Structural and Functional Studies on SCO1815: A β-Ketoacyl-Acyl Carrier Protein Reductase from *Streptomyces coelicolor* A3(2)[∞]

Yinyan Tang,[‡] Ho Young Lee,[‡] Yi Tang,[§,||] Chu-Young Kim,[‡,§,⊥] Irimpan Mathews,[@] and Chaitan Khosla*[,‡,§,+]

Department of Chemistry and Chemical Engineering, Stanford University, Stanford, California 94305, Stanford Synchrotron Radiation Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, and Department of Biochemistry, Stanford University, Stanford, California 94305

Received June 14, 2006; Revised Manuscript Received September 24, 2006

ABSTRACT: Aromatic polyketides are medicinally important natural products produced by type II polyketide synthases (PKSs). Some aromatic PKSs are bimodular and include a dedicated initiation module which synthesizes a non-acetate primer unit. Understanding the mechanism of this initiation module is expected to further enhance the potential for regiospecific modification of bacterial aromatic polyketides. A typical initiation module is comprised of a ketosynthase (KS), an acyl carrier protein (ACP), a malonyl-CoA: ACP transacylase (MAT), an acyl-ACP thioesterase, a ketoreductase (KR), a dehydratase (DH), and an enoyl reductase (ER). Thus far, the identities of the ketoreductase, dehydratase, and enoyl reductase remain a mystery because they are not encoded within the PKS biosynthetic gene cluster. Here we report that SCO1815 from *Streptomyces coelicolor* A3(2), an uncharacterized homologue of a NADPH-dependent ketoreductase, recognizes and reduces the β-ketoacyl-ACP intermediate from the initiation module of the R1128 PKS. SCO1815 exhibits moderate specificity for both the acyl chain and the thiol donor. The X-ray crystal structure of SCO1815 was determined to 2.0 Å. The structure shows that SCO1815 adopts a Rossmann fold and suggests that a conformational change occurs upon cofactor binding. We propose that a positively charged patch formed by three conserved residues is the ACP docking site. Our findings provide new engineering opportunities for incorporating unnatural primer units into novel polyketides and shed light on the biology of yet another cryptic protein in the *S. coelicolor* genome.

Polyketides are structurally diverse and medicinally important natural products (*1*). They are produced as secondary metabolites primarily by bacteria, fungi, and plants (*2*) and comprise many of the widely used drugs in the world (*3*). Polyfunctional aromatic products of *Streptomyces* spp. such as the tetracyclines and doxorubicin (Figure 1) are an important subclass of polyketides. Aromatic polyketides are biosynthesized through repeated decarboxylative condensations between malonyl-CoA-derived building blocks. Their synthesis is catalyzed by type II polyketide synthases (PKSs),[1] which share a common architecture and mechanism with type II fatty acid synthases found in bacteria and plants (*2*). A minimal type II PKS is comprised of four proteins, including the ketosynthase (KS), the chain length factor (CLF), the acyl carrier protein (ACP), and the malonyl-CoA:

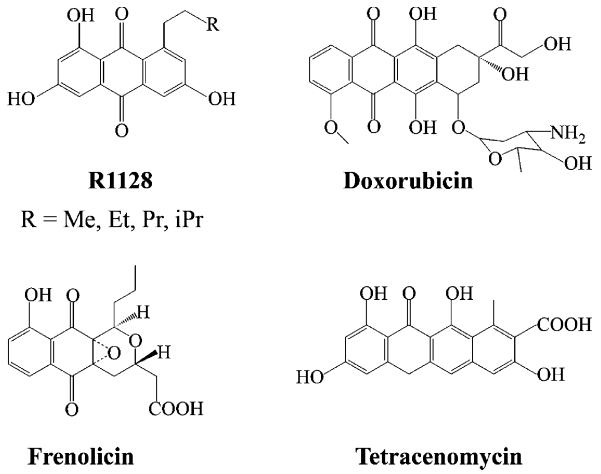

FIGURE 1: Bacterial aromatic polyketides. Doxorubicin and tetracycline are primed with acetate primers. Frenolicin is primed with a butyryl group, and R1128 is primed with a variety of acyl groups as shown.

ACP transacylase (MAT) (*4*−*12*). The KS and the CLF form a heterodimer which catalyzes condensation reactions between successive malonyl units and controls the overall polyketide chain length. The malonyl extender unit is shuttled to the active site of the KS-CLF in the form of a malonyl-ACP, which in turn is derived via MAT-catalyzed acyl transfer from malonyl-CoA.

The enormous diversity of naturally occurring aromatic polyketides is collectively derived from variations in the

[∞] The coordinates for SCO1815 have been deposited as Protein Data Bank entry 2NM0.

* To whom correspondence should be addressed. Phone and fax: (650) 723-6538. E-mail: khosla@stanford.edu.

[‡] Department of Chemistry, Stanford University.

[§] Department of Chemical Engineering, Stanford University.

[||] Current address: Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095.

[⊥] Current address: Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543.

[@] Stanford Synchrotron Radiation Laboratory.

[+] Department of Biochemistry, Stanford University.

[1] Abbreviations: PKS, polyketide synthase; KS, ketosynthase; CLF, chain length factor; ACP, acyl carrier protein; MAT, malonyl-CoA: ACP transferase; KR, ketoreductase; DH, dehydratase; ER, enoyl reductase; CoA, coenzyme A; SDR, short chain dehydrogenese/reductase; rmsd, root-mean-square deviation.

10.1021/bi061187v CCC: $33.50
Published on Web 11/03/2006

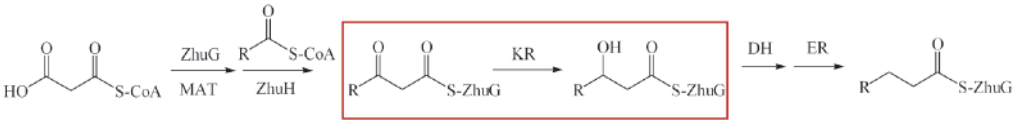

FIGURE 2: Biosynthesis of R1128. The reactions catalyzed by the initiation module are shown. ZhuG is charged with a malonyl group by MAT, and ZhuH catalyzes the condensation between malonyl-ZhuG and an acyl-CoA to form β-ketoacyl-ZhuG. The β-ketoacyl-ZhuG is then reduced to β-hydroxyacyl-ZhuG by a KR, which is further reduced to acyl-ZhuG by a DH and an ER. The elongation module is then primed with acyl-ZhuG. The reaction catalyzed by the KR is highlighted in the red box.

polyketide chain length, the building block used for chain initiation, regioselective cyclizations, and subsequent enzymatic modifications. Over the past decade, PKSs have been targets of extensive manipulations in producing novel polyketides. One example is the replacement of the primer unit of an aromatic polyketide with alternative functional groups (*13*−*16*). Although most of aromatic polyketides are primed with an acetate unit, derived via malonyl-ACP decarboxylation, there are a few exceptions such as doxorubicin, R1128, frenolicin, and enterocin (*17*−*21*).

Recent studies have revealed that a set of enzymes, termed the initiation module, are dedicated to non-acetate priming for some of these polyketides (Figure 2) (*22*−*24*). By coupling one of these initiation modules with different minimal PKSs, novel polyketides with alternative primer units have been synthesized (*24*). The initiation module of the R1128 PKS is comprised of a homodimeric KS (ZhuH), an ACP (ZhuG), an MAT (identical to the one used by the minimal PKS described above), and an editing acyl-ACP thioesterase (ZhuC). ZhuH dictates primer unit selectivity, whereas ZhuC suppresses acetate priming in favor of non-acetate priming by selectively hydrolyzing acetyl-ACP (*23*). In addition, the R1128 initiation module is also expected to recruit a ketoreductase (KR), a dehydratase (DH), and an enoylreductase (ER), although their identities remain a mystery, since their genes are absent from the R1128 gene cluster. A similar situation also exists in the frenolicin biosynthetic pathway.

Because the intact R1128 gene cluster, when introduced into *Streptomyces coelicolor*, can yield the expected congeners of this natural product, we attempted to identify the KR, DH, and ER from this heterologous host since its genome has been completely sequenced (*25*). Drawing upon our previous experience with the MAT (*26*, *27*), which was shown to be shared between the type II PKSs and the housekeeping fatty acid synthase of *S. coelicolor*, we hypothesized that the KR, DH, and ER responsible for R1128 biosynthesis in *S. coelicolor* may also be shared between the polyketide and fatty acid pathways. Rather than taking a reverse genetic approach, however, this time we elected to take a forward genomics approach to identifying these proteins. Here we report the results of our efforts to biochemically and structurally characterize SCO1815, a putative gene product from the *S. coelicolor* genome.

MATERIALS AND METHODS

Cloning, Expression, and Purification of SCO1815. The gene encoding SCO1815 was amplified from *S. coelicolor* genomic DNA by PCR using the primers 5′-CGAAAGGGTGATCATATGAGCCGCTCGGTTCT-3′ and 5′-CCGGTGATCAGGACAAGCTTGCCCTCGAGAAT-3′ (restriction sites underlined). The amplified insert was digested with *Nde*I and *Hind*III and cloned into the pET28a expression vector (Novagen), yielding a construct encoding SCO1815 with an N-terminal His_6 tag. *Escherichia coli* BL21(DE3) cells were transformed with the pYT3 resulting plasmid, and the resulting transformant was grown in LB medium at 37 °C to an OD_{600} of 0.6. The culture was then cooled to 18 °C, induced with 0.2 mM isopropyl β-D-galactopyranoside, and incubated for an additional 14 h at 18 °C. The cells were centrifuged (5000 rpm for 15 min), resuspended in lysis/wash buffer [50 mM phosphate (pH 7.6), 300 mM NaCl, and 10 mM imidazole], and lysed via sonication. The cell debris was removed by centrifugation, and the supernatant was applied to a nickel−NTA agarose (Qiagen) column. After the column had been extensively washed with 10 volumes of lysis/wash buffer, the protein was eluted with 3 volumes of elution buffer [50 mM phosphate (pH 7.6), 100 mM NaCl, and 150 mM imidazole]. Final purification was achieved on a Hi-TrapQ anion exchange column (Amersham-Pharmacia) with an increasing linear gradient of 0 to 1 M NaCl. SCO1815 eluted at approximately 190 mM NaCl. The purified protein was exchanged into a buffer containing 20 mM HEPES (pH 7.6) and concentrated to 10 mg/mL.

Mutant V129A was constructed via QuickChange mutagenesis (Stratagene) and purified as described above.

Substrate Preparation. Coenzyme A derivatives were purchased from Sigma and used without further purification. (3*S*)-Hydroxybutyryl-SNAc, (3*R*)-hydroxybutyryl-SNAc, 3-hydroxypentanoyl-SNAc, and (3*R*)-2-methyl-3-hydroxypentanoyl-SNAc were prepared by established procedures (*13*, *28*, *29*). Apo-ZhuG was expressed and purified as previously reported (*30*), and (3*R*)-hydroxybutyryl-ZhuG was prepared as described previously for malonyl-ZhuG (*30*).

Ketoreductase Assay. Our assay monitors oxidation of the β-hydroxyl group of appropriate thioester substrates by measuring an increase in NADPH or NADH fluorescence using a Hitachi F-4500 fluorescence spectrophotometer. The excitation wavelength was 340 nm, and the emission wavelength was 455 nm. Reactions were performed at room temperature in 100 mM phosphate buffer (pH 7.2). Reaction mixtures involving CoA derivatives or SNAc thioesters contained 1 mM cofactor ($NADP^+$ or NAD^+), 100 or 200 μM substrate, and 4 μM SCO1815 in a total reaction volume of 100 μL. To quantify the substrate specificity of SCO1815 for (3*R*)-hydroxybutyryl-ZhuG, steady-state kinetic measurements were obtained in the presence of variable (3*R*)-hydroxybutyryl-ZhuG concentrations (0−200 μM), 1 mM $NADP^+$, and 50 nM SCO1815.

Crystallization and Data Collection. The crystals of the apo form of SCO1815 were grown at room temperature via the hanging-drop vapor diffusion method, using 2 μL of protein (10 mg/mL) and 2 μL of mother liquor. Several initial crystallization conditions were identified using crystal screens 1 and 2 from Hampton Research. Crystals used for data collection were grown in 20% PEG MME 2000, 10 mM

Table 1: Data Collection and Structure Refinement Statistics for SCO1815

data collection statistics	
wavelength (Å)	1.03155
resolution (Å)	2.00
no. of reflections (observed/ unique)	551154/37985
completeness (%)	100 (100)
redundancy	14.4 (14.4)
$I/\sigma(I)$	62.0 (7.8)
R_{merge}	0.080 (0.497)
refinement statistics	
space group	$P3_121$
unit cell parameters	$a = 113.993$ Å, $b = 113.993$ Å, $c = 74.307$ Å, $\gamma = 120°$
no. of monomers per AU	2
Matthews coefficient	2.8
resolution (Å)	2.0
no. of protein atoms	3269
no. of water molecules	304
no. of missing residues	26
R_{free} (%)	23.2
R_{crys} (%)	19.8
rmsd for bonds (Å)	0.005
rmsd for angles (deg)	1.17
Ramachandran plot (%)	
most favored	94.3
favored	5.7
generously allowed	0

$NiCl_2$, and 100 mM Tris (pH 8.5). These crystals belonged to space group $P3_121$ and contained two monomers per symmetric unit. They were found to have a solvent content of 55%. The unit cell had the following dimensions: $a = 113.993$ Å, $b = 113.993$ Å, $c = 74.307$ Å, and $\gamma = 120°$. Prior to being cryo-cooled, crystals were soaked in a solution of 10% glycerol. Diffraction data sets were collected on beamline 11-1 at the Stanford Synchrotron Radiation Laboratory (SSRL) to 2.0 Å and processed using HKL2000 (*31*). A summary of the crystallographic data is presented in Table 1.

Several attempts were also made to grow diffracting crystals of SCO1815 bound to $NADP^+$, but these crystals failed to yield interpretable electron densities.

Structure Refinement and Model Building. Initial phases were obtained by molecular replacement using Molrep in CCP4i (*32*). The structure of the FabG protein from *E. coli* (Protein Data Bank entry 1Q7B) was used as the search model. The initial solution contained two monomers per asymmetric unit with a correlation coefficient of 26.7% and an *R* factor of 55.5%. The SCO1815 sequence was introduced into the model, and the structure was further refined using CNS (*33*). In addition, O was used for manual fitting (*34*). Noncrystallographic symmetry restraints were applied for initial rounds of refinement but were removed for the final stages of refinement. Water molecules were added using CNS, followed by visual inspections with O at the final stage. Ten percent of the reflections were excluded from the refinement and constituted the R_{free} set. The final geometry was assessed using PROCHECK (*35*).

Molecular Docking Simulations. Docking models of SCO1815 bound to $NADP^+$ and the thioester substrate were generated manually, followed by energy minimization using CNS (*33*). Coordinates for ZhuG were generated using the WHATIF homology server and were based upon the solution NMR structure of the closely related *act* ACP (*36*, *37*). Docking between ZhuG and the $NADP^+$·SCO1815 complex was performed on the Patchdock server (http://bioinfo3d.cs.tau.ac.il/PatchDock/index.html) and the Cluspro server (http://nrc.bu.edu/cluster/). In both cases, the $NADP^+$·SCO1815 complex was defined as the receptor and ZhuG was defined as the ligand. The best docking models returned by both servers were very similar, identifying the same recognition site for ZhuG on SCO1815.

RESULTS AND DISCUSSION

Cloning, Expression, and Purification of SCO1815 and ZhuG. The gene encoding SCO1815 was amplified from genomic DNA of *S. coelicolor* and expressed and purified as described in Materials and Methods. The typical yield of SCO1815 was 100 mg/L of culture volume. The cloning and purification of ZhuG have been described previously (*30*).

SCO1815 Can Function as the Ketoreductase Component of the Initiation Module of the R1128 PKS. The high degree of sequence similarity between SCO1815 and the FabG component of the *E. coli* fatty acid synthase suggested that SCO1815 could be a β-ketoacyl ACP reductase. We therefore wished to test whether SCO1815 could reduce β-ketoacyl substrates tethered to the ACP (ZhuG) from the R1128 initiation module. Initial attempts to assay SCO1815 employed 3-keto-butyryl-ZhuG as a substrate; however, this substrate was readily hydrolyzed under the reaction conditions, even in the absence of any enzyme. Therefore, the ketoreductase activity of SCO1815 was assayed through the reverse reaction using β-hydroxybutyryl-ZhuG as a substrate. The reaction was monitored by an increase in fluorescence as $NADP^+$ or NAD^+ was reduced. As shown in Figure 3A, SCO1815 was able to oxidize β-hydroxybutyryl-ZhuG. Analogous CoA thioester and SNAc thioester substrates were also examined. Under comparable reaction conditions, the rates of oxidation of (3*R*)-hydroxybutyryl-SNAc or (3*R*)-hydroxybutyryl-CoA were negligible compared to the rate at which (3*R*)-hydroxybutyryl-ZhuG was oxidized (Figure 3A), suggesting that the extended pantetheinyl arm and especially the protein portion of holo-ACP contribute significantly to its specificity. Steady-state kinetic experiments with (3*R*)-hydroxybutyryl-ZhuG yielded a k_{cat}/K_M of 0.29 μM^{-1} min^{-1} (Figure 3B). Although the enzyme could not be saturated when the reverse reaction was monitored, its bimolecular rate constant was fast enough to suggest that β-ketoreduction would not limit the overall rate of polyketide biosynthesis. [Under steady-state conditions, the typical turnover number of an aromatic PKS is less than 1 min^{-1} (*38*).]

Substrate Specificity and Stereospecificity of SCO1815. The cofactor preference of SCO1815 was investigated using both $NADP^+$ and NAD^+. SCO1815 exclusively utilized $NADP^+$ as its cofactor; no reaction was observed in the presence of NAD^+ (data not shown). To gain insight into the stereospecificity of SCO1815, enantiomers of β-hydroxybutyryl-SNAc were assayed in lieu of β-hydroxybutyryl-ACP substrates. Both (3*S*)- and (3*R*)-hydroxybutyryl-SNAc were tested as substrates. The fluorescence of the reaction mixture increased in the presence of (3*R*)-hydroxybutyryl-SNAc, while it remained unchanged when (3*S*)-hydroxybutyryl-SNAc was present (data not shown). Thus, SCO1815 is (*R*)-stereospecific. To examine whether SCO1815 accepts α-sub-

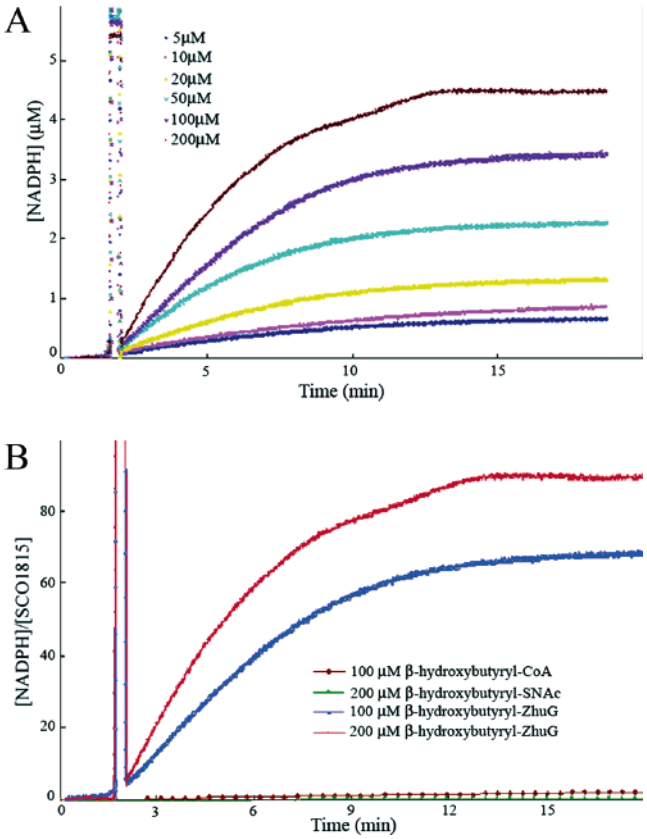

FIGURE 3: Enzymatic activity of SCO1815. (A) Steady-state kinetic assay of (3*R*)-hydroxybutyryl-ZhuG. Reactions were performed in the presence of 5, 10, 20, 50, 100, and 200 μM (3*R*)-hydroxybutyryl-ZhuG. (B) Assay with 100 μM (3*R*)-hydroxybutyryl-CoA, 200 μM (3*R*)-hydroxybutyryl-SNAc, and 100 or 200 μM (3*R*)-hydroxybutyryl-ZhuG.

stituted substrates, both diastereomers of (3*R*)-2-methyl-3-hydroxypentanoyl-SNAc were synthesized and assayed. No activity was detectable for either diastereomer (data not shown), suggesting that SCO1815 discriminates against α-branched substrates. Last, 3-hydroxypentanoyl-SNAc, a longer chain substrate, was also recognized and oxidized by SCO1815. This is consistent with the observation that the R1128 PKS can be primed with a variety of non-acetate groups (Figure 1) (*18*).

Overall Structure of Apo SCO1815. The apo form of SCO1815 was crystallized in space group $P3_121$ with two monomers per asymmetric unit. Phasing was carried out via molecular replacement using the 48% identical *E. coli* FabG as the search model. All of the polypeptide chains could be traced through the electron density map except residues 84−86 and 176−184 and C-terminal residue His234. A Ramachandran plot of the structure showed that all residues lie within the favored regions (94.3% in the most favored regions and 5.7% in favored regions).

Like other members of the short chain dehydrogenase/reductase (SDR) family, each SCO1815 monomer is comprised of two subdomains. The large subdomain adopts a Rossmann fold (Figure 4A), which consists of a twisted, parallel seven-stranded β-sheet (βA, βB, βC, βD, βE, βF, and βG) in the center, flanked by two short helices (αB and αG) on one side and three long helices (αD, αE, and αF) on the other. The small subdomain (αFG1 and αFG2) is disordered, as is the case in many SDR enzyme structures

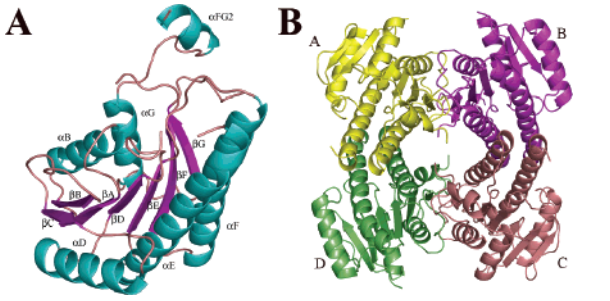

FIGURE 4: Crystal structure of apo SCO1815. (A) Structure of the monomer. α-Helices are colored cyan, β-strands magenta, and loops brown. (B) Structure of the tetramer. Each monomer is colored with a different color.

(*39*, *40*). αFG1 is invisible in both monomers, whereas αFG2 is only partially visible; the temperature factors are more than 20 Å^2 higher than that of the remaining protein.

A tetrameric structure is formed along the 2-fold crystallographic axis between the A−B dimer and the C−D dimer (Figure 4B). The tetramer has an approximate diameter of 85 Å. It has 222 symmetry, and there are two different dimer interfaces. The first interface, between monomer A and monomer B, is comprised of αG−αG′ and βG−βG′ interactions. A total of approximately 1400 Å^2 of solvent-accessible surface area is buried in each monomer. At the αG−αG′ interface, the interactions are mainly hydrophobic in nature with the two aromatic rings of Phe210 from each monomer stacked upon each other. In the loop region, the side chain oxygen of Tyr218 is hydrogen bonded to the backbone carbonyl oxygen of Arg197 from the other monomer, and the backbone carbonyl oxygen of Tyr218 forms a hydrogen bond with the backbone nitrogen of Gly228 from the other monomer. At the βG−βG′ interface, the interactions are also mostly hydrophobic. At the second interface between monomer A and monomer D, a four-helix bundle is formed by αE, αF, αE′, and αF′, and approximately 1000 Å^2 of solvent-accessible surface area is buried in each monomer. At this interface, αE primarily interacts with αE′, whereas αF interacts with αF′. Both hydrophobic and hydrophilic interactions are observed at the αE−αE′ interface. The side chain oxygen of Thr101 forms a hydrogen bond to the main chain carbonyl oxygen of Gln90 from the other monomer; additionally, a salt bridge is observed between Arg105 and Asp91. The interactions between αF and αF′ are largely van der Waals in nature, as these two helices have a stretch of alanine residues that sterically complement each other. In addition, the oxygen atom of the Ser134 side chain forms hydrogen bonds with the Ser153 and Glu157 side chains.

Comparison of SCO1815 to Other Ketoreductases. The residues that are known to be important for cofactor binding, catalysis, and the structural integrity of SDR enzymes are conserved in SCO1815. For example, the N-terminal cofactor binding motif (Thr8-Gly9-X_3-Gly13-X-Gly15), the catalytic tetrad (Asn99-Ser127-Tyr140-Lys144), and the C-terminal Pro170-Gly171 motif are conserved in SCO1815. The central Asn-Asn-Ala-Gly sequence, believed to play a role in cofactor binding, is present in a slightly modified form as the Ala74-Asn75-Ala76-Gly77 sequence (*41*) The predicted property of these residues agrees well with our $NADP^+$• SCO1815 model, in which Asn75 hydrogen bonds to the cofactor (vide infra) (Figure 5).

FIGURE 5: Multiple-sequence alignment of SCO1815 with other KRs. The cofactor binding sites are denoted with blue circles; the catalytic tetrad residues are denoted with red stars, and the ACP docking site is denoted with orange squares. Residue V129, which may play a role in controlling substrate specificity of SCO1815, is highlighted.

A PSI-BLAST search of the Protein Data Bank revealed that the three most homologous structures are ketoreductases from fatty acid synthases: MabA from *Mycobacterium tuberculosis* and the FabG proteins from *E. coli* and *Brassica napus*, with levels of sequence identity of 57, 48, and 42%, respectively. The C-9 ketoreductase from the actinorhodin PKS (PDB entries 1XR3, 1X7G, 1X7H, and 1W4Z) from *S. coelicolor* is more distantly related to SCO1815, with a

37% identical sequence. The structure of apo SCO1815 is mostly identical to the structures of these apo KRs, but subtle differences can be observed. For example, helix αE is straight in SCO1815 but kinked in other KRs at the position equivalent to Asn99 (a residue in the catalytic tetrad). This suggests that helix αE of SCO1815 might undergo a conformational change during catalysis. Another prominent difference lies between strands βB and βC. This region adopts a loop conformation in SCO1815 due to a 12-residue truncation. In contrast, it is replaced with a helix (αC) in most other SDR enzymes. This helix to loop switch has also been observed in the structure of apo-MabA (*42*). Superposition of the SCO1815 structure onto the structures of apo-MabA (PDB entry 1UZL) and *E. coli* apo-FabG (PDB entry 1OI1) shows rmsd values of 1.06 and 1.05 Å for 186 C_α and 177 C_α atoms, respectively. Regions which lack good electron density in the apo SCO1815 structure are missing in the apo-MabA and apo-FabG structures as well. When apo SCO1815 is compared with the structures of $NADP^+$·KR binary complexes, more prominent differences are observed. The disordered regions in the apo structures are well-defined in the cofactor-bound structures. For example, superposition of the apo SCO1815 structure onto the structures of $NADP^+$·FabG complexes from *E. coli* (PDB entry 1Q7B) and *B. napus* (PDB entry 1EDO) shows rmsd values of 1.25 and 1.25 Å for 190 C_α and 186 C_α atoms, respectively. The higher rmsd arises from conformational changes induced upon cofactor binding. Among KR enzymes from fatty acid synthases, the structure of FabG from *E. coli* has been determined in both the apo form (*39*) and the $NADP^+$-bound form (*43*). Through these two structures, it has been established that the loop between αF and βE undergoes a conformational change upon cofactor binding. It has also been suggested that the cofactor binds to the enzyme through a two-stage process (*43*). The adenine moiety of the $NADP^+$ first binds to the protein, which then pushes the loop back to reveal the nicotinamide binding cavity. As a consequence of these conformational changes, the active site residues are brought into catalytically active positions. By careful comparison of the structure of apo SCO1815 with the structures of apo and $NADP^+$-bound KRs, it appears that a similar conformational change is also likely to occur when $NADP^+$ binds to SCO1815. As there is not enough space for the nicotinamide ring, when $NADP^+$ binds to SCO1815 the C_α atom of the catalytic residue Ser127 and the C_α atom of Val128 would be expected to move 4.9 and 8 Å, respectively (Figure 6).

Modeling of Binding of the Cofactor to SCO1815. SCO1815 utilizes NADP(H) as its cofactor. However, our attempts to grow crystals of NADP(H)-bound SCO1815 failed to yield crystals with interpretable electron densities. To gain further insight, we modeled the binding of NADP(H) to SCO1815 using the *E. coli* FabG·$NADP^+$ coordinates. As shown in Figure 7, NADP(H) is predicted to bind to the SCO1815 protein in an extended conformation. The nicotinamide ring, which is in the trans conformation, inserts into the cavity created by βE and βF. Its amide oxygen forms a hydrogen bond to the backbone amide nitrogen of Gly171. The nicotinamide ribose is predicted to be involved in four hydrogen bonds. The 2′-hydroxyl presumably forms a hydrogen bond to the side chain oxygen of Tyr140, and the 3′-hydroxyl is predicted to bond to the backbone oxygen

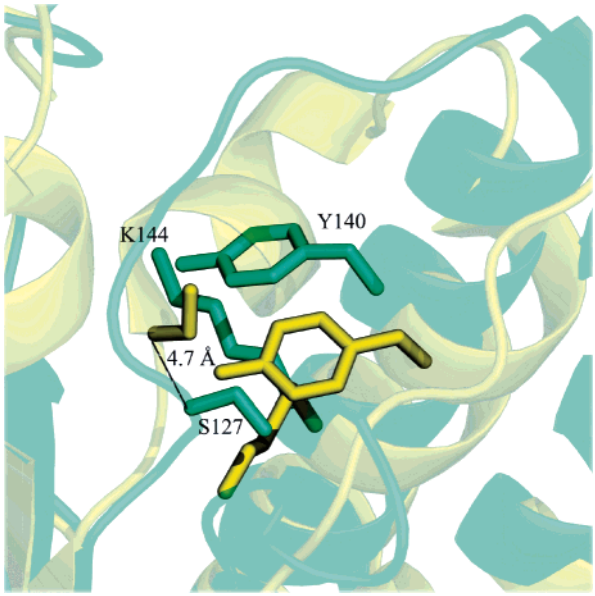

FIGURE 6: Superposition of the active sites of apo SCO1815 and the $NADP^+$·FabG complex. Apo SCO1815 is colored cyan, and the $NADP^+$·FabG complex is colored yellow. S127, Y140, and K144 in the catalytic tetrad are shown in stick representation. The putative conformational change induced by cofactor binding leads to the reorganization of catalytic residues, which results in S127 moving approximately 4.7 Å.

(OD1) of Asn75, the backbone oxygen of Ile125, and the side chain nitrogen of Lys144. The pyrophosphate group is likely anchored on the SCO1815 protein through a hydrogen bond and a salt bridge interaction with Arg12, whereas the adenine ring and its associated ribose are predicted to reside in a pocket formed by four N-terminal loops and αE. A hydrogen bond is likely present between the primary amine nitrogen at position N6 and the side chain of Asp48. The 3′-hydroxyl of adenine ribose is presumably secured through hydrogen bonding to side chain oxygen OD1 of Asn11 and the backbone oxygen of Gly9, while the 2′-phosphate is held in place by forming hydrogen bonds with ND2 of Asn11 and the Ser35 oxygen. Not surprisingly, nearly all the residues thought to be involved in cofactor binding are highly conserved among KR enzymes.

A structural comparison among SDR enzymes (*44*) shows that $NADP^+$ specific SDRs have at least one positively charged residue at position 12 or 35 (SCO1815 numbering), interacting directly with the phosphate group. In contrast, NAD^+ specific SDRs contain a negatively charged Asp residue at position 34. A double mutant D34S/M35D of XDH reverses its cofactor specificity from NAD^+ to $NADP^+$ (*44*). By analogy, we propose that Arg12 in SCO1815 indicates that it is $NADP^+$ specific. The side chain of this residue is predicted to undergo a conformational change, allowing a salt bridge to form with the adenine ribose phosphate group.

Active Site. In the active site of KR enzymes, a tyrosine, a lysine, a serine, and an asparagine form the catalytic tetrad. The tyrosine and lysine contribute to cofactor binding by forming hydrogen bonds with the cofactor, whereas the serine orients the substrate by forming a hydrogen bond with the oxygen of the β-ketone group. In the apo SCO1815 structure, Tyr140 and Lys144 in the catalytic tetrad are located on helix αF, pointing to a nonproductive orientation. Upon binding of NADPH (but not NADH), the catalytic tetrad is predicted to reorient into a reactive conformation.

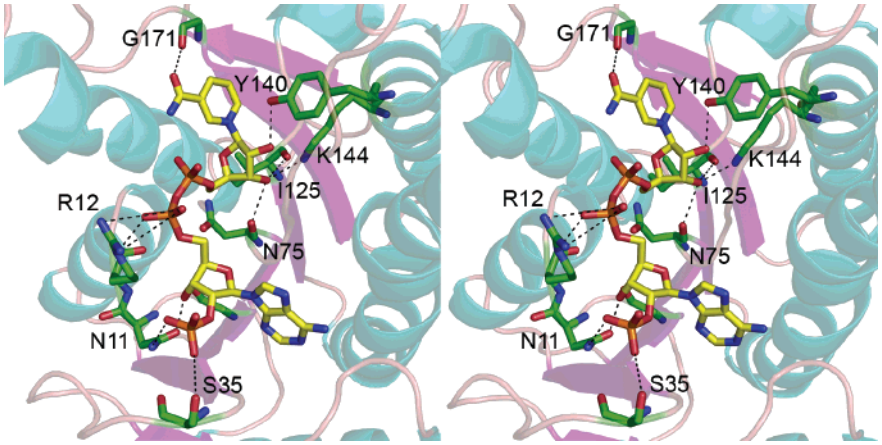

FIGURE 7: Model for binding of $NADP^+$ to SCO1815. $NADP^+$ C−C bonds are colored yellow, SCO1815 C−C bonds colored green, and hydrogen bonds depicted as dashed lines.

Possible Substrate Binding and ZhuG Docking Sites. The crystal structures of several SDR enzymes, such as the 7α-hydroxysteroid dehydrogenase from *E. coli* (PDB entry 1AHI) and tropinone reductase II (PDB entry 2AE2), have been determined as cofactor- and substrate-bound ternary complexes. Although neither shares a high degree of sequence identity with SCO1815, the three-dimensional structures are strikingly similar to the structure of SCO1815 reported here. Guided by these two ternary structures, we readily modeled a 3-ketobutyryl-CoA substrate in the active site. The β-ketone group is within hydrogen bonding range of Ser127 and Tyr140 and is also close to the C-4 atom of the nicotinamide ring from which a hydride is transferred. The relative orientation of the remainder of the substrate determines the catalytic stereospecificity of SCO1815. As described above, SCO1815 is (*R*)-stereospecific. Thus, the substrate can only enter into the active site from the same side as the cofactor-binding pocket (Figure 8). A similar substrate binding mode was also proposed for *B. napus* FabG (*45*) and *M. tuberculosis* MabA (*42*).

The stereospecificity of KR domains in type I PKSs is determined by specific motifs that can be recognized from the primary sequences of these enzymes. KR domains with LDD and PXXXN motifs yield products with (*R*)-stereochemistry, whereas those which lack these signature motifs have opposite stereospecificity (*46*). Mutagenesis of the signature motifs has successfully switched the stereospecificity of the KR from module 1 of the 6-deoxyerythronolide B synthase (*47*, *48*). However, none of these motifs are conserved in SCO1815.

It has been recognized that charge−charge interactions are essential for ACPs to interact with other domains in type II PKSs and fatty acid synthases (*49*). Inspection of the surface of SCO1815 reveals a positively charged patch formed by three residues, Arg12, Arg34, and Lys80 (Figure 8). Interestingly, these residues are located at the substrate entrance in our docking model discussed above and are also conserved among type II KR proteins. This is consistent with the fact that the products of the reactions catalyzed by type II KRs have (*R*)-stereochemistry. Different KR−ACP binding models were proposed for the C-9 ketoreductase of the actinorhodin PKS (*50*, *51*). In one model, residues Val152, Tyr202, Trp206, Glu212, and Arg220 are important in ACP docking (*50*); in the other model, three arginine residues, Arg38, Arg65, and Arg93, form a positively charged pocket that interacts with a negatively charged Glu47 from the ACP (*51*). Neither set of residues is conserved in SCO1815, suggesting that SCO1815 interacts with ZhuG in a different manner. It has been shown that two conserved Arg residues (Arg129 and Arg172) in *E. coli* FabG are essential for the KR−ACP interaction (*52*). Mutation of these Arg residues abrogates fatty acid biosynthesis. These two residues are also conserved in SCO1815, as Lys118 and Arg161, respectively. However, Lys118 is located in the loop region between helix αE and sheet βE, while Arg161 is located in the loop region between helix αF and sheet βF. These residues are ca. 30 Å from the catalytic tetrad. Since the substrate is presented to the active site through the 18 Å long phosphopantetheine arm of ZhuG, it is unlikely that the substrate could reach the active site if ZhuG docks at these residues.

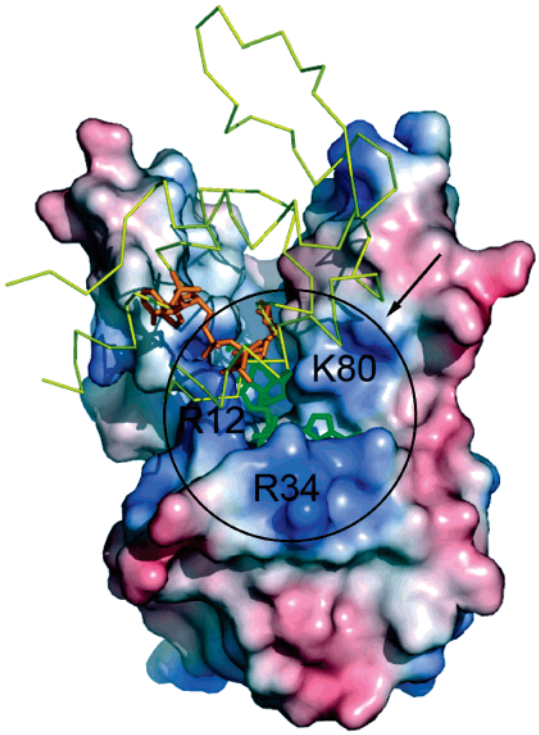

FIGURE 8: Model for docking of ZhuG on SCO1815. SCO1815 is shown with a surface model. ZhuG is shown as yellow lines; 3-ketobutyryl-CoA substrate and $NADP^+$ are shown as orange and green sticks, respectively. The positive patch formed by Arg12, Arg34, and Lys80 is circled and denoted with an arrow. The surface electrostatic potential was generated by PyMOL.

We have shown that 2-methyl-substituted acyl-SNAc substrates are not accepted by SCO1815. Therefore, the substrate docking model was used to pinpoint the structural basis for the substrate specificity of SCO1815. Our model suggests that a (2*R*)-methyl group would clash with the nicotinamide ring, whereas a (2*S*)-methyl group would clash with the side chain of Val129. In contrast to SCO1815, the substrates of many ketoreductase domains from type I PKSs are α-substituted acyl-ACPs. Multiple-sequence alignment indicates that the residue in a type I PKS at the position corresponding to Val129 is a highly conserved alanine residue (Figure 5). Thus, Val129 was hypothesized to play a role in controlling the substrate specificity of SCO1815. The V129A mutant, however, is similar to wild-type SCO1815. Although it retains wild-type oxidative activity against (3*R*)-hydroxybutyryl-SNAc, it failed to reveal activity against either diastereomer of (3*R*)-2-methyl-3-hydroxypentanoyl-SNAc (data not shown). This suggests that the selectivity filter which prevents binding of 2-methyl-substituted acyl-SNAc substrates is comprised of additional residues within the SCO1815 protein.

SUMMARY

Polyketides are among the most important secondary metabolites. Precursor-directed biosynthesis has been shown to be a powerful tool for rational syntheses of structural analogues of type I polyketides, such as erythromycin, avermectin, and rapamycin (*13*−*16*). The discovery of bimodular type II PKSs prompted a similar approach with aromatic PKS systems. By coexpressing the R1128 initiation module with the actinorhodin and tetracenomycin minimal PKSs, polyketides primed with pentanoyl, hexanoyl, and isohexanoyl groups were efficiently synthesized by *S. coelicolor* (*24*). Successful exploitation of this approach requires a detailed understanding of the initiation module. In this work, SCO1815, a FabG homologue from the genome of *S. coelicolor*, was shown to function as the β-ketoacyl ACP reductase component of the R1128 initiation module. Its crystal structure suggests that NADP(H) and ACP interact with the enzyme through a positive patch formed by Arg12, Arg34, and Lys80. These studies provide a foundation for future efforts aimed at exploiting SCO1815 for biosynthetic engineering.

ACKNOWLEDGMENT

We thank Sheryl Tsai, Alice Chen, Rongsheng Jin, and Pavel Strop for helpful discussions, Dan Pinkas and Christian Ridley for proofreading, and David Cane for providing the (3*R*,2*R*)-2-methyl-3-hydroxypentanoyl-SNAc substrate.

REFERENCES

1. O'Hagan, D. (1991) *The polyketide metabolites*, Ellis Horwood, Chichester, U.K.
2. Hopwood, D. A. (1997) Genetic Contributions to Understanding Polyketide Synthases, *Chem. Rev. 97*, 2465−98.
3. Weissman, K. J., and Leadlay, P. F. (2005) Combinatorial biosynthesis of reduced polyketides, *Nat. Rev. Microbiol. 3*, 925−36.
4. McDaniel, R., Ebert-Khosla, S., Hopwood, D. A., and Khosla, C. (1993) Engineered biosynthesis of novel polyketides: Manipulation and analysis of an aromatic polyketide synthase with unproven catalytic specificities, *J. Am. Chem. Soc. 115*, 11671−5.
5. McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A., and Khosla, C. (1994) Engineered biosynthesis of novel polyketides: Influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase, *Proc. Natl. Acad. Sci. U.S.A. 91*, 11542−6.
6. Fu, H., Hopwood, D. A., and Khosla, C. (1994) Engineered biosynthesis of novel polyketides: Evidence for temporal, but not regiospecific, control of cyclization of an aromatic polyketide precursor, *Chem. Biol. 1*, 205−10.
7. Fu, H., McDaniel, R., Hopwood, D. A., and Khosla, C. (1994) Engineered biosynthesis of novel polyketides: Stereochemical course of two reactions catalyzed by a polyketide synthase, *Biochemistry 33*, 9321−6.
8. Fu, H., Alvarez, M. A., Khosla, C., and Bailey, J. E. (1996) Engineered biosynthesis of novel polyketides: Regiospecific methylation of an unnatural substrate by the tcmO O-methyltransferase, *Biochemistry 35*, 6527−32.
9. Li, Q., Khosla, C., Puglisi, J. D., and Liu, C. W. (2003) Solution structure and backbone dynamics of the holo form of the frenolicin acyl carrier protein, *Biochemistry 42*, 4648−57.
10. Yu, T. W., and Hopwood, D. A. (1995) Ectopic expression of the *Streptomyces coelicolor* whiE genes for polyketide spore pigment synthesis and their interaction with the act genes for actinorhodin biosynthesis, *Microbiology 141* (Part 11), 2779−91.
11. Shen, B., and Hutchinson, C. R. (1993) Enzymatic synthesis of a bacterial polyketide from acetyl and malonyl coenzyme A, *Science 262*, 1535−40.
12. Dreier, J., Shah, A. N., and Khosla, C. (1999) Kinetic analysis of the actinorhodin aromatic polyketide synthase, *J. Biol. Chem. 274*, 25108−12.
13. Jacobsen, J. R., Hutchinson, C. R., Cane, D. E., and Khosla, C. (1997) Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase, *Science 277*, 367−9.
14. Marsden, A. F., Wilkinson, B., Cortes, J., Dunster, N. J., Staunton, J., and Leadlay, P. F. (1998) Engineering broader specificity into an antibiotic-producing polyketide synthase, *Science 279*, 199−202.
15. Long, P. F., Wilkinson, C. J., Bisang, C. P., Cortes, J., Dunster, N., Oliynyk, M., McCormick, E., McArthur, H., Mendez, C., Salas, J. A., Staunton, J., and Leadlay, P. F. (2002) Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase, *Mol. Microbiol. 43*, 1215−25.
16. Lowden, P. A. W. B., Bohm, G. A., Handa, S., Floss, H. G., et al. (2001) Origin and true nature of the starter unit for the rapamycin polyketide synthase, *Angew. Chem., Int. Ed. 40*, 777−9.
17. Moore, B. S., and Hertweck, C. (2002) Biosynthesis and attachment of novel bacterial polyketide synthase starter units, *Nat. Prod. Rep. 19*, 70−99.
18. Marti, T., Hu, Z., Pohl, N. L., Shah, A. N., and Khosla, C. (2000) Cloning, nucleotide sequence, and heterologous expression of the biosynthetic gene cluster for R1128, a non-steroidal estrogen receptor antagonist. Insights into an unusual priming mechanism, *J. Biol. Chem. 275*, 33443−8.
19. Bibb, M. J., Sherman, D. H., Omura, S., and Hopwood, D. A. (1994) Cloning, sequencing and deduced functions of a cluster of *Streptomyces* genes probably encoding biosynthesis of the polyketide antibiotic frenolicin, *Gene 142*, 31−9.
20. Piel, J., Hertweck, C., Shipley, P. R., Hunt, D. M., Newman, M. S., and Moore, B. S. (2000) Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate '*Streptomyces maritimus*': Evidence for the derailment of an aromatic polyketide synthase, *Chem. Biol. 7*, 943−55.
21. Grimm, A., Madduri, K., Ali, A., and Hutchinson, C. R. (1994) Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase, *Gene 151*, 1−10.
22. Tang, Y., Lee, T. S., Kobayashi, S., and Khosla, C. (2003) Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity, *Biochemistry 42*, 6588−95.
23. Tang, Y., Koppisch, A. T., and Khosla, C. (2004) The acyltransferase homologue from the initiation module of the R1128 polyketide synthase is an acyl-ACP thioesterase that edits acetyl primer units, *Biochemistry 43*, 9546−55.
24. Tang, Y., Lee, T. S., and Khosla, C. (2004) Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases, *PLoS Biol. 2*, E31.
25. Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G.,

Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), *Nature 417*, 141−7.

26. Revill, W. P., Bibb, M. J., and Hopwood, D. A. (1995) Purification of a malonyltransferase from *Streptomyces coelicolor* A3(2) and analysis of its genetic determinant, *J. Bacteriol. 177*, 3946−52.
27. Carreras, C. W., and Khosla, C. (1998) Purification and in vitro reconstitution of the essential protein components of an aromatic polyketide synthase, *Biochemistry 37*, 2084−8.
28. Cane, D. E., Kudo, F., Kinoshita, K., and Khosla, C. (2002) Precursor-directed biosynthesis: Biochemical basis of the remarkable selectivity of the erythromycin polyketide synthase toward unsaturated triketides, *Chem. Biol. 9*, 131−42.
29. Harris, R. C., Cutter, A. L., Wessman, K. J., Hanefelf, U., Timoney, M. C., and Staunton, J. (1998) Enantiospecific Synthesis and Analogs of the Diketide Intermediate of the Erythromycin Polyketide Synthase (PKS), *J. Chem. Res. 283*, 1230−47.
30. Meadows, E. S., and Khosla, C. (2001) In vitro reconstitution and analysis of the chain initiating enzymes of the R1128 polyketide synthase, *Biochemistry 40*, 14855−61.
31. Otwinowski, Z. M. W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−26.
32. Collaborative Computational Project, Number 4 (1994) Collaborative Computational Project, Number 4, *Acta Crystallogr. D50*, 760−3.
33. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54* (Part 5), 905−21.
34. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47* (Part 2), 110−9.
35. Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−91.
36. Crump, M. P., Crosby, J., Dempsey, C. E., Parkinson, J. A., Murray, M., Hopwood, D. A., and Simpson, T. J. (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from *Streptomyces coelicolor* A3(2), *Biochemistry 36*, 6000−8.
37. Rodriguez, R., Chinea, G., Lopez, N., Pons, T., and Vriend, G. (1998) Homology modeling, model and software evaluation: Three related resources, *Bioinformatics 14*, 523−8.
38. Zawada, R. J., and Khosla, C. (1999) Heterologous expression, purification, reconstitution and kinetic analysis of an extended type II polyketide synthase, *Chem. Biol. 6*, 607−15.
39. Price, A. C., Zhang, Y. M., Rock, C. O., and White, S. W. (2001) Structure of β-ketoacyl-[acyl carrier protein] reductase from *Escherichia coli*: Negative cooperativity and its structural basis, *Biochemistry 40*, 12772−81.
40. Keatinge-Clay, A. T., and Stroud, R. M. (2006) The structure of a ketoreductase determines the organization of the β-carbon processing enzymes of modular polyketide synthases, *Structure 14*, 737−48.
41. Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B., and Jornvall, H. (2003) Short-chain dehydrogenases/reductases (SDR): The 2002 update, *Chem.-Biol. Interact. 143*−*144*, 247−53.
42. Cohen-Gonsaud, M., Ducasse, S., Hoh, F., Zerbib, D., Labesse, G., and Quemard, A. (2002) Crystal structure of MabA from *Mycobacterium tuberculosis*, a reductase involved in long-chain fatty acid biosynthesis, *J. Mol. Biol. 320*, 249−61.
43. Price, A. C., Zhang, Y. M., Rock, C. O., and White, S. W. (2004) Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG, *Structure 12*, 417−28.
44. Ehrensberger, A. H., Elling, R. A., and Wilson, D. K. (2006) Structure-guided engineering of xylitol dehydrogenase cosubstrate specificity, *Structure 14*, 567−75.
45. Fisher, M., Kroon, J. T., Martindale, W., Stuitje, A. R., Slabas, A. R., and Rafferty, J. B. (2000) The X-ray structure of *Brassica napus* β-keto acyl carrier protein reductase and its implications for substrate binding and catalysis, *Structure 8*, 339−47.
46. Caffrey, P. (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases, *ChemBioChem 4*, 654−7.
47. Baerga-Ortiz, A., Popovic, B., Siskos, A. P., O'Hare, H. M., Spiteller, D., Williams, M. G., Campillo, N., Spencer, J. B., and Leadlay, P. F. (2006) Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase, *Chem. Biol. 13*, 277−85.
48. O'Hare, H. M., Baerga-Ortiz, A., Popovic, B., Spencer, J. B., and Leadlay, P. F. (2006) High-throughput mutagenesis to evalutate models of setereochemical control in ketoreductase domains from the erythromycin polyketide synthase, *Chem. Biol. 13*, 287−96.
49. Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) Identification and analysis of the acyl carrier protein (ACP) docking site on β-ketoacyl-ACP synthase III, *J. Biol. Chem. 276*, 8231−8.
50. Korman, T. P., Hill, J. A., Vu, T. N., and Tsai, S. C. (2004) Structural analysis of actinorhodin polyketide ketoreductase: Cofactor binding and substrate specificity, *Biochemistry 43*, 14529−38.
51. Hadfield, A. T., Limpkin, C., Teartasin, W., Simpson, T. J., Crosby, J., and Crump, M. P. (2004) The crystal structure of the actIII actinorhodin polyketide reductase: Proposed mechanism for ACP and polyketide binding, *Structure 12*, 1865−75.
52. Zhang, Y. M., Wu, B., Zheng, J., and Rock, C. O. (2003) Key residues responsible for acyl carrier protein and β-ketoacyl-acyl carrier protein reductase (FabG) interaction, *J. Biol. Chem. 278*, 52935−43.

BI061187V