Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Enzyme evolution in natural products biosynthesis: target- or diversity-oriented?

Lianet Noda-Garcia[1] and Dan S. Tawfik[2]

Abstract

Natural product biosynthesis (NPB) is the Panda's Thumb of evolutionary biochemistry. Arm races between organisms, and ever-changing environments, result in relentless innovation. This review focusses on enzyme evolution in NPB. First, we review cases of *de novo* emergence, whereby a completely new enzymatic activity arose in a ligand-binding protein, or a new enzyme emerged including a completely new scaffold. Second, we briefly review the current models for enzyme evolution, and how they explain the inherent promiscuity of NPB enzymes and their tendency to produce multiple related products. We thus suggest that NPB enzymes a priori evolved to generate a specific product; they are, however, trapped in a multifunctional, generalist evolutionary state and thereby produce a diversity of products.

Addresses

[1] Department of Plant Pathology and Microbiology, Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel

[2] Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot, 76100, Israel

Corresponding author: Tawfik, Dan S (dan.tawfik@weizmann.ac.il)

Current Opinion in Chemical Biology 2020, **59**:147–154

This reviews comes from a themed issue on **Mechanistic Biology**

Edited by **Katherine Ryan** and **Tomohisa Kuzuyama**

For a complete overview see the Issue and the Editorial

Available online 6 August 2020

https://doi.org/10.1016/j.cbpa.2020.05.011

Keywords

Enzyme evolution, Natural products biosynthesis, Chemical diversity, Promiscuity, Generalist.

Abbreviations

Natural product biosynthesis, NPB.

Introduction

A myriad of natural products has been and continue to be identified, mainly in microorganisms and plants. Studying how these molecules are produced (the biosynthetic pathways) and used (their biological function) led to the realization that the enzymes and pathways that mediate natural product biosynthesis (NPB) differ fundamentally from their anabolic counterparts in central metabolism. Indeed, NPB has become the Panda's Thumb of evolutionary biochemistry — it vividly demonstrates the powers of evolutionary tinkering and how new products arise by rapid recruitment of preexisting components.

NPB enzymes differ fundamentally from those of central metabolism

In central metabolism, the identity of metabolites and their physiological roles are relatively well established. The enzymes have diverged at early evolutionary stages and largely remained unchanged ever since. Thus, although evolutionary relationships between enzymes can be inferred, the trajectories that led to their divergence are rarely tractable. In contrast, the selection pressures shaping NPB are constantly changing. Thus, the evolutionary history of NPB enzymes and pathways, that emerged relatively recently, can be unraveled in detail. In fact, most of our detailed knowledge of enzyme evolution comes from enzymes of specialized metabolism, and of NPB in particular.

The enzymes mediating central metabolism differ from those of NPB not only in their evolutionary history. In addition, in central metabolism, demands for high flux and the high cost of breaches of specificity, led to catalytically efficient and highly selective enzymes, often hitting the physicochemical limits of molecular recognition [1]•. The high connectivity of enzymes and metabolites in central metabolism (substrates and intermediates shared by multiple pathways) [2], also hinders changes in substrate or/and reaction specificity. NPB stands in contrast — it evolves under rapidly changing environments (the chemical arm race). Moreover, NPB pathways are active only in specific environmental conditions, and typically, at low flux. Accordingly, enzymes in secondary metabolism exhibit on average 30-fold lower turnover number (kcat) than those in central metabolism [3]. NPB pathways are also more singular than those of central metabolism — their intermediates and products are more rarely used by other enzymes in the same organism (low connectivity). These features, low flux and low connectivity, result in the cost of poor selectivity being minimal.

The outline of this review

Broadly speaking, in both central and natural products metabolism, the evolution of enzymes occurs in two

steps: *recruitment* and *shaping*. The first part of this review relates to *recruitment*. Typically, a preexisting enzyme is recruited to catalyze the same, or a similar reaction, on a new substrate (discussed later, in 'The shaping of new enzymatic activities'). Occasionally, however, an enzyme emerges 'out-of-the-blue' — *e.g.* from a noncatalytic ligand-binding protein. These events that mark the birth of a new enzyme are tractable in NPB. The following section, '*De novo* enzyme emergence in NPB', describes few such intriguing cases.

The second step is shaping or reshaping. Following the recruitment of an enzyme, mostly via a latent promiscuous activity, the enzyme adapts to its new role, namely, evolves towards higher catalytic efficiency and selectivity. In "The shaping of new enzymatic activities" we briefly describe the evolutionary forces and constraints under which the catalytic efficiency and selectivity of enzymes are shaped. We show that these constraints explain the multifunctionality and the high levels of promiscuity that characterize NPB enzymes, and thereby their tendency to generate a multitude of products, sometimes dozens of related products. We therefore argue that NPB enzymes did not a priori evolve to generate a diversity of products as suggested by the 'screening hypothesis' [4]. Rather, we propose that NPB enzymes and pathways generally evolved as target-oriented, namely, under selection to generate a specific product that provides a distinct selective advantage. Yet, the mechanism that shape enzyme selectivity, and the evolutionary constraints acting on them, resulted in diversity-generating pathways.

Once recruited and shaped for a new activity, individual enzymes are assembled in pathways. Several recent reviews have focused on pathway evolution [2], including in NPB [5,6]. This review, however, focuses on the evolution of individual enzymes.

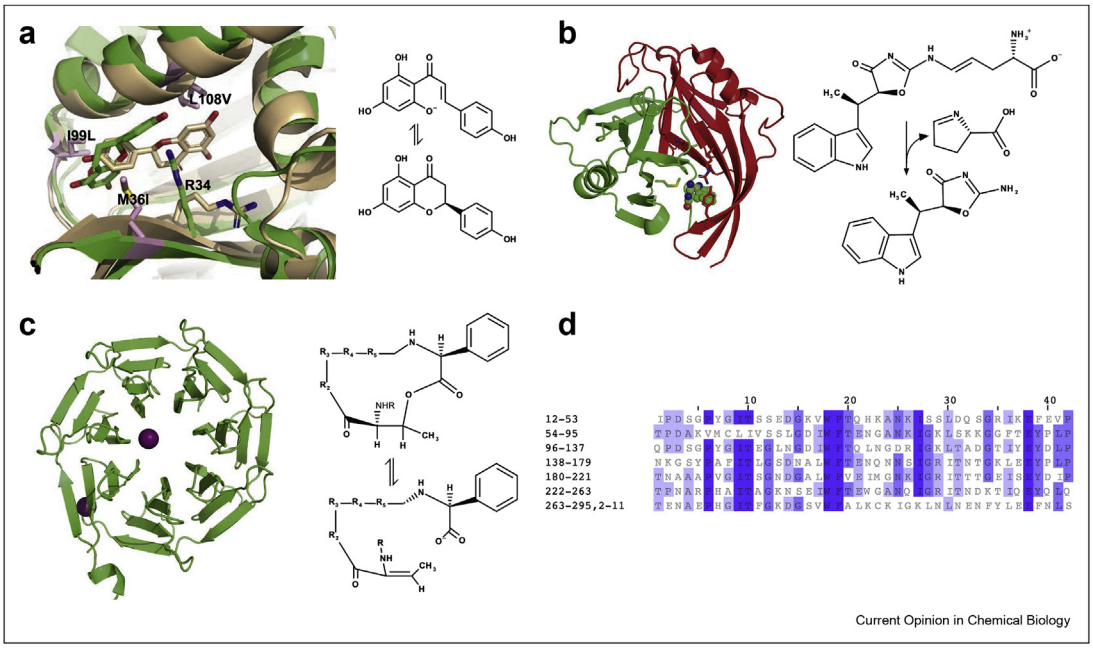
De novo enzyme emergence in NPB

Broadly speaking, NPB enzymes have various origins: preexisting NPB pathways, or other pathways, either in the same organism (typically, from central metabolism; for example see Ref. in a study by Carrington et al.[7]) or a different one (horizontal gene transfer [8]). Typically, in the recruitment step, a promiscuous activity, namely a latent enzymatic activity with no physiological relevance, becomes physiologically relevant, and hence 'visible' to natural selection (elaborated further in the next section). It therefore follows that in most cases, the recruited enzyme applies the same chemistry on a different substrate, or with minor modifications of chemistry [9]. The recruited enzyme may even retain its original reaction and substrate and be merely rewired in terms or regulation or/and localization [10–12]. Nonetheless, combined with other enzymes, a new pathway generating a new natural product may emerge. While this mix and match dominates the evolution of NPB, there are some unique cases of enzymes that emerged 'out-of-the-blue'. In central metabolism these events typically occurred at the early stages of evolution and are largely intractable. However, NPB, owing to ever-changing nature, presents multiple such examples. '*De novo* emergence' may involve a new enzymatic active site that emerges in a preexisting fold, for example, in a ligand-binding protein devoid of any catalytic activity (Figure 1a and b), or even a completely new protein, with a new scaffold, as well as a new active site (Figure 1c).

One case that has been studied in detail is chalcone isomerase. This enzyme catalyzes a Michael-addition leading to (2S)-Naringenin, a key intermediate in plant flavonoid biosynthesis. Uncharacteristically, this isomerase diverged from a nonenzymatic ancestor, a long-lost fatty acid–binding protein [13]•. This ancestor could nonetheless be reconstructed in the laboratory, and as anticipated, exhibits no isomerase activity. It binds the substrate, yet with the 'business end' facing the active site exit (Figure 1a). Reconstruction of the trajectory that potentially led from this inactive ancestor to the contemporary enzyme indicated that single point mutations of a surprisingly subtle nature (effectively adding or removing a methyl) induce stereoselective isomerase activity. These mutations slightly reshape the active site pocket allowing the substrate to bind in the right mode, and also realign the side chain of Arg34. This residue which is the key to the isomerase activity was already present in that fatty acid–binding ancestor, and the mutations merely realigned it to enable catalysis [13]• (for a similar case of a ligand-binding protein diverging to an enzyme Ref. [14]•).

Two other notable examples regard the pathogen-antibiotics arms race, a race that seems to boost enzyme innovation. Both cases bear the hallmarks of *de novo* emergence, although this has not been explicitly noted in the original publications. The first is a recently discovered enzyme, PluN2, that catalyzes imine hydrolysis in a pathway leading to an indolmycin antibiotic [15]• (Figure 1b). The enzyme was found in a unique pathway that evolved independently of the known indolmycin biosynthesis pathway found in *Streptomyces*. Although most of the pathway's enzymes are orthologs of the *Streptomyces* ones, PluN2 seems to have emerged independently. The active site lies at the dimer interface of the versatile VOC fold [16]. This fold comprises mostly metallo-enzymes, yet some family members diverged into a nonmetal-binding protein, giving rise to so-called bleomycin resistance proteins that confer resistance by stoichiometric binding of the antibiotic. It seems that PluN2, an antibiotics synthesis enzyme, comprises an 'orphan' enzyme that arose from a bleomycin resistance protein, while using the periphery of

Figure 1



Examples of *de novo* emergence of NPB enzymes. (**a**) Chalcone isomerase diverged from an ancestor that exhibits no isomerase activity [13]•. The ancestor (in green; PDB:5WKR) binds the substrate, albeit in the opposite mode, as illustrated by the comparison to a contemporary chalcone isomerase with the bound substrate (in pink; PDB: 1EYQ; note the different sidechain of Arg34 the rotamer of which differs in the two structures). Highlighted in pink are the mutations that reshaped the ancestral binding pocket to give a stereoselective enzyme. (**b**) PluN2 catalyzes the hydrolysis of an imine leading to an indolmycin antibiotic (PDB: 6P29). The active site marked by the substrate in spheres lies at the dimer interface [15]•. (**c**) Along the same vein, virginiamycin B lyase mediates antibiotics resistance (PDB: 2qc5) [17]. The enzyme adopted a beta-propeller that emerged *de novo*, with the active site residing at the top of the propeller (facing the reader). (**d**) The internal sequence identity between the seven repeats of virginiamycin B lyase indicates its *de novo* emergence, by co-option of a 45-residue fragment, and its duplication and fusion to give an intact β-propeller. The 7th repeat that carries most of the active site residues seems to have diverged the most [17].

the original antibiotic binding site for the new enzymatic active site.

The second case is virginiamycin B lyase, an enzyme that mediates antibiotics resistance (Figure 1c). The enzyme catalyzes a retro-Michael (beta-elimination) reaction that leads to lactone opening and inactivation of virginiamycin [17]. It is a beta-propeller with a clear fingerprint of high sequence identity between its seven blades (Figure 1d), thus indicating *de novo* emergence via duplication and fusion of a short segment co-opted from another protein [18,19].

A common feature of these three examples of *de novo* emerged enzymes is that they catalyze 'easy' reactions that have low activation energy barrier and hence occur spontaneously at considerable rate. These reactions may have initially occurred without an enzyme, and once their product became advantageous an enzyme emerged that catalyzes this step. Indeed, spontaneously occurring reaction are rare in central metabolism yet quite common in NPB [2]. In some cases, the role of the evolving enzyme may have been not in rate acceleration per se, but, for example, in directing the reaction's stereoselectivity [13]•.

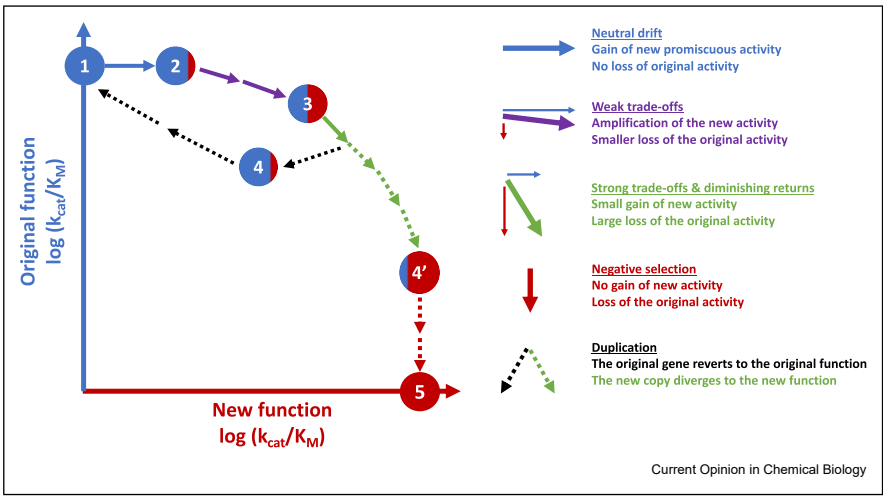
Other interesting cases that demonstrate the remarkable powers of evolutionary tinkering in NPB enzymes regard drastic changes in the catalyzed reaction. For example, an S-adenosylmethionine-dependent methyltransferase that diverged into a decarboxylase, or an oxygenase [6]•. Other oxygenases in NPB unexpectedly diverged from pyridoxal phosphate–dependent enzymes that typically catalyze isomerizations or decarboxylations [20]•.

The shaping of new enzymatic activities

In a nutshell

The starting point is typically a promiscuous activity, namely a preexisting activity that generates a product that is beneficial upon a new challenge. Promiscuous activities appear (and disappear) coincidently, as gene sequences drift under purifying selection, that is, while maintaining the original activity [21,22] (Figure 2). After the initial recruitment of an enzyme, mutations that amplify the newly recruited activity would be

Figure 2



The shaping of a new enzymatic activity — a schematic representation of a characteristic mutational trajectory, its selection driving forces, trade-offs and diminishing returns. Each arrow indicates a single step, in principle, a single mutation along the trajectory leading from the starting point (*1*) to the end point (*5*). The trajectory begins with an enzyme that carries a given function (the original function, plotted along the Y-axis, in blue). Mutations that are neutral with respect to the original function (*blue arrow*) may accumulate by chance (drift, polymorphism, and so on) and give rise to an enzyme variant with a coincidental, promiscuous activity that is typically orders of magnitude weaker than the original one (X-axis, in red) (*2*). A change in the environment results in this promiscuous activity has a beneficial effect, and thus becoming under selection for higher rate (positive selection). The early mutations (*purple arrows*) are responsible for the acquisition of higher levels of the new activity; they compromise the original function, yet typically to a smaller extent (weak trade-off). This first stage therefore yields a bifunctional, generalist intermediate (*3*) that may evolve further. However, the next stage (*green arrows*) differs from the first one in two respects: (*i*) the improvement per mutation in the new activity becomes increasingly smaller (diminishing returns); (*ii*) the trade-off intensifies — improvements in the new activity come at increasingly large cost with respect to the original one. At this stage, demands for higher activity for the original, as well as the new function, and a stronger trade-off, would typically enforce duplication — the original gene reverts to the original function (*4*, black dashed arrows), whereas the new copy keeps evolving toward the new function (green dashed arrows). As the trajectory progresses, diminishing returns are strong, and thus multiple mutations each with a small effect are needed to obtain a highly efficient enzyme (*4′*). To complete the shaping of the new specialist (*5*), 'negative selection' would typically act in parallel (*red arrows*), leading to the selection of mutations that abolish the original activity.

enriched by selection, and eventually fixed, thus optimizing the new enzyme toward the target substrate and reaction.

Gene duplication typically takes place at some stage. It may occur early on, before, or immediately after the recruitment of a promiscuous activity [23]. But it may also occur at much later stages, or not occur at all. Divergence trajectories have been studied in detail, and reproduced in laboratory experiments, indicating that in many cases the new activity is amplified by mutations, while the original activity is retained [22,24]. Thus, a newly diverging gene would typically become a 'generalist' — a bi-, or even multifunctional enzyme, carrying both the original and a new function. This phenomenon of 'gene sharing' has been observed early on [25,26], and many reconstructed ancestors of what are now multiple paralogous enzymes, were shown to be generalists [27] (for an exception, and further discussion of this issue, see a study by Rauwerdink et al. [28]). These generalist ancestors were eventually duplicated, and subfunctionalized to give two or more enzymes, each specializing in one of the ancestral functions. Indeed, generalists are widely deemed an intermediate step, typically followed by gene duplication and specialization.

The generalist state may persist

In many cases, however, duplication and specialization does not occur — the same gene continues to fulfill multiple functions [26]. Another common scenario is that although duplication had occurred, the enzymatic function remained largely unchanged, and divergence related to changes in regulation [10], localization [11], or in moonlighting functions [12].

Why? The major driving force for specialization is trade-offs: mutations that improve the new activity inevitably come at the expense of the original one. If the latter drops to an extent that compromises fitness, duplication becomes inevitable — the original gene maintains the original function, while the new copy, freed from the burden of carrying the original function, can readily diverge toward the new function. This trade-off is, however, often weak, at least at the early stages of divergence (Figure 2). Laboratory evolution experiments consistently show that promiscuous activities are highly evolvable: mutations that amplify them tend to have

weak effects on the original function even when the latter is not under selection [22,29,30]. In the laboratory, a latent promiscuous activity can be increased by up to 1000-fold with only several-fold decrease in the original function [31]. Thus, if trade-offs are weak on the one hand, and the target, new activity need not be highly efficient on the other, a 'generalist' intermediate is a sustainable long-term solution.

Persistence of the 'generalist' state also relates to the fact that optimization of new activities is subject to strong diminishing returns. The early mutations provide large improvements, whereas the subsequent ones (that also exhibit large trade-offs) provide increasingly smaller improvements (Figure 2). Fixation of a mutation strongly depends on the magnitude of its fitness benefit. If these are small, mutations that drive the generalist intermediate toward specialization may never fixate (especially if population size is limited) [24]. In specialized metabolism, fluxes are low as natural products are usually produced at low quantities. In addition, the contributions of natural products to organismal growth and survival tend to be small. Hence, the selection forces may be too weak to drive specialization.

Specialization usually demands 'negative selection'

Specialization is driven by demands for adequate regulation and localization, as well as for high catalytic efficiency with the target substrate. The latter comprises a 'positive selection' that shapes the enzyme's active site toward a specific substrate and transition state. If a strong trade-off prevails, positive selection will also lead to lower activity with the original substrate (Figure 2, *green arrows*) and possibly reduce activity with other potential substrates (lower promiscuity). It appears, however, that positive selection rarely yields high selectivity ([24] and [1]• for a more detailed discussion). Indeed, selectivity is typically shaped by 'negative selection', namely selection to exclude certain substrates (Figure 2, red arrows) [1]•. For example, in a generalist, bifunctional state, the original substrate acts as competitive inhibitor, thus lowering the new activity. Thus, mutations that strongly reduce the original function would be favored, and this reduction might drive duplication and subspecialization [32] (Figure 2, dashed black arrows). Negative selection is also driven by harmful effects of side products. In central metabolism this is evident in the evolution of editing in certain enzymes (polymerases, aminoacyl tRNA synthetases, etc') [1]•, as well as in repair enzymes that evolved to degrade so-called damage metabolites [33,34]. These error-correcting mechanisms also indicate that there are physicochemical constrains as to how selective an enzyme can become. These depend on structural differences between the desirable substrate (which is under positive selection) and the undesirable ones (that drive negative selection), and also on the active site's structure [1]•. These physicochemical constrains also hinder specialization and may thus lead to the persistence of generalist states.

Generalists are inherently promiscuous and evolvable

This may sound purely semantic, yet for an evolutionary biochemist, there exists a fundamental difference between bi-, multi-or broad-functionality (terms that are largely synonymous) and promiscuity. The former is shaped under selection, whereby two or more functions are physiologically relevant. In contrast, promiscuity is coincidental and regards latent activities with no physiological relevance [29]. That said, these two phenomena are interlinked. As shown by laboratory evolution, when a given promiscuous activity is enhanced via mutation and selection, the resulting enzyme variants exhibit not only the original function and the newly evolved one (bifunctionality), but also a range of new promiscuous activities that were neither present in the starting enzyme, nor selected for. Thus, the transition from a specialist to a generalist yields not only bifunctional enzymes, but also inherently promiscuous and evolvable ones — such generalists are therefore fewer mutations away from a new function compared with a specialized enzyme [35], as also demonstrated with reconstructed ancestors [27].

NPB enzymes are trapped in the generalist state

NPB usually demands low flux, and thus may not impose selection pressures that are sufficiently high to enforce specialization. The physiological costs of low selectivity, namely of acceptance of alternative substrates, and the subsequent reduction in the target activity and/or production of non-desirable products, may be minimal compared with central metabolism. Finally, NPB evolves relatively frequent, in response to chemical arm races and rapidly changing environments. This means relatively frequent events of enzyme reshaping, and in turn, these enzymes would tend to remain in the generalist state. Along the same vein, NPB enzymes are frequently subject to horizontal gene transfer, and landing in a new organism, let alone serving in a new pathway, demands changes in activity. Thus, given the constrains acting on specialization on the hand, and the low incentive to specialize on the other, most NPB enzymes might remain in the generalist state (Figure 2, step 3). Indeed, multifunctional NPB enzymes are widely observed which either mediate few consecutive steps (*e.g.* a series of methylations of different groups of the same substrate) or even entirely different reactions [6]•. When combined in a pathway, the action of individual generalist, promiscuous enzymes can accumulate to give dozens of related products (for example, if each enzyme generates two products that can be taken as substrates

by the next one, a pathway comprising *n* enzymes would result in 2^n products).

The ancestors of central metabolism enzymes were likely similar. Peptidoglycan biosynthesis, for example, includes four consecutive condensation steps, each adding one additional amino acid to the growing glycopeptide chain. These steps are mediated by four different enzymes that diverged from a common ancestor that probably recursively catalyzed all these, or at least a subset of these steps [2]. Divergence and specialization of this generalist ancestor enabled high rate of synthesis of a single product, as expected for a product that directly relates to growth (cell wall formation) and hence has an immediate effect on fitness.

Target- or diversity-oriented evolution?

Our current knowledge of how enzyme evolves calls for a reexamination of the widely accepted 'screening hypothesis' [4]. That NPB pathways typically generate a multitude of products is beyond question, what is in question, however, is why. This section examines two opposing 'why' hypotheses:

(1) The chemical diversity observed in NPB is an explicitly evolved trait. In other words, NPB enzymes are diversity oriented [36], they have a priori evolved as multifunctional such that they generate a myriad of products, including products that have no biological role; further, these enzymes are constantly maintained under selection to remain so. Selection is therefore acting to 'prevent the evolution of highly efficient and specific enzymes because such a high specialization state could be detrimental for increasing chemical diversity' [6]•.
(2) NPB enzymes are target oriented [36], namely, they evolved in response to a specific challenge, a challenge that preceded their emergence, and are also maintained under selection to produce a specific product that meets a specific challenge. In other words, the emergence of an enzyme, or a pathway, is driven by an imminent fitness advantage that its product provides, and so are changes in specificity that may occur in preexisting pathways. By this hypothesis, the tendency of NPB enzymes to generate a diversity of products is a mere side-effect of being trapped in a generalist state.

We surmise that (2) should be the default hypothesis. As detailed above, NPB enzymes tend to be in the generalist state owing to weak specialization forces, and also because chemical arms race, and changing environments, drive relatively frequent changes in enzyme specificity [37] (*Generalists are inherently promiscuous and evolvable*).

Note, however, that 'frequent changes' is on an evolutionary, multigeneration time scale, not within a single generation as is the case with the immune system — the analogy provided by Firn and Jones [4] for a diversity-oriented system. The immune system produces a huge repertoire of antibodies that are not targeted towards a specific antigen. This so-called naïve repertoire may include up to 10^{12} antibodies, each exhibiting a different binding specificity, such that upon arrival of a new challenge, a suitable antibody is immediately available. Along the same rationale, NPB pathways evolved to produce a large diversity of compounds regardless of their biological activity [4]. Preexisting diversity certainly provides an evolutionary advantage, as does enzyme promiscuity. However, the deduction that product diversity in NPB is an explicitly evolved trait, as in the immune system, is Panglossian (glasses are held on noses, it does not mean noses evolved to hold glasses) [38]. For a start, in the immune system, a mechanism for generating diversity per se is available (V(D)J recombination, mutator enzymes, and so on). However, as for today, evidence for a similar mechanism of generating diversity in NPB regardless of biological function is scarce (further discussed in the following context).

Evidence for target- vs. diversity-oriented evolution

The biological function of natural products is often invisible. Nonetheless, it became clear that even if the molecular function of certain natural products remains unknown, there are environmental conditions, typically highly specific, which led to their emergence [5]. There are multiple evidences indicating that most NPB pathways, evolved to make a specific product, even if the precise biological role of it may remain unknown [39–41]. Other indications include, for example, that antibiotic resistance evolved concomitantly with the emergence of the synthesis of these antibiotics. This suggests that these pathways evolved a priori toward a specific molecule with a defined and beneficial biological function [39].

What about evidence in favor of diversity-oriented evolution? Natural products typically come in families that have a shared scaffold made by a set of relatively conserved enzymes. The scaffold is 'decorated' by multiple different modifications. The latter is equivalent to 'lead optimization' in drug development, and is hence subject to frequent evolutionary changes, and also leads to a diversity of products [6•,35]. Might the decorating enzymes evolve to generate diversity per se?

One case that comes closest to the generation of diversity pe se is the extensive radiation of lanthipeptides in cyanobacteria [42]•. Here, a single synthetase that exhibits very broad substrate acceptance can generate a whole range of cyclic peptides. A metagenomic survey

indicated that the genes encoding the precursor peptides (prochlorosins) have intensely diversified. Collectively, the examined natural cyanobacteria populations have the potential to produce thousands of different peptides that widely vary in length, sequence and ring topologies. However, whether this system is diversity oriented remains an open question [42]•. More recent work showed that the occurrence and genomic location of the prochlorosin genes are tightly linked with a transposase, suggesting the evolution of a mechanism that promotes diversification [43]. The cyanobacterial lanthipeptides, however, is an exception. There might be other exceptions. However, in most NPB pathways, product diversity is not an evolved trait but rather an outcome of the way by which enzymes evolve.

Concluding remarks

NPB is possibly the best and wildest demonstration of biochemical evolution in action. It provides stunning examples of the powers of evolutionary innovation; how few mutations can generate a completely new metabolite that in turn initiates a new pathway that generates dozens of different products. It provides ample opportunities to examine the evolutionary trajectories and mechanisms that shape enzymes. At the same time, this extravagant act of evolution can be misleading. Panglossian rationale is intuitively appealing: enzyme promiscuity (and other sources of noise and infidelity) is critical to evolutionary innovation. Yet should we then deduce that enzymes a priori evolved, or/and are maintained by selection, to be promiscuous? Similarly, although evolution depends on mutations, polymerases did not evolve to be error prone [44]. Evidence in favor of mechanisms of 'evolved evolvability' is scarce. Rather, as elaborated here, promiscuity is the outcome of evolutionary and physicochemical constrains, and of how enzymes emerge and are further shaped. NPB enzymes also comprise a powerful demonstration of these constrains.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgements

D.S.T. is the Nella and Leo Benoziyo Professor of Biochemistry. Funding by the Abney Foundation is gratefully acknowledged. L.N.G. acknowledges the funding support of the CONACYT (Mexico) grant #203740 and the Martin Kushner Fellowship, from the Weizmann Institute of Science, Israel. The authors thank Wilfred van der Donk, Mikko Metsä- Ketelä, and Pablo Cruz-Morales for their insightful comments.

References

1. Tawfik DS, Gruic-Sovulj I: **How evolution shapes enzyme selectivity - lessons from aminoacyl-tRNA synthetases and other amino acid utilizing enzymes**. *FEBS J* 2020, **287**: 1284–1305, https://doi.org/10.1111/febs.15199.

This review provides background reading to this article, with more detailed discussion of the evolutionarily forces that shape enzyme selectivity and of the physicochemical constraints of selectivity.

2. Noda-Garcia L, Liebermeister W, Tawfik DS: **Metabolite-enzyme coevolution: from single enzymes to metabolic pathways and networks**. *Annu Rev Biochem* 2018, **87**:187–216, https://doi.org/10.1146/annurev-biochem-062917-012023.

3. Bar-Even A, Noor E, Savir Y, *et al.*: **The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters**. *Biochemistry* 2011, **50**:4402–4410, https://doi.org/10.1021/bi2002289.

4. Firn RD, Jones CG: **The evolution of secondary metabolism - a unifying model**. *Mol Microbiol* 2000, **37**:989–994, https://doi.org/10.1046/j.1365-2958.2000.02098.x.

5. Chevrette MG, Gutiérrez-García K, Selem-Mojica N, *et al.*: **Evolutionary dynamics of natural product biosynthesis in bacteria**. *Nat Prod Rep* 2020, **37**:566–599, https://doi.org/10.1039/c9np00048h.

6. • Fewer DP, Metsä-Ketelä M: **A pharmaceutical model for the molecular evolution of microbial natural products**. *FEBS J* 2020, **287**:1429–1449, https://doi.org/10.1111/febs.15129.

A recent, informative review that addresses enzyme and pathway evolution in NPB, provides insightful examples of evolutionary tinkering, and also a new evolutionary model.

7. Carrington Y, Guo J, Le CH, *et al.*: **Evolution of a secondary metabolic pathway from primary metabolism: shikimate and quinate biosynthesis in plants [published online ahead of print, 2018 Jun 12]**. *Plant J* 2018, https://doi.org/10.1111/tpj.13990.

8. Glasner ME, Truong DP, Morse BC: **How enzyme promiscuity and horizontal gene transfer contribute to metabolic innovation**. *FEBS J* 2020, **287**:1323–1342, https://doi.org/10.1111/febs.15185.

9. Furnham N, Dawson NL, Rahman SA, Thornton JM, Orengo CA: **Large-scale analysis exploring evolution of catalytic machineries and mechanisms in enzyme superfamilies**. *J Mol Biol* 2016, **428**:253–267, https://doi.org/10.1016/j.jmb.2015.11.010.

10. Nguyen Ba AN, Strome B, Hua JJ, *et al.*: **Detecting functional divergence after gene duplication through evolutionary changes in posttranslational regulatory sequences**. *PLoS Comput Biol* 2014, **10**:e1003977, https://doi.org/10.1371/journal.pcbi.1003977. Published 2014 Dec 4.

11. Qiu Y, Tay YV, Ruan Y, Adams KL: **Divergence of duplicated genes by repeated partitioning of splice forms and subcellular localization**. *New Phytol* 2020, **225**:1011–1022, https://doi.org/10.1111/nph.16148.

12. Espinosa-Cantú A, Ascencio D, Barona-Gómez F, DeLuna A: **Gene duplication and the evolution of moonlighting proteins**. *Front Genet* 2015, **6**:227, https://doi.org/10.3389/fgene.2015.00227. Published 2015 Jul 7.

13. • Kaltenbach M, Burke JR, Dindo M, *et al.*: **Evolution of chalcone isomerase from a noncatalytic ancestor**. *Nat Chem Biol* 2018, **14**:548–555, https://doi.org/10.1038/s41589-018-0042-3.

A reconstruction of the evolutionary trajectory leading from a ligand binding to a stereoselective enzyme that is a key to plant flavonoid biosynthesis. It describes in atomic how subtle changes in active site residues can turn a noncatalytic protein into a stereoselective enzyme (Figure 1A).

14. • Clifton BE, Kaczmarski JA, Carr PD, Gerth ML, Tokuriki N, Jackson CJ: **Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein**. *Nat Chem Biol* 2018, **14**: 542–547, https://doi.org/10.1038/s41589-018-0043-2.

A parallel case to Ref. [13] whereby a ligand binding protein diverged into an enzyme.

15. • Du YL, Higgins MA, Zhao G, Ryan KS: **Convergent biosynthetic transformations to a bacterial specialized metabolite**. *Nat Chem Biol* 2019, **15**:1043–1048, https://doi.org/10.1038/s41589-019-0331-5.

An interesting case of convergent evolution (making the same natural product via a different pathway) that was enabled by the de *novo*

emergence of an enzyme, likely from a ligand binding protein (Figure 1B).

16. He P, Moran GR: **Structural and mechanistic comparisons of the metal-binding members of the vicinal oxygen chelate (VOC) superfamily**. *J Inorg Biochem* 2011, **105**:1259–1272, https://doi.org/10.1016/j.jinorgbio.2011.06.006.

17. Lipka M, Filipek R, Bochtler M: **Crystal structure and mechanism of the Staphylococcus cohnii virginiamycin B lyase (Vgb)**. *Biochemistry* 2008, **47**:4257–4265, https://doi.org/10.1021/bi7015266.

18. Smock RG, Yadid I, Dym O, Clarke J, Tawfik DS: **De novo evolutionary emergence of a symmetrical protein is shaped by folding constraints**. *Cell* 2016, **164**:476–486, https://doi.org/10.1016/j.cell.2015.12.024.

19. Afanasieva E, Chaudhuri I, Martin J, *et al.*: **Structural diversity of oligomeric β-propellers with different numbers of identical blades**. *Elife* 2019, **8**:e49853, https://doi.org/10.7554/eLife.49853. Published 2019 Oct 15.

20. • Hoffarth ER, Rothchild KW, Ryan KS: **Emergence of oxygen- and pyridoxal phosphate-dependent reactions**. *FEBS J* 2020, **287**:1403–1428, https://doi.org/10.1111/febs.15277.
A recent, informative review with multiple illustrations of the powers of evolutionary tinkering in natural products biosynthesis: how old enzymes evolve towards completely new reactions.

21. Amitai G, Gupta RD, Tawfik DS: **Latent evolutionary potentials under the neutral mutational drift of an enzyme**. *HFSP J* 2007, **1**:67–78, https://doi.org/10.2976/1.2739115/10.2976/1.

22. Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C, Tawfik DS: **The 'evolvability' of promiscuous protein functions**. *Nat Genet* 2005, **37**:73–76, https://doi.org/10.1038/ng1482.

23. Näsvall J, Sun L, Roth JR, Andersson DI: **Real-time evolution of new genes by innovation, amplification, and divergence**. *Science* 2012, **338**:384–387, https://doi.org/10.1126/science.1226521.

24. Tokuriki N, Jackson CJ, Afriat-Jurnou L, Wyganowski KT, Tang R, Tawfik DS: **Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme**. *Nat Commun* 2012, **3**:1257, https://doi.org/10.1038/ncomms2246.

25. Hughes AL: **The evolution of functionally novel proteins after gene duplication**. *Proc Biol Sci* 1994, **256**:119–124, https://doi.org/10.1098/rspb.1994.0058.

26. Piatigorsky J: **Gene sharing, lens crystallins and speculations on an eye/ear evolutionary relationship**. *Integr Comp Biol* 2003, **43**:492–499, https://doi.org/10.1093/icb/43.4.492.

27. Gumulya Y, Gillam EM: **Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering**. *Biochem J* 2017, **474**:1–19, https://doi.org/10.1042/BCJ20160507.

28. Rauwerdink A, Lunzer M, Devamani T, *et al.*: **Evolution of a catalytic mechanism**. *Mol Biol Evol* 2016, **33**:971–979, https://doi.org/10.1093/molbev/msv338.

29. Khersonsky O, Tawfik DS: **Enzyme promiscuity: a mechanistic and evolutionary perspective**. *Annu Rev Biochem* 2010, **79**:471–505, https://doi.org/10.1146/annurev-biochem-030409-143718.

30. Kaltenbach M, Emond S, Hollfelder F, Tokuriki N: **Functional trade-offs in promiscuous enzymes cannot Be explained by intrinsic mutational robustness of the native activity**. *PLoS Genet* 2016, **12**:e1006305, https://doi.org/10.1371/journal.pgen.1006305. Published 2016 Oct 7.

31. Khersonsky O, Roodveldt C, Tawfik DS: **Enzyme promiscuity: evolutionary and mechanistic aspects**. *Curr Opin Chem Biol* 2006, **10**:498–508, https://doi.org/10.1016/j.cbpa.2006.08.011.

32. Copley SD: **Evolution of new enzymes by gene duplication and divergence**. *FEBS J* 2020, **287**:1262–1283, https://doi.org/10.1111/febs.15299.

33. Bommer GT, Van Schaftingen E, Veiga-da-Cunha M: **Metabolite repair enzymes control metabolic damage in glycolysis**. *Trends Biochem Sci* 2020, **45**:228–243, https://doi.org/10.1016/j.tibs.2019.07.004.

34. Niehaus TD, Hillmann KB: **Enzyme promiscuity, metabolite damage, and metabolite damage control systems of the tricarboxylic acid cycle**. *FEBS J* 2020, **287**:1343–1358, https://doi.org/10.1111/febs.15284.

35. Rockah-Shmuel L, Tawfik DS: **Evolutionary transitions to new DNA methyltransferases through target site expansion and shrinkage**. *Nucleic Acids Res* 2012, **40**:11627–11637, https://doi.org/10.1093/nar/gks944.

36. Fischbach MA, Clardy J: **One pathway, many products**. *Nat Chem Biol* 2007, **3**:353–355, https://doi.org/10.1038/nchembio0707-353.

37. Pichersky E, Raguso RA: **Why do plants produce so many terpenoid compounds?** *New Phytol* 2018, **220**:692–702, https://doi.org/10.1111/nph.14178.

38. Gould SJ, Lewontin RC: **The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme**. *Proc R Soc Lond B Biol Sci* 1979, **205**:581–598, https://doi.org/10.1098/rspb.1979.0086.

39. Waglechner N, McArthur AG, Wright GD: **Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance**. *Nat Microbiol* 2019, **4**:1862–1871, https://doi.org/10.1038/s41564-019-0531-5.

40. Joynt R, Seipke RF: **A phylogenetic and evolutionary analysis of antimycin biosynthesis**. *Microbiology* 2018, **164**:28–39, https://doi.org/10.1099/mic.0.000572.

41. Garcia-Pichel F, Lombard J, Soule T, Dunaj S, Wu SH, Wojciechowski MF: **Timing the evolutionary advent of cyanobacteria and the later great oxidation event using gene phylogenies of a sunscreen**. *mBio* 2019, **10**:e00561-19, https://doi.org/10.1128/mBio.00561-19. Published 2019 May 21.

42. • Cubillos-Ruiz A, Berta-Thompson JW, Becker JW, van der Donk WA, Chisholm SW: **Evolutionary radiation of lanthipeptides in marine cyanobacteria**. *Proc Natl Acad Sci U S A* 2017, **114**:E5424–E5433, https://doi.org/10.1073/pnas.1700990114.
The paper describes the extensive and rapid radiation of genes encoding the precursors for lanthipeptides in cyanobacteria. This unique system may be diversity-targeted (see text, and Ref [43] below).

43. Laurenceau R, Raho N, Cariani Z, Bliem C, Osman MAM, Chisholm SW: **Association of lanthipeptide genes with TnpAREP transposases in marine picocyanobacteria**. *BioRxiv* 2020, https://doi.org/10.1101/2020.03.09.984088.

44. Lynch M, Ackerman MS, Gout JF, *et al.*: **Genetic drift, selection and the evolution of the mutation rate**. *Nat Rev Genet* 2016, **17**:704–714, https://doi.org/10.1038/nrg.2016.104.