Chemistry & Biology
Previews

Tailor-Made Peptide Synthetases

Hajo Kries[1] and Donald Hilvert[1,*]
[1]Laboratory of Organic Chemistry, ETH Zürich, CH-8093 Zurich, Switzerland
*Correspondence: donald.hilvert@org.chem.ethz.ch
DOI 10.1016/j.chembiol.2011.10.006

Harnessing the modular architecture of non-ribosomal peptide synthetases for combinatorial biosynthesis is a longstanding goal in chemical biology. Several recent reports illustrate how computational design and directed evolution can be used to tailor the specificity of these assembly-line enzymes.

The serendipitous discovery of penicillin 80 years ago revolutionized the treatment of dangerous bacterial infections. It also showcased a structurally diverse class of natural products (i.e., non-ribosomal peptides) that exhibit a broad spectrum of valuable biological and pharmacological activities. In addition to many clinically important antibiotics, these secondary metabolites include potent cytostatic agents, immunosuppressants, siderophores, and toxins.

While penicillin is still in use today, many bacteria are now resistant to the original drug. Chemists have responded to this challenge by synthesizing novel penicillin derivatives that prolong life and good health, but the need for new and improved antibiotics (and other agents) is never-ending. Biological approaches may soon be able to help; elucidation of the chemical logic underlying the biological assembly of non-ribosomal peptides over the last two decades has made biosynthesis of non-natural derivatives in microorganisms an enticing alternative to total synthesis in the laboratory. Combinatorial manipulation of these pathways has significant potential to expand the molecular diversity of structurally complex natural products, contributing both to the discovery and optimization of pharmacological leads (Cane et al., 1998).

Non-ribosomal peptides are produced biologically by gigantic mega-enzymes called non-ribosomal peptide synthetases (NRPSs). These enzymes consist of multiple protein modules, strung together like beads on a string, that function much like an assembly line (Sieber and Marahiel, 2005). Each module is responsible for incorporating one amino acid into the growing peptide chain, with the length and composition of the final natural product determined by the nature, number, and order of the individual modules. The substrate specificity of each module is dictated by an adenylation domain, or A domain, which recognizes and activates the appropriate amino acid and tethers it covalently to the assembly line.

The modular architecture of NRPSs suggests that it should be possible to create "unnatural" products simply by replacing or modifying the A domains in an existing NRPS (Figure 1). Protein engineers can take advantage of the large set of natural A domains that activate hundreds of different amino acids and swap individual domains or even entire modules in NRPSs to make new peptides. Because intermodule communication is still poorly understood, however, the resulting constructs often suffer from poor overall activity. If an effective screen or selection is available, laboratory evolution can be applied to increase the efficiency of these hybrid systems, as demonstrated for novel NRPSs that produce andrimid derivatives, potent non-ribosomal peptide/polyketide antibiotics (Fischbach et al., 2007).

Reengineering existing A domains represents an alternative strategy for altering NRPS specificity that, at least in principle, enables exploration of substrate preferences not exhibited by natural A domains. Using a structurally characterized PheA domain as a guide, Stachelhaus et al. (1999) identified a set of ten residues in the amino acid binding pocket as potential determinants of substrate specificity. The sequence variability at these positions in biochemically characterized adenylation domains suggested a set of general rules, the so-called "Stachelhaus code," for predicting substrate preferences directly from primary sequence. These predictions have also been used for rational redesign of A domain selectivity, but changes achievable by single point mutations are typically conservative (Asp→Asn, Glu→Gln, Phe→Leu) (Stachelhaus et al., 1999; Eppelmann et al., 2002). Not surprisingly, because of phylogenetic bias and relatively small data sets, sequence-based approaches cannot fully capture the subtleties underlying amino acid recognition.

The substrate preferences of many proteins have been successfully altered by more far-reaching combinatorial searches of sequence space. Structure-based computational methods are particularly attractive in this regard. Algorithms have been developed that enable rapid evaluation of thousands of different sequences and hundreds of millions of conformations. Applying this approach to the first adenylation domain in gramicidin S, Chen et al. (2009) achieved a 2000-fold switch in substrate specificity from L-phenylalanine to L-leucine. Notable in this context was the identification of a novel set of mutations not encountered in natural A domains, including bolstering mutations that do not directly contact the substrate. Although currently limited by the simplifying assumptions of the computational models, this approach to substrate redesign has enormous potential and will undoubtedly increase in importance as the methods become faster and more accurate.

In the interim, as Villiers and Hollfelder (2011) report in this issue of *Chemistry & Biology*, experimental evolutionary searches can also be quite effective. For example, using a medium-throughput screen and several rounds of iterative saturation mutagenesis followed by recombination of beneficial mutations, they successfully repurposed the TycA module of tyrocidine synthetase. Normally, the A domain of TycA activates L-phenylalanine five to six orders of magnitude more efficiently than small, hydrophilic substrates

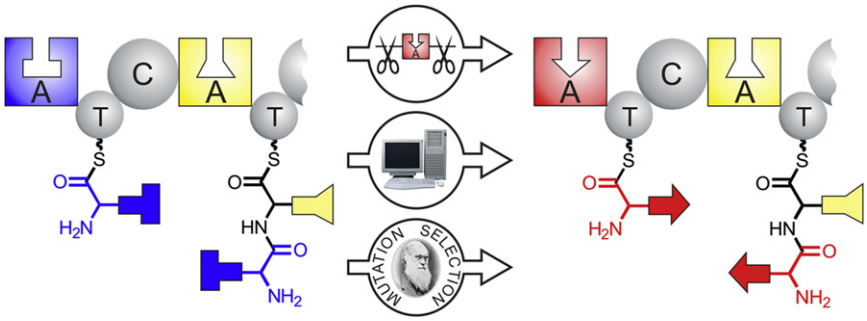

Figure 1. Tailoring the specificity of NRPSs. Novel peptides can be made by domain splicing, computational design, or directed evolution of NRPSs. A, adenylation domain; T, thiolation domain; C, condensation domain.

like L-alanine and L-threonine. Over the course of directed evolution, however, the weak activity toward the latter could be significantly enhanced. Relatively few changes to the specificity determining residues sterically restricted access to the active site, leading to a 170-fold increase in the k_{cat}/K_m value for L-Ala and a concomitant 10^3-fold decrease in the k_{cat}/K_m value for L-Phe. This 10^5-fold change in specificity is impressive and substantially larger than the changes typically achieved by implementing the predictions of the "Stachelhaus code." Nevertheless, the modified TycA variant is unlike highly evolved A domains in Nature, since it activates large and small substrates comparably well. As such, it is more like the promiscuous intermediates postulated to be the evolutionary precursors of modern enzymes. It will be interesting to see whether this functional plasticity can be effectively channeled through additional laboratory evolution to achieve more specialized amino acid recognition.

Recent work by Evans et al. (2011) indicates that this basic strategy is not restricted to isolated A domains. They generated libraries of the AdmK module of the andrimid biosynthetic gene cluster, targeting three of the specificity determining active site residues of the A domain for saturation mutagenesis. These were expressed in the native host, *Pantoea agglomerans*, and the >14,000-member library was screened directly for the production of andrimid-like metabolites using a highly sensitive mass spectrometric technique. Several clones were identified that produced four bioactive andrimid derivatives, three of which were new. Interestingly, the AdmK variants all exhibited broadened substrate specificity relative to the wild-type NRPS; two even had promiscuous activity resembling that of the TycA variant just discussed, inserting mixtures of L-alanine and L-phenylalanine in place of the normal substrate L-valine. Although only a single round of mutagenesis and screening was performed, this integrated in vivo approach will facilitate further evolution and should be readily extendable to larger libraries and other biosynthetic enzymes.

The proof-of-concept experiments described in this short perspective highlight important progress toward the redesign of NRPS specificity. While reliable, on-demand biosynthesis of complex natural product analogs remains a distant goal, it can be expected that in silico design and experimental evolution, either individually or, more powerfully, in combination, will significantly expand the toolkit for constructing non-ribosomal peptides. Indeed, advances in this area will likely fuel design and manufacture of novel therapeutic agents of many kinds in the years to come.

REFERENCES

Cane, D.E., Walsh, C.T., and Khosla, C. (1998). Science *282*, 63–68.

Chen, C.Y., Georgiev, I., Anderson, A.C., and Donald, B.R. (2009). Proc. Natl. Acad. Sci. USA *106*, 3764–3769.

Eppelmann, K., Stachelhaus, T., and Marahiel, M.A. (2002). Biochemistry *41*, 9718–9726.

Evans, B.S., Chen, Y.Q., Metcalf, W.W., Zhao, H.M., and Kelleher, N.L. (2011). Chem. Biol. *18*, 601–607.

Fischbach, M.A., Lai, J.R., Roche, E.D., Walsh, C.T., and Liu, D.R. (2007). Proc. Natl. Acad. Sci. USA *104*, 11951–11956.

Sieber, S.A., and Marahiel, M.A. (2005). Chem. Rev. *105*, 715–738.

Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999). Chem. Biol. *6*, 493–505.

Villiers, B.R.M., and Hollfelder, F. (2011). Chem. Biol. *18*, this issue, 1290–1299.