JOC Perspective

pubs.acs.org/joc

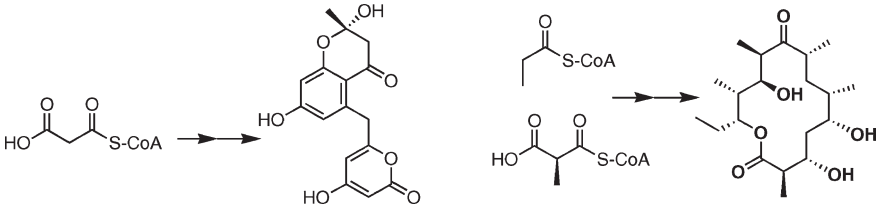
Structures and Mechanisms of Polyketide Synthases

Chaitan Khosla*

Stanford University, Stanford, California 94305-5080

khosla@stanford.edu

Received June 8, 2009



Nearly a quarter-century ago, the advent of molecular genetic tools in the field of natural product biosynthesis led to the remarkable revelation that the genes responsible for the biosynthesis, regulation, and self-resistance of complex polyketide antibiotics were clustered in the genomes of the bacteria that produced these compounds. This in turn facilitated rapid cloning and sequencing of genes encoding a number of polyketide synthases (PKSs). By now, it is abundantly clear that, notwithstanding extraordinary architectural and biocatalytic diversity, all PKSs are evolutionarily related enzyme assemblies. As such, understanding the molecular logic for the biosynthesis of literally thousands of amazing polyketide natural products made by nature can benefit enormously from detailed investigations into a few "model systems". For nearly the past two decades, our laboratory has focused its efforts on two such PKSs. One of them synthesizes two polyketides in approximately equal ratios, SEK4 and SEK4b, and both shunt products from the pathway that leads to the biosynthesis of the pigmented antibiotic actinorhodin. The other synthesizes 6-deoxyerythronolide B, the first isolable intermediate in the biosynthetic pathway for the widely used antibacterial agent erythromycin. Our present-day knowledge of the structures and mechanisms of these two PKSs is summarized here.

Nearly a quarter-century ago, the advent of molecular genetic tools in the field of natural product biosynthesis led to the remarkable revelation that the genes responsible for the biosynthesis, regulation, and self-resistance of complex polyketide antibiotics were clustered in the genomes of the bacteria that produced these compounds.[1] This in turn facilitated rapid cloning and sequencing of genes encoding a number of polyketide synthases (PKSs). By now, it is abundantly clear that, notwithstanding extraordinary architectural and biocatalytic diversity, all PKSs are evolutionarily related enzyme assemblies. As such, understanding the molecular logic for the biosynthesis of literally thousands of amazing polyketide natural products made by nature can benefit enormously from detailed investigations into a few "model systems". For nearly the past two decades, our laboratory has focused its efforts on two such PKSs. One of them synthesizes two polyketides in approximately equal ratios, SEK4 and SEK4b,[2] and both shunt products from the pathway that leads to the biosynthesis of the pigmented antibiotic actinorhodin (Figure 1). The other synthesizes 6-deoxyerythronolide B (Figure 2),[3,4] the first isolable intermediate in the biosynthetic pathway for the widely used antibacterial agent erythromycin. Our present-day knowledge of the structures and mechanisms of these two PKSs is summarized below. Historically, they have been referred to as type II and type I PKSs, respectively,[5] although more recently it has been argued that such nomenclature may be overly simplistic.[6]

(1) Hopwood, D. A. *Chem. Rev.* **1997**, *97*, 2465–2498.

(2) Fu, H.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 4166–4170.

(3) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. *Nature* **1990**, *348*, 176–178.

(4) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Science* **1991**, *252*, 675–679.

(5) Hopwood, D. A.; Sherman, D. H. *Annu. Rev. Genet.* **1990**, *24*, 37–66.

(6) Shen, B. *Curr. Opin. Chem. Biol.* **2003**, *7*, 285–295.

Published on Web 07/10/2009 DOI: 10.1021/jo9012089

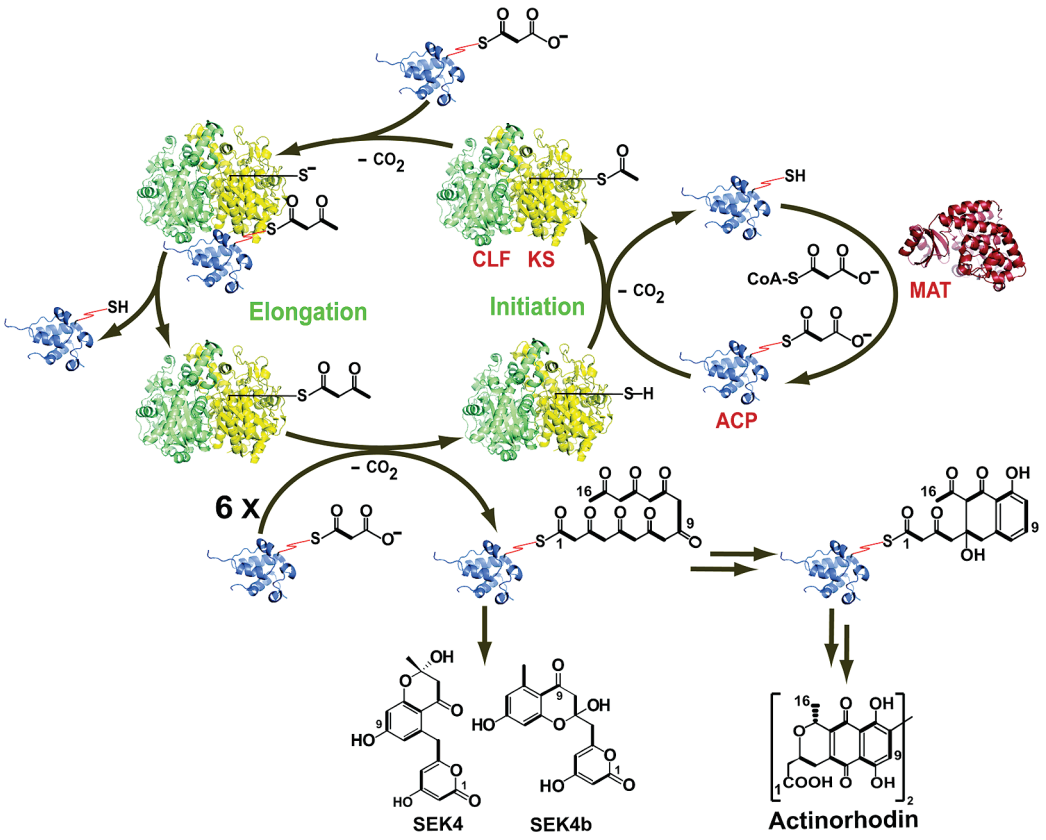


FIGURE 1. Schematic representation of SEK4 and SEK4b biosynthesis by the minimal PKS from the actinorhodin biosynthesis pathway. Actinorhodin is formed by the further action of several downstream enzymes on a common ACP-bound intermediate. The cysteine thiol of the KS and the pantetheinyl thiol of the ACP are explicitly shown. Chain growth is initiated via decarboxylation of malonyl-ACP, followed by transfer of the resultant acetyl group onto the KS. Following seven rounds of chain elongation, the ACP-bound octaketide is released from the KS−CLF and cyclized to yield SEK4 and SEK4b. The atomic structures of all proteins shown have been solved: KS, ketosynthase; CLF, chain length factor; ACP, acyl carrier protein; MAT, malonyl-CoA:ACP transacylase.

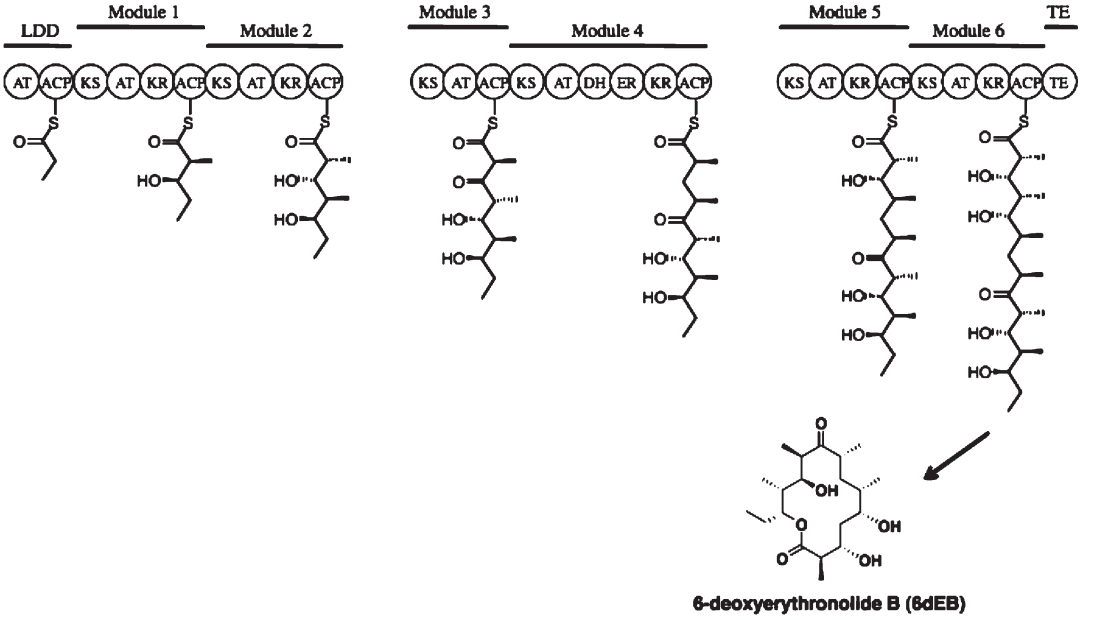


FIGURE 2. Heptaketide macrocycle 6-deoxyerythronolide B (6-dEB) is assembled from one propionyl-CoA and six methylmalonyl-CoA precursors. Chain elongation occurs minimally through the combined action of the ketosynthase (KS), acyl transferase (AT), and acyl carrier protein (ACP) domains. The final oxidation state of the β-carbon is controlled by the specific combination of ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains present in a given module. Once processed, the polyketide chain is either passed to the KS domain of the downstream module or cyclized and released by the thioesterase (TE) domain at the C-terminus of the polyketide synthase. The loading didomain (LDD) is responsible for the selection and subsequent loading of the appropriate priming unit.

Actinorhodin PKS. A ketosynthase (KS), chain length factor (CLF), and acyl carrier protein (ACP) are encoded within the actinorhodin gene cluster.[7] Together with the malonyl-CoA:ACP transacylase (MAT), which is borrowed from the endogenous fatty acid synthase, these proteins comprise a minimal PKS that is necessary and sufficient for catalyzing SEK4 and SEK4b synthesis in the presence of malonyl-CoA as the sole substrate.[8] Cell-free synthesis of SEK4 and SEK4b by a protein preparation from a recombinant strain of *Streptomyces coelicolor* paved the way for reconstituting and biochemically characterizing this minimal PKS. As summarized in Figure 1, it is now well established that the two C_{16} products of the actinorhodin PKS are primed by decarboxylation of malonyl-S-ACP, followed by an interthiol transfer of the resulting acetyl group onto the active site of the KS. This is followed by seven rounds of chain elongation, each of which involves decarboxylative C−C bond formation between a nucleophilic malonyl-S-ACP and the growing polyketide chain bound to the KS via an electrophilic thioester linkage. The full-length poly-β-ketoacyl chain is released from the ACP to yield SEK4 and SEK4b. It must be emphasized that the relative timing of formation of the three 6-membered rings of each product are speculative at this point in time. Presumably they are the result of interactions of the full-length polyketide chain with the PKS proteins as well as its intrinsic "folding propensity", although the relative contribution of the two factors remains to be established.

A combination of biosynthetic, enzymological, and structural biological approaches have been deployed in the context of the actinorhodin PKS. The goals of these investigations have been to obtain a better understanding those features of PKSs that are conserved across most systems as well as those that vary from synthase to synthase. Three questions in particular have been of keen interest to our laboratory in the context of this prototypical PKS. How does the ketosynthase control product structure (especially chain length)? In addition to recognizing the growing polyketide chain, does the ketosynthase also exhibit specificity toward its partner acyl carrier protein? And how is the back-and-forth movement of the growing polyketide chain controlled in this synthase? As summarized below, not only are the answers to these questions providing interesting insights into the catalytic chemistry of type II PKSs, but they also provide a framework for understanding similar mysteries in the context of more complex, multimodular PKSs.

X-ray crystallography revealed that the actinorhodin KS and CLF proteins form a tight heterodimer with an elaborate pocket at their interface into which the growing polyketide chain is extruded.[9] The "depth" of this pocket, dictated by specific residues in the CLF, determines the chain length of the polyketide product. This model has been verified by structure based protein engineering, where changes at selected CLF residues lining this pocket yielded a PKS with a preference for synthesizing longer products.[10] The same conclusion also emerges if chain synthesis is primed with a nonacetyl starter unit; in this situation the KS−CLF heterodimer compensates by catalyzing correspondingly fewer rounds of chain elongation such that the modified backbone is of similar length as the natural product.[11] Indeed, in nature, backbone chain length is the primary variable that dictates the structural variety observed within polycyclic aromatic polyketides from bacteria, and phylogenetic analysis suggests that this property has diversified repeatedly during evolution.[12] It remains to be seen to what extent can the protein engineer manipulate this integral feature of PKSs in the laboratory.

In contrast to the tight association of the KS and CLF, KS−ACP interactions are weak but specific. This specificity is critical in order to prevent different ACPs that coexist in the same cell from transferring biosynthetic intermediates to undesignated KS partners. For example, *Streptomyces coelicolor*, which produces actinorhodin, also produces a closely related polyketide as a spore pigment.[13] The KS−ACP pairs from the corresponding minimal PKSs have orthogonal protein−protein recognition features, and are therefore able to keep the appropriate biosynthetic intermediates compartmentalized within separate enzyme assemblies. Similarly, bacteria that produce antibiotics such as R1128 and frenolicin harbor two distinct ACPs in the same cell, one belonging to each of two modules of a bimodular PKS. KS−ACP interactions are orthogonal between these two modules.[14] As discussed below, KS−ACP specificity is also important in more complex PKSs such as the 6-deoxyerythronolide B synthase.

One consequence of weak KS−ACP interactions is that multiple KS−ACP exchanges occur during the formation of a single molecule of SEK4 or SEK4b. After the formation of each successive C−C bond, the growing chain is transferred back from the ACP to the KS before the proteins dissociate.[15] This timely back-transfer is the essential difference between the actinorhodin PKS and the 6-deoxyerythronolide B synthase. As discussed below, rather than transferring the growing chain back to the KS that forms each C−C bond, the ACPs within the latter PKS deliver intermediates to a successive KS. Understanding the precise mechanistic basis for this difference promises to be of fundamental chemical and evolutionary significance.

In summary, the actinorhodin PKS can be viewed as the "hydrogen atom" of this family of enzyme assemblies. As a single "module" of four proteins that include only two catalytic centers, this synthase captures many of the universal features common to all PKSs. Given the foundational structural and mechanistic knowledge that is now available regarding this system, along with the broad range of chemical and biological tools that can be harnessed to interrogate it, the stage is set to explore the limits of its catalytic potential and, by extension, that of other related PKSs.

6-Deoxyerythronolide B Synthase. For more than half a century, extensive studies into the biosynthesis of

(7) Fernandez-Moreno, M.; Martinez, E.; Boto, L.; Hopwood, D. A.; Malpartida, F. *J. Biol. Chem.* **1992**, *267*, 19278–19290.
(8) Carreras, C. W.; Khosla, C. *Biochemistry* **1998**, *37*, 2084–2088.
(9) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. *Nature Struct. Mol. Biol.* **2004**, *11*, 888–893.
(10) Tang, Y.; Tsai, S. C.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 12708–12709.
(11) Tang, Y.; Lee, T. S.; Khosla, C. *PLoS Biol.* **2004**, *2*, 227–238.
(12) Ridley, C. P.; Lee, H. Y.; Khosla, C. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4595–4600.
(13) Lee, T. S.; Khosla, C.; Tang, Y. *J. Antibiot.* **2005**, *58*, 663–666.
(14) Tang, Y.; Lee, T. S.; Kobayashi, S.; Khosla, C. *Biochemistry* **2003**, *42*, 6588–6595.
(15) Dreier, J.; Khosla, C. *Biochemistry* **2000**, *39*, 2088–2095.

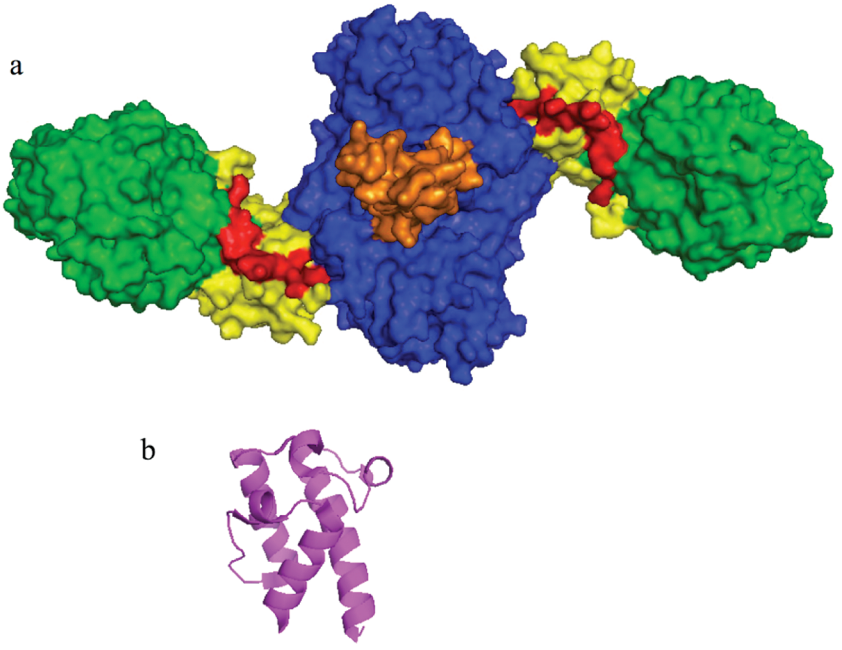


FIGURE 3. (a) Space-filling model of the homodimeric 190 kD fragment of DEBS module 3 comprising both the catalytically active domains (KS and AT shown in blue and green, respectively) and the linkers (shown in gold, yellow, and red). (b) Ribbon diagram of the ACP domain of DEBS module 2.

erythromycin have formed the basis for the biosynthetic paradigm for complex polyketide natural products. Through incorporation experiments with [^{14}C]-, [^{13}C]-, [^{18}O]-, and [^{2}H]-labeled substrates and intermediate analogues, a wealth of indirect information was gained regarding the underlying chemistry of erythromycin biosynthesis.[16] In particular, these isotope labeling studies demonstrated that the carbon chain backbones of the 14-membered macrocyclic core of erythromycin is derived through C−C bond formation between propionate building block units in a manner that is analogous to acetate-derived fatty acids. Furthermore, incorporation of exogenously added analogs of putative biosynthetic intermediates into the natural product unequivocally proved that the biosynthetic machinery for macrolide biosynthesis acts via a processive mechanism in which the oxidation level and stereochemistry of the growing polyketide chain is adjusted immediately after each step of polyketide chain elongation. With the advent of molecular genetic tools came the remarkable discovery of the multimodular assembly line known as the 6-deoxyerythronolide B synthase (DEBS, Figure 2). If the actinorhodin PKS can be thought of as an iterative, unimodular system, DEBS is a hexamodular synthase where each module gets just one turn at elaborating the growing polyketide chain. As summarized below, our investigations into this marvelous assembly line have primarily focused on questions whose answers are not apparent from the schematic shown in Figure 2.

Although the structure of the complete DEBS complex (MW ~2 MDa) remains elusive, high-resolution structures of several domain and multidomain components of the overall assembly line have been solved. Indeed, prototypical high-resolution structures are now available for every domain of DEBS with the exception of the ER domain of module 4.[17] Guided by these structural insights and the high sequence identity between homologous domains of DEBS, it is now possible to reconstruct a coherent structural model for the entire DEBS system (see the issue cover). A major corollary that immediately emerges from the distance constraints of this model is that the ACP domain is more than simply an anchor for the flexible phosphopantetheine swinging arm but must itself undergo considerable motion, presumably along with reactive intermediates, during the course of 6-deoxyerythronolide B synthesis. The mechanisms underlying this phenomenon are, in large part, a complete mystery at the present time. However, to gain a better appreciation of the dynamic considerations, consider the catalytic cycle of module 3 of DEBS, arguably the simplest module in this assembly line (Figure 2).

The X-ray crystal structure of the homodimeric 190 kD portion of module 3 comprising both its catalytically active domains is shown in Figure 3a.[18] Like all of its counterparts, this KS domain has two small molecule substrates (the electrophilic triketide chain synthesized by module 2 and the nucleophilic methylmalonyl extender transferred by the AT domain) and two protein substrates (the ACP domain of module 2 that donates the triketide and the ACP domain of module 3). It also catalyzes two reactions—an interthiol chain transfer from ACP2 onto itself and decarboxylative C−C bond formation. The solution NMR structure of ACP2 has been solved (Figure 3b)[19] and can be used to construct a reliable model for ACP3 as well, in light of their small sizes and high sequence identities. In turn, this inferred structure

(16) O'Hagan, D. *The Polyketide Metabolites*; Ellis Horwood: Amsterdam, 1990.

(17) Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E. *Annu. Rev. Biochem.* **2007**, *76*, 195–221.

(18) Tang, Y.; Chen, A. Y.; Kim, C. Y.; Cane, D. E.; Khosla, C. *Chem. Biol.* **2007**, *14*, 931–943.

(19) Alekseyev, V. Y.; Liu, C. W.; Cane, D. E.; Puglisi, J. D.; Khosla, C. *Protein Sci.* **2007**, *16*, 2093–2107.

of ACP3 has facilitated development of a docking model between the AT and ACP partners as well.

Recent studies on DEBS have revealed that the rates of virtually every reaction within and between individual modules are controlled by a combination of protein−substrate and protein−protein interactions. For example, whereas intermodular chain transfer specificity is heavily influenced by interactions between the donor ACP and acceptor KS,[20] the impact of the transferred polyketide chain itself is relatively modest.[21] In contrast, both the electrophilic thioester substrate[21] and the ACP to which the methylmalonyl extender unit is bound[22] contribute significantly to the specificity of the KS-catalyzed C−C bond-forming reaction, but recognition for the extender unit itself is relatively weak in this reaction.[22] AT-catalyzed transfer of the extender unit from the corresponding coenzyme A thioester to the ACP is greatly influenced by the α-substituent on the malonyl group[23] and less so by the identity of the ACP substrate.[24] What is most remarkable about the contributions of protein−protein interactions is their ability to influence both the selectivity (k_{cat}/K_M) as well as the maximum rate constants (k_{cat}) of individual reactions.[20,22] This suggests that interactions at protein−protein interfaces have the ability to influence the energetics of transition states in the active sites of individual enzyme domains, a principle whose mechanism is of critical significance to the biosynthetic engineer.

A cursory examination of Figure 3a also reveals that a large fraction of this module is comprised of structurally well-defined sequences that link adjacent domains (and modules) to each other but are entirely overlooked in cartoons such as that shown in Figure 2. With the exception of one family of linkers—the docking domains that facilitate intermodular chain transfer between noncovalently interacting modules such as modules 2 and 3 (or alternatively modules 4 and 5)[25,26]—the mechanistic contributions of these linkers to assembly line catalysis is poorly understood, but nevertheless promises to be a frontier for many groundbreaking discoveries in the future.

In summary, by analogy with the ribosome, DEBS (and scores of other related multimodular megasynthases) can be thought of as a highly versatile assembly line for chemical catalysis. The diversity of controlled chemistry accomplished by nature on such assembly lines is immense, yet the principles for its operation appear to be relatively conserved. Understanding these mechanistic principles will undoubtedly open the door in the future to entirely new opportunities for regio- and stereoselective biosynthesis of natural product analogs, and DEBS is likely to remain the prototypical system for such investigations for the foreseeable future.

Acknowledgment. I am grateful to my co-workers and collaborators who, over the years, have made important contributions to the studies summarized above. Our studies on the structures and mechanisms of polyketide synthases have been supported by grants from the National Institutes of Health (CA 66736 and GM 87934).

(20) Wu, N.; Cane, D. E.; Khosla, C. *Biochemistry* **2002**, *41*, 5056–5066.
(21) Wu, J.; Kinoshita, K.; Khosla, C.; Cane, D. E. *Biochemistry* **2004**, *43*, 16301–16310.
(22) Chen, A. Y.; Schnarr, N. A.; Kim, C. Y.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **2006**, *128*, 3067–3074.
(23) Marsden, A. F. A.; Caffrey, P.; Aparicio, J. F.; Loughran, M. S.; Staunton, J.; Leadlay, P. F. *Science* **1994**, *263*, 378–380.
(24) Chen, A. Y.; Cane, D. E.; Khosla, C. *Chem. Biol.* **2007**, *14*, 784–792.
(25) Gokhale, R. S.; Tsuji, S.; Cane, D. E.; Khosla, C. *Science* **1999**, *284*, 482–485.
(26) Tsuji, S. Y.; Cane, D. E.; Khosla, C. *Biochemistry* **2001**, *40*, 2326–2331.