Minireview

Principles of microbial alchemy: insights from the Streptomyces coelicolor genome sequence
Charles J Thompson, Doris Fink and Liem D Nguyen

Address: Biozentrum, University of Basel, Division of Molecular Microbiology, 70 Klingelbergstrasse, 4056 Basel, Switzerland.

Correspondence: Charles J Thompson. E-mail: Charles-J.Thompson@unibas.ch

Published: 26 June 2002

Genome Biology 2002, 3(7):reviews1020.1–1020.4

The world's most creative producers of natural pharmaceutical compounds are soil-dwelling bacteria classified as Streptomyces. The availability of the recently completed Streptomyces coelicolor genome sequence provides a link between the folklore of antibiotics and other bioactive compounds to underlying biochemical, molecular genetic and evolutionary principles.

Streptomyces, which belong to the bacterial order Actinomycetales, are similar in appearance to fungi and undergo a program of colonial morphogenesis that is coordinated with the excretion of bioactive compounds that are often colored (Figure 1) [1]. The biosynthesis of these compounds, many of which are antibiotics, occurs typically in response to a breakdown of balanced metabolism that interrupts or arrests growth [2,3]. On solid media this triggers formation of aerial filaments that septate and mature into a chain of spores.

Streptomyces species are considered exceptionally well endowed for 'chemical warfare', presumably allowing them to eliminate bacterial and fungal competitors in soil ecosystems. This was first realized over fifty years ago, and since then the majority of known antibiotics have been isolated from Streptomyces. Although thousands of antibiotics have been described, these are thought to represent only a small fraction of the repertoire of bioactive compounds produced by Streptomyces [4,5]. In addition, empirical screening using various assays has revealed that Streptomyces culture supernatants contain other pharmaceutically active compounds, such as anti-viral and anti-cancer compounds, modulators of immune responses, and various enzyme inhibitors, as well as herbicides, insecticides, and anti-parasitic compounds [4,6,7]. Furthermore, chemical screening methods have uncovered the remarkable structural diversity of these compounds. Genome sequences of Streptomyces species are now providing an overview of the genetic elements responsible for this metabolic diversity, both within one bacterium and the entire genus. Analysis of the completed genome sequence of Streptomyces coelicolor [8] and its comparison with other species will give further insights into availability and nature of potentially useful bacterial metabolites.

Dozens of different Streptomyces species (Figure 1a) [1] can be selectively isolated from virtually any rich soil sample. Tens of thousands of Streptomyces strains are screened annually by pharmaceutical companies as potential sources of novel chemical compounds. High-throughput chromatographic analyses typically reveal that each strain has its own unique profile of secondary metabolites (metabolites that are secondary to the growth and maintenance of the cell), mostly compounds that are not found in synthetic chemical handbooks or combinatorial chemical libraries. The next step of this commercial process, namely increasing yields to make large-scale production of the compound feasible, has relied on screening for better culture conditions and over-producing mutants. Mutations that result in higher yields of certain metabolites can now be combined into one organism by genomic shuffling [9], thus reducing the time needed for developing industrial production strains. Although such empirical approaches have been very fruitful for individual secondary metabolites, the improvements are strain-specific, so the acquisition of useful knowledge is not cumulative.
chain length by condensation with malonyl CoA and fully reduce the carboxyl residues at each step of extension. These have evolved into two different types of polyketide synthases: type I polyketide synthases, which catalyze a reaction in which the growing acyl chain is sequentially transferred to a series of modular reaction centers of the enzyme, typically generating macrocyclic lactone rings characteristic of so-called macrolide antibiotics; and type II polyketide synthases, which typically condense malonyl CoA iteratively, using the same active condensation site to generate aromatic compounds. Polypeptide-based antibiotics are often synthesized by non-ribosomal peptide synthases, which can condense more than 15 amino acids or unusual analogs into linear or cyclic structures.

Rapid progress in understanding these complex multifunctional enzymes was achieved through a combination of genetic and chemical approaches [10]. It is now possible to predict and engineer the structure of both polypeptides [11-14] and polyketides [11-13,15]. The modules of non-ribosomal peptide synthase and type I polyketide synthase enzymes are arranged in an order that is colinear with the component residues in the mature polymer. Furthermore, these enzyme modules can be genetically engineered to reprogram the order and modifications of residues within the polymers. Using this technology, dozens of derivatives of the polyketide erythromycin have been synthesized by reprogramming the erythromycin polyketide synthase [16]. Domain swapping between peptide synthases has demonstrated the feasibility of engineering peptide antibiotics [14]. Engineering additional domain modules will increase the repertoire available for introducing random changes or designing new antibiotics. It remains to be seen, however, whether these approaches will rival the diversity and specificity of compounds that have evolved in biological systems over hundreds of millions of years. For example, natural products are already specifically preselected to avoid negative interactions with other protein domains that might make them toxic to bacteria that produce them and likewise to humans. So how can genome sequencing help provide better access to natural products?

Thus, new *Streptomyces* secondary metabolites are still accessible only through expensive and time-consuming screening techniques.

The diversity of secondary metabolite structures defies simple chemical classification, but many of the most useful and best understood compounds are synthesized by two families of multifunctional enzymes that can assemble unusual carbon and peptide chains. Polyketide and peptide synthases catalyze condensation reactions that can be used to assemble acetyl CoA or amino-acid derivatives into polyketide and polypeptide structures, respectively. Polyketide synthases are structurally and functionally homologous to fatty acid synthases. ‘Classical’ fatty acid synthases initiate the assembly of fatty-acid chains with acetyl CoA, increase
its secondary metabolites, namely prodigiosins (red pigmented antibiotics), two type II polyketides (the antibiotic actinorhodin as well as a gray spore pigment encoded by the whiE locus), and the non-ribosomal peptide calcium-dependent antibiotic (CDA). The S. coelicolor genome sequence now predicts 18 additional secondary metabolites, including polyketides synthesized by polyketide synthases of type I and type II and by a third type found primarily in plants, peptides assembled by non-ribosomal peptide synthases, siderophores, hopanoids (sterol-like bacterial lipids), butyrolactones (bacterial pheromones), terpene, geosmin (the predominant natural odor of soil), and others [8].

The genes of a secondary metabolic pathway are typically coregulated, clustered within the bacterial chromosome, and functional in heterologous species. Early mapping studies [18] revealed synteny (conserved order and orientation) of genes within homologous gene clusters encoding compounds of antibiotic biosynthesis pathways in different species. This probably reflects common ancestry and common adaptive functions for these gene clusters. Similarly, early linkage analyses of Streptomyces chromosomes suggested synteny of auxotrophic markers in different species [19]. Antibiotic biosynthetic clusters were not found in the same places in different Streptomyces chromosomes, however, suggesting that these clusters were incorporated later. Sequence analysis of Streptomyces chromosomes supports these initial findings and provides further evidence that secondary metabolic islands (SMILES), such as the antibiotic biosynthetic clusters, were acquired by horizontal transmission. Like pathogenicity islands - genomic segments in bacterial pathogens that encode virulence factors and may have evolved through horizontal transfer - genes for illegitimate recombination, such as those encoding integrases and transposases, are occasionally identified within or adjacent to antibiotic gene clusters in S. avermitilis [17] and S. coelicolor [8]. For example, genes found associated with the secondary metabolite clusters whiE and cda in S. coelicolor are such neighbors. Because the GC content of these clusters is indistinguishable from that of the adjacent chromosome sequences and their regulatory systems are incorporated with those of other Streptomyces genes, integration of SMILES in the genome may have occurred long ago. The retention of apparently intact clusters, which contain no obvious pseudogenes, suggests that their products are at present important but are synthesized only when primary metabolic precursor pools can no longer support efficient growth.

The existence of so many different secondary metabolites probably indicates specific adaptions to different habitats. Mycobacteria, which are also members of the Actinomycetales, are also primarily saprophytic soil bacteria. But the genus also includes Mycobacterium tuberculosis and the leprosy-causing Mycobacterium leprae, two of the most dangerous pathogens known to man. The M. tuberculosis and S. coelicolor chromosomes show synteny within a region that is defined as the central core sequence in S. coelicolor (about 5 Megabases) [8]. Interestingly, the M. tuberculosis chromosome contains 18 polyketide synthase gene clusters [20]. The fact that six of these clusters were also found in the severely reduced M. leprae chromosome suggests that they serve indispensable functions in the bacteria [21]. Although many of the clusters may provide cell-wall components, in Mycobacterium ulcerans (the causative agent of a tropical skin disease), one produces mycolactone, a toxin involved in pathogenicity [22]. Mycobactin, a siderophore synthesized by a combination of polyketide synthases and non-ribosomal peptide synthases, plays an essential role in mycobacterial virulence [23]. The fact that Streptomyces produce potent immunomodulators, such as tacrolimus [24] (also called FK506 [25]) and rapamycin [26], suggests that secondary metabolites may provide other virulence functions, such as the immunosuppression that accompanies mycobacterial infections. Such compounds may have enormous therapeutic value because lethal pathogenicity effectors could potentially be turned into life-saving pharmaceutical compounds. Similarly, other actinomycetes that have not been as extensively exploited because of isolation and cultivation difficulties may also produce large numbers of unusual SMILE-encoded compounds with undiscovered and unanticipated biological activities [1,4,7] that may reflect each species’ interactions with soil metazoans.

Because the S. coelicolor genome sequence is now available, large-scale methods of gene expression and site-directed mutagenesis can be exploited to identify even more secondary metabolites and to better understand the regulatory mechanisms that coordinate pathways of primary and secondary metabolism. In an early proteomic analysis using two-dimensional gel electrophoresis of Streptomyces hygroscopicus cultures, that had been pulse-labeled at various time points during growth, we used statistical analyses of global gene-expression patterns to identify secondary metabolic enzymes [27]. Analyses of the patterns displayed by 400 spots on the two-dimensional gels showed that the proteins encoded within the SMILE for biosynthesis of the antibiotic and herbicide bialaphos fell into discrete kinetic groups. Similar, but more extensive two-dimensional gel analyses have now been initiated in S. coelicolor by us and others, and several hundred protein spots have been identified [28,29]. In addition, Cohen and colleagues [30] were able to correlate transcriptional expression patterns with chromosome position by using powerful gene chip studies of S. coelicolor cultures, coupled with a knowledge-based algorithm. This approach allowed them to identify contiguous genes in four SMILES corresponding to known polyketide synthase or non-ribosomal peptide synthase gene clusters. The same principle might be used to predict SMILES producing novel compounds that are not encoded by polyketide or peptide synthases. In fact, kinetic analysis also revealed other blocks of coordinately regulated genes. These genes may be involved in the synthesis of previously unrecognized
secondary metabolites in *S. coelicolor*. Although there are many secondary metabolites that are not dependent on polyketide synthases or non-ribosomal peptide synthases, few corresponding enzymes have been identified.

The genome of *S. coelicolor* provides an overview of the genetic framework that defines one streptomyecete and its SMILLEs. In future, the availability of genome sequences of other Streptomyces species will allow us to identify common features that define their shared genetic heritage and links with acquired secondary metabolic genes in different species. Statistical analyses of global gene-expression patterns, in combination with ‘metabolomic’ studies of the pools of metabolic intermediates and products, will facilitate both understanding and manipulation of the primary and secondary metabolic pathways of these bacteria thus ensuring their continued importance to man.

Acknowledgements

Many thanks to Tobias Kieser, Celia Bruton and Jennifer Tenor for providing the photograph of the Streptomyces colonies.

References

1. Miyadoh S, Hamada M, Hotta K, Kudo T, Seino A, Vobis G, Yokota A: *Atlas of Actinomycetes*: Tokyo: Asakura Publishing Co. Ltd.; 1997.
2. Süsstrunk U, Pidoux J, Taubert S, Ullmann A, Thompson CJ: A short review (1988-1992). *Res Microbiol* 1993, 144:633-642.
3. Vezina C, Kudelski A, Sehgal SN: Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomyecete and isolation of the active principle. *J Antibiot* (Tokyo) 1975, 28:721-726.
4. Vezina C, Kudelski A, Sehgal SN: Rapamycin (AY-22,989), a new antifungal antibiotic. II. Production and isolation of the active principle. *J Antibiot* (Tokyo) 1975, 28:727-733.
5. Friend EJ, Hopwood DA: *Atlas of Actinomycetes* Vol 3 No 7.

4. Bérdy J: *Atlas of Actinomycetes*: Tokyo: Asakura Publishing Co. Ltd.; 1997.

5. Watve MG, Tickoo R, Jog MM, Bhole BD: Pleiotropic effects of cAMP on germination, antibiotic biosynthesis, and morphological development in Streptomyces coelicolor. *Molec Microbiol* 1998, 30:33-46.

6. Chater KF, Bibb MJ: Regulation of bacterial antibiotic production. In *Products of Secondary Metabolism*. Edited by Pühler A. Weinheim: VCH; 1997: 57-105.

7. Omar S, Ikeda H, Ishikawa J, Hanamato A, Takahashi C, Shinose M, Takahashi Y, Horikawa H, Nakazawa H, Osono T, et al.: *Genome sequence of an industrial microorganism Streptomyces avermitilis*: deducing the ability of producing secondodifiers metabolites. *Proc Natl Acad Sci USA* 2001, 98:12215-12220.

8. Staunton J, Wilkinson BB: Combinatorial biosynthesis of polyketides and nonribosomal peptides. *Curr Opin Chem Biol* 2001, 5:159-164.

9. Stachelhaus T, Schneider A, Marahiel MA: Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* 1995, 269:69-72.

10. Cane DE, Walsh CT, Khosla C: Bialaphos biosynthetic genes of Streptomyces viridochromogenes: cloning, heteroprosopstic expression, and comparison with the genes of Streptomyces hygroscopicus. *J Gen Microbiol* 1991, 137:351-359.

11. Chung WS, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eglmeier K, Gas S, Barry CE, 3rd, et al.: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 1999, 393:337-344.

12. Cole ST, Eglmeier K, Parkhill J, James KD, Thompson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, et al.: Massive gene decay in the leprosy bacillus. *Nature* 2001, 409:1007-1011.

13. George KM, Chatterjee D, Gunawardana G, Welty D, Hayman J, Lee R, Small PL: Mycolactone: a polyketide toxin from *Mycobacterium ulcerans* required for virulence. *Science* 1999, 283:854-857.

14. Quadri LE, Sello J, Keating TA, Weinreb PH, Walsh CT: Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem Biol* 1998, 5:631-645.