reviews

molecular pharmaceutics

Diversity of Polyketide Synthases Found in the *Aspergillus* and *Streptomyces* Genomes

James F. Sanchez,[†] Yi-Ming Chiang,[†,‡] and Clay C. C. Wang*[,†,§]

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, 1985 Zonal Avenue, Los Angeles, California 90089, Graduate Institute of Pharmaceutical Science, Chia Nan University of Pharmacy and Science, Tainan 71710, Taiwan, ROC, and Department of Chemistry, College of Letters, Arts, and Sciences, University of Southern California, Los Angeles, California 90089

Received October 20, 2007; Revised Manuscript Received February 19, 2008; Accepted February 22, 2008

Abstract: Natural products produced by biological organisms have played an important role in human health. The genomes of several *Aspergillus* species and several *Streptomyces* species have recently been sequenced. Interestingly the genomes revealed a large range of secondary metabolite genes, for many of which the products are currently unknown, suggesting that there is a wealth of secondary metabolites yet to be discovered. This brief review will discuss the current knowledge in polyketides produced by the various different classes of polyketide synthases (PKSs) present in the *Aspergillus* and *Streptomyces* genomes.

Keywords: Fungal natural product; *Aspergillus*; polyketide synthase; polyketides; secondary metabolite

Secondary Metabolites as Pharmaceutically Important Natural Products. Natural products remain one of the most important sources of chemotherapeutic agents currently in clinical use.[1,2] They are secondary metabolites from various organisms that play nonessential roles in the life-sustaining, or primary, metabolism of the organisms. These organisms have developed sophisticated biosynthetic systems for the production of natural products, partly because these metabolites provide the producing organisms a survival advantage over nonproducing counterparts. Many such natural products possess potent biological activities, such as antimicrobial, antiviral, and antitumor properties, that make them particularly valuable as pharmaceutical agents. Natural products isolated from a variety of biological species have played an important role in human medicine. Well-known examples include erythromycin (**1**), an antibiotic,[3–5] rapamycin (**2**), an immunosuppressant,[6] and lovastatin (**3**), a cholesterol-lowering medicine (Figure 1).[7]

Genome Sequencing of Natural Product Producing Organisms. What is the origin of the diverse polyketide structures currently found in medicinal uses? This is being answered as genomes of natural producing organisms are sequenced in a rapid pace. The genomes of two *Streptomyces* species, *Streptomyces avermitilis*[8] and *Streptomyces coeli-*

* Corresponding author: Clay Wang, Assistant Professor of Pharmaceutical Sciences and Chemistry, Departments of Pharmacology and Pharmaceutical Sciences and Chemistry, University of Southern California,1985 Zonal Ave PSC 716, Los Angeles, CA 90089. Tel: 323-442-1670. Fax: 323-442-1365. E-mail: clayw@usc.edu.

† School of Pharmacy, University of Southern California.

‡ Chia Nan University of Pharmacy and Science.

§ College of Letters, Arts, and Sciences, University of Southern California.

(1) Newman, D. J.; Cragg, G. M.; Snader, K. M. Natural products as sources of new drugs over the period 1981–2002. *J. Nat. Prod.* **2003**, *66* (7), 1022–37.

(2) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **2007**, *70* (3), 461–77.

(3) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **1990**, *348* (6297), 176–8.

(4) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. Modular organization of genes required for complex polyketide biosynthesis. *Science* **1991**, *252* (5006), 675–9.

(5) Kao, C. M.; Katz, L.; Khosla, C. Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science* **1994**, *265* (5171), 509–12.

10.1021/mp700139t CCC: $40.75
Published on Web 03/14/2008

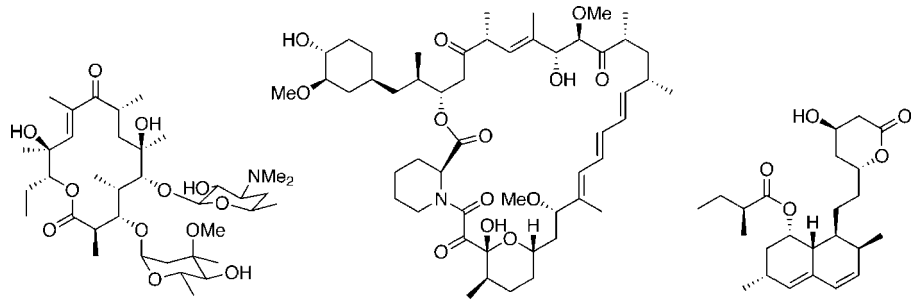

Figure 1. Chemical structures of medicinally important natural products: erythromycin A (**1**), rapamycin (**2**), and lovastatin (**3**).

color,[9] revealed that each organism contains clusters of genes responsible for the assembly, modification, and regulation of a particular natural product and its related compounds. What is interesting from a drug discovery perspective is the fact that the large number of gene clusters in these two organisms that contain polyketide synthases or nonribosomal peptide synthases greatly exceeds the number of known natural products from these organisms. This disparity between known compounds and number of gene clusters is not unique among *Streptomyces* species. Analysis of the genomes of eight *Aspergillus* species, namely, *Aspergillus clavatus, Aspergillus flavus, Aspergillus fumigatus*,[10] *Aspergillus nidulans*,[11] *Aspergillus niger*,[12] *Aspergillus oryzae*,[13] *Aspergillus terreus*,[14] and *Neosartorya fischeri* (also known as *Aspergillus fischeri*), also revealed that a large number of gene clusters contain polyketide synthases and nonribosomal peptide synthetases. The number greatly exceeds the number of known natural products from these organisms. Another interesting observation from all the genome sequencing is the high number of gene clusters that are unique among the various genomes. Therefore these natural product producing organisms have evolved to produce a wide variety of compounds, and interestingly they also evolved to produce different natural products to suit their unique biological requirements. Exploitation of this genomic information creates an exciting new avenue for drug discovery.[15,16]

Secondary Metabolite Genes in *Aspergillus* and *Streptomyces*. As mentioned in the previous paragraph, analysis of the eight species of the fungal genus *Aspergillus*

(6) 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. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* (17), 7839–43.

(7) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **1999**, *284* (5418), 1368–72.

(8) Omura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. Genome sequence of an industrial microorganism Streptomyces avermitilis: Deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (21), 1221520. Epub 2001 Sep 25.

(9) 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.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* **2002**, *417* (6885), 141–147.

(10) Nierman, W. C.; Pain, A.; Anderson, M. J.; Wortman, J. R.; Kim, H. S.; Arroyo, J.; Berriman, M.; Abe, K.; Archer, D. B.; Bermejo, C.; Bennett, J.; Bowyer, P.; Chen, D.; Collins, M.; Coulsen, R.; Davies, R.; Dyer, P. S.; Farman, M.; Fedorova, N.; Fedorova, N.; Feldblyum, T. V.; Fischer, R.; Fosker, N.; Fraser, A.; Garcia, J. L.; Garcia, M. J.; Goble, A.; Goldman, G. H.; Gomi, K.; Griffith-Jones, S.; Gwilliam, R.; Haas, B.; Haas, H.; Harris, D.; Horiuchi, H.; Huang, J.; Humphray, S.; Jimenez, J.; Keller, N.; Khouri, H.; Kitamoto, K.; Kobayashi, T.; Konzack, S.; Kulkarni, R.; Kumagai, T.; Lafton, A.; Latge, J. P.; Li, W. X.; Lord, A.; Majoros, W. H.; May, G. S.; Miller, B. L.; Mohamoud, Y.; Molina, M.; Monod, M.; Mouyna, I.; Mulligan, S.; Murphy, L.; O'Neil, S.; Paulsen, I.; Penalva, M. A.; Pertea, M.; Price, C.; Pritchard, B. L.; Quail, M. A.; Rabbinowitsch, E.; Rawlins, N.; Rajandream, M. A.; Reichard, U.; Renauld, H.; Robson, G. D.; de Cordoba, S. R.; Rodriguez-Pena, J. M.; Ronning, C. M.; Rutter, S.; Salzberg, S. L.; Sanchez, M.; Sanchez-Ferrero, J. C.; Saunders, D.; Seeger, K.; Squares, R.; Squares, S.; Takeuchi, M.; Tekaia, F.; Turner, G.; de Aldana, C. R. V.; Weidman, J.; White, O.; Woodward, J.; Yu, J. H.; Fraser, C.; Galagan, J. E.; Asai, K.; Machida, M.; Hall, N.; Barrell, B.; Denning, D. W. Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus. *Nature* **2005**, *438* (7071), 1151–1156.

(11) Galagan, J. E.; Calvo, S. E.; Cuomo, C.; Ma, L. J.; Wortman, J. R.; Batzoglou, S.; Lee, S. I.; Basturkmen, M.; Spevak, C. C.; Clutterbuck, J.; Kapitonov, V.; Jurka, J.; Scazzocchio, C.; Farman, M.; Butler, J.; Purcell, S.; Harris, S.; Braus, G. H.; Draht, O.; Busch, S.; D'Enfert, C.; Bouchier, C.; Goldman, G. H.; Bell-Pedersen, D.; Griffiths-Jones, S.; Doonan, J. H.; Yu, J.; Vienken, K.; Pain, A.; Freitag, M.; Selker, E. U.; Archer, D. B.; Penalva, M. A.; Oakley, B. R.; Momany, M.; Tanaka, T.; Kumagai, T.; Asai, K.; Machida, M.; Nierman, W. C.; Denning, D. W.; Caddick, M.; Hynes, M.; Paoletti, M.; Fischer, R.; Miller, B.; Dyer, P.; Sachs, M. S.; Osmani, S. A.; Birren, B. W. Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae. *Nature* **2005**, *438* (7071), 1105–15.

and the two species of *Streptomyces* have revealed that these species have secondary metabolism pathways that are far richer than previously known and thus are potentially valuable sources of new and useful natural products. Genes involved in secondary metabolism are highly conserved, and many of them can be identified in genomes from their sequences.[17–20] Interestingly, in fungi including *Aspergillus* species and bacteria including *Streptomyces* species, genes of particular secondary metabolism pathways are generally clustered such that all genes whose products are involved in the synthesis of a particular secondary metabolite are adjacent to each other in the genome.[21] The clustering of biosynthesis genes is important for the understanding of the mechanism of how the secondary metabolites are assembled and how their biosynthesis is regulated.[21–23] Whereas it is not generally possible to predict the precise product of a cluster by inspection of the gene sequences alone, it is possible to predict the class of compounds each cluster produces.[17,19] Genomic analysis in *A. nidulans* identified 27 polyketide synthase (PKS) and 14 nonribosomal peptide synthetase (NRPS) products, but thus far only sterigmatocystin (**4**), penicillin (**5**), terrequinone (**6**), triacetylfusarinine(**7**), ferricrocin (**8**), aspoquinolone (**9**), and aspyridone (**10**) are known secondary metabolites produced by this species (Figure 2).[16,24–26] The low homology between the secondary metabolite gene clusters in the eight aforementioned species suggests that each species makes its own unique set of secondary metabolites to complement its own biological traits, and thus a large number of compounds await discovery in *Aspergillus* as a whole. Genome sequencing revealed that *S. coelicolor* contains 21 natural product clusters and *S. avermitilis* contains 25 natural product clusters (type I modular and both type I and type II iterative PKS, chalcone synthase, NRPS, and terpene cyclase).[8,9,27]

(12) Pel, H. J.; de Winde, J. H.; Archer, D. B.; Dyer, P. S.; Hofmann, G.; Schaap, P. J.; Turner, G.; de Vries, R. P.; Albang, R.; Albermann, K.; Andersen, M. R.; Bendtsen, J. D.; Benen, J. A.; van den Berg, M.; Breestraat, S.; Caddick, M. X.; Contreras, R.; Cornell, M.; Coutinho, P. M.; Danchin, E. G.; Debets, A. J.; Dekker, P.; van Dijck, P. W.; van Dijk, A.; Dijkhuizen, L.; Driessen, A. J.; d'Enfert, C.; Geysens, S.; Goosen, C.; Groot, G. S.; de Groot, P. W.; Guillemette, T.; Henrissat, B.; Herweijer, M.; van den Hombergh, J. P.; van den Hondel, C. A.; van der Heijden, R. T.; van der Kaaij, R. M.; Klis, F. M.; Kools, H. J.; Kubicek, C. P.; van Kuyk, P. A.; Lauber, J.; Lu, X.; van der Maarel, M. J.; Meulenberg, R.; Menke, H.; Mortimer, M. A.; Nielsen, J.; Oliver, S. G.; Olsthoorn, M.; Pal, K.; van Peij, N. N.; Ram, A. F.; Rinas, U.; Roubos, J. A.; Sagt, C. M.; Schmoll, M.; Sun, J.; Ussery, D.; Varga, J.; Vervecken, W.; van de Vondervoort, P. J.; Wedler, H.; Wosten, H. A.; Zeng, A. P.; van Ooyen, A. J.; Visser, J.; Stam, H. Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88. *Nat. Biotechnol.* **2007**, *25* (2), 221–31.

(13) Machida, M.; Asai, K.; Sano, M.; Tanaka, T.; Kumagai, T.; Terai, G.; Kusumoto, K.; Arima, T.; Akita, O.; Kashiwagi, Y.; Abe, K.; Gomi, K.; Horiuchi, H.; Kitamoto, K.; Kobayashi, T.; Takeuchi, M.; Denning, D. W.; Galagan, J. E.; Nierman, W. C.; Yu, J.; Archer, D. B.; Bennett, J. W.; Bhatnagar, D.; Cleveland, T. E.; Fedorova, N. D.; Gotoh, O.; Horikawa, H.; Hosoyama, A.; Ichinomiya, M.; Igarashi, R.; Iwashita, K.; Juvvadi, P. R.; Kato, M.; Kato, Y.; Kin, T.; Kokubun, A.; Maeda, H.; Maeyama, N.; Maruyama, J.; Nagasaki, H.; Nakajima, T.; Oda, K.; Okada, K.; Paulsen, I.; Sakamoto, K.; Sawano, T.; Takahashi, M.; Takase, K.; Terabayashi, Y.; Wortman, J. R.; Yamada, O.; Yamagata, Y.; Anazawa, H.; Hata, Y.; Koide, Y.; Komori, T.; Koyama, Y.; Minetoki, T.; Suharnan, S.; Tanaka, A.; Isono, K.; Kuhara, S.; Ogasawara, N.; Kikuchi, H. Genome sequencing and analysis of Aspergillus oryzae. *Nature* **2005**, *438* (7071), 1157–61.

(14) Askenazi, M.; Driggers, E. M.; Holtzman, D. A.; Norman, T. C.; Iverson, S.; Zimmer, D. P.; Boers, M. E.; Blomquist, P. R.; Martinez, E. J.; Monreal, A. W.; Feibelman, T. P.; Mayorga, M. E.; Maxon, M. E.; Sykes, K.; Tobin, J. V.; Cordero, E.; Salama, S. R.; Trueheart, J.; Royer, J. C.; Madden, K. T. Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat. Biotechnol.* **2003**, *21* (2), 150–6.

(15) Jones, M. G. The first filamentous fungal genome sequences: Aspergillus leads the way for essential everyday resources or dusty museum specimens. *Microbiology* **2007**, *153* (Pt 1), 1–6.

(16) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. Genomics-driven discovery of PKS-NRPS hybrid metabolites from Aspergillus nidulans. *Nat. Chem. Biol.* **2007**, *3* (4), 213–7.

(17) Yadav, G.; Gokhale, R. S.; Mohanty, D. SEARCHPKS: a program for detection and analysis of polyketide synthase domains. *Nucleic Acids Res.* **2003**, *31* (13), 3654–3658.

(18) Cane, D. E.; Walsh, C. T.; Khosla, C. Harnessing the biosynthetic code: Combinations, permutations, and mutations. *Science* **1998**, *282* (5386), 63–8.

(19) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res.* **2005**, *33* (18), 5799–808.

(20) Udwary, D. W.; Merski, M.; Townsend, C. A. A method for prediction of the locations of linker regions within large multifunctional proteins, and application to a type I polyketide synthase. *J. Mol. Biol.* **2002**, *323* (3), 585–98.

(21) Brown, D. W.; Yu, J. H.; Kelkar, H. S.; Fernandes, M.; Nesbitt, T. C.; Keller, N. P.; Adams, T. H.; Leonard, T. J. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in Aspergillus nidulans. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (4), 1418–22.

(22) MacCabe, A. P.; van Liempt, H.; Palissa, H.; Unkles, S. E.; Riach, M. B.; Pfeifer, E.; von Dohren, H.; Kinghorn, J. R. Delta-(L-alpha-aminoadipyl)-L-cysteinyl-D-valine synthetase from Aspergillus nidulans. Molecular characterization of the acvA gene encoding the first enzyme of the penicillin biosynthetic pathway. *J. Biol. Chem.* **1991**, *266* (19), 12646–54.

(23) MacCabe, A. P.; Riach, M. B.; Kinghorn, J. R. Identification and expression of the ACV synthetase gene. *J. Biotechnol.* **1991**, *17* (1), 91–7.

(24) Bok, J. W.; Hoffmeister, D.; Maggio-Hall, L. A.; Murillo, R.; Glasner, J. D.; Keller, N. P. Genomic mining for Aspergillus natural products. *Chem. Biol.* **2006**, *13* (1), 31–7.

(25) Eisendle, M.; Oberegger, H.; Zadra, I.; Haas, H. The siderophore system is essential for viability of Aspergillus nidulans: functional analysis of two genes encoding l-ornithine N 5-monooxygenase (sidA) and a non-ribosomal peptide synthetase (sidC). *Mol. Microbiol.* **2003**, *49* (2), 359–75.

(26) Scherlach, K.; Hertweck, C. Discovery of aspoquinolones A-D, prenylated quinoline-2-one alkaloids from Aspergillus nidulans, motivated by genome mining. *Org. Biomol. Chem.* **2006**, *4* (18), 3517–20.

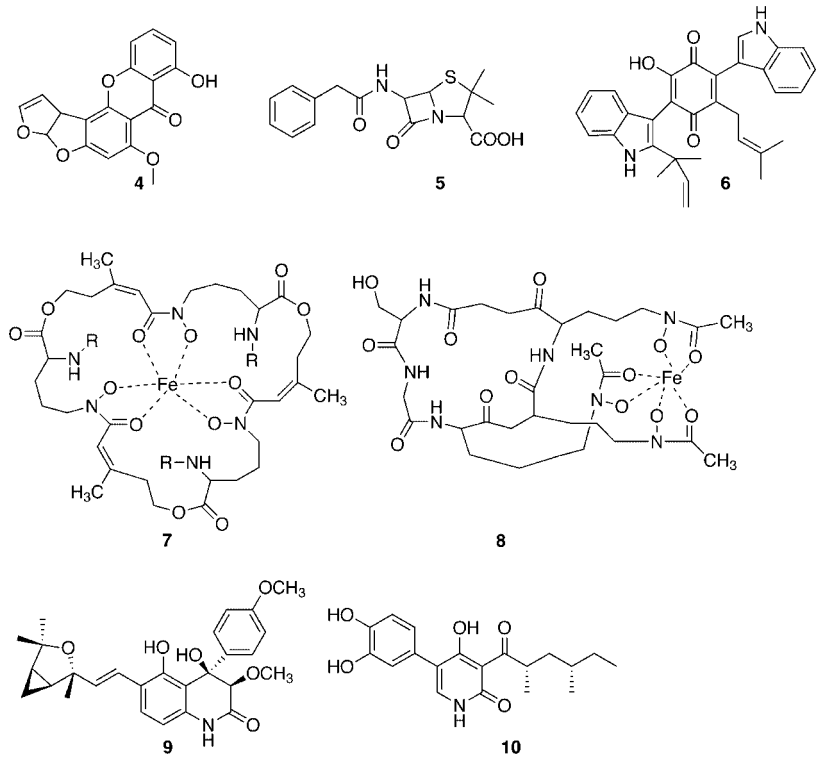

Figure 2. Chemical structures of known *A. nidulans* secondary metabolites.

There have been a number of excellent recent reviews concerning polyketides and their biosynthesis by Hill, Walsh, Cox, and Keller.[28–31] This review will not seek to provide a comprehensive evaluation on the subject; instead, the focus of this brief review will be on the different types of PKS genes found in both *Aspergillus* and *Streptomyces*. The goal is to examine PKS clusters where the products are known to serve as a roadmap for the understanding of the vast number of currently still unexplored polyketide pathways provided by the genome projects.

Organization of PKS. Polyketide synthases are responsible for the biosynthesis of polyketides.[20,32,33] These enzymes catalyze repetitive Claisen condensations to form the elongating polyketide chain. The enzyme systems are classified according to their architectural organization. In type I systems, the constituent catalytic components are covalently linked into large multifunctional proteins while in type II systems the catalytic components are free-standing. In bacterial genomes such as *Streptomyces* both modular type I and iterative type II systems exist. Fungal type I PKSs are distinct from both bacterial PKS types and are characterized by one single large multidomain enzyme similar to bacterial type I PKSs, but can be used repetitively, similar to bacterial type II PKSs. At a minimum PKSs contain ketosynthase (KS), acyl transferase (AT), and acyl carrier protein (ACP) domains. In addition, optional β-keto processing reactions may be catalyzed by the ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains. Further modifying domains include cyclase (CYC) and methyl transferase (MT) domains.[33]

In a bacterial type I system, best exemplified by erythromycin, one set of active sites, a module, is used to carry out either initiation or one round of chain extension. Each module contains KS, AT, ACP, plus modifying domains to determine the choice of substrate and degree of reduction or dehydration (Figure 3).[28,34] The number of modules determines the length of the polyketide chain length; therefore, typically type I PKS clusters are very large in size. The bacterial type II PKS contains several small enzymes, each responsible for one

(27) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nat. Biotechnol.* **2003**, *21* (5), 526–531.

(28) Hill, A. M. The biosynthesis, molecular genetics and enzymology of the polyketide-derived metabolites. *Nat. Prod. Rep.* **2006**, *23* (2), 256–320.

(29) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106* (8), 3468–96.

(30) Cox, R. J. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org. Biomol. Chem.* **2007**, *5* (13), 2010–26.

(31) Hoffmeister, D.; Keller, N. P. Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat. Prod. Rep.* **2007**, *24* (2), 393–416.

(32) Schumann, J.; Hertweck, C. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *J. Biotechnol.* **2006**, *124* (4), 690–703.

(33) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (26), 15670–5.

(34) Khosla, C. Harnessing the biosynthetic potential of modular polyketide synthases. *Chem. Rev.* **1997**, *97* (7), 2577–2590.

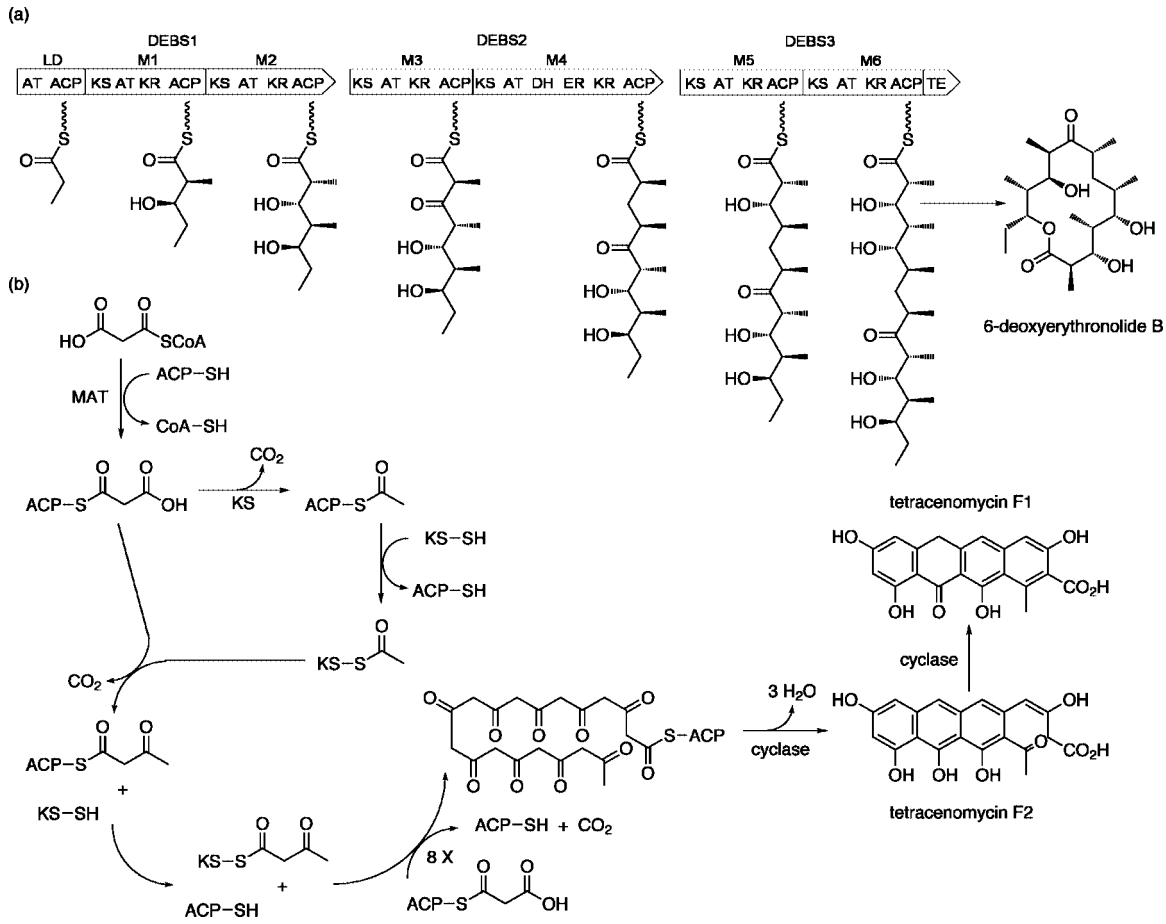

Figure 3. (a) Organization of bacterial modular type I PKS. Biosynthesis of 6-deoxyerythronolide B. (b) Organization of bacterial iterative type II PKS. Biosynthesis of tetracenomycin F1. Abbreviations: DEBS, 6-deoxyerythronolide B synthase; LD, loading domain; M, module; KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; ACP, acyl carrier protein; TE, thioesterase; MAT, malonyl-CoA ACP acyltransferase.

chemical reaction.[35,36] These enzymes are used iteratively to produce mainly aromatic polyketides such as tetracenomycin.[37,38] According to the variation in these optional processing domains the fungal PKS can be divided into three classes: nonreducing or aromatic (NR-PKS), partially reducing (PR-PKS), and highly reducing (HR-PKS) enzymes (Figure 4). Surrounding the PKS genes are additional genes coding for other modifying enzymes. In certain cases the biosynthetic cluster contains additional PKS or NRPS genes that act together to produce the natural product. By examining the organization of the PKS, it is possible to predict the class of polyketides that the PKS produces.

Bacterial Modular Type I Polyketide Synthase. The importance of modular polyketides in human health has prompted tremendous effort in understanding the mechanism for the assembly of this class of natural products. Erythromycin is the first modular polyketide of which its biosynthesis genes were discovered, by Leonard Katz and Peter Leadlay, and has served as the model for the understanding of modular polyketides (Figure 3a).[3,4] The polyketide core is assembled by three very large polypeptides denoted DEBS1, DEBS2, and DEBS3.[5] The subunit contains two unique modules defined as a set of covalently linked enzymes that are responsible for one round of polyketide chain elongation and subsequent processing of the backbone. In addition DEBS 1 contains domains necessary for the loading of the primer unit, and DEBS 3 contains a thioesterase domain for the catalytic release of the polyketide chain from the enzyme. A signature feature of a modular PKS is that adjacent modules are used sequentially in an assembly line manner, therefore analysis of bacterial genomes can easily identify a type I PKS. In addition the bioinformatic programs are available that can predict from sequence homology to the selectivity of the acyl transferase domains of each module. However because of module skipping, iterative uses

(35) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* **1995**, *375* (6532), 549–54.
(36) Khosla, C.; Harbury, P. B. Modular enzymes. *Nature* **2001**, *409* (6817), 247–52.
(37) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of Aspergillus nidulans. *Chem. Biol.* **2001**, *8* (2), 189–97.
(38) Hutchinson, C. R. Biosynthetic Studies of Daunorubicin and Tetracenomycin C. *Chem. Rev.* **1997**, *97* (7), 2525–2536.

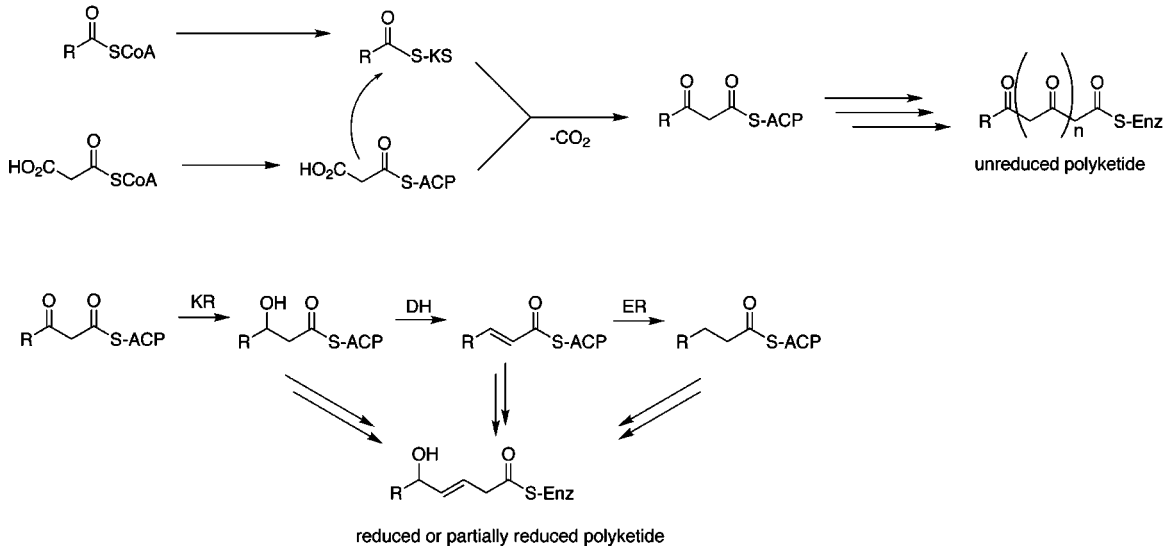

Figure 4. Chemical reaction catalyzed by iterative fungal PKS.

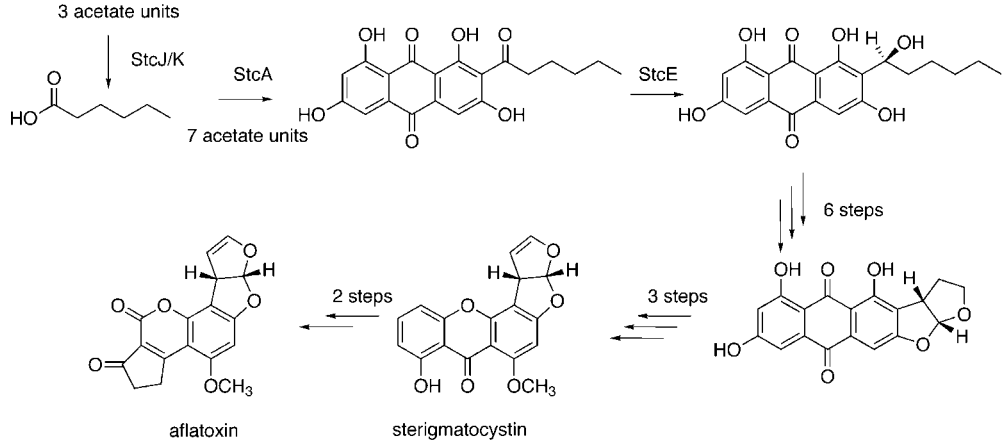

Figure 5. Biosynthetic pathway for sterigmatocystin and aflatoxin.

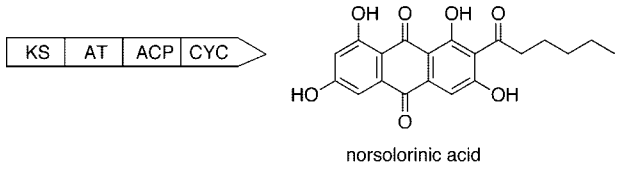

Figure 6. Domain structure of the norsolorinic acid PKS, an example of a NR-PKS.

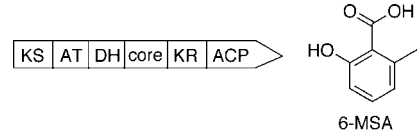

Figure 7. Domain structure of the 6-methylsalicylic acid PKS, an example of a PR-PKS.

of certain modules, and additional modifications in some cases, it is still not possible to 100% predict the precise structure of the product of a cryptic type I polyketide synthase cluster.

Bacterial Aromatic Type II Polyketide Synthase. The biosynthetic pathway for tetracenomycin C an antibiotic/antitumor produced by *S. glaucescens*. As a typical bacterial type II PKS the components of the PKS are encoded by a series of small individual genes (Figure 3b).[38] The linear decaketide is synthesized by five genes, *tcmJ*, *tcmK*, *tcmL*, and *tcmM*.[38–40] The two genes *tcmK*[α] and *tcmL*[*β*] encode for the two KS components. The gene *tcmM* encodes for the ACP. The addition of the cyclase TcmN to the other PKS components forms the fused ring TcmF2. A second cyclase TcmI then further converts TcmF2 to TcmF1. Several additional steps then catalyze the formation of the final product tetracenomycin C.

Fungal Nonreducing Polyketide Synthase (NR-PKS). The biosynthetic pathway for aflatoxin/sterigmatocystin exemplifies a typical nonreducing PKS pathway (Figure 5).[41,42] Aflatoxin is a well-known toxin, and sterigmatocystin is an intermediate in its biosynthesis. Aflatoxin is a product

(39) Shen, B.; Hutchinson, C. R. Enzymatic synthesis of a bacterial polyketide from acetyl and malonyl coenzyme A. *Science* **1993**, *262* (5139), 1535–40.

(40) Thompson, T. B.; Katayama, K.; Watanabe, K.; Hutchinson, C. R.; Rayment, I. Structural and functional analysis of tetracenomycin F2 cyclase from Streptomyces glaucescens. A type II polyketide cyclase. *J. Biol. Chem.* **2004**, *279* (36), 37956–63.

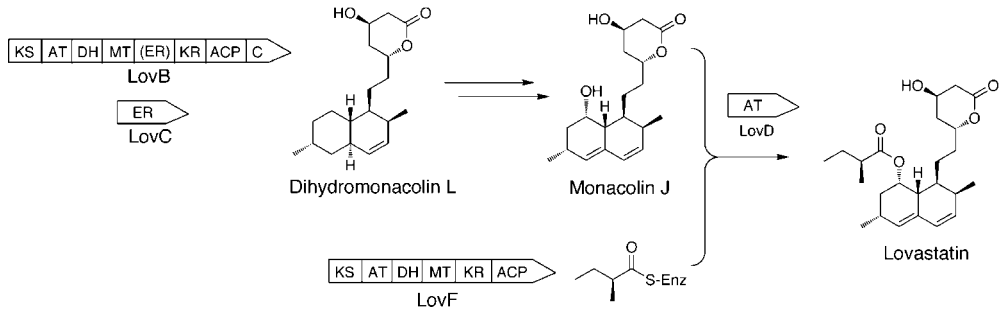

Figure 8. Domain structure of the lovastatin PKS, examples of HR-PKS.

of *A. flavus* and *A. parasiticus* while sterigmatocystin is an ultimate product of *A. nidulans.*[43] Due to the biological relevance of aflatoxin, an enormous body of literature on aflatoxin/sterigmatocystin biosynthesis has been accumulated. The initial biosynthetic step is the fatty-acid catalyzed formation of hexanoic acid by AflA/B for *A. parasiticus* and by StcJ/K for *A. nidulans,* which is then extended to norsolorinic acid by the norsolorinic acid (NA) PKS (Figure 5).[44,45] Using the UMA algorithm, the Townsend group has shown that the NA PKS in *A. parasiticus* contains the minimally required ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP) domains (Figure 6).[20] The lack of KR, DH, and ER domains is a signature of nonreducing polyketide synthases. The *A. nidulans* genome contains several nonreducing PKSs such as AN6000.3 and AN0150.3, where the polyketide product is not known, suggesting that several aromatic polyketides have yet to be identified in *A. nidulans* alone.

Fungal Partially Reducing Polyketide Synthase (PR-PKS). The *atX* gene from *Aspergillus terreus* encodes for the PKS of 6-methylsalicyclic acid.[46,47] The partially reducing polyketide synthase (PR-PKS) contains KS, AT, DH, KR, and ACP domains (Figure 7). The DNA sequence of the KS domain is distinguishable from NR-PKS. Interestingly there is a high degree of homology between fungal PR-PKS and iterative type I PKS found in bacteria. The *calO5* gene in the calicheamicin gene cluster from *Micromonospora echinospora* and the *aviM* gene in the avilamycin gene cluster from *Streptomyces viridochromogeneses* Tu57 are two such bacterial genes with high homology to fungal PR-PKSs.[48,49] Although there are several PR-PKSs found in *Aspergillus* genomes, thus far the only compound ascribed to a gene cluster is 6-methylsalicyclic acid.

Fungal Highly Reducing Polyketide Synthase (HR-PKS). Lovastatin, a product of *A. terreus*, is a potent inhibitor of HMG CoA reductase and therefore a valuable therapeutic agent for cholesterol lowering.[50,51] The biosynthetic gene cluster for lovastatin has been deciphered. It contains 17 genes spanning over 64 kb.[7,52] The polyketide is assembled by two separate HR-PKSs, LovB and LovF, working in concert. LovB, also called lovastatin nonaketide synthase (LNKS), contains seven active sites: KS, AT, DH, MT, ER, KR, and ACP domains (Figure 8).[53,54] Expression of LNKS in *A. nidulans* provides evidence that the ER domain in LNKS is not active and an accessory protein, LovC, contains the necessary ER catalytic activity. A second type I PKS,

(41) Yabe, K.; Nakajima, H. Enzyme reactions and genes in aflatoxin biosynthesis. *Appl. Microbiol. Biotechnol.* **2004**, *64* (6), 745–55.
(42) Minto, R. E.; Townsend, C. A. Enzymology and molecular biology of aflatoxin biosynthesis. *Chem. Rev.* **1997**, *97* (7), 2537–2556.
(43) Yu, J.; Chang, P. K.; Cary, J. W.; Wright, M.; Bhatnagar, D.; Cleveland, T. E.; Payne, G. A.; Linz, J. E. Comparative mapping of aflatoxin pathway gene clusters in Aspergillus parasiticus and Aspergillus flavus. *Appl. Environ. Microbiol.* **1995**, *61* (6), 2365–71.
(44) Watanabe, C. M.; Townsend, C. A. Initial characterization of a type I fatty acid synthase and polyketide synthase multienzyme complex NorS in the biosynthesis of aflatoxin B(1). *Chem. Biol.* **2002**, *9* (9), 981–8.
(45) Watanabe, C. M.; Wilson, D.; Linz, J. E.; Townsend, C. A. Demonstration of the catalytic roles and evidence for the physical association of type I fatty acid synthases and a polyketide synthase in the biosynthesis of aflatoxin B1. *Chem. Biol.* **1996**, *3* (6), 463–9.
(46) Fujii, I.; Ono, Y.; Tada, H.; Gomi, K.; Ebizuka, Y.; Sankawa, U. Cloning of the polyketide synthase gene atX from Aspergillus terreus and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression. *Mol. Gen. Genet.* **1996**, *253* (1–2), 1–10.
(47) Moriguchi, T.; Ebizuka, Y.; Fujii, I. Analysis of subunit interactions in the iterative type I polyketide synthase ATX from Aspergillus terreus. *ChemBioChem* **2006**, *7* (12), 1869–74.
(48) Ahlert, J.; Shepard, E.; Lomovskaya, N.; Zazopoulos, E.; Staffa, A.; Bachmann, B. O.; Huang, K.; Fonstein, L.; Czisny, A.; Whitwam, R. E.; Farnet, C. M.; Thorson, J. S. The calicheamicin gene cluster and its iterative type I enediyne PKS. *Science* **2002**, *297* (5584), 1173–6.
(49) Gaisser, S.; Trefzer, A.; Stockert, S.; Kirschning, A.; Bechthold, A. Cloning of an avilamycin biosynthetic gene cluster from Streptomyces viridochromogenes Tu57. *J. Bacteriol.* **1997**, *179* (20), 6271–8.
(50) Sutherland, A.; Auclair, K.; Vederas, J. C. Recent advances in the biosynthetic studies of lovastatin. *Curr. Opin. Drug Discovery Dev.* **2001**, *4* (2), 229–36.
(51) Chakravarti, R.; Sahai, V. Compactin-a review. *Appl. Microbiol. Biotechnol.* **2004**, *64* (5), 618–24.
(52) Abe, Y.; Suzuki, T.; Ono, C.; Iwamoto, K.; Hosobuchi, M.; Yoshikawa, H. Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in Penicillium citrinum. *Mol. Genet. Genomics* **2002**, *267* (5), 636–46.
(53) Ma, S. M.; Tang, Y. Biochemical characterization of the minimal polyketide synthase domains in the lovastatin nonaketide synthase LovB. *FEBS J.* **2007**, *274* (11), 2854–64.
(54) Xie, X.; Watanabe, K.; Wojcicki, W. A.; Wang, C. C.; Tang, Y. Biosynthesis of lovastatin analogs with a broadly specific acyltransferase. *Chem. Biol.* **2006**, *13* (11), 1161–9.

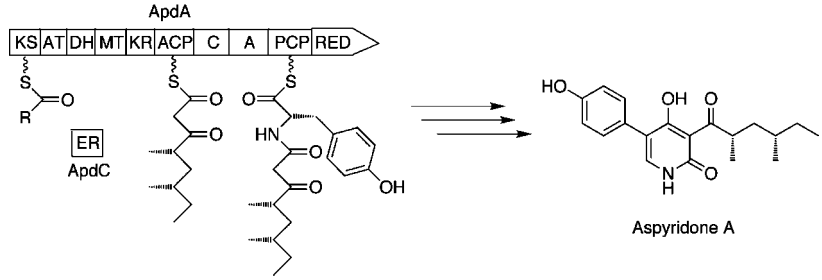

Figure 9. Domain structure of the aspyridone NRPS/PKS.

LovF, provides lovastatin's diketide moiety, 2-methylbutyrate. Although no compound of structure similar to lovastatin has been reported in *A. nidulans*, the genome contains two adjacent NR-PKS genes, AN3610.3 and AN3612.3, suggesting that a compound of a similar biosynthetic pathway to lovastatin has yet to be discovered.

Hybrid PKS/NRPS Gene. Aspyridone is an example of a secondary metabolite in *A. nidulans* biosynthesized by a hybrid PKS/NRPS gene (Figure 9).[16] By replacing the promoter of the C6 transcription factor gene found in the aspyridone cluster with an inducible promoter, the Hertweck group induced the entire pathway, allowing the chemical elucidation of the product of this specific gene cluster. The hybrid *apdA* gene contains domains commonly found in HR-PKSs such as KS, AT, DH, MT, and KR domains. The gene also contains domains found in NRPS genes such as a condensation domain (C), adenylation domain (A), peptidyl carrier protein (PCP), and reductase domain (RED).

The Road Ahead. The completed *Aspergillus* and *Streptomyces* genomes clearly demonstrate the richness of polyketides that could potentially be produced by these organisms. The task ahead will be to chemically match the numerous unexplored PKS clusters in these genomes with their chemical products. The availability of shared resources such as the Broad Institute Fungal Genome Initiative (http://www.broad.mit.edu/annotation), Comprehensive Microbial Resource (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi) and the Database for NRPS and PKS (http://www.nii.res.in/search/html) have made the task easier.[17] Investment in the development of better tools for functional analysis of secondary metabolite genes and in the understanding of the regulation of secondary metabolism will also be important in order to translate the sequence information into potential therapeutics.

MP700139T