Biochemistry

Article

Cite This: *Biochemistry* 2019, 58, 2978−2986

pubs.acs.org/biochemistry

Structural Insights into the Substrate Specificity of Acyltransferases from Salinomycin Polyketide Synthase

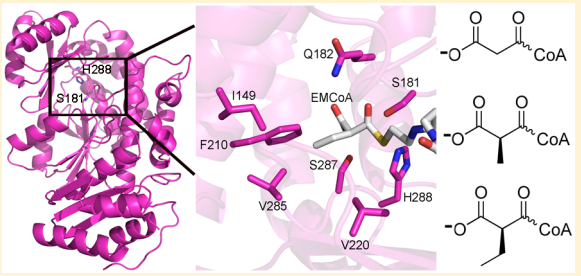
Fa Zhang,[†] Ting Shi,[†] Huining Ji,[†] Imtiaz Ali,[†] Shuxin Huang,[†] Zixin Deng,[†] Qing Min,[‡] Linquan Bai,[†] Yilei Zhao,*[,†] and Jianting Zheng*[,†]

[†]State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

[‡]Pharmacy School, Hubei University of Science and Technology, Hubei, Xianning 437100, China

S *Supporting Information*

ABSTRACT: Salinomycin with antibacterial and anticoccidial activities is a commercial polyether polyketide widely used in animal husbandry as a food additive. Malonyl-CoA (MCoA), methylmalonyl-CoA (MMCoA), and ethylmalonyl-CoA (EMCoA) are used as extension units in its biosynthesis. To understand how the salinomycin modular polyketide synthase (PKS) strictly discriminates among these extension units, the acyltransferase (AT) domains selecting MCoA, MMCoA, and EMCoA were structurally characterized. Molecular dynamics simulations of the AT structures helped to reveal the key interactions involved in enzyme−substrate recognitions, which enabled the engineering of AT mutants with switched specificity. The catalytic efficiencies (k_{cat}/K_m) of these AT mutants are comparable with those of the wild-type AT domains. These results set the stage for engineering the AT substrate specificity of modular PKSs.



Modular polyketide synthases (PKSs) are multidomain enzymatic assembly lines responsible for the biosynthesis of polyketides that exhibit a wide range of useful pharmacological properties.[1] They construct the macrocyclic aglycones of polyketides through repeated decarboxylative condensations from short-chain carboxylate building blocks activated by coenzyme A (CoA). In each module, an acyltransferase (AT) domain recruits a specific extension unit and transfers the carboxyacyl group to the phosphopantetheine arm of an acyl carrier protein (ACP) domain, while a ketosynthase (KS) domain accepts a polyketide intermediate from the upstream module and catalyzes the condensation between the growing intermediate and the ACP-bound extension unit. Optional ketoreductase, dehydratase, and enoylreductase domains in the module catalyze the stepwise reduction of the initially formed *β*-keto group to control the final oxidation state of the growing polyketide acyl-ACP intermediate.

AT domains have long been attractive targets for the engineering of modular PKSs because the use of different extension units is an important source of the structural diversity of polyketides.[2] Multiple approaches, including domain swapping and complementation of inactivated *cis*-AT domains by *trans*-ATs, have been applied in the engineering of AT domains.[3−5] Due to the potential to minimize the deleterious side effects of engineering, targeted mutagenesis of AT domains is becoming an increasingly popular approach to regiospecifically incorporating diverse types of extension units into polyketides.[6−10] Although the apo structures of EryAT3 [Protein Data Bank (PDB) entry 2QO3][11] and EryAT5 (PDB entry 2HG4)[12] and the complex structures of SpnD-AT with a benzylmalonyl (PDB entry 5YDM) or pentynylmalonyl (PDB entry 5YDL)[13] extender unit have helped reveal active site residues that show strong correlations with substrate specificity, engineering the specificity of AT domains is still challenging.

An AT domain loads a carboxy acyl group onto the phosphopantetheine thiol of an ACP domain by a ping-pong bi-bi mechanism using a His-Ser catalytic dyad.[11,12] The ability of an AT domain to strictly discriminate between the extension units supplied by a producing strain has been revealed by numerous studies. AT domains of erythromycin PKS are archetypical examples specific for a methylmalonyl-CoA (MMCoA) extension unit. Detailed kinetic studies of EryAT3 suggest the specificity on MMCoA lies in the formation of the acyl-AT intermediate,[13] but even with the structures of several AT domains in hand, the molecular basis for extension unit selection is still obscure. Phylogenetic analyses of AT domains reveal divergence based on extension unit specificity instead of species of origin.[14] Conserved motifs correlated with MMCoA or malonyl-CoA (MCoA) substrate specificity have been identified in the active site of AT domains by structure-based sequence alignments.[15] However, mutation from the MMCoA-specific YASH motif to the MCoA-specific

Received: April 5, 2019
Revised: June 3, 2019
Published: June 14, 2019

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

HAFH motif in EryAT1, EryAT4, and EryAT6 results in mutants with relaxed specificity that can incorporate both extension units, rather than a switch of specificity from MMCoA to MCoA.[8,15,16] Kinetic analysis of the YASH to HAFH mutant of EryAT3 reveals a drastic decrease in catalytic efficiency toward both MMCoA and MCoA, suggesting that the observed incorporation of non-native extension units *in vivo* is due to the diminished catalytic activity toward the native substrate rather than the enhanced catalytic activity for the non-native substrate.[13] A mixed HASH motif is observed in the active site of EpoAT3 that utilizes both MCoA and MMCoA in epothilone biosynthesis. The YASH mutant of EpoAT3 exhibits diminished activity for MMCoA, while the HAFH mutant is catalytically inactive.[17]

Salinomycin, an important commercial polyether polyketide with antibacterial and anticoccidial activities, is widely used in animal husbandry as a food additive. Salinomycin PKS from *Streptomyces albus* XM211 catalyzes the formation of salinomycin from six MCoAs, six MMCoAs, and three ethylmalonyl-CoAs (EMCoA) (Figure 1).[18,19] To understand

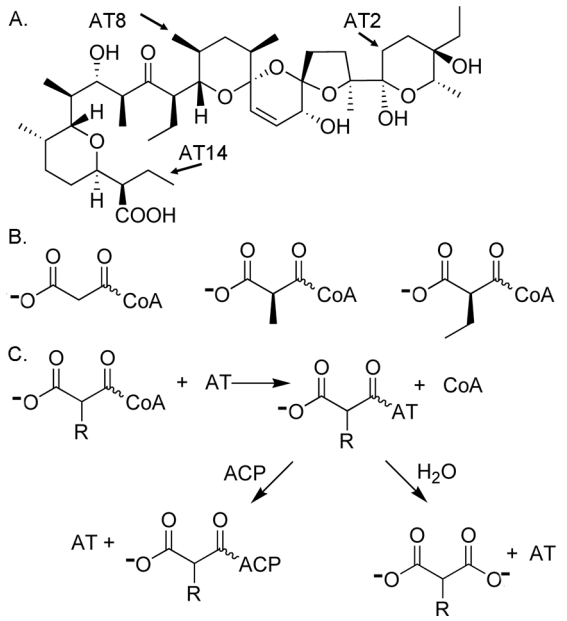


Figure 1. Extension units incorporated by AT domains during salinomycin biosynthesis. (A) Structure of salinomycin. The extension units incorporated by SalAT2, SalAT8, and SalAT14 are labeled. (B) Malonyl-CoA, methylmalonyl-CoA, and ethylmalonyl-CoA. (C) AT-catalyzed reactions.

how AT domains of salinomycin PKS discriminate among these α-carboxyacyl-CoA extension units, crystal structures of MCoA-specific SalAT2, MMCoA-specific SalAT8, and EMCoA-specific SalAT14 were determined. Molecular dynamics (MD) simulations were used to reveal the enzyme–substrate interactions involved in substrate binding. The hydrogen bond interactions stabilizing the substrate free carboxylate and the thioester carbonyl groups are highly conserved. As previously reported, engineering the MCoA-specific HAFH and MMCoA-specific YASH motifs significantly diminished the activity of AT domains. However, catalytically efficient mutants with switched substrate specificity were successfully generated by engineering the hydrophobic residues contacting α substituents and exhibited ~70% of the wild-type enzyme/native substrate activity.

METHODS

Protein Expression and Purification. The fragments encoding SalAT2 (AEZ53945.1), SalAT8 (AEZ53949.1), and SalAT14 (AEZ53953.1) were amplified from *Streptomyces albus* XM211 genomic DNA and cloned into pET28a by using the primers listed in Table S5. The resulting plasmids were transformed into *Escherichia coli* BL21(DE3). The transformed cells were grown to an OD_{600} of 0.4 in LB medium supplemented with 50 mg/mL kanamycin at 37 °C and induced with 0.3 mM isopropyl β-D-1-thiogalactopyranoside at 16 °C for 12 h. The cells were harvested by centrifugation at 4 °C, resuspended in lysis buffer containing 500 mM NaCl and 50 mM Tris (pH 7.5), and lysed by sonication on ice. Cell debris was removed by centrifugation at 15000*g* for 40 min at 4 °C. The supernatant was loaded onto nickel-NTA resin equilibrated with lysis buffer followed by washing with lysis buffer containing 15 mM imidazole (pH 7.5). The His-tagged protein was eluted with lysis buffer containing 300 mM imidazole and polished with a Superdex 200 gel filtration column equilibrated with 150 mM NaCl and 10 mM Tris (pH 7.5). The resulting protein was exchanged into a buffer containing 25 mM NaCl, 10 mM Tris (pH 7.5), and 1 mM DTT (pH 7.5) and concentrated to 10 mg/mL. Holo-ACP domains were expressed in *E. coli* BAP1 that contains 4′-phosphopantetheinyl transferase from *Bacillus subtilis* in the chromosome.

Kinetic Analysis. The release of CoA catalyzed by AT domains and their mutants was coupled to the generation of NADH that can be monitored spectrophotometrically at 340 nm.[13] All reactions were carried out at 25 °C in 100 μL of sodium phosphate buffer (50 mM, pH 7.5) containing 1 μM AT, 5 μM to 3.2 mM α-carboxyacyl-CoA, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM tris(2-carboxyethyl)phosphine (TCEP), 0.4 mM NAD^+, 0.4 milliunit/μL α-ketoglutarate dehydrogenase (αKGDH), 0.4 mM thiamine pyrophosphate (TPP), 0.05 mg/mL *N*,*O*-bis(trimethylsilyl) acetamide (BSA), 2 mM α-ketoglutaric acid, and 10% glycerol. The cognate holo-ACPs (100 μM) were supplemented to assay substrate specificity in AT-catalyzed transacylation reactions. Kinetic parameters were deduced by nonlinear regression analysis based on Michaelis–Menten kinetics using Origin 9.0.

Crystallization and Structure Determination. Crystals of AT domains were grown in sitting drops containing 2 μL of a protein solution and 1 μL of a precipitant solution [SalAT2, 0.1 M Tris (pH 7.5), 0.2 M NaCl, and 23% PEG 3350; SalAT8, 0.056 M NaH_2PO_4 and 1.544 M K_2HPO_4 (pH 8.2); SalAT14, 0.2 M potassium acetate (pH 8.0) and 23% PEG 3350] at 20 °C. The resulting crystals were soaked in a crystallization solution containing 20% (v/v) glycerol before being frozen in liquid nitrogen. Data were collected at Shanghai Synchrotron Radiation Facility beamlines BL17U, BL18U, and BL19U and processed with HKL2000.[20] The DEBS AT5 structure (PDB entry 2HG4)[21] was used as a search model for molecular replacement in Phaser. The models were built in Coot and refined through Refmac. Structure coordinates have been deposited in the PDB as entries 6IYO (SalAT2), 6IYR (SalAT8), and 6IYT (SalAT14).

Mutagenesis. The mutants were generated using the GeneTailor Site-Directed Mutagenesis System (Invitrogen) following the manufacturer's instructions. The oligonucleotides used for mutagenesis are listed in Table S4.

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978–2986

MD Simulations. AutoDock was used to prepare the structures of AT-acyl-CoA complexes.[22] All MD simulations were performed by using the Amber software suite.[23] Protonation states of titratable residues of the AT domains were assigned at pH 7.0 using the H++ web server and visually inspected.[24] To prepare the parameters for the acyl-CoA extension units, Gaussian 09 was used for the conformational optimization at the level of HF/6-31G(d) and the computations of the electrostatic surface potential (ESP) charge.[25] The restrained electrostatic potential (RESP) charge was fitted on the substrate by using the antechamber program. The parmchk2 program was used to prepare the missing parameters. Topology and coordinate files for the AT−substrate complexes were prepared by using the tleap module. Following parametrization, the AT−substrate complexes were solvated in a cubic box of TIP3P water molecules with a water thickness extending 10 Å from the protein surface. The systems were neutralized by addition of sodium ions.

MD simulations were performed using the PMEMD.cuda program of the AMBER 18 package. A cutoff of 10 Å was used for nonbonded interactions. Long-range electrostatic interactions were calculated by the particle mesh Ewald (PME) method. The lengths of bonds involving hydrogen atoms were fixed with the SHAKE algorithm. Periodic boundary conditions simulated the effects of a larger system size. Minimization was performed in two steps to correct any possibly unrealistic arrangements. The first step involved the relaxation of the water molecules only, while the second step involved the minimization of the whole system. The Langevin dynamics with a collision frequency of 2 ps^{-1} was then used to gradually increase the system temperature from 0 to 300 K over 50 ps. Prior to production simulations, the system was subjected to a 50 ps equilibration with constant pressure and temperature (*NPT*) to adjust the density followed by a second equilibration step of 1 ns. Finally, 50 ns production simulations without any restraint were performed under *NPT* conditions. Each simulation was repeated three times with a different random number, and one of them was extended to a 150 ns simulation. An integration time step of 2 fs was utilized with structural snapshots being extracted every 1000 steps. The simulation trajectory was analyzed by the cpptraj in Amber tools18. The hydrogen bonding analysis was performed with default options. Root-mean-square deviation (RMSD)-based clustering was performed to extract the representative structures for comparisons.

RESULTS

Specificity in Hydrolytic Reactions. Discrete SalAT2 (residues 3201−3642 of SlnA1), SalAT8 (residues 480−929 of SlnA5), and Sal14 (residues 465−910 of SlnA9), including corresponding KS−AT linkers, were overexpressed and purified to homogeneity.[18] MCoA, MMCoA, and EMCoA were used to evaluate AT specificity by monitoring the release of free CoA.[13] As expected, they all show a preference for their native substrates (Figure 2, Figure S1, and Tables S1−S3). The k_{cat}/K_m values are 4−20-fold higher for the native substrate than for non-native substrates. SalAT2 and SalAT8 prefer substrates with smaller α substituent groups, while SalAT14 prefers substrates with larger α substituent groups. All of the tested AT domains have K_m values for their cognate substrates in the range of 13−20 μM, lower than the concentration of MCoA (35 μM) in exponentially growing *E. coli*.[26] The K_m values increase 10−50-fold with these ATs working on non-native substrates except that SalAT2 has similar K_m values for its native MCoA and non-native MMCoA. All three AT domains show promiscuity toward the tested extender units. However, because the CoA release assays were carried out in the absence of cognate holo-ACPs, the promiscuity observed here may not reflect the characteristics of AT domains in transacylation reactions. The salinomycin analogues with altered extender units are not observed in the fermentation products of the producing strain.

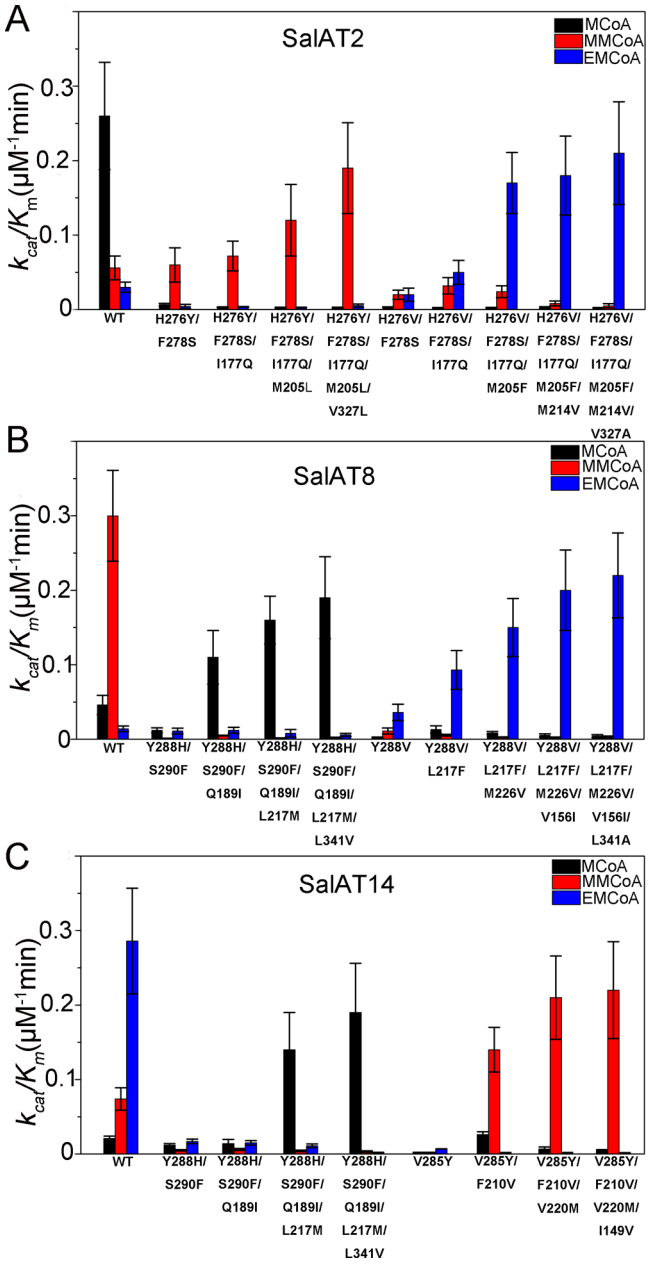


Figure 2. k_{cat}/K_m values of the ATs and their mutants. (A) SalAT2 and its mutant. (B) SalAT8 and its mutants. (C) SalAT14 and its mutants. WT, wild type. The WT enzymes show a preference for their native substrates. The k_{cat}/K_m values are 4−20-fold higher for the native substrate than for non-native substrates. The substrate specificity can be switched by engineering AT active site residues.

Crystal Structures of AT Domains. A moderate level of overall sequence identity was observed between SalAT2 and SalAT8 (38%), SalAT2 and SalAT14 (34%), and SalAT8 and SalAT14 (50%) (Figure S2). To understand the molecular

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

basis for their substrate specificity, crystal structures of three AT domains were determined by molecular replacement using the structure of EryAT5 as a search model and refined to 2.20 Å (SalAT2), 2.05 Å (SalAT8), and 1.78 Å (SalAT14) resolution (Table 1). The 47 N-terminal residues of each

Table 1. Data Collection and Refinement Statistics (molecular replacement)

	SalAT2	SalAT8	SalAT14
	Data Collection		
space group	$P22_12_1$	$P4_22_12$	$C2$
a, b, c (Å)	49.093, 73.367, 248.000	123.244, 123.244, 78.532	120.818, 78.474, 106.714
α, β, γ (deg)			117.00
resolution (Å)	50−2.20	50−2.05	50−1.78
R_{merge}	0.087 (0.51)	0.087 (0.66)	0.074 (0.26)
$I/\sigma I$	21.7 (4.7)	56.4 (5.3)	17.0 (4.2)
$CC_{1/2}$	0.997 (0.691)	0.997 (0.838)	0.993 (0.873)
completeness (%)	90.8 (87.9)	100 (100)	99.9 (99.7)
redundancy	7.2 (7.5)	14.5 (14.7)	3.4 (3.2)
	Refinement		
resolution (Å)	50−2.20	50−2.05	50−1.78
no. of reflections	40302	35404	80596
R_{work}/R_{free}	0.22/0.26	0.20/0.24	0.18/0.21
no. of atoms			
protein	6231	3098	6316
PO_4		5	
water	36	88	440
B-factor			
protein	59	38	14
PO_4		63	
water	35	36	17
RMSD			
bond lengths (Å)	0.009	0.012	0.013
bond angles (deg)	1.146	1.516	1.054
Ramachandran plot (%)			
favored	98.0	95.7	97.0
allowed	1.9	4.0	2.8
outliers	0.1	0.3	0.2

SalAT2 monomer in the asymmetric unit fold onto the other molecule, resulting in the close contact of the KS−AT linker domains (Figure S3). Similar nonphysiological domain swapping has been observed in the structure of the KR−ER didomain from the second module of the spinosyn PKS.[27] The overall structures of the three AT domains closely resemble those of EryAT3 and EryAT5, containing a large α/β-hydrolase subdomain and a small ferredoxin-like subdomain (Figure 3).[21,28] The Ser-His catalytic dyads at the interface of the two subdomains are orientated as observed in previously reported AT domains. The catalytic serine (Ser176 in SalAT2, Ser188 in SalAT8, and Ser181 in SalAT14) in the highly conserved GHSXG motifs is located in a sharp turn between a β-sheet and an α-helix, while the catalytic histidine (His279 in SalAT2, His291 in SalAT8, and His288 in SalAT14) in the substrate-specific motif (^{276}HAFH^{279} in SalAT2, ^{288}YASH^{291} in SalAT8, and ^{285}VASH^{288} in SalAT14) is positioned in a loop connecting two subdomains (Figure 3B−D). The similarly

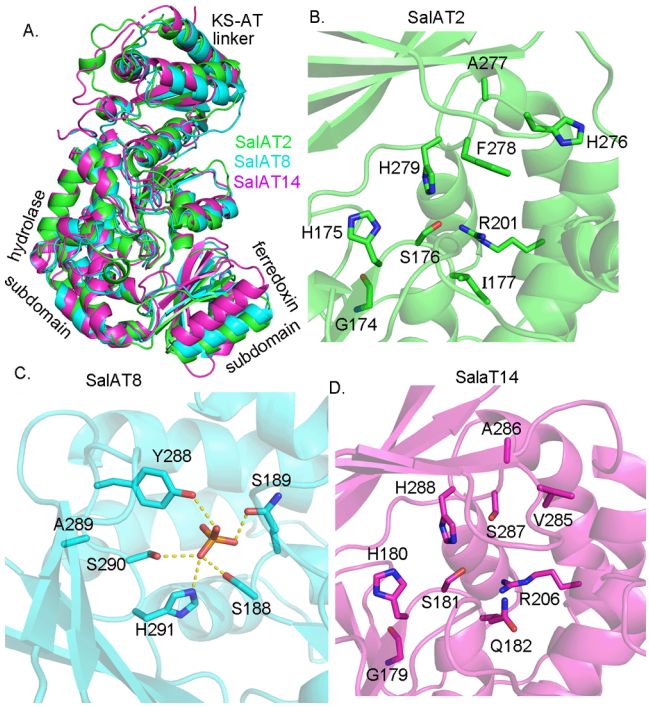


Figure 3. Structures of AT domains. (A) Overall structures. Active sites of (B) SalAT2, (C) SalAT8, and (D) SalAT14. The active site residues are shown as sticks. The hydrogen bonds involved in the binding of the phosphate group are shown as dashed lines.

arranged His-Ser catalytic dyads suggest a catalytic mechanism analogous to other PKS AT domains.

A phosphate group from the crystallization buffer is observed in the active site of SalAT8 (Figure 3C). This molecule of phosphate makes contacts with several active site residues, including S188, Q189, Y288, S290, H291, and R213. The substrate-specific HAFH and YASH motif are clustered in the substrate binding pocket. The H276 imidazole ring of SalAT2 is placed outside the substrate binding pocket and interacts with the other loop connecting two subdomains, while the bulky benzyl group of F278 is placed inside the substrate binding pocket and obviously reduces the space available to accommodate the substrate (Figure 3B). Both Y288 and S290 are placed in the interior of the SalAT8 active site. The YASH motif is replaced by the VASH motif in SalAT14 that has a larger pocket to accommodate the ethyl side chain (Figure 3D).

MD Simulations of AT−Acyl−CoA Complexes. Attempts to obtain crystal structures of AT−acyl−CoA complexes were unsuccessful. Although crystals of the wild-type enzymes and the SalAT2S176A, SalAT2S176C, SalAT8-S188A, SalAT8S188C, SalAT14S181A, and SalAT14S181C mutants could be grown in or soaked in millimolar concentrations of substrates, no density for the corresponding substrate was observed in the electron density maps. The structure of fatty acid synthase AT (FabD) complexed with MCoA was used to guide the modeling of MCoA into the active site of SalAT2.[29] The resulting model of SalAT2 with MCoA was then used as the template for preparing other SalAT8−MMCoA and SalAT14−EMCoA complexes. All of these complexes were subjected to MD simulations to further explore the enzyme−substrate interactions. The RMSD of the backbone atoms during the entire MD trajectory was analyzed

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

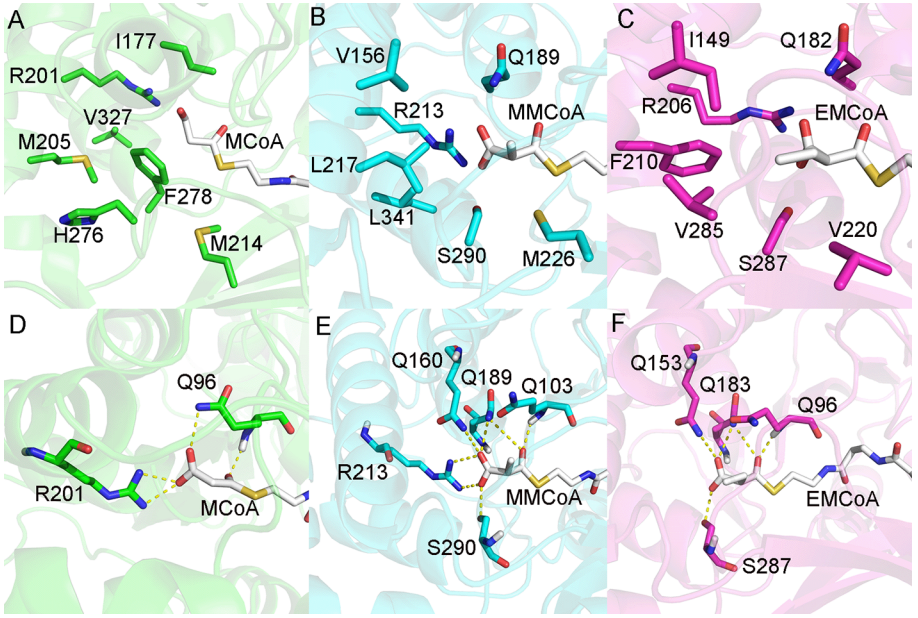


Figure 4. Enzyme−substrate interactions revealed by MD simulation showing (A) the binding of MCoA in SalAT2, (B) the binding of MMCoA in SalAT8, and (C) the binding of EMCoA in SalAT14. The residues involved in engineering are labeled. Hydrogen bond interactions between (D) SalAT2 and MCoA, (E) SalAT8 and MMCoA, and (F) SalAT14 and EMCoA are shown as dashed lines.

to confirm the stability of the MD simulation (Figure S4). We focused on the molecular interactions between the substrate acyl group and the AT active site residues. As expected, the MD simulations show that the conserved hydrogen bond interactions with the free carboxylate and the thioester carbonyl are the main factors governing the correct positions of the substrate acyl groups (Figure 4). Intriguingly, the SalAT14 R206 side chain sways away from the acyl binding pocket and is not involved in the stabilization of the free carboxylate group. Although crystal structures show this arginine of SalAT14 is in an unusual orientation compared to those of other ATs (Figure S5), a substrate-bound structure is necessary to reveal the role of this arginine. The bulky side chain of F278 is located on top of the Cα atom of MCoA, consistent with its role in dominating the substrate specificity of SalAT2. The α substituents of MMCoA and EMCoA are similarly positioned in SalAT8 and SalAT14. The methyl group is placed in the hydrophobic pocket formed by Q103, Q160, L217, M226, V286, Y288, S290, and L341 in SalAT8, while the ethyl group is placed in the hydrophobic pocket formed by Q96, I149, Q153, F210, V220, T284, V285, S287, and C338 in SalAT14. Obviously, the tyrosine (Y288 in SalAT8) to valine (V220 in SalAT14) mutation provides more space to accommodate the larger ethyl group of EMCoA.

Previous efforts have revealed that engineering the HAFH and YASH motifs is insufficient to switch the specificity of AT domains and constantly results in promiscuous mutants with significantly diminished activity.[8,13,15−17] Extensive interactions between the substrate acyl group and the AT active site are revealed by the MD simulations of the salinomycin AT domains. Residues involved in hydrogen bond interactions are highly conserved, suggesting that the residues contacting the α substituents are the main factors discriminating between different acyl-CoA substrates (Figure S2). We envisioned that the substrate preference of AT domains could be altered by replacing some of these residues with the corresponding residues from an AT with expected specificity.

Engineering Specificity of SalAT14. SalAT14 represents the first AT structure specific for EMCoA to the best of our knowledge and therefore is the first target we tried to engineer. Wild-type SalAT14 shows a moderate specificity for its native substrate EMCoA (k_{cat}/K_m = 0.29 ± 0.096 μM^{-1} min^{-1}; k_{cat} = 5.11 ± 1.38 min^{-1}; K_m = 17.67 ± 3.37 μM) with a k_{cat}/K_m value ~14.5-fold higher than that of MCoA and ~4.0-fold higher than that of MMCoA. To switch its substrate specificity from EMCoA to MMCoA, V285 was mutated to tyrosine to restore the YASH motif that extensively interacts with MMCoA. Kinetic analysis of the V285Y mutant revealed drastically decreased activity for all three substrates (Figure 2C and Table S3). We next turned to inspect the hydrophobic residues revealed by MD simulations. F210 contacts the ethyl α substituent of EMCoA in wild-type SalAT14. In the V285Y mutant, it likely affects the binding of MMCoA by orienting the side chain of the newly introduced tyrosine (Figure 4C). Encouragingly, a V285Y/F210V double mutant showed an ~64-fold increase in k_{cat}/K_m for MMCoA, an ~12-fold increase in k_{cat}/K_m for MCoA, and an ~4-fold decrease in k_{cat}/K_m for EMCoA. M226 of SalAT8 contacting the methyl group of MMCoA has a larger side chain compared to that of the corresponding V220 of SalAT14. Therefore, a V285Y/F210V/V220M triplet mutant of SalAT14 was generated and showed a further ~1.5-fold increase in k_{cat}/K_m with MMCoA but an ~4-fold decrease in k_{cat}/K_m with MCoA. The effect of the V220M mutation on the k_{cat}/K_m with EMCoA was negligible. To further increase the specificity for MMCoA, the I149 residue that contacts Y285 was mutated to a valine with a smaller side chain that was observed in SalAT8 (V156). The V285Y/F210V/V220M/I149V quadruple mutant showed an ~147-fold higher k_{cat}/K_m for MMCoA than for EMCoA and an ~49-fold higher k_{cat}/K_m for MMCoA than for MCoA. The k_{cat}/K_m value of the quadruple mutant for MMCoA (k_{cat}/K_m = 0.22 ± 0.088 μM^{-1} min^{-1}; k_{cat} = 4.61 ± 1.4 min^{-1}; K_m = 20.65 ± 5.64 μM) is comparable to that of wild-type SalAT8 for its native MMCoA substrate (k_{cat}/K_m = 0.30 ± 0.078 μM^{-1} min^{-1}; k_{cat} =

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

5.82 ± 1.1 min^{-1}; K_m = 19.68 ± 3.47 μM) (Figure 2C and Table S3).

To switch the substrate specificity of SalAT14 from EMCoA to MCoA, the V285H and S287F mutations were introduced into the active site to restore the HAFH motif correlated to MCoA and significantly decreased the activity for all three substrates (Figure 2C and Table S3). Q182 of SalAT14 contacts both the carboxylate group and the thioester carbonyl group of EMCoA (Figure 4C). In SalAT2, the corresponding position is occupied by I177. However, the introduction of the Q182I mutation resulted in negligible effects on enzyme activity. F210 might affect the side-chain orientation of the introduced F287, and a methionine is observed at the corresponding position of SalAT2 (M205) (Figure 4A,C). A V285H/S287F/Q182I/F210M quadruple mutant was generated and had an ~10-fold increase in k_{cat}/K_m for MCoA. To further increase the activity for MCoA, V220 of SalAT14 was mutated to a methionine residue that is observed in SalAT2 (M214). The generated V285H/S287F/Q182I/F210M/V220M pentad mutant shows an ~58-fold higher k_{cat}/K_m for MCoA than for MMCoA and an ~136-fold higher k_{cat}/K_m for MCoA than for EMCoA. The k_{cat}/K_m value of the pentad mutant of SalAT14 for MCoA (k_{cat}/K_m = 0.19 ± 0.089 μM^{-1} min^{-1}; k_{cat} = 4.62 ± 1.6 min^{-1}; K_m = 24.24 ± 7.28 μM) is close to that of wild-type SalAT2 for its native MCoA substrate (k_{cat}/K_m = 0.26 ± 0.082 μM^{-1} min^{-1}; k_{cat} = 3.57 ± 0.77 min^{-1}; K_m = 13.59 ± 3.12 μM) (Figure 2C and Table S3).

Engineering Specificity of SalAT8. The MMCoA-specific SalAT8 shows a k_{cat}/K_m value ~6.5-fold higher than that for MCoA and ~21.4-fold higher than that for EMCoA (Figure 2 and Table S2). Kinetic assays of SalAT14 mutants revealed the involvement of three hydrophobic residues (SalAT14 numbering, 149, 210, and 220) in AT substrate specificity and encouraged us to manipulate the corresponding residues in SalAT8 to alter its specificity (Figure 4B,C). The Y288V mutation was introduced to provide more space to accommodate the ethyl side chain. The mutant exhibited an ~2.6-fold increase in k_{cat}/K_m with EMCoA but a >20-fold decrease in k_{cat}/K_m with MCoA and MMCoA (Figure 2B and Table S2). The L217F, M226V, and V156I mutations were successively introduced to optimize the hydrophobic enzyme−substrate interactions and switched the specificity of SalAT8 from MMCoA to EMCoA. The L341 that might affect the orientation of the introduced F217 has a larger side chain compared to that of the corresponding C338 of SalAT14. The introduction of the L341A mutation further increased the k_{cat}/K_m for EMCoA slightly. The resulting pentad mutant (Y228V/L217F/M226V/V156I/L341A) shows at least 50-fold specificity for EMCoA (k_{cat}/K_m = 0.22 ± 0.067 μM^{-1} min^{-1}; k_{cat} = 4.55 ± 0.97 min^{-1}; K_m = 20.54 ± 4.48 μM) than for MCoA and MMCoA (Figure 2B and Table S2).

The aforementioned kinetic analysis of AT mutants suggests that four hydrophobic residues (SalAT14 numbering, 149, 210, 220, and 338) are involved in the substrate specificity of AT domains in addition to the conserved HAFH/YASH motifs (Figure 4C). We then tried to switch the specificity of SalAT8 from MMCoA to MCoA by engineering the corresponding residues (Figure 4A,B). As observed in engineering the specificity of SalAT14 for MCoA, theY288H/S290F mutant has attenuated activity for all three substrates. Successive introduction of Q189I, L217M, and L341V mutations increased the activity and specificity for MCoA (k_{cat}/K_m = 0.19 ± 0.065 μM^{-1} min^{-1}; k_{cat} = 4.97 ± 1.24 min^{-1}; K_m = 27.16 ± 6.48 μM) close to that of wild-type SalAT2 for MCoA (Figure 2B and Table S2).

Engineering the Specificity of SalAT2. WT SalAT2 shows a moderate specificity for its native substrate MCoA with a k_{cat}/K_m value ~4.6-fold higher than that for MMCoA and ~8.7-fold higher than that for EMCoA. The H276Y/F278S mutant has a slightly increased activity for MMCoA but a significantly decreased activity for MCoA and EMCoA. Successive introduction of I177Q, M205L, and V327L mutations significantly increased the activity and specificity for MMCoA (k_{cat}/K_m = 0.19 ± 0.099 μM^{-1} min^{-1}; k_{cat} = 4.19 ± 1.45 min^{-1}; K_m = 22.55 ± 8.78 μM) (Figure 2A and Table S1).

The H276V/F278S double mutant showed attenuated activity for all three substrates. Successive introduction of I177Q, M205F, M214V, and V327A mutations increased the activity and specificity for EMCoA (k_{cat}/K_m = 0.21 ± 0.089 μM^{-1} min^{-1}; k_{cat} = 4.60 ± 1.51 min^{-1}; K_m = 22.10 ± 5.90 μM) comparable to that of wild-type SalAT14 for EMCoA (Figure 2A and Table S1).

Specificity in the Presence of Cognate Holo-ACP. The cognate holo-ACPs were supplemented to quantify the substrate specificity of the ATs and their mutants in the presence of cognate holo-ACPs. As shown in Figure 5, all wild-type AT domains had >8-fold specificity for their cognate acyl-CoA substrates. Of the three, SalAT8 had the highest substrate specificity with a k_{cat}/K_m for MMCoA nearly 20-fold higher than the k_{cat}/K_m for either MCoA or EMCoA (Table S4). Encouragingly, the AT mutants with switched specificity in hydrolytic reactions did show the expected specificity for the acyl-CoA substrates in the presence of cognate holo-ACPs. They all had 3−14-fold higher K_m values compared to those of the wild-type enzymes, suggesting that further optimizations were necessary. However, quantifying the amounts of acylated ACPs is necessary for measuring the substrate specificity in the transacylation reactions because the release of CoA reflects only the acyl transfer from the CoA to the active site serine of an AT.

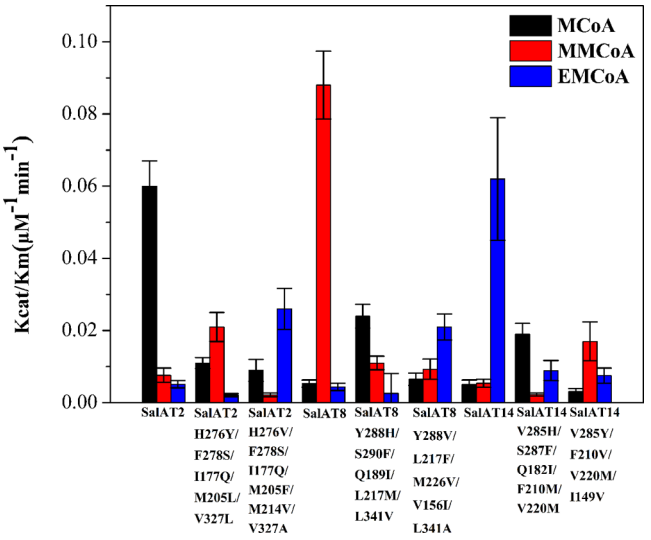


Figure 5. Specificity of ATs and their mutants in the presence of cognate holo-ACPs. The wild-type ATs have >8-fold specificity for their cognate CoA substrates. The AT mutants show expected specificity.

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

DISCUSSION

AT domains of modular PKS control the structural diversity of the resulting polyketides by selecting and incorporating monomeric carboxylate extension units and are therefore attractive engineering targets for generating novel, therapeutically active polyketides by site-specific incorporations of non-native extension units. Swapping of entire AT domains for homologues with different specificities is a useful approach to regiospecifically installing alternative extender units into polyketides, whereas chimeric PKSs often display significantly reduced activity compared to those of their wild-type counterparts, likely due to the disruption of proper protein−protein interactions required for kinetically efficient polyketide chain elongation or the inability of downstream domains to process altered polyketide intermediates.[2,5] An alternative strategy explores active site mutagenesis to shift the substrate specificity of AT domains toward non-native extension units. AT domain swapping of DEBS module 4 fails to produce the expected derivatives, while mutagenesis in the active site of DEBS AT4 results in a functional enzyme that produces both 6-deoxyerythronolide B (6-dEB) and the desired 6-desmethyl derivative.[15] Encouragingly, non-natural extension units, for example, propargylmalonyl-CoA, can be installed in the expected position of a polyketide product by DEBS carrying an engineered AT6.[6−8]

Most AT domains display tight substrate specificity in their natural environment. Previous active site engineering efforts to shift AT specificity have focused on the substrate-specific motifs that are identified by multiple-sequence alignments. An HAFH motif indicates a malonyl-CoA-specific AT domain, while a YASH motif indicates a methylmalonyl-CoA-specific AT domain. Crystal structures of AT domains confirm the involvement of HAFH and YASH motifs in substrate binding. However, the manipulation of YASH and HAFH motifs consistently results in promiscuous mutants that can incorporate both extension units.[8,15,16] Kinetic assays reveal that the relaxed specificity is caused by a drastic decrease in catalytic efficiency toward the native substrate rather than enhanced catalytic efficiency toward the non-native substrate, suggesting further improvements are necessary for the current active site engineering strategy to alter AT substrate specificity.[13] A better understanding of the specificity determinants would greatly assist the engineering of AT domains.

We opted for a combined structural and computational approach to seek new insights into the molecular basis for extension unit selection exercised by AT domains. Crystal structures have been used to guide previous engineering efforts of AT domains but failed to pinpoint new specificity determinants in addition to these that have already been identified by multiple-sequence alignments. A residue that constrains the side-chain size of an extension unit in DEBS AT6 is revealed by MD simulations followed by quantum mechanics/molecular mechanics optimizations.[7,8] A non-natural substrate with a larger side chain, 2-propargylmalonyl-SNAC, is successfully incorporated by mutating this residue from valine to alanine. A Y189R mutant of DEBS AT6 with switched specificity from MMCoA to propargylmalonyl-CoA is identified by saturation mutagenesis of the YASH motif.[6]

MCoA, MMCoA, and, to a lesser extent, EMCoA are the most commonly used extension units and provide hydrogen, methyl, and ethyl substituents, respectively, in the assembly of polyketide chains. We determined the crystal structures of SalAT2 that selects MCoA, SalAT8 that selects MMCoA, and SalAT14 that selects EMCoA to help understand how these AT domains efficiently discriminate the three substrates that differ by only one carbon. These structures enabled us to carefully compare the enzyme−substrate interactions of AT domains with different extension unit specificities by using MD simulations. The hydrophobic enzyme−substrate interactions obviously play important roles in discriminating between different acyl-CoA substrates because the hydrogen bond interactions are highly conserved among AT domains with different substrate specificities. Four hydrophobic residues (SalAT14 numbering, I149, F210, V220, and C338) were revealed to be the substrate specificity determinants in addition to the previously reported YASH/HAFH motifs. As previously reported, manipulation of the YASH/HAFH motifs was insufficient to switch AT specificity from one extension unit to another.[13] Here, we consistently generated mutants with decreased activity for all three tested substrates by engineering these motifs. However, the specificity and activity for a targeted extension unit could be successively increased by engineering the specificity determinants identified in the study presented here. The introduced residues were selected by comparing with AT domains that show the expected specificity. By this strategy, the specificity of AT domains for MCoA, MMCoA, and EMCoA can be rationally designed. Kinetic analysis confirmed that the k_{cat}/K_m values of these mutants were close to that of wild-type enzymes for their native substrates. The difference in substrate specificity seems to emerge mainly from the binding step (related to K_m), although contributions from the chemical step (related to k_{cat}) cannot be ruled out. Because AT domains of modular PKSs exhibit significant sequence and structural similarity, the strategy reported here could potentially be a generalized strategy for engineering the specificity of AT domains toward MCoA, MMCoA, and EMCoA extension units (Figure S2).

In summary, understanding the molecular basis for the substrate specificity of AT domains is crucial to introducing non-native extension units into polyketides by rational engineering of modular PKSs. In this work, we identified additional specificity determinants of PKS AT domains by combined structural and computational approaches. Catalytically efficient AT mutants with switched specificity toward MCoA, MMCoA, and EMCoA were generated by engineering these specificity determinants and previously reported motifs simultaneously, encouraging the view that polyketide derivatives with desired site-specific modifications could be efficiently generated by modular PKSs with engineered active sites.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00305.

Five figures and five tables (PDF)

Accession Codes

The atomic coordinates and structure factors have been deposited in the RCSB PDB (entries 6IYO, 6IYR, and 6IYT).

AUTHOR INFORMATION

Corresponding Authors

*E-mail: yileizhao@sjtu.edu.cn.

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

*E-mail: jtzheng@sjtu.edu.cn.

ORCID

Ting Shi: 0000-0003-3921-4412
Yilei Zhao: 0000-0003-4687-7847
Jianting Zheng: 0000-0003-1250-3556

Funding

This work was supported by the National Natural Science Foundation of China (31570056 and 31770068).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Crystal diffraction data were collected at Shanghai Synchrotron Radiation Facility beamlines BL17U, BL18U, and BL19U.

■ REFERENCES

(1) Robbins, T., Liu, Y. C., Cane, D. E., and Khosla, C. (2016) Structure and mechanism of assembly line polyketide synthases. *Curr. Opin. Struct. Biol. 41*, 10−18.

(2) Dunn, B. J., and Khosla, C. (2013) Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *J. R. Soc., Interface 10*, 20130297.

(3) Nigam, A., Almabruk, K. H., Saxena, A., Yang, J., Mukherjee, U., Kaur, H., Kohli, P., Kumari, R., Singh, P., Zakharov, L. N., Singh, Y., Mahmud, T., and Lal, R. (2014) Modification of Rifamycin Polyketide Backbone Leads to Improved Drug Activity against Rifampicin-resistant Mycobacterium tuberculosis. *J. Biol. Chem. 289*, 21142−21152.

(4) Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E., and Khosla, C. (2014) Comparative analysis of the substrate specificity of trans- versus cis-acyltransferases of assembly line polyketide synthases. *Biochemistry 53*, 3796−3806.

(5) Yuzawa, S., Deng, K., Wang, G., Baidoo, E. E., Northen, T. R., Adams, P. D., Katz, L., and Keasling, J. D. (2017) Comprehensive in Vitro Analysis of Acyltransferase Domain Exchanges in Modular Polyketide Synthases and Its Application for Short-Chain Ketone Production. *ACS Synth. Biol. 6*, 139−147.

(6) Koryakina, I., Kasey, C., McArthur, J. B., Lowell, A. N., Chemler, J. A., Li, S., Hansen, D. A., Sherman, D. H., and Williams, G. J. (2017) Inversion of Extender Unit Selectivity in the Erythromycin Polyketide Synthase by Acyltransferase Domain Engineering. *ACS Chem. Biol. 12*, 114−123.

(7) Bravo-Rodriguez, K., Klopries, S., Koopmans, K. R., Sundermann, U., Yahiaoui, S., Arens, J., Kushnir, S., Schulz, F., and Sanchez-Garcia, E. (2015) Substrate Flexibility of a Mutated Acyltransferase Domain and Implications for Polyketide Biosynthesis. *Chem. Biol. 22*, 1425−1430.

(8) Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., and Schulz, F. (2013) Enzyme-Directed Mutasynthesis: A Combined Experimental and Theoretical Approach to Substrate Recognition of a Polyketide Synthase. *ACS Chem. Biol. 8*, 443−450.

(9) Kalkreuter, E., CroweTipton, J. M., Lowell, A. N., Sherman, D. H., and Williams, G. J. (2019) Engineering the Substrate Specificity of a Modular Polyketide Synthase for Installation of Consecutive Non-Natural Extender Units. *J. Am. Chem. Soc. 141*, 1961−1969.

(10) Vogeli, B., Geyer, K., Gerlinger, P. D., Benkstein, S., Cortina, N. S., and Erb, T. J. (2018) Combining Promiscuous Acyl-CoA Oxidase and Enoyl-CoA Carboxylase/Reductases for Atypical Polyketide Extender Unit Biosynthesis. *Cell Chem. Biol. 25*, 833−839.

(11) 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.

(12) Paiva, P., Sousa, S. F., Ramos, M. J., and Fernandes, P. A. (2018) Understanding the Catalytic Machinery and the Reaction Pathway of the Malonyl-Acetyl Transferase Domain of Human Fatty Acid Synthase. *ACS Catal. 8*, 4860−4872.

(13) Dunn, B. J., Cane, D. E., and Khosla, C. (2013) Mechanism and Specificity of an Acyltransferase Domain from a Modular Polyketide Synthase. *Biochemistry 52*, 1839−1841.

(14) Jenke-Kodama, H., Sandmann, A., Muller, R., and Dittmann, E. (2005) Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol. 22*, 2027−2039.

(15) Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R., and McDaniel, R. (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry 40*, 15464−15470.

(16) Del Vecchio, F., Petkovic, H., Kendrew, S. G., Low, L., Wilkinson, B., Lill, R., Cortes, J., Rudd, B. A., Staunton, J., and Leadlay, P. F. (2003) Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol. 30*, 489−494.

(17) Petkovic, H., Sandmann, A., Challis, I. R., Hecht, H. J., Silakowski, B., Low, L., Beeston, N., Kuscer, E., Garcia-Bernardo, J., Leadlay, P. F., Kendrew, S. G., Wilkinson, B., and Muller, R. (2008) Substrate specificity of the acyl transferase domains of EpoC from the epothilone polyketide synthase. *Org. Biomol. Chem. 6*, 500−506.

(18) Jiang, C., Wang, H., Kang, Q., Liu, J., and Bai, L. (2012) Cloning and characterization of the polyether salinomycin biosynthesis gene cluster of Streptomyces albus XM211. *Appl. Environ. Microbiol. 78*, 994−1003.

(19) Yurkovich, M. E., Tyrakis, P. A., Hong, H., Sun, Y. H., Samborskyy, M., Kamiya, K., and Leadlay, P. F. (2012) A Late-Stage Intermediate in Salinomycin Biosynthesis Is Revealed by Specific Mutation in the Biosynthetic Gene Cluster. *ChemBioChem 13*, 66−71.

(20) Wang, Q. S., Zhang, K. H., Cui, Y., Wang, Z. J., Pan, Q. Y., Liu, K., Sun, B., Zhou, H., Li, M. J., Xu, Q., Xu, C. Y., Yu, F., and He, J. H. (2018) Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. *Nucl. Sci. Tech. 29*, 68.

(21) Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7-Angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U. S. A. 103*, 11124−11129.

(22) Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem. 30*, 2785−2791.

(23) Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E. I., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Ghoreishi, D., Gilson, M. K., Gohlke, H., Goetz, A. W., Greene, D., Harris, R., Homeyer, N., Izadi, S., Kovalenko, A., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Mermelstein, D. J., Merz, K. M., Miao, Y., Monard, G., Nguyen, C., Nguyen, H., Omelyan, I., Onufriev, A., Pan, F., Qi, R., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Smith, J. A., Salomon-Ferrer, R., Swails, J., Walker, R. C., Wang, J., Wei, H., Wolf, R. M., Wu, X., Xiao, L., York, D. M., and Kollman, P. A. (2018) *AMBER 2018*, University of California, San Francisco.

(24) Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. (2012) H++3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res. 40*, W537−W541.

(25) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N.,

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986

Keith, T., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., and Fox, D. J. (2009) *Gaussian 09*, Gaussian, Inc., Wallingford, CT.

(26) Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol. 5*, 593−599.

(27) Zheng, J., Gay, D. C., Demeler, B., White, M. A., and Keatinge-Clay, A. T. (2012) Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nat. Chem. Biol. 8*, 615−621.

(28) 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.

(29) 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., Sect. D: Biol. Crystallogr. 62*, 613−618.

DOI: 10.1021/acs.biochem.9b00305
Biochemistry 2019, 58, 2978−2986