Chemistry & Biology, Vol. 12, 1145–1153, October, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.chembiol.2005.08.017

Molecular Basis of Celmer's Rules: Stereochemistry of Catalysis by Isolated Ketoreductase Domains from Modular Polyketide Synthases

Alexandros P. Siskos,[1] Abel Baerga-Ortiz,[1,3] Shilpa Bali,[1,3] Viktor Stein,[1] Hassan Mamdani,[1] Dieter Spiteller,[2] Bojana Popovic,[1] Jonathan B. Spencer,[2] James Staunton,[2] Kira J. Weissman,[1] 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

A system is reported for the recombinant expression of individual ketoreductase (KR) domains from modular polyketide synthases (PKSs) and scrutiny of their intrinsic specificity and stereospecificity toward surrogate diketide substrates. The eryKR$_1$ and the tylKR$_1$ domains, derived from the first extension module of the erythromycin PKS and the tylosin PKS, respectively, both catalyzed reduction of (2*R, S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine thioester, with complete stereoselectivity and stereospecificity, even though the substrate is not tethered to an acyl carrier protein or an intact PKS multienzyme. In contrast, and to varying degrees, the isolated enzymes eryKR$_2$, eryKR$_5$, and eryKR$_6$ exercised poorer control over substrate selection and the stereochemical course of ketoreduction. These data, together with modeling of diketide binding to KR$_1$ and KR$_2$, demonstrate the fine energetic balance between alternative modes of presentation of ketoacylthioester substrates to KR active sites.

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 giant, multifunctional enzymes that catalyze the assembly line biosynthesis of structurally complex and clinically important polyketide natural products [1, 2]. 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. Numerous strategies have been successfully developed to engineer modular PKSs, and also downstream post-PKS enzymes, to produce chemically altered products for use in drug discovery [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.

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, normally from either malonyl-CoA or 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. For almost all modules, there is a convincing correlation between the predicted constituent domains and the chemical structure of the corresponding chain extension unit in the growing polyketide chain. Although noncanonical modular PKSs, in which the AT activity is carried on a discrete protein or in which the order of domains in the multienzyme is unusual [8–11], have recently been described, the mechanisms for assuring chemical and stereochemical control are believed to remain the same.

The methyl centers at C-2 and the hydroxyl centers at C-3 generated in many newly added chain extension units have either an *R* or *S* configuration. As early as 1965, Celmer noted that 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 [12]. The AT domains of the erythromycin PKS DEBS have all been shown to recruit the (2*S*), and not the (2*R*), isomer of methylmalonyl-CoA as the extension unit [13, 14]; thus, AT domains are not thought to play a primary role in the control of stereochemistry at methyl-branched centers in the polyketide product. This has focused attention on the ketosynthase (KS) and ketoreductase (KR) domains as potentially determining the outcome of polyketide chain extension that underlies Celmer's Rules.

In mechanistic terms, the chirality at the hydroxyl group must be determined by the direction of hydride addition to the keto group (*re* or *si* face) within the active site of the KR domain. It has recently been found to be possible to correlate motifs in the amino acid sequence of PKS KR domains with the stereochemical outcome of the reductions they catalyze [15, 16], suggesting that the direction of KR-catalyzed reduction may be "hard-wired."

*Correspondence: pfl10@mole.bio.cam.ac.uk
[3]These authors contributed equally to this work.

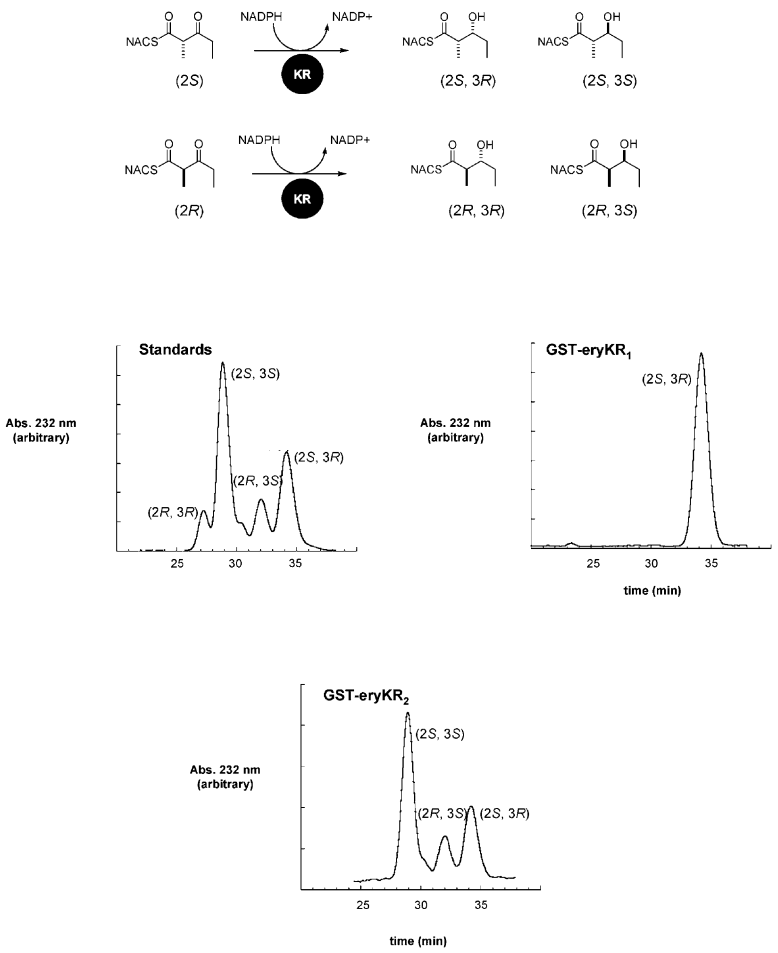


Figure 1. The Stereochemical Outcome of Reduction of 2-Methyl-3-Oxopentanoic Acid NAC Thioester by Erythromycin PKS Ketoreductases eryKR$_1$ and eryKR$_2$

Analysis by HPLC shows that GST-eryKR$_1$ produces a single product with the (2*S*, 3*R*) configuration. In contrast, GST-eryKR$_2$ yields three products.

A few examples of individual KR domains have already been successfully swapped between modules of different PKSs resulting in the predicted products, apparently indicating that the stereospecificity of reduction is dictated by the KR domain and is transferable ([17] and L. Kellenberger and P.F.L., unpublished data). For example, the KS domain in the first extension module of DEBS is capable of catalyzing the epimerization at the methyl center of the initially formed (2*R*) isomer of the keto ester to the (2*S*) isomer [18], and the eryKR$_1$ domain selects only the (2*S*) isomer from the mixture. In a chimaeric diketide synthase that combined KS$_1$ with the reductive enzyme corresponding to module 2 [19], a diketide was produced in vivo in which the stereochemistry was altered at both C-2 and C-3; this supports the idea that both enantiomeric keto esters are formed by KS$_1$ and that the eryKR$_2$ domain selects correctly between them. However, other evidence suggests that the true situation is more complex and that the direction of reduction can be influenced by the nature of the substrate [20]. For example, when KS$_1$ was engineered in place of eryKS$_5$ in a triketide lactone synthase model system based on DEBS 3, giving rise to both the (2*R*) and (2*S*) isomers of the diketide intermediate, the KR domains (eryKR$_5$ and eryKR$_6$) carried out reduction at both the *re* and *si* faces of their substrates [20].

In vitro studies on purified enzymes have been undertaken in an attempt to clarify some of these issues, and to uncover a structural basis for the observed stereocontrol by KR domains. Studies have been made of the effects of mutations in eryKR$_6$ on the rate of extension of a diketide thioester substrate by recombinant DEBS module 6-thioesterase (TE) expressed in *E. coli* [15], and to measure the kinetics of chain extension on diketide synthases containing only DEBS module 2 [19] or DEBS module 1 [21]. Unfortunately, in such experiments, only the overall kinetic behavior of a multienzyme in which stereocontrol is potentially exercized at more than one step is measured. We have previously shown that the KR activities can be specifically assayed (as for fatty acid synthases [22, 23]) when expressed as part of a PKS [20, 21, 24], by using (2*R, S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine (NAC) thioester (Figure 1). In these experiments, eryKR$_1$ was found to be both stereoselective and stereospecific, as it was able to select the expected (2*S*) stereoisomer out of a racemic mixture and make the expected (2*S*, 3*R*) product. The KR$_2$ was less active toward the racemic substrate, but it was still able to make the expected (2*R*, 3*S*) isomer. However, when the same model substrate was reduced in vitro by purified DEBS 3 (which contains two KR domains, eryKR$_5$ and eryKR$_6$, both producing the same stereochemical outcomes as eryKR$_2$ during normal chain extension), comparable amounts of all four possible stereoisomers of the product alcohol were reported to be obtained [20], implying a loss of

stereoselectivity and specificity. These experiments have raised interesting (and as yet unresolved) questions as to how the (stereo)specificity of an individual KR domain is influenced by both its context within the PKS and the tethering of the substrate to an acyl carrier protein.

We describe here the expression and purification of individual KR domains of PKSs as soluble and enzymatically active proteins in recombinant *E. coli*, in the absence of other PKS components. This advance has allowed measurement of the kinetics and stereochemistry of ketoreduction of model substrates for eryKR$_1$, eryKR$_2$, eryKR$_5$, eryKR$_6$, and tylosin KR$_1$ (tylKR$_1$). We have also undertaken detailed molecular modeling (A.B.-O. et al., submitted) of representative KR domains and of the way in which a diketide substrate might be accommodated in the active site. Our results reveal that at least some isolated KR domains are intrinsically capable of showing a very high specificity and stereospecificity toward the untethered surrogate substrates used here; however, other KR domains are less discriminating. The energetic differences between alternative binding modes that lead to different chiral products are clearly rather small.

Results and Discussion

Expression of PKS Ketoreductase Domains

Individual KR domains were expressed as glutathione-*S*-transferase (GST) fusion proteins in *E. coli*. For eryKR$_1$, the fragment was designed to contain the KR domain (170 amino acid residues) and also the "linker" regions flanking the KR, consisting of an N-terminal region of some 225 amino acid residues up to the preceding AT domain, and a C-terminal region of some 75 amino acid residues up to the start of the ACP domain. The N terminus was selected to coincide with a major site for limited proteolysis of DEBS [25], and the C terminus was selected to coincide with a site used in previously successful KR engineering and domain swaps ([21]; L.E. Kellenberger and P.F.L., unpublished data). For eryKR$_2$, eryKR$_5$, eryKR$_6$, and tylKR$_1$ the respective N and C termini of the recombinant domain were selected by ClustalW sequence alignment [26] of their sequences with that of eryKR$_1$. The respective nucleic acid fragments were amplified from cloned PKS templates by using mutagenic PCR, which introduced flanking BamHI and EcoRI sites and a stop codon just upstream of the EcoRI site.

Initially, each fragment was cloned into the expression vector pGEX-4T-3 (Amersham), according to the manufacturer's instructions, from which the GST fusion protein was in each case expressed in *E. coli* as a soluble protein and was purified to homogeneity by affinity chromatography on glutathione-agarose. Thrombin was used to cleave eryKR$_1$, eryKR$_2$, eryKR$_5$, and eryKR$_6$ from their respective fusion proteins, and the KR$_1$ domain was purified in further chromatographic steps to apparent homogeneity. The purified eryKR$_1$ domain behaved as a monomer on gel filtration (expected M_r, 54,900; observed apparent M_r, 58,800) (data not shown). In contrast, release of the other KR domains from the respective N-terminal GST fusion protein was slow, and the use of increased concentrations of thrombin led to multiple products of nonspecific cleavage. Therefore, the coding regions for the KR domains were all recloned, as BamHI-EcoRI fragments, into the related expression plasmid pGEX-6P-1, in which the N-terminally fused GST is released by use of the highly specific rhinovirus 2C proteinase (PreScission Protease [Amersham]), rather than by thrombin. After expression and affinity purification of the fusion proteins as described in Experimental Procedures, each KR domain was efficiently released by using this proteinase and was purified to near homogeneity (>90%–95%, the balance being a small amount of residual GST contaminant) by subsequent ion-exchange chromatography. In each case, the KR domain was obtained as a single protein species of exactly the predicted molecular mass, as judged by electrospray MS analysis (data not shown).

Table 1. Kinetic Parameters for Reduction of (9*RS*)-*Trans*-1-Decalone by PKS Ketoreductases

Ketoreductase	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
eryKR$_1$ (DKS) [21]	0.53 (± 0.02)	1.2 (± 0.3)	442 (± 128)
eryKR$_1$ (DEBS1-TEΔKR2) [21]	0.44 (± 0.05)	1.10 (± 0.01)	400 (± 49)
GST-eryKR$_1$	0.26 (± 0.01)	2.2 (± 0.2)	118 (± 15)
eryKR$_1$	0.13 (± 0.01)	2.8 (± 0.4)	46 (± 10)
GST-eryKR$_2$	0.16 (± 0.01)	5 (± 1)	32 (± 8)
eryKR$_2$	0.16 (± 0.01)	5 (± 1)	32 (± 8)
GST-eryKR$_5$	0.12 (± 0.01)	12 (± 3)	10 (± 3)
eryKR$_5$	0.12 (± 0.01)	10 (± 3)	12 (± 4)
GST-eryKR$_6$	0.17 (± 0.01)	3.7 (± 0.8)	46 (± 12)
eryKR$_6$	0.28 (± 0.04)	1.5 (± 0.7)	187 (± 114)
eryKR$_5$ + eryKR$_6$ (DEBS 3)	0.13 (± 0.01)	2.8 (± 0.4)	46 (± 10)
GST-tylKR$_1$	0.18 (± 0.01)	1.4 (± 0.4)	128 (± 43)

All experiments were carried out in triplicate. Data are reported as mean ± SEM.

Kinetic Analysis of Individual KR Domains

Kinetic analysis was performed as detailed in Experimental Procedures to determine the steady-state kinetic constants for the NADPH-linked reduction of (9*R*, *S*)-*trans*-1-decalone [22], which we have previously shown to be a useful surrogate substrate for PKS KR domains [21]. The results are given in Table 1 and are compared with those obtained previously for KR$_1$ in the context of the recombinant DEBS-derived diketide synthase (DKS) [21] and the triketide synthase (DEBS1-TE [ΔKR2]) [21]. Assays were also conducted with the intact GST-KR fusion proteins. The recombinant KR domains were also assayed by using (2*R*, *S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine (NAC) thioester as substrate [21] (Table 2). All experiments were carried out in triplicate.

From the results in Tables 1 and 2, it is evident that recombinant KR domains derived from modular PKSs are capable of folding to enzymatically active enzymes in *E. coli*. Reassuringly, the purified recombinant eryKR$_1$ has values for k_{cat} and K_m (for decalone or for diketide NAC thioester) that are not dramatically different from those found previously for the eryKR$_1$ domain in the context of a modular PKS [21]. Further, the presence of an N-terminal GST domain did not alter the behavior of

Table 2. Steady-State Kinetic Parameters for Ketoreductase-Catalyzed Reduction of (2*R, S*)-2-Methyl-3-Ketopentanoic Acid, *N*-Acetylcysteamine Thioester

Ketoreductase	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} M^{-1})
GST-eryKR$_1$	0.26 ± 0.01	35 ± 4	7 ± 1
eryKR$_1$ (DKS) [21]	0.27 ± 0.01	13 ± 2	22 ± 4
GST-eryKR$_2$	0.03 ± 0.01	13 ± 3	2 ± 1
GST-eryKR$_5$	0.05 ± 0.01	116 ± 28[a]	0.4 ± 0.2
GST-eryKR$_6$	0.18 ± 0.02	144 ± 19[a]	1.3 ± 0.3
GST-tylKR$_1$	0.61 ± 0.04	13 ± 2	47 ± 10

All experiments were carried out in triplicate. Data are reported as mean ± SEM.
[a] Due to insolubility of the substrate at concentrations in excess of 150 mM, these data represent the lower limits for the K_m values.

any of the KR domains with decalone substrate. There is similarly good agreement between the values obtained for eryKR$_1$ with diketide NAC thioester substrate, whether measured as a GST fusion protein or as a domain within an engineered diketide synthase (DKS) [21].

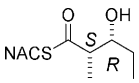
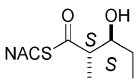
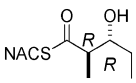
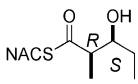
The specificity constants (k_{cat}/K_m) in Table 2 also show that the diketide NAC thioester is a better substrate for eryKR$_1$ and tylKR$_1$, which normally process diketide substrates, than for the other KRs; this difference is due to a substantial decrease in the rate of turnover (k_{cat}) by KR$_2$, and an increase in the K_m for KRs 5 and 6. In agreement with this, it has been previously shown that the purified DEBS 1-TE multienzyme (which contains both eryKR$_1$ and eryKR$_2$ domains) reduces the diketide NAC thioester largely via the eryKR$_1$ active site [24]. Overall, these kinetic data serve to reassure us that the recombinant enzymes are all active against plausible surrogate substrates.

Stereochemistry of Ketoreduction

We next wished to determine the stereochemical outcome of the reaction of the recombinant KR domains with (2*R, S*)-2-methyl-3-oxopentanoic acid *N*-acetylcysteamine (NAC) thioester as substrate. No attempt was made to present the enzymes with enantiomerically pure substrate, as spontaneous racemization under the assay conditions is expected to be rapid. Each enzyme is therefore presented with two alternative enantiomeric substrates, each of which can in principle be reduced from either face of the 3-keto group; thus, there are four possible stereoisomeric products. As indicated in Figure 1, these alcohol products can be readily separated by using chiral HPLC. The GST fusion proteins were used for these experiments, and the reaction products (normally after about 10%–20% of the initial keto substrate had been reduced), were separated and analyzed as described in Experimental Procedures. (In preliminary control experiments, in which the reaction was stopped after 2%, 10%, and 50% conversion, respectively, it was confirmed that the ratio of stereoisomers found was independent of the extent of reduction, as expected if the concentration ratio of 2*R*:2*S* substrates remains 1:1 throughout.) The results of these experiments are summarized in Table 3, in which the percentages found for the expected product (based on the apparent specificity of the KR domain in its native context) are indicated in bold type.

A striking feature of the data in Table 3 is the very different behavior of individual KR domains toward the diketide substrate. First, eryKR$_1$ selects only one enantiomer of the substrate and reduces it efficiently from only one face to give the anticipated (2*S*, 3*R*) product. Likewise, tylKR$_1$ selects the opposite enantiomer of the substrate and reduces it almost cleanly to the anticipated (2*R*, 3*R*) product. From these results alone, it is evident that for eryKR$_1$ and tylKR$_1$, tethering of the diketide keto substrate, via thioester linkage to the 4′-phosphopantetheinyl prosthetic group of an acyl carrier protein, is not a requirement for stereospecific and stereoselective reduction, nor is the context of the PKS essential for the integrity of this stereochemical outcome. If the tylosin PKS resembles DEBS [13, 14] and recruits only (2*S*)-methylmalonyl-CoA as extension units, and not the (2*R*) isomer, and if (as with DEBS [18]) the initial carbon-carbon bond formation on the ketosynthase proceeds with retention of configuration at C-2, then the initially formed diketide would be the (2*R*) isomer. Unlike the DEBS, the tylosin PKS therefore has no need of an epimerization step to convert this isomer into the (2*S*) isomer before ketoreduction can occur. Because of this, we initially considered it possible that the tylosin KR$_1$ might not in its usual context be confronted with the (2*S*) isomer, and that as an isolated domain, it might not be able to discriminate effectively against this substrate. The results in Table 3 do not support that view, since only 1% of the products arise from reduction of the (2*S*) isomer.

Table 3. Percentage of Each 2-Methyl-3-Hydroxy Stereoisomer Produced by Reduction of (2*R, S*)-2-Methyl-3-Oxopentanoic Acid by PKS Ketoreductases

	Percentage of Each Stereoisomer[a]			
Ketoreductase				
GST-eryKR$_1$	**100**	0	0	0
GST-eryKR$_2$	20 (± 3)	74 (±4)	0	**6 (± 1)**
GST-eryKR$_5$	2 (± 0.7)	74 (± 3)	0	**24 (± 1)**
GST-eryKR$_6$	5 (± 2)	70 (± 4)	0	**25 (± 2)**
eryKR$_5$ + eryKR$_6$ (DEBS 3)	4 (± 1)	54 (± 3)	0	**42 (± 2)**
GST-tylKR$_1$	1 (± 0.4)	0	**99 (± 1)**	0

[a] The stereoisomer predicted to be the unique product of each KR, based on its role in the PKS, is shown in bold type.

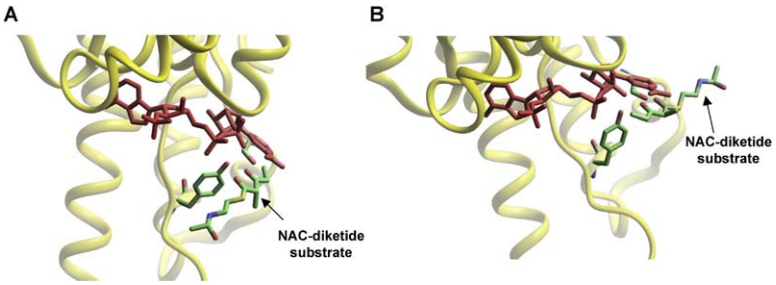


Figure 2. Homology Model of the Ketoreductase eryKR$_1$ Showing Two Alternative Binding Modes of the Synthetic Substrate (2*S*)-2-Methyl-3-Oxopentanoyl-NAC Thioester

(A and B) The orientation in (A) yields the (2*S*, 3*R*) diasteroisomer of the product alcohol, while the orientation in (B) yields the (2*S*, 3*S*) diastereoisomer. The NADPH cofactor is shown in red, and active site residues Ser136 and Tyr149, which polarize the substrate carbonyl group, are highlighted. For further details, refer to Figure 3 and the text. The derivation of the homology model will be presented separately (A.B.-O. et al., submitted).

In contrast, those KR domains that would not normally process diketide substrates gave mixtures of stereoisomeric products in these assays. For example, the multienzyme DEBS 3 (which houses both eryKR$_5$ and eryKR$_6$ domains), which we purified as a recombinant protein from extracts of *E. coli* as previously described [27], gave only 42% of the expected (2*R*, 3*S*) product. We previously reported that diketide ketoreduction on this enzyme gave roughly equal amounts of all four isomers [20], but our present reanalysis showed that, under the conditions used here, the only other major product was the (2*S*, 3*S*) isomer, in which the opposite enantiomer is accepted but nevertheless reduced from the correct face at C-3. Consistent with this, when the individual eryKR$_5$ and eryKR$_6$ domains were assayed, the same pattern was found, except that the expected isomer formed only 25% of the total product mixture. These data could be taken as evidence that the specificity of these domains can be altered and heightened by their incorporation into a multienzyme assembly line, or by tethering the substrates to an appropriate ACP domain. The obvious caveat is that these enzymes might also have shown full integrity of stereochemical outcome had they been presented with substrates more closely resembling their natural substrates.

The results obtained with ketoreduction by eryKR$_2$ (Table 3) are even more important. Here, the predicted (2*R*, 3*S*) product represents only 6% of the product mixture, and, in this case, the "wrong" stereoisomer not only serves as a substrate, but gives rise to significant amounts of products arising by hydride delivery from NADPH to either face at C-3. Taken together, these data show once again that it cannot be assumed that an unnatural substrate will always be correctly processed by the KR domains in an engineered hybrid PKS [20]. The stereospecificity of eryKR$_2$ is less evidently hardwired, although it still catalyzes reduction from the "correct" face of the keto group 80% of the time.

Two Alternative Binding Modes of Diketide Substrates in the Active Sites of PKS Ketoreductases

It has been previously shown that in KRs 1, 2, 5, and 6 the pro-*S* hydrogen is specifically delivered to the keto substrate from the C-4′ of NADPH [28, 29]. The KR domains of modular PKSs have high sequence similarity to each other, and to the KR domain of fatty acid synthase, and homology modelling and site-directed mutagenesis [15] has confirmed that they are members of the family of short-chain dehydrogenase/reductase (SDR) enzymes [30], with an active site containing highly conserved lysine, serine, and tyrosine residues (and possibly also asparagine) [15]. We have independently obtained a detailed homology model of both eryKR$_1$ and eryKR$_2$ (A.B.-O. et al., submitted) that shows that eryKR$_1$ and eryKR$_2$ have exactly the same overall fold and a highly similar active site geometry. In particular, the positioning of the NADPH coenzyme and of critical active site residues Ser136 and Tyr149 (numbering according to [16]), which are likely to be involved in the polarization of the keto group [15], are predicted to be essentially identical for KR$_1$ and for KR$_2$ (Figure 2).

Figure 3 summarizes how we propose the different stereochemical outcomes of reduction by eryKR$_1$, eryKR$_2$, and the other KRs studied arise: from two alternative binding modes of each enantiomeric diketide substrate to the active site. Although confirmation of this idea must await detailed structural studies on KR domains with bound substrate, it accounts straightforwardly for the observed range of behavior of the various KRs. As the results in Table 3 show, a single enantiomer in only one of the two binding modes may be exclusively favored (eryKR$_1$ and tylKR$_1$). In the other cases studied, when the correct enantiomer is bound, the reduction does take place specifically from the predicted face (eryKR$_2$, eryKR$_5$, and eryKR$_6$). For these three enzymes, however, the reduction of the "wrong" enantiomer is kinetically favored and, in the case of KR$_2$, gives rise to significant reduction even from the "wrong" face of the keto group.

The idea of two alternative binding modes, and the observation of product mixtures in our experiments, leads to the prediction that the detailed arrangement of relatively few amino acids in the respective active sites of individual KRs might be a key determinant of stereochemical outcome. It has already been reported [15, 16] that certain amino acid residues are characteristic of natural modular PKS KRs that reduce from one or the other face of a ketoacylthioester substrate. In particular, Caffrey [16] has carefully documented two amino acid motifs, which he terms A- and B-side selective, whose respective absence or presence are hallmarks of these two groups of enzymes (Figure 4). As will be discussed in detail separately (A.B.-O. et al., submitted), the residues of the Caffrey motifs both lie in the KR active site predicted from homology modeling (Figure 2). Recently, two research groups [31, 32] have independently determined crystal structures for the

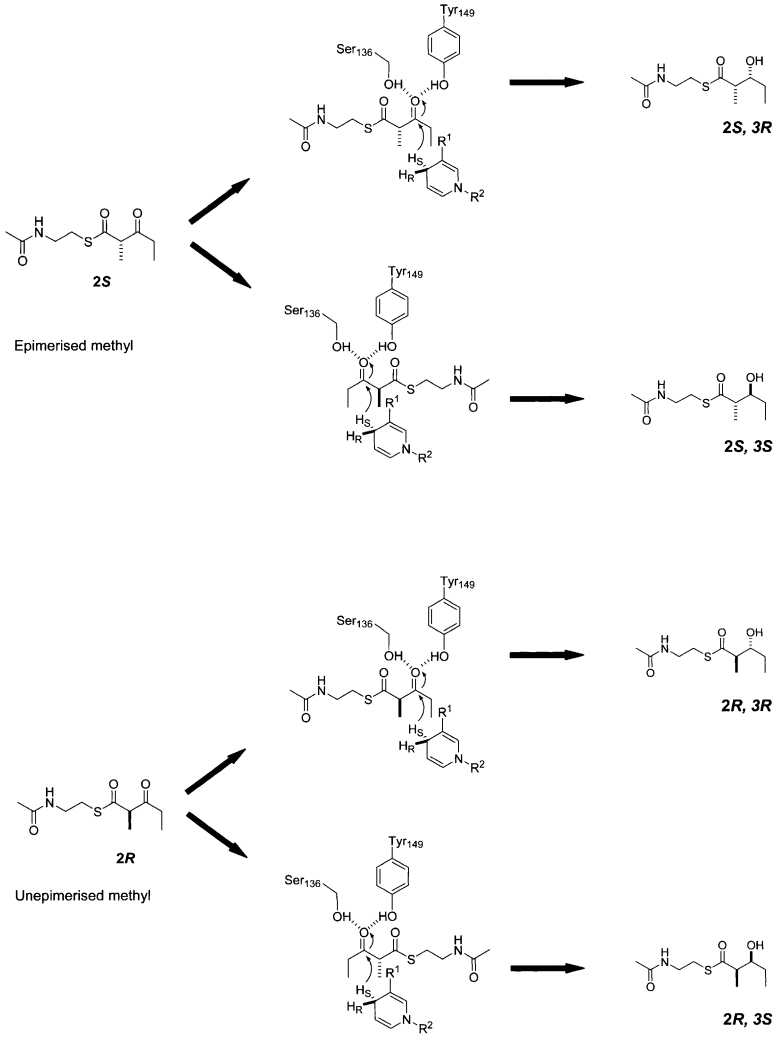


Figure 3. Alternative Orientations of 2-Methyl-3-Oxopentanoic Acid NAC Thioesters in the Active Site of PKS Ketoreductases

The relative positions of the NADPH cofactor and the active site catalytic residues required to polarize the keto group (Ser136 and Tyr149) remain fixed, and, therefore, the orientation of the substrate within the active site determines the observed direction of ketoreduction. eryKR$_1$ yields only the (2*S*, 3*R*) product, and tylKR$_1$ almost exclusively yields the (2*R*, 3*R*) product. In contrast, eryKR$_2$, eryKR$_5$, and eryKR$_6$ give mixtures of products.

actinorhodin KR, which specifically reduces the C-9 keto group during the biosynthesis of this aromatic polyketide. These structures confirm that the Caffrey motifs are predicted to lie within the KR active site, but these sequence motifs are not conserved in aromatic KRs [32], and no clear prediction could be made from inspection of the crystal structure as to which stereoisomer of the alcohol would be produced [32]. The respective roles of individual amino acid residues in controlling stereospecificity and stereoselectivity therefore remain to be determined, for both aromatic and modular PKS KRs. It remains an attractive idea

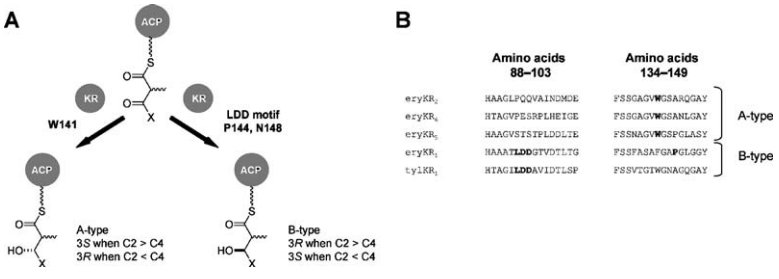


Figure 4. Amino Acid Motifs, which Are Putatively Involved in the Control of Stereochemistry by KR Domains from the Erythromycin and Tylosin PKSs

Amino acid motifs have been identified by Caffrey [16].

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

(B) Sequence alignment of KR domains from the erythromycin and tylosin PKSs. The active site motifs, where present, are indicated in bold (residues are numbered according to [16]). Note that tylKR$_1$ has W141 although it is a B-type KR.

that engineered alteration of the amino acid residues at the active site, and especially the residues comprising the Caffrey motifs, might trigger a switch of stereochemical outcome in modular PKS KRs. There is clear precedent for this in the recent demonstration that the stereoselectivity of tropinone reductases that reduce, respectively, (+)- and (−)-tropinone can be switched by site-directed mutagenesis [33]. These enzymes are also members of the SDR family of NAD(P)H-linked ketoreductases, and their active site geometry has been shown to be closely related to that of the actinorhodin KR [31]. Now that the modular PKS KR domains can be conveniently studied as individual recombinant enzymes, this idea can be readily tested by using experimental approaches based on both site-directed mutagenesis and molecular evolution.

Significance

Numerous strategies have been successfully developed to engineer modular PKSs, and also downstream post-PKS enzymes, to produce chemically altered polyketide products for use in drug discovery. In contrast, attempts to alter the configuration at specific stereocenters have been hampered by our limited understanding of the fundamental mechanisms by which PKSs exert stereocontrol over polyketide assembly. We have shown that individual KR domains can be expressed as enzymatically active enzymes in *E. coli*, separate from other PKS activities. The KR domains show a range of kinetic and stereochemical behavior against synthetic diketide thioester substrates. KR domains that normally act on diketide substrates showed essentially complete stereocontrol, demonstrating that this is a property intrinsic to the KR. In contrast, the other KR domains discriminated poorly between stereoisomeric substrates and reduced certain substrates from both faces of the keto group. These results show that KR stereocontrol is intrinsic, but also delicately balanced, and so it cannot be assumed that a KR in the context of a hybrid PKS will behave faithfully. On the other hand, this even energetic balance may herald the possibility of engineering a switch in the stereospecificity of a given KR domain, by targeting key active site residues. This could lead to more reliable and efficient engineering of polyketide stereochemistry.

Experimental Procedures

Materials

Routine cloning and transformation procedures were as previously described for *E. coli* [34]. Electrocompetent cells of the *E. coli* DH10B strain were made as described previously [35]. 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 L MQ water.

Cloning Procedures

The DNA fragment corresponding to $eryKR_1$ (accession number Q03131) was amplified by PCR by using Pfu polymerase (Stratagene) and oligonucleotides 5′-GTCCACCGGATCCGACGAG GTTTCCGCGCTGCG-3′ and 5′-GGGCAGGAATT CTCACGCGCCCA CCCGCGGTTCGGC-3′ as primers and by using a pCJR24-based plasmid containing the gene for DEBS 1-TE as the template [36]. The DNA fragment corresponding to $eryKR_2$ was amplified as described above, but by using oligonucleotides 5′-GCACCACCGGA TCCGACGAGCTCGACGGCTGGTTCTAC-3′ and 5′-GCACGCGGCA GGAATTCTCACCGGTCGCGCAGGCTCTC-3′ as primers. For $tylKR_1$ (accession number U78289), oligonucleotides 5′-ACGACCGGATCC AGCCCCACCGATGCCTGGCGCTAC-3′ and 5′-TCAGCATGAATTC TCACCGCGCCAGCTCGGGCGCGC-3′ were used as primers, and cosmid pTB2, containing the PKS genes for the biosynthesis of tylosin, was used as the template. $eryKR_5$ and $eryKR_6$ were PCR amplified from a plasmid (pIB023, I.U. Böhm, personal communication) containing the three DEBS genes by using primers 5′-GGACGGAT CCGACGACTGGCGCTA-3′ and 5′-GCAGGAATTCTCAGAGCCGCT GGGCGAGCGCCGGGCCCTCCT-3′ for $eryKR_5$, and 3-CGACGGA TCCGCCGACAGCCGCTACCGCGTCGACTGG-5′ and 3-ACTCGAA TTCTCACGTCATCTCCCGCGCCGGGGCCGCCT-5′ for $eryKR_6$. All constructs were checked by DNA sequencing, and the resulting DNA fragments were cloned into pGEX4T-3 (Amersham) by digestion with BamHI and EcoRI. With the exception of $tylKR_1$, all of the KRs were subcloned into pGEX-6P-1 (Amersham) by digestion with BamHI and EcoRI.

Protein Expression

One liter of culture in LB medium of *E. coli*-BL21-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 OD_{600} of 0.6, were induced with 0.1 mM IPTG (Melford Laboratories). After growth for an additional 16 hr, cells were collected by centrifugation, and cell pellets were stored at −20ºC.

Protein Purification

Typically, frozen cell pellets were thawed and resuspended in lysis buffer (50 mM NaH_2PO_4 [pH 8.2], 150 mM NaCl, 10% glycerol) containing 1 tablet of protease inhibitors (Complete Roche), 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 were disrupted by sonication (Misonix, Inc.), 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 by resuspending in 10 volumes of lysis buffer and were eluted by incubating in elution buffer (50 mM Tris-HCl [pH 8.0], 10 mM reduced glutathione) for 30 min at 4ºC on a rotor. The elution step was repeated three times, and the fractions were collected. Typical yields for all proteins were 10−13 mg/L of culture, as measured by the Bradford assay.

Cleavage of the GST fusion proteins was carried out by incubation with human rhinovirus C protease (PreScission Protease, Amersham) at a concentration of 20 U/mg GST-KR, in 50 mM Tris-HCl buffer (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM DTT at 4ºC for 6–24 hr. Separation of the GST cleavage products was carried out on an anion exchange column (MonoQ, Hi Prep Q FF [Amersham]). The column was washed with two column volumes of HEPES buffer, and the proteins were eluted by using a linear salt gradient (100–1000 mM NaCl) over four column volumes.

Kinetic Assays

NADPH-linked assays were carried out at 30ºC in the chamber of a Spectramax Plus microplate reader running Softmax Pr (Molecular Devices) in 25 mM HEPES buffer (pH 7.5). 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). KR domains were used at 2–8 μM, and NADPH was used at 1.3 mM; the substrates were dissolved in 3% DMSO, and their concentrations varied between 0 and 50 mM (keeping the total amount of DMSO constant). The total assay volume was 110 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, at least. The raw data were converted into a velocity (rate of consumption of NADPH) for each concentration, and then the velocities were plotted against substrate concentration. The data were fitted to the Michaelis-Menten steady-state equation by using KaleidaGraph (Abelbeck Software).

Product Separation by Chiral HPLC

Reduction of NAC-diketide substrate was carried out in a total volume of 200 μl containing 10 mM NADPH, 15 mM NAC-diketide, 400 mM NaH_2PO_4 (pH 7.5), and 1 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 by using a ChiralCel OC column (Daicel, 25 cm × 4.6 mm) on 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. Control samples were run with each batch of analyses to determine the retention times of the four diastereoisomeric diketides, and, in a few cases, samples were spiked with one of the standards.

Acknowledgments

We thank Professor Sir Tom Blundell for helpful discussions and Drs. Ben Luisi and Fabian Walas for help with the Spectromax plate recorder. We are grateful for the support of this work by a project grant to P.F.L. and J.S. from the UK Biotechnology and Biological Sciences Research Council (BBSRC). A.B.-O. is a Royal Society USA Postdoctoral Research Fellow, K.J.W. is a Royal Society Dorothy Hodgkin Research Fellow, and D.S. is a post-doctoral fellow of the Leopoldina German Academy of Natural Scientists. H.M. was supported by a BBSRC studentship, and S.B. is supported by the Dharam Hinduja Trust and the Nehru Cambridge Commonwealth Trust.

Received: July 7, 2005
Revised: August 1, 2005
Accepted: August 1, 2005
Published: October 21, 2005

References

1. Katz, L., and Ashley, G.W. (2005). Translation and protein synthesis: macrolides. Chem. Rev. *105*, 499–528.
2. Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.
3. Walsh, C.T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science *303*, 1805–1810.
4. Weissman, K.J. (2004). Polyketide biosynthesis: understanding and exploiting modularity. Philos. Trans. R. Soc. Lond. A *362*, 2671–2690.
5. Walsh, C.T. (2002). Combinatorial biosynthesis of antibiotics: challenges and opportunities. ChemBioChem *3*, 125–134.
6. Khosla, C., and Keasling, J.D. (2003). Metabolic engineering for drug discovery and development. Nat. Rev. Drug Discov. *2*, 1019–1025.
7. Mendez, C., and Salas, J.A. (2001). Altering the glycosylation pattern of bioactive compounds. Trends Biotechnol. *19*, 449–456.
8. Piel, J., Butzke, D., Fusetani, N., Hui, D., Platzer, M., Wen, G., and Matsunaga, S. (2005). Exploring the chemistry of uncultivated bacterial symbionts: antitumor polyketides of the pederin family. J. Nat. Prod. *68*, 472–479.
9. Tang, G.L., Cheng, Y.Q., and Shen, B. (2004). Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. Chem. Biol. *11*, 33–45.
10. Rahman, A.S., Hothersall, J., Crosby, J., Simpson, T.J., and Thomas, C.M. (2005). Tandemly duplicated acyl carrier proteins, which increase polyketide antibiotic production, can apparently function either in parallel or in series. J. Biol. Chem. *280*, 6399–6408.
11. Arakawa, K., Sugino, F., Kodama, K., Ishii, T., and Kinashi, H. (2005). Cyclization mechanism for the synthesis of macrocyclic antibiotic lankacidin in *Streptomyces rochei*. Chem. Biol. *12*, 249–256.
12. Celmer, W.D. (1965). Biogenetic, constitutional, and stereochemical unitary principles in macrolide antibiotics. Antimicrob. Agents Chemother. *5*, 144–156.
13. 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.
14. 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.
15. 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.
16. Caffrey, P. (2003). Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. ChemBioChem *4*, 649–662.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.
22. 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.
23. 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.
24. 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.
25. Aparicio, J.F., Caffrey, P., Marsden, A.F.A., Staundon, J., and Leadlay, P.F. (1994). Limited proteolysis and active site studies of the first multienzyme component of the erythromycin-producing polyketide synthase. J. Biol. Chem. *269*, 8524–8528.
26. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876–4882.
27. Weissman, K.J., Hong, H., Oliynyk, M., Siskos, A.P., and Leadlay, P.F. (2004). Identification of a phosphopanetheinyl transferase for erythromycin biosynthesis in *Saccharopolyspora erythraea*. Chembiochem *5*, 116–125.
28. 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.
29. Yin, Y., Gokhale, R., Khosla, C., and Cane, D.E. (2001). Erythromycin biosynthesis. The 4-pro-S hydride of NADPH is utilized

for ketoreduction by both module 5 and module 6 of the 6-deoxyerythronolide B synthase. Bioorg. Med. Chem. Lett. *11*, 1477–1479.

30. Kallberg, Y., Oppermann, U., Jornvall, H., and Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). Eur. J. Biochem. *269*, 4409–4417.
31. 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 (Camb) *12*, 1865–1875.
32. 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.
33. 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.
34. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
35. 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.
36. Rowe, C.J., Cortés, J., Gaisser, S., Staunton, J., and Leadlay, P.F. (1998). Construction of new vectors for high-level expression in actinomycetes. Gene *216*, 215–223.