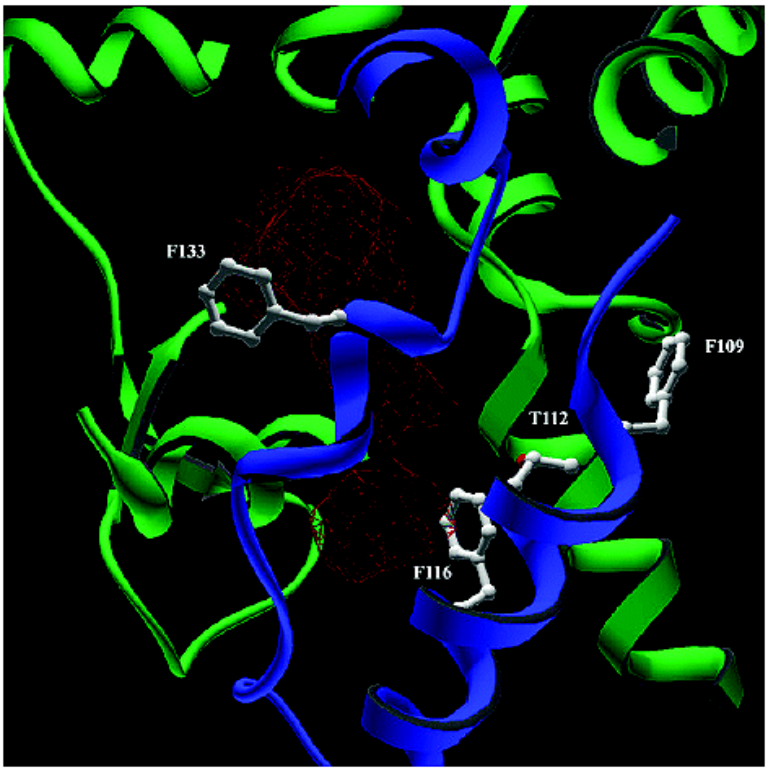
Communication

Polyketide Chain Length Control by Chain Length Factor

Yi Tang, Shiou-Chuan Tsai, and Chaitan Khosla

J. Am. Chem. Soc., **2003**, 125 (42), 12708-12709• DOI: 10.1021/ja0378759 • Publication Date (Web): 27 September 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 1 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

J|A|C|S
COMMUNICATIONS

Published on Web 09/27/2003

Polyketide Chain Length Control by Chain Length Factor

Yi Tang, Shiou-Chuan Tsai,† and Chaitan Khosla*

Departments of Chemistry, Chemical Engineering, and Biochemistry, Stanford University, Stanford, California 94305

Received August 12, 2003; E-mail: ck@chemeng.stanford.edu

Bacterial aromatic polyketides are pharmacologically important natural products.[1] A critical parameter that dictates product structure is the carbon chain length of the polyketide backbone.[2] Systematic manipulation of polyketide chain length represents a major unmet challenge in natural product biosynthesis. Polyketide chain elongation is catalyzed by a heterodimeric ketosynthase (KS/CLF, also referred to as KS_α/KS_β). In contrast to homodimeric ketosynthases found in fatty acid synthase (FAS),[3−5] the active site cysteines are absent from the CLF subunit of this heterodimer. The precise role of the catalytically silent subunit of aromatic polyketide synthases (PKSs) has been actively debated over the past decade.[6−8] The role of the CLF subunit in chain length control was originally illustrated by genetically combining heterologous KS and CLF subunits.[6] Subsequently, the CLF has been proposed to play a role in the decarboxylation of malonyl units.[8] We demonstrate here that this subunit is the primary determinant of polyketide chain length, thereby validating its designation as chain length factor. Using structure-based mutagenesis, we identified key residues in the CLF that could be manipulated to alter polyketide chain length.

A homology model of the well-studied actinorhodin (*act*) KS/CLF (which synthesizes a C_{16} chain from eight malonyl units) revealed an 18 Å long channel that starts at the conserved cysteine in the KS active site and spans across the KS-CLF dimer interface (Figure 1A; deposited at PDB: 1QXG). Residues lining the sides of the channel from both the KS and the CLF subunits are well conserved. The terminal wall of this channel is capped by a helix-turn-loop in the CLF. Sequence alignments of CLF subunits (Figure 1B) revealed that this region is highly conserved, except for residues 112 and 116 (*act* CLF numbering). Remarkably, as chain length specificity increases from C_{16} to C_{24}, these two residues are replaced with less bulky amino acids. In addition, the conserved phenylalanine residues at positions 109 and 133 in C_{16} and C_{20} KS/CLFs are changed to glycine and isoleucine, respectively, in the C_{24} *whiE* KS/CLF. We hypothesized that these four residues at the KS/CLF dimer interface modulate the length of the putative polyketide channel and determine chain length.

We constructed mutants of the *act* CLF bearing large-to-small changes at these positions using site-directed mutagenesis. The mutant CLF genes were inserted into pRM5,[6] which overexpresses the wild-type KS subunit, *act* ketoreductase (KR, *act*III), *act* aromatase (ARO, *act*VII), and *act* cyclase (CYC, *act*IV). The clones were transformed into *Streptomyces coelicolor* strain CH999,[6] and the polyketide products were analyzed by LC/MS. The strain containing the wild-type *act* KS/CLF produced the expected 3,8-dihydroxy-1-methylanthraquinone-2-carboxylic acid (DMAC). Replacing the *act* KS/CLF with wild-type tetracenomycin (*tcm*) KS/CLF (a decaketide synthase) yielded RM20b as the major product.[9] Point mutants (to alanine residues) of CLF at each of the four residues shown in Figure 1B did not yield any products of unnatural

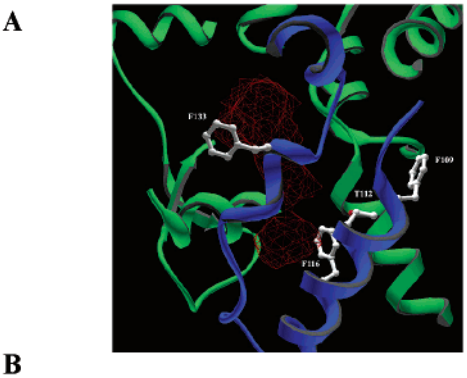


Figure 1. (A) Homology structure of the actinorhodin KS/CLF generated on the basis of the structures of several chain-extending KS homodimers. The sequences of seven KS/CLFs were submitted to the Swiss-model server. The server found 25−40% sequence identity of KS or CLF with 1E5M (KAS II from *Synechocystis* sp.),[3] 1KAS (KAS II from *E. coli*; FabF),[4] and 1F91 (KAS I from *E coli*; FabB).[5] The greatest sequence identity was with the *Synechocystis* enzyme (30−40% for KS and CLF). Models were generated with a combinatorial template of 1E5M/1KAS/1F91 for KS and 1E5M/1KAS for CLF. The four variable residues identified in (B) are shown in balls and sticks. The KS and CLF subunits are shown as green and blue ribbons, respectively. The 18 Å polyketide channel is shown in red. The terminal wall of the channel is defined by residues from the CLF. (B) Multiple sequence alignment of the helix-turn-loop regions shown in (A) from several CLFs with increasing chain length specificity. The variable residues are shown in color and are subjected to mutagenesis.

chain lengths, except for the F116A mutant, which yielded 6% RM20b. In contrast, >65% of the polyketide products of the double mutant F109A/F116A were decaketides (Figure 2A). The combined yield of octaketides and decaketides produced by the strain harboring the F109A/F116A CLF mutant was 30% of DMAC produced by the wild-type strain. Triple mutations (F109A/T112A/F116A) did not alter the specificity further. No products were observed when all four residues were mutated to alanine residues, possibly due to excessive structural perturbation at the KS/CLF interface.

To probe the roles of the corresponding residues in a decaketide CLF, small-to-large mutations at G116 and M120 (*tcm* CLF numbering) were introduced into *tcm* CLF and the resulting mutants were cloned into pSEK15,[10] which contains the minimal *tcm* PKS. The minimal PKS products of an octaketide synthase and a decaketide synthase are SEK4/SEK4b[11] and SEK15/SEK15b, respectively (Scheme 1). A single G116T mutation in the *tcm* CLF yielded >65% octaketide products with an overall polyketide yield comparable to that of wild-type *tcm* KS/CLF (Figure 2B). Thus, manipulating one or two residues located at the helix-turn-loop

† Present address: Department of Molecular Biology & Biochemistry, Department of Chemistry, University of California, Irvine, CA 92697.

10.1021/ja0378759 CCC: $25.00

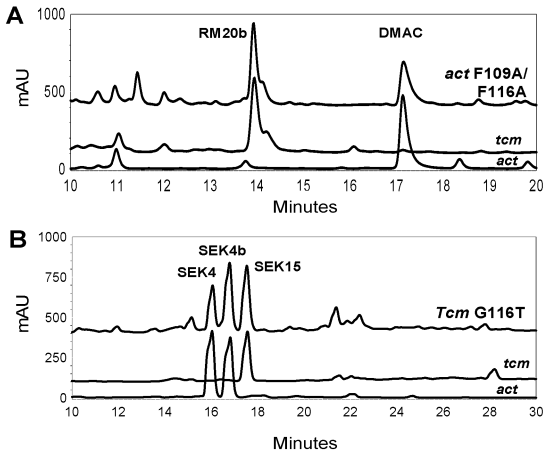
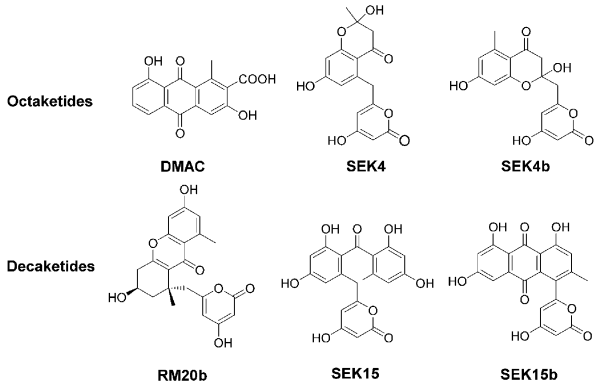


Figure 2. Distribution of polyketides produced by *S. coelicolor* CH999 expressing mutant CLF. (A) *act* ketoreductase, aromatase, and cyclase were coexpressed with the minimal PKS. Wild-type *act* KS/CLF produced DMAC, while wild-type *tcm* KS/CLF produced RM20b as a major product. The *act* CLF F109A/F116A mutant synthesized RM20b and DMAC at a ratio of 2:1. (B) Minimal PKS alone. Wild-type *act* KS/CLF produced SEK4/SEK4b, while wild-type *tcm* produced SEK15 as a major product. The *tcm* CLF G116T mutant synthesized SEK4/4b and SEK15 at a ratio of 2:1.

Scheme 1. Polyketides Observed in This Study



region is sufficient to convert an octaketide synthase into a decaketide synthase, and vice versa. However, the double mutant G116T/M120F yielded no detectable polyketides, suggesting that both residues have additional structural roles in the *tcm* KS/CLF complex.

We then purified the mutant *act* KS/CLFs from *S. coelicolor* and analyzed the distribution of polyketides produced in vitro using the minimal PKS assay.[12] Each KS/CLF mutant (FLAG tagged) was purified to >95% homogeneity.[13] The minimal PKS assay allowed us to study chain length specificity in the absence of octaketide tailoring enzymes (including *act* CYC and *act* ARO) and other cellular proteins, which may interact with KS/CLF and bias the chain length preference toward either an octaketide or a decaketide in vivo.[14] The mutant KS/CLFs exhibited significantly stronger preference toward decaketides under the assay conditions (Table 1). In the absence of downstream enzymes that may interact with KS/CLF, the point mutation F116A was sufficient to increase

Table 1. Distribution of Polyketides Produced by Mutant KS/CLF

	in vivo		in vitro	
	oct	dec	oct	dec
act WT	100	0	99	1
act CLF:F116A	94	6	36	64
act CLF:F109A/F116A	34	66	4	96

the levels of decaketides to 64% of total polyketides, while the F109A/F116A double mutant synthesized decaketides as >95% of total polyketides.

We have shown through rational mutagenesis that CLF exerts polyketide chain length control by defining the size of the polyketide channel, thereby confirming its role as the chain length factor.[6] Residues 109, 112, and 116 in the *act* CLF serve as gatekeepers that terminate the channel at the KS/CLF interface. Reducing the sizes of these residues lengthens the channel and allows two more chain-extending cycles. The channels may extend further into the CLF subunit for the C_{24} *whiE* KS/CLF and are gated by additional, unidentified residues in the C_{20} CLFs. Our results also suggest that, under in vivo conditions, additional proteins that interact with the KS/CLF further bias the formation of polyketides of a particular chain length, highlighting the complexity of protein−protein interactions among the individual type II PKS catalytic units. Our results should lead to novel strategies for the engineered biosynthesis of hitherto unidentified polyketide scaffolds.

Acknowledgment. This work was supported by a grant from the NIH (CA 77248 to C.K.). Y.T. is supported by a NIH postdoctoral fellowship.

Supporting Information Available: PDB coordinates of the homology model for the *act* KS/CLF are deposited at the Protein Data Bank (1QXG). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) O'Hagan, D. *The Polyketide Metabolites*; Ellis Howard: Chichester, UK, 1991.
(2) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *Nature* **1995**, *375*, 549−554.
(3) Moche, M.; Dehesh, K.; Edwards, P.; Lindqvist, Y. *J. Mol. Biol.* **2001**, *305*, 491−503.
(4) Huang, W.; Jia, J.; Edwards, P.; Dehesh, K.; Schneider, G.; Lindqvist, Y. *EMBO J.* **1998**, *17*, 1183−1191.
(5) Olsen, J. G.; Kadziola, A.; von Wettstein-Knowles, P.; Siggaard-Andersen, M.; Larsen, S. *Structure (Camb)* **2001**, *9*, 233−243.
(6) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *Science* **1993**, *262*, 1546−1550.
(7) Bao, W.; Wendt-Pienkowski, E.; Hutchinson, C. R. *Biochemistry* **1998**, *37*, 8132−8138.
(8) Bisang, C.; Long, P. F.; Cortes, J.; Westcott, J.; Crosby, J.; Matharu, A. L.; Cox, R. J.; Simpson, T. J.; Staunton, J.; Leadlay, P. F. *Nature* **1999**, *401*, 502−505.
(9) Fu, H.; McDaniel, R.; Hopwood, D. A.; Khosla, C. *Biochemistry* **1994**, *33*, 9321−9326.
(10) McDaniel, R.; Ebert-Khosla, S.; Fu, H.; Hopwood, D. A.; Khosla, C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11542−11546.
(11) Fu, H.; Ebert-khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 4166−4170.
(12) Carreras, C. W.; Khosla, C. *Biochemistry* **1998**, *37*, 2084−2088.
(13) Tang, Y.; Lee, T. S.; Kobayashi, S.; Khosla, C. *Biochemistry* **2003**, *42*, 6588−6595.
(14) Yu, T. W.; Shen, Y.; McDaniel, R.; Floss, H. G.; Khosla, C.; Hopwood, D.; Moore, B. S. *J. Am. Chem. Soc.* **1998**, *120*, 7749−7759.

JA0378759