Engineered biosynthesis of novel polyketides: evidence for temporal, but not regiospecific, control of cyclization of an aromatic polyketide precursor

Hong Fu¹, David A Hopwood² and Chaitan Khosla^{1*}

¹Department of Chemical Engineering, Stanford University, Stanford, CA 94305-5025, USA and 2Department of Genetics, John lnnes Centre, Norwich NR4 7UH, UK

Background: Aromatic polyketide synthases (PKSs) catalyze the formation and cyclization of polyketide chains of variable lengths, generating a family of compounds of proven medical significance. Initial control over the regiospecificity of cyclization is believed to be exercised by the minimal PKS, composed of the three essential components for polyketide biosynthesis, which catalyzes an intramolecular aldol condensation towards the middle of the chain. Subsequent cyclization reactions are either catalyzed by additional components of the PKS, or occur in the absence of specific catalysts.

Results: Structural and biosynthetic studies on SEK4b, a novel octaketide product of a minimal PKS, revealed an unusual cyclization pattern.The first cyclization (an aldol condensation) occurs at the methyl end of the unreduced polyketide backbone precursor. This is followed by hemiketal formation and lactonization. The overall structure of SEK4b is similar to that of SEK4, a previouslyidentified product of the same genetically-engineered strain, differing only in the positions of a methyl and a pyrone group around a common fused-ring system. The biosynthetic pathways of the two molecules are quite different, however. The yield of SEK4b relative to SEK4 is much higher in the absence of PKS components (aromatases and cyclases) acting later in the pathway Conclusions: In this cyclization pathway, the regiospecificity of cyclization is not directly controlled by the minimal PKS. Instead, we propose that the enzyme influences cyclization by controlling the timing of chain release. Chain release and cyclization may be concurrent with synthesis. Other PKS subunits appear to stabilize the complex of the PKS with the nascent chain, preventing premature release.

Chemistry & Biology December 1994, 1:205-210 Key words: biosynthesis, combinatorial chemistry, polyketide, Streptomyces

Introduction

Polyketides are a large family of structurally-diverse natural products with a broad range of biological activities. These products include antibiotics such as tetracyclines and immunosuppressants such as FK506 and rapamycin. They are synthesized by multifunctional polyketide synthases (PKSs) which catalyze repeated decarboxylative condensations between acylthioesters (usually acetyl, propionyl, malonyl or methylmalonyl). In addition to varying chain length, choice of primer and extender units, and the degree and regiospecificity of reduction of the polyketide backbone, PKSs introduce structural variability into the product by catalyzing regiospecific cyclizations of nascent polyketide chains [1,2].There is considerable interest [3-131 in the possibility that genetically-engineered PKSs could be used to produce rationally-designed polyketides [14] as well as for combinatorial biosynthesis of libraries of novel molecules.

The family of bacterial aromatic PKSs has served as an excellent model system for understanding the mechanistic basis for polyketide diversity. Over the past decade the genes encoding many structurally-related aromatic PKSs have been cloned and sequenced (summarized in Fig. 1; see also [14]). Members of this class of PKSs consist of several subunits, each having distinct functions in chain assembly and modification. These studies provided the basis for the heterologous expression of a number of recombinant aromatic PKSs with various combinations of enzymatic subunits [5-11]. Structural and biosynthetic characterization of the products of such recombinant enzymes (Fig. 2) has provided new insights into the catalytic functions and specificities of aromatic PKS enzymes. In particular, proteins involved in determining carbon chain length [5,6], the degree and regiospecificity of ketoreduction [5-71, and the regiospecificity of cyclization [7,8,10,11] have been identified.

An especially striking feature of the set of polyketides produced using recombinant PKSs is the diversity of cyclization patterns undergone by polyketide chains of similar lengths. (All the polyketides in Fig. 2 are derived from octaketide, nonaketide, or decaketide backbones.) Taken together, these genetic and chemical results have led us to propose three general principles for the cyclization of aromatic polyketide precursors. First, the minimal PKS (which controls chain length and is composed of the ketosynthase, the chain-length factor, and the acyl carrier protein; Fig. 1) is the primary determinant of the regiospecificity of the first cyclization. In most cases this

^{*}Corresponding author.

genes are shown for the PKS gene clusters that produce acti-
norhodin (act) in S. coelicolor [17], tetracenomycin (tcm) in S. diverse with respect to their regiospecificities. Their norhodin (act) in S. coelicolor [17], tetracenomycin (tcm) in S. glaucescens [18-20], and frenolicin (fren) in S. roseofulvus [21].

C-7 carbonyl and the C-12 methylene [5-91. Certain Fig. 1). Alternatively, these cyclizations can occur in the

exhibit a relaxed specificity, and can catalyze a C-9/C-14 aldol condensation in addition to the more commonly observed C-7/C-12 reaction [10]. Second, certain PKS components, such as the TcmN protein (a putative bifunctional aromatase and methyltransferase) and the act ketoreductase (see Fig. l), can interact with the minimal PKS to alter significantly the relative levels of products in a branched cyclization pathway [10], or possibly even generate new patterns of cyclization (for example, a C-5/C-10 aldol condensation, as observed in compound 7 of Fig. 2) [6]. Third, the subsequent cyclization reac-Fig. 1. Bacterial aromatic PKS gene clusters. Arrangements of tions typically involve aldol condensations, hemiketal genes are shown for the PKS gene clusters that produce acti-
formation, or intramolecular lactonization, specificity can be controlled by structurally-diverse cyclases (such as the cyclase encoded in the act gene reaction occurs via an aldol condensation between the cluster and its homologs, TcmI, and possibly TcmJ; see PKSs, however, represented by the tetracenomycin PKS, absence of specific enzymatic components. In the latter

Fig. 2. Polyketide products (derived from octaketide, nonaketide, and decaketide polyketide backbones) generated from genetically engineered aromatic polyketide synthases [5-11]. Products from different recombinant PKSs of (a) the act PKS, (b) the fren PKS and (c) the tcm PKS are shown. Carbon atoms 7-12 are colored to facilitate comparison of cyclization patterns.

^bCarbons are labelled according to their number in the polyketide backbone (see Fig. 4).

situation, although there are theoretically more degrees of freedom for the cyclization of the polyketide intermediates (particularly in the case of longer chains such as decaketides), we have previously noted that the cyclization of the methyl end of the polyketide chain appears to be independent from the carboxyl end [7,8], consistent with earlier biomimetic studies [15].

Here we report the structure and deduced biosynthesis of a novel octaketide produced by two (related) recombinant PKSs. The cyclization pattern of this product is, to our knowledge, different from any natural or engineered bacterial polyketide reported thus far. Our results provide insights into a new and potentially significant way in which molecular diversity can be generated using genetically-engineered PKSs.

Results and discussion

In an attempt to characterize the properties of minimal PKSs, we constructed and analyzed a series of genetically engineered organisms containing different combinations of the PKS subunits from the act gene cluster (Fig. 1) in the Streptomyces coelicolor expression host CH999. The octaketide SEK4 (compound 8 in Fig. 2) was identified as the primary product of the minimal act PKS, and was present in cultures of CH999 cells engineered to express either just the minimal act PKS (CH999/pSEK24) [10] or the minimal act PKS in combination with act aromatase and cyclase (CH999/pSEK4) [7]. Isotopic labeling showed this compound to result from a C-7/C-12 aldol condensation, followed by two (presumably uncatalyzed) cyclization reactions. Based on chromatographic analyses of culture extracts, SEK4 was deduced to be the most abundant component in both these strains. More recently, however, we detected a second product (designated

SEK4b) in both strains.The mobility of SEK4b was identical to that of SEK4 under a variety of different chromatographic conditions, but the two molecules could be separated by differential precipitation $-$ while SEK4 precipitated from a partially purified mixture in methanol/ethyl acetate at 4 "C, SEK4b remained in solution and could be precipitated only on addition of a 1:l mixture of acetonitrile and water (see Materials and methods).The ratio of SEK4 to SEK4b was approximately 1:l in CH999/pSEK4, and 1:5 in CH999/pSEK24.

Structure of SEK4b (15)

The structure of SEK4b was characterized using NMR, high resolution mass spectroscopy (HRMS) and sodium $[1,2-13C₂]$ acetate isotopic labeling experiments. ¹H and ¹³C NMR spectral data for SEK4b are summarized in Table 1. Through a comparison with the ¹H and ¹³C NMR spectra of SEK4 [7], we deduced that both molecules contained a pyrone ring, an aromatic ring with two (uncoupled) hydrogen atoms, a six-membered heterocyclic ring fused to the aromatic ring, and two sets of diastereotopic geminal protons (with coupling constants of 14 and 16, respectively, in SEK4b). High resolution fast atom bombardment (FAB) mass spectroscopy on SEK4b gave a molecular weight of 450.9794 ($M + Cs^{+}$), consistent with the molecular formula $C_{16}H_{14}O_7$ (M + Cs^+ 450.9807).This molecular formula is identical to that of SEK4. D_2O exchange studies led to the identification of three exchangeable hydroxyls. We therefore postulated that the two molecules differed from each other simply with respect to the positions of a methyl, a hydroxyl, and a pyrone group around a common fused ring system. Sodium $[1,2^{-13}C_2]$ acetate feeding experiments indicated that the carbon chain of SEK4b was derived from eight acetate units, and also facilitated ¹³C peak assignment via the identification of adjacent carbons. All of the above data were found to be consistent with the structure for SEK4b shown in Fig. 3, in which the aromatic ring was formed by a C-10/C-15 cyclization, while the hemiacetal was formed between the C-11 hydroxyl and C-7 carbonyl.The pyrone moiety is identical in both SEK4 and SEK4b. Nuclear Overhauser effect (NOE) measurements confirmed the final structure of SEK4b (Fig. 3).

Model for the biosynthesis of SEK4b

As outlined above, studies of several natural and genetically-engineered Streptomyces aromatic polyketides have

Fig. 3. NOE Summary of SEK4b.

led to the conclusion that, while downstream cyclizations can vary tremendously, the regiospecificity of the first cyclization of the nascent polyketide backbone is controlled by the minimal PKS and typically occurs between carbon atoms towards the middle of the chain. In other words the minimal PKS constrains the polyketide chain and prevents it from sampling the extraordinarily large number of regiochemical options which might otherwise be available for cyclization. As a result it is generally believed that the entire carbon chain must be synthesized before the first cyclization reaction occurs. Here, we report the first clear example of a bacterial aromatic polyketide in which all cyclization reactions appear to be uncatalyzed.

The proposed biosynthetic pathways for SEK4 and SEK4b, based on earlier results as well as those presented here, are shown in Fig. 4. In this model for SEK4b biosynthesis, although the first cyclization presumably is not regiospecifically catalyzed by the minimal PKS, we propose that the enzyme does

influence the cyclization process, by exercising temporal control over chain release. In this model, the methyl end of the octaketide intermediate is released first from the enzyme active site and undergoes cyclization (and subsequent spontaneous aromatization via the loss of a water molecule) to form the first of the three thermodynamically stable six-membered rings. As additional carbons from the nascent chain are released into the surrounding medium, the C-7 carbonyl would become available for attack by the C-11 hydroxyl, thereby completing the formation of the fused ring system.The final step in SEK4b biosynthesis (as also in SEK4 biosynthesis) involves the formation of the pyrone ring. Although the above progressive release and cyclization could occur after the chain is complete, it is also possible that chain synthesis and cyclization are concurrent processes in the biosynthesis of this molecule; this would imply that the growing polyketide chain does not remain completely bound to the active site of the enzyme but instead is extruded into the surrounding medium in the process of being synthesized.

Fig. 4. Biosynthesis of SEK4 and SEK4b. The first cyclization is presumed to be catalyzed by the minimal PKS based upon biosynthetic studies on a number of natural and genetically-engineered polyketides ([7,8,10,11] and references therein). Reactions that are presumed to occur in solution following (in some cases partial) release of the nascent polyketide by the enzyme complex are termed 'non-catalytic'. Carbon atoms 7-12 are colored as in Fig 2. E is the PKS enzyme.

Based on the significant differences in the relative yields of SEK4 and SEK4b from CH999/pSEK4 and CH999/pSEK24, it could also be argued that the presence of downstream enzymes, such as aromatase and/or cyclase, enhances the ability of the PKS active site to bind to the nascent polyketide backbone. Recently we presented evidence suggesting that the act aromatase and/or act cyclase can also influence the catalytic specificity of the minimal PKS from the tcm gene cluster in the first cyclization reaction [lO].Taken together, these results indicate the ability of proteins such as aromatases and cyclases to associate directly with the subunits of the minimal PKS and facilitate participation in both binding and catalysis during the conversion of a highly-reactive polyketide backbone into a relatively stable aromatic product.

Finally, the isolation and characterization of SEK4b makes it clear that aromatic polyketides can be even more structurally diverse than previously realized.Our previous studies have shown that chain length, degree and regiospecificity of reduction, and regiospecificity of the first (catalyzed) and subsequent (primarily uncatalyzed) cyclizations can all be varied using genetic engineering. Here we illustrate the existence of a new degree of freedom during polyketide biosynthesis; particularly in the absence of downstream enzymes, the minimal PKS can exercise temporal but not regiospecific control over cyclization, permitting the nascent chain to sample hitherto unrecognized reaction pathways with increased frequency. In retrospect, it could be argued that SEK4b may not be the first example arising out of engineered biosynthesis that follows the above principle. For example, the structures of SEK15 (10) , RM20b (13) , and RM20c (14) (Fig. 2) would also be consistent with biosynthetic pathways that involve an initial cyclization event at the methyl end of the chain. Further analysis of the possible outcomes of cyclization during aromatic polyketide biosynthesis using the combined tools of genetics and chemistry will assist in evaluating the overall potential for combinatorial biosynthesis using this class of enzymes.

Significance

The aromatic polyketide synthases (PKSs) catalyze the biosynthesis of complex molecules with remarkable overall specificity, at the same time generating molecular diversity within a family of compounds of proven medicinal significance. Combinatorial expression of subunits of Streptomyces PKSs is a powerful method for probing the catalytic mechanisms of these enzymes. We have systematically constructed and characterized genetically engineered strains of S. coelicolor CH999 containing various combinations of PKS subunits. These studies have demonstrated that polyketide chain length, degree and regiospecificity of reduction, and regiospecificity of the first (catalyzed) and subsequent (primarily uncatalyzed) cyclizations can be varied in an increasingly predictable way. In the long run, we hope to be able to design new polyketides and produce them by engineering an appropriately modified PKS. For this to be possible, we need to understand the catalytic mechanism of the PKSs and the functions of their different subunits.

Here we have discovered an aromatic polyketide with an unusual cyclization pattern; if this unusual pattern can be reproduced in other polyketide chains, it may be that the range of aromatic polyketides that can be produced is much wider than previously thought. This unusual cyclization appears to result from partial release of the nascent (possibly incomplete) polyketide chain from the enzyme in the same temporal sequence as it is synthesized. Portions of the chain thus released presumably undergo spontaneous cyclizations, thereby leading to the formation of unusually cyclized products. It is noteworthy that a large fraction of the molecules generated using recombinant enzymes have thus far not been identified from natural sources or in man-made chemical libraries. 'Unnatural' natural products of this kind thus represent a novel source of molecular diversity, which could be used in screens for medicinal activities in the same way as conventional natural products and synthetic organic substances.

Materials and methods

Bacterial strains, plasmids, and culture conditions

The construction of strains S. coelicolor CH999/pSEK4 and CH999/pSEK24 has been described earlier [7,10]. The relevant biosynthetic genes on plasmid pSEK4 are: actI open reading frames 1-3 (whose products are the *act* ketosynthase, chain length factor, and acyl carrier protein, respectively), actVII (aromatase), and actIV (cyclase). Plasmid pSEK24 is similar, except that it lacks the actVII and actIV genes. For polyketide production, strains were grown on R2YE agar plates [16] rather than in liquid media because of the apparently more abundant production of metabolites on agar media.

Production and purification of SEK4b

CH999/pSEK4 was grown on 90 agar plates (\sim 34 ml/plate) at 30 "C for 7 days.The agar was chopped and extracted with ethyl acetate/methanol (4/l) in the presence of 1 % acetic acid (3 x 1000 ml). Following removal of the solvent under vacuum, 200 ml of ethyl acetate containing 1 % acetic acid were added.The precipitate was filtered and discarded, and the solvent was evaporated to dryness. The product mixture in ethyl acetate was applied to a Florisil column (Fisher Scientific), and eluted with ethyl acetate containing 3 % acetic acid.The first 100 ml fraction was collected and evaporated to dryness. Five ml of 50 % acetonitrile in water was added to cause the precipitation of SEK4b. The precipitate was collected by filtration and washed with ethyl acetate to give ~850 mg of SEK4b. $R_f = 0.33$ (ethyl acetate with 1 % acetic acid). Results from NMR spectroscopy on SEK4b are reported in Table 1. FAB HRMS, $M + Cs^{+}$, calculated m/e 450.9807, observed m/e 450.9794.

$[1,2-13C_2]$ acetate feeding experiments

Two 2-l flasks, each containing 400 ml of modified NMP medium [16], were inoculated with spores of S. coelicolor CH999/pSEK4. At 72 and 96 h post-inoculation, 50 mg of sodium $[1,2-^{13}C_2]$ acetate (Aldrich) was added to each flask. After 120 h, the cultures were extracted with two 500 ml volumes of ethyl acetate/l % acetic acid. The organic phase was kept, and purification proceeded as described above. Approximately 5 mg of SEK4b was obtained. 13C NMR data (2048 scans) indicated \sim 0.5-1 % enrichment (estimated by comparing peak areas to the natural abundance 13 C peak area).

Mass and NMR spectroscopy

HRMS were recorded on aVG ZAB-ZSE mass spectrometer under FAB conditions. NMR spectra were recorded on a Varian XL-400. 13 C spectra were acquired with continuous broadband proton decoupling. For NOE studies, the onedimensional difference method was employed. All compounds were dissolved in DMSO- d_6 (Sigma, 99+ atom % D) and spectra were referenced internally to the solvent. Hydroxyl resonances were identified by adding D,O (Aldrich, 99 atom % D) and checking for disappearance of signal.

Acknowledgements: We thank Koji Ichinose and Robert McDaniel for critical reading of this manuscript.This research was supported in part by a National Science FoundationYoung Investigator Award and a David and Lucile Packard Fellowship for Science and Engineering to C.K. D.A.H. acknowledges financial support from the Biotechnology and Biological Sciences Research Council and the John Innes Foundation.

References

- 1. O'Hagan, D. (1992). Biosynthesis of polyketide metabolites. Nat. Product. Reports 9, 447-479.
- 2. Katz, L. & Donadio, S. (1993). Polyketide synthesis: prospects for hybrid antibiotics. Annu. Rev. Microbiol. 47, 875-912.
- 3. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. Science 252, 675-679.
- 4. Donadio, S., McAIoine, I.B., Sheldon, P.I., lackson, M. & Katz, L. (1993). An erythromycin analog produced by reprogramming of polyketide synthesis. Proc. Natl. Acad. Sci. USA 90, 7119-7123.
- 5. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1993). Engineered biosynthesis of novel polyketides. Science 262, 1546-l 550.
- 6. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1993). Engineered biosynthesis of novel polyketides: manipulatio and analysis of an aromatic polyketide synthase with unproven catalytic specificities. J. Am. Chem. Soc. 115, 11671-11675.
- 7. Fu, H., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1994). Engineered biosynthesis of novel polyketides: dissection of the

catalytic specificity of the act ketoreductase. J. Am. Chem. Soc. 116. 41664170.

- 8. Fu, H., McDaniel, R., Hopwood, D.A. & Khosla, C. (1994). Engineered biosynthesis of novel polyketides: Stereochemical course of two reactions catalyzed by a polyketide synthase. Biochemistry 33, 9321-9326.
- 9. Fu, H., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1994). Relaxed specificity of the oxytetracycline polyketide synthase for an acetate primer in the absence of a malonamyl primer. /. Am. Chem. Soc.116,6443-6444.
- IO. McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D.A. & Khosla, C. (I 994). Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. Proc. Natl. Acad. Sci. USA in press.
- Il. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1994). Engineered biosynthesis of novel polyketides: actVll and actlV genes encode aromatase and cyclase enzymes, respectively. /. Am. Chem. Soc. in press.
- 12. Kao, C.M., Katz, L. & Khosla, C. (1994). Engineered biosynthesis of a complete macrolactone in a heterologous host. Science 265, 509-512.
- 13. Kao, C.M., Luo, C., Katz, L., Cane, D.E. & Khosla, C. (1994). Engineered biosynthesis of a triketide lactone from an incomplete modular polyketide synthase. J. Am. Chem. Soc. in press
- 14. Hopwood, D. (1993). Genetic engineering of Streptomyces to create hybrid antibiotics. Curr. Opin. Biotechnol. 4, 531-537.
- 15. Griffin, D.A., Leeper, F.j. & Staunton, I. (1984). Biomimetic syntheses of polyketide aromatics from pyrylium salts. J. Chem. Soc. Perkin Trans. I 1035-1045.
- 16. Hopwood, D.A., et al., & Schrempf, H. (1985). Genetic Manipulation of Streptomyces. A Laboratory Manual. The John lnnes Foundation, Norwich.
- 17. Fernandez-Moreno, M.A., Martinez, E., Boto, L., Hopwood, D.A. & Malpartida, F. (I 992). Nucleotide sequence and deduced functions of a set of co-transcribed genes of Streptomyces coelicolor A3(2) including the polyketide synthase for the antibiotic actinorhodin. J. Biol. Chem. 267, 19278-19290.
- 18. Bibb, M.J., Biro, S., Motamedi, H., Collins, J.F. & Hutchinson, C.R. (1989). Analysis of the nucleotide sequence of the Strepfomyces glaucescens tcm genes provides key information about the enzymology of polyketide tetracenomycin C antibiotic biosynthesis. EMBOJ. 8,2727-2736.
- 19. Summers, R. G., Wendt-Pienkowski; E., Motamedi, H. & Hutchinson, C.R. (1992). Nucleotide sequence of the tcmll-tcmlV region of the tetracenomycin C biosynthetic gene cluster of Streptomyces glaucescens and evidence that the tcmN gene encodes a multifunctional cyclase-dehydratase-O-methyltransferase. I.Bacterio/. 174, 6807-6814.
- 20. Summers, R. G., Wendt-Pienkowski, E., Motamedi, H. & Hutchinson, C.R. (1993). The tcmVl region of the tetracenomvcin C biosynthetic gene cluster of Streptomyces glaucescens encodes the tetracenomycin Fl monooxygenase,tetracenomycin F2 cyclase, and, most likely, a second cyclase. J.Bacteriol. 175, 7571-7580.
- 21. Bibb, M.J., Sherman, D.H., Omura, S. & Hopwood, D.A. (1994). Cloning, sequencing, and deduced function of a cluster of Streptomyces genes probably encoding the biosynthesis of the polyketide antibiotic frenolicin. Gene 142, 31-40.

Received: 4 Nov 1994; revisions requested: 17 Nov 1994; revisions received: 18 Nov 1994. Accepted: 18 Nov 1994.