Chemistry & Biology 13, 277–285, March 2006 DOI 10.1016/j.chembiol.2006.01.004

Directed Mutagenesis Alters the Stereochemistry of Catalysis by Isolated Ketoreductase Domains from the Erythromycin Polyketide Synthase

Abel Baerga-Ortiz,[1] Bojana Popovic,[1] Alexandros P. Siskos,[1] Helen M. O'Hare,[1,2,3] Dieter Spiteller,[1,4] Mark G. Williams,[1,5] Nuria Campillo,[1,6] Jonathan B. Spencer,[2] and Peter F. Leadlay[1,*]

[1] Department of Biochemistry
University of Cambridge
80 Tennis Court Road
Cambridge CB2 1GA
United Kingdom

[2] Department of Chemistry
University of Cambridge
Lensfield Road
Cambridge CB2 1EW
United Kingdom

Summary

The ketoreductase (KR) domains eryKR$_1$ and eryKR$_2$ from the erythromycin-producing polyketide synthase (PKS) reduce 3-ketoacyl-thioester intermediates with opposite stereospecificity. Modeling of eryKR$_1$ and eryKR$_2$ showed that conserved amino acids previously correlated with production of alternative alcohol configurations lie in the active site. eryKR$_1$ domains mutated at these positions showed an altered stereochemical outcome in reduction of (2*R*, *S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine thioester. The wild-type eryKR$_1$ domain exclusively gave the (2*S*, 3*R*)-3-hydroxy-2-methylpentanoic acid *N*-acetylcysteamine thioester, while the double mutant (F141W, P144G) gave only the (2*S*, 3*S*) isomer, a switch of the alcohol stereochemistry. Mutation of the eryKR$_2$ domain, in contrast, greatly increased the proportion of the wild-type (2*R*, 3*S*)-alcohol product. These data confirm the role of key residues in stereocontrol and suggest an additional way to make rational alterations in polyketide antibiotic structure.

Introduction

Modular type I polyketide synthases (PKSs), for example the 6-deoxyerythronolide B synthase (DEBS) responsible for synthesis of the aglycone core of the macrolide antibiotic erythromycin A, are multienzymes that catalyze the assembly line biosynthesis of structurally complex and clinically important polyketide natural products [1, 2], starting from simple carboxylic acid building blocks. Polyketide chain assembly resembles fatty acid biosynthesis in that acyl-CoA starter units and extender units are condensed together, without the intermediates being released into solution. In contrast to fatty acid biosynthesis, modular PKSs utilize a wider variety of starter and extender units, and the β-keto functionality created after each condensation step is not necessarily fully reduced before the next cycle of chain extension, leading to a far greater potential for chemical and stereochemical diversity in the products. A number of strategies have been successfully used to engineer modular PKSs, and also downstream post-PKS enzymes, to produce rationally altered analogs of bioactive polyketide products [3–7]. In contrast, attempts to alter the configuration at specific stereocenters have so far been hampered by our limited understanding of the fundamental mechanisms by which PKSs exert stereocontrol over polyketide assembly.

As illustrated in Figure 1 for the first two extension modules of DEBS, each PKS extension module contains a ketosynthase (KS) domain, which catalyzes the formation of a carbon-carbon bond, and an acyltransferase (AT) domain, which recruits the chain extension unit, here from (2*S*)-methylmalonyl-CoA. The acyl carrier protein (ACP) cooperates in the carbon-carbon bond formation to form a β-ketoester and then carries the growing chain to domains involved in reductive processing of the keto group before the next cycle of chain extension. These reductive activities are optionally present depending on the module and comprise: a β-ketoreductase (KR) domain, which catalyzes the reduction of the initially formed β-ketoester to a β-hydroxyester; a dehydratase (DH) domain, which dehydrates the β-hydroxyester; and an enoyl reductase (ER) domain, which reduces the double bond. There is a good correlation between the predicted constituent domains and the chemical structure of the corresponding chain extension unit in the growing polyketide chain.

The methyl centers at C-2 and the hydroxyl centers at C-3 can have either an *R* or *S* configuration. As first noted by Celmer, there is a strong position-specific structural and stereochemical homology in families of macrolide polyketides, suggesting a genetic origin for stereochemical control at such centers [8]. AT domains are not thought [9, 10] to control stereochemistry at methyl-branched centers in the polyketide product. Rather, it appears that KS domains in only some extension modules can catalyze the epimerization, at the methyl center of the initially formed (2*R*) isomer of the ketoester, to the (2*S*) isomer [11]. If such extension modules also contain a KR that selects only the (2*S*) isomer from the mixture, then, effectively, the KR specificity determines the configuration at the methyl center [12, 13].

The chirality at the β-hydroxyl group is determined by the direction of hydride addition to the keto group (*re* or *si* face) within the active site of the KR domain. In vivo studies have shown in a number of cases that swapping individual KR domains between modules of different PKSs results in the inversion of alcohol stereochemistry

*Correspondence: pfl10@mole.bio.cam.ac.uk
[3] Present address: Ecole Polytechnique Fédérale de Lausanne, Institute of Chemical Sciences and Engineering, 1015 Lausanne, Switzerland.
[4] Present address: Max-Planck-Institute for Chemical Ecology, Department of Bioorganic Chemistry, Hans-Knöll-Strasse 8, 07745 Jena, Germany.
[5] Present address: Astex Therapeutics, Ltd., Cambridge CB4 0WE, United Kingdom.
[6] Present address: Instituto de Quimica Organica Lora Tamayo, C/Juan de la Cierva, N3 Madrid, Spain.

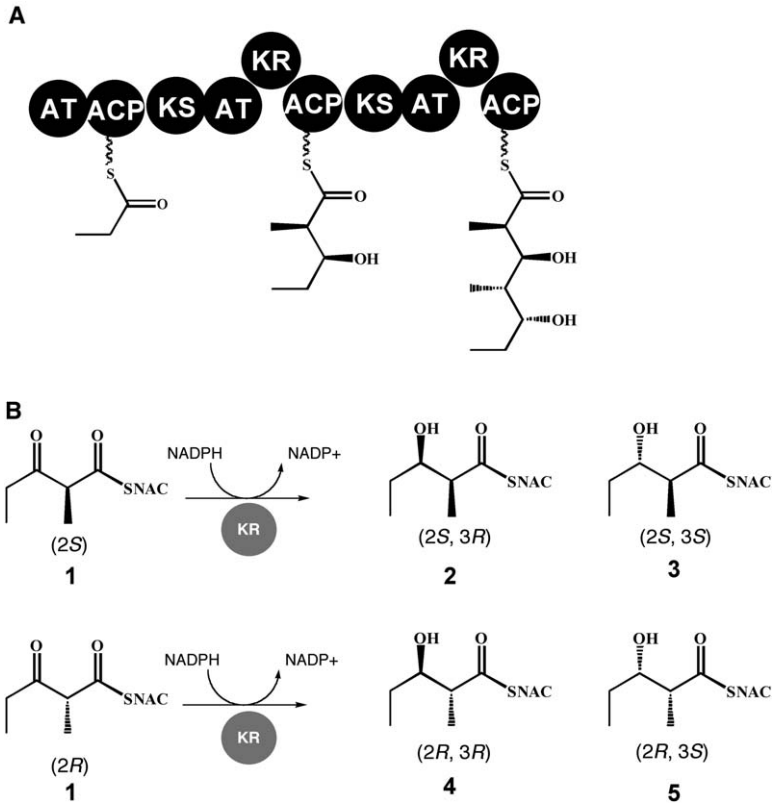

Figure 1. The Stereochemistry of Polyketide Chain Formation on the First Two Extension Modules of DEBS, the Erythromycin-Producing Modular PKS

(A) NADPH-dependent reduction by ketoreductase domains eryKR$_1$ and eryKR$_2$ produces the opposite configuration in the alcohol intermediates.

(B) Reduction of a racemic mixture of (2*R*, *S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine ester **1** to four possible diastereoisomeric alcohols **2**–**5**. Isomers **2** and **5** are the expected products from eryKR$_1$ and eryKR$_2$, respectively, based on the absolute configuration of erythromycin.

KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; ACP, acyl carrier protein; DEBS, 6-deoxyerythronolide B synthase; NAC, *N*-acetylcysteamine.

at that position in the resulting natural product, indicating that the stereospecificity of reduction is dictated by the KR domain and is transferable ([14] and L. Kellenberger and P.F.L., unpublished data). However, when a KR domain in a PKS is presented with an unnatural substrate, the stereochemical outcome may be unexpectedly altered [15]. To gain a better understanding of the factors at work, we have therefore sought to analyze the intrinsic (stereo)specificity of KR domains.

We have previously shown that the KR activities can be specifically assayed (as for fatty acid synthases [16, 17]) by using the surrogate substrate (2*R*, *S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine thioester **1** (Figure 1B), when expressed either as part of a PKS [13, 15, 18] or as soluble and enzymatically active proteins in recombinant *Escherichia coli*, in the absence of other PKS components [19]. In the latter experiments, eryKR$_1$ was found to be intrinsically both stereoselective and stereospecific, as it was able to select the expected (2*S*) enantiomer from the racemic mixture and make from it exclusively the expected (2*S*, 3*R*)-3-hydroxy-2-methylpentanoic acid *N*-acetylcysteamine thioester **2**. The eryKR$_2$ was less active toward **1**, and the expected (2*R*, 3*S*)-3-hydroxy-2-methylpentanoic acid *N*-acetylcysteamine thioester **5** was a minor product; the majority of the product consisted of **3** and **2**, obtained by preferential reduction of the "wrong" enantiomer of **1** [19]. These data support the idea that alternative modes of binding to the KR active site are responsible for different stereochemical outcomes, and they reveal that the energetic differences between these binding modes are rather small. Therefore, alteration of a few key residues at the active site might suffice to induce a change in the stereochemistry of KR-mediated reduction. There is clear precedent for this in recent work on two mutually homologous plant tropinone reductases that act on the 3-keto group of the alkaloid to form two enantiomeric products, and whose stereospecificity has successfully been switched by site-directed mutagenesis [20, 21].

It has previously been reported [22, 23] that certain amino acid residues are characteristic of natural modular PKS KRs that reduce from one or the other face of a 3-ketoacylthioester substrate. In particular, Caffrey [23] has noted, among 68 PKS KR domains where the configuration of the product is known, 2 amino acid motifs (referred to here as motif I and motif II, respectively) that are hallmarks of these 2 groups of enzymes, which he terms A- and B-side selective (Figure 2). We report here the site-specific mutagenesis of some of these amino acids and the effects of these mutations on the stereochemistry of the products of KR-mediated ketoreduction by both eryKR$_1$ and eryKR$_2$. In the absence of an X-ray crystal structure, we also have undertaken homology modeling of both eryKR$_1$ and eryKR$_2$ (which, as previously reported [22], belong to the short-chain dehydrogenase/reductase [SDR] enzyme family) to analyze for consistency of active-site architecture and also to elucidate whether the residues of the two motifs are situated close enough to the KR active site that they might plausibly contact the substrate and play a direct role in stereocontrol. Our kinetic and stereochemical analysis showed that exchange of only two of these amino acids led to an effective switch of alcohol stereochemistry in eryKR$_1$. The results of mutation of eryKR$_2$, which is not stereoselective against diketide substrates, were less clear-cut, but they also confirmed the importance of

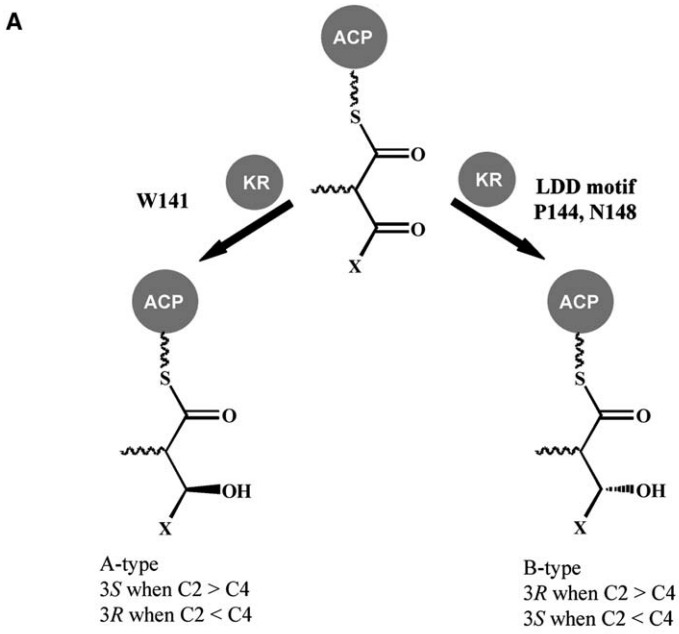

Figure 2. Amino Acid Motifs that Are Putatively Involved in the Control of Stereochemistry by KR Domains from Modular PKSs

Amino acid motifs have been identified by Caffrey [23].

(A) The strongest indicator for B-type KR domains is an LDD motif (motif I) in the region between amino acids 88 and 103, which is absent from A-type KR domains [22]. Additional amino acids in the 134–149 region (motif II), specifically P144 and N148, which indicate B-type KRs and W141 in A-type domains, support this assignment.

(B) Sequence alignment of KR domains showing the mutants constructed in this work. The active site motifs, and the changes made, are indicated in bold (residues are numbered according to [23]). The consensus sequences for A- and B-type KRs are also shown. h, aliphatic hydrophobic residue; x, unspecified residue.

B

	Motif I	Motif II
KR_1 wt	88 HAAAT**LDD**GTVDTLTG	134 FSSFASA**F**GA**P**GLG**GY**
KR_1-PQS	88 HAAAT**PQS**GTVDTLTG	134 FSSFASA**F**GA**P**GLG**GY**
KR_1-WGG	88 HAAAT**LDD**GTVDTLTG	134 FSSFASAW**G**A**G**GLG**GY**
KR_1-5M	88 HAAAT**PQS**GTVDTLTG	134 FSSFASAW**G**A**G**GLG**GY**
B-type	88 HTAGh**LDD**GxhxxhTP	134 FSSAAG**x**h**G**x**P**GQG**NY**
KR_2 wt	88 HAAGL**PQQ**VAINDMDE	134 FSSGAGVW**G**S**A**RQG**AY**
KR_2-LDD	88 HAAGL**LDD**VAINDMDE	134 FSSGAGVW**G**S**A**RQG**AY**
KR_2-LPN	88 HAAGL**PQQ**VAINDMDE	134 FSSGAGV**L**G**S**P**R**QG**NY**
KR_2-6M	88 HAAGL**LDD**VAINDMDE	134 FSSGAGV**L**G**S**P**R**QG**NY**
A-type	88 HAAGh**xxxx**Phxxxx**E**	134 FSSxAGh**W**G**S**GxQG**x**Y

these motifs. In an accompanying paper [24] (this issue of *Chemistry & Biology*), we demonstrate that the methods of molecular evolution can be applied to identify numerous active site mutants of both $eryKR_1$ and $eryKR_2$ that are active against surrogate substrates, but which have altered stereocontrol. Making such tailored mutations in KR domains within intact PKS multienzymes would offer a new route to engineered polyketide drugs with altered chirality.

Results and Discussion

Homology Modeling of PKS Ketoreductase Domains Predicts that Key Amino Acids Proposed to Be Involved in Stereoselection Are Located at the Active Site

As previously reported [22], searches of publicly available databases readily show that the ketoreductase domains of modular PKS multienzymes are members of the short-chain NAD(P)H-dependent dehydrogenase/reductase (SDR) enzyme family. These enzymes are characterized by a well-defined Rossmann fold and by key tyrosine, serine, lysine, and asparagine residues at the active site [22, 25]. Homology models of both $eryKR_1$ and $eryKR_2$ were constructed based on the SDR enzyme porcine carbonyl reductase as a template (28% sequence identity), as described in Experimental Procedures. The overall fold predicted for $eryKR_1$ and $eryKR_2$ was essentially identical, and served to reveal key conserved features of their predicted active sites (Figure 3). The position of the bound NADPH cofactor could also be modeled with confidence. The carboxyamide group of the nicotinamide ring of the cofactor is anchored by hydrogen bonding to the side chain of S182 in $eryKR_1$ (S181 in $eryKR_2$). This directs the A face of the nicotinamide ring toward the putative substrate binding site, consistent with the observed specificity of both $eryKR_1$ and $eryKR_2$ for hydride transfer of the pro-S hydrogen from the C-4′ of the reduced nicotinamide [26]. The position of the critical residues Y149 and S136 (Caffrey numbering [23]), which is thought to assist hydride transfer by polarization of the keto group of the substrate, is essentially identical in both $eryKR_1$ and $eryKR_2$, consistent with the idea that the observed differences in the stereochemistry are determined by different modes of substrate binding to a conserved catalytic apparatus [23]. The amino acid sequence motif I [23], containing the conserved motif LDD at residues 93–95 and also

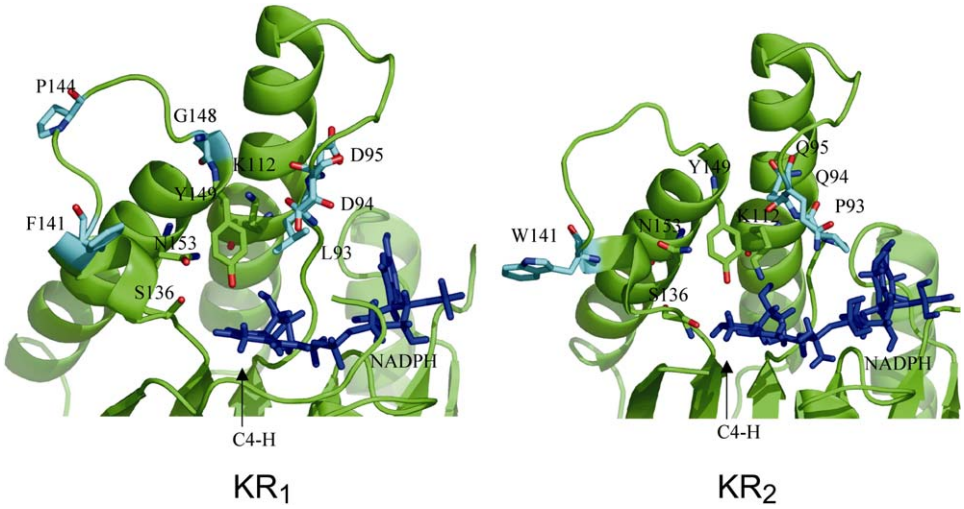

Figure 3. Homology Models of eryKR$_1$ and eryKR$_2$ Indicating the Location of Conserved Amino Acids Proposed to Influence the Stereospecificity of Reduction

In eryKR$_1$, the active site residue Y149 positions the substrate carbonyl for hydride transfer from NADPH (blue). Motif residues are shown in green as follows: F141, P144, and G148 are part of a loop directly adjacent to catalytic residue Y149, while LDD93–95 belong in a loop adjacent to the active site. In eryKR$_2$, W141 and 93–95PQQ are the counterparts of F141 and 93–95LDD in eryKR$_1$. The numbering of key residues has been done to facilitate comparison with Figure 2, based on Caffrey [23]. Because of single amino acid insertions in the sequences of eryKR$_1$ and eryKR$_2$, the true active site residue numbers derived from the model differ very slightly from those shown, and they are given in the Supplemental Data. The coordinates of the homology models are available upon request from the authors.

noted by Reid et al. [22], is located in a loop region adjacent to the active site Y149 (Figure 3) and to the nicotinamide ring of NADPH. The sequence motif II previously described [23] and including (for eryKR$_1$) residues F141, P144, and G148 lies in a second loop directly upstream of the active site Y149 (Figure 3). This region apparently has no definite secondary structure, but it is flanked by predicted α helices. In the absence of a crystal structure for a PKS KR domain, these models provide useful support for the hypothesis that directed mutagenesis of the motif residues could directly affect stereocontrol by modulating the binding of substrates. Recently, the crystal structure of the discrete KR involved in biosynthesis of the aromatic polyketide actinorhodin has been determined [27, 28]. At least for the analogous KR from the enterocin pathway, it appears that ketoreduction may take place during polyketide chain elongation [29]. These enzymes also belong to the SDR family [23]. Although exact details of substrate-enzyme binding have proved elusive, inspection of the active site architecture also supports the conclusion, from modeling of the eryKR domains, that motifs I and II are at or near the active site in such enzymes.

Mutagenesis, Expression, and Relative Ketoreductase Activities of eryKR$_1$ and eryKR$_2$ Domains

As shown in Figure 2, eryKR$_1$ has LDD at residues 93–95, as in the motif I consensus for B-type KR domains. In motif II, the domain lacks W141 and has P144 as expected [23], but it has G at 148 rather than the consensus N148. To switch the amino acid residues of the motif to those typical of A-side KRs, mutant KR$_1$-PQS was designed to alter all of residues 93–95, and mutant KR$_1$-WGG was designed to replace F141 and P144 by W141 and G144 respectively. In a third mutant, KR$_1$-5M, all five amino acid changes were introduced. Examination of the eryKR$_2$ sequence indicates the absence of LDD at residues 93–95 and the presence of W141, as expected for A-type KR domains. In mutant KR$_2$-LDD, residues 93–95 were altered to LDD, and, in an additional mutant (KR$_2$-LPN), all of W141, A144, and A148 were replaced by L141, P144, and N148, respectively. Finally, in mutant KR$_2$-6M, all six of these amino acid substitutions were combined. Individual mutant KR domains were expressed as glutathione-*S*-transferase (GST) fusion proteins in *E. coli*, and were purified as previously described for the wild-type domains [19]. The fusion proteins were used directly for additional experiments, since we have previously shown [19] that the presence of the GST domain does not affect the kinetics or stereochemistry of KR-mediated reduction.

In order to assess the effect of the amino acid substitutions on overall KR activity, the rate of consumption of NADPH was measured at high (presumed to be saturating) [19] concentrations of both **1** (30 mM) and NADPH (1 mM) and was compared to the activity of wild-type KR domains (Figure 4). These data show that substitutions in eryKR$_1$ result in an ~5-fold decrease in overall activity, to a level closer to that of wild-type eryKR$_2$. In contrast, all of the eryKR$_2$ mutants have a slightly higher activity than the wild-type domain.

Site-Specific Mutations in Motifs I and II Alter Stereocontrol by Isolated KR Domains

We next determined the effects of these mutations on the stereochemical outcome of ketoreduction by separating **2**–**5** from **1** by using chiral HPLC (Figures 5 and 6). Under these conditions, the wild-type eryKR$_1$ domain produced almost exclusively (>98%) the (2*S*, 3*R*) isomer **2** [19] (Figure 5A). Surprisingly, the double mutant altered in motif II (KR$_1$-WGG), although it also showed

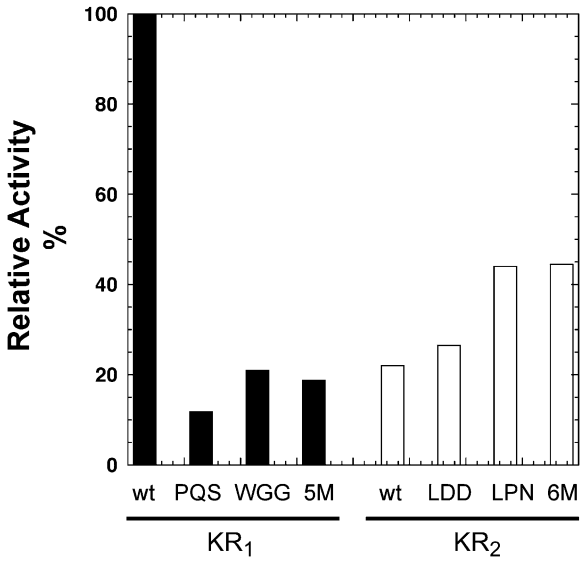

Figure 4. Ketoreductase Activity of Recombinant KR Domains

Mutant and wild-type KR domains, purified as their GST fusion proteins, were assayed for NADPH-dependent ketoreductase activity with (2*R*, *S*)-2-methyl-3-oxopentanoic acid NAC ester (IUPAC name: *S*-2-acetamidoethyl 3-hydroxy-2-methyl-pentanethioate) **1** as substrate.

specificity for the (2*S*) substrate, reduced it almost exclusively to the (2*S*, 3*S*) stereoisomer **3**, in which the configuration at C-3 is switched (Figure 5B), together with very small amounts of both the (2*S*, 3*R*) and the (2*R*, 3*R*) isomers **2** and **4**. The KR_1 containing the triple mutation in motif I (KR_1-PQS) also favored selection of the correct stereoisomer of the substrate, but it produced a 1:1 mixture of **2** and **3**, again with a small amount of **4** (Figure 5C), showing that, in this case, reduction occurred without stereochemical preference, from either face of the keto group. The mutant with all five substitutions (KR_1-5M) also showed switched specificity in comparison to the wild-type (Figure 5D), yielding almost exclusively the (2*S*, 3*S*) stereoisomer **3**. None of the eryKR$_1$ mutants in this study showed altered stereoselectivity since they all preferentially selected the *S* isomer of **1**.

Detailed interpretation of these striking results in terms of specific enzyme-substrate interactions must await the determination of the crystal structure of wild-type and mutant KR domains in the presence of a suitable substrate analog. However, these results are consistent with the previously postulated role of conserved active site residues in determining the stereochemistry of reduction, as evidenced by sequence alignments of KR domains [22, 23]. They support the proposal that the energetic differences between alternative modes of presentation of 3-oxoacyl-NAC substrates (for example **1**) to KR domains are rather small [19]. They also demonstrate the feasibility of using site-specific mutation to engineer a precise change in alcohol product stereochemistry.

The consequences of mutation of eryKR$_2$ on the stereocontrol of ketoreduction were less clear-cut (Figure 6). The wild-type domain under these conditions produces three products, but the (2*S*, 3*S*) isomer **3** and the (2*S*, 3*R*) isomer **2** are the predominant forms (Figure 6A). The predicted (2*R*, 3*S*) alcohol **5** represents only 6%–15% of the product mixture [19, 24] (Figure 6A). Presumably, this lack of expected stereocontrol is a reflection of the fact that the normal substrate for this enzyme is a triketide rather than a diketide. Unfortunately, the use of chemically synthesized triketide thioesters as substrates is frustrated by their spontaneous and rapid cyclization to the corresponding lactone. Three mutants of eryKR$_2$ were expressed and assayed. The eryKR$_2$ domain containing the triple mutation in motif I (KR_2-LDD) produced only two products, an ~1:1 mixture of **3** and **5** (Figure 6B). A similar product mixture was also produced by the triple mutant in motif II (KR_2-LPN) (data not shown) and by the KR_2 with all six substitutions (KR_2-6M) (Figure 6C). These results must be considered together with the data of Figure 4, showing that the mutant enzymes are apparently more efficient than wild-type in the NADPH-linked assay. Rather than prompting the expected switch in the configuration at C-3 of the alcohol product, these mutations actually increase the proportion of the wild-type (2*R*, 3*S*) stereoisomer. Again, in the absence of structural information, no definite interpretation can yet be given for these results. One possible explanation is that the altered active site disfavors competition from the (2*S*) enantiomer of substrate **1**, and/or from nonproductive binding modes of the (2*R*) enantiomer. However, as pointed out in the accompanying paper [24], the same alterations in the product ratio are seen in mutants of eryKR$_2$ where bulky residues (P93 or W141 or both) are replaced by smaller ones, which would seem to indicate a more open active site in such mutants. Whatever the detailed explanation, the present results confirm the importance of amino acids in motifs I and II in binding substrate at the active site of KR domains. They also demonstrate for the first time that mutagenesis of only a few residues can shift the substrate specificity of a PKS KR with respect to the configuration at the methyl-branched (C-2) position of the substrate, as well as the stereochemistry at C-3.

Several models have been proposed to explain the stereospecificity of ketoreductases in type I PKS. One is that the position and orientation of the active site Y149, which activates the keto group via coordination of the carbonyl oxygen, is different in A-type and B-type KR domains, resulting in a different orientation of reduction [25]. This hypothesis is consistent with the results of our mutations in the 141–148 (motif II) loop of KR, which is directly adjacent to the active site tyrosine. Substitutions in this loop of eryKR$_1$ resulted in switched specificity, whereas substitutions in the same loop of eryKR$_2$ resulted in the apparent tightening of the active site. Both results are consistent with a slight movement of the active site tyrosine as a result of having to accommodate the amino acid substitutions. This model, however, does not easily explain the effect of substitutions in the LDD motif I of eryKR$_1$, as these amino acids do not seem to have any direct contact with the active site tyrosine in our homology models of KR.

Alternatively, it has been proposed that the active site architecture is invariant, with respect to the positioning of key residues and the face of bound NADPH, which is exposed to substrate, and this hypothesis is strongly supported by our homology modeling experiments and those of Reid et al. [22]. The differences in

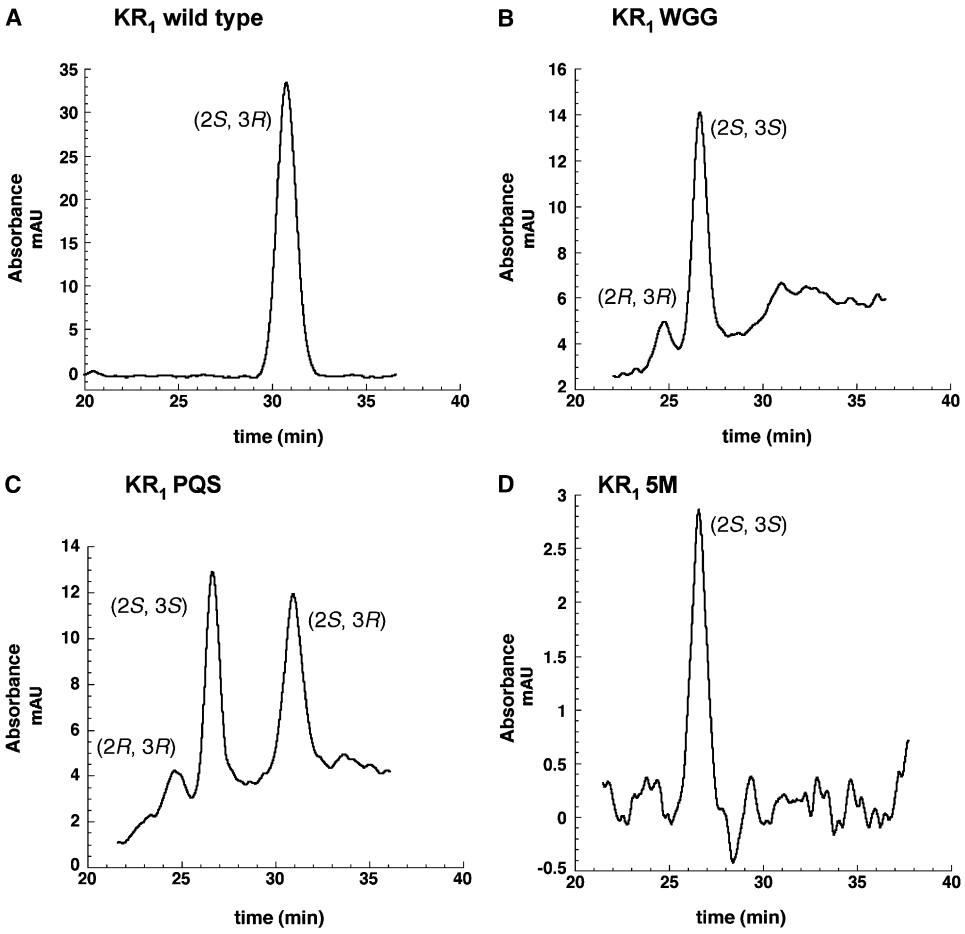

Figure 5. Stereochemical Outcome of Reduction by Wild-Type and Mutant eryKR$_1$ Domains

(A–D) The products of ketoreduction of (2*R*, *S*)-2-methyl-3-oxopentanoic acid NAC ester 1 were separated on a Chiracel OD chiral HPLC column after incubation with (A) wild-type eryKR$_1$, (B) KR$_1$-WGG, (C) KR$_1$-PQS, or (D) KR$_1$-WGG/PQS (also referred to in the text as KR$_1$-5M).

stereochemical outcome between A-type and B-type KR domains would then arise because the entering substrate has available two alternative binding modes, related by a 180º rotation about the axis of the target carbonyl group, and the residues of motifs I and II affect the energetic balance between the two. We currently favor this model, which also accounts for the range of kinetic and stereochemical behavior seen with diketide substrates on different recombinant KR domains [19].

Configurational Switching in Polyketide Biosynthesis

The stereochemistry of ketoreduction in polyketide biosynthesis is tightly controlled in nature by KR domains. Previous efforts to harness the intrinsic stereospecificity of KR domains have involved the replacement of an entire KR domain by another one of opposite stereochemistry within a PKS multienzyme ([14] and J.L. Kellenberger and P.F.L., unpublished data). Although some of these chimeric constructs generated the product with altered alcohol (C-3) stereochemistry, the yields of the desired novel compound were relatively low. One plausible explanation for the mixed success of domain-swap (or module-swap) strategies is that swapping a KR domain for another from a different natural modular PKS not only could alter the stereochemistry at the site of reduction, but might also reduce the overall activity of the PKS, through the introduction of unfavorable protein:protein interactions. Reid et al. have previously demonstrated that ablation of KR activity in extension module 6 of the erythromycin-producing PKS by site-specific change of the essential tyrosine residue gave a slightly better yield of macrocyclic oxolactone product, both in vivo and in vitro, than when the engineering was done by a domain swap [22]. In this paper, we have identified several amino acids in eryKR$_1$ that, when changed simultaneously, allow it to retain its activity but actually switch its intrinsic stereospecificity. Previously, two groups have reported independently that site-directed mutagenesis of PKS AT domains led to a relaxation of extender unit specificity, falling short of a complete switch [30, 31]. If a PKS multienzyme with such KR active site substitutions generated novel compounds efficiently and with specifically altered configuration, it would offer an attractive additional route to engineered polyketides. Clearly, to make the best use of the strategy, we need to know far more about how changes in nonessential KR active site residues affect

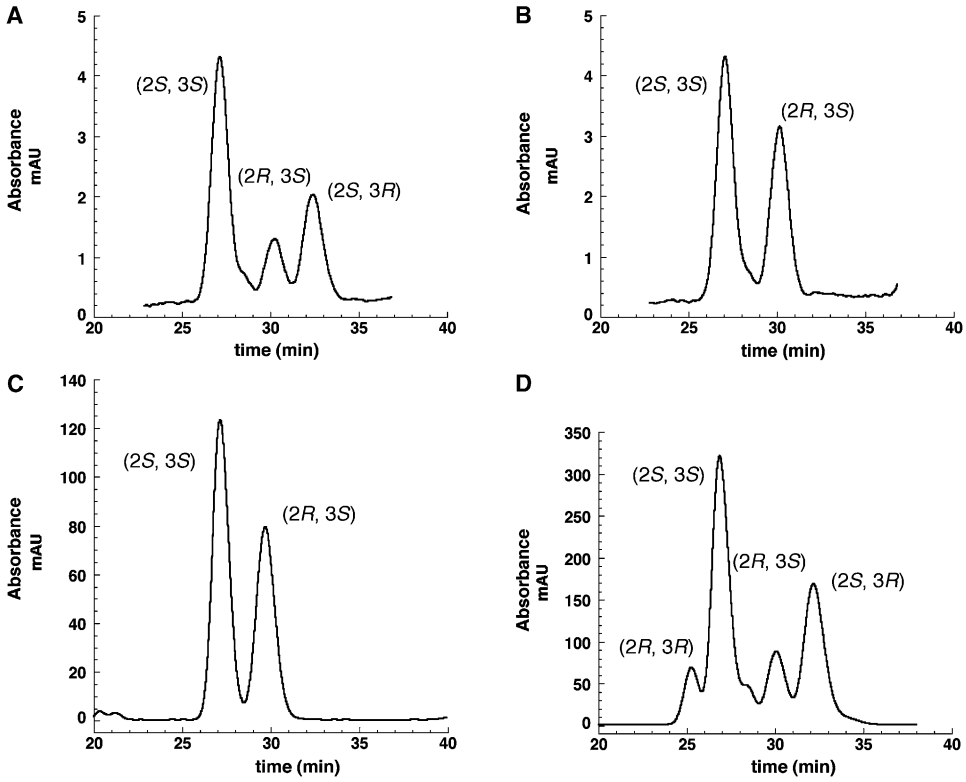

Figure 6. Stereochemical Outcome of Reduction by Wild-Type and Mutant eryKR$_2$ Domains

(A–C) The products of ketoreduction of (2*R*, *S*)-2-methyl-3-oxopentanoic acid NAC ester 1 were separated on a Chiracel OD chiral HPLC column after incubation with (A) wild-type eryKR$_2$, (B) KR$_2$-LDD, or (C) KR$_2$-LDD/LPN (also referred to in the text as KR$_2$-6M).

(D) The separation of a mixture of synthetic standards is also shown.

activity and stereoselectivity. In the accompanying paper [24], we demonstrate how high-throughput mutagenesis of isolated KR domains can be used rapidly to explore the effects of replacing such residues, and to develop a more detailed model for stereochemical control in these enzymes.

Significance

Attempts to alter the configuration at specific stereocenters in polyketide natural products by engineering the biosynthetic genes have been hampered by a limited understanding of the fundamental mechanisms by which modular polyketide synthases (PKSs) exert stereocontrol over polyketide assembly. Previous work has indicated that ketoreductase (KR) domains exercise control over the configuration of the alcohol product by favoring one of two alternative modes of binding of the 3-oxoacyl substrate, and empirical correlations have been detected between the presence of certain amino acid sequence motifs in natural PKS KR domains and the stereoselectivity of the reduction. Here, we have used homology modeling to confirm that these sequence motifs are at the KR active site, where they might interact directly with bound substrate. In one case, the substitution of two of these amino acid residues gave a mutant enzyme with reduced activity, but with a switch in the configuration of the alcohol product. Optimization of this approach, and its extension to KR domains within intact PKS multienzymes, could lead to more reliable and efficient engineering of polyketide stereochemistry, with potentially wide-ranging benefits for rationally engineered biosynthesis of polyketide therapeutics.

Experimental Procedures

Materials

Routine cloning and transformation procedures were as previously described for *E. coli* [32]. Electrocompetent cells of the *E. coli* DH10B strain were made as described previously [33]. All antibiotics were bought from Sigma. PCR reactions were performed on a programmable Robo Cycler Gradient 96 (Stratagene). Automated DNA sequencing was carried out on double-stranded DNA templates by using an automated ABI Prism 3700 DNA Analyzer (Applied Biosystems). LB medium contained 10 g tryptone (Duchefa), 5 g yeast extract (Duchefa), and 10 g NaCl in 1 liter of water.

Cloning Procedures

The plasmid used as a template for site-directed mutagenesis of eryKR$_1$ was pAPS28 [19], which contains the DNA encoding the eryKR$_1$ domain (plus flanking regions) cloned into the expression vector pGEX-4T-3 (Stratagene) to create a C-terminal fusion protein with glutathione-*S*-transferase (GST) [19]. The corresponding plasmid used as a template for mutagenesis of eryKR$_2$ was pVS1, which likewise contains the DNA encoding the eryKR$_2$ domain (plus flanking regions) cloned into the expression vector pGEX-4T-3 (Stratagene) [19]. Mutagenesis was carried out by overlap PCR by using Pfu DNA polymerase (Stratagene) and specific mutagenic oligonucleotides in two separate PCR reactions, as detailed in Table S1 (see the Supplemental Data available with this article online). The PCR products from reactions 1 and 2 were mixed together and

used as a template for a third PCR reaction by using the appropriate flanking primers. The resulting fragment was in each case a full-length KR domain that contained the desired mutations. These fragments were digested with BamHI and EcoRI and cloned into expression plasmid pGEX-4T-3, previously cut with the same enzymes. The KR_1 mutant with five substitutions across both motifs I and II, and the KR_2 mutant with six substitutions across motifs I and II (Figure 3), were made by substituting the StyI/EcoRI fragment from eryKR$_1$-WGG and eryKR$_2$-LPN into the corresponding unique sites of the eryKR$_1$-PQS and eryKR$_2$-LDD, respectively, resulting in the eryKR$_1$-PQS-WGG (eryKR$_1$-5M) and eryKR$_2$-LDD-LPN (eryKR$_2$-6M) housed in the pGEX expression vector. All constructs were verified by DNA sequencing.

Protein Expression

One liter of culture in LB medium of *E. coli*-BL21(DE3)-CodonPlus-RP (Stratagene) harboring the recombinant plasmids was grown at 30ºC to an OD_{600} of 0.4, under carbenicillin (50 μg/ml) and chloramphenicol (34 μg/ml) selection. The cultures were cooled to 22ºC and, upon reaching an A_{600} of 0.6, induced with 0.1 mM IPTG (Melford Laboratories); the cell growth continued for 16 hr. Cells were harvested by centrifugation, and cell pellets were stored at −20ºC.

Protein Purification

Typically, frozen cell pellets (about 3 g wet weight) were thawed and resuspended in 20 ml lysis buffer (100 mM NaH_2PO_4 [pH 8.2], 150 mM NaCl, 1 mM dithiothreitol, 10% glycerol) containing 1 tablet of protease inhibitors (Complete Roche, Indianapolis, IN) per 25 ml, RNase (25 mg/l, Sigma), DNase (100 mg/l, Sigma), and lysozyme (1 mg/ml, Sigma). The cell suspensions were incubated on ice for 40 min and disrupted by passage through a French pressure cell (SLM Instruments, Urbana, IL) operated at 1000 psi, and the insoluble material was collected by centrifugation. The resulting supernatant was applied to a 50% suspension of glutathione-agarose beads (Sigma), prepared according to the manufacturer's instructions, and incubated for 45 min at 4ºC. The beads were washed several times by resuspending in 10 volumes of 25 mM HEPES-KOH buffer (pH 7.5) containing 150 mM NaCl, and eluted with the same buffer containing 20 mM reduced glutathione. The elution step was repeated three times, and the fractions collected. The combined protein-containing fractions were concentrated in a spin concentrator to a final volume of 2.5 ml, and the glutathione was removed by using a PD10 (Pharmacia) column equilibrated with 25 mM HEPES-KOH buffer (pH 7.5), containing 150 mM NaCl. Typical yields for all proteins were 6–10 mg/l of culture, as measured by the bicinchoninic acid (BCA) assay (Pierce, Rockford, IL).

Kinetic Assays

NADPH-linked assays were carried out at 30ºC in the chamber of a Spectramax Plus 96-well microplate reader equipped with Softmax Pro data acquisition software (Molecular Devices, Sunnyvale, CA). The reactions were followed by monitoring the change in absorbance at 340 nm at 30 s intervals for 7 min in a UV transparent 96-well plate (Greiner Bio-One, Kremsmuenster, Austria). Protein samples (0.1–0.3 mg/ml) were incubated with saturating amounts of cofactor and substrate (1.0 mM NADPH, 30 mM (2*R*, *S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine thioester **1**) in 400 mM potassium phosphate buffer (pH 7.5) at 30ºC. Protein concentration was determined after the assay by the BCA method. The total assay volume was 200 μl. The enzyme was preincubated with NADPH for 20 min, the substrate was added, and the reaction was mixed by shaking for 10 s before taking the first reading, and for 3 s prior to each subsequent reading. All readings were corrected for the change in absorbance at 340 nm in the absence of enzyme, and each assay was repeated in triplicate.

Product Separation by Chiral HPLC

Reduction of (2*R*, *S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine thioester **1** was carried out in a total volume of 200 μl containing 10 mM NADPH, 15 mM **1**, 200 mM NaH_2PO_4 (pH 7.5), and either wild-type (0.1 mg/ml, 2 μM) or mutant (1.0 mg/ml, 20 μM) KR enzymes. Reactions were incubated at 22ºC for 16 hr and extracted with 3 ml ethyl acetate. Extracts were evaporated and resuspended in 50 μl isopropanol. The reaction products were separated with a ChiraCel OC column (Daicel, 25 cm × 4.6 mm), connected to an Agilent HPLC system at a flow rate of 0.8 ml/min in an isocratic mobile phase of 93% isohexane/7% ethanol at 40ºC for 2 hr. The unreacted **1** still present served as an internal standard for the calibration of traces. Control samples were run with each batch of analyses to determine the retention times of the four diastereoisomeric products **2**–**5** (Figure 6D). Additionally, some samples were spiked with one of the standards to aid identification.

Homology Modeling of the KR_1 and KR_2 Structures

The sequences of erythromycin PKS ketoreductase 1 and 2 (eryKR$_1$ and eryKR$_2$) were compared with all protein sequences deposited in SWISS-PROT, by using the Blast v2.2 program [34]. A PSI-BLAST search in each case produced an alignment between the eryKR domain and homologous proteins, which highlighted conserved residues in the short-chain dehydrogenase/reductase (SDR) family to which the eryKR domains were found to belong, as previously reported [22]. Homologous proteins with known structure were identified by using the FUGUE [35] homology recognition server. FUGUE searches for homologs in the structural profile library derived from the structure-based alignments in the HOMSTRAD database [35] and uses the environment-specific substitution tables to generate, automatically, the best alignments for the top hits. The homolog found with the highest percentage of identity (28%) to both eryKR domains was porcine carbonyl reductase (pdb code: 1N5D). The alignments produced by FUGUE for this highest-scoring hit were formatted with JOY [36] and analyzed visually to highlight the conservation of structurally and functionally important residues. The alignment was further improved by using SEAVIEW [37]. The models of eryKR$_1$ and eryKR$_2$ were constructed with MODELLER [38] by using porcine carbonyl reductase as the template. The models were improved by using the loop-building programs RAPPER [39] and CODA [40]. Final energy and structure minimization were performed by using the SYBYL (Tripos, Inc., St. Louis, MO, USA) force field. Finally, the models were validated by PROCHECK [41], VERIFY3D [42], and by visual inspection with 3D graphics software. The results from all of these programs revealed that the models did not require further modification.

Supplemental Data

Supplemental Data including the methods used to construct the homology models and the sequences of oligonucleotide primers used for mutagenesis are available at http://www.chembiol.com/cgi/content/full/13/3/277/DC1/.

Acknowledgments

We thank Professor Sir Tom Blundell and Dr. Kira Weissman for helpful discussions, and Drs. Ben Luisi and Fabien Walas for help with the Spectramax plate recorder. We are grateful for the support of this work by project grants to P.F.L. and J.S. from the United Kingdom Biotechnology and Biological Sciences Research Council (BBSRC) and from the BBSRC to J.B.S. and Professor Blundell (for B.P). A.B.-O. is a Royal Society USA Postdoctoral Research Fellow. D.S. thanks the Deutschen Akademie der Naturforscher Leopoldina for a postdoctoral fellowship (BMBF-LPD 9901/8-90).

Received: September 26, 2005
Revised: January 10, 2006
Accepted: January 10, 2006
Published online: March 24, 2006

References

1. Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.
2. Walsh, C.T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science *303*, 1805–1810.
3. Weissman, K.J. (2004). Polyketide biosynthesis: understanding and exploiting modularity. Philos. Trans. R. Soc. Lond. A *362*, 2671–2690.
4. Walsh, C.T. (2002). Combinatorial biosynthesis of antibiotics: challenges and opportunities. ChemBioChem *3*, 125–134.

5. Khosla, C., and Keasling, J.D. (2003). Metabolic engineering for drug discovery and development. Nat. Rev. Drug Discov. *2*, 1019–1025.
6. Weissman, K.J., and Leadlay, P.F. (2005). Combinatorial biosynthesis of reduced polyketides. Nat. Rev. Microbiol. *3*, 925–936.
7. Mendez, C., and Salas, J.A. (2001). Altering the glycosylation pattern of bioactive compounds. Trends Biotechnol. *19*, 449–456.
8. Celmer, W.D. (1965). Biogenetic, constitutional, and stereochemical unitary principles in macrolide antibiotics. Antimicrob. Agents Chemother. *5*, 144–156.
9. Marsden, A.F.A., Caffrey, P., Aparicio, J.F., Loughran, M.S., Staunton, J., and Leadlay, P.F. (1994). Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. Science *263*, 378–380.
10. Wiesmann, K.E.H., Cortés, J., Brown, M.J.B., Cutter, A.L., Staunton, J., and Leadlay, P.F. (1995). Polyketide synthesis *in vitro* on a modular polyketide synthase. Chem. Biol. *2*, 583–589.
11. 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.
12. Böhm, I., Holzbaur, I.E., Hanefeld, U., Cortés, J., Staunton, J., and Leadlay, P.F. (1998). Engineering of a minimal modular polyketide synthase, and targeted alteration of the stereospecificity of polyketide chain extension. Chem. Biol. *5*, 407–412.
13. Holzbaur, I.E., Harris, R.C., Bycroft, J., Cortés, J., Bisang, C., Staunton, J., Rudd, B.A.M., 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.
14. 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.
15. Holzbaur, I.E., Ranganathan, A., Thomas, I.P., Kearney, D.J.A., Reather, J., Rudd, B.A.M., 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.
16. Joshi, A.K., and Smith, S. (1993). Construction of a cDNA encoding the multifunctional animal fatty acid synthase and expression in *Spodoptera frugiperda* cells using baculoviral vectors. Biochem. J. *296*, 143–149.
17. Rangan, V.S., Joshi, A.K., and Smith, S. (1998). Mapping the functional topology of the animal fatty acid synthase by mutant complementation *in vitro*. J. Biol. Chem. *273*, 34949–34953.
18. Østergaard, L.H., Kellenberger, L., Cortés, J., Roddis, M.P., Deacon, M., Staunton, J., and Leadlay, P.F. (2002). Stereochemistry of catalysis by the ketoreductase activity in the first extension module of the erythromycin polyketide synthase. Biochemistry *41*, 2719–2726.
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., et al. (2005). Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. Chem. Biol. *12*, 1145–1153.
20. Nakajima, K., Yamashita, A., Akama, H., Nakatsu, T., Kato, H., Hashimoto, T., Oda, J., and Yamada, Y. (1998). Crystal structures of two tropinone reductases: different reaction stereospecificities in the same protein fold. Proc. Natl. Acad. Sci. USA *95*, 4876–4881.
21. Nakajima, K., Kato, H., Oda, J., Yamada, Y., and Hashimoto, T. (1999). Site-directed mutagenesis of putative substrate-binding residues reveals a mechanism controlling the different stereospecificities of two tropinone reductases. J. Biol. Chem. *274*, 16563–16568.
22. 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.
23. Caffrey, P. (2003). Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. ChemBioChem *4*, 649–662.
24. O'Hare, H.M., Baerga-Ortiz, A., Popovic, B., Spencer, J.B., and Leadlay, P.F. (2005). High-throughput mutagenesis of active site residues to evaluate models of stereochemical control in ketoreductase domains from the erythromycin polyketide synthase. Chem. Biol. *13*, this issue, 287–296.
25. Kallberg, Y., Oppermann, U., Jornvall, H., and Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). Eur. J. Biochem. *269*, 4409–4417.
26. McPherson, M., Khosla, C., and Cane, D.E. (1998). Erythromycin biosynthesis: the β-ketoreductase domains catalyze the stereospecific transfer of the 4-*pro*-*S* hydride of NADPH. J. Am. Chem. Soc. *120*, 3267–3268.
27. 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*, 1865–1875.
28. 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*, 14529–14538.
29. Hertweck, C., Xiang, L., Kalaitzis, J.A., Cheng, Q., Palzer, M., and Moore, B.S. (2004). Context-dependent behavior of the enterocin iterative polyketide synthase; a new model for ketoreduction. Chem. Biol. *11*, 461–468.
30. Reeves, C.D., Ashley, G.W., Piagentini, M., Hutchinson, C.R., and McDaniel, R. (2001). Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. Biochemistry *40*, 15464–15470.
31. Del Vecchio, F., Petkovic, H., Kendrew, S.G., Low, L., Wilkinson, B., Lill, R.E., Cortés, J., Rudd, B.A.M., Staunton, J., and Leadlay, P.F. (2003). Active-site residue, domain and module swaps in modular polyketide synthases. J. Ind. Microbiol. Biotechnol. *30*, 489–494.
32. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
33. Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1998). High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. *16*, 6127–6145.
34. Altschul, S.F., Madden, T.L., Schafer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
35. Shi, J., Blundell, T.L., and Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. *310*, 243–257.
36. Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S., and Overington, J.P. (1998). JOY: protein sequence-structure representation and analysis. Bioinformatics *14*, 617–623.
37. Galtier, N., Gouy, M., and Gautier, C. (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. *12*, 543–548.
38. Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.
39. de Bakker, P.I., DePristo, M.A., Burke, D.F., and Blundell, T.L. (2003). *Ab initio* construction of polypeptide fragments: accuracy of loop decoy discrimination by an all-atom statistical potential and the AMBER force field with the Generalized Born solvation model. Proteins *51*, 21–40.
40. Deane, C.M., and Blundell, T.L. (2001). CODA: a combined algorithm for predicting the structurally variable regions of protein models. Protein Sci. *10*, 599–612.
41. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
42. Luthy, R., Bowie, J.U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature *356*, 83–85.