Invited Contribution: 2008 Professional Progress Award

Modular Biocatalysts

Chaitan Khosla
Dept. of Chemical Engineering, Stanford University, Stanford, CA 94305
Dept. of Chemistry, Stanford University, Stanford, CA 94305
Dept. of Biochemistry, Stanford University, Stanford, CA 94305

DOI 10.1002/aic.11981
Published online June 16, 2009 in Wiley InterScience (www.interscience.wiley.com).

Modularity is a highly sought after feature in engineering design. A modular catalyst is a multi-component system whose parts can be predictably interchanged for functional flexibility and variety. Over the past two decades, much of the research in our laboratory has focused on understanding the modularity of a class of multifunctional enzymes called polyketide synthases (PKSs). PKSs catalyze the biosynthesis of a broad range of complex natural products in microorganisms, including many well-known and emerging antibiotics. A better understanding of the fundamental principles governing their modular chemistry promises to create powerful opportunities for engineering new medicines, and may even open the door to radically new catalytic processes for functionally dense, chiral synthons. © 2009 American Institute of Chemical Engineers *AIChE J,* 55: 1926–1929, 2009
Keywords: biocatalysts, modular catalyst

Modularity is a highly sought after feature in engineering design. Large-scale integrated circuits, automobile assembly lines, and multipurpose chemical plants are just some examples of the power of modular engineering. (In contrast, notwithstanding its exquisite engineering elegance, a Swiss watch is not known for its modularity.) A modular system may be defined as a multi-component system that can be divided into smaller subsystems, which interact with each other and can be predictably interchanged for *functional* flexibility and variety. The word ''functional'' merits particularly close attention—the modularity of a device is not as much about its method of preparation or its architecture as it is about its function. This is particularly relevant in the context of catalysis because a cursory scan of the broad literature on inorganic, organic, and biological catalysts might suggest otherwise.

My interest in modular biocatalysts was motivated approximately two decades ago by the discovery of the genes encoding the biosynthesis of seemingly unrelated polyketide antibiotics such as granaticin,[1] tetracenomycin,[2] and erythromycin[3,4] (Figure 1). Previous chemical analysis had established that, notwithstanding their structural differences, the biosynthesis of these natural products bears close mechanistic relationships.[5] Not only did the new genetic analysis powerfully reinforce that conclusion, but it also revealed that polyketide synthases (PKSs) are evolutionarily related multienzyme systems with potentially modular features. This was nothing short of astounding. Back then, protein engineering was in its infancy. Its application to a number of ''simple'' enzymes was teaching us that, whereas the preparative modularity of protein catalysts was boundless, their catalytic properties were not especially modular. Imagine then the surprise at the postulate that, all along while evolution had been introducing serious constraints into the modularity of monofunctional catalysts in most of the living world, it had also being playing a large-scale game of molecular Lego in a relatively esoteric metabolic niche called polyketide biosynthesis. The questions were: Was there any truth to this tantalizing hypothesis? And if yes, how simple was this game of Lego going to be?

Fast-forward 20 years. Much is known today about the scope and limitations of PKS modularity. How did this knowledge emerge? How has it influenced the engineering of new polyketide antibiotics? And what questions remain

Correspondence concerning this article should be addressed to C. Khosla at khosla@stanford.edu

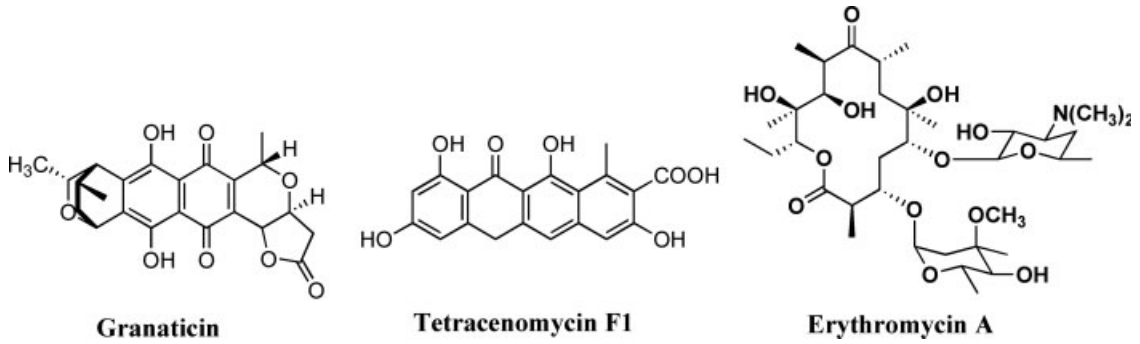
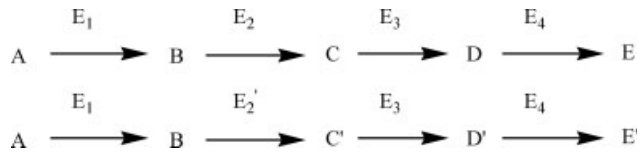
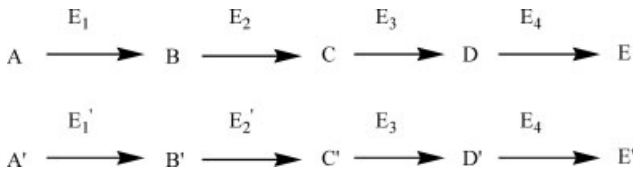


Figure 1. Structures of selected antibiotics discussed in this article.



Scheme 1. A modular biosynthetic pathway in which one catalyst (E_2) is replaced to alter a functional group or stereocenter.



Scheme 2. A modular biosynthetic pathway in which a subset of catalysts (E_1 and E_2) are replaced to enable processing of an alternative substrate (A′).

unanswered? Some general answers are presented later; for an in-depth analysis of the same issues, the reader is directed to recent reviews by the same author.[6–8]

Let us start with PKS modularity. Guided by the above definition of a modular device, the functional flexibility of a catalyst would either refer to its range of chemical transformations or the scope of its substrate tolerance. Schemes 1 and 2 illustrate two different forms of functional modularity that one desires in a modular PKS. In both schemes E_1, E_2, E_3, and E_4 are sequentially acting enzymes, and B, C, and D are intermediates in the polyketide biosynthetic pathway. In Scheme 1, E_2 is replaced with a different catalyst E_2' to alter a targeted functional group or stereocenter without affecting the rest of the natural product or the PKS turnover rate. Several examples of such PKS modifications have been reported in the literature. For example, by replacing one ketoreductase (KR) domain of the PKS system shown in Figure 2 with a KR having different stereospecificity, one predictably obtains a triketide product with one of four stereocenters inverted.[9]

Scheme 2 illustrates a distinct modular principle, in which E_1 and E_2 are replaced by E_1' and E_2', respectively, to accommodate the processing of an alternative functional group introduced via the substrate A′. Such PKS manipulations also have extensive precedence in the literature. For example, if one wishes to engineer the DMAC synthase shown in Figure 3 so that it incorporates a C_4 or C_5 substituent in place of the methyl substituent of DMAC, one must not only introduce an initiation module that synthesizes the desired acyl-ACP intermediate, but also replace the downstream KS-CLF heterodimer of the DMAC synthase with a homologous enzyme that has a longer chain length specificity (Figure 3).[10] In the example of Figure 3, it should be noted that, whereas the enzymes required for chain initiation and elongation exhibit modular characteristics, the KR, aromatase,

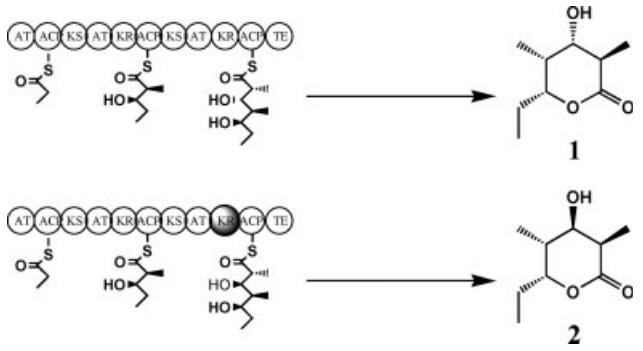


Figure 2. A truncated derivative of the erythromycin polyketide synthase is shown above with a chain initiation module [comprised acyl transferase (AT) and acyl carrier protein (ACP) domains], two chain elongation modules [each comprised ketosynthase (KS), AT, ketoreductase (KR) and ACP domains], and a terminal thioesterase that releases the polyketide chain.

Also shown above are the biosynthetic intermediates and the final lactone product (**1**) of this modular PKS. Below, a variant of the same PKS is shown in which one KR domain is replaced by a KR domain from a different PKS with opposite stereospecificity. The modified PKS exclusively synthesizes a diastereomeric lactone (**2**).

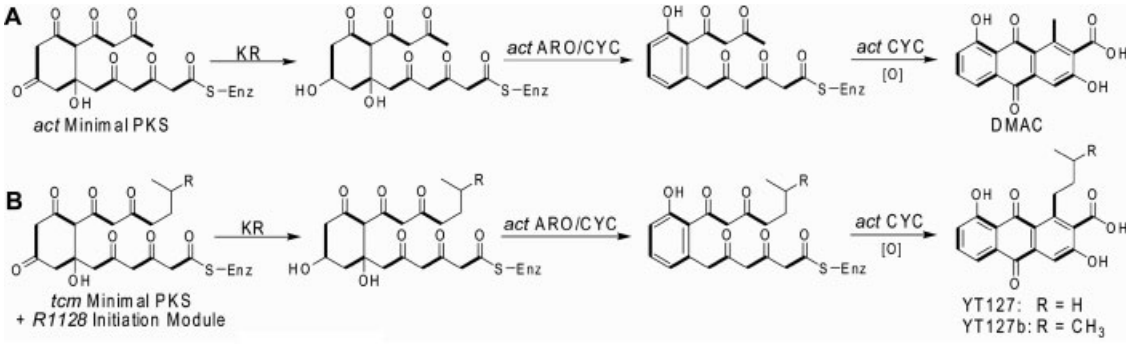


Figure 3. Aromatic polyketides produced by (A) the *act* minimal PKS, and (B) *R*1128 initiation module and *tcm* minimal PKS.

The *act* minimal PKS synthesizes a C_{16} carbon chain backbone that is primed by acetyl group. The five- and six-carbon primer unit in YT127 and YT127b, respectively, are derived from the initiation module of the R1128 PKS. In conjunction with this initiation module, the *tcm* minimal PKS therefore synthesizes an octaketide backbone rather than a decaketide backbone, so that the chain length of the resulting polyketide product remains constant. For details see (Ref. 10).

and cyclase should simply be regarded as enzymes with broad substrate tolerance; these PKS components need not be modular.

How did current knowledge of PKS modularity emerge? The earliest insights came from empirical mix-and-match experiments, such as those summarized in Figures 2 and 3. In these experiments, one introduced modifications into a target PKS at the genetic level and analyzed the implications of these modifications via fermentation and product analysis of the resulting recombinant bacterium. Such experiments proved enormously powerful, especially in the early years of molecular analysis of PKS structure and function. However, notwithstanding their power, they were limited with regards to the nature of mechanistic insights that could be derived. Two other types of experiments soon emerged as critical complements. First, PKSs were reconstituted and characterized in vitro (i.e., without the constraints of cell biology). The power of directly interrogating an isolated catalyst with a set of well-defined reagents is obvious and needs no further elaboration. Equally important was the application of high-resolution methods from macromolecular structural biology to PKSs. Together, these technological approaches have brought us to the threshold of an era when at least some types of analogs of naturally occurring polyketide antibiotics can be designed and produced in a relatively predictable fashion. Examples of antibiotics that have been extensively modified via genetic engineering include anthracycline antibiotics, macrolide antibiotics, certain ansamycin antibiotics, and the newly discovered epothilones.

Notwithstanding the above advances in our understanding of PKS modularity, much remains to be learned about the catalytic scope and limitations of these remarkable catalysts. At an architectural level, three important questions are within reach now. First, can the assembly line be deconstructed into structurally intact subsystems? Second, to what extent are subsystems from heterologous sources architecturally compatible with each other? And third, how universal are the connectors linking subsystems that we wish to interchange in order to achieve functional flexibility? Similar challenges are also encountered when one contemplates the functional modularity of PKSs. The subsystems of a digital signaling device may be architecturally modular and the connectors that interface these subsystems may also be universal. However, if the replacement of a subsystem with a functionally equivalent unit results in impedance mismatching, signal transfer across the modified interface is incomplete, and some of the signal is reflected back. By analogy, architectural modularity of a modular PKS is a necessary but of itself insufficient condition for the design of kinetically competent hybrid PKSs. Developing methods for accurately predicting the scope and limitations to the functional modularity of PKS domains and modules, as illustrated in Schemes 1 and 2, are arguably the most far-reaching goal for this field of research.

In summary, whereas a number of studies over the past two decades have highlighted the immense modular potential of naturally occurring PKSs, the foundational knowledge for their rational exploitation is only just beginning to emerge, even for prototypical systems that have been extensively investigated. Given the embarrassment of riches that awaits exploitation within nature's vast repertoire of modular PKSs, one clearly cannot wait for a case-by-case dissection of individual PKS structures and properties. The challenge therefore lies in the development of predictive algorithms that exploit this modularity. This is where chemical engineers can be expected to shine in the coming decades.

Acknowledgments

Research on polyketide synthases in the author's lab has been supported by grants from the National Institutes of Health (CA 66736 and CA 77248).

Literature Cited

1. Sherman DH, Malpartida F, Bibb MJ, Kieser HM, Bibb MJ, Hopwood DA. Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tu22. *EMBO J*. 1989;8:2717–2725.
2. Bibb MJ, Biro S, Motamedi H, Collins JF, Hutchinson CR. Analysis of the nucleotide sequence of the *Streptomyces glaucescens* tcmI genes provides key information about the enzymology of polyketide antibiotic biosynthesis. *EMBO J*. 1989;8:2727–2736.

3. Cortes J, Haydock SF, Roberts GA, Bevitt DJ, Leadlay PF. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature*. 1990;348:176–178.
4. Donadio S, Staver MJ, McAlpine JB, Swanson SJ, Katz L. Modular organization of genes required for complex polyketide biosynthesis. *Science*. 1991;252:675–679.
5. O'Hagan D. *The Polyketide Metabolites*. Chichester, U.K.: Ellis Horwood, 1993.
6. Khosla C, Tang Y, Chen AY, Schnarr NA, Cane DE. Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem*. 2007;76:195–221.
7. Khosla C, Kapur S, Cane DE. Revisiting the modularity of modular polyketide synthases. *Curr. Opin. Chem. Biol*. 2009;13:135–143.
8. Das A, Khosla C. Biosynthesis of aromatic polyketides in bacteria. *Acc. Chem. Res*. In press.
9. Kao CM, McPherson M, McDaniel RN, Fu H, Cane DE, Khosla C. Alcohol stereochemistry in polyketide backbones is controlled by the β-ketoreductase domains of modular polyketide synthases. *J Am Chem Soc*. 1998;120:2478–2479.
10. Tang Y, Lee TS, Khosla C. Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases. *PLoS Biol*. 2004;2:227–238.

Manuscript received Apr. 21, 2009