The *Streptomyces glaucescens tcmKL* Polyketide Synthase and *tcmN* Polyketide Cyclase Genes Govern the Size and Shape of Aromatic Polyketides

B. Shen,[†] R. G. Summers,[†,§] E. Wendt-Pienkowski,[†] and C. R. Hutchinson*[,†,‡]

Contribution from the School of Pharmacy and Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706

Received February 7, 1995[⊗]

Abstract: The mechanism of a type II polyketide synthase was analyzed by combinatorially expressing components of the tetracenomycin (*tcm*) and the actinorhodin (*act*) polyketide synthase genes in various mutants or heterologous hosts. Structural analysis of metabolites produced by the recombinant organisms provided evidence to dissect the function of individual components of a type II PKS. Complementation studies with the *S. glaucescens* TcmK$^-$ and TcmL$^-$ mutants and constructs harboring the *tcmK* and *actI-ORF2* or *actI-ORF1* and *tcmL* genes demonstrated that a heterologous pair of genes encoding a β-ketoacyl:acyl carrier protein synthase and chain length factor is often (but not always) nonfunctional. Isolations of the octaketides **6** and **7** from strains bearing *actI-ORF1*, *actI-ORF2*, and *tcmM* (pWHM766) or *tcmJ*, *actI-ORF1*, *actI-ORF2*, and *tcmM* (pWHM768), and the decaketides **8** and **9** from strains bearing *tcmK*, *tcmL*, and *tcmM* (pELE37) or *tcmJ*, *tcmK*, *tcmL*, and *tcmM* (pWHM731) established that the β-ketoacyl:acyl carrier protein synthase and chain length factor proteins determine the chain length of the polyketide. While the addition of the *tcmJ* gene to the polyketide synthase core proteins consisting of the β-ketoacyl:acyl carrier protein synthase, chain length factor and acyl carrier protein had no effect on the structures of the resulting metabolites, adding the *tcmN* gene to either pWHM766 and pWHM768 or pELE37 and pWHM731 resulted in the synthesis of octaketide **5** or decaketide **2**, respectively. TcmN thus can alter the regiospecificity for the first aldol cyclization from C-7/C-12 to C-9/C-14, suggesting that cyclases like TcmN determine the folding pattern of the linear polyketide intermediate. These activities, along with the choice of the starter unit, the loading of the extender unit to the PKS complex, and the function of TcmJ are discussed in an attempt to provide a rationale for the engineered biosynthesis of novel polyketides.

Polyketide metabolites are a large and diverse family of secondary metabolites found in bacteria, fungi, and plants.[1] Many of them are clinically valuable antibiotics or chemotherapeutic agents or have other useful pharmacological activities (immunosuppressive, antiparasitic, insecticidal, etc).[2] Ever since Birch[3] outlined his "acetate hypothesis" in the early 1950s that polyketides could be biosynthetically derived from short fatty acids such as acetate, propionate, or butyrate, activated as acylthioesters, chemists and biochemists have been searching for a unified biosynthetic mechanism to account for the huge structural diversity found in polyketides.

From the results of isotope labeling experiments, it has been evident for some time that the mechanism of polyketide biosynthesis is analogous to that of long-chain fatty acid biosynthesis catalyzed by the fatty acid synthases and that both polyketide and fatty acid biosynthesis utilize largely the same precursors. However, direct evidence supporting the analogy between polyketide and fatty acid biosynthesis has come only recently from the molecular genetic and biochemical studies of antibiotic biosynthesis in *Streptomyces* species, principally. Sequence analysis of several sets of bacterial genes (plus a few from fungi) encoding polyketide synthases has revealed a highly conserved gene organization and a high degree of amino acid sequence similarity among the enzymes. Comparison of the polyketide synthases and fatty acid synthases has led to the conclusion that these systems must share a common mechanism of carbon chain assembly from similar if not identical precursors.[4−9] Yet, the details of how a polyketide synthase assembles the polyketide carbon skeleton remain obscure. Most of the mechanistic insights have been deduced from the effects of introducing mutations into the chromosomal copy of the polyketide synthase genes or by overexpressing the native or mutant polyketide synthase genes in various backgrounds.[9] Consequently, our understanding of the enzymology and biochemistry of polyketide synthases is rudimentary, and the purification or reconstitution of a complete polyketide synthase complex from a bacterium has so far not been reported, although significant progress has recently been made in this direction.[10−15]

* Corresponding author: C. R. Hutchinson, School of Pharmacy, University of Wisconsin, 425 North Charter St., Madison, WI 53706. Phone: (608)262-7582. Fax: (608)262-3134. email: crhutchi@facstaff.wisc.edu.

[†] School of Pharmacy.

[‡] Department of Bacteriology.

[§] Current address: Department 47N, Abbott Laboratories AP9A, Abbott Park, IL 60064.

[⊗] Abstract published in *Advance ACS Abstracts*, June 15, 1995.

(1) O'Hagen, D. *The Polyketide Metabolites*; Ellis Horwood: Chichester, U.K., 1991.

(2) Monaghan, R. L.; Tkacz, J. S. *Annu. Rev. Microbiol.* **1990**, *44*, 271−301.

(3) Birch, A. J.; Donovan, F. W. *Aus. J. Chem.* **1953**, *6*, 360−368, 373−378; Birch, A. J.; Elliot, P. *Aus. J. Chem.* **1953**, *6*, 369−372; Birch, A. J. *Science* **1967**, *156*, 202−206.

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

(5) Katz, L.; Donadio, S. *Annu. Rev. Microbiol.* **1993**, *47*, 875−912.

(6) Hopwood, D. A.; Khosla, C. *Ciba Found. Symp.* **1992**, *171*, 88−112.

(7) Hopwood, D. A. *Curr. Opin. Biotechnol.* **1994**, *4*, 531−537.

(8) Vining, L. C.; Stuttard, C. *Genetics and Biochemistry of Antibiotic Production*; Butterworth-Heinemann: Boston, 1995; pp 323−498.

(9) Hutchinson, C. R.; Fujii, I. *Annu. Rev. Microbiol.* **1995**, *47*, 201−238.

0002-7863/95/1517-6811$09.00/0

Polyketide synthases found in bacteria and fungi have been classified into two groups. Type I polyketide synthases, catalyzing the biosynthesis of highly reduced polyketides such as macrolide and polyether antibiotics, are large multifunctional proteins that harbor a distinct enzyme activity for every step catalyzed and function largely nonreiteratively. Type II polyketide synthases, catalyzing the biosynthesis of aromatic polyketides, in contrast, are multienzyme complexes that carry a single set of reiteratively used activities and consist of several, largely monofunctional proteins.[4-9] It appears that the means to determine the assembly and processing of the nascent polyketide chain for a type I polyketide synthase was to evolve a series of enzyme activities possibly arranged in the order of the steps in polyketide biosynthesis, as typified by the 6-deoxyerythronolide B synthase from *Saccharopolyspora erythraea*.[16-19] In this case, a set of domains (known as a module) containing the requisite number of active sites for substrate loading and the condensation between the starter unit or acyl-coenzyme A (CoA) intermediates and the chain-extender units [R(COOH)COS-CoA] plus the subsequent processing events (carbonyl or double bond reduction and alcohol dehydration) associated with each step of polyketide construction can be recognized in the protein sequence. The mechanism of a type II polyketide synthase is not as decipherable since the latter only has one active site for each type of reaction, as exemplified by the tetracenomycin (Tcm) polyketide synthase from *Streptomyces glaucescens*.[13,20-22] Therefore, the principal challenges to understanding the catalytic mechanism of the type II polyketide synthases are to determine how the starter unit is specified, how the length of the growing polyketide chain is controlled, how the nascent polyketide chain is folded into the correct conformation for cyclization, and how the carbonyl and methylene groups are directed to react to form the six-membered rings of the final cyclic compounds.

We have been studying the biosynthesis of Tcm C, **1**, a polyketide antitumor antibiotic produced by *S. glaucescens*, as a model of the type II polyketide synthase (Figure 1). On the basis of sequence analysis[20-22] and in vitro studies,[10,11,13] we reported that the enzymes produced by the *tcmK*, *tcmL*, and *tcmM* genes[23] are responsible for the synthesis of the nascent, enzyme bound linear decaketide from one acetyl-CoA and nine malonyl-CoA molecules, and that this decaketide intermediate is subsequently folded and cyclized by the TcmN[23] protein alone or in combination with the TcmJ protein to form Tcm F2, **2**, the earliest isolable product released from the Tcm polyketide synthase complex (Figure 1).[21,24] We also demonstrated that the Tcm polyketide synthase has a high substrate specificityfor acetyl-CoA as the starter unit.[13] Moreover, we have suggested[13,25,26] a linkage between polyketide and fatty acid biosynthesis in *S. glaucescens* and have proposed that the loading of the malonyl-CoA extender unit to the TcmM acyl carrier protein of the Tcm polyketide synthase is catalyzed by a malonyl CoA:acyl carrier protein acyltransferase believed to be part of the fatty acid synthase cluster in the same organism.[25,26]

To extend our investigation of the mechanism of the Tcm polyketide synthase, we describe here an approach, complementary to the cell-free system we reported previously,[13] which involves combinatorially expressing components of the Tcm and actinorhodin (Act) polyketide synthases in a suitable *Streptomyces* host. This approach was first used by Khosla, Hopwood, and their collaborators,[27] who showed that hybrid polyketide synthases, made by combining genes for components of two different polyketide synthases, are often functional. The use of *S. glaucescens* TcmK$^-$ and TcmL$^-$ mutants[23] blocked in the early steps of Tcm C biosynthesis provide an opportunity to examine the interactions between polyketide synthase components from chromosomal and plasmid borne gene products, while the heterologous host *Streptomyces lividans*[28] and the *S. glaucescens* TcmIc[29,30] mutant devoid of the normal Tcm polyketide synthase proteins minimize the uncertainties associated with the background activities. Structural analysis of metabolites produced by recombinant bacteria bearing novel combinations of *tcm* and *act* polyketide synthase genes was expected to provide clues from which to deduce the specific function of the individual components of the Tcm polyketide synthase. We chose the combination between the Tcm polyketide synthase and the Act polyketide synthase from *Streptomyces coelicolor*,[31-33] which forms the shunt products aloesaponarin II, **3a**, and its 2-carboxy derivative, **3b**[34] (derived from the pathway leading to the main *S. coelicolor* metabolite actinorhodin **4**), because the Tcm polyketide synthase synthesizes **2**, a decaketide resulting from a linear polyketide folded at C-11,

(10) Gramajo, H.; White, J.; Hutchinson, C. R.; Bibb, M. J. *J. Bacteriol.* **1991**, *173*, 6475-6483.
(11) Shen, B.; Gramajo, H.; Bibb, M. J.; Hutchinson, C. R. *J. Bacteriol.* **1992**, *174*, 3818-3821.
(12) Roberts, G. A.; Staunton, J.; Leadlay, P. F. *Eur. J. Biochem.* **1993**, *214*, 305-311.
(13) Shen, B.; Hutchinson, C. R. *Science* **1993**, *262*, 1535-1540.
(14) Marsden, A. F. A.; Caffrey, P.; Aparicio, J. F.; Loughran, M. S.; Staunton, J.; Leadlay, P. F. *Science* **1994**, *263*, 378-380.
(15) Aparicio, J. F.; Caffrey, P.; Marsden, A. F. A.; Staunton , J.; Leadlay, P. F. *J. Biol. Chem.* **1994**, *269*, 8524-8528.
(16) Bevitt, D. J.; Cortes, J.; Haydock, S. F.; Leadlay, P. F. *Eur. J. Biochem.* **1992**, *204*, 39-49.
(17) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. *Nature* **1990**, *346*, 176-178.
(18) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Science* **1991**, *252*, 675-679.
(19) Donadio, S.; Katz, L. *Gene* **1992**, *111*, 51-60.
(20) Bibb, M. J.; Biro, S.; Motamedi, H.; Collins, J. F.; Hutchinson, C. R. *EMBO J.* **1989**, *8*, 2727-2736.
(21) Summers, R. G.; Wendt-Pienkowski, E.; Motamedi, H.; Hutchinson, C. R. *J. Bacteriol.* **1992**, *174*, 1810-1820.
(22) Summers, R. G.; Wendt-Pienkowski, E.; Motamedi, H.; Hutchinson, C. R. *J. Bacteriol.* **1993**, *175*, 7571-7580.
(23) Glossary: *tcm* , *act* and *dps* designate genes encoding polyketide synthase or cyclase enzymes; Tcm, Act and Dps designate the protein products of the corresponding genes; and TCM or ACT indicate the intermediates or products of tetracenomycin and actinorhodin biosynthesis, respectively. For instance, the *tcmJ*, *tcmK*, *tcmL*, *tcmM*, and *tcmN* genes (also designated as *tcmJKLMN*) encode the TcmJ, TcmK, TcmL, TcmM, and TcmN proteins; and *tcmK* or TcmK$^-$ and *tcmL* or TcmL$^-$ indicate the geneotypes and phenotypes, respectively, of mutants that lack either of these enzymes. Plasmid vectors for expression of *tcm* or *act* genes are designated by pWHM and bacterial strains by WMH prefixes, respectively.
(24) Shen, B.; Nakayama, H.; Hutchinson, C. R. *J. Nat. Prod.* **1993**, *56*, 1288-1293.
(25) Summers, R. G.; Ali. A.; Shen, B.; Wessel, W. A.; Hutchinson, C. R. *Biochemistry*, **1995**, in press.
(26) Hutchinson, C. R.; Decker, H.; Motamedi, H.; Shen, B.; Summers, R. G.; Wendt-Pienkowski, E.; Wessel, W. L. in *Industrial Microorganisms: Basic and Applied Molecular Genetics*; Baltz, R. H., Hegenman, G. D., Skatrud, P. L., Eds.; American Society for Microbiology: Washington, D.C., 1993; pp 203-216.
(27) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *Science* **1993**, *262*, 1546-1550.
(28) Hopwood, D. A.; Kieser, T.; Wright, H. M.; Bibb, M. J. *J. Gen. Microbiol.* **1983**, *129*, 2257-2269.
(29) Motamedi, H.; Wendt-Pienkowski, E.; Hutchinson, C. R. *J. Bacteriol.* **1986**, *167*, 575-580.
(30) Decker, H.; Hutchinson, C. R. *J. Bacteriol.* **1993**, *175*, 3887-3892.
(31) Fernandez-Moreno, M. A.; Martinez, E.; Boto, L.; Hopwood, D. A.; Malpartida, F. *J. Biol. Chem.* **1992**, *267*, 19278-19290.
(32) Sherman, D. H.; Bibb, M. J.; Simpson, T. J.; Johnson, D.; Malpartida, F.; Fernandez-Moreno, M.; Maetinez, E.; Hutchinson, C. R.; Hopwood, D. A. *Tetrahedron* **1991**, *47*, 6029-6043.
(33) Hallam, S. E.; Malpartida, F.; Hopwood, D. A. *Gene* **1988**, *74*, 305-320.
(34) Bartel, P. L.; Zhu, J. S.; Lampel, J. S.; Dosch, D. C.; Connors, N. C.; Strohl, W. R.; Beale, J. M.; Floss, H. G. *J. Bacteriol.* **1990**, *172*, 4816-4826.

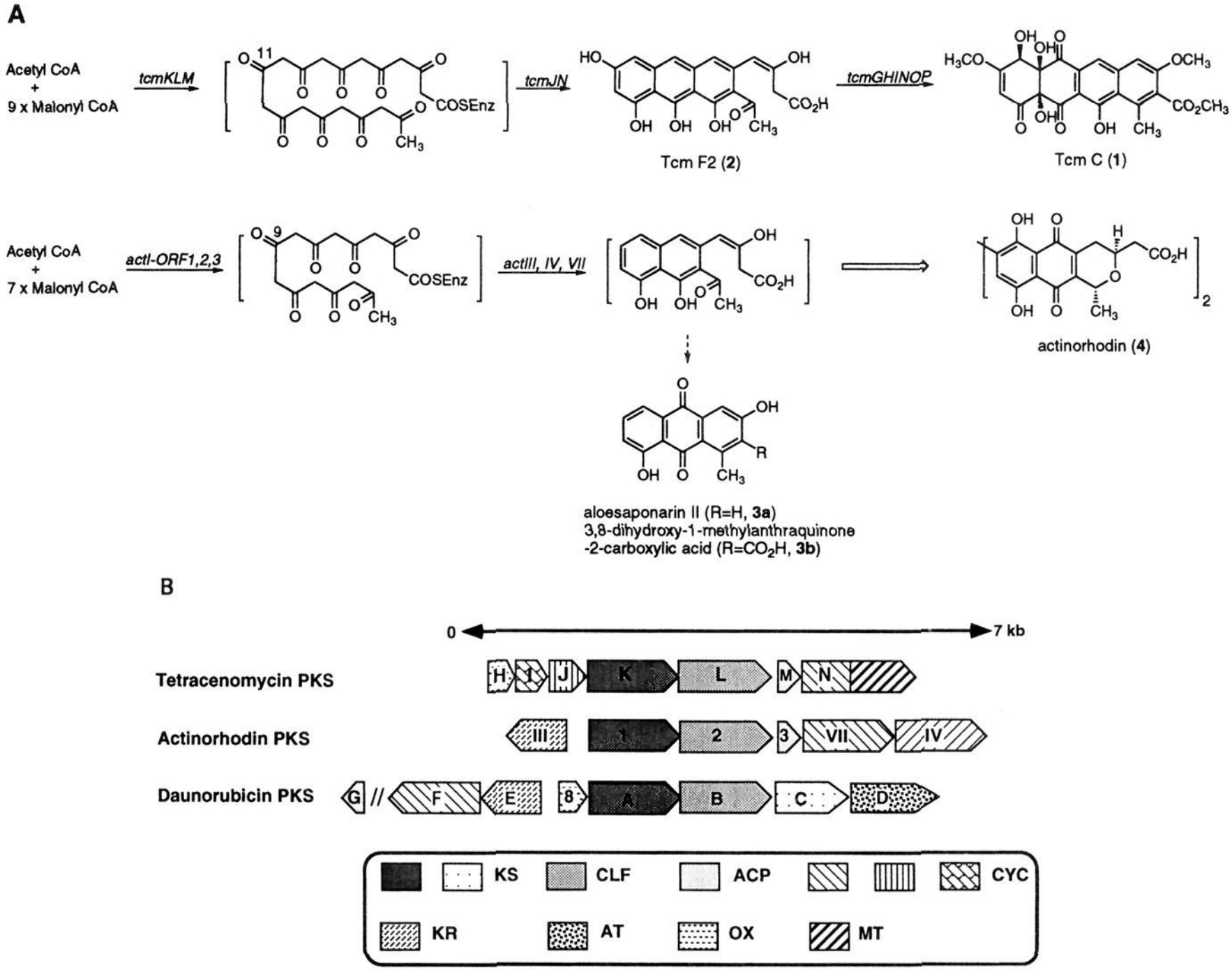

Figure 1. 1. Biosynthetic pathways (**A**) and the gene organizations (**B**) for the production of Tcm C, **1**, in *S. glaucescens*, actinorhodin, **4**, in *S. coelicolor*, and daunorubicin in *S. peucetius*. The abbreviations used are KS, β-ketoacyl:acyl carrier protein synthase; CLF, chain length factor; ACP, acyl carrier protein; CYC, polyketide cyclase; KR, β-ketoacyl reductase; AT, acyltransferase, OX, oxygenase; MT, methyl transferase.

while the Act polyketide synthase produces **3a** and **3b**, octaketides derived by folding the linear polyketide at C-9 prior to the first cyclization (Figure 1A). On the basis of the presumption that functional hybrid polyketide synthases would result from expression of the *tcm* and *act* genes in the same bacterium,[27] our studies were aimed at determining the factors that control the chain length of the polyketide intermediate and the regiospecificity of the subsequent cyclizations to form the aromatic products. Five compounds, UWM1, **5**, SEK4, **6**, SEK4b, **7**, SEK15, **8**, and SEK15b, **9**, in addition to **2**, were identified from the recombinant organisms (Figure 3). From the relationship between the composition of the hybrid polyketide synthase and the compounds produced, we conclude that the *tcmKL* genes (or genes whose products are quite similar to TcmK and TcmL) determine the length of the nascent polyketide chain made from acetate and malonate and that cyclases like that encoded by the *tcmN* gene dictate the regiospecificity of the first of a set of cyclization steps of the linear polyketide intermediate. Thus, the *tcmKLMN* genes govern the size and shape of the aromatic polyketides normally produced by *S. glaucescens*. Our results not only corroborate the findings of Khosla, Hopwood, and co-workers reported over the past year (*vide infra*) but also reveal significant differences among the behavior of type II polyketide synthases in different experimental systems.

Results

The Chain Length Factor Alone Is Not Sufficient To Control the Chain Length. A typical type II polyketide synthase consists of three core proteins, the β-ketoacyl:acyl carrier protein synthase and its homologous companion, recently named the "chain length factor",[27] and the acyl carrier protein, plus unique proteins called a polyketide cyclase, a β-ketoacyl reductase, and an acyltransferase, some of which may be absent in an individual type II polyketide synthase (Figure 1B). Since it had been demonstrated early that interchanges of the acyl carrier protein component have little influence on the structures of the polyketides produced,[27,35-38] we used the TcmM acyl carrier protein throughout this study. Hybrid Tcm/Act polyketide synthases provided by heterologous pairs of genes (i.e., combinations of genes from different species of bacteria) encoding the β-ketoacyl:acyl carrier protein synthase/chain length factor proteins were constructed in a derivative of plasmid pIJ486 where the transcription of the polyketide synthase genes is under

(35) Khosla, C.; McDaniel, R.; Ebert-Khosla, S.; Torres, R.; Sherman, D. H. *J. Bacteriol.* **1993**, *175*, 2197−2204.
(36) Khosla, C.; Ebert-Khosla, S.; Hopwood, D. A. *Molec. Microbiol.* **1992**, *6*, 3237−3249.
(37) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1993**, *115*, 11671−11675.
(38) McDaniel, R.; Ebert-Khosla, S.; Fu, H.; Hopwood, D. A.; Khosla, C. *Proc. Natl. Acad. Sci.* U.S.A. **1994**, *91*, 11542−11546.

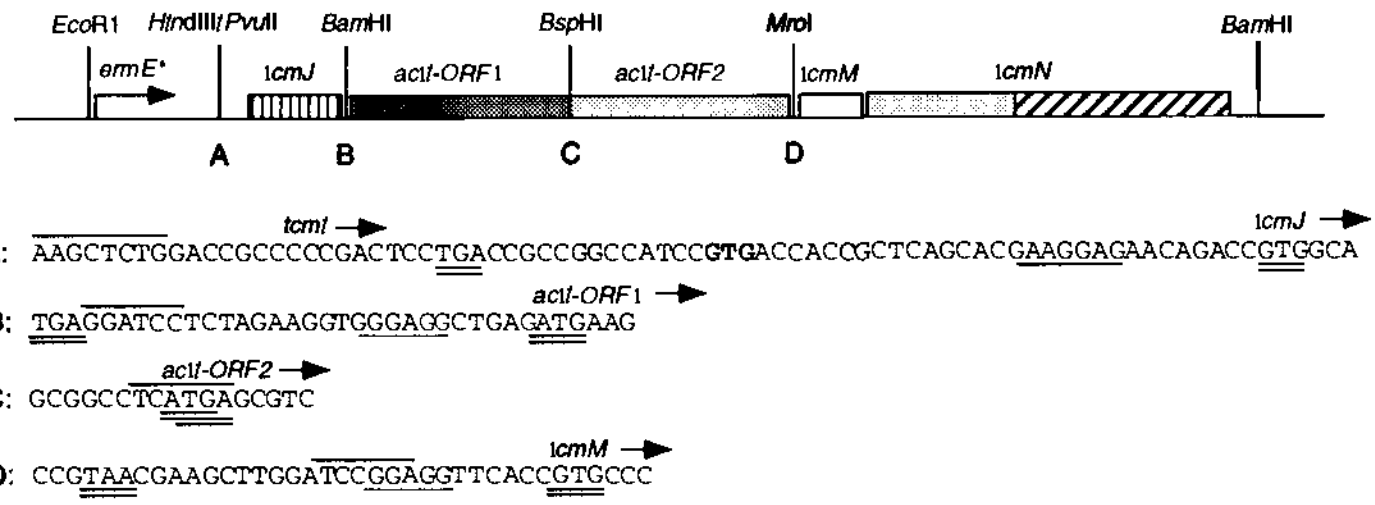

Figure 2. Genetic arrangement of pWHM769. The genes cloned in the *Eco*R1 and *Bam*H1 sites of plJ486 are depicted by boxes shaded according to Figure 1B, and the junctions between the genes are lettered A–D. The *ermE** promoter is indicated by a bent arrow pointing in the direction of gene transcription. Only the restriction sites of interest are shown and are not necessarily unique. Gene sequence data not shown are available under GenBank accession numbers M80674 and X63449 for the *tcm* genes and the *act* genes, respectively. The genes for the Act β-ketoacyl:acyl carrier protein synthase and chain length factor (*actI-ORF1* and *actI-ORF2*, respectively) were obtained by PCR amplification from *S. coelicolor* M111,[59] and the sequence of the genes was confirmed by DNA sequencing. The *tcm* and *act* genes were assembled at junction sites A–D. To maintain the possibility for translational coupling between the *act* β-ketoacyl:acyl carrier protein synthase and either the *act* or *tcm* chain length factor genes (junction C), the codon for the last residue of *actI-ORF1* was changed from alanine to serine. This alteration appears to be phenotypically silent. To ensure efficient expression in *S. glaucescens*, the initiating TTG codon of *actI-ORF1* was changed to an ATG codon (junction B). The DNA sequence of each of the fusion junctions is shown beneath the physical map of the cloned genes. Putative ribosome binding sites and the relevant start and stop codons for the cloned genes are singly and doubly underlined, respectively. The restriction enzyme sites shown in the physical map are overscored. An alternative GTG start codon for *tcmJ* (shown in boldface type) is suggested by the results of site-directed mutagenesis experiments (E. Wendt-Pienkowski and C. R. Hutchinson, unpublished data).

the control of the strong, constitutive *ermE** promoter (Table 1).[21] Figure 2 shows details of the construction of pWHM769,[23] a typical example of plasmids containing *act* and *tcm* polyketide synthase genes. The plasmids were introduced by transformation into the Tcm C non-producing strains *S. glaucescens* WMH1061,[23] -WMH1068, and -WMH1077 to examine the relationship between the gene combinations and the production of known or new metabolites. *S. glaucescens* WMH1061 has a point mutation in the *tcmK* gene,[22,29] *S. glaucescens* WMH1068 has a deletion mutation in the *tcmL* gene,[22,29] and *S. glaucescens* WMH1077 has a point mutation in the promoter region of the *tcmGHIJKLMNO* operon,[22,29,31] which prevents significant expression of these nine *tcm* genes.

Table 2 summarizes the metabolites identified from the recombinant bacteria. Introduction of pWHM732 (*tcmJKLMN*) into either the WMH1061 or WMH1068 strains restored the production of **1** and served as a positive control, demonstrating that the plasmid borne gene products are able to interact effectively with those from the chromosomal genes. As anticipated, the hybrid polyketide synthases resulting from introduction of pWHM750, pWHM751, and pWHM752, each of which carry the normal *tcmK* gene along with *actI-ORF2*, restored the production of **1** to WMH1061. However, despite the high sequence similarity between TcmK and ActI-Orf1, complementation was not observed when either pWHM762, pWHM763, or pWHM764, each carrying the *actI-ORF1* gene along with *tcmL*, was introduced into WHM1061; only a trace amount of **1** was observed with pWHM765 that has the productivity-enhancing *tcmJ* gene discussed below. The complementation results obtained with the WMH1068 strain that lacks TcmL were different. Not only the plasmids harboring the *tcmL* gene, such as pWHM762, pWHM763, pWHM764, and pWHM765, but also those in which the *tcmL* gene was replaced by *actI-ORF2*, such as in pWHM750, pWHM751, and pWHM752, restored production of **1** to WMH1068. When these cassettes of hybrid polyketide synthase genes were introduced into WMH1077, no significant amount of any metabolite was observed except in the case of pWHM752, which contains *tcmJ*, *tcmK*, *actI-ORF2*, *tcmM*, and *tcmN* and caused the production of a mixture of **2** and **5** in small quantity, in sharp contrast to the positive control pWHM732 bearing *tcmJKLMN* that produced **2** exclusively. Taken together, these results show that a heterologous pair of β-ketoacyl:acyl carrier protein synthase/chain length factor proteins can be nonfunctional, as evident in the combinations of *actI-ORF1/tcmL* that failed either to complement the mutation in WMH1061 (i.e., to restore Tcm C synthesis) or to synthesize any metabolite in WMH1077. However, the chain length factor seems to have a relaxed fidelity because the combination of *tcmK/actI-ORF2* was able to complement WMH1068, albeit less effectively than *tcmK/tcmL* on the basis of the relative yields of **1**. Since the Act polyketide synthase specifies an octaketide intermediate, the fact that *tcmK/actI-ORF2* restored the production of **1** to WMH1068, via a decaketide intermediate, strongly suggests that the chain length factor alone is not sufficient to control the chain length. In fact, populations of both C_{16} and C_{20} metabolites were identified as **2** and **5** in the case of the WMH1077 transformant carrying *tcmJK/actI-ORF2/tcmMN*. These results differ from the behavior of the chain length factor encoding genes reported by Sherman, Hopwood, and co-workers[39] where only the *gral-ORF2* gene (encoding a protein functionally identical to Act-Orf1) among the three tested was capable of restoring actinorhodin production to an *actI-ORF2* mutant but are consistent with the report of McDonald et al.[27] that the *tcmL* gene in combination with *actI-ORF1* did not produce a functional polyketide synthase.

The Joint Activities of the β-Ketoacyl:Acyl Carrier Protein Synthase and Chain Length Factor Determine the Chain Length. Since the above results show that a heterologous pair of β-ketoacyl:acyl carrier protein synthase/chain length factor proteins can be nonfunctional in certain cases, we next constructed cassettes of hybrid polyketide synthase genes with the β-ketoacyl:acyl carrier protein synthase/chain length factor-encoding pair from the same bacterium. We initially introduced these plJ486-based cassettes into both *S. glaucescens* WMH1077 and *S. lividans* to examine their abilities to synthesize polyketide metabolites. Upon thin layer chromatography (TLC) and/or high performance liquid chromatography (HPLC) analysis of

(39) Kim, E.-S.; Hopwood, D. A.; Sherman, D. H. *J. Bacteriol.* **1994**, *176*, 1801–1804.

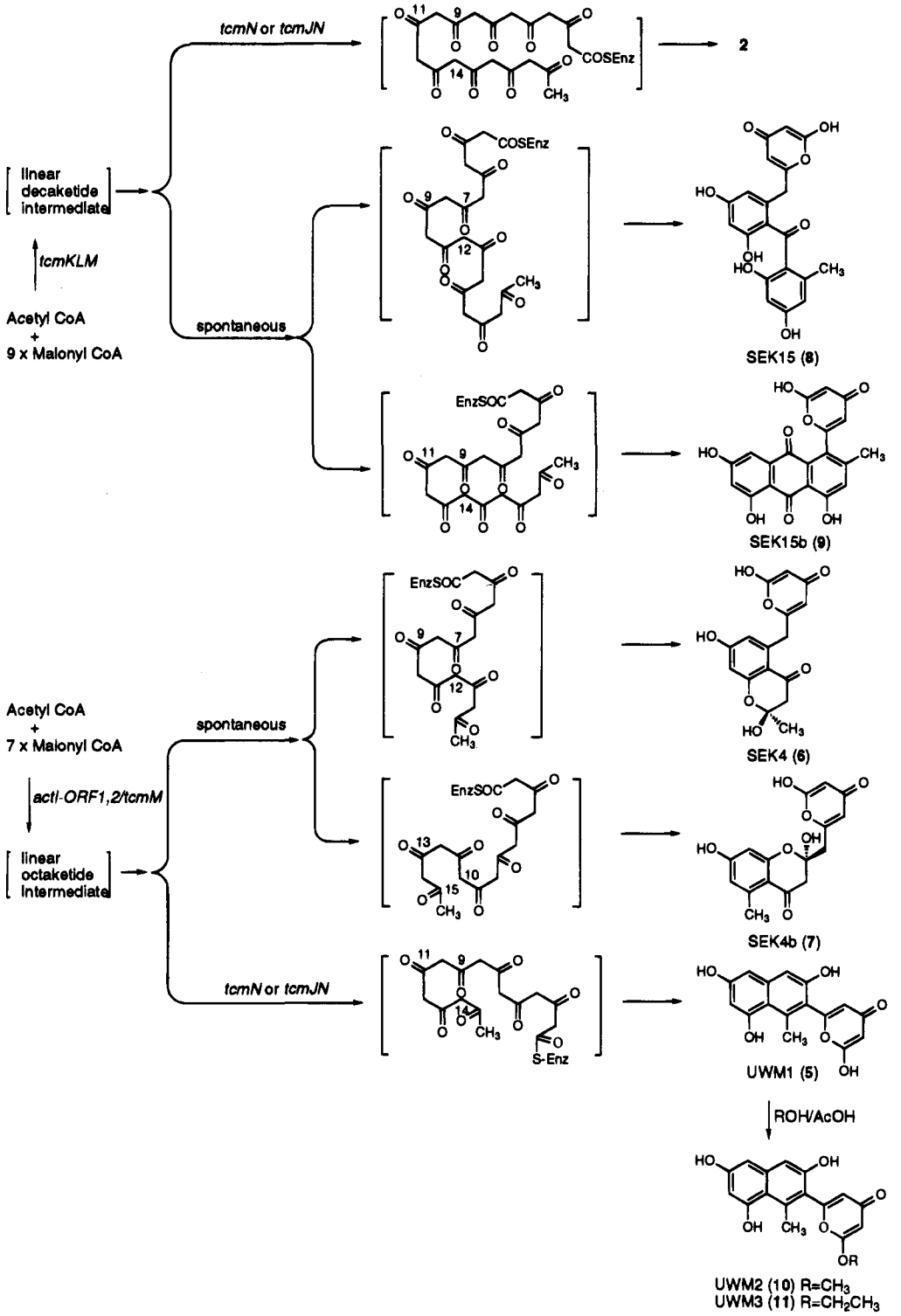

Figure 3. Proposed biosynthetic mechanisms and structures of identified polyketides produced by the *act/tcm* hybrid polyketide synthase plasmids in *S. glaucescens* or *S. lividans*.

the crude extracts from both hosts, we found that the product distributions in both systems were almost identical. Since the WMH1077 strain consistently gave better yields, it subsequently became the host of choice.

Table 3 summarizes the major polyketide metabolites identified from the recombinant bacteria. Although the productivity varied between batches even for the same construct, the comparison of yields between different constructs are significant because they represent the average results of many fermentations. Mass spectral (MS) analysis of these metabolites led us to classify the recombinant bacteria into two groups: the octaketide producers with pWHM766, pWHM767, pWHM768, and pWHM769 containing *actI-ORF1* and *-ORF2* combined with different *tcm* genes and the decaketide producers with pELE37, pWHM722, pWHM731, and pWHM732 containing only *tcm* genes. The two groups correlated with the origin of the β-ketoacyl:acyl carrier protein synthase/chain length factor pair in each construction (Table 1): the octaketide producers utilize the Act β-ketoacyl:acyl carrier protein synthase/chain length factor proteins and the decaketide producers contain the Tcm β-ketoacyl:acyl carrier protein synthase/chain length factor proteins.

Table 1. Plasmids Used in This Study

plasmid	genotype	ref
pWHM762	*actI-ORF1/tcmLM*	this work
pWHM763	*actI-ORF1/tcmLMN*	this work
pWHM764	*tcmJ/actI-ORF1/tcmLM*	this work
pWHM765	*tcmJ/actI-ORF1/tcmLMN*	this work
pWHM750	*tcmK/actI-ORF2/tcmM*	this work
pWHM751	*tcmK/actI-ORF2/tcmMN*	this work
pWHM752	*tcmJK/actI-ORF2/tcmMN*	this work
pWHM766	*actI-ORF1/actI-ORF2/tcmM*	this work
pWHM767	*actI-ORF1/actI-ORF2/tcmMN*	this work
pWHM768	*tcmJ/actI-ORF1/actI-ORF2/tcmM*	this work
pWHM769	*tcmJ/actI-ORF1/actI-ORF2/tcmMN*	this work
pELE37	*tcmKLM*	21, 22
pWHM722	*tcmKLMN*	21, 22
pWHM731	*tcmJKLM*	21, 22
pWHM732	*tcmJKLMN*	this work

The structures of the metabolites isolated were established mainly by ^1H and ^{13}C NMR spectroscopic analyses and were confirmed by isotope labeling experiments with [1-^{13}C]-, [2-^{13}C]-, and [1,2-$^{13}C_2$]acetate. The metabolite, **2**, synthesized by strains with *tcmKLMN* or *tcmJKLMN*, was identified as Tcm F2.[24] The major metabolite, **8**, isolated from strains bearing *tcmKLM* or *tcmJKLM*, was clearly a decaketide on the basis of the high resolution fast atom bombardment (FAB) MS analysis, which yielded a $[M + H]^+$ ion at 385.0920 ($C_{20}H_{16}O_8$, calc. 384.0845). ^1H, $^1H/^1H$ COSY, and ^{13}C NMR analyses of **8** revealed that it was identical to SEK15.[40] The minor metabolite, **9**, copurified along with **8**, was produced in such a low quantity that no significant NMR data were collected. However, it behaved identically to authentic SEK15b upon either TLC or HPLC analysis, leading us to assign **9** as SEK15b[38] (Figure 3).

Similar methods were applied to establish the structures of the octaketide metabolites. The molecular formulas of the two major compounds, **6** and **7**, isolated from strains bearing *actI-ORF1/actI-ORF2/tcmM* or *tcmJ/actI-ORF1/actI-ORF2/tcmM*, were established to be $C_{16}H_{14}O_7$ (calc. 318.0740) by high resolution FAB MS analysis, which gave almost identical $[M + H]^+$ peaks at 319.0822 and 319.0823, respectively, although the fragmentation patterns of the two compounds were different. Upon NMR analysis, **6** and **7** displayed identical ^1H, $^1H/^1H$ COSY, ^{13}C, and $^1H/^{13}C$ HETCOR data to those of SEK4[40] and SEK4b,[41] respectively. The structure of **5** was more elusive. Electron spray MS analysis of **5** revealed a molecular weight of 300, indicative of $C_{16}H_{12}O_6$. Moreover, when treated with CH_3OH or CH_3CH_2OH in the presence of a trace of AcOH, **5** formed two adducts, UWM2, **10**, and UWM3, **11**. High resolution electron ionization (EI) MS analyses of **10** and **11** yielded molecular weights of 314.0818 and 328.0970, suggesting molecular formulas of $C_{17}H_{14}O_6$ (calc. 314.0790) and $C_{18}H_{16}O_6$ (calc. 328.0947), respectively. Thus, it was clear that **5** is a octaketide metabolite. Therefore, these results show that plasmids harboring *actI-ORF1/actI-ORF2* synthesized exclusively octaketides and that plasmids harboring *tcmK/tcmL* synthesized exclusively decaketides, suggesting that it is the joint activities of the β-ketoacyl:acyl carrier protein synthase/chain length factor proteins that determine the chain length of a polyketide made by a type II polyketide synthase (Figure 3). Consequently, there is a much stronger correlation between the size of the polyketide produced when the β-ketoacyl:acyl carrier protein synthase and chain length factor proteins of the polyketide synthase are from the same bacterium than when they are encoded by heterologous pair of genes encoding these proteins if the latter results in a functional enzyme complex.

Cyclases Like TcmN Appear To Determine the Folding Pattern of the Linear Polyketide Intermediates. The isolation of **8**, along with small amounts of **9**, from strains bearing the *tcmKLM* genes encoding the polyketide synthase core proteins is remarkable because **8** apparently results from folding the linear decaketide intermediate at C-9 instead of C-11, the position intrinsic to the formation of **2** prior to the subsequent ring-forming cyclizations. This unexpected regiospecificity for the first cyclization, although precedented,[38,41] was completely diverted to the C-11 based folding pattern by the addition of the *tcmN* or *tcmJN* genes to the same construct, as evident in the isolation of **2** from strains carrying pWHM722 (*tcmKLMN*) or pWHM732 (*tcmJKLMN*) (Figure 3). Do these results imply that the β-ketoacyl:acyl carrier protein synthase, chain length factor, and acyl carrier protein proteins of a given type II polyketide synthase only specify the synthesis of the linear polyketide intermediate, whose subsequent cyclizations are governed completely by the polyketide cyclases such as TcmJ and TcmN to give a particular aromatic carbon skeleton? If so, in the absence of such cyclases, the linear decaketide intermediate formed by the TcmK, -L, and -M proteins must undergo spontaneous cyclizations driven by either its kinetic reactivity or a thermodynamically-determined property of the cyclized products to give a mixture of different carbon skeletons, such as **8** and **9**. This idea is consistent with the isolation of **6** and **7** from strains bearing pWHM766 containing *actI-ORF1/actI-ORF2/tcmM* because each of the latter compounds arises from a different folding pattern of the nascent octaketide.

To answer this question, we added *tcmN* or *tcmJ* plus *tcmN* to the pWHM766 construct to test if the folding pattern of the linear octaketide could be diverted from C-9 as in **6** or C-13 as in **7** to C-11. Indeed, both pWHM767 containing *actI-ORF1/actI-ORF2/tcmMN* and pWHM769 containing *tcmJ/actI-ORF1/actI-ORF2/tcmMN* synthesized a new metabolite, **5**, which reacted readily with CH_3OH or CH_3CH_2OH to form adducts **10** and **11**, respectively. The mass spectra of **5**, **10**, and **11** have been discussed above and Table 4 summarizes the ^1H and ^{13}C NMR data of **10** and **11**. The [1-^{13}C]- and [2-^{13}C]acetate labeling experiments specifically enriched **11** at eight alternate carbons with the two carbons at 55.05 and 18.51 ppm unlabeled, confirming that **5** is derived from eight acetate equivalents with the two unlabeled carbons in **11** derived from CH_3CH_2OH. The one bond $^{13}C-^{13}C$ coupling constants for **11** were determined through a [1,2-$^{13}C_2$]acetate feeding experiment that resulted in the assignment of the eight pairs of acetate units (Figure 4A). The $^1H/^1H$ and $^1H/^{13}C$ HETCOR spectra of **11** enabled the assignment of the H−C connectivities. These assignments were unequivocally confirmed by heteronuclear multiple bond correlation (HMBC) studies (Figure 4B), resulting in the structure of **11** as shown. It is apparent that **5** will readily react with alcohols at its pyrone moiety to form the corresponding ketals such as **10** or **11**, particularly in the presence of acids like AcOH, which was used during the isolation. The structure of **5**, therefore, supports the hypothesis that a type II polyketide cyclase can determine the folding pattern of the linear polyketide intermediate, as observed with the TcmN cyclase that initiated the C-11 folding pattern for both a decaketide and an octaketide intermediate.

Discussion

Studies of the genetics and biochemistry of the biosynthesis of Tcm C (**1**) in *S. glaucescens*, along with parallel studies of actinorhodin biosynthesis in *S. coelicolor*, have provided a paradigm for the biosynthesis of aromatic polyketide metabolites

(40) Fu, H.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 4166−4170.
(41) Fu, H.; Hopwood, D. A.; Khosla, C. *Chem Biol.* **1994**, *1*, 205−210.

Table 2. Complementation of Tcm PKS Mutants by Cassettes of Hybrid *act/tcm* PKS Genes

	pWHM762	pWHM763	pWHM764	pWHM765	pWHM750	pWHM751	pWHM752	pWHM732
WMH1061	–	–	–	**1** (trace)	**1**	**1**	**1**	**1**
WMH1068	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
WMH1077	–	–	–	–	–	–	**2/5** (trace)	**2**

Table 3. Engineered Biosynthesis of Polyketides by *S. glaucescens* and *S. lividans* Strains Bearing Hybrid *act/tcm* PKS Genes

plasmid	polyketide identified (mg/L)
pELE37	**8** (1),[a] **9** (trace)
pWHM766	**6**, **7** (trace)
pWHM731	**8** (5),[b] **9** (<1)[b]
pWHM768	**6** (100–150),[b] **7** (15–25)[b]
pWHM722	**2** (1–3)[b]
pWHM767	**5** (trace)
pWHM732	**2** (30–50)[a]
pWHM769	**5** (100–150)[b]

[a] Yield was estimated on the basis of the isolated yield by HPLC analysis. [b] Isolated yield.

in general and for the enzymatic mechanism of a type II polyketide synthase in particular.[4–9] Our isolation of a putative fatty acid synthase cluster from *S. glaucescens* and the subsequent characterization of the FabD malonyl CoA:acyl carrier protein acyltransferase led us to propose a linkage between polyketide and fatty acid biosynthesis.[13,25,26] The recent discovery of a similar enzyme in *S. coelicolor*,[42] along with the implication that it is vital for growth as would be expected of a fatty acid biosynthetic enzyme, supports this hypothesis. Both of these malonyl-CoA acyltransferases, genes for which apparently are lacking in all type II polyketide synthases characterized to date,[4–9] are thought to be responsible for loading the malonyl-CoA extender unit onto the acyl carrier protein of the polyketide synthase complex.

Although acetate is the most often used starter unit for the type II polyketide synthases, as in the case of **2** and **4**, other starter units have been identified.[4–9] For example, the biosynthesis of daunorubicin in *Streptomyces peucetius*[43] or oxytetracycline in *Streptomyces rimosus*[43] utilizes propionate or a presumed malonamide as the respective starter units. Our in vitro studies of the Tcm polyketide synthase have previously demonstrated that it has a high specificity for acetyl-CoA as the starter unit.[13] Nonetheless, a "defaulty" starter unit specificity has been observed by Fu et al.[44] for the core proteins of the oxytetracycline polyketide synthase because expression of their genes in the *S. coelicolor* CH999 strain resulted in the synthesis of decaketides beginning with an acetate (or possibly acetoacetate) unit, despite the fact that oxytetracycline biosynthesis normally appears to require a malonamide starter unit. We have observed similar results by replacing the TcmK and TcmL proteins of the Tcm polyketide synthase with the corresponding DpsA β-ketoacyl:acyl carrier protein synthase and DspB chain length factor components of the daunorubicin polyketide synthase of *S. peucetius*.[45] The resulting hybrid polyketide synthase used an acetate starter unit to synthesize **2** exclusively in *S. glaucescens* WMH1077,[46] reinforcing the idea that the specificity for the starter unit is not determined by the core proteins (β-ketoacyl:acyl carrier protein synthase, chain length factor and acyl carrier protein) of a type II polyketide synthase. (Other genes unique to the daunorubicin polyketide synthase gene clusters in *S. peucetius*[45] and *Streptomyces sp.* strain C5[47] are thought to determine the choice of the starter unit.)

The function of the β-ketoacyl:acyl carrier protein synthase component of the type II polyketide synthases has been established by site-directed mutagenesis studies on both the *tcmK*[48] and the *actI-ORF1*[49] genes. However, it still is not certain that the chain length factor alone determines the chain length, i.e., the number of times that the malonyl-CoA chain extender unit reacts with the acyl-*S*Enz intermediate presumed to be bound to the β-ketoacyl:acyl carrier protein subunit (TcmK or ActI-Orf1). An *actI-ORF1* mutant was complemented *in trans* by the β-ketoacyl:acyl carrier protein synthases from several other type II polyketide synthases,[39,50] but the granaticin *graI-ORF2* biosynthesis gene was the only one among three chain length factor-encoding genes tested that complemented an *actI-ORF2* mutant.[39] The latter result is not surprising because the earliest steps of actinorhodin and granaticin biosynthesis are identical, and thus a hybrid polyketide synthase constituted from such heterologous subunits is likely to be functional. In contrast, MacDonald and co-workers have reported that hybrid polyketide synthases formed from *tcmK* and *actI-ORF2* or *actI-ORF1* and *tcmL* in *S. coelicolor* CH999 are nonfunctional[27] and that plasmid pRM8,[37] which contains the *fren-ORFI* gene encoding a possible octa- and nonaketide-specific β-ketoacyl:acyl carrier protein synthase plus *actI-ORF2*, resulted in the production of **3a** and **3b** in *S. coelicolor*. Consequently, the isolation here of decaketide **2** in addition to octaketide **5** from a plasmid bearing the *tcmK* and *actI-ORF2* genes (pWHM752, Table 2) is very intriguing because it represents a relaxed chain length specificity of the heterologous β-ketoacyl:acyl carrier protein synthase/chain length factor pair different from the systems described by McDonald et al.[27,37] Complementation of the *tcmL* mutation (in which a major portion of the coding region is deleted) with plasmids containing the *tcmK/actI-ORF2* pair supports this observation because the production of **1** must have resulted from a decaketide, not an actinorhodin-like octaketide intermediate. We therefore conclude, in agreement with Khosla, Hopwood, and co-workers,[27] that the hybrid polyketide synthase constructs harboring the β-ketoacyl:acyl carrier protein synthase/chain length factor-encoding genes from the same gene cluster in a bacterium consistently synthesize a polyketide of a defined length, even though the final structure of the polyketide metabolite initially formed can vary as the function of additional activities such as a β-ketoacyl reductase or polyketide cyclase. However, we emphasize that the collective results suggest that it is the joint activities of β-ketoacyl:acyl carrier protein synthase and chain length factor proteins that determine the size of the initial poly-β-ketone intermediate, rather just the chain length factor alone, although the behavior of the TcmK/ActI-Orf2 pair shows that the latter type of protein certainly has a major role in determining the chain length, as originally noted.[27] We presume that its

(42) Revill, P. W.; Bibb, M. J.; Hopwood, D. A. *J. Bacteriol.* In press.
(43) Hutchinson, C. R. in ref 8, pp 331–357.
(44) Fu, H.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 6443–6444.
(45) Grimm, A.; Madduri, K.; Ali, A.; Hutchinson, C. R. *Gene* **1994**, *151*, 1–10.
(46) Meurer, G.; Hutchinson, C. R., *J. Am. Chem. Soc.*, **1995**, *117*, 000–000.
(47) Ye, J.; Dickens, M. L.; Plater, R.; Li, Y.; Lawrence, J.; Strohl, W. R. *J. Bacteriol.* **1994**, *176*, 6270–6280.
(48) Meurer, G.; Hutchinson, C. R. *J. Bacteriol.* **1995**, *177*, 477–481.
(49) Kim, E.-S.; Cramer, K. D.; Shreve, A. L.; Sherman, D. H. *J. Bacteriol.* **1995**, *177*, 1202–1207.
(50) Sherman, D. H.; Kim, E.-S.; Bibb, M. J.; Hopwood, D. A. *J. Bacteriol.* **1992**, *174*, 6184–6190.

Table 4. ^{1}H and ^{13}C NMR Data of **10** and **11** in DMSO-d_6

	10		**11**		
carbon[a]	^{13}C (75.5 MHz) δ	^{1}H (300.1 MHz) δ, m, J (Hz)	^{13}C (75.5 MHz) δ	J_{cc} (Hz)	^{1}H (300.1 MHz) δ, m, J (Hz)
1	165.42		166.39	80.25	
2	88.63	5.30 s, 1H	87.33	80.25	4.88 s, 1H
3	172.64		176.65	51.99	
4	105.98	6.05 s, 1H	108.63	52.56	5.74 s, 1H
5	157.76		157.83	67.81	
6	120.18		120.95	67.25	
7	152.71		52.94	67.82	
8	105.98	6.78 s, 1H	05.83	67.82	6.69 s, 1H
9	138.92		138.72	58.21	
10	99.79	6.39 s, 1H	99.56	57.65	6.30 s, 1H
11	156.56		156.49	63.86	
12	99.94	6.36 s, 1H	99.83	64.42	6.30 s, 1H
13	160.32		158.83	66.68	
14	113.20		113.17	66.68	
15	136.44		136.08	42.38	
16	20.47	2.69, 3H	20.44	42.95	2.63 s, 3H
1′	48.56	3.2 s, 3H	55.05		3.44 q, 6.95, 2H
2′			18.51		1.05, t, 6.95, 3H
					9.88 br s, OHs

[a] Carbons are labeled according to Figure 4A.

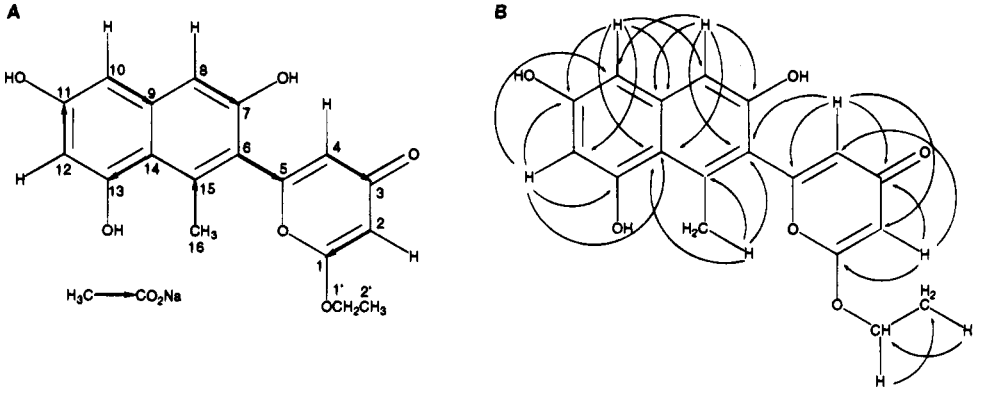

Figure 4. Specific labeling pattern of UWM3, **11**, resulting from feeding [1,2-$^{13}C_2$]acetate (**A**, boldfaced bond represents intact incorporation of an acetate unit) and the HMBC diagram of **11** (**B**).

effect is mediated through an interaction with the β-ketoacyl: acyl carrier protein synthase component in a way we do not yet understand, since no recognizable active site motif has been identified in the known chain length factor proteins. Disruption of such interaction by mismatching the β-ketoacyl:acyl carrier protein synthase/chain length factor proteins, as in the non-functional cases noted above and elsewhere,[27,37] can completely abolish the polyketide synthase activity.

Fu et al.[40,51] and McDaniel et al.[38] have postulated that the β-ketoacyl reductase enzymes that are often part of polyketide metabolism do not influence the regiospecificity of the first cyclization of the linear polyketide intermediate (contraindications of this idea are reported in the accompanying paper[52]), although a β-ketoacyl reductase protein was not always present in the type II polyketide synthases studied. They further speculated[38,40,51] that the "minimal polyketide synthase", consisting of the core proteins defined above, can dictate the regiospecificity of this cyclization, even though it was shown[38] that the presence of the *actVII* polyketide cyclase (also called AROM[53]) gene in pSEK15, which contains the *tcmKLM* and *actVII* genes, causes the synthesis of **8** and **9** in the CH999 strain to change from an approximately equal amount by the *tcmKL* and *actI-ORF3* genes in pSEK33 to the production of predominantly **8**. Hence, the polyketide cyclase must have had a profound effect on the cyclization pattern of the poly-β-ketone intermediate.[38,41] Our data favor the idea that the polyketide cyclase, not the core proteins, dictates the regiospecificity of the first cyclization of the linear polyketide. In fact, the respective minimal polyketide synthases expressed from the *tcmKLM* or *actI-ORF1, -ORF2,* and *tcmM* genes in pELE37 or pWHM766 produced a complex mixture of substances from which only **8** and **9** or **6** and **7** were isolated, respectively. The possibility that the still unidentified compounds could represent other types of foldings cannot be excluded. Therefore, we propose that the β-ketoacyl:acyl carrier protein synthase, chain length factor, and acyl carrier protein of a type II polyketide synthase are responsible only for the synthesis of the linear polyketide that, in the absence of a polyketide cyclase, cyclizes spontaneously to give a mixture of compounds whose distribution may be determined by the kinetic reactivity of the poly-β-ketone intemediate or the thermodynamic property of the products. This idea is consistent with model studies of polyketide chemistry where free polyketides, generated in situ upon deprotection, readily underwent spontaneous cyclizations.[54]

Khosla, Hopwood, and co-workers have favored the idea that the minimal polyketide synthase plays a major role in catalyzing

(51) Fu, H.; McDaniel, R.; Hopwood, D. A.; Khosla, C. *Biochemistry* **1994**, *33*, 9321–9326.

(52) McDaniel, R.; Hutchinson, C. R.; Khosla, C. *J. Am. Chem. Soc.* **1995**, *117*, 6805–6810.

(53) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 10855–10859.

(54) Harris, T. M.; Harris, C. M. *Pure Appl. Chem.* **1986**, *58*, 283–294.

the cyclization of the first ring, because in many cases the majority of the polyketides isolated from *S. coelicolor* CH999 bearing the *actI-ORF1, -ORF2,* and *-ORF3* (or functionally similar) genes, sometimes along with β-ketoacyl reductase or polyketide cyclase genes, appear to have resulted from an initial cyclization with a regiospecificity identical to that found in the biosynthesis of **3a**, **3b**, and **4** by *S. coelicolor*; for instance, as in pSEK24[38] that produced **6**. However, formation of **6** and **7** by the ActI-ORF1, -ORF2, and -ORF3[40,41] proteins, and **8** and **9** by the TcmKLM[38,40] proteins, the only examples to date of polyketides resulting from two different cyclizations in the absence of a polyketide cyclase, indicate that the initial aldol condensation can have a relaxed regiospecificity. It is difficult to tell whether this results from a mixture of enzyme-catalyzed and spontaneous events (Fu et al.[41] speculated that **7** is formed spontaneously). Therefore, even though the polyketide cyclases examined clearly have a major influence on the regiospecificity of the first cyclization, further insight is needed before we can understand how much of the biochemistry is enzyme-mediated and precisely when the first intramolecular aldol condensation takes place. The outcome of the first cyclization is likely to influence significantly the direction of subsequent cyclizations because it will constrain the number of possible choices.

We have previously speculated that TcmN and TcmJ act additively and that TcmN catalyzes later steps than TcmJ in the biosynthesis of **2** based on experiments carried out in vivo.[22] Our current data support the additivity of TcmJ and TcmN catalysis, even when the *tcmK* and *tcmL* genes are replaced by *actI-ORF1* and *actI-ORF2*, as evidenced by the effect of adding *tcmJ* to pWHM722 or pWHM767. The resulting constructs pWHM732 or pWHM769 (Table 2) synthesized at least 10-fold more **2** or **5** than pWHM722 or pWHM767, respectively. However, the addition of *tcmJ* to either the *tcmKLM* or *actI-ORF1/actI-ORF2/tcmM* genes did not alter the nature of the metabolites produced, although the yields of **8** and **9** or **6** and **7** were significantly increased. A similar effect was observed by adding *tcmJ* to either pWHM722 or pWHM767. The resulting constructs produced **2** or **5** but in a significantly increased quantity. In contrast, addition of *tcmN* to either the *tcmKLM* and *tcmJKLM* or to *actI-ORF1/actI-ORF2/tcmM* and *tcmJ/actI-ORF1/actI-ORF2/tcmM* completely altered the structures produced. The resulting plasmids synthesized **2** exclusively (pWHM722 and pWHM732) or **5** predominantly (pWHM767 and pWHM769). These data clearly indicate that TcmN directly affects the regiospecificity of the first cyclization, presumably by folding either a decaketide or octaketide intermediate at C-11 followed by an aldol condensation between C-7 and C-14. This supports our hypothesis that the polyketide cyclase component of a type II polyketide synthase determines the regiospecificity of the initial cyclization.

In the accompanying paper, McDaniel et al.[52] show by using a hybrid type II polyketide synthase produced from the *actI-ORF1, -ORF2, -ORF3,* and *tcmN* genes that the TcmN protein governs the regiospecificity of the first aldol condensation during the formation of RM77, an oxidized form of **5**. They also establish that TcmN can catalyze the aromatization of the second ring in forming the fused-ring compounds RM80 and RM80b. Taken together with the results we described earlier[21,22] and in this paper, TcmN appears to be able to catalyze stepwise each of the cyclizations needed to produce bi- and tricyclic fused-ring aromatic compounds like RM77, RM80, **5**, and **2**. It is unlikely that TcmN governs only the second ring cyclization (a conceivable reason for the formation of RM80 and RM80b[52]), but it is not clear whether TcmN can play a selective role in such cyclizations in the absence of TcmJ.

The function of TcmJ remains a mystery. Our present results are consistent with the observation[22] made in a study of a *S. glaucescens tcmJ* mutant that the *tcmJ* gene can influence the amount of **1** produced but is not required for its biosynthesis. Since TcmJ seems not to be essential for the synthesis of any metabolite in all constructs tested although it affects the amount produced significantly, we speculate that TcmJ provides optimal activity either for the polyketide synthase core proteins by facilitating their complexation or for the polyketide cyclase by bridging TcmN to the core protein complex.

The prospect of making novel compounds by metabolic pathway engineering is very appealing,[5,7,9,55] and the type II polyketide synthase is the system in which significant progress has been made recently.[27,37,38,40,41,44,46,51-53] By choosing a combination of a β-ketoacyl:acyl carrier protein synthase and chain length factor pair, an acyl carrier protein, one or more polyketide cyclases to define the length of the polyketide and the subsequent cyclization, and a β-ketoacyl reductase or genes specific for the starter unit, it is beginning to be possible to engineer the formation of novel compounds rationally. Furthermore, the results described here show that such work can be effectively pursued with different types of vectors and host strains, which should expand its scope and versatility.

Experimental Section

General Methods. All NMR spectra were taken on a Bruker AC300 spectrometer, except for the HMBC experiments that were performed on a Bruker AM600 spectrometer, and were referenced internally to the solvent. EI MS was carried out on a Kratos MS-80RFA spectrometer, electron spray MS on a Sciex API3 spectrometer, and FAB MS on a VG AutoSpec M mass spectrometer. Analytic TLC was done with precoated Keiselgel 60 F_{254} glass plate (0.25 mm) and visualized by long- and/or short-wave UV light. HPLC was done on a Waters 501 instrument (Marlborough, MA) with a Waters 484 tunable absorbance detector and a Waters Radi-Pak C_{18} (Novapak, 4 μM, 8 × 100 mm) column for analysis and a Prep Nova-Pak HR C_{18} (6 μM, 25 × 100 mm) and a guard cartridge (25 × 10 mm) of the same material for preparative scale isolation.

Bacteria Strains, Plasmids, and Other Materials. The *S. glaucescens* type strains WMH1061,[22,29] WMH1068,[22,29] and WMH1077,[22,29,30] *S. lividans* 1326,[28] and plasmids pIJ702[56] and pIJ486[57] are described elsewhere. The pGEM plasmids were purchased from Promega Corporation (Madison, WI). Thiostrepton was obtained from Sal Lucania at the Squibb Institute for Medical Research (Princeton, NJ). Unless specified, common chemicals, restriction enzymes, DNA ligase, and other materials for recombinant DNA procedures were purchased from standard commercial sources and used as provided.

Constructions of Act/Tcm Hybrid Polyketide Synthase Cassettes. Agarose gel electrophoresis, plasmid DNA isolation from *Escherichia coli*, restriction enzyme digestion, DNA ligation, and preparation of competent *E. coli* DH5α cells and their transformation were performed by established methods.[58] Plasmid DNA isolation and protoplast preparation and trans-

(55) Hutchinson, C. R. *Bio/Technology* **1994**, *12*, 375-380.
(56) Katz, E.; Thompson, C. J.; Hopwood, D. A. *J. Gen. Microbiol.* **1983**, *129*, 2703-2714.
(57) Ward, J. M.; Janssen, G. R.; Kieser, T.; Bibb, M. J.; Buttner, M. J.; Bibb, M. J. *Mol. Gen. Genet.* **1986**, 468-478.
(58) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1985.

formation in *Streptomyces* spp. were performed according to the methods of Hopwood et al.[59]

Table 1 lists plasmids used in this study. The constructions of pELE37,[10] pWHM722,[21] and pWHM732[22] were reported earlier. pWHM732 was a derivative of pWHM722 by replacing the *ermE** and the *tcmK* gene fragment with its homolog from pWHM726[22] that carried the *ermE** promoter and the *tcmJK* genes. A similar strategy was used to construct the *act/tcm* hybrid polyketide synthase plasmids of pWHM750 to pWHM769 (Figure 2) from pELE37, pWHM722, pWHM731, or pWHM732. If necessary, the junctions between the *act* and *tcm* genes were modified by the polymerase chain reaction following standard procedures.[30] These junction sequences are shown in Figure 2.

Complementation Analysis. Protoplasts of strains WMH1061 and WMH1068 were transformed with plasmids pWHM762 to pWHM765, pWHM750 to pWHM752, and pWHM732. Recombinant strains were then grown and screened for the Tcm C^+ phenotype by bioassay against a thiostrepton resistant derivative of *S. coelicolor* M111 as described previously.[21,22,29] Colonies from each transformation were also grown in 5 mL of R2YENG medium[21,22,29] containing thiostrepton (10 μg/ml) at 30 °C and 300 rpm for 3−5 days and assayed for the presence of **1** and its fluorescent biosynthetic intermediates by TLC or HPLC analysis of the EtOAc extract of the culture, according to Summers et al.[21,22,29]

Production, Isolation, and Structural Elucidation of Polyketide Metabolites. Protoplasts of *S. glaucescens* WMH-1077 and *S. lividans* 1326 were transformed with plasmids listed in Table 1, and colonies from each transformation were grown in R2YENG medium containing thiostrepton (10 μg/mL) as described above. EtOAc extracts of individual cultures were analyzed by TLC or HPLC. The TLC plate was developed in $CH_3Cl/CH_3OH/AcOH$ (87.5/10/2.5), and the C_{18} analytic HPLC column was developed with a linear gradient of CH_3CN in 0.1% AcOH aqueous solvent (see below for details).

Tcm F2 (2). The isolation and identification of **2** from WMH1077(pWHM722) and WMH1077(pWHM732) were done as reported previously.[24]

UWM1 (5), UWM2 (10), and UWM3 (11). Frozen cell suspensions of *S. glaucescens* WMH1068(pWHM769) were used to inoculate 50 mL of R2YENG medium containing thiostrepton (10 μg/mL) in a 250-mL baffled Erlenmeyer flask and incubated for 2 days at 30 °C and 300 rpm in a rotary shaker to produce a vegetative inoculum. The R2YENG fermentation medium (500 mL) with thiostrepton (10 μg/mL) in a 2-L baffled Erlenmeyer flask was inoculated with the above seed inoculum (20 mL of inoculum/500 mL of medium) and incubated under the same conditions for 3 days. The fermentation culture (usually 1−3 L) was centrifuged [4 °C, 20 min, 17 700 × g in a Sorvall RC-5B refrigerated centrifuge (Newtown, CT)] to separate the mycelia from the broth. The mycelia were extracted with acetone, and the acetone extract was vacuum concentrated (<40 °C) to give an aqueous residue. This residue was combined with the broth and extracted with EtOAc (1% AcOH) three times. The combined EtOAc extract was vacuum concentrated (<40 °C) to near dryness and loaded onto a Florisil (Aldrich, Milwaukee, WI) column (2 × 30 cm) developed in $EtOAc/CH_3OH/AcOH$ (95:5:1). Fractions containing **5** were combined and vacuum concentrated (< 40 °C) to yield a brown oil. This material was loaded on a Sephadex LH-20 (Pharmacia Biotech Inc.) column (1.5 × 75 cm), developed with CH_3OH. Fractions containing **5** were vacuum concentrated to give a light yellow powder (100−150 mg/L) as **10**: high resolution EI MS, M^+ 314.0818 ($C_{17}H_{14}O_6$, calc. 314.0790); ^1H and ^{13}C NMR data are listed in Table 4. If CH_3OH was omitted in the Florisil chromotography, the final product was purified as **5**: MS (electron spray), $[M+H]^+$ 301. If anhydrous CH_3CH_2OH was used during the purification to remove AcOH and H_2O, the final product was purified as **11**: high resolution EI MS, M^+ 328.0970 ($C_{18}H_{16}O_6$, calc. 328.0947); ^1H and ^{13}C NMR data are listed in Table 4.

[1−^{13}C]-, [2−^{13}C]-, and [1,2−$^{13}C_2$]Acetate Labeling of 11 in *S. glaucescens* WMH1068(pWHM769). [^{13}C]-Sodium acetate (Aldrich, Milwaukee, WI) was dissolved in 4 mL of H_2O and sterilized by filtration. Two batches (500 mL) of the R2YENG fermentation medium with thiostrepton (10 μg/ml) in 2-L baffled Erlenmeyer flask were inoculated with the seed inoculum and incubated for 3 days as described above. The [^{13}C]-acetate solution was added to the fermentation in equal portions at 0 and 12 h after inoculation, and the ^{13}C-enriched **11** was purified as described above. Percentage enrichments were estimated from carbon signals by comparison with that of the natural abundance. [1−^{13}C]Acetate feeding (1.5 g/L) resulted in the isolation of **11** (108 mg/L) with 2.5−3% enrichment specifically at carbons 1, 3, 5, 7, 9, 11,13, and 15. [2-^{13}C]Acetate feeding (2 g/L) yielded **11** (66.8 mg/L) with 9−11% enrichment specifically at carbons 2, 4, 6, 8, 10, 12, 14, 16. [1,2−$^{13}C_2$]Acetate feeding (2 g/L) led to the isolation of **11** (86 mg/L) with 4−5% enrichment labeled at carbons 1−16 with the coupling constants shown in Table 4.

SEK4 (6) and SEK4b (7). Frozen cell suspensions of *S. lividans* 1326(pWHM768) were used to inoculate 50 mL of R2YE medium[59] containing thiostrepton (10 μg/mL) in a 250-mL baffled Erlenmeyer flask and incubated for 3 days at 30 °C and 300 rpm in a rotary shaker to produce a vegetative inoculum. The R2YE fermentation medium (500 mL) in a 2-L baffled Erlenmeyer flask with thiostrepton (10 μg/mL) was inoculated with the above seed inoculum (20 mL of inoculum/500 mL of medium) and incubated under the same conditions for 3 days. The culture was extracted with EtOAc (1% AcOH), and the resulting extracts were fractionated by chromatography on the Florisil and the LH-20 columns sequentially as described above to give a mixture of **6** and **7**. This mixture was finally resolved on a preparative C_{18} column, developed with a linear gradient of CH_3CN from 15% to 35% in 43 min in 0.1% AcOH−H_2O at a flow rate of 5 mL/min to yield as colorless fluffy powders **6** (100 mg/L) and **7** (15 mg/L), respectively. On the analytic C_{18} column developed with a linear gradient of CH_3CN from 14% to 35% in 15 min in 0.1% AcOH−H_2O at a flow rate of 1.5 mL/min, **6** and **7** have retention times of 6.43 and 7.37 min, respectively. **6**: high resolution FAB MS, $[M+H]^+$ 319.0822 ($C_{16}H_{14}O_7$, calc. 318.0740); ^1H NMR (300.1 MHz, DMSO-d_6) δ 11.0 (br s, 2H), 6.91 (br s, 1H), 6.33 (s, 1H), 6.27 (s, 1H), 5.63 (s, 1H), 5.17 (s, 1H), 4.17 (d, J = 15.9 Hz, 1H), 4.07 (d, J = 15.8 Hz, 1H), 2.93 (d, J = 15.9 Hz, 1H), 2.55 (d, J = 16.2 Hz, 1H), 1.57 (s, 3H); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 191.1, 171.0, 165.3, 164.0, 163.0, 161.9, 138.7, 113.0, 111.4, 102.9, 100.7, 100.0, 88.1, 49.3, 37.6, 27.5. **7**: high resolution FAB MS, $[M+H]^+$ 319.0823 ($C_{16}H_{14}O_7$, calc. 318.0740); ^1H J = (300.1 MHz, DMSO-d_6) δ 11.0 (br s, 2H), 7.19 (br s, 1H), 6.31 (s, 1H), 6.21 (s, 1H), 6.05 (s, 1H), 5.15 (s, 1H), 3.10 (d, J = 14.0 Hz, 1H), 3.04 (d, J = 16.4 Hz, 1H), 3.00 (d, J = 14.0 Hz, 1H), 2.59 (d, J = 16.4 Hz, 1H), 2.50 (s, 3H); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 190.3, 172.2, 164.2, 162.7, 161.0, 160.0, 142.3, 112.6, 111.8, 104.6, 101.6, 100.2, 88.5, 47.5, 44.2, 22.5.

(59) Hopwood, D. A.; Bibb, M. J.; Chater, K. F.; Kieser, T.; Bruton, C. J.; Kieser, H. M.; Lydiate, D. J.; Smith, C. P.; Ward, J. M.; Schrempf, H. *Genetic Manipulation of Streptomyces: A Laboratory Manual*; The John Innes Foundation: Norwich, UK; 1985.

SEK15 (8) and SEK15b (9). Frozen cell suspensions of *S. glaucescens* WMH1077(pWHM731) were used to inoculate 50 mL of R2YENG medium containing thiostrepton (10 μg/mL) in a 250-mL baffled Erlenmeyer flask and incubated for 2 days at 30 °C and 300 rpm in a rotary shaker to produce a vegetative inoculum. The R2YENG fermentation medium (500 mL) with thiostrepton (10 μg/mL) in a 2-L baffled Erlenmeyer flask was inoculated with the above seed inoculum (20 mL of inoculum/500 mL of medium) and incubated under the same conditions for 3 days. The purification procedures for **8** and **9** were identical to that of **6** and **7** and preparative C_{18} HPLC finally yielded colorless fluffy powders **8** (5 mg/L) and a yellowish residue **9** (< 1 mg/L), respectively. Under the same HPLC conditions as for **6** and **7**, **8** and **9** have retention times of 9.43 and 10.43 min, respectively. **8**: high resolution FAB MS, $[M + H]^+$ 385.0920 ($C_{20}H_{16}O_8$, calc. 384.0845); ^{1}H NMR (300.1 MHz, DMSO-d_6) δ 12.13 (br s, 1H), 10.24 (br s, 1H), 6.24 (br s, 1H), 6.20 (s, 1H), 6.12 (br s, 1H), 6.09 (s, 1H), 5.66 (s, 1H), 5.07 (s, 1H), 3.50 (s, 2H), 1.83 (s, 3H); ^{13}C NMR (75.5 MHz, DMSO-d_6) δ 200.2, 171.1, 163.8, 163.6, 163.0, 162.1, 159.7, 157.3, 141.8, 135.3, 121.0, 117.1, 111.6, 110.9, 109.0, 101.6, 101.5, 88.05, 36.7, 21.0.

Acknowledgment. We thank C. Khosla at Stanford University for authentic samples of SEK4, SEK4b, SEK15, and SEK15b. The HMBC experiments were carried out in the National Magnetic Resonance Facility at Madison and the equipment in the facility was purchased with funds from the University of Wisconsin, the NSF Biological Instrumentation Program (Grant DMB-841504), the NIH Biomedical Research Technology Program (Grant RR02301), NIH Shared Instrumentation Program (Grant RR02781), and the U.S. Department of Agriculture. This study was supported by National Institutes of Health (Grant CA35381).

JA950406P