J|A|C|S ARTICLES

Published on Web 11/23/2009

The Biochemical Basis for Stereochemical Control in Polyketide Biosynthesis

Chiara R. Valenzano,[†] Rachel J. Lawson,[†] Alice Y. Chen,[‡] Chaitan Khosla,[‡,§,||] and David E. Cane*[,†]

Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912-9108, and Departments of Chemical Engineering, Chemistry, and Biochemistry, Stanford University, Stanford, California 94305

Received September 29, 2009; E-mail: David_Cane@brown.edu

Abstract: One of the most striking features of complex polyketides is the presence of numerous methyl- and hydroxyl-bearing stereogenic centers. To investigate the biochemical basis for the control of polyketide stereochemistry and to establish the timing and mechanism of the epimerization at methyl-bearing centers, a series of incubations was carried out using reconstituted components from a variety of modular polyketide synthases. In all cases the stereochemistry of the product was directly correlated with the intrinsic stereospecificity of the ketoreductase domain, independent of the particular chain elongation domains that were used, thereby establishing that methyl group epimerization, when it does occur, takes place after ketosynthase-catalyzed chain elongation. The finding that there were only minor differences in the rates of product formation observed for parallel incubations using an epimerizing ketoreductase domain and the nonepimerizing ketoreductase domain supports the proposal that the epimerization is catalyzed by the ketoreductase domain itself.

The macrolides methymycin, erythromycin, picromycin, and tylosin are closely related 12-, 14-, and 16-membered-ring lactone members of a large family of several hundred branched-chain, polyoxygenated polyketides, many of which are widely used in human and veterinary medicine as antibiotic, immunosuppresant, antitumor, antifungal, and antiparasitic agents (Figure 1). The biosynthesis of their parent macrocyclic aglycones **1**−**4** is controlled by dedicated modular polyketide synthases (PKSs), a group of unusually large, multifunctional enzymes that catalyze the multistep assembly of these extraordinarily complex metabolites.[1] Each homodimeric PKS subunit contains from 1 to 10 PKS modules of 150−250 kDa, each of which is responsible in turn for a single round of polyketide chain elongation and functional group modification. Within every module are a minimum of three core domains consisting of a β-ketoacyl-ACP synthase (the KS domain) that catalyzes the key chain-building decarboxylative condensation of a malonyl-, methylmalonyl-, or occasionally ethymalonyl-ACP building block with the polyketide chain that has been generated by the upstream PKS module; an acyl transferase (AT) domain that loads the correct malonyl-CoA or methylmalonyl-CoA extender unit onto the flexible 18-Å phosphopantetheinyl arm of the acyl-carrier protein (ACP) domain; and the ACP domain itself, that carries the polyketide biosynthetic intermediates from domain to domain and then delivers the resulting product to the KS domain of the next module. Most modules also harbor additional domains that together are responsible for modification of the oxidation state and stereochemistry of the growing ACP-bound intermediates: a β-ketoacyl-ACP reductase (KR) domain, a dehydratase (DH) domain, and an enoyl reductase (ER) domain. At the N-terminus of the most upstream module is an additional loading didomain or tridomain that primes the KS1 domain. Finally, release and cyclization of the full-length macrocyclic polyketide are controlled by a dedicated thioesterase (TE) domain that is located at the C-terminus of the most downstream module. Subsequent glycosylations, oxidations, methylations, and other post-PKS modifications generate the mature, biologically active polyketide natural product. The chain length, substitution pattern, and oxidation level of the initially generated, full-length polyketide, such as 6-deoxyerythronolide B (6-dEB, **1**), the erythromycin macrolide aglycone, are the direct consequence of the number of PKS modules as well as the domain composition of each module. For example, 6-dEB synthase (DEBS) consists of three bimodular proteins that together catalyze the six cycles of polyketide chain elongation and functional group modification that generate the macrocyclic heptaketide, 6-dEB (Figure 1a).[2] The 14-membered ring of 6-dEB displays an impressive aggregate of 10 stereogenic centers, corresponding to three D-methyl, three-L-methyl, one D-hydroxy, and three L-hydroxy substituents. As originally noted by Celmer and is evident in the aglycone structures **1**−**4**, similar

[†] Brown University.
[‡] Department of Chemical Engineering, Stanford University.
[§] Department of Chemistry, Stanford University.
[||] Department of Biochemistry, Stanford University.

(1) (a) Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E. *Annu. Rev. Biochem.* **2007**, *76*, 195–221. (b) Khosla, C.; Kapur, S.; Cane, D. E. *Curr. Opin. Chem. Biol.* **2009**, *13*, 135–143. (c) Hopwood, D. A., Ed. *Methods in Enzymology. Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates*; Academic Press: 2009; Vol. 459.

(2) (a) Donadio, S.; Katz, L. *Gene* **1992**, *111*, 51–60. (b) Bevitt, D. J.; Cortes, J.; Haydock, S. F.; Leadlay, P. F. *Eur. J. Biochem.* **1992**, *204*, 39–49. Caffrey, P.; Bevitt, D. J.; Staunton, J.; Leadlay, P. F. *FEBS Lett.* **1992**, *304*, 225–228.

10.1021/ja908296m

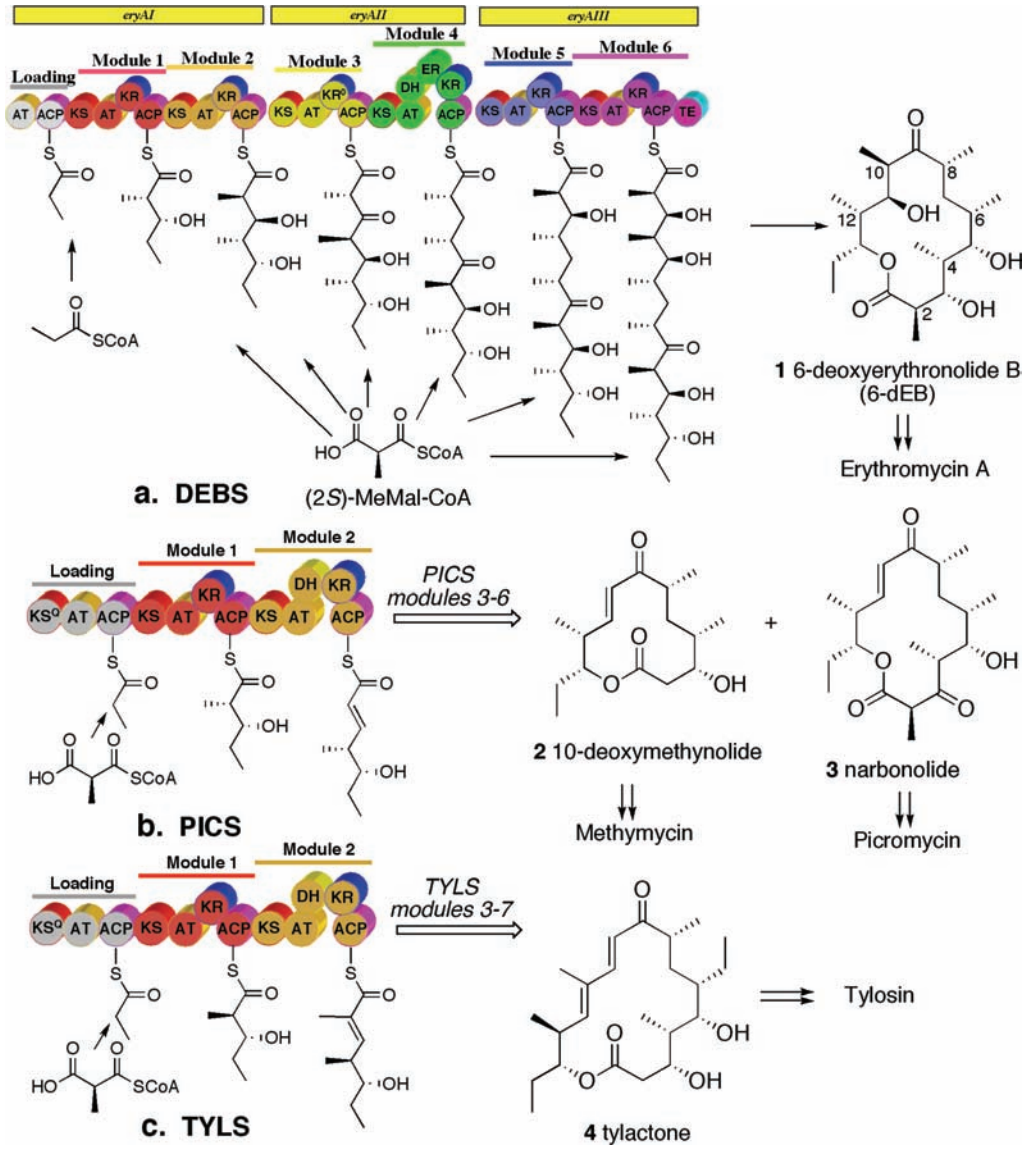


Figure 1. Modular organization of (a) the 6-deoxyerythronolide B (6-dEB, **1**) synthase (DEBS), (b) the picromycin/methymycin (10-deoxymethynolide (**2**)/narbonolide (**3**)) synthase (PICS), and (c) the tylactone (**4**) synthase (TYLS). Only the first two modules and the loading domains of PICS and TYLS are illustrated. In addition to the three core catalytic domains—the β-ketoacyl-ACP synthase (KS), the acyltransferase (AT), and the acyl carrier protein (ACP) domains—individual extension modules carry varying combinations of tailoring ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains. Loading didomains or tridomains prime module 1 of each synthase with the propionyl starter unit and a thioesterase (TE) domain at the C-terminus of the furthest downstream module catalyzes release and cyclization of the full-length polyketide giving the parent macrolide aglycone. Dedicated oxygenases, glycosyl transferases, and methyl transferases then generate the mature antibiotic.

stereochemical patterns are found in a wide range of closely related 12-, 14-, and 16-membered ring macrolides.[3] In spite of dramatic progress in the elucidation of the biochemistry and molecular genetics of polyketide biosynthesis, however, the mechanistic basis for the remarkable stereochemical complexity of the polyketides generated by modular PKSs is still incompletely understood.

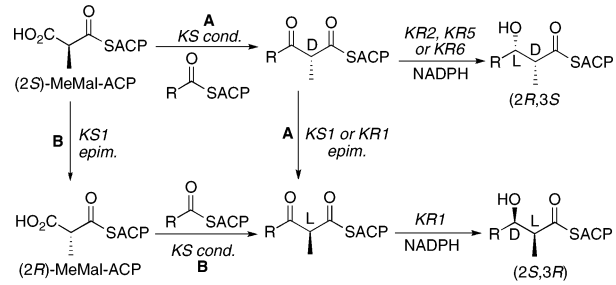
The stereochemical course of each β-ketoacyl-ACP reduction is an intrinsic property of the responsible KR domain and is independent of either modular context or substrate structure, including chain length and substitution pattern.[4] The DEBS KR1, KR2, KR5, and KR6 domains all utilize the 4*si* (4-pro-*S*) hydride of the NADPH cofactor, in common with the closely related KR domains of both Type I and Type II fatty acid synthases.[5] Indeed, based on the high level of sequence conservation, this 4*si* stereospecificity is probably a universal property of all PKS KR domains. There is also an intriguing correlation between the amino acid sequences of PKS KR domains and the known stereospecificity of ketoreduction, with a highly conserved Leu-Asp-Asp triad being strongly correlated with, if not necessarily the sole determinant of, the generation of D-hydroxyl stereochemistry in the product.[6] In the absence of this conserved motif, the observed reduction product has the epimeric L-hydroxy stereochemistry. Crystal structures of the KR1 domain from module 1 of DEBS and of the homologous

(3) Celmer, W. D. *J. Am. Chem. Soc.* **1965**, *87*, 1801–1802. Celmer, W. D. *Pure Appl. Chem.* **1971**, *28*, 413–453.
(4) Kao, C. M.; McPherson, M.; McDaniel, R. N.; Fu, H.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1998**, *120*, 2478–2479.
(5) McPherson, M.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **1998**, *120*, 3267–3268. Yin, Y.; Gokhale, R.; Khosla, C.; Cane, D. E. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1477–1479.

domain from module 1 of tylactone synthase (TYLS) have suggested that the loop harboring the Leu-Asp-Asp motif plays a key role in mediating access of the phosphopantetheinyl-bound β-ketoacyl thioester substrate to the common substrate-binding groove, thereby controlling the orientation of the substrate carbonyl relative to the enzyme-bound NADPH cofactor, resulting in formation of the observed D-hydroxy product.[7] In KR domains that lack this Leu-Asp-Asp motif but harbor instead a conserved Trp at an alternative site, the phosphopantetheinyl-bound β-ketoacyl substrate is thought to be constrained to enter the substrate-binding groove from the opposite end, thereby giving rise to the observed L-hydroxy product by presenting the opposite face of the ketone to the cofactor. In contrast with this emerging picture of the protein structural basis for control of hydroxyl group stereochemistry, the biochemical basis for the control of D- or L-methyl group stereochemistry is considerably more obscure.[8] All six AT domains of the 6-dEB synthase are completely stereospecific for (2*S*)-methylmalonyl-CoA,[9] while the decarboxylative condensations catalyzed by the various KS domains all take place with inversion of configuration at C-2 of the methylmalonyl unit, irrespective of the eventual D- or L-configuration of the resulting 2-methyl-3-ketoacyl-ACP product.[10] Generation of the L-methyl groups at C-12 and C-8 of 6-dEB (**1**), mediated by DEBS module 1 and 3, respectively, therefore requires an epimerization at some stage in the catalytic cycle of chain elongation and functional group modification. (The L-configuration of the C-6 methyl group of 6-dEB, generated by DEBS module 4, is unrelated to the stereochemistry of condensation or ketoreduction but is the direct consequence of the ER-catalyzed stereospecific reduction of the corresponding (*E*)-2,3-enoyl-ACP pentaketide intermediate.) Unlike modular nonribosomal peptide synthetases, however, which utilize discrete epimerization domains to catalyze the conversion of ACP-bound L-amino acid building blocks to the epimeric D-amino acid derivatives,[11] neither DEBS nor any related PKS harbors a readily discernible epimerization domain. The mechanism by which polyketide synthases control the configuration at methyl-bearing centers therefore remains one of the major, unsolved biochemical questions of polyketide biosynthesis. Indeed, neither the precise timing of methyl group epimerization nor the responsible catalytic domain has as yet been unambiguously identified. Various investigators have suggested that either the KS domain[12] or the KR domain[7b,13] might be responsible for the methyl group epimerization. In the chain elongation cycle mediated by DEBS module 1, since the AT1 domain is strictly specific for (2*S*)-methylmalonyl-CoA, the inversion of the methyl-bearing center must occur at some as yet unspecified stage after the original methylmalonyl unit has been loaded onto the ACP domain. In principle this epimerization could occur subsequent to formation of the 2-D-methyl-3-ketoacyl-ACP intermediate (Scheme 1, pathway A) or prior to the KS-catalyzed condensation reaction (Scheme 1, pathway B). In pathway A, the KS domain would first catalyze the decarboxylative condensation of the (2*S*)-methylmalonyl-ACP with the acyl thioester provided by the upstream module to generate the D-2-methyl-3-ketoacyl-ACP intermediate. For a module with an associated epimerase activity, such as DEBS module 1, this 2-methyl-3-ketoacyl-ACP thioester then would undergo epimerization followed by diastereoselective reduction mediated by the KR domain, thereby fixing the configuration at both C-2 and C-3 of the resultant 2-methyl-3-hydroxyacyl-ACP chain elongation intermediate. In principle, the key epimerization could be catalyzed either by the KS domain, with the KR domain then stereospecifically reducing only the resulting 2-L-methyl-3-ketoacyl-ACP diastereomer, or by the KR domain itself, which might harbor *both* the epimerase and reductase activities, with diastereospecific reduction of only the epimerized, KR-bound 2-L-methyl-3-ketoacyl-ACP substrate. By contrast, in pathway B, the KS domain would first epimerize (2*S*)-methylmalonyl-ACP to (2*R*)-methylmalonyl-ACP and then catalyze the decarboxylative condensation with net inversion of configuration to give the L-2-methy-3-ketoacyl-ACP intermediate. In pathway B, therefore, control of the methyl stereochemistry would be vested exclusively in the KS domain, with the KR domain simply reducing the initially generated ACP-bound diketide intermediate.

We have recently described a sensitive and versatile analytical protocol for the determination of the stereochemical course of β-ketoacyl-ACP synthase- and/or β-ketoacyl-ACP reductase-catalyzed reactions (Scheme 2).[14] In this method, pairs of recombinant ketosynthase-acyl transferase [KS][AT] didomains and discrete acyl carrier protein (ACP) domains are incubated

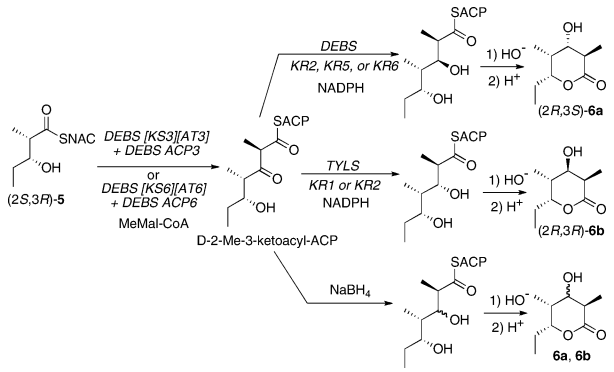
Scheme 1. Proposed Pathways for Methyl Group Epimerization[a]



[a] In Pathway A, epimerization follows KS-catalyzed condensation, with the KR domain reducing only the corresponding diastereomer of the 2-methyl-3-ketoacyl-ACP intermediate. In Pathway B, KS-catalyzed epimerization precedes chain elongation, with the KR domain reducing the resulting intermediate of the correct configuration.

(6) (a) Reid, R.; Piagentini, M.; Rodriguez, E.; Ashley, G.; Viswanathan, N.; Carney, J.; Santi, D. V.; Hutchinson, C. R.; McDaniel, R. *Biochemistry* **2003**, *42*, 72–79. (b) Caffrey, P. *ChemBioChem* **2003**, *4*, 654–657.
(7) (a) Keatinge-Clay, A. T.; Stroud, R. M. *Structure* **2006**, *14*, 737–748. (b) Keatinge-Clay, A. T. *Chem. Biol.* **2007**, *14*, 898–908.
(8) (a) Siskos, A. P.; Baerga-Ortiz, A.; Bali, S.; Stein, V.; Mamdani, H.; Spiteller, D.; Popovic, B.; Spencer, J. B.; Staunton, J.; Weissman, K. J.; Leadlay, P. F. *Chem. Biol.* **2005**, *12*, 1145–1153. (b) Bali, S.; O'Hare, H. M.; Weissman, K. J. *ChemBioChem* **2006**, *7*, 478–484. (c) Bali, S.; Weissman, K. J. *ChemBioChem* **2006**, *7*, 1935–1942. (d) Ostergaard, L. H.; Kellenberger, L.; Cortes, J.; Roddis, M. P.; Deacon, M.; Staunton, J.; Leadlay, P. F. *Biochemistry* **2002**, *41*, 2719–2726.
(9) (a) Marsden, A. F.; Caffrey, P.; Aparicio, J. F.; Loughran, M. S.; Staunton, J.; Leadlay, P. F. *Science* **1994**, *263*, 378–380. (b) Wiesmann, K. E.; Cortes, J.; Brown, M. J.; Cutter, A. L.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **1995**, *2*, 583–589.
(10) (a) Weissman, K. J.; Timoney, M.; Bycroft, M.; Grice, P.; Hanefeld, U.; Staunton, J.; Leadlay, P. F. *Biochemistry* **1997**, *36*, 13849–13855. (b) Cane, D. E.; Liang, T. C.; Taylor, P. B.; Chang, C.; Yang, C. C. *J. Am. Chem. Soc.* **1986**, *108*, 4957–4964.
(11) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. *Chem. Rev.* **1997**, *97*, 2651–2674.
(12) Bohm, I.; Holzbaur, I. E.; Hanefeld, U.; Cortes, J.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **1998**, *5*, 407–412. Holzbaur, I. E.; Ranganathan, A.; Thomas, I. P.; Kearney, D. J.; Reather, J. A.; Rudd, B. A.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **2001**, *8*, 329–340.
(13) Starcevic, A.; Jaspars, M.; Cullum, J.; Hranueli, D.; Long, P. F. *ChemBioChem* **2007**, *8*, 28–31.
(14) (a) Castonguay, R.; He, W.; Chen, A. Y.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2007**, *129*, 13758–13769. (b) Castonguay, R.; Valenzano, C. R.; Chen, A. Y.; Keatinge-Clay, A.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2008**, *130*, 11598–11599.

Scheme 2. Stereochemistry of Triketide Lactone Formation Catalyzed by Dissected DEBS [KS][AT] and ACP Domains in Combination with Recombinant KR Domains



in combination with recombinant ketoreductase (KR) domains from a range of parent PKS systems in the presence of the diketide substrate analogue, (2*S*,3*R*)-2-methyl-3-hydroxypentanoic acid *N*-acetylcysteamine (SNAC) thioester (**5**), along with methylmalonyl-CoA and NADPH. After release of the resultant reduced, ACP-bound triketide by base-catalyzed hydrolysis and cyclization, the stereochemistry of the derived triketide lactone **6** is then assigned by GC-MS analysis of the corresponding trimethylsilyl (TMS) ether and direct comparison with the retention time and mass spectrum of synthetic standards of each of the four possible 2-methyl-3-hydroxy diastereomers of **6**. In this manner, we have shown that the combination of the DEBS module 6 [KS6][[AT6] didomain and the discrete DEBS ACP6 domain with DEBS KR6 generates the triketide lactone **6a**. Use of either DEBS KR2 or KR5 in the incubation mixture gives the identical product **6a**. The observed configurations of the unepimerized D-2-methyl group and of the L-3-hydroxyl group of **6a** are fully consistent with the stereochemistry of the corresponding natural triketide, hexaketide, and heptaketide chain elongation intermediates generated by DEBS modules 2, 5, and 6, respectively. Identical results were also obtained using DEBS KR2, KR5, or KR6 in combination with the alternative chain elongation pair, DEBS [KS3][AT3] plus ACP3. By contrast, incubation of TYLS KR1 or TYLS KR2 in combination with either DEBS [KS6][AT6]+ACP6 or DEBS [KS3][AT3]+ACP3 gave exclusively the diastereomeric triketide lactone **6b**, carrying a D-3-hydroxyl group but an unepimerized D-2-methyl group.[14] The observed stereochemistry of the TYLS KR1 product corresponds to the natural *anti*-(2*R*,3*R*)-2-methyl-3-hydroxy configuration of the ACP-bound tylactone diketide intermediate produced by TYLS module 1 (Figure 1),[8a,15] while the formation of **6b** by TYLS KR2 revealed for the first time the previously cryptic stereochemistry of the reduced, ACP-bound triketide intermediate that normally serves as the substrate for the dehydratase domain of TYLS module 2. Control experiments in which the ACP-bound 2-methyl-3-ketoacyl triketide intermediates were trapped by reductive quenching with aqueous sodium borohydride gave a mixture of the 3-hydroxy diastereomers **6a** and **6b**, thereby establishing that the decarboxylative condensation catalyzed by DEBS [KS6][AT6]+ACP6 or [KS3][AT3]+ACP3 takes place in each case with net inversion of configuration, giving exclusively the product with an unepimerized D-2-methyl group (Scheme 2). The initially generated D-2-methyl-3-ketoacyl-ACP triketide did not undergo significant epimerization over a period of more than 1 h under the standard incubation conditions.

We now report results that firmly establish that polyketide methyl group epimerization occurs subsequent to the KS-catalyzed condensation step and before reduction of the 2-methyl-3-ketoacyl-ACP intermediate. We also provide evidence indicating that the KR domain itself catalyzes this methyl group epimerization and thereby controls the configuration of both the newly generated 3-hydroxyl group and its vicinal 2-methyl substituent.

Results

Expression of Discrete Domains from DEBS and PICS Module 1. None of the KR domains and neither of the [KS][AT] didomains examined in the previously reported combinatorial PKS incubation experiments are derived from a parent PKS module that mediates the formation of a *reduced* polyketide with an *epimerized* methyl group. To establish the origin and timing of methyl group epimerization, we therefore turned our attention to the catalytic domains from two PKS modules that normally produce a reduced polyketide chain elongation product carrying an epimerized L-methyl group. Both DEBS module 1 and picromycin/methymycin synthase (PICS) module 1 catalyze the first round of polyketide chain elongation in the biosynthesis of their respective aglycone end products, 6-dEB (**1**) and 10-deoxymethynolide (**2**) or narbonolide (**3**), giving rise to the identical ACP-bound diketide, *syn*-(2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-ACP (Figure 1).[16]

For heterologous expression of the DEBS [KS][AT1] didomain, we modified pPK156, a pET28-derived plasmid in which the native DEBS AT-ACP loading didomain, located just upstream of DEBS module 1, had been replaced by the Type II ACP ZhuG.[17] The fusion of ZhuG upstream of the KS1 domain facilitated expression of stable, active PKS protein, consistent with the known preference for an N-terminal linker peptide upstream of the KS domain.[18] The KR1 and ACP1 domains of DEBS module 1 were deleted from pPK156 by excision of an ~2000-bp *Asc*I-*Eco*RI fragment that spanned from just upstream of the consensus 3′-end of the AT1 domain to the end of the module. This segment was then replaced by a pair of 48-bp synthetic oligonucleotides that restored the C-terminus of the AT1 domain up to the consensus **YR**I**EW** peptide junction site.[1a,19] The resulting plasmid, pAYC19, was used to express and purify DEBS [KS1][AT1] carrying an N-terminal His_6-tag, using procedures analogous to those employed for other [KS][AT] didomains.[19]

To express the homologous PICS [KS1][AT1] didomain, a synthetic gene optimized for expression in *E. coli* was designed in which the 44-amino acid N-terminal docking domain of PICS module 5 [M5(N)] was fused to the universally conserved Ala-Cys-Arg tripeptide junction at the N-terminus of the PICS KS1

(15) Yue, S.; Duncan, J. S.; Yamamoto, Y.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1987**, *109*, 1253–1255.

(16) (a) Cane, D. E.; Yang, C.-C. *J. Am. Chem. Soc.* **1987**, *109*, 1255–1257. (b) Cane, D. E.; Lambalot, R. H.; Prabhakaran, P. C.; Ott, W. R. *J. Am. Chem. Soc.* **1993**, *115*, 522–526.

(17) Marti, T.; Hu, Z.; Pohl, N. L.; Shah, A. N.; Khosla, C. *J. Biol. Chem.* **2000**, *275*, 33443–33448.

(18) Khosla, C.; Gokhale, R. S.; Jacobsen, J. R.; Cane, D. E. *Annu. Rev. Biochem.* **1999**, *68*, 219–253.

(19) (a) Kim, C. Y.; Alekseyev, V. Y.; Chen, A. Y.; Tang, Y.; Cane, D. E.; Khosla, C. *Biochemistry* **2004**, *43*, 13892–13898. (b) Chen, A. Y.; Schnarr, N. A.; Kim, C. Y.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **2006**, *128*, 3067–3074. (c) Tang, Y.; Kim, C. Y.; Mathews, I. I.; Cane, D. E.; Khosla, C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11124–11129.

domain site and ending at the C-terminus of PICS AT1 at the conserved Tyr-Arg junction.[1a,19] The resulting construct was ligated into the *Nco*I and *Xho*I sites of the high copy number plasmid pETBlue-2 to give the expression vector pRJL13. The corresponding protein, PICS [M5(N)][KS1][AT1], was expressed in *E. coli* Tuner with an C-terminal His_6-tag, yielding 6 mg of soluble protein/L of culture that was readily purified by Ni^{2+}-affinity chromatography.

The synthetic gene for PICS ACP1, spanning the typical ACP domain boundaries[1a,19] from Gly2477 to Gly2572 of PICS module 1, was cloned into the *Nde*I and *Eco*RI sites of pET28a to give pRJL04. The resulting recombinant PICS ACP1 carrying an N-terminal His_6-tag was obtained as the holo-enzyme after expression in *E. coli* BL21(DE3) BAP1 which harbors an integrated *sfp* gene encoding the *Bacillus subtilis* surfactin phosphopantetheinyl transferase. MALDI-TOF analysis confirmed the presence of the phosphopantetheinyl side chain, giving *m/z* 12 667 ± 12 (calcd for α-*N*-gluconoyl-His_6 tag ACP1 12,661).[20]

A synthetic gene with optimized *E. coli* codons encoding the 492 amino acid recombinant KR1, using the consensus domain boundaries[1a,19] from Ala1997 to Ala2488 of PICS module 1, was cloned into pET28a. The resulting vector pRJL01 was used to transform *E. coli* BL21(DE3) to give PICS KR1 with an N-terminal His_6-tag. Using the model substrate (±)-*trans*-decalone and NADPH, PICS KR1 displayed a k_{cat} of 0.22 ± 0.03 s^{-1}, similar to that observed for both DEBS KR1 and TYLS KR1, with a K_m of 14.8 ± 2.5 mM for *trans*-decalone. Reduction of the substrate analogue (±)-2-methyl-3-ketopentanoyl-SNAC showed a k_{cat} of 1.5 ± 0.3 s^{-1} and a K_m of 7.3 ± 2.4 mM, comparable to the steady-state kinetic parameters measured for DEBS KR1.[8a]

Determination of Ketoreductase Stereospecificity. The KS domains of both DEBS and PICS module 1 are normally primed by a propionyl-ACP thioester derived from free propionyl-CoA or methylmalonyl-CoA, respectively (Figure 1). The activity of recombinant DEBS [KS1][AT1] plus DEBS ACP1 was initially demonstrated by incubating the substrate analogue propionyl-SNAC, [2-^{14}C]methylmalonyl-CoA, and NADPH in combination with either DEBS KR1 or DEBS KR6. After quenching with aqueous base and acidification, TLC phosphorimaging revealed formation of a single radioactive product, R_f 0.59, that cochromatographed with the authentic diketide *syn*-2-methyl-3-hydroxypentanoic acid (**7a** or its enantiomer **7b**) (Supporting Information, Figure S1a). PICS KR1 gave the same apparent *syn*-diketide acid product, while TYLS KR1 gave a new product, R_f 0.57, corresponding to the predicted diastereomer, *anti*-2-methyl-3-hydroxypentanoic acid (**8a** or **8b**). Similar results were obtained with DEBS KR6, which gave a product, corresponding to **7b**, with the same R_f as that from DEBS KR1 or PICS KR1. Analogous incubations of propionyl-SNAC with PICS [KS1][AT1] plus PICS ACP1 and either PICS KR1 or TYLS KR1 also gave the predicted *syn*- or *anti*-diketide acids (**7a/7b, 8a/8b**) respectively. Interestingly, both DEBS KS3 and KS6, which normally process longer chain polyketide substrates, could also be primed by propionyl-SNAC. Thus incubation of propionyl-SNAC with either DEBS [KS3][AT3] plus ACP3 or DEBS [KS6][AT6] plus ACP6 in the presence of DEBS KR1 gave the expected *syn*-diketide acid (**7a** or **7b**), while the combination of DEBS [KS6][AT6], ACP6, and DEBS KR6 also gave the *syn*-diketide (**7a** or **7b**) (Figure S1b). Finally the combination of DEBS [KS6][AT6] plus ACP6 and TYLS KR1 gave the predicted *anti*-diketide acid (**8a** or **8b**).

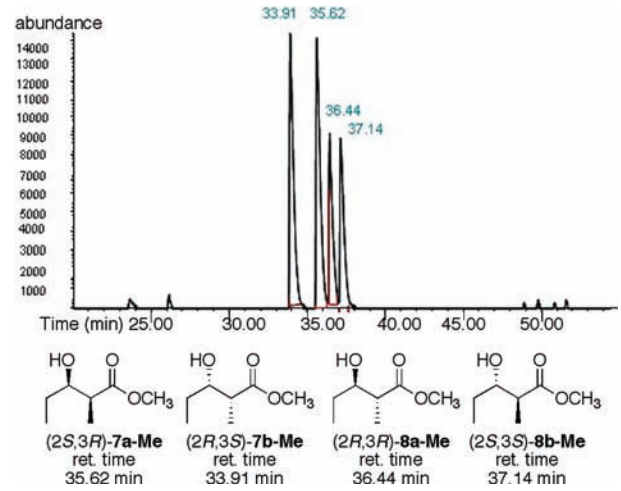


Figure 2. Chiral GC-EI-MS (XIC at *m/z* 88) of the mixture of (±)-2-methyl-3-hydroxypentanoate methyl esters and retention times of the individual diastereomers (**7a-Me**, **7b-Me**, **8a-Me**, and **8b-Me**).

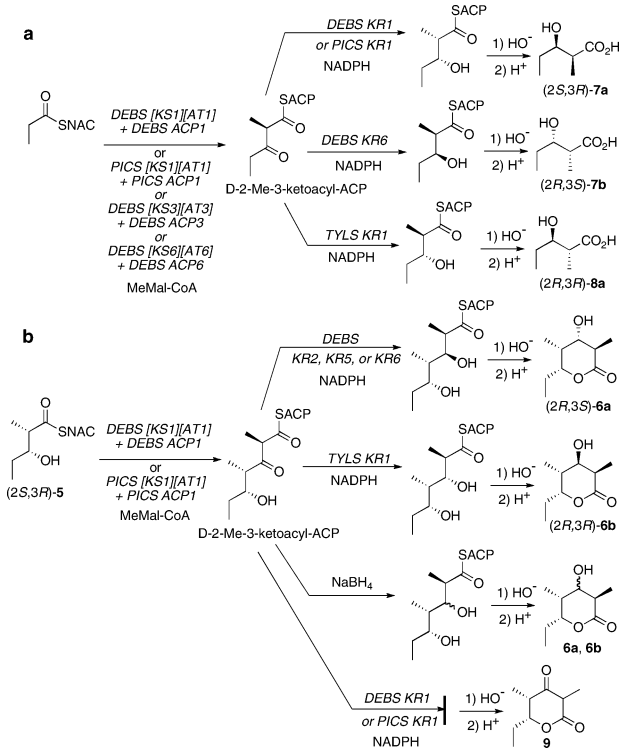
To determine the absolute configuration of each of the enzymatically generated diketide acids, we developed a convenient and sensitive protocol based on chiral GC-MS analysis of the derived methyl esters. Comparison of the experimentally measured retention times obtained using a capillary CP-Chirasil-DEX CB column as well as the corresponding mass spectra with those of synthetic samples of each of the four methyl 2-methyl-3-hydroxypentanoate diastereomers, **7a**, **7b**, **8a**, and **8b** (Figure 2)[21] allowed determination of the absolute configuration of the diketides resulting from each reconstituted PKS incubation mixture, as well as approximate measurement of the relative yield of each incubation reaction.

We then carried out a series of incubations using combinations of [KS][AT] didomains and their cognate ACP domains with a set of recombinant KR domains from both the parent module and heterologous PKS modules (Scheme 3, Table 1). In a typical incubation, DEBS [KS1][AT1] (40 μM) was preincubated for 1 h at 25 °C with 5 mM propionyl-SNAC, followed by addition of a mixture of DEBS ACP1 (200 μM), methylmalonyl-CoA (300 mM), DEBS KR1 (300 μM), and NADPH (2 mM). After an additional 1 h, the reaction was quenched by addition of aqueous base. Following hydrolysis and reacidification, the resulting diketide acid was treated with trimethylsilyl diazomethane, and the derived methyl ester was subjected to chiral GC-MS analysis, revealing the formation of the predicted *syn*-(2*S*,3*R*)-2-methyl-3-hydroxypentanoate (**7a**) as the major product (Scheme 3a, Table 1, Figure 3). Use of the corresponding PICS KR1 domain gave exclusively the same *syn*-(2*S*,3*R*)-diketide **7a**, while DEBS KR6 yielded predominantly the enantiomeric *syn*-(2*R*,3*S*)-diketide **7b**. As predicted, TYLS KR1 produced the diastereomeric *anti*-(2*R*,3*R*)-diketide acid **8a**. Similarly, when the PICS [KS1][AT1] plus PICS ACP1 pair was used, both DEBS KR1 and PICS KR1 generated the exclusively same *syn*-

(20) Geoghegan, K. F.; Dixon, H. B.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. *Anal. Biochem.* **1999**, *267*, 169–184.

(21) (a) Smith, A. B., III; Levenberg, P. A. *Synthesis* **1981**, 567–570. (b) Evans, D. A.; Kim, A. S.; Metternich, R.; Novack, V. J. *J. Am. Chem. Soc.* **1998**, *120*, 5921–5942. (c) Vicario, J. L.; Badia, D.; Dominguez, E.; Rodriguez, M.; Carrillo, L. *J. Org. Chem.* **2000**, *65*, 3754–3760. (d) Raimundo, B. C.; Heathcock, C. H. *Synlett.* **1995**, 1213–1214.

Scheme 3. Stereochemistry of KR-Catalyzed Reduction of 2-Methyl-3-ketoacyl-ACP Intermediates Generated by Combinations of PKS [KS][AT] and ACP Domains, As Determined by Chiral GC-MS Analysis of the Resulting Diketides or Triketides[a]



[a] (a) Use of propionyl-SNAC as primer. (b) Use of diketide-SNAC **5** as primer.

(2*S*,3*R*)-diketide acid **7a**, while DEBS KR6 and TYLS KR1 yielded the respective *syn*-(2*R*,3*S*)-**7b** and *anti*-(2*R*,3*R*)-**8a** diketides. Analogous incubations were also carried out using the paired DEBS [KS3][AT3]+ACP3 or [KS6][AT6]+ACP6 proteins in combination with the same set of recombinant KR domains. In each case, the stereochemistry of the exclusive or predominant diastereomer of the diketide acid that was formed was directly correlated with the specificity of the relevant KR domain, independent of which [KS][AT]+ACP combination was used to generate the initial 2-methyl-3-ketopentanoyl-ACP intermediate. Depending on the specific [KS][AT]+KR combination, quantitation of the yields of diketide acid from each incubation indicated catalysis of 2−4 turnovers by each KS domain, corresponding to 40−80% of the theoretical maximum, based on the concentration of the ACP domain, which was present in 5-fold excess over the [KS][AT] didomain.

Interestingly, we found that both DEBS [KS1][AT1] and PICS [KS1][AT1] could also utilize the diketide (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-SNAC (**5**) as a surrogate substrate (Scheme 3b). Thus when either the DEBS or the PICS [KS1][AT1] didomain and its cognate ACP1 domain was incubated with **5**, methylmalonyl-CoA, and NADPH in the presence of DEBS KR2, KR5, or KR6 followed by GC-MS analysis of the derived triketide lactones, as previously described,[14] the exclusively formed triketide lactone product was the predicted **6a**, while TYLS KR1 gave the diastereomeric triketide lactone **6b**. Neither DEBS KR1 nor PICS KR1 was able to reduce the ACP-bound 3-ketoacyltriketide intermediates, giving only unreduced triketide ketolactone **9**, consistent with the previously observed inability of DEBS KR1 to reduce triketide substrates.[22] In control experiments, sodium borohydride trapping of the unreduced ACP-bound triketide that was generated from diketide **5** by the DEBS [KS1][AT1]+ACP1 system followed by analysis of the derived triketide lactones **6a** and **6b** established that the intermediate consisted of >85% unepimerized 2-D-methyl-3-ketoacyl triketide (Scheme 3b). These results were similar to those obtained for both DEBS [KS3][AT3] and DEBS [KS6][AT6].[14a] Attempted borohydride trapping of the corresponding 2-methyl-3-ketodiketide intermediate generated when propionyl-SNAC was used as the substrate was unsuccessful.

To address the actual timing of the methyl group epimerization, we also carried out a series of time course experiments using propionyl-SNAC and the DEBS [KS6][AT6]+ACP6 pair in combination with either DEBS KR6 or DEBS KR1 as the reductase component. Several aliquots from each incubation mixture were withdrawn and quenched during the nominally linear reaction phase of the reaction, between 4 and 15 min. The remaining reaction mixture was quenched at the end of 60 min to determine total product formation for each incubation. The yield of the derived methyl esters of (2*R*,3*S*)-**7b** or (2*S*,3*R*)-**7a** from each quench was analyzed by calibrated GC-MS (Figures S2 and S3). In spite of the fact that DEBS [KS6][AT6] initially generates exclusively the ACP-bound D-2-methyl-3-ketopentanoyl intermediate and does not itself support epimerization,[14a] there was no more than a 2-fold difference between the initial rate of product formation for the DEBS KR6 domain and that with the DEBS KR1 domain. Moreover, after 60 min the total yields of the respective reduced diketide acid products, **7a** and **7b**, were also within a factor of 2 of one another.

Discussion

The discovery that the biosynthesis of many complex polyketides is controlled by modular polyketide syntheses has already provided numerous invaluable insights into the molecular basis for the programming of polyketide biosynthesis, providing an explanation for the number, type, and organization of polyketide building blocks as well as the control of the functionality and oxidation level at each stage of polyketide chain elongation.[18] Nonetheless, until recently, the biochemical basis for the control of one of the most remarkable structural features in all macrolides and polyethers, the number and distribution of stereogenic centers in each polyketide, has remained obscure.

As summarized in Schemes 2 and 3 and detailed in Table 1, the stereochemistry of the 2-methyl-3-hydroxy diketide or triketide observed for each reconstituted PKS incubation is strictly correlated with the intrinsic diastereospecificity of the KR domain, independent of the origin of the [KS][AT]+ACP chain elongation pair. Thus the DEBS KR1 and PICS KR1 domains both generate the reduced *syn*-(2*S*,3*R*)-2-methyl-3-hydroxypentanoate **7a** having an epimerized L-2-methyl group while DEBS KR2, KR5, and KR6 each produce the enantiomeric *syn*-(2*R*,3*S*)-2-methyl-3-hydroxypentanoate (**7b**) with an unepimerized D-2-methyl. Similarly, TYLS KR1 generates the diastereomeric *anti*-(2*R*,3*R*)-2-methyl-3-hydroxypentanoate (**8a**), again with an unepimerized methyl group. In all these cases, the experimentally observed configurations of both the hydroxyl and the methyl groups for any recombinant KR domain correspond to those in the natural reduced polyketide intermedi-

(22) Chen, A. Y.; Cane, D. E.; Khosla, C. *Chem. Biol.* **2007**, *14*, 784–792.

Table 1. Incubation of Propionyl-SNAC with Combinations of [KS][AT], ACP, and KR Domains in the Presence of Methylmalonyl-CoA and NADPH[a]

PKS domains			diketide acid			
[KS][AT]	**ACP**	**KR**	**(2*S*,3*R*)-7a (%)**	**(2*R*,3*S*)-7b (%)**	**(2*R*,3*R*)-8a (%)**	**(2*S*,3*S*)-8b (%)**
DEBS [KS1][AT1]	DEBS ACP1	DEBS KR1	78	10	8	4
PICS [KS1][AT1]	PICS ACP1	DEBS KR1	98	2	0	0
DEBS [KS3][AT3]	DEBS ACP3	DEBS KR1	>99	<1	0	0
DEBS [KS6][AT6]	DEBS ACP6	DEBS KR1	92	8	0	0
DEBS [KS1][AT1]	DEBS ACP1	PICS KR1	>99	<1	0	0
PICS [KS1][AT1]	PICS ACP1	PICS KR1	>99	<1	0	0
DEBS [KS3][AT3]	DEBS ACP3	PICS KR1	>99	<1	0	0
DEBS [KS6][AT6]	DEBS ACP6	PICS KR1	>99	<1	0	0
DEBS [KS1][AT1]	DEBS ACP1	DEBS KR6	0	88	0	12
DEBS [KS3][AT3]	DEBS ACP3	DEBS KR6	0	92	0	8
DEBS [KS6][AT6]	DEBS ACP6	DEBS KR6	0	96	0	4
PICS [KS1][AT1]	PICS ACP1	DEBS KR6	0	99	0	1
DEBS [KS1][AT1]	DEBS ACP1	TYLS KR1	12	0	88	0
DEBS [KS3][AT3]	DEBS ACP3	TYLS KR1	6	0	94	0
DEBS [KS6][AT6]	DEBS ACP6	TYLS KR1	5	0	95	0
PICS [KS1][AT1]	PICS ACP1	TYLS KR1	17	<1	79	4

[a] Stereoisomer distribution of diketide acid product determined by chiral GC-XIC-MS analysis.

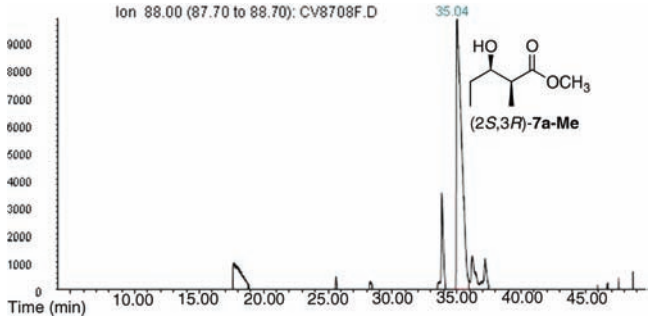


Figure 3. Chiral GC-EI-MS (XIC at *m*/*z* 88) of methyl (2*S*,3*R*)-2-methyl-3-hydroxypentanoate (**7a-Me**, ret. time 35.04 min) produced by incubation of propionyl-SNAC with DEBS [KS1][AT1], DEBS ACP1, and DEBS KR1 in the presence of methylmalonyl-CoA and NADPH.

ate generated by the parent PKS module. Each KR domain therefore exercises control over both the *choice* of which epimer of the ACP-bound 2-methyl-3-ketoacyl substrate to reduce and the *stereochemistry* of the reduction itself, thereby completely controlling which diastereomer of the reduced product is generated.

Since DEBS KS1, KS3, and KS6 and PICS KS1 each initially generate only an unepimerized D-2-methyl-3-ketoacyl-ACP intermediate, methyl group epimerization, if and when it does occur, must therefore take place *following* KS-catalyzed condensation (Scheme 1, Pathway A). Which domain is responsible for catalysis of the actual epimerization? Several lines of evidence now point strongly to the KR domain itself.

In the absence of an added recombinant KR domain, the ACP-bound D-2-methyl-3-ketoacyl intermediates generated by the various KS-AT didomains are configurationally remarkably stable. No more than 5−15% conversion to L-2-methyl epimer can be detected by sodium borohydride trapping, even after 1 h of further incubation.[14a] The estimated first-order rate of epimerization of ACP-bound substrate in the presence of the DEBS KS1, KS3, or KS6 domains should therefore have an upper limit of 0.01 min^{-1}, considerably slower than the rate of 0.15 min^{-1} measured for buffer-catalyzed deuterium exchange of untethered methyl 2-methylacetoacetate.[14a] This enhanced configurational stability of ACP-bound ketoacyl intermediates is most likely due to sequestering of the bound polyketide in the cleft between ACP helices 2 and 3.[23] Whatever the explanation, however, the background epimerization rate of the ACP-bound 2-methyl-3-ketoacylthioester is therefore 10−100-fold slower than the typical k_{cat} of ∼1 min^{-1} for KS-catalyzed

(23) Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. *J. Mol. Biol.* **2007**, *365*, 135–145.

condensation.[24] Although the rates of reduction of ACP-bound substrates by free KR domains have not been determined directly, the k_{cat} for reduction of the 2-methyl-3-ketopentanoyl-SNAC analogue by either DEBS KR1 or PICS KR1 is ~50 min^{-1}, 25−50 times that for the KS-catalyzed condensation reaction and 10^3-10^4 times faster than background epimerization of the ACP-bound 2-methyl-3-ketoacyl triketide intermediate. The net rate constant for the coupled condensation and reduction reaction mediated by the DEBS [KS6][AT6] plus DEBS KR6 system should therefore correspond to the limiting rate of KS-catalyzed condensation. By contrast, in the absence of catalysis of epimerization by the KR1 domain itself, the net rate for the coupled condensation, epimerization, and reduction reactions associated with the DEBS [KS6][AT6] plus DEBS KR1 system would be dominated by the rate of the slowest step, background buffer-catalyzed epimerization. The fact that there is no more than a 2-fold difference observed in the initial rate of formation of the unepimerized and the epimerized reduction products **7a** and **7b** in the presence of DEBS KR1 and KR6, respectively (Figures S2 and S3), strongly suggests that DEBS KR1 indeed catalyzes both the epimerization of ACP-bound D-2-methyl-3-ketopentanoate and the subsequent reduction of the L-2-methyl-3-ketopentanoyl-ACP.

This postulated dual role of the KR1 domain in mediating epimerization is also consistent with independent analysis of ketoreductase structures. Keatinge-Clay has determined the structures of both DEBS KR1 and TYLS KR1 and has modeled epimerized (2*S*,3*R*)-diketide and the nicotinamide portion of the NADPH cofactor into the active sites of both proteins to define the most important residues involved in substrate binding and in catalysis.[7] These key protein regions were then used to identify homologous amino acid sequences in a set of ~20 PKS KR domains that were grouped according to which of the four possible 2-methyl-3-hydroxyacyl-ACP diastereomers is normally generated by each native KR domain. Based on this analysis, a small number of residues in the catalytic region were tentatively correlated with the formation of products with epimerized methyl groups, consistent with the notion that it is indeed the KR domain that is responsible for catalysis of the epimerization itself. The interesting suggestion was also made that the nominally reductase-inactive KR3° domain of DEBS module 3 may be responsible for epimerization of the methyl group in the corresponding ACP3-bound D-2-methyl-3-ketoacyl tetraketide product, thereby setting the configuration of the L-8-methyl of 6-dEB.[7b] We have previously reported, however, that addition of recombinant KR3° to an incubation mixture containing DEBS [KS6][AT6]+ACP6 followed by quenching with $NaBH_4$ gave the same mixture of triketide lactones **6a** and **6b** as were obtained in the absence of KR3°, with no detectable additional epimerization of the C-2 methyl group.[14a] Although these results do not support the proposed role of KR3° as an epimerase, we had no independent method of knowing whether the recombinant protein was indeed catalytically competent. The origin of the epimerase activity apparently associated with DEBS module 3 therefore remains unresolved.

Our results conclusively establish that in the biosynthesis of complex polyketides the KS-catalyzed condensation and chain elongation reaction take place prior to any eventual epimerization of the methyl substituent. The configurations of the vicinal methyl and hydroxyl groups in the observed reduction product from each combinatorial PKS incubation are strictly correlated with the origin of the specific KR domain that mediates the reduction. Each ketoreductase domain is therefore completely diastereoselective, reducing only one of two possible methyl epimers of its 2-methyl-3-ketoacyl-ACP substrate so as to yield a single diastereomer of the resulting product. Thus, both the DEBS [KS1][AT1] and PICS [KS1][AT1] didomains originate from modules that normally mediate the formation of a reduced diketide product with an epimerized L-2-methyl group (Figure 1). The DEBS [KS3][AT3] didomain is derived from a module that normally generates an unreduced tetraketide intermediate that at some undetermined stage undergoes epimerization to give an epimerized L-methyl group, while the [KS6][AT6] didomain corresponds to a module that naturally yields a heptaketide product with an unepimerized D-2-methyl group. Regardless of their modular origins, each of the individual [KS][AT] didomains initially generates exclusively D-2-methyl-3-ketoacyl-ACP diketide and triketide intermediates when incubated with its cognate ACP domain and propionyl-SNAC or (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-SNAC, respectively. Moreover, these initially generated D-2-methyl-3-ketoacyl-ACP products undergo negligible epimerization in the absence of an added recombinant KR domain, even after a 1 h interval, indicating that KS1, KS3, or KS6 alone is not able to catalyze methyl group epimerization. This critical epimerization is most likely catalyzed by the KR1 domain itself. Further experiments are in progress to confirm the proposed epimerase activity of the KR1 domains and to elucidate the still unknown mechanism of epimerization.

Our findings also have important implications for the rational engineering of polyketide synthases.[1] Early efforts to construct chimeric PKS modules relied primarily on the utilization of presumptive domain boundaries inferred from analysis of multiple sequence alignments as well as on the exploitation of convenient nearby restriction sites.[18] More recently, the determination of the structures of several PKS domains and the demonstration of the biochemical activity of individual recombinant dissected domains have provided more direct insights into actual PKS domain boundaries.[7,19,25] In addition, both biochemical and structural studies have led to a more sophisticated understanding of the protein structural basis for intermodular communication.[1,26] The newest findings as to how and when the complex stereochemistry of polyketide biosynthesis is controlled represent another important step toward the goal of rationally designing chimeric PKS domains with targeted biosynthetic function.

Experimental Section

Materials. DL-[2-^{14}C]-Methylmalonyl-CoA was from American Radiolabeled Chemicals. Isopropylthio-D-galactopyranoside (IPTG) was purchased from Invitrogen. All other reagents and substrates of the highest grade available were purchased from Sigma-Aldrich and were used without further purification. (2*S*,3*R*)-2-Methyl-3-hydroxypentanoyl-*N*-acetylcysteamine thioester (**5**) was prepared according to established methods.[27] (±)-2-Methyl-3-ketopentanoyl-*N*-acetylcysteamine thioester was synthesized following the reported

(24) Wu, N.; Kudo, F.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **2000**, *122*, 4847–4852.

(25) (a) Tang, Y.; Chen, A. Y.; Kim, C. Y.; Cane, D. E.; Khosla, C. *Chem. Biol.* **2007**, *14*, 931–943. (b) Keatinge-Clay, A. *J. Mol. Biol.* **2008**, *384*, 941–953. (c) Alekseyev, V. Y.; Liu, C. W.; Cane, D. E.; Puglisi, J. D.; Khosla, C. *Protein Sci.* **2007**, *16*, 2093–2107.

(26) Broadhurst, R. W.; Nietlispach, D.; Wheatcroft, M. P.; Leadlay, P. F.; Weissman, K. J. *Chem. Biol.* **2003**, *10*, 723–731.

(27) Cane, D. E.; Tan, W.; Ott, W. R. *J. Am. Chem. Soc.* **1993**, *115*, 527–535.

procedure,[28] and propionyl-*N*-acetylcysteamine thioester was prepared as previously described.[29] Reference standards of each of the four diastereomers of triketide lactone **6** and the corresponding TMS-derivative were prepared and analyzed as described.[14a] Thin-layer chromatography (TLC) plates were from J. T. Baker. SDS PAGE-gradient gels were from BioRad. Ni-NTA affinity resin was from Qiagen. HiTrap-Q anion exchange columns were from Amersham Pharmacia. PD-10 columns were from GE Healthcare, and ultrafiltration devices (Amicon Ultra-15 10000 and 30000 MWCO) were purchased from Millipore. Protein concentration was determined by Bradford assay using bovine serum albumin as standard. The pETBlue-2 vector and *E. coli* Tuner strain were purchased from Novagen. Synthetic genes with codons optimized for expression in *E. coli* were purchased from DNA2.0.

Methods. General methods were as previously described.[14a] ^{1}H and ^{13}C NMR spectra were obtained on Bruker Advance AM300 and AM400 spectrometers. [^{14}C]-Labeled products were visualized using a BioRad FX-Pro Plus Molecular Imager. Optical rotation was measured using a Jasco polarimeter P-1010. Ketoreductase assays were carried out using a Tecan Infinity microplate reader. Protein MALDI-TOF measurements were performed on an Applied Biosystems Voyager DE PRO MALDI-TOF benchtop mass spectrometer.

Synthesis of 2-Methyl-3-hydroxypentanoic Acid Diastereomers. A mixture of all four diastereomers of 2-methy-3-hydroxypentanoic, **7a**, **7b**, **8a**, and **8b**, was prepared by hydrolysis of methyl 2-methyl-3-hydroxypentanoate obtained by aldol condensation between methyl propanoate and propionaldehyde in the presence of lithium diisopropylamide.[21a] *syn*-(2*S*,3*R*)-**7a** and (2*R*,3*S*)-2-methyl-3-hydroxypentanoic acid (**7b**) were each prepared by Evans asymmetric *syn*-selective aldol reaction followed by treatment with lithium hydroperoxide.[21b,30] (2*S*,3*R*)-**7a** and (2*R*,3*S*)-**7b** produced in high yield were each diastereomerically pure, as shown by ^{1}H NMR analysis,[27,31] and enantiomerically homogeneous as confirmed by optical rotation: $[\alpha]^{25}_D$ +4.1 (*c* 2.1, Et_2O) for (2*S*,3*R*)-**7a** and $[\alpha]^{25}_D$ −4.2 (*c* 2.01, Et_2O) for (2*R*,3*S*)-2-methyl-3-hydroxypentanoic acid (**7b**).[21c] The diastereomer (2*R*,3*R*)-**8a** was synthesized according to the established stereoselective route[21d] involving an asymmetric aldol reaction in which the use of 2 equiv of Lewis acid *n*-Bu_2BOTf favored the formation of the *anti* product over the *syn*. (2*R*,3*R*)-**8a** was obtained in a ratio of 4:1 along with (2*R*,3*S*)-**7b**, in agreement with the results of previous studies. The two compounds were easily distinguished by GC-MS analysis and characterized by ^{1}H NMR.

Chiral GC-MS Analysis of Diketide Acid Diastereomers. The synthetic mixture of racemic 2-methy-3-hydroxypentanoic acids was converted to the corresponding methyl esters **7a-Me**, **7b-Me**, **8a-Me**, and **8b-Me**, which could be resolved by chiral capillary GC-MS (Figures 2 and 3). GC-MS spectra were recorded on a GC-MS Hewlett-Packard Series 2 GC-MSD, 70 eV EI in positive ion mode, using a Varian CP-Chirasil-DEX CB (25 m × 0.32 mm) and a temperature program of (1) initial temp 50 °C for 2 min, (2) increase at rate of 1 °C/min up to 90 °C, and then (3) 20 °C/min to final temp of 200 °C. The individual components were rigorously identified by a separate injection of authentic sample of the methyl esters of (2*S*,3*R*)-**7a**, (2*R*,3*S*)-**7b**, and (2*R*,3*R*)-**8a** prepared by dissolving 20 μg of each authentic acid in 100 μL of MeOH and treatment with 20 μL of TMS-diazomethane (2 M in hexanes) (Figure 2).

Cloning of DEBS [KS1][AT1]. Plasmid pPK156, harboring the gene for DEBS module 1 (KS1 start of AQRET) fused downstream of a His_6-tag and the gene for the Type II ACP ZhuG (MGSSHHHHHHSSGLVPRGSHMDPFTLDDLKRLIDACVGTDDAVQLDETGAATPFLDLGLDSLAVYEVVTRIQDERGVAISDDDIDGLETPRDMVAFVNGLLVETAGDA), was doubly digested with *Asc*I and *Eco*RI to excise an ~2000 bp fragment running from just upstream of the consensus end of the AT1 domain (nt 4905 of *eryAI*) to the end of the module. This segment, which had encompassed the KR1 and ACP1 domains, was replaced by annealing two complementary oligonucleotides, 5′-CGCGCCGGTCCACCGAGGTCGACGAGGTTTCCGCGCTGCGCTACCGCATG-3′ and 5′-AATTCATGCGGTAGCGCAGCGCGGAAACCTCGTCGACCTCGGTGGACCGG-3′, to give a 48-bp double-stranded oligo with overhangs complementary to the *Asc*I and *Eco*RI sticky ends which was then cloned into the corresponding *Asc*I and *Eco*RI sites in the digested pPK156 to give pAYC19.

Cloning of DEBS KR5. The DNA sequence from *EryAIII* encoding DEBS KR5 was amplified by PCR using the primers 5′-AAAAACATATGGACGACTGGCGCTACCAGGTCG-3′ and 5′-TTTTGAATTCAAGAGCCGCTGGGCGAGCGCCGG-3′ carrying *Nde*I and *Eco*RI restrictions (underlined), and the resultant amplicon was ligated into the corresponding sites of pET28a to give the KR5 expression vector pAYC61.

Cloning of PICS [KS1][AT1] from PICS Module 1. For the cloning of the PICS [KS1][AT1] didomain from *PikAI* (Genbank accession number AF079138), a synthetic gene with codons optimized for expression in *E. coli* was constructed based on a fusion of the N-terminal docking domain of PICS module 5 [M5(N)] (Met1-Met45 of PikAIII ending in the highly conserved domain boundary PAIVGM) to the PICS [KS1][AT1] didomain (Ala1117 of PikAI at the AlaCysArg trypsin cleavage site to the consensus N-terminal trypsin site Arg2005). The synthetic gene included a 5′-CCATGGGTGCC sequence corresponding to an *Nco*I site that included the ATG start codon while inserting an extra Gly residue so as to encode MetGlyAla in place of the natural MetAla at the N-terminus of the PICS module 5 docking domain. The AlaCysArg tripeptide was encoded by GCCTGCAGA that incorporated a *Pst*I site so as to allow optional swapping of the N-terminal docking domain as required. (Use of the alternative 43 amino acid PICS module 3 N-terminal docking domain [M3(N)] gave significantly lower yields of active protein.) At the 3′-end of the synthetic gene was an *Xho*I site. The *Nco*I- and *Xho*I-digested gene was inserted into the corresponding sites of the pUC-based expression vector pETBlue-2 to give plasmid pRJL13, which was used to transform competent cells of the expression host *E. coli* Tuner.

Cloning of PICS ACP1. The synthetic gene for PICS ACP1 from PICS module 1, optimized for expression in *E. coli*, corresponded to Gly2477 to Gly2572 of the PiKAI protein. An *Nde*I site at the 5′-end appended the requisite N-terminal methionine, and an *Eco*RI site immediately followed the stop codon. The doubly digested construct was inserted into the pET28a vector to generate plasmid pRJL04 which was used to transform competent cells of *E. coli* BL21(DE3) BAP1 harboring the *sfp* gene.[32]

Cloning of PICS KR1. The synthetic gene for PICS KR1 from PICS module 1, optimized for expression in *E. coli*, corresponded to Ala1997 to Ala2488 of the PiKAI protein. An *Nde*I site at the 5′-end appended the requisite N-terminal Met, and an *Eco*RI site immediately followed the stop codon. The doubly digested construct was inserted into the pET28a vector to generate plasmid pRJL01 which was introduced into *E. coli* BL21(DE3).

Expression and Purification of Recombinant PKS Domains. The TYLS KR1 expression plasmid was a gift from Prof. Adrian Keatinge-Clay. The expression and purification of DEBS [KS3][AT3], [KS6][AT6], ACP1, ACP3, ACP6, KR1, KR2, KR6, and TYLS KR1 have been previously described.[7b,19b,22] DEBS KR5, PICS KR1, and PICS ACP1 were each expressed and purified following a procedure analogous to those previously described for other recombinant PKS domains.[14a,19] For expression of PICS [M5(N)]-[KS1][AT1], Luria−Bertani (LB) broth (5 mL) containing 100

(28) Sharma, K. K.; Boddy, C. N. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3034–3037.
(29) Block, M. H.; Cane, D. E. *J. Org. Chem.* **1988**, *53*, 4923–4928.
(30) Evans, D. A.; Britton, T. C.; Ellman, J. A. *Tetrahedron Lett.* **1987**, *28*, 6141–6144.
(31) Gil, S.; Parra, M.; Rodriguez, P.; Sotoca, E. *Synthesis* **2005**, 3451–3455.
(32) Walsh, C. T.; Gehring, A. M.; Weinreb, P. H.; Quadri, L. E.; Flugel, R. S. *Curr. Opin. Chem. Biol.* **1997**, *1*, 309–315.

μg/mL carbenicillin was inoculated with a fresh glycerol stock of *E. coli* Tuner/pRJL13 and incubated at 230 rpm overnight at 37 °C. Fresh LB (8 × 5 mL) containing the appropriate antibiotic was inoculated with 100 μL of the overnight culture and grown under the same conditions for 6 h. These cultures were used to inoculate larger baffled flasks (10 mL per 0.5 L of SuperBroth [3.2% Tryptone, 2.0% g yeast extract, 0.5% NaCl, pH 7.5] supplemented with 100 μg/mL carbenicillin), and these were grown under the same conditions until the OD_{600} reached 0.6. The cultures were cooled to 18 °C, and IPTG was added to a final concentration of 0.2 mM. After 18 h of further incubation, the cells were harvested by centrifugation (5000*g*, 10 min), washed in Tris buffer (50 mM Tris buffer, 1 mM EDTA, pH 8.0), and frozen until required. The frozen cell pellets were thawed on ice in 25 mL of lysis buffer (50 mM NaH_2PO_4, 10 mM imidazole, 300 mM NaCl, 10% glycerol, pH 8.0) containing 18 μL of protease inhibitors (1 mg/mL pepstatin, 1 mg/mL leupeptin, and 50 mg/mL benzamidine). The resuspended cells were passed through a French Press prior to sonication (3 × 30 s pulses of 50% amplitude interspaced by 90 s on ice). The cell debris was removed by centrifugation (10 000*g*, 50 min, 4 °C). The clarified lysate was slurried with pre-equilibrated Ni-NTA resin (5 mL) for 2 h at 4 °C. The slurry was poured into a column, and the nonbound proteins were removed. The column was washed with 5 column volumes of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 40 mM imidazole, 10% glycerol, pH 7.6) prior to eluting the protein in 3 column volumes of elution buffer (50 mM NaH_2PO_4, 50 mM NaCl, 150 mM imidazole, 10% glycerol, pH 7.4). Fractions of 4 mL in size were collected. Fractions containing the protein of interest were collated and buffer exchanged into 50 mM NaH_2PO_4, pH 7.2 via ultrafiltration. The sample was loaded onto a Q-sepharose anion exchange column pre-equilibrated with 3 column volumes of binding buffer (50 mM NaH_2PO_4, pH 7.2, 10% v/v glycerol). After being washed to remove the nonbound fractions (3 column volumes of binding buffer) the protein was eluted with a linear salt gradient (100−400 mM NaCl, 50 mM NaH_2PO_4, pH 7.2, 10% v/v glycerol), and 4.0 mL fractions were collected. Fractions containing the protein were collated and buffer exchanged into storage buffer (100 mM NaH_2PO_4, 10% v/v glycerol, pH 7.2) via ultrafiltration. After further concentration to ~0.25 mL the proteins were flash frozen in liquid nitrogen and stored in 20-μL aliquots at −80 °C until required.

Kinetic Assay of PICS KR1. The change in the absorption of the cofactor NADPH was followed at 340 nm over time by use of a microplate plate reader and Falcon polystyrene 96-well plates. PICS KR1 protein was diluted in assay buffer (50 mM NaH_2PO_4, pH 7.2) to give a stock solution of 30 μM. The substrate *trans*-1-decalone or (±)-(2*RS*)-2-methyl-3-ketopentanoic acid *N*-acetylcysteamine thioester was diluted with DMSO so as to give substrate concentrations of between 0 and 8 mM in the final mixture. The total amount of DMSO was kept constant at 5% v/v (16 μL). In one well, substrate (16 μL) and enzyme stock solution (50 μL) were added to the buffer (total vol. 250 μL). These were mixed by gently pipetting up and down. To initiate the reaction NADPH (50 μL of a 2 mM stock) was added, the solutions were quickly mixed by pipetting up and down 3 times, and then the absorbance was read at 340 nm every 10 s for 30 min. The final volume of 300 μL corresponded to a path length of 1 cm. Initial rates were calculated by fitting the linear portion of the absorbance progress curve (typically in the first 200 s at high substrate concentrations). The steady-state kinetic parameters were calculated by fitting the rate and substrate concentration data to the Michaelis−Menten equation by nonlinear least-squares regression using the Origin program in the Microcal kinetics package. Reported standard deviations in the steady-state kinetic parameters represent the calculated statistical errors in the nonlinear, least-squares regression analysis.

TLC-Phosphorimaging Assay of Enzymatic Formation of Diketide Acids. Radio-TLC assays of the conversion of propionyl-SNAC or (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-SNAC (**5**) to diketides or triketides, respectively, by reconstituted PKS domains were carried out as previously described.[14a] In a typical incubation, a [KS][AT] didomain (40 μM) was incubated with 5 mM propionyl-SNAC (added as stock solution in DMSO) and 2.5 mM of TCEP in a reaction volume of 50 μL (100 mM sodium phosphate buffer, pH 7.2) for 1 h at 25 °C. The paired ACP domain (200 μM) from the same PKS module as the [KS][AT] didomain, plus the KR domain (300 μM) and DL-[2-^{14}C]-methylmalonyl-CoA (300 μM), was then added in the presence of NADPH (2 mM). The reaction mixture was held at 25 °C for an additional 1 h. Addition of a NaOH solution (20 μL, 0.5 M) and incubation for 30 min at 65 °C allowed diketide release from the ACP. The basic mixture was treated with HCl solution (20 μL, 1 M) to allow ethyl acetate (3 × 100 μL) extraction of the enzymatically generated products. After solvent removal by Speed-Vac, the concentrated organic extracts were resuspended in 20 μL of ethyl acetate and spotted onto a TLC plate (silica gel; EtOAc/DCM 9;1, 0.1% AcOH) and visualized by phosphorimaging. Direct comparison with a 2-methyl-3-hydroxy pentanoic acid standard allowed confirmation of the formation of the expected product (Figure S1).

Incubation of Reconstituted Mixtures of PKS Domains. The protocol for assaying reconstituted mixtures of PKS components was based on the previously developed procedures.[28] In a typical incubation procedure, DEBS [KS1][AT1] (40 μM) was preincubated with 5 mM propionyl-SNAC in a total volume of 500 μL of phosphate buffer (100 mM, pH 7.2) containing tris-2-carboxyethyl-phosphine (TCEP, 2.5 mM). After 1 h at 25 °C DEBS ACP1 (200 μM), methylmalonyl-CoA (300 μM), DEBS KR1 (300 μM), and NADPH (2 mM) were added. (The incubation mixture also contained ~1% glycerol from the various protein solutions.) The enzyme mixture was incubated at 25 °C for an additional hour, and the reaction was then quenched by addition of 200 μL of 0.5 M sodium hydroxide solution and heated to 65 °C for 20 min before acidification with 200 μL of 1 M aq. hydrochloric acid. The solution was extracted with ethyl acetate (3 × 600 μL), and the concentrated organic extract was analyzed by GC-MS, as described below. Analogous procedures were also used for incubations using (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-SNAC (**5**) as substrate in place of propionyl-SNAC, with analysis of the derived triketide lactone.

GC-MS Analysis of Enzymatically Generated Diketide Acids. The concentrated organic extracts of the above-described enzymatic incubations, containing the diketide acid products, were dissolved in 100 μL of MeOH and then treated with 20 μL of trimethylsilyldiazomethane (2 M in hexane) for 5 min at 25 °C. The derived methyl esters were directly analyzed by chiral GC-MS as described for the diketide methyl ester standards (Figures 2 and 3). Both the total ion current (TIC) and extracted ion current (XIC) of the chromatograms were analyzed, with the XIC set on the base peak at 88 *m/z* of 2-methyl-3-hydroxypentanoic acid methyl ester. Use of the XIC method minimized interference from any side products resulting from the incubation and subsequent derivatization, allowing facile detection of total diketide methyl ester down to 10 nmol (0.08 nmol per 1-μL injection). Since changes in the concentration of individual samples resulted in small variations in the precise retention times, the identification of each enzymatically generated diketide acid was confirmed by coinjection of a sample of the corresponding authentic diastereomer of **7-Me** or **8-Me**.

Time Course of Diketide Acid Formation. DEBS [KS6][AT6] (40 μM) protein in a total volume of 2 mL of 100 mM phosphate buffer, pH 7.2, was incubated with 5 mM propionyl-SNAC and 2.5 mM TCEP for 1 h at 25 °C to acylate the KS domain. The recombinant KR protein, either DEBS KR1 or DEBS KR6 (300 μM), and ACP6 (30 μM) were then added along with 300 μM methylmalonyl-CoA and 2 mM NADPH. For each time point between 4 and 60 min a 500-μL aliquot was withdrawn and mixed

with 200 μL of 0.5 M sodium hydroxide to quench the reaction, followed by heating the mixture for 20 min at 65 °C to release the free diketide acid. Hydrochloric acid (200 μL of 1 M) was then added, the mixture was extracted with ethyl acetate, and organic solvent was removed by rotary evaporation. The crude extract was dissolved in 100 μL of methanol and treated with TMS diazomethane (20 μL) and analyzed by GC-EI-MS, as above (Figures 2 and 3).

Acknowledgment. We thank Dr. Tun-Li Shen for assistance with the GC-MS analysis. This work was supported by NIH Grants GM22172 (D.E.C.) and CA87394 (C.K.). A.Y.C. was the recipient of a Stanford Graduate Fellowship.

Supporting Information Available: Figures for TLC-phosphorimaging assays and kinetic assays can be found in Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

JA908296M