng, Y., and Shen, B. (2004) Leinamycin Biosynthesis Revealing Unprecedented Architectural Complexity for a Hybrid Polyketide Synthase and Nonribosomal Peptide Synthetase. *Chem. Biol. 11*, 33–45.

(6) Lim, S., Ju, J., Zazopoulos, E., Jiang, H., Seo, J., Chen, Y., Feng, Z., Rajski, S. R., Farnet, C. M., and Shen, B. (2009) Iso-Migrastatin, Migrastatin, and Dorrigocin Production in *Streptomyces platensis* NRRL 18993 is Governed by a Single Biosynthetic Machinery Featuring an Acyltransferase-Less Type I Polyketide Synthase. *J. Biol. Chem. 284*, 29746–29756.

(7) 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.

(8) 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.

(9) Wong, F. T., Chen, A. Y., Cane, D. E., and Khosla, C. (2010) Protein-Protein Recognition between Acyltransferases and Acyl Carrier Proteins in Multimodular Polyketide Synthases. *Biochemistry 49*, 95–102.

(10) Keatinge-Clay, A., Shelat, A. A., Savage, D. F., Tsai, S. C., Miercke, L. J., O'Connell, J., Khosla, C., and Stroud, R. M. (2003) Catalysis, Specificity, and ACP Docking Site of *Streptomyces coelicolor* Malonyl-CoA:ACP Transacylase. *Structure 11*, 147–154.

(11) Oefner, C., Schulz, H., D'Arcy, A., and Dale, G. E. (2006) Mapping the Active Site of *Escherichia coli* Malonyl-CoA-Acyl Carrier Protein Transacylase (FabD) by Protein Crystallography. *Acta Crystallogr. 62*, 613–618.

(12) Tang, Y., Kim, C., 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.

(13) Tang, Y., Chen, A., Kim, C., Cane, D., 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.

(14) 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.

(15) Kabsch, W. (2010) XDS. *Acta Crystallogr. D66*, 125–132.

(16) Perrakis, A., Morris, R. M., and Lamzin, V. S. (1999) Automated Protein Model Building Combined with Iterative Structure Refinement. *Nat. Struct. Biol. 6*, 458–463.

(17) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D53*, 240–255.

(18) Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc. 5*, 725–738.

(19) Zhang, Y. (2008) I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinf. 9*, 40.

(20) Johnson, M. A., Peti, W., Herrmann, T., Wilson, I. A., and Wuthrich, K. (2006) Solution Structure of Asl1650, an Acyl Carrier Protein from *Anabaena* sp. PCC 7120 with a Variant Phosphopantetheinylation-Site Sequence. *Protein Sci. 15*, 1030–1041.

(21) 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.

(22) Duhovny, D., Nussinov, R., and Wolfson, H. (2002) Efficient Unbound Docking of Rigid Molecules. In *Proceedings of the 2nd Workshop on Algorithms in Bioinformatics (WABI) Lecture Notes in Computer Science* (Guigo, R., and Gusfield, D., Eds.) Vol. 2452, pp 185−200, Springer-Verlag, Rome.

(23) 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.

(24) Andrusier, N., Nussinov, R., and Wolfson, H. J. (2007) FireDock: Fast Interaction Refinement in Molecular Docking. *Proteins 69*, 139–159.

(25) 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–W232.

(26) 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.

(27) Dreier, J., Li, Q., and Khosla, C. (2001) Malonyl-CoA:ACP Transacylase from *Streptomyces coelicolor* has Two Alternative Catalytically Active Nucleophiles. *Biochemistry 40*, 12407–12411.

(28) Pazirandeh, M., Chirala, S. S., and Wakil, S. J. (1991) Site-directed mutagenesis studies on the recombinant thioesterase domain of chicken fatty acid synthase expressed in *Escherichia coli*. *J. Biol. Chem. 266*, 20946–20952.

(29) Higaki, J. N., Evnin, L. B., and Craik, C. S. (1989) Introduction of a Cysteine Protease Active Site into Trypsin. *Biochemistry 28*, 9256–9263.

(30) Hur, G. H., Meier, J. L., Baskin, J., Codelli, J. A., Bertozzi, C. R., Marahiel, M. A., and Burkart, M. D. (2009) Crosslinking Studies of Protein-Protein Interactions in Nonribosomal Peptide Biosynthesis. *Chem. Biol. 16*, 372–381.

(31) Kapur, S., Worthington, A., Tang, Y., Cane, D. E., Burkart, M. D., and Khosla, C. (2008) Mechanism Based Protein Crosslinking of Domains from the 6-Deoxyerythronolide B Synthase. *Bioorg. Med. Chem. Lett. 18*, 3034–3038.

(32) Diederichs, K., and Karplus, P. A. (1997) Improved R-Factors for Diffraction Data Analysis in Macromolecular Crystallography. *Nat. Struct. Mol. Biol. 4*, 269–275.

(33) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X Version 2.0. *Bioinformatics 23*, 2947–2948.

(34) Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: Multiple Sequence Alignments in PostScript. *Bioinformatics 15*, 305–308.

(35) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037–10041.

(36) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.