

Polyketide and non-ribosomal peptide synthases: Falling together by coming apart

C. Richard Hutchinson*

Kosan Biosciences, 3832 Bay Center Place, Hayward, CA 94545

Microorganisms have invented many devious ways to thwart their foes and survive in adverse environments. Non-motile microbes, especially ones that inhabit soil and marine environments, are noted for an ability to produce a wide range of chemicals (1). These are usually called “secondary metabolites,” because of their seeming dispensability for the organism’s ontogeny. It is obvious from the complex structures of such metabolites that their production involves unusual biochemistry as well as complex genetics devoted to the property of self-defense, intercellular communication, and other aspects of microbial life.

The polyketide synthases (PKSs) and non-ribosomal peptide synthases (NRPSs) (2) constitute a class of multifunctional proteins that govern complex enzymology with fascinating mechanisms and great commercial appeal, because the enzymes create a multitude of secondary metabolites, many of which have become important drugs. These enzymes represent some of the largest proteins known, and, as a consequence of their multifunctional character, a single protein can catalyze dozens of discrete biochemical reactions. For example, the rapamycin PKS (3) and cyclosporin NRPS (4) catalyze 51 and 40 steps in assembly of their respective products. Before the revelation about the leinamycin PKS (5) featured in this issue, type I modular PKSs (6, 7) were known to have the following characteristics. Integral to all of them is the ability to recognize simple fatty acid CoA esters and transfer them, by the acyltransferase (AT) domain, onto an acyl carrier protein (ACP) domain; and to join two thioester intermediates, one bound to the ACP and the other to the ketosynthase (KS) domain, coincident with decarboxylation of the ACP-bound substrate, to form an enzyme-bound β -keto thioester. If the product of these reactions is not released from the PKS, usually by the action of a thioesterase (TE) domain situated at the C terminus of a module, it can undergo a series of optional reactions: reduction of the β -keto group by a ketoreductase (KR) domain, dehydration of the resulting β -hydroxyl group by a dehydratase (DH) domain, and further reduction of

the resulting α,β -unsaturated thioester intermediate by an enoyl reductase (ER) domain. Methylation of the α -carbon of the thioester by a methyltransferase (MT) domain, probably at the stage of the β -keto thioester intermediate, is also possible. The key feature of modular PKSs is the organization of their domains, from the basic three (KS + AT + ACP) enzymatic activities to as many as eight activities for the full set of possibilities, into distinct modules. In turn, from one to six modules are organized into protein subunits that make up the PKS complex.

Is this new paradigm for modular PKSs truly novel and unanticipated?

Evolution presumably favored development of the multifunctional modular PKSs, because tethering of all intermediates to the PKS subunits should facilitate intramolecular transfer of substrates from one domain to the next [“substrate channeling” (8)] in the sequence of biochemical reactions.

Type I PKSs are common in bacteria; e.g., whole genome sequencing has shown that *Streptomyces avermitilis* (9) and *Streptomyces coelicolor* (10) contain eight and three type I PKSs, respectively. The 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea* (refs. 6 and 7; Fig. 1) is the archetype of modular PKSs. One of its modules has only three domains [KS, AT and ACP, believed until the work of Cheng *et al.* (5) to be the minimal number for a functional module], and others have from four to six domains, as shown. (The N-terminal didomain containing and AT and ACP is considered part of module 1.) In this PKS, the ATs all select the same substrate, methylmalonyl-CoA (mM-CoA), and as a consequence of the processive nature of the PKS, utilization of this substrate causes the polyketide chain to grow by two carbons per cycle and creates the branched, methyl-bearing carbons found in the product, 6-deoxyerythronolide B. (These

processes are highlighted for DEBS1 in Fig. 1.) Other kinds of modular PKSs use one or more of a set of closely related acyl-CoA substrates [e.g., mM-CoA, malonyl-CoA (m-CoA), methoxymalonyl-CoA (mX-CoA), and ethylmalonyl-CoA (eM-CoA)] in the same way, as in the formation of the lactone precursor of midecamycin from these substrates (Fig. 1). Sequence comparisons have established unique amino acid signatures within the AT domains for recognition of the m-CoA, mM-CoA (11, 12), and mX-CoA substrates (R. Reid, personal communication).

Cheng *et al.* (5) report the unexpected finding for leinamycin biosynthesis by *Streptomyces atroolivaceus* that modular PKSs can function without any of the modules having the cognate AT domain. Instead, a universal AT domain, present in a bifunctional protein, LnmG, with both AT and putative oxidation properties (the latter is incidental to the function of the PKS), serves both the substrate discrimination and loading functions for the entire PKS. That is, the mobile AT services all modules within the leinamycin PKS, including those that have KR, DH, and MT domains in addition to the KS and ACP domains.

This observation, supported by enzymatic experiments showing that the expressed LnmG protein catalyzes the transfer of m-CoA to each of the six discrete ACP domains of the leinamycin PKS, but not to the peptidyl carrier protein (PCP) of its NRPS component (5), reveals that modular PKSs do not have to have dedicated AT domains within each module, as had been assumed heretofore. In this case, the need for a mM-CoA transferase at the sixth step of the assembly process is obviated by the MT domain built into module six of the PKS (5). Introduction of the branched, methyl-bearing carbon at position six of leinamycin by loading of an mM-CoA substrate thus is superseded, a simplification that may have come about because all of the acyltransferase steps of this PKS require only m-CoA. The authors point out, on the basis of existing and some newly uncovered data, that a

See companion article on page 3149.

*E-mail: hutchinson@kosan.com.

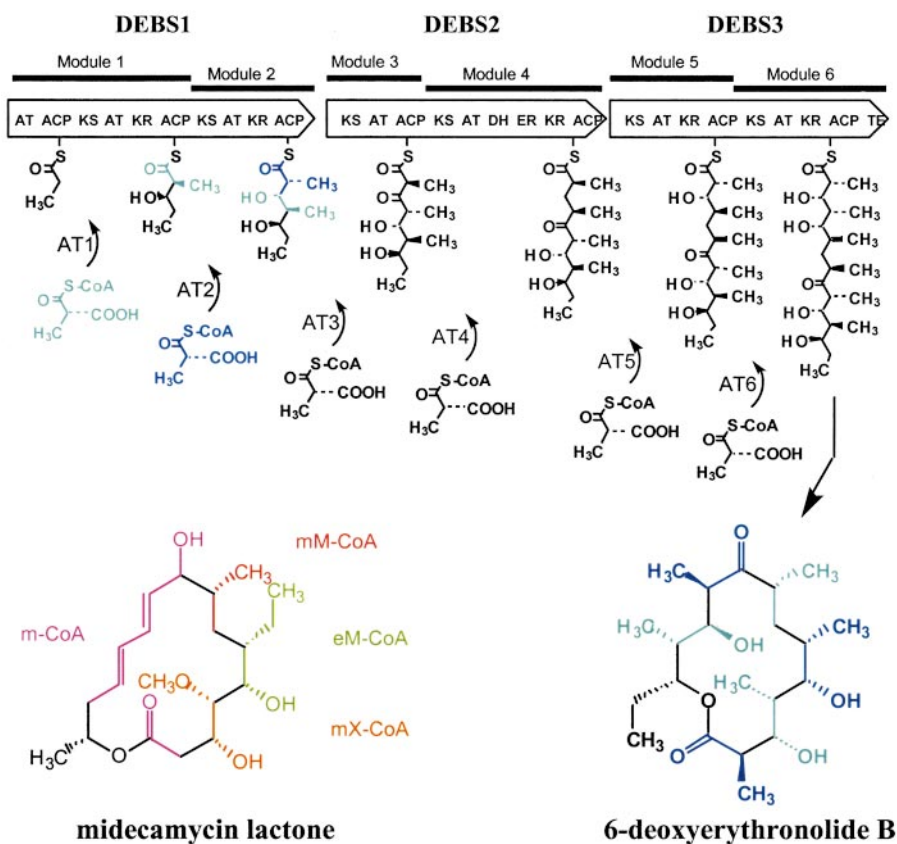


Fig. 1. The form and function of a typical modular PKS. Each of the three DEBS subunits is depicted by the open wedges, in which the type and order of the individual domains are indicated from the left (N-terminal) to right (C-terminal) ends. Modules are defined by the solid black lines above the wedges. The key intermediates of polyketide assembly are shown attached to the ACP at the point before intramolecular transfer to the succeeding KS domain. Blue color shading shows which portions of the product of DEBS1 are provided by the respective mM-CoA substrates, catalyzed by the two ATs indicated; the remainder of the enzyme-bound product is built correspondingly from four more mM-CoAs and ATs. The TE domain catalyzes the release and intramolecular cyclization of this product as 6-deoxyerythronolide B. Color coding is also used to indicate the portions of the midecamycin lactone built from each acyl-CoA substrate by a different modular PKS.

few other modular PKSs also do not have recognizably distinct AT domains within each module.

Is this new paradigm for modular PKSs (that they can consist of monofunctional proteins in contradistinction to their more characteristic polyfunctional nature) truly novel and unanticipated? No, because monofunctional proteins already had been noted as components of the NRPS class of multifunctional, modular enzymes and certain mixed NRPS/PKS systems, as well as the type II class of PKSs. For example, the enterobactin (13), pyoluteorin (14), and vibriobactin (15) NRPSs contain monofunctional adenylation (A), condensation (C), and/or PCP domains, and the bleomycin (16) and mycobactin (17) mixed NRPS/PKS systems contain monofunctional A, C, KS, or PCP domains. (“A” domains recognize amino acids and load them onto the PCP domains by formation of intermediate adenylylates, whereas “C” domains per-

form the same function as the KS domains of PKSs, joining of enzyme-bound acylthioester substrates where the

α -position of the substrates commonly contains an amino group.) In the yersiniabactin NRPS/PKS, a single A domain services three PCP domains (18). Nonetheless, the paper by Cheng *et al.* (5) contains the first bona fide evidence for how an “AT-less” modular PKS functions and supports the implication, from sequence data, that other types of monofunctional proteins can be components of such PKSs. Table 1 lists representative examples of other cases among PKSs and NRPSs where monofunctional proteins are known through experimental work or speculated to exist on the basis of sequence data. The “pksX” system, whose product still is unknown, was the first example of PKS modules that do not contain an AT domain (19), and others have been uncovered in recent years, including one from a microbial symbiont of an insect (20).

What is the significance of the new information about modular PKSs? On the one hand, it shows that modularity, as defined by the presence of at least three enzyme activities (KS, AT, ACP) within the same protein, is not the sine qua non feature of a type I PKS. This fact had been anticipated by the few modular NRPS systems where monofunctional proteins had also been verified by biochemical studies. It also brings to mind the relationship between type I, type II, and type III PKSs. In the type II case, one finds a collection of monofunctional proteins performing many of the functions ascribed to modular PKSs (AT, KS, ACP, KR, and DH activities), as in the actinorhodin (21) and tetracenomycin (22) PKSs, and for the type III class, bifunctional proteins that catalyze reactions similar to those of the type II KS and AT enzymes (23). It will be interesting to see whether

Table 1. NRPS or PKS systems with monofunctional proteins

Product	NRPS or PKS	Discrete domain	Ref.
Hypothetical compound X	PKS	ATs (encoded by <i>pksD</i> and <i>pksE</i>), KS (by <i>pksF</i>)	19
Pederin	PKS/NRPS	ATs (encoded by <i>pedC</i> and <i>pedD</i>)	20
Barbamide	NRPS/PKS/NRPS	A (encoded by <i>barD</i>), PCP (by <i>barA</i>)	32
Bacillibactin	NRPS	A (encoded by <i>dhbE</i>), ArCP (by <i>dhbB</i>)	33
Pyoluteorin	PKS/NRPS	A (encoded by <i>pltF</i>), PCP (by <i>pltL</i>)	14
Bleomycin	NRPS/PKS/NRPS	Cs (encoded by <i>blmII</i> and <i>blmXI</i>), PCP (by <i>blmI</i>)	16
Mycobactin	NRPS/PKS/NRPS	A (encoded by <i>mbtA</i>), KS (by <i>mbtC</i>)	17
Enterobactin	NRPS	A (encoded by <i>entE</i>), ArCP (by <i>entB</i>)	13
Vibriobactin	NRPS	A (encoded by <i>vibE</i>), PCP (by <i>vibB</i>), C (by <i>vibH</i>)	15
C-1027	PKS/NRPS	A (encoded by <i>sgcC1</i>), C (by <i>sgcC5</i>), PCP (by <i>sgcC2</i>)	29
Eicosapentaenoic acid	PKS*	AT (encoded by <i>orf6</i>), ER (by <i>orf8</i>)	34
Nonactin	PKS*	KSs (encoded by <i>nonJ</i> and <i>nonK</i>)	31

*These systems bear the characteristics of both type I and type II PKSs.

LnmG and other wandering ATs have sequence motifs that specify how these proteins complex with the rest of the PKS subunits, as has been found for the DEBS proteins themselves (24).

On the other hand, the work of Cheng *et al.* (5) reveals new ways to enhance or modulate the overall function of a modular PKS. An overabundance of the LnmG protein, achieved by introducing multiple copies of its gene into *S. atroolivaceus*, resulted in a 3- to 5-fold increase in leinamycin production (5). One can imagine that substitution of the gene for a particular monofunctional AT by its homolog that specifies a different substrate might result in formation of a product with the presence or absence of a branched carbon atom at one or more positions. For exam-

ple, the LnmG protein can recognize mM-CoA instead of m-CoA and catalyze its transfer to the ACP domains of the leinamycin PKS (B. Shen, personal communication). Of broader importance, however, is the prospect that the interchange (“swapping”) of different kinds of monofunctional PKS and NRPS proteins by gene engineering will be a fruitful route to the formation of analogs of known natural products and hence improved drugs for the treatment of human and animal diseases. Examples of this possibility have already been reported for certain PKS and NRPS systems that have been engineered to make novel products, largely by means of domain replacements within the multifunctional proteins (e.g., refs. 25–28).

The story of the fascinating PKS and NRPS systems continues to unfold in

remarkable ways. Genome sequencing will surely expedite further development of this area by uncovering many more “spare parts” to use for engineering new enzymes (9, 10). The promise of the so-called “combinatorial biosynthesis,” felt to be a new frontier in natural product drug discovery (25–28), depends on learning much more about the fundamental properties of such enzymes, as highlighted here, and how they can be exploited commercially. It is comforting to realize that recent discoveries about the novelty of the leinamycin (5), C-1027 (29), calicheamicin (30), and nonactin (31) PKSs forecast a rich and exciting future for this area of research into microbial biochemistry and genetics.

- Berdy, J. (1987) *CRC Handbook of Antibiotic Compounds: Microbial Metabolites* (CRC Press, Boca Raton, FL), Parts 1, 2, and 3.
- Schwarzer, D. & Marahiel, M. A. (2001) *Naturwissenschaften* **88**, 93–101.
- Schwecke, T., Aparicio, J. F., Molnar, I., Konig, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortes, J., Lester, J. B., *et al.* (1995) *Proc. Natl. Acad. USA* **92**, 7839–7843.
- Weber, G., Schorgendorfer, K., Schneider-Scherzer, E. & Leitner, E. (1994) *Curr. Genet.* **26**, 120–125.
- Cheng, Y.-Q., Tang, G.-L. & Shen, B. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 3149–3154.
- Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. (1991) *Science* **252**, 675–679.
- Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J. & Leadlay, P. F. (1990) *Nature* **348**, 176–178.
- Huang, X., Holden, H. M. & Raushel, F. M. (2001) *Annu. Rev. Biochem.* **70**, 149–180.
- Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98**, 12215–12220.
- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., *et al.* (2002) *Nature* **417**, 141–147.
- Haydock, S. F., Aparicio, J. F., Molnar, I., Schwecke, T., Khaw, L. E., Konig, A., Marsden, A. F., Galloway, I. S., Staunton, J. & Leadlay, P. F. (1995) *FEBS Lett.* **374**, 246–248.
- Kakavas, S. J., Katz, L. & Stassi, D. (1997) *J. Bacteriol.* **179**, 7515–7522.
- Gehring, A. M., Mori, I. & Walsh, C. T. (1998) *Biochemistry* **37**, 2648–2659.
- Thomas, M. G., Burkart, M. D. & Walsh, C. T. (2002) *Chem. Biol.* **9**, 171–184.
- Keating, T. A., Marshall, C. G. & Walsh, C. T. (2000) *Biochemistry* **39**, 15522–15530.
- Du, L., Sanchez, C., Chen, M., Edwards, D. J. & Shen, B. (2000) *Chem. Biol.* **7**, 623–642.
- Quadri, L. E., Sello, J., Keating, T. A., Weinreb, P. H. & Walsh, C. T. (1998) *Chem. Biol.* **5**, 631–645.
- Miller, D. A., Luo, L., Hillson, N., Keating, T. A. & Walsh, C. T. (2002) *Chem. Biol.* **9**, 333–344.
- Albertini, A. M., Caramori, T., Scoffone, F., Scotti, C. & Galizzi, A. (1995) *Microbiology* **141**, 299–309.
- Piel, J. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 14002–14007.
- Hopwood, D. A. (1997) *Chem. Rev.* **97**, 2465–2498.
- Hutchinson, C. R. (1997) *Chem. Rev.* **97**, 2525–2536.
- Jez, J. M., Bowman, M. E. & Noel, J. P. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 5319–5324.
- Gokhale, R. S., Tsuji, S. Y., Cane, D. E. & Khosla, C. (1999) *Science* **284**, 482–485.
- Rodriguez, E. & McDaniel, R. (2001) *Curr. Opin. Microbiol.* **4**, 526–534.
- Cropp, T. A., Kim, B. S., Beck, B. J., Yoon, Y. J., Sherman, D. H. & Reynolds, K. A. (2000) *Biotechnol. Genet. Eng. Rev.* **19**, 159–172.
- Walsh, C. T. (2002) *ChemBiochem* **3**, 125–134.
- Mootz, H. D., Schwarzer, D. & Marahiel, M. A. (2002) *ChemBiochem* **3**, 490–504.
- Liu, W., Christenson, S. D., Standage, S. & Shen, B. (2002) *Science* **297**, 1170–1173.
- Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R. E., *et al.* (2002) *Science* **297**, 1173–1176.
- Kwon, H. J., Smith, W. C., Scharon, A. J., Hwang, S. H., Kurth, M. J. & Shen, B. (2002) *Science* **297**, 1327–1330.