molecular systems biology

NEWS AND VIEWS

Combinatorial polyketide biosynthesis at higher stage

Holger Jenke-Kodama and Elke Dittmann

Department of Molecular Ecology, Institute of Biology, Humboldt University Berlin, Berlin, Germany

Modular polyketide synthases (PKS) of bacteria provide an amazing molecular assembly line for the biosynthesis of complex polyketides. This biosynthesis system has, since its discovery, attracted the attention of scientists and pharmaceutical companies owing to its combinatorial potential. It has provoked the idea of constructing 'unnatural' product libraries containing a myriad of compounds with all possible lengths and combinations of carbon units. The recent study by Menzella *et al* (2005) represents an important milestone on the way to construct such libraries, and highlights both the potential and the limits of biocombinatorial strategies for a complete reorganisation of natural enzyme units. The group at Kosan Biosciences Inc. has adopted and improved existing engineering techniques and describes for the first time a high-throughput methodology for the generation of synthetic triketides.

Natural polyketides derived from giant PKS complexes were mostly isolated from microbes in the soil, but were increasingly discovered also in the oceans and have served as lead products for some of the most important pharmaceuticals currently on the market (Staunton and Weissman, 2001). They include antibiotics (e.g. erythromycin A), anticancer drugs (e.g. epothilone), antifungals (e.g. nystatin) or immunosuppressants (e.g. rapamycin). Multienzymes determining the order and configuration of a given polyketide are composed of repetitive modules consisting of sets of domains carrying the active sites for the successive activation, modification and elongation of single carbon building blocks. The chemical steps of chain extension and correspondingly the enzymatic activities of polyketide synthesis resemble those of fatty acid biosynthesis (Hopwood and Sherman, 1990). The extraordinary high diversity of polyketide products, however, is achieved by an optional use of domains for the modification of ketogroups and by the use of different substrates for chain initiation and extension. It has been calculated that a PKS system comprising six modules is theoretically able to produce over 100 000 possible structures (Gonzalez-Lergier *et al*, 2005).

The modules used in the study by Menzella *et al* come in two types—'loading' modules and 'extender' modules. Both are provided on individual plasmids enabling a heterologous expression of the enzyme cassettes in the fast growing *Escherichia coli* host. Earlier studies (e.g. Gokhale *et al*, 1999) have revealed the high importance of intermodular linkers for the successful cooperation of PKS units and have suggested an active participation of the linkers in polyketide chain transfer. A focus of the study was therefore the computer-based design of optimal and universal linkers, provided as repeated sets of flanking restriction sites. This brought about the development of automated cloning strategies for the facile interchange of polyketide gene cassettes. This strategy was intrusive, but before this study laborious and time-consuming gene cloning steps were needed for single combinatorial experiments. The innovation of the strategy comes from the combination of universal linker technology with a rapid and efficient long gene synthesis, which consists of two steps (Kodumal *et al*, 2004). First, sequences of about 500 bp in length are constructed by polymerase cycling assembly (Stemmer *et al*, 1995) and then these DNA fragments are connected using the so-called ligation by selection technology (Kodumal and Santi, 2004). This allowed for an integration of gene cassettes from any desirable host with a concurrent optimisation of the codon usage and an easy introduction of restriction sites.

The authors have validated the efficiency of their technology for 154 bimodular combinations assembled from 14 synthesised modules of seven different polyketide biosynthesis pathways of streptomycetes and myxobacteria. About half of the module pairs yielded detectable amounts of the triketide lactone product. Yet, the objective of this systematic approach was not primarily to study the productivity itself, but the identification of donor and acceptor enzyme components displaying a very high cooperative flexibility for future combinatorial experiments. Indeed, individual 'loading' and 'extender' modules were found to differ in their promiscuity. For a wider use of the strategy, however, we still need to understand the molecular basis behind these differences. The study by Menzella and colleagues is thus only the starting point towards the design of truly universal synthetic PKS cassettes.

Aside from its enormous biotechnological relevance, this work also directs attention to some fundamental questions about the evolution of multimodular synthase systems. In constructing new combinations of modules or new domain arrangements within modules, researchers obviously try to imitate nature. Bacteria face the same problems: How to recombine the elements of this modular machinery to produce new compounds without demolishing the system's operability? Evolutionary studies in this field have focused on the distribution of PKS among bacteria, duplication patterns and the involvement of horizontal gene transfer events (Ginolhac *et al*, 2005; Jenke-Kodama *et al*, 2005). What is still missing, however, is an inspection of the mechanism and evolutionary forces, which act upon the formation of this type of gene

clusters. Cursory examination suggests that in addition to duplications, many recombination events are necessary to build up multienzyme systems showing tremendously diverse module compositions. It would be of interest to see how organisms handle the potential of modularity. Do they use gene cassettes? Is there an exchange of whole modules or certain domain groups? Where are the break points that guide the remodelling of a cluster? The growing number of available sequences of both PKS gene clusters and whole genomes of polyketide-producing bacteria provide a promising basis for such studies. The answers to these questions would be important by themselves, but it may be possible to work out principles that could help optimise the biotechnological approach. It could be helpful to have a look at the manual rather than just taking the parts of the construction kit from nature.

References

Ginolhac A, Jarrin C, Robe P, Perriere G, Vogel TM, Simonet P, Nalin R (2005) Type I polyketide synthases may have evolved through horizontal gene transfer. *J Mol Evol* **60:** 716–725

Gokhale RS, Tsuji SY, Cane DE, Khosla C (1999) Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284:** 482–485

Gonzalez-Lergier J, Broadbelt LJ, Hatzimanikatis V (2005) Theoretical considerations and computational analysis of the complexity in polyketide synthesis pathways. *J Am Chem Soc* **127:** 9930–9938

Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu Rev Genet* **24:** 37–66

Jenke-Kodama H, Sandmann A, Müller R, Dittmann E (2005) Evolutionary implications of bacterial polyketide synthases. *Mol Biol Evol* **22:** 2027–2039

Kodumal SJ, Patel KG, Reid R, Menzella HG, Welch M, Santi DV (2004) Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc Natl Acad Sci USA* **101:** 15573–15578

Kodumal SJ, Santi DV (2004) DNA ligation by selection. *Biotechniques* **37:** 34–40

Menzella HG, Reid R, Carney JR, Chandran SS, Reisinger SJ, Patel KG, Hopwood DA, Santi DV (2005) Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes. *Nat Biotechnol* **23:** 1171–1176

Staunton J, Weissman KJ (2001) Polyketide biosynthesis: a millennium review. *Nat Prod Rep* **18:** 380–416

Stemmer WP, Crameri A, Ha KD, Brennan TM, Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164:** 49–53