Chemistry & Biology
Article

A Tylosin Ketoreductase Reveals How Chirality Is Determined in Polyketides

Adrian T. Keatinge-Clay[1,*]
[1]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA
*Correspondence: adriankc@msg.ucsf.edu
DOI 10.1016/j.chembiol.2007.07.009

SUMMARY

Because it controls the majority of polyketide stereocenters, the ketoreductase (KR) is a central target in engineering polyketide synthases (PKSs). To elucidate the mechanisms of stereocontrol, the structure of KR from the first module of the tylosin PKS was determined. A comparison with a recently solved erythromycin KR that operates on the same substrate explains why their products have opposite α-substituent chiralities. The structure reveals how polyketides are guided into the active site by key residues in different KR types. There are four types of reductase-competent KRs, each capable of fixing a unique combination of α-substituent and β-hydroxyl group chiralities, as well as two types of reductase-incompetent KRs that control α-substituent chirality alone. A protocol to assign how a module will enforce substituent chirality based on its sequence is presented.

INTRODUCTION

Polyketides are an important class of pharmaceuticals that are primarily produced by streptomycetes, a type of soil bacteria [1–5]. The antibiotic tylosin is synthesized by *Streptomyces fradiae* using a megasynthase that can be described as a molecular assembly line [6]. The seven modules of the tylosin polyketide synthase (PKS) condense one ethylmalonyl, four methylmalonyl, and two malonyl extender units onto a priming propionyl unit to yield the 23-carbon macrocycle, tylactone, which is subsequently modified through oxidations and glycosylations into tylosin (Figure 1A).

Within a module, an extender unit is selected by an acyltransferase (AT), shuttled by an acyl carrier protein (ACP) via an ~18 Å long phosphopantetheinyl arm, and condensed to the growing polyketide chain by a ketosynthase (KS) [4, 7]. Three β-carbon processing enzymes can act on the β-keto group formed after each condensation: a ketoreductase (KR) uses NADPH to stereospecifically reduce it to a hydroxyl group, a dehydratase (DH) removes the hydroxyl group to create a double bond, and an enoylreductase (ER) uses NADPH to stereospecifically reduce the double bond. A methylene functionality is produced by a module containing each of these enzyme activities in a reaction sequence that parallels that of the related animalian fatty acid synthase (FAS) [8, 9]. To create a more functionalized ketide unit, enzymes acting on the β-carbon are either inactive or absent from the module that controls its addition to the growing polyketide chain. The combinations of the substituents that branch from the acyl chain and their chiralities make polyketides one of the most diverse classes of molecules.

Since the discovery of PKSs, it was apparent that AT controls the identity of the majority of the polyketide substituents through its selection of extender units larger than a malonyl group [10]; however, the mechanisms that set the chiralities of those substituents, as well as any hydroxyl substituents, are only now being deciphered. Sequence motifs, or fingerprints, have been detected within KRs that are predictive of hydroxyl group stereochemistry: A-type KRs employ a conserved tryptophan to produce a hydroxyl group of "S" stereochemistry; B-type KRs employ an LDD motif to produce a hydroxyl group of "R" stereochemistry (quotation marks are placed around R and S, as the convention used to label chiral centers in polyketides can deviate from the RS system: when discussing chirality at the β position, the γ position is given the lowest priority after the hydrogen; when discussing chirality at the α position, the α-substituent is given the lowest priority after the hydrogen) [11, 12]. The mechanisms used by KRs to set β-hydroxyl group stereochemistry are currently being elucidated [13]; however, even the identities of the enzymes controlling α-substituent stereochemistry have not been established [14, 15].

The chirality of the α-substituent (usually a methyl, ethyl, or methoxy group) has been studied at various stages of polyketide synthesis (Figure 1B): ATs only select extender units with an α-substituent of "S" stereochemistry, such as (2*S*)-methylmalonyl-CoA [16]. During the condensation reaction, KSs reverse the α-substituent chirality to yield a polyketide chain with an α-substituent of "R" stereochemistry [14, 16]. In studies of DEBS1-TE, which contains the first and second modules of the erythromycin PKS, a triketide lactone product revealed that a deuterium attached to the α-carbon of (2*S*)-methylmalonyl-CoA was not incorporated by the first module, but was incorporated by the second module. This result indicated that another reversal of the α-substituent chirality occurs after the

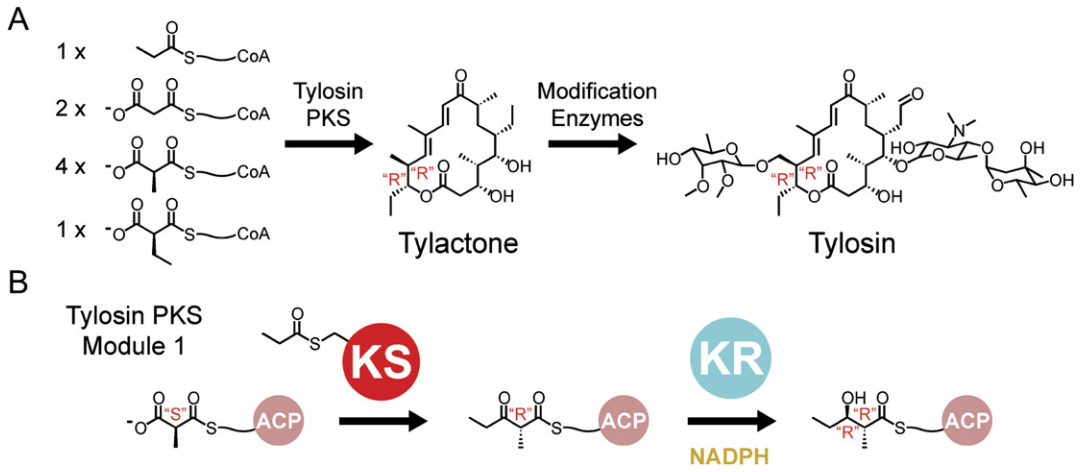

Figure 1. Tylosin Stereocenters

(A) Primer and extender units are selected, fused, and reduced in a stereocontrolled manner by the tylosin PKS to yield tylactone, the precursor to the antibiotic tylosin. TylKR1 sets the stereochemistry at the labeled stereocenters.

(B) Polyketide intermediates during catalysis by the first module of the tylosin PKS. The product of TylKS1 is the "R" diketide. TylKR1 acts on this diketide epimer, thereby setting the α-substituent chirality of the product. The stereospecific reduction also sets the chirality of the resulting β-hydroxyl group.

condensation reaction in the first module to yield an α-substituent of "S" stereochemistry.

Four theories have been presented to explain how α-substituent stereochemistry is set. (1) Some KSs are capable of catalyzing epimerization after condensation to yield a mixture of epimers, and a downstream enzyme selects the epimerized polyketide [17]. (2) Some modules allow spontaneous epimerization to occur in water, and a downstream enzyme selects the epimerized polyketide [17]. (3) KR catalyzes a cryptic epimerization, in which the polyketide exists in the enol form during hydride attack, and α-substituent chirality is determined by the side of the enol that the α-carbon acquires a proton [15]. (4) Some KRs have evolved the ability to catalyze epimerization before catalyzing the reduction reaction [18].

Recently, KRs isolated from their PKSs have aided in distinguishing between these possibilities. EryKR1 and TylKR1 were shown to maintain stereocontrol of the α-substituent and the β-hydroxyl group during in vitro reactions with polyketide substrates (the nomenclature describes a domain by its PKS and module of origin) (Figure 2A) [19]. Both enzymes can reduce a racemic mixture of diketide substrate analogs into the anticipated products (analogous to their diketide products in their native PKSs), which differ only in the stereochemistry of the α-methyl group. A racemic substrate mixture may not seem representative of the polyketides naturally presented to KR, as KS may only produce the unepimerized polyketide; however, KRs need to contend with the spontaneous epimerization that occurs at some frequency to unreduced polyketide intermediates, as with all substituted dicarbonyl compounds exposed to water [20].

The structure of EryKR1 revealed the active site groove that polyketides enter from the left side in A-type KRs and from the right side in B-type KRs, resulting in opposite β-hydroxyl group chiralities [18]. However, because features surrounding the EryKR1 active site were not highly ordered, the structure did not elucidate how a polyketide could be differentially guided into the groove. It was also unclear how KRs are specific for the epimer they reduce. In order to determine the mechanisms that set α- and β-substituent chirality, the atomic resolution structure of the KR active site as it is arranged during catalysis is required.

TylKR1 was solved to 1.95 Å resolution (Figure 2B). It has the same fold as EryKR1; however, the active site is more ordered and may represent the conformation adopted during catalysis. Most notably, αFG and the loop that precedes it, referred to as the "lid" in related short-chain dehydrogenase/reductase (SDR) enzymes, help form the substrate binding site [21, 22]. Through modeling the nicotinamide portion of NADPH and a diketide substrate into the active site, the mechanisms by which key residues control α- and β-substituent chirality are elucidated. Structural data, sequence alignments, and substituent chiralities observed in polyketides, analyzed together, reveal that at least six types of KRs exist. A protocol to deduce the ketide unit that is added by a module based on its sequence is presented.

RESULTS

Overall Structure

As with EryKR1, TylKR1 is composed of two subdomains, each resembling a Rossmann fold (Figure 2B) [18]. The structural subdomain does not have a dinucleotide binding site, and its principal role is apparently to stabilize the catalytic subdomain for catalysis. The catalytic residues

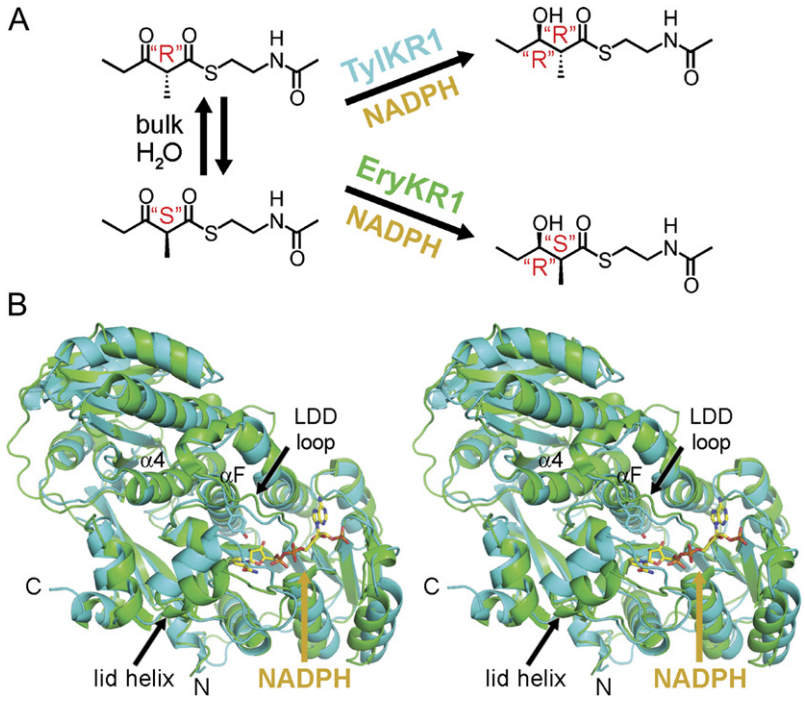

Figure 2. Isolated KRs

(A) Both TylKR1 and EryKR1, separated from their PKSs, maintain stereocontrol of the α and β positions. The enzymes choose opposite epimers for the reduction reaction. The diketide substrate analog exists as a racemic mixture, as dicarbonyl compounds spontaneously epimerize in water.
(B) A superposition of TylKR1 (blue) and EryKR1 (green) reveals the 7 Å shift of the N-terminal end of the lid helix. The order around the active site indicates that TylKR1 is in the closed, or active, conformation. NADPH is from the EryKR1 structure.

cooperate as observed in related SDR enzymes [22]: the catalytic tyrosine, Y383, and a strictly conserved serine, S370, position the carbonyl that is to be reduced adjacent to the reactive NADPH hydrogen (numbering is based on the first observable residue in the structure; to obtain actual numbering, add 1961) (Figure 3A). Y383 is activated into a general acid by a neighboring lysine, K345, to donate its proton to the carbonyl oxygen after hydride transfer. Mutation of the catalytic tyrosine to a phenylalanine prevented EryKR6 from catalyzing the reduction reaction [12].

The overall structures of TylKR1 and EryKR1 are quite similar (1.70 Å C_α root-mean-square deviation), but it is their differences that are informative (Figure 2B). The N-terminal end of the lid helix is 7 Å closer to the active site than observed in EryKR1 (it was distant from the active site in one crystal form and unobservable in another; Protein Data Bank [PDB] ID codes 2FR0 and 2FR1). Residues from the lid helix make specific interactions with active site residues: E424 hydrogen bonds with Q380; L426 contacts L411 and L431; and the G422 carbonyl hydrogen bonds with N377. The residues of the LDD motif, which lie on the "LDD loop," are also more ordered than observed in the EryKR1 structure (Figure 3A). The first aspartate caps αF via the catalytic tyrosine amide, while the second aspartate caps the lid helix. Compared to EryKR1, a conserved methionine on the loop preceding the lid helix, M417, is more ordered and closer to where the nicotinamide portion of NADPH binds.

TylKR1 crystallized in the absence of NADPH, enabling a comparison of its empty binding site with the bound site of EryKR1. The dinucleotide binding motif (consensus: TGG[S,T][G,A][G,A,T,V][V,I,L][G,A]), which makes contact with the adenine ribose and associated phosphoryl groups in EryKR1, remains well ordered in the absence of NADPH and is in the same conformation observed in EryKR1. However, several residues that would have made contact with the adenine ring (R271 and D299-E302) have elevated temperature factors.

The Ternary Complex and ACP Docking

The well-ordered, specific interactions of active site residues with the lid helix and the LDD loop indicate that TylKR1 is in a closed conformation representative of its active state during catalysis. To detail how substituent chiralities are set by KR, the nicotinamide portion of NADPH and the unepimerized diketide substrate were modeled into the active site (Figure 3B). The nicotinamide half of NADPH was placed in the orientation observed in the ternary complex of the related tropinone reductase, with the hydroxyl groups of the nicotinamide ribose hydrogen bonding to the catalytic tyrosine, Y383, and neighboring lysine, K345 (PDB ID code 1IPF) [23]. The β-carbonyl of the diketide substrate was positioned between S370 and Y383. The polyketide was oriented to enter the active site groove from the right side, as observed in related enzymes that operate on phosphopantetheinyl-bound substrates

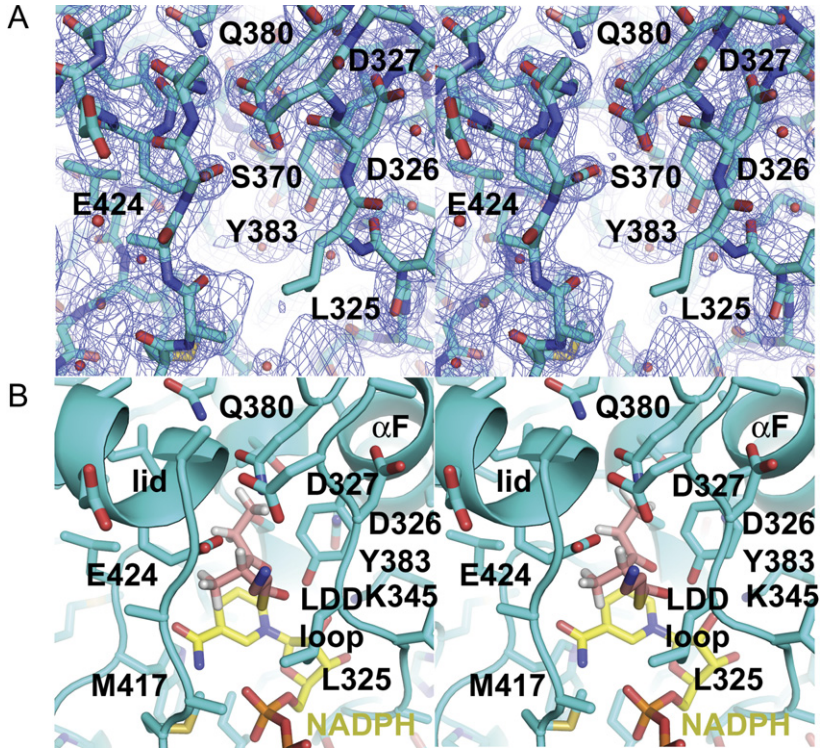

Figure 3. The KR Active Site

(A) The $2F_o - F_c$ electron density map (1.2 σ) reveals the order of key residues in the lid and LDD loop that surround the active site.

(B) NADPH and the diketide substrate are modeled in the active site. L325 guides the polyketide from the right side of the active site groove by making hydrophobic contact with the terminal portion of the phosphopantetheinyl arm. The β-carbonyl accepts a hydrogen bond from Y383. The α-hydrogen interacts with the E424 carboxylate. The epimerized diketide might not bind due to a clash between the α-methyl group and the E424 carboxylate.

(e.g., FabI; PDB ID code 1BVR) [24]. This places the thioester carbonyl adjacent to the nicotinamide ribose 2′-hydroxyl group and the terminal hydrophobic portion of the phosphopantetheinyl arm adjacent to L325 from the LDD motif.

In the ternary complex, TylKR1 makes several favorable contacts to the α-hydrogen and α-methyl group of the unepimerized diketide (Figure 3B). E424 from the lid helix, positioned by Q380 and the G421 amide, hydrogen bonds with the partially positive α-hydrogen. The γ-methylene unit of E424 makes a hydrophobic contact with the α-methyl group. An epimerized diketide would not fit due to a steric clash between its α-methyl group and the E424 carboxylate.

ACP must dock to KR in order to swing the polyketide via the phosphopantetheinyl arm into the KR active site (Figure 4A). Adjacent to the phosphopantetheinylated serine of ACP is a conserved hydrophobic residue that is hypothesized to interact with a hydrophobic patch on the KR surface during docking [25]. Experiments on the related enzyme FabG suggest this surface is near the C-terminal end of α4 [26]. Using an FAS ACP as a model (PDB ID code 2FAE), TylACP1 was docked to TylKR1 while keeping its N terminus in proximity to the TylACP1 C terminus, as they are joined by ~20 residues [27]. From this docking site, the phosphopantetheinyl arm is in a position to enter the active site groove from the left side in A-type KRs and from the right side in B-type KRs.

KR Types

Until now, KRs have only been classified by the β-hydroxyl group stereochemistry of their products: A-type KRs produce an "S" hydroxyl group and are characterized through the absence of an LDD motif and the presence of a highly conserved tryptophan. B-type KRs produce an "R" hydroxyl group and are characterized through the presence of the LDD motif (the leucine is occasionally replaced by valine or isoleucine; the first aspartate can be replaced by several residues; the second aspartate is strictly conserved) (Figure 5).

As KRs maintain stereocontrol when isolated from their PKSs, it follows that they are responsible for setting both the resulting β-hydroxyl group and α-substituent stereochemistries. Because several combinations of these stereochemistries can be achieved, KRs are not fully described as either A-type or B-type. From a comparison of the EryKR1 and TylKR1 structures and sequence alignments of KRs that catalyze distinct reactions, previously undetected fingerprints were identified for six KR types. It is clear that there are four types of reductase-competent KRs and two types of reductase-incompetent KRs. Therefore, the current nomenclature has been expanded by

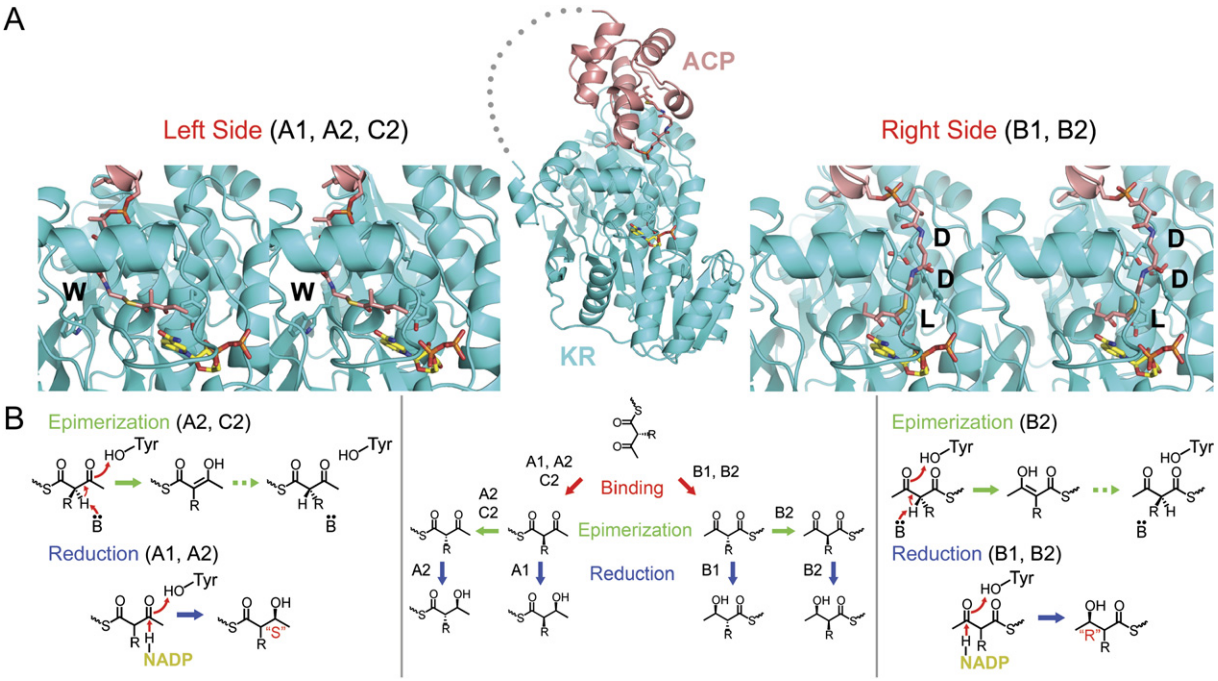

Figure 4. KR Catalysis

(A) The polyketide chain may be protected from epimerization within ACP until docking to KR. In A- and C2-type KRs, the phosphopantetheinyl arm slips behind the lid helix into the active site groove to make contact with a conserved tryptophan, so that the polyketide enters the active site from the left side. In B-type KRs, the phosphopantetheinyl arm encounters the LDD motif, which prevents it from slipping behind the lid helix, so that the polyketide enters the active site from the right side. Only the nicotinamide half of NADPH is shown.

(B) Center: through a combination of binding, epimerization, and reduction events, each KR type controls how a polyketide is processed. Left and right: catalysis proceeds similarly whether the polyketide enters a KR active site from the left side or from the right side. Epimerization is likely a combination of a catalyzed enolization and an uncatalyzed tautomerization back to the keto form. The catalytic base cannot catalyze enolization of the epimerized polyketide because the acidic hydrogen is inaccessible. In A2- and C2-type KRs, the base is either a lid helix residue or water; in B2-type KRs, the base may be the catalytic tyrosine. Reduction occurs through attack of the β-carbon by the NADPH hydride along with the transfer of a proton from the catalytic tyrosine to the carbonyl oxygen.

appending 1 if the α-substituent is not epimerized and 2 if the α-substituent is epimerized, and by including reductase-incompetent KRs as "C-type" KRs.

The roles of conserved residues (indicated by bold numbering in Figure 5) that enable each KR type to uniquely select and process a polyketide substrate are described as follows (Figures 4 and 5).

A1-Type KRs

The unepimerized polyketide enters the active site groove from the left side, guided by the conserved tryptophan (**2**). A glutamine or leucine (**3**) is in a position to interact with the unepimerized α-substituent (PhoKR6 employs a serine at this position to accommodate an ethyl substituent). The reduction reaction yields an "S" hydroxyl group.

A2-Type KRs

The unepimerized polyketide enters the active site groove from the left side, guided by the conserved tryptophan (**2**). The polyketide is enolized when the acidic α-hydrogen is abstracted by a base (either a lid helix residue or a water molecule). An uncatalyzed tautomerization back to the keto form eventually yields the epimerized polyketide. The epimerized polyketide is no longer a substrate for enolization, as its α-hydrogen is inaccessible to the base. A conserved histidine (**3**) is in position to help select against the unepimerized polyketide by sterically clashing with its α-substituent. The reduction yields an "S" hydroxyl group.

B1-Type KRs

The unepimerized polyketide enters the active site groove from the right side, guided by the leucine of the LDD motif (**1**). As in TylKR1, a glutamine (**3**) can collaborate with lid helix residues to prevent a spontaneously epimerized polyketide from being reduced. The reduction yields an "R" hydroxyl group.

KRs are usually B1 type when they work in concert with the other β-carbon processing enzymes DH and ER. This is readily observed in the avermectin PKS: the catalytic subdomains of AveKR7 and AveKR9 are identical in sequence and are highly likely to produce the same combination of stereochemistries at the α- and β-carbons. The chiralities set by AveKR9 are erased through a subsequent dehydration reaction catalyzed by AveDH9. However, AveKR7 is observed to produce an "R" α-substituent and an "R" β-hydroxyl group, making it a B1-type KR. Thus, AveKR9 and most KRs that produce a substrate for a DH are B1-type KRs.

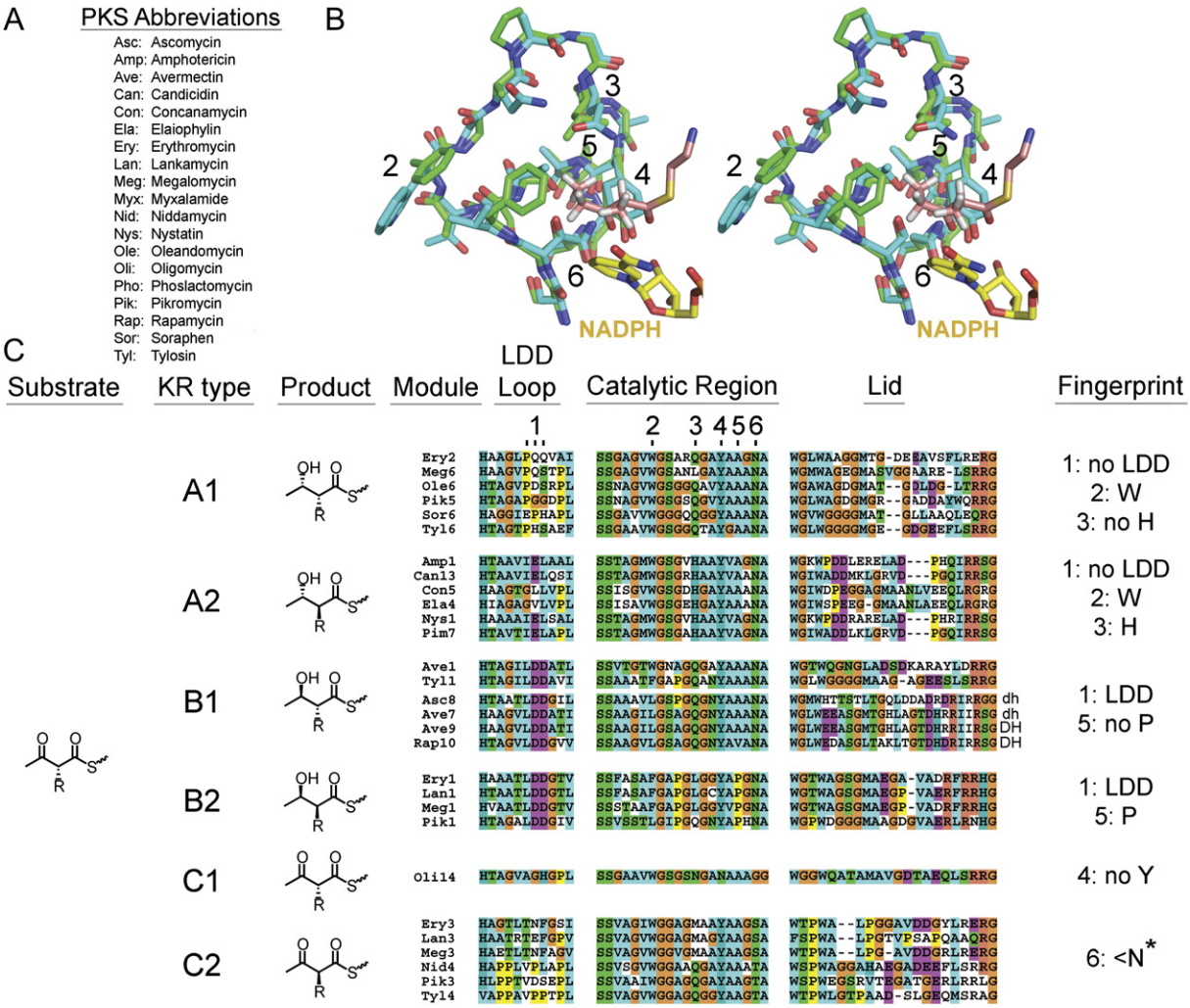

Figure 5. KR Fingerprints

(A) PKS abbreviations.

(B) An alignment of the TylKR1 and EryKR1 catalytic regions illustrates the architectural differences between KRs. The NADPH and substrate are from the model of the TylKR1 ternary complex.

(C) Each of the KR types uniquely processes a similar substrate through controlling the chirality of α-substituents and β-hydroxyl groups. Several regions from representative KRs are aligned to illustrate the fingerprints that can be observed in their sequences. The lid is extremely variable in sequence and length, indicating that it may make specific interactions with a bound polyketide. The lid is longer and less variable when an accompanying DH is present. DH, functional; dh, nonfunctional; *, or obscured dinucleotide binding motif.

Interestingly, most of the B1-type KRs possess a highly conserved and slightly longer lid, probably consisting of two α helices, as in the related SDR enzymes FabG and FabI [21, 24]. This lid is always present when KR is accompanied by a DH. If modular PKSs are descended from FASs [5], which contain B-type KRs, it is plausible that all PKS KRs evolved from B-type KRs.

B2-Type KRs

The unepimerized polyketide enters the active site groove from the right side, guided by the leucine of the LDD motif (**1**). The catalytic tyrosine, freed from αF by a conserved proline (**5**), may have sufficient flexibility to abstract the acidic α-hydrogen to enolize the polyketide [18]. A productive tautomerization back to the keto form generates the epimerized polyketide, which cannot be enolized by KR because the α-hydrogen is inaccessible to the tyrosine. A leucine or a glutamine (**3**) can collaborate with the lid helix to help select the epimerized polyketide for reduction. In EryKR1, the combination of a leucine at this position and a valine from the lid helix creates a hydrophobic pocket for the epimerized α-methyl group. The reduction yields an "R" hydroxyl group.

C1-Type KRs

These rare, nonfunctional KRs lack the catalytic tyrosine (**4**). Because an unepimerized, unreduced ketide unit can be added by modules completely lacking a KR, there may be no evolutionary pressure for modules to maintain these KRs.

C2-Type KRs

The unepimerized polyketide enters the active site groove from the left side, guided by the conserved tryptophan (**2**). The polyketide is enolized when the acidic α-hydrogen is abstracted by a base (either a lid helix residue or a water molecule). An uncatalyzed tautomerization back to the keto form eventually yields the epimerized polyketide. The epimerized polyketide is no longer a substrate for enolization, as its α-hydrogen is inaccessible to the base. In these KRs, the dinucleotide binding motif is obscured and the conserved asparagine (**6**) is typically replaced by a smaller residue.

Protocol to Assign Substituent Chirality

From the fingerprints of the six KR types, how a given KR will enforce substituent chiralities can be deduced from its sequence (Figure 6). By combining this predictive scheme in the framework of how a module produces a ketide unit, a flowchart has been created that can be used to deduce what addition a module will make to a growing polyketide chain. This methodology has been developed using every sequenced PKS that produces a structurally validated polyketide.

For the protocol to be completely predictive of substituent chirality from modular sequence information alone, the fingerprints of different types of DHs and ERs need to be identified. Several PKSs use DHs as epimerases. Although the conservation of putative DH active site residues is suggestive of a dehydration-competent enzyme, sequence information cannot predict whether a DH is an epimerase, as the associated fingerprint has not yet been determined. However, if a DH is known not to perform the dehydration reaction and its catalytic residues are conserved, it is likely to be an epimerase. Because there are so few examples of ERs that reduce α-substituted polyketides, the fingerprints of "epimerizing" and "nonepimerizing" ERs have yet to be established.

Mutational Analysis of Stereocontrol

The kinetic parameters of EryKR1 and TylKR1 operating on an unreduced diketide substrate have previously been measured (Figure 2A) [19]. The β-hydroxyl group of both products has the same "R" stereochemistry; however, TylKR1 enforces an "R" stereochemistry on the α-methyl group, and EryKR1 enforces an "S" stereochemistry on the α-methyl group. The ternary model suggests that Q380 and E424 are important in selecting against the epimerized diketide in TylKR1. Because the residue equivalent to Q380 in EryKR1 is a leucine, the Q380L mutant was engineered and assayed for stereocontrol (see the Supplemental Data available with this article online for diketide substrate synthesis). The mutant retained the ability to select the unepimerized diketide, indicating that E424 from the lid helix is the most important residue in setting α-substituent chirality. Thus, the E424A mutant was engineered and assayed for stereocontrol. This mutant was compromised, selecting both epimers for reduction, although it favored its natural epimer 4:1.

DISCUSSION

Stereocontrol

TylKR1 was crystallized in the active conformation adopted during catalysis, in which the lid helix and the LDD motif make specific interactions with active site residues (Figure 3). Modeling of the nicotinamide portion of NADPH and the diketide substrate into the active site shows that the diketide is completely surrounded by NADPH and protein during catalysis and suggests that the lid helix interacts with bound substrates to ensure a complementary fit.

How B-type KRs use the LDD motif to guide polyketides into the active site groove is made apparent by considering how the phosphopantetheinyl arm swings from a docked ACP into the KR active site groove (Figure 4A). Because the invariant second aspartate caps the lid helix, sealing the LDD loop and the lid helix together, the phosphopantetheinyl arm is prevented from slipping between them into the active site groove. To enter the groove, the arm must swing over the aspartates so that its hydrophobic end interacts with the leucine. Polyketides are thus positioned by B-type KRs to enter the active site from the right side.

The active sites of A- and C2-type KRs accept polyketides from the left side. These KRs lack the LDD motif but possess a conserved tryptophan on the left side of the groove. ACPs probably dock to these KRs very similarly to how they dock to B-type KRs. However, when the phosphopantetheinyl arm swings toward the active site, it does not encounter the seal between the lid helix and the LDD loop that is made in B-type KRs, and can slip between them. The arm inserts into the KR active site groove until it contacts the conserved tryptophan. The tryptophan makes space on the left side of the groove for the phosphopantetheinyl arm by burying its side chain in the hydrophobic core (as observed in TylKR1, one of the rare B-type KRs that possess this tryptophan) (Figure 5B). Polyketides are thus positioned by A- and C2-type KRs to enter the active site from the left side.

The interaction of the lid helix with polyketide α-substituents in A- and B-type KRs helps select the epimer that is to be reduced. As revealed by the TylKR1 Q380L and E424A mutants, the primary determinant of α-substituent stereochemistry is the lid helix residue E424. In EryKR1, a conserved valine from the lid helix is in a location to cooperate with the leucine three residues before the catalytic tyrosine in creating a hydrophobic pocket for the epimerized methyl group. In general, the variability of the lid helix prevents any strong fingerprints in this region from being detected.

Epimerization

From the deuterium-labeling experiments, it is apparent that the triketide formed by EryKS2 does not epimerize appreciably [14], suggesting two possibilities: either spontaneous epimerization does not occur on the timescale of polyketide production or the polyketide is shielded from spontaneous epimerization. As polyketides are synthesized by PKSs on the order of minutes and model diketide

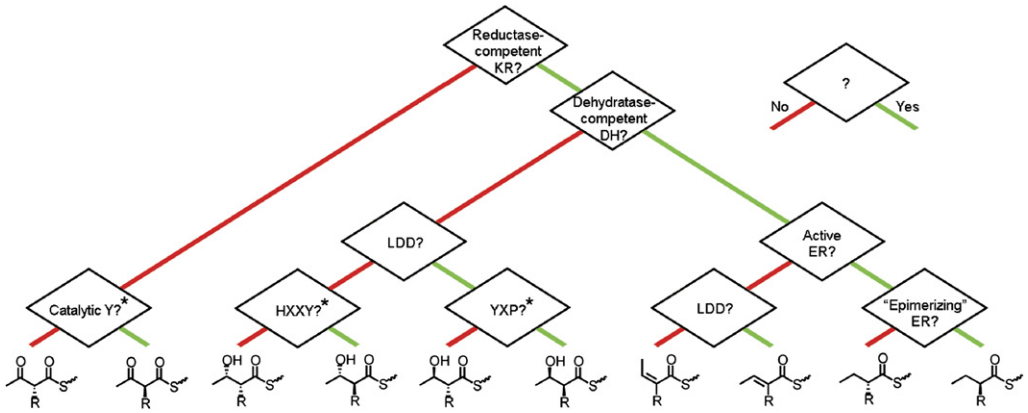

Figure 6. The Substituent Flowchart

By combining the KR fingerprints, indicative of how α- and β-substituents are set, with how DH and ER can process intermediates, a protocol has been developed to assign which of these ten ketide units will be added by a module. The first six are direct outcomes of the KR type present in the module; the last four rely on the types of DH and ER present. The R groups can be determined by fingerprints in the accompanying AT domain [29]. *, DH epimerase activity?

substrates epimerize in water at least that quickly, the second model is preferred [28]. As with fatty acid chains in an FAS, polyketide chains may be sequestered inside the ACP [27]. Within the ACP helical bundle, the unreduced polyketide would be protected from bulk solvent and, hence, spontaneous epimerization, as it is shuttled between the KS and KR active sites (Figure 4A).

That epimerizing KRs completely epimerize a polyketide substrate seems to challenge the doctrine that enzymes cannot alter the natural ratio between substrate and product [18] (Figure 4B). However, an epimerizing KR is probably best defined as an enolase, so its reaction product is actually an enol. Based on the available structures, the most likely enolization mechanism is that a base abstracts the polyketide α-hydrogen and the catalytic tyrosine donates its hydrogen to the polyketide β-carbonyl oxygen. The resulting enol can spontaneously tautomerize to the lower-energy epimerized polyketide and be kinetically trapped as such in the KR active site. The epimerized polyketide may not be a substrate for enolization because its α-hydrogen is inaccessible to the catalytic base.

In B2-type KRs, the putative epimerization mechanism is that the tyrosine, freed from αF by a proline, acts as the catalytic base, abstracting the acidic α-hydrogen and donating its hydrogen to the β-carbonyl oxygen [18] (Figure 4B). After a productive tautomerization, the epimerized polyketide is no longer a substrate for the epimerization reaction, as its α-hydrogen is on the other side of the polyketide, inaccessible to the tyrosine. In A2- and C2-type KRs, the reaction must be catalyzed differently, as the unreduced, unepimerized polyketide enters the active site from the left side. The α-hydrogen is most likely abstracted by a lid helix residue or water molecule to form the enol. The epimerized polyketide, formed when the enol productively tautomerizes, is not a substrate for enolization, as the base can no longer access the α-hydrogen.

Cryptic epimerization is not supported by current biochemical and structural data to be a general mechanism [15]. The deuterium-labeling experiments demonstrate that the α-hydrogen is not lost from the polyketide as it passes through EryMod2, indicating that the polyketide is not in the enol form in the A1-type EryKR2. A polyketide enol cannot be formed in the B1-type TylKR1, due to a steric clash that would result with E424. The mechanism is also not sterically feasible in the B2-type EryKR1, as a hydride and a proton would need to be added concurrently from the same face of the polyketide at the β- and α-carbons. The mechanism is only a formal possibility in A2-type KRs, which possess a conserved histidine in the necessary location to donate a proton to the α-carbon.

Several PKSs contain DHs that catalyze epimerization. DHs may be more natural at this reaction than KRs, as the first step in the dehydration mechanism is the abstraction of the α-proton. There are several examples: MyxDH2 performs epimerization prior to a reduction catalyzed by an A-type KR; AscDH2, RifDH6, and RifDH7 perform epimerizations prior to a reduction catalyzed by a B-type KR; and RapDH3 and RapDH6 perform epimerizations that are not followed by reduction.

Protocol to Assign Substituent Chiralities

By combining the fingerprints indicative of how α- and β-substituent chiralities are set by KRs with how DHs and ERs can further process the resulting intermediates, a protocol has been generated to assign polyketide substituent chirality based on the sequence of a module (Figure 6). The first six possibilities are direct outcomes of the KR type that acts on a polyketide in a given module (unless an epimerizing DH is present). The next two outcomes depend on both the KR type and the presence of an active DH. In rare cases, such as RifMod10, a DH accompanies an A1-type KR to produce a *cis* double bond [12]. The final two outcomes depend on whether ER is "epimerizing"

(producing an α-substituent of "S" stereochemistry). More examples of ERs must be discovered for these fingerprints to be proposed. To determine the identity of an α-substituent, fingerprints in the accompanying AT can be examined: an HAFH motif indicates that it selects a malonyl group so there is no α-substituent, whereas a YASH motif indicates that it selects a methylmalonyl group so the α-substituent is a methyl group. The fingerprints for other AT selectivities are not as well defined, as there are fewer examples of them [29].

The protocol presented here can be used to both verify and derive polyketide structures. Small-molecule NMR and crystallography do not always yield correct stereochemical assignments, hence polyketide substituent chiralities are often incompletely and even incorrectly annotated [30, 31]. Dictyostatin is a 22-membered macrolide that exhibits Taxol-like properties. Before attempting its synthesis, chemists were wise to thoroughly examine its reported structure—it was misassigned at seven chiral centers [30, 32]. If the dictyostatin PKS sequence were available to the chemists, the protocol would have presented an easier route to verifying the stereocenters. The PKSs for many polyketides are available. By applying the protocol to the recently reported structure of meridamycin using the associated PKS sequence, more than half of the stereocenters appear to be misassigned [33]. A careful examination of its structure should be pursued. Although the protocol is not infallible, its utility is undeniable.

Engineering Polyketide Synthases

The obvious route of engineering PKSs through swapping KR domains has only yielded a few successes [34, 35]. This technique may suffer from the incompatability between the inserted KR and the native ACP. An alternative approach, which does not disrupt domain interactions, is to alter KR selectivity through site-directed mutagenesis [13]. Recently, an EryKR1 mutant with a tryptophan engineered into the left side of the groove was shown to exclusively produce a reduced diketide with a β-hydroxyl group of "S" stereochemistry. Thus, a B2-type KR was successfully converted into an A2-type KR. If such engineered KRs are functional inside PKSs, then libraries of novel polyketides can be generated through the combinatorial mutagenesis of KRs within PKSs.

SIGNIFICANCE

In order to engineer PKSs that produce novel polyketides of medicinal value, the mechanisms of their component enzymes must be deciphered. The KR is a principal target in PKS engineering, as it controls the majority of polyketide stereocenters. By comparing the structure of the tylosin KR reported here with a recently solved erythromycin KR, residues involved in polyketide binding, epimerization, and reduction, and how they cooperate to control substituent stereochemistries, were elucidated. The nomenclature for KR types, which is currently only reflective of how KRs set the hydroxyl group stereochemistry, was expanded to reflect how KRs also set the neighboring acyl-substituent stereochemistry. A methodology was created to deduce the identities and chiralities of substituents added by a PKS module based on the sequences of its component enzymes. Together, these advances will give thrust to the development of novel polyketide pharmaceuticals.

EXPERIMENTAL PROCEDURES

Cloning

DNA encoding the KR fragment was amplified from *S. fradiae* genomic DNA with primers 5′-GCAGATATACATATGAGCCCCACCGATGCCTG GCGC-3′ and 5′-GTGGTGCTCGAGTCATCAGGCTGCCGTCAGGGC CTCCCG-3′, digested with NdeI and XhoI, and inserted into pET28b between the NdeI and XhoI sites. To create the Q380L mutant, QuikChange (Stratagene) was performed with 5′-GGCAACGCCGGCCTG GGTGCGTACGCC-3′ and 5′-GGCGTACGCACCCAGGCCGGCGTT GCC-3′. To create the E424A mutant, QuikChange was performed with 5′-GGCGGGTGCGGGCGCGGAGAGTCTGTCGC-3′ and 5′-GCGACAGACTCTCCGCGCCCGCACCCGCC-3′.

Protein Expression, Purification, and Crystallization

For each protein, *Escherichia coli* BL21(DE3) cells were transformed with the appropriate plasmid. The cells were grown in Luria broth at 37°C to OD_{600} = 0.4, cooled to 15°C, and induced with 1 mM IPTG. After 14 hr, the cells were harvested, resuspended in lysis buffer (0.5 M NaCl, 30 mM Tris [pH 7.4]), and lysed by sonication. After centrifugation, the lysate was poured over a nickel-NTA column equilibrated with lysis buffer. The bound protein was washed with 15 mM imidazole in lysis buffer and eluted with 150 mM imidazole in lysis buffer. The protein was passed through a Superdex 200 gel-filtration column equilibrated with 150 mM NaCl, 10 mM Tris (pH 7.4) and concentrated to 10 mg/ml.

TylKR1 crystallized in 1.5 M ammonium sulfate, 0.1 M Tris (pH 8.0) by hanging-drop vapor diffusion at 22°C with a protein to crystallization buffer ratio of 1:1. Crystals were cryoprotected in 20% glycerol, 1.8 M ammonium sulfate, 0.1 M Tris (pH 8.0) for 5 s prior to freezing in liquid nitrogen. Cocrystallizing and soaking with NADPH and diketides were attempted; however, no density from these substrates was observed.

Data Processing and Refinement

Data collected at Advanced Light Source beamline 8.3.1 were processed with HKL2000 [36] (Table 1). The EryKR1 structure (PDB ID code 2FR0) was used as a search model for molecular replacement in Phaser [37]. Several rounds of refinement in CNS and model building in Coot were performed [38, 39]. Two large loops were not observable in the electron density maps corresponding to 24 residues between β6 and α5 and 11 residues between β1 and β2.

Enzymatic Activity Assays

Enzymes were incubated with 5 mM NADPH and 0.5 mM diketide substrate analog in 300 mM Na_2PO_4 (pH 8.0) at 22°C for 1 hr. Ethyl acetate extracts of the reactions were analyzed by LC/MS with a C_{18} reverse-phase column (20%–30% acetonitrile/water gradient).

Supplemental Data

Supplemental Data contain the synthesis of the diketide substrate with NMR spectral data and one figure and are available at http://www.chembiol.com/cgi/content/full/14/8/898/DC1/.

Table 1. Crystallographic Data and Refinement Statistics

Data Collection	
Space group	*P2(1)*
Cell dimensions	
a, *b*, *c* (Å)	66.2, 49.8, 68.8
β (°)	109.2
Resolution (Å)	50–1.95
R_{merge}	0.066 (0.498)
I/σ(I)	19.6 (1.7)
Completeness (%)	93.4 (61.4)
Redundancy	6.4 (3.5)
Refinement	
Resolution (Å)	50–1.95
Unique reflections	28,533
R_{work}/R_{free}	0.259/0.287
Number of atoms	
Protein	3,303
Water	193
B factors	
Protein	43
Water	46
Rms deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.1

Statistics in parentheses refer to the highest resolution shell (1.95–2.02 Å).

ACKNOWLEDGMENTS

Ralph Reid, Chris Reeves, and Hugo Menzella shared their knowledge of PKS enzymology. Robert Stroud and Janet Finer-Moore assisted in interpreting experimental results. Research was supported by National Cancer Institute, National Institutes of Health grant CA63081 (to Robert Stroud).

Received: April 7, 2007
Revised: June 9, 2007
Accepted: July 9, 2007
Published: August 24, 2007

REFERENCES

1. Staunton, J., and Wilkinson, B. (1997). Biosynthesis of erythromycin and rapamycin. Chem. Rev. *97*, 2611–2630.
2. O'Hagan, D.O. (1991). The Polyketide Metabolites (Chichester, UK: Ellis Horwood).
3. Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.
4. 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.
5. Sherman, D.H., and Smith, J.L. (2006). Clearing the skies over modular polyketide synthases. ACS Chem. Biol. *1*, 505–509.
6. Cundliffe, E., Bate, N., Butler, A., Fish, S., Gandecha, A., and Merson-Davies, L. (2001). The tylosin-biosynthetic genes of *Streptomyces fradiae*. Antonie Van Leeuwenhoek *79*, 229–234.
7. Tang, Y., Kim, C.Y., Mathews, I.I., Cane, D.E., and Khosla, C. (2006). The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. Proc. Natl. Acad. Sci. USA *103*, 11124–11129.
8. Smith, S., Witkowski, A., and Joshi, A.K. (2003). Structural and functional organization of the animal fatty acid synthase. Prog. Lipid Res. *42*, 289–317.
9. Maier, T., Jenni, S., and Ban, N. (2006). Architecture of mammalian fatty acid synthase at 4.5 Å resolution. Science *311*, 1258–1262.
10. Donadio, S., and Katz, L. (1992). Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*. Gene *111*, 51–60.
11. Caffrey, P. (2005). The stereochemistry of ketoreduction. Chem. Biol. *12*, 1060–1062.
12. Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D.V., Hutchinson, C.R., and McDaniel, R. (2003). A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. Biochemistry *42*, 72–79.
13. 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*, 277–285.
14. Weissman, K.J., Timoney, M., Bycroft, M., Grice, P., Hanefeld, U., Staunton, J., and Leadlay, P.F. (1997). The molecular basis of Celmer's rules: the stereochemistry of the condensation step in chain extension on the erythromycin polyketide synthase. Biochemistry *36*, 13849–13855.
15. Starcevic, A., Jaspars, M., Cullum, J., Hranueli, D., and Long, P.F. (2007). Predicting the nature and timing of epimerisation on a modular polyketide synthase. ChemBioChem *8*, 28–31.
16. Cane, D.E., Liang, T.C., Taylor, P.B., Chang, C., and Yang, C.C. (1986). Macrolide biosynthesis. 3. Stereochemistry of the chain-elongation steps of erythromycin biosynthesis. J. Am. Chem. Soc. *108*, 4957–4964.
17. Holzbaur, I.E., Ranganathan, A., Thomas, I.P., Kearney, D.J., Reather, J.A., Rudd, B.A., Staunton, J., and Leadlay, P.F. (2001). Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. Chem. Biol. *8*, 329–340.
18. 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*, 737–748.
19. Siskos, A.P., Baerga-Ortiz, A., Bali, S., Stein, V., Mamdani, H., Spiteller, D., Popovic, B., Spencer, J.B., Staunton, J., Weissman, K.J., and Leadlay, P.F. (2005). Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. Chem. Biol. *12*, 1145–1153.
20. Burdett, G.L., and Rogers, M.T. (1964). Keto-enol tautomerism in β-dicarbonyls studied by nuclear magnetic resonance spectroscopy 1. Proton chemical shifts and equilibrium constants of pure compounds. J. Am. Chem. Soc. *86*, 2105–2109.
21. 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*, 339–347.

22. 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*, 247–253.

23. Yamashita, A., Endo, M., Higashi, T., Nakatsu, T., Yamada, Y., Oda, J., and Kato, H. (2003). Capturing enzyme structure prior to reaction initiation: tropinone reductase-II-substrate complexes. Biochemistry *42*, 5566–5573.

24. Rozwarski, D.A., Vilcheze, C., Sugantino, M., Bittman, R., and Sacchettini, J.C. (1999). Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD^+ and a C_{16} fatty acyl substrate. J. Biol. Chem. *274*, 15582–15589.

25. 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*, 8231–8238.

26. 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*, 52935–52943.

27. Roujeinikova, A., Simon, W.J., Gilroy, J., Rice, D.W., Rafferty, J.B., and Slabas, A.R. (2007). Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. J. Mol. Biol. *365*, 135–145.

28. Holzbaur, I.E., Harris, R.C., Bycroft, M., Cortes, J., Bisang, C., Staunton, J., Rudd, B.A., and Leadlay, P.F. (1999). Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase. Chem. Biol. *6*, 189–195.

29. Yadav, G., Gokhale, R.S., and Mohanty, D. (2003). Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. J. Mol. Biol. *328*, 335–363.

30. Nicolaou, K.C., and Snyder, S.A. (2005). Chasing molecules that were never there: misassigned natural products and the role of chemical synthesis in modern structure elucidation. Angew. Chem. Int. Ed. Engl. *44*, 1012–1044.

31. Kitamura, M., Nakano, K., Miki, T., Okada, M., and Noyori, R. (2001). Stereochemistry of aldols: configuration and conformation of aldols derived from cycloalkanones and aldehydes. J. Am. Chem. Soc. *123*, 8939–8950.

32. Paterson, I., Britton, R., Delgado, O., and Wright, A.E. (2004). Stereochemical determination of dictyostatin, a novel microtubule-stabilising macrolide from the marine sponge *Corallistidae* sp. Chem. Commun. (Camb.) (6), 632–633.

33. Sun, Y., Hong, H., Samborskyy, M., Mironenko, T., Leadlay, P.F., and Haydock, S.F. (2006). Organization of the biosynthetic gene cluster in *Streptomyces* sp. DSM 4137 for the novel neuroprotectant polyketide meridamycin. Microbiology *152*, 3507–3515.

34. Kao, C.M., McPherson, M., McDaniel, R.N., Fu, H., Cane, D.E., and Khosla, C. (1998). Alcohol stereochemistry in polyketide backbones is controlled by the β-ketoreductase domains of modular polyketide synthases. J. Am. Chem. Soc. *120*, 2478–2479.

35. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., and Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. Proc. Natl. Acad. Sci. USA *96*, 1846–1851.

36. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In Methods in Enzymology, Volume 276: Macromolecular Crystallography, Part A, C.W. Carter, Jr. and R.M. Sweet, eds. (New York: Academic Press), pp. 307–326.

37. CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

38. 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., et al. (1998). Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

39. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Accession Numbers

The atomic coordinates of TylKR1 have been deposited in the Protein Data Bank under ID code 2Z5L.

