One pathway, many products

Michael A Fischbach & Jon Clardy

Biosynthetic pathways for secondary metabolites usually make many products, not just one. In this Commentary, we consider why molecular promiscuity might be an evolutionarily advantageous feature of these pathways.

The biosynthetic pathways that synthesize small molecules use one of two distinct strategies. Primary metabolic pathways, which are usually active and synthesize the small molecules used by most organisms most of the time, make single products. The tryptophan pathway makes tryptophan, the cholesterol pathway makes cholesterol, and neither pathway makes side products. How has such perfection come about? Presumably, evolution—successive rounds of mutation, selection and amplification—has honed the enzymes in each pathway to maximize their yield and avoid the buildup of unnecessary and potentially toxic side products.

Secondary metabolic pathways, which are turned on in response to specific cues and make natural products, typically make a variety of products. Some pathways make only one or two products, and some make more than 100. In the language of laboratory synthesis, primary metabolic pathways are target-oriented, whereas secondary pathways are diversity-oriented. Is biosynthetic molecular promiscuity a bug or a feature? In this Commentary, we consider why natural product biosynthetic pathways may have evolved to favor molecular diversity. In keeping with the thematic focus of this issue, we will focus on terpenoid natural products, but similar arguments can be made for the other categories of secondary metabolites.

Views on how natural product pathways evolved reflect the dichotomy in biosynthetic strategies described above. The better-known view, which was proposed by D. Williams in 1989, is that natural products confer a survival benefit through their ability to bind to cellular targets in competing organisms[1,2]. Evolutionary pressure on natural product biosynthetic pathways yields widespread "natural product/receptor interactions of sophistication comparable to enzyme/substrate interactions"[2]. Today, one is tempted to describe this model as the 'target-based' model of natural product evolution, because a molecule is carefully sculpted by evolution to bind tightly and specifically to its target. The model has compelling supporting evidence from numerous structural studies that show the exquisite fit between a natural product and its cellular target. A personal favorite is the complex of the immunosuppressive natural product rapamycin with two proteins: FKBP12 and the FRB domain of the mammalian target of rapamycin (mTOR)[3] (**Fig. 1**). Such crystal structures invariably reveal the exquisite fit of natural products and their protein receptors. These structural analyses, along with the potency and specificity of natural product–receptor pairs, reinforce a view of natural products as pinnacles of evolutionary perfection—evolution's greatest small-molecule hits.

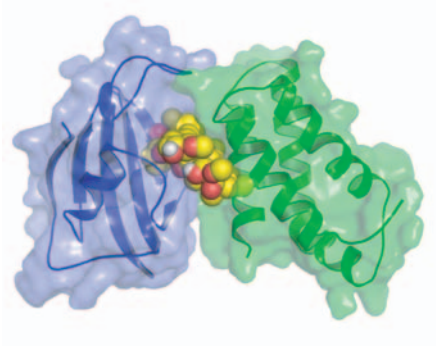

Figure 1 The immunosuppressive compound rapamycin bound to FKBP12 and the FRB domain of mTOR. FKBP12 is represented by the blue ribbon diagram on the left, the FRB domain of mTOR is represented by the green ribbon diagram on the right, and rapamycin is shown as individual spheres. Rapamycin simultaneously binds these two proteins by having different parts of the molecule fill hydrophobic pockets in each protein while the more polar region of rapamycin is concentrated on a small solvent-exposed surface between the two proteins.

An alternative view, formulated by Firn and Jones in 1991, is based on the observation that potent biological activity is a rare property for any molecule (including natural products) to have, and that an organism needs the ability to make multiple molecules in order to hit upon the rare potent ones[4–6]. Thus "evolution would favor organisms that could generate and retain chemical diversity at low cost. Organisms that make and 'screen' a large number of chemicals will have an increased likelihood of enhanced fitness simply because the greater the chemical diversity, the greater the chances of producing the rare chemical with a useful, potent biological activity"[5]. Though Firn and Jones referred to their model as the 'screening' model, it might be more appropriate to describe it as the 'diversity-based' model to emphasize the nature of the biosynthetic pathways rather than the way in which their products are used. Several features of natural product biosynthesis seem to support the diversity-based model, especially the large number of natural products with no known activity (or at least no known potent activity), the tendency of natural product pathways to produce a suite of molecules, and the widespread use of branched and matrix biosynthetic pathways to share metabolic and genetic costs.

To examine the phenomenon of biosynthetic promiscuity in light of the target-based and diversity-based models, we consider a pathway that produces at least 136 products: the gibberellin-family diterpenes (http://www.plant-hormones.info/). The gibberellins are a

Michael A. Fischbach and Jon Clardy are in the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA.
e-mail: jon_clardy@hms.harvard.edu

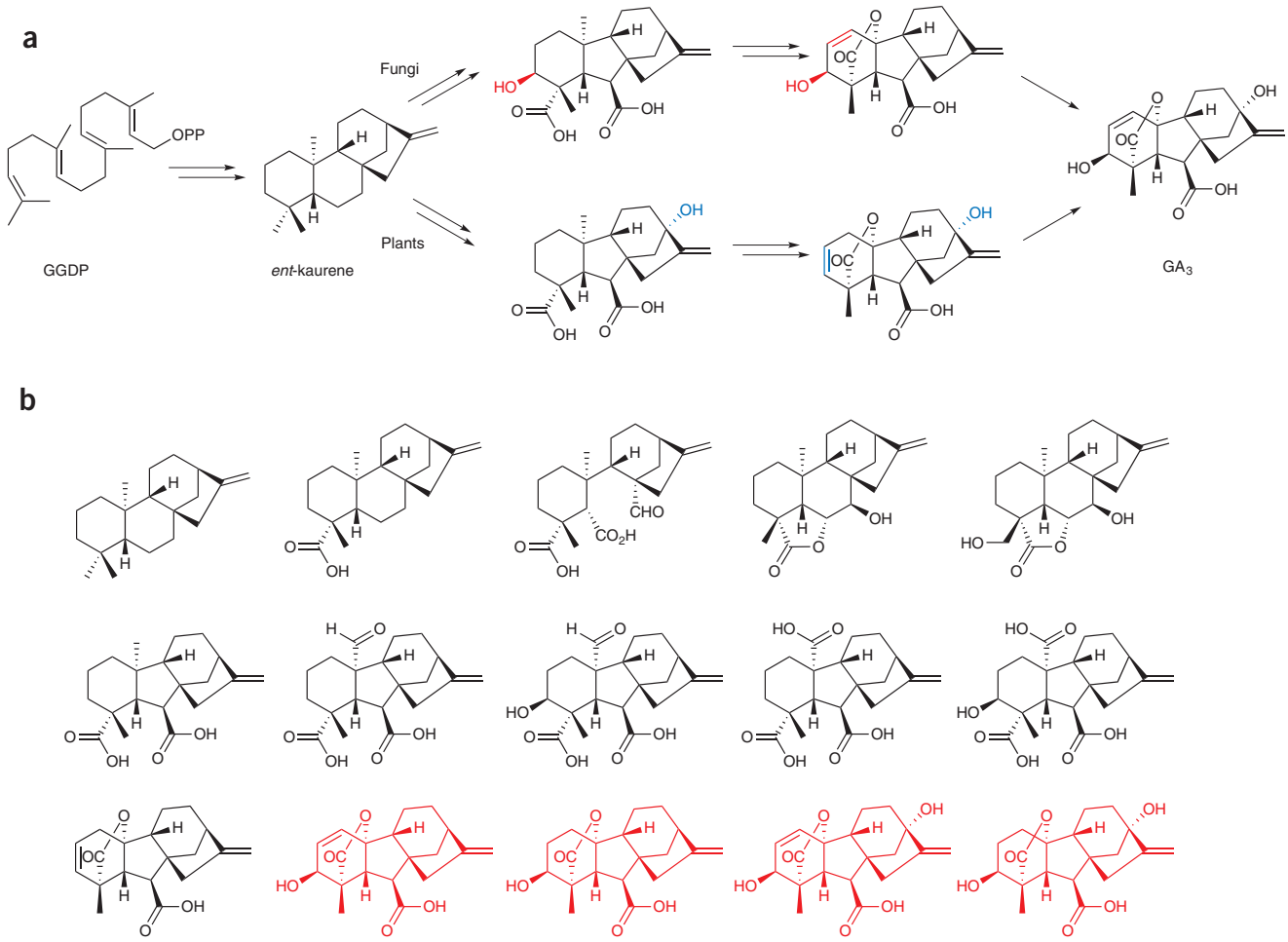

Figure 2 The gibberellin biosynthetic pathway. (**a**) Fungi and plants convert the carbon skeleton *ent*-kaurene to gibberellic acid 3 (GA_3) by a different series of oxidative transformations, thus showing that these pathways are evolutionarily convergent. Fungal transformations are shown in red and plant transformations are shown in blue. (**b**) The 15 most abundant gibberellin-family metabolites isolated from a culture of *Fusarium fujikuroi*[7]. Those with a known biological activity are shown in red.

fascinating example of convergent biosynthetic evolution[7,8], as they are produced by at least three sources: higher plants, which use them as hormones to regulate growth and other developmental processes; fungi, which use them to deregulate plant growth and render plants more susceptible to fungal infection; and bacteria, for which their biological role is not well characterized. We focus on two questions: how does a pathway produce so many gibberellin derivatives, and why has evolution selected a pathway with such profound promiscuity?

Gibberellin biosynthesis comprises two phases that are common to most terpene pathways. In the first phase, five-carbon building blocks are joined together and then cyclized to form a basic hydrocarbon skeleton; in the second phase, this carbon skeleton is modified by oxidative enzymatic transformations (**Fig. 2**). We focus here on the fungal gibberellin pathway because it has been characterized in more detail than the plant and bacterial pathways.

In the first phase of gibberellin assembly, four five-carbon units are joined to form the linear twenty-carbon polymer geranylgeranyl diphosphate (GGDP), and GGDP is cyclized by two successive cyclase enzymes to form a carbon skeleton called *ent*-kaurene, which is the precursor to all the gibberellins. There are possibilities for structural diversity in this first stage, because biosynthetic enzymes can join the five-carbon units together in more than one way[9] or fold the linear polymer into more than one kind of carbon skeleton[10–12]. But the gibberellin pathway seems to form *ent*-kaurene exclusively.

The second phase of gibberellin biosynthesis by oxidative modifications is the source of all of its diversity. The *ent*-kaurene carbon skeleton is successively oxidized by four cytochrome P450 monooxygenases, which are enzymes that incorporate one atom of oxygen from O_2 into a C-H bond in the molecule. In the gibberellin pathway, some of the cytochrome P450 enzymes insert an oxygen into more than one position on the molecule—but only some of the time. Because several of these extra oxygenation events occur on independently chosen subsets of the gibberellin pool, the result is like any combinatorial chemistry effort, in which diversity increases exponentially with the number of differential operations. Importantly, unlike the programmed diversity of a combinatorial chemistry scheme, the gibberellin pathway makes use of a stochastic diversification process. Plants adopt a similar strategy of diversification through late-stage oxidations but use very different tactics. For example, a key oxidation at C3 is an early P450 monooxygenase–catalyzed step in the fungal pathway but a late 2-oxoglutarate–dependent dioxygenase-catalyzed step in the plant pathway (see GA_3 in **Fig. 2a**). These and many other differences make a strong case for convergent evolution—two different organisms hitting upon the same overall strategy to make the

same collection of small molecules using unrelated gene products[8].

Why has a pathway with such florid molecular diversity emerged not once but twice? Only a few of the 100-plus gibberellins have a known biological activity, but those few that are active are potent at nanomolar concentrations. The rest? One possibility is that terpene biosynthetic pathways are intrinsically 'sloppy'; in other words, they are not capable of evolving to produce a single product. However, the fidelity of the cholesterol pathway and of cytochromes P450 in myriad other pathways is a convincing counterargument, and we can reject the possibility that terpene pathways are inherently incapable of high fidelity.

There are three other explanations for molecular promiscuity that bear consideration. First, all of the minor products could have an important biological activity, so the presence of any pathway product implies that evolution has selected for it specifically. This level of complexity seems unlikely; although three or four products from a pathway might have important (and different) activities, it is unlikely that every one of the 136 products is important (**Fig. 2**). Second, selective pressure to improve fidelity through the pathway might be weak or absent. Because the potency of the active gibberellin derivatives requires only a low carbon flux through the pathway, the quantity of side products produced is miniscule, so higher fidelity would not confer much of a selective advantage. This explanation requires that the side products are neither toxic to the producing cell nor capable of interfering with the active products, which would be the case if molecular discrimination occurred at the level of the plant gibberellin receptor, as is known for many other hormones, pheromones and semiochemicals[13].

A third possible explanation is that evolution has selected promiscuous terpene pathways. Because selection acts at the level of the gene, a set of genes encoding a terpene pathway that makes minor side products might be particularly evolvable; that is, this set of genes would be fewer mutational steps away from making a new natural product that meets a new selective need than a more rigid pathway would be. More importantly, the evolutionary changes would be quantitative, not qualitative, and many recent analyses highlight the importance of quantitative changes in evolutionary diversification[14]. A tiny amount of the new product is all that is required as a starting point for quantitative evolutionary improvement. If such facile evolvability were a common feature of terpene pathways, one might imagine that this would have favored their propagation (horizontally among microbes, and by gene duplication in plants) relative to other, more rigid pathways.

Is the gibberellin pathway representative of other terpene biosynthetic pathways? At first glance, it seems that the gibberellin pathway must be at one extreme of the fidelity-promiscuity spectrum. But this appearance might be deceptive because the gibberellin pathway has been studied in much greater detail than other terpene pathways—gibberellins are important items of commerce, and improved fungal strains can now produce in excess of 300 mg l^{-1}. If we studied other terpene pathways[15] in the same detail, would we find a similarly large spectrum of minor side products? Several recent reports on terpenes from well-studied plants suggest that terpene families resemble a set of Russian dolls, in which a wooden figure is pulled apart to reveal a similar figure, which in turn reveals a similar figure inside it, and so on. Reexamination with new analytical techniques and larger quantities of starting material invariably reveals new products from what seem to be diversity-based biosynthetic pathways. And finally, given that we have focused on secondary metabolic pathways for terpene biosynthesis, how much (or little) would these arguments change for other natural product pathways, such as those that produce polyketides and nonribosomal peptides? In this regard, it is worth noting that many familiar therapeutic compounds—amphotericin B, cyclosporin A, the bacitracin complex and even rapamycin—are members of families of closely related molecules. We have made great progress in understanding the structures, biosynthetic pathways, and roles of natural products, and we now know enough to begin connecting these structures, pathways, and roles into a broader evolutionary perspective. Making this evolutionary connection is the next big challenge for natural products chemistry, and it will require systematic studies on genomes, genetically encoded small molecules, and their ecological and physiological roles.

ACKNOWLEDGMENTS

We dedicate this paper to A.I. Scott for his pioneering efforts in natural product biosynthesis. We thank C. Walsh for helpful discussions. J.C. is supported by US National Institutes of Health grants CA59021 and CA24487, and M.A.F. is supported by a fellowship from the Hertz Foundation.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

1. Stone, M.J. & Williams, D.H. *Mol. Microbiol.* **6**, 29–34 (1992).
2. Williams, D.H., Stone, M.J., Hauck, P.R. & Rahman, S.K. *J. Nat. Prod.* **52**, 1189–1208 (1989).
3. Choi, J., Chen, J., Schreiber, S.L. & Clardy, J. *Science* **273**, 239–242 (1996).
4. Firn, R.D. & Jones, C.G. *Mol. Microbiol.* **37**, 989–994 (2000).
5. Firn, R.D. & Jones, C.G. *Nat. Prod. Rep.* **20**, 382–391 (2003).
6. Firn, R.D. & Jones, C.G. *Trends Plant Sci.* **11**, 112–113 (2006); author reply **11**, 113–114 (2006).
7. Tudzynski, B. *Appl. Microbiol. Biotechnol.* **66**, 597–611 (2005).
8. Hedden, P., Phillips, A.L., Rojas, M.C., Carrera, E. & Tudzynski, B. *J. Plant Growth Regul.* **20**, 319–331 (2001).
9. Thulasiram, H.V., Erickson, H.K. & Poulter, C.D. *Science* **316**, 73–76 (2007).
10. Phillips, D.R., Rasbery, J.M., Bartel, B. & Matsuda, S.P. *Curr. Opin. Plant Biol.* **9**, 305–314 (2006).
11. Yoshikuni, Y., Ferrin, T.E. & Keasling, J.D. *Nature* **440**, 1078–1082 (2006).
12. Zhang, H., Shibuya, M., Yokota, S. & Ebizuka, Y. *Biol. Pharm. Bull.* **26**, 642–650 (2003).
13. Richards, D.E., King, K.E., Ait-Ali, T. & Harberd, N.P. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 67–88 (2001).
14. Carroll, S.B. *Cell* **101**, 577–580 (2000).
15. Brady, S.F., Bondi, S.M. & Clardy, J. *J. Am. Chem. Soc.* **123**, 9900–9901 (2001).