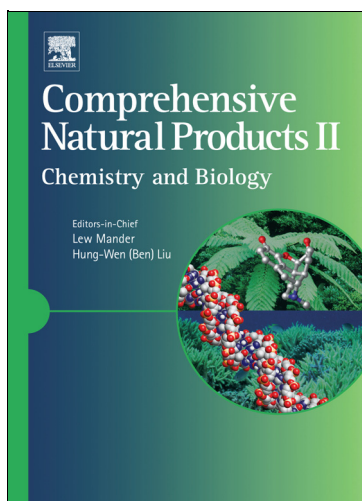


Provided for non-commercial research and educational use only.
Not for reproduction, distribution or commercial use.

This article was originally published in the *Comprehensive Natural Products II Chemistry and Biology* published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for noncommercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Olga Khersonsky and Dan S. Tawfik. In *Comprehensive Natural Products II Chemistry and Biology*; Mander, L., Liu, H.-W., Eds.; Elsevier: Oxford, 2010; volume 8, pp. 48–90.

8.03 Enzyme Promiscuity – Evolutionary and Mechanistic Aspects

Olga Khersonsky and Dan S. Tawfik, Weizmann Institute of Science, Rehovot, Israel

© 2010 Elsevier Ltd. All rights reserved.

8.03.1	Introduction	48
8.03.2	Promiscuity – The Rule or an Exception	48
8.03.3	The Definitions of Promiscuity	50
8.03.3.1	Native Function	50
8.03.3.2	Multispecificity or Broad Specificity	51
8.03.3.3	Substrate and Cofactor Ambiguity	51
8.03.3.4	Promiscuity or Catalytic Promiscuity	51
8.03.3.5	Moonlighting	51
8.03.4	Quantifying the Degree and Magnitude of Promiscuity	52
8.03.4.1	The Degree of Promiscuity	52
8.03.4.2	Assessing the Degree of Promiscuity with EC Numbers	52
8.03.4.3	The Magnitude of Promiscuity	56
8.03.5	Predicting Promiscuity	56
8.03.6	Mechanistic Aspects of Promiscuity	57
8.03.6.1	How Do Specificity and Promiscuity Coincide within the Same Active Site?	57
8.03.6.1.1	Conformational diversity	57
8.03.6.1.2	Accommodating alternative substrates	57
8.03.6.1.3	Different protonation states	59
8.03.6.1.4	Different subsites within the same active site	60
8.03.6.1.5	Promiscuity due to alternative cofactors	61
8.03.6.1.6	Water-assisted promiscuity	61
8.03.6.2	Deciphering Enzyme Mechanisms by Studying Promiscuous Functions	61
8.03.6.3	Mechanistic Origins of Differences in the Catalytic Parameters for Native Versus Promiscuous Functions	61
8.03.7	Promiscuity and the Divergence of Enzyme Superfamilies	62
8.03.8	Evolutionary Aspects of Promiscuity	66
8.03.8.1	The Evolvability of Promiscuous Enzyme Functions	67
8.03.8.1.1	Promiscuity as a starting point – the three basic assumptions	67
8.03.8.2	Promiscuous Functions Can Provide an Immediate Advantage	67
8.03.8.3	Negative Trade-offs and the Evolvability of Promiscuous Functions	72
8.03.8.4	Exceptions to Weak Negative Trade-offs	74
8.03.8.4.1	Size and charge considerations	74
8.03.8.4.2	Stability trade-offs	74
8.03.8.4.3	Targeted versus random mutagenesis	75
8.03.8.5	On- and Off-Pathway Evolutionary Intermediates	75
8.03.8.6	Promiscuity and the Mechanisms for the Divergence of New Gene Functions	79
8.03.8.6.1	Ohno's model	79
8.03.8.6.2	Gene sharing	80
8.03.8.6.3	Divergence prior to duplication	80
8.03.8.6.4	Duplication is positively selected	81
8.03.8.6.5	Subfunctionalization	81
References		84

8.03.1 Introduction

Traditionally, enzymes are referred to as remarkably fast and specific catalysts. The fact that many enzymes are capable of catalyzing other reactions, besides the one they physiologically specialize in, or evolved for, is definitely not new. Since a long time, breaches of specificity, or promiscuity, of enzymes have been recognized. Early examples of enzyme promiscuity include pyruvate decarboxylase,¹ carbonic anhydrase,² pepsin,³ chymotrypsin,⁴ and L-asparaginase.⁵ However, these early discussions of enzymatic versatility were scarce, and until recently, the wider implications of this 'darker' side of enzymes were largely ignored. During the last decade, protein, and especially enzyme promiscuity, received considerable attention, and their importance in various contexts was systematically studied. Reviews by O'Brien and Herschlag,⁶ and later Copley,⁷ were the first to highlight the mechanistic and evolutionary implications of promiscuity. Other, more recent reviews have focused on the practical implications of promiscuity in organic synthesis,^{1,8,9} on promiscuity and divergence in certain enzyme families,^{10–13} on mechanistic aspects of promiscuity,^{1,14} and on promiscuity in the context of protein evolution¹⁴ and design.¹⁵

The primary focus of this chapter is the role of promiscuity in the evolution of new enzyme functions. New enzymes have constantly emerged throughout the natural history of this planet. Over the past decades, enzymes that degrade synthetic chemicals were introduced to the biosystem,^{16–20} and enzymes associated with drug resistance,^{21–24} provide vivid examples of how rapid the evolution of new enzymatic functions can be. The first direct connection between protein evolution and promiscuity was made in 1976 by Jensen.²⁵ In his landmark review, Jensen formalized the hypothesis that the starting points for evolution were provided by broad specificity, or promiscuity, of the ancestral enzymes. Jensen proposed that unlike modern enzymes that tend to specialize in one substrate and reaction, the primordial, ancient enzymes possessed very broad specificities, and thus few enzymes could perform many functions. Divergence of specialized enzymes, through duplication, mutation, and selection, led to the current diversity of enzymes, and to increased metabolic efficiency.

Extensive research has been carried out since Jensen provided ample evidence for the idea that promiscuity is a key factor in the evolution of new protein functions. Here, we attempt to survey this accumulated knowledge. We focus on several aspects of promiscuity (and enzyme promiscuity, in particular) with an emphasis on its mechanistic aspects, and its role in enzyme evolution. In contrast to journal articles, this volume allows a very detailed discussion, including numerous examples and sidetracks to several related issues. We therefore provide a comprehensive treatise that begins with the question of generality – are promiscuous activities a rare exception, or should they be considered an accompanying trait of all enzymes (Section 8.03.2)? We continue with an attempt to define promiscuity, and its variable facets, in a more rigorous way, and quantify its levels and magnitude (Sections 8.03.3 and 8.03.4). Following that at present, promiscuity is a widely studied phenomenon, we also discuss the prospects of predicting it (Section 8.03.5). In the next sections, we discuss the structural and mechanistic aspects of promiscuity (Section 8.03.6), and the evolutionary implications (Sections 8.03.7 and 8.03.8). The latter begins with reviewing evidence for the role of promiscuity in the divergent evolution of enzyme families and superfamilies (Section 8.03.7), and continues with a broader discussion of various factors, and mechanisms, that drive the evolution of new enzymatic functions (Section 8.03.8).

8.03.2 Promiscuity – The Rule or an Exception

How wide is the phenomenon of promiscuity? The paradigm of absolute specificity: 'one enzyme = one substrate' dominates our textbooks. But this view is obviously schematic, and is mostly valid as a first approximation. Numerous examples for enzyme promiscuity are currently known, but one might argue that these are largely anecdotal and do not provide a general picture. The paragraphs that follow detail several arguments in favor of the notion that promiscuity, at different levels and magnitudes, is a wide phenomenon, and should thus be treated as a rule, rather than an exception.

Specificity bears a high cost, in substrate binding energies, and hence in k_{cat} .²⁶ Even the most specific enzymes, for example, enzymes involved in DNA replication or protein synthesis, exhibit measurable substrate infidelities, often at surprisingly high rates.²⁶ When necessary, high fidelity is achieved through proof-reading mechanisms that involve energy costs, and essentially reverse the process, and repeat it, to correct possible errors. For example, the proof-reading domain of many polymerases is an exonuclease that digests parts of the extending strand.

Specificity is shaped by natural selection. It is therefore context dependent. Cross-reactivities and promiscuous activities that are harmful were selected against. Consider aminoacyl-tRNA synthetases for example. Undoubtedly, their selectivity with respect to tRNAs and amino acids is under tight selection. Because of the close similarity of certain amino acids, proof-editing mechanisms have evolved whereby formation of a noncognate aminoacyl-tRNA is followed by its rapid hydrolysis.²⁷ Isoleucyl-tRNA synthetase, for example, favors the reaction with isoleucine over valine only by ~ 100 -fold, and valine concentration *in vivo* is fivefold higher than that of isoleucine. However, the proof-editing mechanism, which involves the rapid hydrolysis and removal of mis-incorporated valine-tRNA, makes the error rate decrease to 1 in 3000.²⁶ Occasionally, as is the case of D-tyrosine, proofreading is performed by another enzyme that hydrolyzes the misacylated tRNA.²⁸ Nonetheless, unlike the case of D-tyrosine, selectivity with respect to unnatural amino acids that have never been present in living cells is very low. For example, the unnatural analogue 4-hydroxy-phenylalanine is bound in the same mode as the native phenylalanine substrate.²⁹

Many enzymes perform secondary tasks,^{30,31} and some of the more illuminating examples include enzymes that have been under intense selection for high specificity, such as aminoacyl-tRNA synthetases. These enzymes exhibit secondary functions, for example, the biosynthesis of signaling molecules such as Ap4A (2 adenosines linked through 4 phosphates) is mediated by lysyl-tRNA synthetases.^{32,33} In addition, certain aminoacyl-tRNA synthetases bind DNA or mRNA and thus regulate transcription, splicing, and translation, or act as cofactors in RNA trafficking.³⁴ It is probable that these functions were recruited well after the primary function had emerged (loading of a specific tRNA with the cognate amino acid), from promiscuous, or side products, of the primary reactions (see also Section 8.03.8.6.2 on gene sharing). Once recruited, such functions remained under selection, and therefore became a native function of the enzyme (Section 8.03.3.1).

High-throughput screens for binding cross-reactivities revealed a very clear pattern, whereby the number of identified cross-reactants (e.g., small molecules or proteins), and their affinity, increase exponentially with the number of tested ligands, or binding sites (e.g., number of different antibodies).^{35–37} Several theoretical models account for these observations^{38–40} (for further discussion of these aspects see Griffiths and Tawfik⁴¹ and James and Tawfik⁴²). A screen using a sufficiently large diversity of substrates (and reactions that can be performed on these substrates) is probable to reveal that every enzyme exhibits a pattern of promiscuous functions. Some of these may be related to the enzyme's native function, and others might have a weak relation. Unfortunately, contrary to binding, exhaustive screens for catalytic promiscuity are technically challenging, not the least because different detection methods need to be applied for different substrates and reactions.

Few, and often none, of the promiscuous activities found *in vitro* (or those that could be found by systematic screens) bear a physiological or evolutionary meaning. But even those that might, are not necessarily relevant *in vivo*. The main reason is regulation. Of the entire enzyme diversity available to organisms, only a small fraction is accessible, and active, at a given time and cellular location. In this way, many of the undesirable outcomes of enzyme promiscuity are prevented. Regulation at the level of expression is obviously directed to prevent the spending of unnecessary resources.^{34,43,44} Different regulation regimes are the key to control enzyme activity, especially with enzymes whose specificity is broad. For example, *E. coli* has 23 different members of the HAD superfamily (haloacid dehalogenase (HAD)-like hydrolases). Most of these enzymes are phosphatases that show remarkably broad and overlapping substrate spectra.^{45,46} However, these paralogs that are biochemically similar operate under different regulation schemes, and specificity is achieved through regulation and not by the more familiar manner of controlling enzyme reactivity.⁴⁶

Regulation can also occur at the protein level, such as allosteric regulation that prevents the wasteful conversion of costly metabolites. Such regulation is expected to be, and in many cases is, product-controlled. However, in many cases, the substrate also comprises of an allosteric regulator of its own enzyme – in other words, in the absence of its substrate, the enzyme active site becomes inactive. Preventing the active sites from promiscuously reacting with other, undesirable substrates could be one of the driving forces for the evolution of such mechanism.

There exists evidence indicating that despite the action of natural selection to increase enzyme selectivity by various means, ranging from shaping the active site itself to regulation of enzyme expression and activity, numerous cross-reactions and breaches of specificity occur, not just *in vitro*, but in particular within living cells. Such cross-reactivities are often unraveled by the analysis of auxotrophic knockout strains that lack a crucial enzyme. These deficiencies are often complemented by other enzymes, or even other enzyme pathways,

sometimes in an unexpected manner. For example, knockouts of the *phn* operon in *E. coli* that is known to utilize phosphite (HPO_3^{2-}) led to the identification of promiscuous phosphite-dependent hydrogenase activity in alkaline phosphatase (see **Figure 2(a)** in Section 8.03.6.1.2). This activity enables the growth of the *phn* knockout strain with phosphite as the sole phosphorous source.⁴⁷ Other examples of ‘metabolic plasticity’, or ‘underground metabolism’ were reviewed by Jensen,^{10,25} and later by D’Ari and Casadesus.⁴⁸

The fact that cross-reactivities between different metabolic pathways are highly expected was also indicated by an *in silico* experiment that attempted to dock 125 common metabolites into the active sites of 120 key metabolic enzymes. Numerous potential cross-reactions were found amongst these 15 000 potential pairs, cross-reactions that were often stronger than the cognate interactions. Although docking has obvious limitations, this study further highlights the potential for promiscuity, and underground metabolism, within numerous metabolic pathways.⁴⁹ A systematic survey of complementation of deficient *E. coli* strains, by selection from a library of *E. coli*'s own genes under overexpression, revealed a similar picture.⁵⁰ The deleted gene and its suppressor were in most cases unrelated. Complementation was achieved through the promiscuous action of another enzyme, through increased transport (and not necessarily of the deficient metabolite), and most often, through the opening of an alternative metabolic pathway and/or more global changes through the overexpression of a regulatory factor. Thus, promiscuity is not necessarily a phenomenon limited to the single enzyme level, but often whole pathways can act promiscuously, namely, outside their routine functional scope.

The above observations and considerations led to new hypotheses which suggest that genetic and metabolic pathways are inherently probabilistic and ambiguous. By these hypotheses, the well-studied linear pathways described in textbooks can be cross-wired in a variety of unexpected ways. Evolution may capitalize on these unexpected cross-wirings, in a way of adaptive plasticity, to generate new metabolic capabilities.⁵¹ Phenomena similar to ‘underground metabolism’ were also observed in genetic analyses where the observed phenotypes turned out to be correlated with changes in many different genes, including genes from unrelated pathways. These studies demonstrate a remarkable flexibility in the way genomes respond to changes. As is the case with enzymes and metabolic pathways, genome flexibility is the outcome of the limited specificity, or promiscuity, of gene action and of intergenic interactions.^{52,53} Thus, it appears that, beyond the linear, well-defined pathways which are extensively studied, there exist ‘flexible genomes’,⁵² as well as ‘flexible proteomes’, and ‘flexible metabolomes’, the contribution of which to evolutionary adaptation requires further study.

8.03.3 The Definitions of Promiscuity

The term ‘enzyme promiscuity’ was imprinted through a review by O’Brien and Herschlag⁶ that highlighted a topic that, at the time, was relatively ignored by enzymologists. However, the term promiscuity is not very well defined and has been used to describe a wide range of fundamentally different phenomena that are not necessarily related. We thus propose the following terminology, or glossary:

8.03.3.1 Native Function

The substrate, and chemical transformation, for which an enzyme has evolved, and that is relevant to physiology of the organism in which this enzyme resides. By definition, the native, or primary function is maintained under selection, and mutations that harm it may affect organismal fitness. Many enzymes have more than one physiological function,³¹ and in this case, all these functions should be defined as native, even though some of these were obviously recruited at much later evolutionary stage. The ‘primary function’ describes the function that underlines these enzymes. For example, the primary function of aminoacyl-tRNA synthetases is activating and loading amino acids onto the cognate tRNA, whereas the generation of Ap4A is clearly ‘a secondary function’ executed only in certain organisms and under specific circumstances.³² The ‘original, or existing, function’, are complementary terms that refer to the native, or primary, function in the context of the divergence of new functions.

8.03.3.2 Multispecificity or Broad Specificity

Many enzymes evolved to perform a certain reaction on a whole range of similar substrates, rather than on a single substrate. Their broad specificity is therefore an inherent, evolved function, and they should be regarded as multispecific enzymes, and not as promiscuous enzymes. Enzymes known for their broad substrate specificity include mammalian detoxifying enzymes, such as glutathione *S*-transferases (GSTs) and cytochrome P450s,^{54–58} certain enzymes in terpenoid biosynthesis such as sesquiterpene cyclases,^{59,60} and methane monooxygenase that hydroxylates over 150 various substrates.⁶¹ Other examples include exonucleases that, contrary to restriction endonucleases, cleave double-stranded, and often also single-stranded DNA, with almost no sequence specificity. In cases of multispecificity, the reactions with the various substrates are expected to exhibit similar kinetic parameters (K_M , k_{cat}).

8.03.3.3 Substrate and Cofactor Ambiguity

This definition applies to enzymes that evolved to transform one well-defined substrate. The enzyme's activity with other substrates is purely accidental, or promiscuous. However, in cases where the structure and chemical nature of the alternative promiscuous substrates resemble the native substrate, this breach of specificity is best defined as substrate ambiguity. Common examples of substrate ambiguity include lipases,^{62,63} whose natural substrates are glyceryl esters of long-chain carboxylic acids (fatty acids), and that catalyze the hydrolysis of a broad range of esters. More recent examples include asparagine synthetase that can complement a deficiency in glutamine synthetase.⁵⁰ Another type of substrate ambiguity is when an enzyme uses various nucleophiles to react with the same substrate, as in case of halohydrin dehalogenase. Normally, this enzyme uses a water molecule to open an epoxide ring and form a diol product. However, a whole range of monovalent anions, such as Br^- , Cl^- , CN^- , and N_3^- , can be promiscuously applied by the enzyme to generate a broad range of products.⁶⁴ The related category of 'cofactor ambiguity' includes enzymes that can utilize coenzymes, or cofactors, other than the one they have evolved with. For example, D-2-hydroxyacid dehydrogenase from *H. mediterranei* can use both NADH and NADPH as cofactors.⁶⁵ Another example is regarding metallo-enzymes that have been shown to catalyze promiscuous reactions when the naturally occurring active-site metal is replaced (Section 8.03.6.1.5).

8.03.3.4 Promiscuity or Catalytic Promiscuity

This category refers to enzymes that catalyze different reactions (and not just different substrates) than the one they evolved for. As is the case with substrate and coenzyme ambiguity, the enzyme's activity with these alternative substrates is purely accidental, and is under no selection, and is therefore promiscuous by definition. As suggested,⁸ these cases include chemical transformations where the bonds that are broken, or formed, are different than those in the native substrate and reaction, and/or transformations that proceed through a different transition state. As discussed later, the promiscuous chemical transformations can be performed by the same catalytic side chains, and by essentially the same mechanism, as the native enzymatic function (Section 8.03.6). But there are also cases in which the enzyme utilizes different subsets of active-site residues, and somewhat different mechanisms, for the native and promiscuous functions (Section 8.03.6.1.4).

8.03.3.5 Moonlighting

In contrast to promiscuity that occurs within the same active site as the primary, native function, moonlighting relates to the utilization of protein parts outside the active site for other functions, mostly regulatory and structural,^{7,66} but sometimes enzymatic ones.^{67,68} Such activities can be recruited at later evolutionary stages, as indicated by the classical example of crystallins whereby metabolic enzymes were recruited later in evolution as structural components of eye lenses^{30,69} (see Section 8.03.8.6.2 on gene sharing).

8.03.4 Quantifying the Degree and Magnitude of Promiscuity

A more refined definition of promiscuity should include a quantitative measure for the degree and magnitude of promiscuity. The *degree of promiscuity* refers to the level to which enzyme specificity is breached (or the degree of multispecificity for enzymes that react with a broad range of different substrates). Namely, how different are the native and promiscuous functions. The *magnitude of promiscuity* refers to the kinetic parameters for the promiscuous activity relative to the native one.

8.03.4.1 The Degree of Promiscuity

Previous work suggested that the degree of promiscuity can be qualitatively assessed along two dimensions whereby, one dimension describes differences in the type of bonds that are being formed or broken, and the other dimension describes differences in the mechanism of the catalyzed reaction.⁸ A more recent work aimed at a quantitative measure dubbed 'index of promiscuity' that is calculated by mapping the substrate structure and quantifying the degree of variability between different substrates of the same enzyme.⁵⁸ However, this method does not take into account substrate chemistry, and assumes that the same chemical transformation occurs on all substrates. As such, it is more suitable for the quantitative analysis of cross-reactivity with multispecific enzymes such as GSTs (as originally demonstrated), and largely inapplicable to quantify catalytic promiscuity whereby, as defined in Section 8.03.3.4, the major differences pertain to the chemistry and not the substrate structure.

8.03.4.2 Assessing the Degree of Promiscuity with EC Numbers

Here we propose a simple, and relatively objective, way of assessing the degree of promiscuity using a comparison of Enzyme Commission numbers (EC) for the native and promiscuous activities. In enzymes exhibiting multispecificity, or substrate ambiguity, EC numbers for the various substrates should be the same, or differ only by the 4th digit that generally distinguishes between enzymes of the same class. Catalytic promiscuity should refer to cases in which the EC numbers of the various substrates and reactions catalyzed by the same enzyme differ in the 2nd, or the 3rd, digits that refer to different chemistries, and different classes of substrates, or even by the 1st digit that indicates a completely different reaction category.

Table 1 lists several examples for substrate ambiguity and catalytic promiscuity. The table indicates the EC numbers for the native and promiscuous activities, and thus categorizes their promiscuity accordingly to differences in EC numbers. Typical examples where differences in EC numbers reflect the degree of promiscuity, and cases where they might not are discussed below.

Almost all cases of substrate ambiguity and multispecificity (as defined in Section 8.03.3) are manifested in differences in the 4th digit. Examples for multispecific, or broad-specificity enzymes include sulfotransferases and GSTs (**Table 1**, entries 1, 2).

Examples for substrate ambiguity include enzymes such as sugar kinases, amino acid transferases, glycosidases, and methyltransferases, which can perform the same chemical transformation on substrates other than their native one (**Table 1**, entries 3–6).

Lipases (EC 3.1.1.3; **Table 1**, entry 7) comprise a clear example where EC numbers seem to reflect differences in the degree of promiscuity for a whole range of promiscuous activities. Their native substrates are triglycerides (EC 3.1.1.3), and their ability to promiscuously hydrolyze aryl esters of various carboxylic acids is manifested by a difference in the 4th digit only (3.1.1.X) and thus is correctly defined as substrate ambiguity. The promiscuous amide hydrolysis concerns the cleavage of a different bond (C–N vs C–O) and is manifested by differences in the 2nd digit (EC 3.4.X.X). Lipases were also shown to promiscuously catalyze nonhydrolytic reactions such as aldol condensations and Michael additions,^{78–80} and these belong to the 4th EC category (EC 4.X.X.X). In these cases, both the mechanisms and the bonds that are being formed or broken differ, and these differences are manifested in the 1st digit.

Alkaline phosphatase (EC 3.1.3.1; **Table 1**, entry 8) also possesses a wide range of promiscuous activities. Some of them differ from the native phosphate monoesters hydrolysis in the 3rd digit (sulfatase,

Table 1 Examples for classifying the degree of promiscuity based on differences in EC numbers

<i>Enzyme</i>	<i>Native activity (EC number)</i>	<i>Promiscuous activity (EC number)</i>	<i>Promiscuity type</i>	<i>Reference(s)</i>
1 Human cytosolic sulfo-transferases (hSULTs)	Broad-specificity enzymes, sulfonate group transfer from PAPS to the various substrates EC 2.8.2.X.	Sulfonate group transfer from PAPS to the various substrates EC 2.8.2.X.	Multispecificity	70
2 Glutathione transferases	Broad substrate specificity GST of A-class 2.5.1.18.	Glutathione coupling with various ligands EC 2.5.1.18.	Multispecificity	57
3 <i>N</i> -acetyl-D-mannosamine kinase (NanK)	Phosphorylation of <i>N</i> -acetyl-D-mannosamine EC 2.7.1.60	Phosphorylation of glucose EC 2.7.1.2.	Substrate ambiguity	71, 72
Fructose kinase (YajK)	Phosphorylation of fructose EC 2.7.1.4	Phosphorylation of glucose EC 2.7.1.2.		
Allose kinase (Alsk)	Phosphorylation of allose EC 2.7.1.55	Phosphorylation of glucose EC 2.7.1.2.		
4 Aspartate amino-transferase (AATase)	Transamination of dicarboxylic substrates EC 2.6.1.1.	Transamination of tyrosine and phenylalanine EC 2.6.1.X.	Substrate ambiguity)	73
5 Beta-glucuronidase (GUS)	Hydrolysis of beta-glucuronides EC 3.2.1.31	Hydrolysis of pNP-galactoside EC 3.2.1.23	Substrate ambiguity	74, 75
6 <i>HaeIII</i> methyltransferase	Methylation of GGCC sites EC 2.1.1.X.	Methylation of AGCC sites EC 2.1.1.X.	Substrate ambiguity	76
7 Lipases	Triglyceride hydrolysis EC 3.1.1.3.	Hydrolysis of aryl esters of various carboxylic acids 3.1.1.X.	Substrate ambiguity	62, 63
		Amide bond hydrolysis EC 3.4.X.X.	Catalytic promiscuity	77
		Aldol reaction (C–C bond formation) EC 4.1.X.X.	Catalytic promiscuity	78
		Michael-type additions EC 2.5.1.18. (as in GST) EC 4.4.X.X. (as in lyases)	Catalytic promiscuity	79, 80
		Oligomerization of siloxanes	Catalytic promiscuity	81
8 Alkaline phosphatase	Hydrolysis of monophosphate esters EC 3.1.3.1.	Unnatural reaction for which no native enzyme in known		
		Phosphodiesterase EC 3.1.4.X.	Catalytic promiscuity	82
		Sulfate esters hydrolysis EC 3.1.6.X.	Catalytic promiscuity	83, 84
		Phosphite-dependent dehydrogenase EC 1.1/2.X.X	Catalytic promiscuity	47
9 Muconate lactonizing enzyme (MLE)	Cycloisomerization EC 5.5.1.1.	OSBS (β -elimination) EC 4.2.1.113	Catalytic promiscuity	85

(Continued)

Table 1 (Continued)

<i>Enzyme</i>	<i>Native activity (EC number)</i>	<i>Promiscuous activity (EC number)</i>	<i>Promiscuity type</i>	<i>Reference(s)</i>
10 Phosphotri-esterase from <i>P. diminuta</i> (PTE)	Phosphotriester hydrolysis EC 3.1.8.1.	Phosphodiesterase EC 3.1.4.X. Esterase EC 3.1.1.X. Lactonase EC 3.1.1.X.	Catalytic promiscuity	19, 86, 87
11 PLLs (PTE-like lactonases)	Hydrolysis of quorum sensing lactones EC 3.1.1.X.	Phosphotriester hydrolysis EC 3.1.8.1.	Catalytic promiscuity	88
12 <i>AiiA</i> from <i>B. thuringiensis</i>	Hydrolysis of quorum sensing lactones EC 3.1.1.X.	Phosphotriester hydrolysis EC 3.1.8.1	Catalytic promiscuity	88, 89; H.-S. Kim, personal communication
13 PONs (serum paraoxonases)	Mammalian lactonases EC 3.1.1.X.	Phosphotriesterase EC 3.1.8.X.	Catalytic promiscuity	90–92
14 Serum albumins	Nonenzymatic proteins	Aryl esterase EC 3.1.1.X. Esterase EC 3.1.1.X. Carbamate hydrolysis EC 3.1.1.X. Kemp elimination (unnatural reaction, for which no native enzyme is known).	Substrate ambiguity Catalytic promiscuity Catalytic promiscuity Catalytic promiscuity	93 94 95
15 Carbonic anhydrase	Hydration of CO ₂ EC 4.2.1.1.	Esterase EC 3.1.1.X. Epoxide synthase (styrene epoxidation, by metal exchange) EC 1.14.X.X.	Substrate ambiguity Catalytic promiscuity, or cofactor ambiguity	2, 96 97

phosphodiesterase), but others (phosphite dehydrogenation) differ in the 1st digit, and represent a higher degree of promiscuity.

Other cases in which the native and promiscuous activities differ in the 1st digit include muconate lactonizing enzyme (MLE; **Table 1**, entry 9), whose native activity is cycloisomerization (EC 5.5.1.1) and possesses a promiscuous OSBS (β -elimination) activity (EC 4.2.1.113).

Examples for a high degree of promiscuity obviously include nonenzymatic proteins such as serum albumins that exhibit promiscuous catalytic activities^{93–95} (**Table 1**, entry 14). Other cases may include catalysis of unnatural reactions, meaning reactions for which, to our knowledge, no natural enzyme has evolved (e.g., the Kemp elimination performed by serum albumin (**Table 1**, entry 14, or siloxane hydrolysis by lipase (**Table 1**, entry 7)).

Another example where EC numbers seem to reflect differences in the degree of promiscuity concerns lactonases. These enzymes that belong to three different superfamilies have all been shown to exhibit promiscuous phosphotriesterase (PTE) activity (**Table 1**, entries 11–13). Lactonase activity involves the hydrolytic cleavage of a C–O bond and is described as EC 3.1.1.X (where X refers to a specific lactone substrate). The hydrolytic cleavage of the P–O bond of phosphotriesters is described by EC number 3.1.8.X. The difference in the 3rd digit reflects the difference in the bond that is being broken (C–O versus P–O), while applying essentially the same mechanism. For comparison, many of these lactonases (but not all of them) also exhibit esterase activity with aryl esters in particular.⁸⁸ This activity is described as EC 3.1.1.X, as is the case with lactonase activity. Because both reactions involve cleavage of C–O bonds, and the lactonase–esterase differences are manifested in the 4th digit, this case is better described as substrate ambiguity.

It should be noted, however, that EC numbers can also be misleading. There are notable cases where, despite considerable similarity in the chemistry of catalysis, the EC numbers differ, and even by the 1st digit. This is primarily because the EC definitions relate not only to the chemistry, but also to the physiological context. A clear example is carbonic anhydrase (**Table 1**, entry 15; EC 4.2.1.1) that exhibits promiscuous aryl esterase activity (3.1.1.X). The EC numbers suggest a totally different chemistry. But although the substrates differ significantly, primarily in size, both reactions involve the attack of a hydroxide ion on a carbonyl (**Figure 1**). The phosphite-dependent hydrogenase activity (EC 1.1/2.X.X) of alkaline phosphatase (**Table 1**, entry 8; EC 3.1.3.1) also represents a borderline case. The very different native and promiscuous activities (as manifested in the different EC categories) actually utilize a similar mechanism (see Section 8.03.6.1.2).

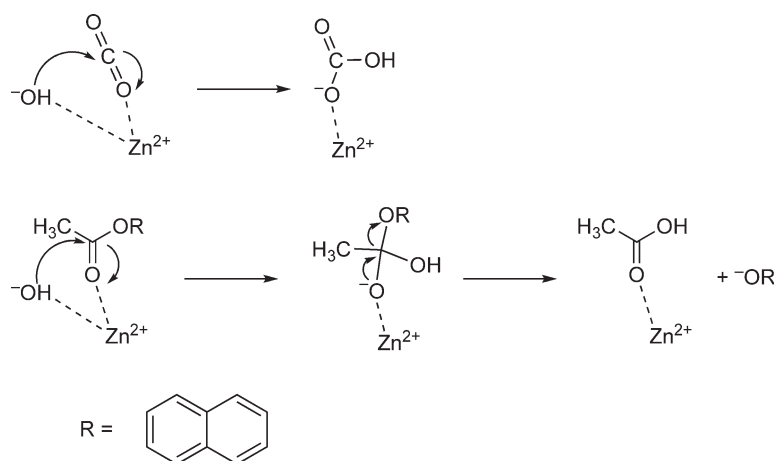


Figure 1 The native reaction of carbonic anhydrase (CO_2 hydration, *top*) and its promiscuous aryl esterase reaction, exemplified with naphthyl acetate (*bottom*). Both reactions proceed by the same mechanism of hydroxide ion attack on a carbonyl, followed by the stabilization of an oxyanion intermediate by the active-site Zn^{2+} . Despite this obvious similarity, the EC numbers of these reactions differ in the first digit (**Table 1**, entry 15).

8.03.4.3 The Magnitude of Promiscuity

Another quantitative measure addresses the magnitude of promiscuous activities, that is, how do the kinetic parameters for the promiscuous reaction/substrate compare with those for the native substrate. Whereas most enzymes exhibit $k_{\text{cat}}/K_{\text{M}}$ values in the order of 10^5 – 10^8 mol⁻¹ s⁻¹ for their native substrates,²⁶ the magnitude of promiscuous activities varies over more orders-of-magnitude, both in absolute terms, and relative to the native activity. In many cases, the promiscuous activities are relatively high and fall within just an order-of-magnitude, or two, from the native function (e.g., **Table 2**, entries 1, 2, 3, 5, 6, 9). Such activities can obviously provide a distinct and immediate selective advantage. These cases may resemble ‘generalist’ intermediates (see Section 8.03.8.3) and their divergence may proceed through a ‘gene sharing’ mechanism, as described in Sections 8.03.8.6.2 and 8.03.8.6.3. In many cases, however, the promiscuous activities are very low, or even barely detectable. Weak promiscuous functions can still provide a selective advantage, certainly under high expression levels.^{47,98,99} Gene duplication is another way by which enzyme levels could be increased, and thus endow a weak promiscuous activity a selective advantage (Section 8.03.8.6.4).

Furthermore, a useful way of assessing the magnitude of promiscuous activities is the rate acceleration ($k_{\text{cat}}/k_{\text{uncat}}$) or catalytic proficiency ($k_{\text{cat}}/K_{\text{M}}/k_{\text{uncat}}$). These parameters are indicative because they take into account the inherent reactivity of the substrate.⁹ In many cases, promiscuous activities occur, or are measured, with highly reactive substrates. Such activities are in a way expected.¹⁰⁰ However, there are many cases in which promiscuous activities take place with substrates that are less activated than the native one. Examples include, the amidase activity of esterases such as lipases (**Table 1**, entry 7), the phosphodiesterase activities of *P. diminuta* PTE and alkaline phosphatase (**Table 1**, entries 10 and 8), and the PTE activities or various lactonases (**Table 1**, entries 11–13; and the notable fact that some of these lactonases do not hydrolyze the more activated aryl esters). In such cases, the chemical challenge posed by a less activated substrate is reflected in the more favorable comparisons of rate accelerations, or catalytic proficiencies, for the native versus the promiscuous substrates.

8.03.5 Predicting Promiscuity

The growing interest in promiscuity, from both a fundamental point of view, and an applicative one (applications of enzymes in organic synthesis), provides an incentive for the development of computational and bioinformatic tools for its prediction. By default, promiscuity is a phenomenon which is unpredictable – and the more interesting and unrelated is the promiscuous function, the harder would be its prediction. The difficulty of prediction is further augmented by the fact that in many cases, promiscuity involves some degree of structural plasticity, namely, when the native and promiscuous functions are mediated by different active-site conformations (see Section 8.03.6.1.1).

Early attempts to assess the frequency and potential for promiscuous enzyme–substrate encounters were made by docking a set of substrates into a set of enzymes and examining the number and distribution of the resulting virtual substrate–enzyme matches.⁴⁹ This study, however, aimed at evaluating the potential for cross-reactivities and had intention, or computational accuracy, only to identify individual cases of promiscuity. Significant improvements in structural predictions and ligand docking, and in particular, successful attempts to dock transition states rather than substrates,¹⁰¹ may enable such studies to be performed with much higher accuracy and scope.

The hope that promiscuity is predictable is also supported by the identification of systematic patterns of promiscuity. For example, lactonases, and in particular lactonases that favor hydrophobic lactones, show a consistent tendency to promiscuously catalyze the hydrolysis of phosphotriesters. This pattern has now been seen in lactonases from three different superfamilies:⁷⁶ PLLs (TIM-barrels from the amidohydrolase superfamily; **Table 1**, entry 11); PONs, or serum paraoxonases (calcium-dependent six-bladed β -propellers; **Table 1**, entry 13); and *AiiA* (a lactonase from the metallo- β -lactamase superfamily; **Table 1**, entry 12). That very different scaffolds and active-sites configurations share the same promiscuity pattern suggests that these reactions share a key feature, probably in the geometry of their transition states. This feature must be distinct, also because many of these lactonases do not hydrolyze esters that are much closer to lactones than phosphotriesters, and should thus be amenable to structural analysis and prediction.

The ongoing enrichment of structure and sequence databases, and the development of novel computational and bioinformatic tools, should also facilitate the prediction of promiscuous functions. In particular, as discussed in Section 8.03.7, the observation that within highly diverse superfamilies, the native function of one enzyme family often comprises of a promiscuous function of another family, and vice versa, also suggests that certain promiscuous functions are predictable.

8.03.6 Mechanistic Aspects of Promiscuity

8.03.6.1 How Do Specificity and Promiscuity Coincide within the Same Active Site?

A frequently asked question is how the very same active-site and catalytic machinery can show exquisite specificity with respect to the native substrate (and thus avoid catalyzing other, closely related substrates), but still catalyze other, and often completely unrelated, functions in a promiscuous manner. The answer to this question is somewhat complex, because different scenarios, or mechanisms, seem to account for the coexistence of specificity and promiscuity within the very same active site. Several scenarios are outlined in the next section, and typical examples are described.

8.03.6.1.1 Conformational diversity

Despite the rigid, lock-and-key images of active sites obtained by X-ray crystallography, active sites, and even their protein scaffolds, are remarkably flexible. The role of structural plasticity in facilitating enzyme action, and evolution, has been discussed in several reviews.^{102–104} In many cases promiscuity is linked to conformational diversity, whereby the native and the promiscuous functions are mediated by different active-site configurations.

A notable example of the role of conformational plasticity comes from α -lytic protease, where a single amino acid substitution increased the activity toward promiscuous substrates by a factor of 10^5 , whereas the native activity was reduced by only twofold.¹⁰⁵ This large shift in the selectivity of this enzyme and its other family members is allowed by the structural flexibility of the substrate binding loops.^{106,107} The mobility of active-site loops was also demonstrated to play a key role in mediating promiscuity in isopropylmalate isomerase, an enzyme with dual substrate specificity, where a loop structure is dependent on the substrate presence.¹⁰⁸ Other examples include human sulfotransferase SULT1A1, where conformational changes enable the same enzyme to accommodate different substrates.¹⁰² The importance of conformational plasticity in mediating functional promiscuity was also demonstrated in a study about glutathione-S-transferase isoforms GSTA1-1 and GSTA4-4.⁵⁷ In an evolved aminoacyl-tRNA synthetase, the disruption of an α -helix, introduced structural plasticity to the enzyme's active site and thus enabled it to accept a relatively broad range of unnatural amino acid substrates.¹⁰⁹

It should be noted that cases of different active-site configurations can include substantial backbone rearrangements, in some cases, and only different side-chain rotamers in others. In principle, however, in all these cases promiscuity makes use of a structurally modified active site. That alternative active-site conformations can mediate alternative promiscuous functions also has interesting implications regarding the evolvability of promiscuous functions.^{103,104} (see Section 8.03.8.1).

8.03.6.1.2 Accommodating alternative substrates

In many cases, promiscuous activities share the main active-site features (and often the same active-site configuration) with the native activity, and besides differences such as substrate positioning, their mechanism is largely the same. For example, in the family of guanidine-transferring enzymes, three mutually inter-promiscuous enzymes PaADI, PaAgDI, and PaDDAH utilize the same catalytic triad (Cys-His-Asp), in their action on various derivatives of arginine.¹¹⁰ However, in this case of substrate ambiguity, the active-site residues that bind the C α -carboxyl and the guanidino-NH₂ of these different substrates are different.

Other examples include cases in which the enzyme applies nucleophilic catalysis, and the same active-site nucleophile is utilized in both the promiscuous and the native function. For example, alkaline phosphatase is a highly proficient ($k_{\text{cat}}/K_{\text{M}} > 10^7 \text{ mol}^{-1} \text{ s}^{-1}$) phosphate monoesterase that promiscuously hydrolyzes phosphodiester, phosphoamides, and sulfate esters,^{82–84} as well as phosphite (while actually reducing water to release hydrogen; **Figure 2(a)**).⁴⁷ The catalytic mechanism is presumed to be similar for all these reactions, and involves nucleophilic attack by Ser102, and stabilization of the negatively charged intermediate by the

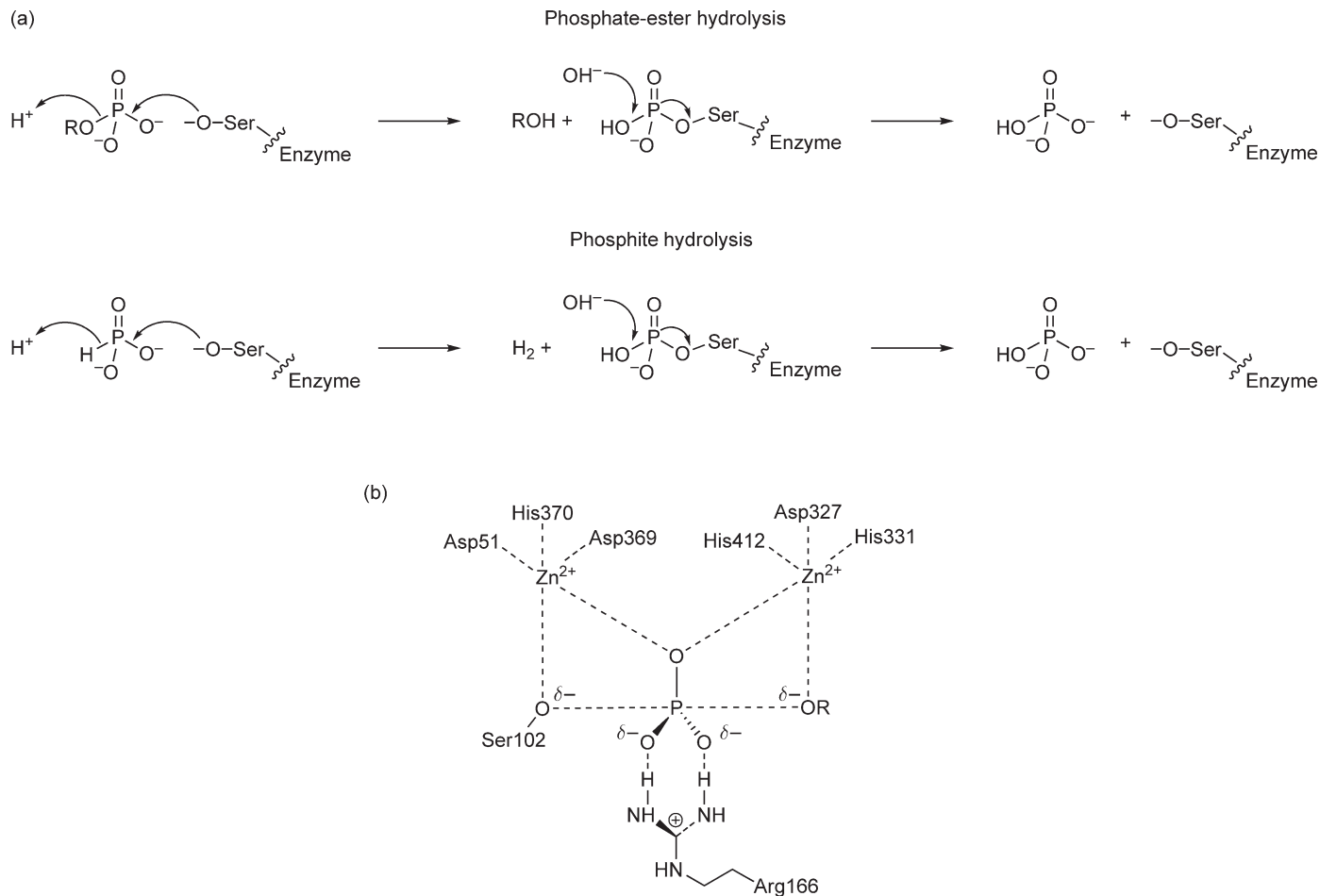


Figure 2 (a) The native monoester phosphatase activity, and the promiscuous phosphite oxidation reactions catalyzed by alkaline phosphatase. Adapted from K. Yang; W. W. Metcalf, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7919–7924. (b) The active site arrangement of alkaline phosphatase with a bound transition state model. Adapted from I. Catrina; P. J. O'Brien; J. Purcell; I. Nikolic-Hughes; J. G. Zalatan; A. C. Hengge; D. Herschlag, *J. Am. Chem. Soc.* **2007**, *129*, 5760–5765.

active-site Zn^{2+} ions and Arg166^{47,82} (Figure 2(b)). Comparison between the phosphate monoesterase, phosphodiesterase, and sulfatase activities of alkaline phosphatase revealed that, whereas these substrates all bind in a similar mode, the interactions with both Zn^{2+} ions and Arg166 are much more favorable for the native phosphate monoester substrates than for other, promiscuous substrates.^{82,83} This difference accounts for the orders-of-magnitude higher rates and catalytic proficiencies of the native substrates versus the promiscuous ones.

There are also cases where the network of hydrogen bonds is the main feature that differs the native reaction from the promiscuous one. D-2-keto-3-deoxy-gluconate aldolase (KDGA) from hyperthermophilic *Sulfolobus solfataricus* reacts with both gluconate and galactonate substrates with similar rates.¹¹¹ The mechanism with both substrates involves Schiff base formation by Lys155, and subsequent hydration and cleavage. The differences between gluconate and galactonate are in the hydrogen bonds formed with KDGA's active site, and in particular in the manner by which the 5' and 6' hydroxyl groups are bound (Figure 3).¹¹²

It therefore seems that the very same active site can offer numerous modes of interactions, and some of these might be utilized by promiscuous substrates. It should be noted, however, that most of the above describes cases analyzed by kinetics and site-directed mutagenesis. Very few structures of the enzyme–substrate, or enzyme transition-state complexes, exist for both the native and promiscuous substrates. And thus, small, or even significant, changes in active-site configuration cannot be excluded in the described cases.

8.03.6.1.3 Different protonation states

There are cases in which the same active-site residue acts in two different protonation states in the native compared to the promiscuous function. In the tautomerase superfamily, various enzymes share the catalytic Pro residue at the enzyme N-terminus, but the mechanism of catalysis depends on its $\text{p}K_a$. In 4-oxalocrotonate tautomerase (4-OT) that catalyzes the 1,5-keto–enol tautomerization of 2-hydroxymuconate to 2-oxo-3-hexendioate, the $\text{p}K_a$ of Pro1 is ~ 6.4 , and it acts as a general base. In another tautomerase family member, CaaD, which catalyzes the hydrolytic halogenation of chloro- and bromoacrylates, Pro1 is protonated ($\text{p}K_a \sim 9.2$) and serves as a general acid.^{113–115} Since in 4-OT little proportion of Pro1 is present in the correct protonation state for general acid catalysis, its promiscuous hydratase activity is quite low ($2.6 \times 10^{-2} \text{ mol}^{-1} \text{ s}^{-1}$). The knowledge of 4-OT and CaaD catalytic mechanisms was used to elucidate the catalytic mechanism of yet another member of tautomerase family, malonate semialdehyde decarboxylase (MSAD). MSAD has a substantial hydratase activity (with even higher turnover number than CaaD, 5.8 s^{-1} vs. 0.7 s^{-1}), and it was proposed that its Pro1 is protonated and serves as a general acid also in the mechanism of the native MSAD activity.^{116,117}

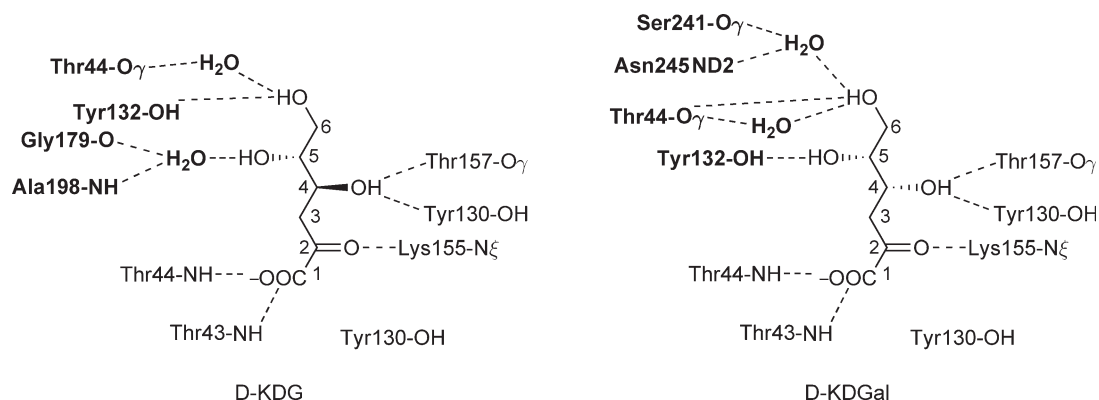


Figure 3 Schematic summary of the different interactions made in the active site of D-2-keto-3-deoxy-gluconate aldolase (KDGA) with its two substrates: D-2-keto-3-deoxy-gluconate (D-KDG) and D-2-keto-3-deoxy-galactonate (D-KDGal). Adapted from A. Theodossis; H. Walden; E. J. Westwick; H. Connaris; H. J. Lamble; D. W. Hough; M. J. Danson; G. L. Taylor, *J. Biol. Chem.* **2004**, 279, 43886–43892.

8.03.6.1.4 Different subsites within the same active site

In probably fewer cases, although both the original and promiscuous activities reside in the same active site, and rely on its major feature (e.g., an oxyanion hole), other parts of the catalytic machinery may differ significantly. One such example is serum paraoxonase (PON1), a mammalian lactonase with promiscuous esterase and PTE activities. All these activities depend on a calcium ion that serves as PON1's 'oxyanion hole', but the general base, which activates a water molecule, differs¹¹⁸ (Figure 4(a)). The hydrolysis of lactones and esters is mediated by a His115–His134 dyad, that deprotonates a water molecule to generate the attacking hydroxide. The promiscuous PTE activity appears to be mediated by another set of residues. Indeed, mutations of both histidine residues increase the promiscuous phosphotriesterase activity, and even shift selectivity in some cases. For example, the His115Trp mutation dramatically shifts the selectivity in favor of P–S versus P–O bonds.^{120,121} Coordination of the phosphoryl oxygen to the active-site calcium is one feature that is shared with the lactonase mechanism. However, the His115–His134 has no role as a base in the hydrolysis of PTEs. It has been suggested that the PTE mechanism of PON1 is analogous to the mechanism of the squid diisopropyl fluorophosphatase (DFPase),¹¹⁹ and involves a nucleophilic attack of Asp269 that comprises one of the Ca²⁺-ligating residues.^{119,122} However, this mechanism remains to be established by manifestations of a phosphoryl-enzyme intermediate, such as burst kinetics (that have not been observed with any of PON1's substrates), and the isolation of a phosphorylated enzyme species. Nonetheless, the mutagenesis data clearly indicate that key features of the lactonase and PTE mechanisms differ.

An analogous example is *Candida antarctica* lipase B (CALB) whose native activity (lipids hydrolysis) is mediated by a Ser105–His224–Asp187 catalytic triad. Using its oxyanion hole, formed by Gln106 and Thr40, CALB also catalyzes various carbon–carbon bond formation reactions, such as Michael additions and aldol condensations, with various ketone and aldehyde substrates.^{78–80} However, in these reactions, the nucleophilic serine – the key part of the catalytic triad, plays no role, and the acid–base transfer is thought to be mediated by His224 in conjunction with Asp187 (Figure 4(b)). Indeed, as in PON1, the Ser105Ala mutant exhibits higher promiscuous activities than wild-type (WT) CalB.

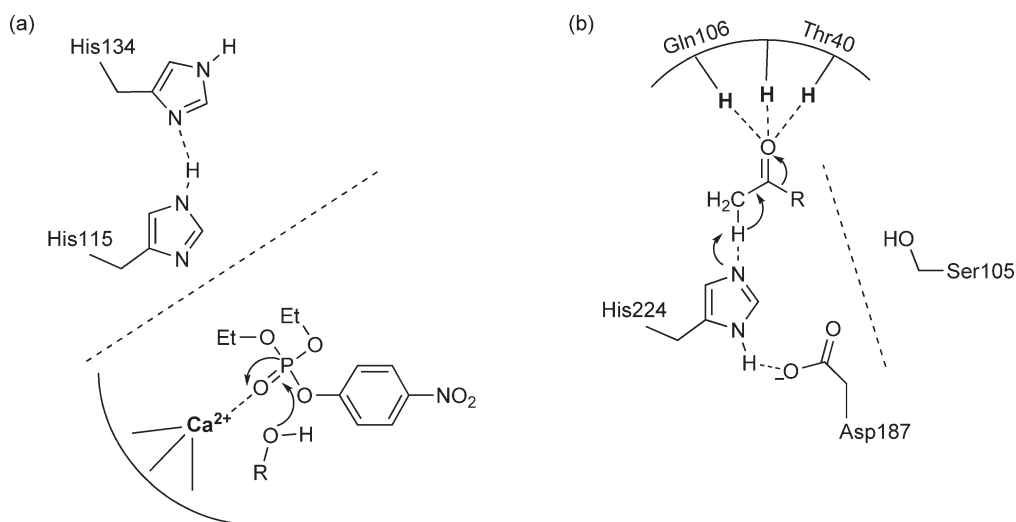


Figure 4 Different subsites within the same active site. (a) The main active site feature of the serum paraoxonase PON1 is the catalytic calcium ion, which lies at the bottom of a deep and hydrophobic active site, and is thought to act as the 'oxyanion hole' of PONs. The native function, hydrolysis of lactones, is mediated by a His115–His134 dyad, that deprotonates a water molecule to generate the attacking hydroxide. Although the same dyad appears to mediate the promiscuous arylesterase activity of PON1, the promiscuous phosphotriesterase activity (shown here for paraoxon) appears to be independent, and mediated by other residues that act as a base, or nucleophile.^{118,119} Indeed, mutations of both His residues may increase the promiscuous phosphotriesterase activity by >300-fold with certain organophosphate substrates.^{120,121} (b) A similar scenario has been described for the lipase CalB. Its native activity (lipid hydrolysis) is mediated by the Ser105–His224–Asp187 triad and the negative charge of the transition states, and the acyl-enzyme intermediate is stabilized by its 'oxyanion hole'. CalB also catalyzes promiscuous C–C bond formation reactions. In these promiscuous activities, the oxyanion hole is also utilized for negative charge stabilization (shown here). However, the catalytic serine takes no part, and acid–base transfer is thought to be mediated by His224 in conjunction with Asp187.^{78–80}

8.03.6.1.5 Promiscuity due to alternative cofactors

Changes in chemical selectivity can also be induced by metal substitution. Such changes can be defined as *cofactor ambiguity*. Following the pioneering work by Kaiser,¹²³ the introduction of copper ions (Cu^{2+}) has been shown to introduce promiscuous oxidase activities in several hydrolytic enzymes.^{124,125} In carbonic anhydrase, substitution of the Zn^{2+} by Mn^{2+} conferred the enzyme with enantioselective epoxidation of styrene.⁹⁷ Incorporating selenocysteine into the active sites of subtilisin,¹²⁶ glyceraldehyde-3-phosphate dehydrogenase GADPH,¹²⁷ and GST,¹²⁸ provided these enzymes with novel peroxidase activities.

8.03.6.1.6 Water-assisted promiscuity

Although the native substrate may interact directly with active-site residues, it is anticipated that water molecules probably play an important role in the indirect mediation of promiscuous interactions. Water molecules can bridge between a promiscuous substrate and active-site groups thus forming weak promiscuous hydrogen bonds. They can partially neutralize opposing dipoles and charges between the substrate and active-site residues, or act as acid, base, or nucleophile, in catalyzing the promiscuous reaction. Indeed, spatially defined active-site water molecules have catalytic power that is not fundamentally inferior to amino acid residues. Such water molecules may have played a key role in primordial enzymatic active sites,²⁶ and therefore probably participate in promiscuous activities. The current evidence for water-mediated promiscuity is slim, primarily because of the few structures of enzymes complexed with promiscuous substrates. However, the *Bacillus subtilis* esterase, which is also capable of amide hydrolysis, comprises an example whereby water molecules may assist the catalysis of a promiscuous substrate. A molecular dynamics study suggested that amide hydrolysis is affected by a network of hydrogen bonds consisting of water molecules. The esterase reaction is not influenced by these hydrogen bonds, due to the fact that esters lack the N-H amide group.¹²⁹

8.03.6.2 Deciphering Enzyme Mechanisms by Studying Promiscuous Functions

Enzymologists have discovered that a systematic research of the 'hidden skills' of enzymes can provide valuable insights regarding their catalytic mechanisms (see also Section 8.03.6.1.3). For example, the promiscuous hydrolysis of phosphonate diester by *Tetrahymena thermophila* ribozyme provided key insights regarding the relative importance of transition state geometry versus charge.¹³⁰ The native activity of this ribozyme is the phosphodiester hydrolysis, which differs from the promiscuous aminoacyl esterase activity by both the geometry and the charge of the transition state. Phosphonate diester was hydrolyzed by the ribozyme with similar turnover and analogous mechanism to the phosphodiester hydrolysis, thus demonstrating that the geometry of the transition state plays a more important role than its charge.

In another study, the promiscuous chorismate mutase activity of isochorismate pyruvate-lyase (PchB) was used to derive mechanistic insights into its native activity (isochorismate pyruvate lyase).¹³¹ Presumed key active-site residues were randomized, and the resulting variants of PchB were selected for the promiscuous chorismate mutase activity. Consequently, a common mechanism was proposed for both functions of PchB, with the rare [1,5]-sigmatropic rearrangement for the lyase activity, being distinct from other pyruvate lyases.

8.03.6.3 Mechanistic Origins of Differences in the Catalytic Parameters for Native Versus Promiscuous Functions

As noted above, the magnitude of promiscuous activities varies over many orders-of-magnitude, both in absolute terms, and relative to the native activity (Section 8.03.4.3). The differences in reactivity between the native and promiscuous substrates can be manifested in differences in both k_{cat} and K_{M} . The schematic view is that the energetics of substrate binding are reflected in the K_{M} , and catalysis by k_{cat} . It is therefore expected that promiscuous substrates that bind poorly, due to steric hindrance, for example, will exhibit high K_{M} values. However, many promiscuous substrates are characterized by low k_{cat} values. For example, a systematic analysis of >50 substrates for the enzyme PON1, the primary function of which is lipophilic lactonase, indicated that the promiscuous aryl ester, and phosphotriester, substrates all exhibit K_{M} values in the millimolar range (0.8–5 mmol l^{-1}).⁹² This is despite the fact that their $k_{\text{cat}}/K_{\text{M}}$ values vary over three orders of magnitude. The differences in reactivity are therefore primarily due to k_{cat} values that vary by >1000-fold. The probable

reason is that for the promiscuous aryl ester and phosphotriester substrates, substrate binding is driven primarily by nonspecific hydrophobic forces with the deep and hydrophobic active site of PON1. It appears that, in many cases, promiscuous substrates are inadequately positioned relative to the catalytic machinery, and therefore exhibit very low k_{cat} values. Interestingly, for the lactones that comprise that native substrate of this enzyme, K_M values vary by ~ 200 -fold (from about 0.1 up to 20 mmol l^{-1}), whereas the variations in k_{cat} values are orders of magnitude lower (~ 10 – 200 s^{-1}).

It is therefore anticipated, and often observed, that the mode of binding of the native substrate – that is typically mediated by several independent, enthalpy-driven interactions – is fundamentally different from that of the promiscuous substrates where hydrophobic and other entropy-driven interactions play a key role.

8.03.7 Promiscuity and the Divergence of Enzyme Superfamilies

Enzyme superfamilies include numerous enzymes that although distant in sequence, share the same fold and the same catalytic mechanism. Members of such diverse superfamilies catalyze different chemical transformations of many different substrates, but share a common motive of catalysis.¹¹ Analysis of enzyme families and superfamilies provides the most solid and convincing body of evidence for the role of promiscuity in the evolution of new functions. Specifically, the identification of promiscuous activities, or cross-reactivities, between different members of the same enzyme family or superfamily, and the directed evolution of these activities, provide important hints regarding evolutionary, structural, and mechanistic relationships within enzyme families (Figure 5).

Several examples of the promiscuous catalytic activities within enzyme families and superfamilies are listed in Table 2. Key observations based on these data are summarized below:

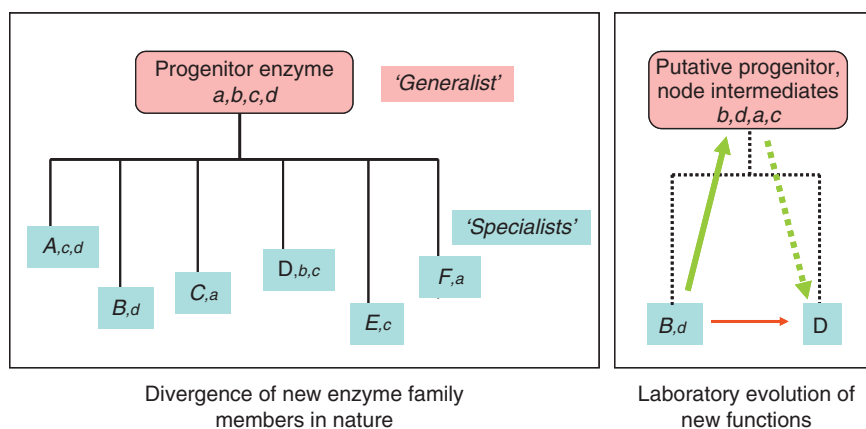


Figure 5 Experimental evidence favors the model of divergence of a 'generalist' progenitor enzyme to a family of 'specialist' enzymes. *Left panel:* Jensen's hypothesis²⁵ surmises that, in nature, an ancestor protein displaying a low level of a range of activities (denoted as a , b , c , d) had been subjected to selection pressures for those activities, thus duplicating and diverging into a family of potent and highly specialized enzymes of the kind seen today (denoted A, B, etc.). Today's 'specialists' may still retain some of the functions of the common ancestor (denoted in lower case), as low levels of promiscuous activities. Indeed, several reports indicate a low level of shared activities within a family, and in particular that the native activity of one member is the promiscuous activity of another, and vice versa (Table 2). *Right panel:* Additional support to the above model comes from the results of many directed evolution experiments. Direct switches of specificity, for example, from B to D (red arrow) are rare, and are typically seen following a parallel selection for an increase in the target activity and elimination of the original one. Upon mutation and selection for an increase of a promiscuous activity (green arrow), the resulting variants usually show significant increases in the target activity, and a smaller decrease in the original one, thus yielding, in effect, a 'generalist' intermediate exhibiting both d and b at relatively high levels (the 'weak negative trade-off' line in Figure 6). Such intermediates are often observed in the laboratory; some even gain other activities, never selected for (denoted a , c), and may therefore resemble the progenitor of this enzyme family, or node intermediates along past routes of its divergence.

Table 2 Examples for promiscuous activities within enzyme families and superfamilies

	<i>Family/ Superfamily</i>	<i>Enzymes</i>	<i>Native activity; substrate (k_{cat}/K_M in $\text{mol}^{-1} \text{I s}^{-1}$)</i>	<i>Promiscuous activity; substrate (k_{cat}/K_M in $\text{mol}^{-1} \text{I s}^{-1}$)</i>	<i>Reference(s)</i>
1	Mammalian Paraoxonases (PONs)	PON1 (serum paraoxonase) PON2 PON3	Lipo-lactonase – aliphatic 5-, 6-membered ring lactones with lipophylic side chains (γ -dodecanoic lactone, 1.2×10^5)	PON1: Aryl esterase (phenyl acetate, $\sim 6 \times 10^5$) Phosphotriesterase (paraoxon, 6×10^3) PON2: barely detectable aryl esterase; no phosphotriesterase PON3: low aryl esterase; barely detectable phosphotriesterase	91, 92
2	Tautomerase superfamily	Malonate semialdehyde decarboxylase (MSAD) 4-oxalocrotonate tautomerase (4-OT) YwhB tautomerase (4-OT analogue) <i>trans</i> -3-chloroacrylic acid dehalogenase (CaaD)	Decarboxylation of malonate semialdehyde (2.2×10^7) Tautomerization of 2-hydroxymuconate to 2-oxo-3-hexenoate ($>10^5$) Tautomerization of 2-hydroxymuconate to 2-oxo-3-hexenoate ($\sim 10^5$) (1.3×10^5) Hydrolytic dehalogenation (hydratase) of 3E-haloacrylates (1.2×10^5 , 3E- chloroacrylate)	Hydration of 2-oxo-3-pentynoate (6×10^2) Hydration of 3E-chloroacrylate (CaaD activity) (2.6×10^{-2}) Hydration of 3E-chloroacrylate (CaaD activity) (4.4×10^{-2}) Hydration of 2-oxo-3-pentynoate (6.4×10^3)	113–117
3	ROK family (repressor, open reading frame, kinase)	NanK YajF YcfX AlsK	<i>N</i> -acetyl-D-mannosamine kinase (2.7×10^5) Fructose kinase (1.1×10^4) Unknown Allose kinase, (6.5×10^4)	Glucose kinase (5.1×10^2) Glucose kinase (2×10^2) Glucose kinase (2.4×10^3) Glucose kinase (15)	71, 72
4	Enolase superfamily: MLE (muconate lactonizing enzyme) subgroup	<i>o</i> -succinylbenzoate synthase (OSBS)	Dehydration of SHCHC (2-succinyl-6R- hydroxy-2,4-cyclohexadiene-1R- carboxylate) (2.5×10^5)	<i>N</i> -acylaminoacid racemase (NAAAR) reaction with <i>N</i> -acetyl methionine isomers ($4.9\text{--}5.9 \times 10^2$)	13, 132

(Continued)

Table 2 (Continued)

	<i>Family/ Superfamily</i>	<i>Enzymes</i>	<i>Native activity; substrate (k_{cat}/K_M in $\text{mol}^{-1} \text{I s}^{-1}$)</i>	<i>Promiscuous activity; substrate (k_{cat}/K_M in $\text{mol}^{-1} \text{I s}^{-1}$)</i>	<i>Reference(s)</i>
5	Amidohydrolase superfamily	Phosphotriesterase from <i>P. diminuta</i> (PTE)	Phosphotriesterase (paraoxon, 4×10^7)	Aryl esterase (2-naphthyl acetate, 500); lactonase (dihydrocoumarin, 6.5×10^5)	88, 87
		Phosphotriesterase homology protein (PHP)	Unknown	Aryl esterase (2-naphthyl acetate, 70)	
		Dihydroorotase (DHO)	Dihydroorotic acid hydrolysis (1.2×10^6)	Phosphotriesterase (paraoxon, 2.8)	
		AhIA; a member of the PLL family (PTE-like lactonases)	Lactonase; <i>N</i> -3-oxooctanoyl L- homoserine lactone (0.7×10^6)	Phosphotriesterase (paraoxon, 0.5)	
		PPH; a member of the PLL family (PTE-like lactonases)	Lactonase; <i>N</i> -3-oxooctanoyl L- homoserine lactone (0.55×10^5)	Phosphotriesterase (paraoxon, 8.6)	
		PTE-like lactonase SsoPox	Lactonase; <i>N</i> -3-oxooctanoyl L- homoserine lactone ($>10^6$)	Aryl esterase (naphthyl acetate, 400); phosphotriesterase (paraoxon, 4000)	
6	Orotidine 5' monophosphate decarboxylase suprafamily (OMPDC)	3' keto L-gluconate 6-phosphate decarboxylase (KGPDS)	Decarboxylation of 3' keto L-gluconate 6-phosphate (7.7×10^4)	Aldol condensation of D-ribose 5-phosphate and formaldehyde (HPS activity, 8.2×10^{-2})	133
		D- <i>arabino</i> -hex-3-ulose 6-phosphate synthase (HPS)	Aldol condensation of D-ribose 5-phosphate and formaldehyde (1.6×10^4)	Decarboxylation of 3' keto L- gluconate 6-phosphate (KGPDS activity, 2.3×10^3)	
7	<i>N</i> -acetyl-neuraminate lyase (NAL) family, pyruvate-dependent aldolases	<i>N</i> -acetyl-neuraminate lyase (NAL)	Cleavage of <i>N</i> -acetyl-neuraminate (3.1×10^3)	Aldol condensation of pyruvate and L-aspartate – β - semialdehyde (DHDPS activity) (20)	134
		Dihydrodi-picolinate synthase (DHDPS)	Aldol condensation of pyruvate and L-aspartate – β -semialdehyde		
8	Alkaline phosphatase superfamily	Alkaline phosphatase	Phosphomonoesters hydrolysis (<i>p</i> -nitrophenyl phosphate, 3.3×10^7)	Phosphodiester hydrolysis (bis- <i>p</i> - nitrophenyl phosphate, 5×10^{-2}) Sulfate ester hydrolysis (<i>p</i> - nitrophenyl sulfate, 1×10^{-2})	82–84
		Nucleotide pyrophosphatase/ Phosphodiesterase (NPP)	Phosphodiester hydrolysis (thymidine 5'-monophosphate 4-nitrophenyl ester, 1.6×10^6)	Phosphomonoesters hydrolysis (<i>p</i> -nitrophenyl phosphate, 1.1)	

9	Guanidino-modifying enzyme superfamily (GMSF), hydrolase branch	Arginine deiminase (PaADI)	Arginine hydrolysis (4.5×10^4)	N^w, N^w -dimethylarginine hydrolysis (PaDDAH activity) (1.8×10^3)	110
10	C–C hydrolase family (branch of α/β hydrolase superfamily)	Agmatine deiminase (PaAgDI) N^w, N^w -dimethyl-arginine dimethyl-aminohydrolase (PaDDAH) C–C hydrolase <i>MhpC</i> from <i>E. coli</i>	Agmatine hydrolysis (7×10^3) N^w, N^w -dimethylarginine hydrolysis (1.8×10^3) C–C bond cleavage (2-hydroxy 6-ketona-2,4-dienoic acid, 28 units)	Arginine hydrolysis (PaDDAH activity) (1.8) Esterase (mono-ethyl adipate, 0.0027 units) Thioesterase (thioethyl adipate, 0.46 units) Hydroxamic acid formation (mono-ethyl adipate + NH_4OH , 0.013 units)	135–137
		Haloperoxidase/Esterase ThcF from <i>Rhodococcus erythropolis</i> Lactonase from <i>Acinetobacter calcoaceticus</i>	Haloperoxidase (monochlorodimedon, $V_{\max} = 0.45 \text{ nmol min}^{-1}$) Lactonase (3,4-dihydrocoumarin, $V_{\max} = 4760$ units)	Esterase (<i>p</i> -nitrophenyl acetate, $V_{\max} = 2.58 \text{ nmol min}^{-1}$) Haloperoxidase (monochlorodimedon, $V_{\max} = 199$ units)	

- (i) The same promiscuous activity is often shared between more than one family member (**Table 2**, entries 2, 3, 5, 10). For example, several members of the tautomerase family share a promiscuous hydratase activity (although its efficiency varies greatly),^{116,117} and all known kinases of the ROK family can utilize glucose as a promiscuous substrate.^{71,72} That the same promiscuous activity is shared by more than one family member can hint toward the existence of yet unidentified family members in which this promiscuous activity comprises the native activity, and also provide new starting points for directed evolution.^{87,88}
- (ii) Promiscuous functions can also appear in one family member but not in the others (**Table 2**, entries 1, 5, 9). For example, in the guanidino-modifying enzyme superfamily (GMSF), no promiscuity was observed in agmatine deiminase (PaAgDI), whereas the other two family members exhibit promiscuous activities.¹¹⁰ In the mammalian paraoxonases family, the promiscuous PTE activity is quite efficient in PON1 ($k_{\text{cat}}/K_{\text{M}} \sim 10^4 \text{ mol}^{-1} \text{ s}^{-1}$), barely detectable in PON3, and nonexistent in PON2. Indeed, the consistency of the lactonase function in all PON family members, and the haphazardness of the others activities (i.e., the paraoxonase and aryl esterase observed only in some family members; **Table 2**, entry 1), enabled the identification of the lactonase as the native function of PONS.^{91,92} This trend is also seen in a recently identified family of bacterial lactonases within the amidohydrolase superfamily. The magnitude (in terms of $k_{\text{cat}}/K_{\text{M}}$) of promiscuous phosphotriesterase activity varies from $0.5 \text{ mol}^{-1} \text{ s}^{-1}$ in one member, to $4000 \text{ mol}^{-1} \text{ s}^{-1}$ in another.⁸⁸ The above patterns are consistent with promiscuous activities being under no selection, and also with the observation that promiscuous activities show large increases and decreases in response to one or few mutations which are neutral with respect to the primary function.⁹⁰
- (iii) The primary, or native, function of one family member is often identified as a promiscuous activity in other family members (**Table 2**, entries 2, 5–9). In the tautomerase superfamily, the promiscuous activity of 4-OT, and its homologue YwhB tautomerase, is the primary native activity of *trans*-3-chloroacrylic acid dehalogenase (CaaD).¹¹³ A similar picture is observed in the orotidine 5' monophosphate decarboxylase (OMPDC) family, the amidohydrolase superfamily, the *N*-acetyl-neuraminatase lyase (NAL) family, the guanidino-modifying enzyme family, and in the alkaline phosphatase superfamily.¹³³
- (iv) Following the above, promiscuous activities may comprise a vestige of the progenitor of an enzyme, or enzyme family (**Figure 5**). This principle was recently demonstrated in an attempt to trace the origins of a bacterial PTE (from *P. diminuta*) – an enzyme thought to have evolved toward the degradation of the synthetic insecticide paraoxon that has only been introduced in the twentieth century. It was found that PTE possesses a promiscuous lactonase activity,⁸⁷ and assumed that this activity could comprise a vestige of its progenitor. Indeed, three homologues from the same superfamily (amidohydrolase) turned out to be representatives of a new group of microbial lactonases dubbed PTE-like lactonases (PLLs). These enzymes proficiently hydrolyze lactones, and in particular *N*-acyl homoserine quorum-sensing lactones, and exhibit weaker promiscuous PTE activities. PLLs share key sequence and active-site features with PTE, and differ primarily by an insertion in one active-site loop. Following their biochemical and biological function, PLLs are probable to have existed for many millions of years. We therefore suggested that PTE could have evolved from a member of the PLL family while utilizing its latent promiscuous paraoxonase activity as an essential starting point.⁸⁸
- (v) Laboratory evolution of one promiscuous activity often leads, indirectly, to the appearance of other promiscuous activities (e.g., **Table 3**, entries 1 and 3) thus yielding 'generalist' intermediates.⁷⁴ Some of the latter might appear in other family members, as either their native, or promiscuous function.^{73,87}

8.03.8 Evolutionary Aspects of Promiscuity

As mentioned in Section 8.03.7, the studies of divergent evolution within enzyme families and superfamilies provide support for the hypothesis that, throughout evolution, promiscuous activities served as the starting points for the divergence of new functions, and for Jensen's hypothesis that broad-specificity enzymes served as progenitors for the whole families and superfamilies of today's specialized enzymes.²⁵ At present, however, this evidence is largely circumstantial, and provides little insight as to the actual mechanisms, and mutational paths, that underline the processes of divergence. Discussed below are several issues that relate to the evolution of new enzymatic functions, and various models that describe the divergent evolution of new genes carrying new enzymatic functions.

8.03.8.1 The Evolvability of Promiscuous Enzyme Functions

Accumulating experience in the laboratory indicates that there are very few cases in which ‘something could be evolved out of nothing’, namely that a completely novel activity was evolved, or engineered, in the laboratory. For example, the emergence of an enzymatic function in a noncatalytic fold demanded the exploration of vast libraries, the genetic diversity of which exceeded natural genetic diversities by many orders of magnitude.¹³⁸ In another case, although the starting point was an enzyme from the same superfamily, and the key active-site catalytic features (a bi-metallo catalytic center in this case) were maintained, the incorporation of a novel function demanded major sequence alterations such as simultaneous deletion and insertion of few active-site loops, a series of engineered point mutations, and the parallel exploration of random mutations.¹³⁹ Computational design has also been applied toward the generation of novel enzymes, but the introduction of novel enzymatic functions also involved a large number (>8) simultaneous amino acid changes.^{140,141} Most notably, these cases, in which a novel activity was introduced, all involve starting points, and/or intermediates, that possess no activity whatsoever. Evolution, however, is a gradual and smooth process. It involves discrete steps of one mutation at a time (be it a point mutation, an insertion, or a deletion) that yields a folded and functional protein. Thus, all intermediates must be functional, at least to some degree.¹⁴²

8.03.8.1.1 Promiscuity as a starting point – the three basic assumptions

Taking for granted the demand for smooth transitions, it is probable that natural evolution routinely takes advantage of promiscuous activities as starting points for the divergence of new enzymes. However, for promiscuity to benefit, let alone lead, the divergence of new enzyme functions, three basic prerequisites must be assumed.

- (i) Once a promiscuous function becomes available, it can be easily improved through one or just few mutations. Indeed, almost all laboratory evolution projects (or at least the successful ones that are reported) aim at further evolving a promiscuous activity, typically for a substrate, or a reaction, that bears some resemblance to the original function. The conclusion from hundreds of such works is that promiscuous functions exhibit high ‘plasticity’ – few mutations can readily increase a promiscuous activity, typically by 10–10³-fold, and 10⁴–10⁶-fold improvements in response to a single mutation were also reported.^{85,143} More examples, and other aspects related to the evolvability of promiscuous functions, are listed in **Table 3**, and discussed in Sections 8.03.8.3–8.03.8.5.
- (ii) Weak promiscuous activities can provide an immediate advantage, and thus become under selection.
- (iii) The divergence path can be completed to give a newly specialized enzyme for which, the promiscuous activity became the native one.

These points are discussed in detail in Sections 8.03.8.2–8.03.8.5.

8.03.8.2 Promiscuous Functions Can Provide an Immediate Advantage

Several reports indicate that, when necessary, weak promiscuous activities can provide an immediate selective advantage to an organism. This has been often seen in the emergence of promiscuous functions following a deficiency created by genetic manipulation in the laboratory. Several examples, including a systematic study conducted by Patrick *et al.*,⁵⁰ are discussed in Section 8.03.2. In another study, performed with an *E. coli* strain deficient of glucokinase activity, several sugar kinases were found that promiscuously phosphorylate glucose.⁷¹ Some of these promiscuous activities are notably weak (**Table 2**); for example, the $k_{\text{cat}}/K_{\text{M}}$ values of the promiscuous sugar kinase *YajF* is the range of 10² mol⁻¹ s⁻¹, and is ~10⁴ lower than that of the primary *E. coli* glucokinase (*Glk*). In all these cases, the promiscuous function complemented a deficiency in a native enzyme by overexpression of the promiscuous enzyme. Clearly, low catalytic efficiency can be compensated by higher enzyme levels.¹⁵⁴ However, the levels of overexpression from a multiple-copy plasmid, and a powerful promoter, are usually not comparable with expression levels from chromosomal copies, and weak promoters, under which most natural enzymes are expressed.

Table 3 Examples for directed evolution of promiscuous enzyme functions and their trade-offs with the native function^a

Enzyme	Native activity (k_{cat}/K_M of wild type, mol^{-1} / s^{-1})	Promiscuous activity under selection (k_{cat}/K_M of wild type, mol^{-1} / s^{-1})	Mutations in selected variants	Effect on native activity ($k_{cat}/K_M^{variant} / k_{cat}/K_M^{wt}$)	Effect on the evolved promiscuous activity ($k_{cat}/K_M^{variant} / k_{cat}/K_M^{wt}$)	Comments	Reference(s)
1 Aspartate aminotransferase (AATase) from <i>E. coli</i>	Transamination of dicarboxylic substrates (9.1)	Transamination of tyrosine (0.055) and phenylalanine (0.012) (TATase activity)	Pro13Thr Asn69Ser Gly72Asp Arg129Gly Thr167Ala Ala293Val Asn297Ser Asn339Ser Ala381Val Asn396Asp Ala398Val Glu323Gly	1.2-fold higher	130- and 270-fold higher, respectively	This work provides a clear example of a 'generalist' intermediate. The <i>in vitro</i> evolved enzyme exhibits wild-type-like AATase activity, and TATase activity that is >10% that of wild-type TATase.	73
2 Muconate lactonizing enzyme (MLE II) from <i>Pseudomonas</i> sp. P51	Cycloisomerization (2×10^4)	β -Elimination (<i>o</i> -succinylbenzoate synthase, OSBS activity). No detectable promiscuous activity (nondetectable) ($<1.5 \times 10^{-3}$)		15-fold lower	>1.2 million-fold higher	The corresponding mutation when engineered in a homologous enzyme (Asp297Gly, in AEE) decreased the native function far more significantly (see Vick <i>et al.</i> ¹⁴⁴).	85
3 Galactokinase (GalK) from <i>E. coli</i>	Phosphorylation of <i>D</i> -galactose to produce α - <i>D</i> -galactose-1-phosphate (860)	Phosphorylation of C5- or C6-substituted sugars (9.8 for <i>D</i> -fucose, and nondetectable for the other substrates)	Tyr371His	1.3-fold lower	21-fold higher for <i>D</i> -fucose, and higher improvements for the other target substrates	This variant expanded the spectrum of substrates to substrates that were not used in the screen. Although the Y317H mutation retains the stringent requirement for the C-4 galactose architecture, it exhibits enhanced substrate flexibility at all other positions.	145
4 β -Glucuronidase (GUS) from <i>E. coli</i>	Hydrolysis of β -glucuronides (8.3×10^5)	Hydrolysis of pNP-galactoside (2.3)	Ile12Val Phe365Ser Trp529Leu Ser557Pro Ile560Val	8.3-fold lower	16-fold higher	Larger increases in the evolving promiscuous galactosidase function of <i>E. coli</i> GUS, with smaller changes of the native function, and acquisition of specificities not selected for were previously described. ⁷⁴	75

5	SinI DNA-methyltransferase from bacteriophage	Methylation of the internal cytosine of the GG(A/T)CC sequence (2.9×10^5)	Relaxation of sequence specificity toward GG(N)CC (2×10^3)	Leu214Ser Tyr229His	4.5-fold lower	18.5-fold higher for the GG(G/C)CC sequence	Similar trends of specificity broadening were observed with <i>HaeIII</i> methyltransferase. ⁷⁶	146
6	Phosphotriesterase from <i>Pseudomonas diminuta</i> (PTE)	Phosphotriesterase (e.g., paraoxon, 4×10^7)	Ester hydrolysis (e.g., 2-naphthyl acetate, 480)	His254Arg Phe306Cys Pro342Ala	Threefold lower	13-fold higher	Up to 150-fold higher activity was observed with esters not selected for.	86, 87
7	Human carbonic anhydrase (hCAII)	Bicarbonate dehydration (3×10^7)	Esterase (e.g., <i>p</i> -nitrophenyl acetate, 2×10^3)	Ala65Val, Asp110Asn Thr200Ala	Twofold lower	40-fold higher	Mutations in conserved regions of the protein did not affect the highly proficient native activity despite the absence of a purifying selection for bicarbonate dehydration.	86, 96
8	Mammalian serum Paraoxonase (PON1)	Lipo-lactonase ^b (e.g., δ -valerolactone, 1.3×10^5 ; and γ -heptanolide, 2×10^4)	Thiolactonase (e.g., γ -butyryl thiolactone, ⁹⁴)	Ile291Leu Thr332Ala	Approx. no change ^b	80-fold higher	The selected mutations are all located on surface loops that comprise the substrate-binding pocket.	86, 147
			Esterase (e.g., 2-naphthyl octanoate, 1.5×10^3)	Phe292Val Tyr293Asp	Approx. no change	31-fold higher		
			Esterase (e.g., 7-acetoxy coumarin, 1.2×10^5)	Phe292Ser Val346Met	~22-fold lower	62-fold higher		
			Phosphotri-esterase (7-diethyl-phosphoro 4-cyano-7-hydroxycoumarin, 9×10^3)	Leu69Val Ser138Leu Ser193Pro Asn287Asp	2.6-fold lower	155-fold higher		

(Continued)

Table 3 (Continued)

<i>Enzyme</i>	<i>Native activity</i> (k_{cat}/K_M of wild type, $mol^{-1} s^{-1}$)	<i>Promiscuous activity</i> <i>under selection</i> ($k_{cat}/$ K_M of wild type, $mol^{-1} s^{-1}$)	<i>Mutations in</i> <i>selected</i> <i>variants</i>	<i>Effect on native</i> <i>activity</i> ($k_{cat}/$ $K_M^{variant}/k_{cat}/K_M^{wt}$)	<i>Effect on the</i> <i>evolved</i> <i>promiscuous</i> <i>activity</i> ($k_{cat}/$ $K_M^{variant}/k_{cat}/K_M^{wt}$)	<i>Comments</i>	<i>Reference(s)</i>
9 Deacetoxy cephalosporin C synthase (DAOCS) from <i>Streptomyces clavoligerus</i>	Ring expansion of penicillin N into deacetoxy cephalosporin C (2.2×10^4)	Ring expansion of penicillin G into phenylacetyl-7-aminodeacetoxy-cephalosporanic acid ¹⁸	Val275Ile Ile305Met	1.1-fold higher	32-fold higher		148
			Cys155Tyr Tyr184His Val275Ile Cys281Tyr	42-fold lower	41-fold higher		
11 β -Lactamase TEM-1	Ampicillin hydrolysis (4.18×10^7)	Cefotaxime hydrolysis (2.07×10^3)	Gly238Ser	6.2-fold lower	86-fold higher	These mutants evolved resistance in the clinic and were later reproduced in the laboratory	23
		Ceftazidime hydrolysis (32.1)	Gly238Ser		19-fold higher		
		Cefotaxime hydrolysis (2.07×10^3)	Gly238Ser E104K	29-fold lower	806-fold higher		
12 Extended-spectrum β -lactamase CTX-M	Hydrolysis of cephalothin and cefotaxime ($4 \times 10^6 - 2 \times 10^7$)	Ceftazidime hydrolysis (32.1)	Gly238Ser Glu104Lys		284-fold higher		149
		Hydrolysis of ceftazidime (3.3×10^3)	Gln87Leu His112Tyr Thr230Ile, Ala231Val Asp240Gly Arg276His	1.4-fold higher, and 1.4-fold lower, for cephalo- thin and cefotaxime, respectively	24-fold higher	These mutants evolved resistance in the clinic and were later reproduced in the laboratory	
13 <i>NotI</i> from <i>Nocardia otitis-caviarum</i>	Recognition and cleavage of GCGGCCGC DNA sequence (5×10^5 U mg ⁻¹ enzyme)	Recognition and cleavage of altered 8-bp sequence (no detectable star activity)	Met91Val Glu156Gly	23-fold lower	>32-fold higher than the Glu156Gly intermediate with GCTGCCGC sequence	Although a considerable reduction in the rate of cleavage of the original sequence is reported, the cleavage specificity of M91V/E156 appears to be relaxed toward a whole set of 8 bp sequence targets, with a distinct preference for the original target.	150

14	D-allose kinase (AlSK),	AlSK – phosphorylation of D-allose (2.5×10^5)	AlSK – phosphorylation of D-glucose (3.4×10^2)	Ala73Gly	1.25-fold lower	62-fold higher	151	
	N-acetyl D-mannosamine kinase (NanK)	NanK – phosphorylation of N-acetyl D-mannosamine (1.5×10^5)	NanK – phosphorylation of D-glucose (3.4×10^3)	Phe145Leu Leu84Pro	1.28-fold higher Twofold lower	11.4-fold higher 11.8-fold higher		
15	ProFAR isomerase (HisA)	Isomerization of N'-[(5'-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide (1.2×10^6)	Isomerization of phosphoribosyl-lanthranilate = TrpF activity (ND)	Val138Met Asp127Val	1.25-fold lower $\sim 10^4$ -fold lower	6.4-fold higher The wild-type activity is below detection limits	Almost all the original HisA activity was lost	152

^a Shown are examples from the last few years for which kinetic parameters are available for both the promiscuous activity under selection, and the original activity. For more examples see Supplementary Table 8 in Aharoni *et al.*⁸⁶ Because the above analysis aims at providing insights on the evolution of new enzyme functions in nature, the examples selected involve selection for only one parameter – increase in a promiscuous activity, and make use of gene libraries prepared by mutagenesis in a completely random manner (point mutations or shuffling) and throughout the genes.

^b Since the publication of this work Aharoni *et al.*,⁸⁶ it has been established that serum paraoxonase (PON1) is a lipo-lactonase, and its preferred substrates are five- and six-membered ring lactones, typically with aliphatic side chains.^{94,95,153} In the original article,⁹⁰ data for trade-offs with the native activity were presented with both the aromatic lactone dihydrocoumarin, and aliphatic lactones. However, more recent works indicated that dihydrocoumarin does not bind PON1's active site in the same mode as aliphatic lactones.^{95,118} Thus, the trade-offs presented here are the average values of two aliphatic lactones (δ -valerolactone and γ -heptanolide).

A notable exception is the case of alkaline phosphatase, the promiscuous phosphite oxidation of which can complement the deficiency of *E. coli* knockout strains (see Section 8.03.6.1.2). Complementation in this case, and the ability to grow on phosphite as the sole source of inorganic phosphorous, occurred through the chromosomal copy of alkaline phosphatase, owing to the extremely high expression levels of the native alkaline phosphatase under phosphate starvation.⁴⁷ In other cases, changes in regulation leading to higher expression were observed. Indeed, few reports describe changes in regulation in cases where a promiscuous enzyme activity came under selection. For example, Hall's classical experiment of the emergence of an alternative β -galactosidase (*egb*) was performed on the *E. coli* chromosome, and not through complementing plasmids. Besides mutations that increased a weak promiscuous β -galactosidase activity in the *egb* glycosylase (whose native function remains unknown), the first mutation that occurred dramatically increased the expression of *egb* by removing its repressor.¹⁵⁵ Miller and Raines also observed a promoter mutation in a complementing plasmid that led to \sim 100-fold increase in expression level of the promiscuous glucokinase *YajF* mentioned above.⁷² As described in Section 8.03.8.6, gene duplication is another relatively abundant event that can lead to an increase in enzyme levels. Thus, if and when a new activity becomes necessary, the combination of a weak promiscuous activity with an increase in enzyme levels can provide the organism an immediate advantage.

8.03.8.3 Negative Trade-offs and the Evolvability of Promiscuous Functions

As discussed above, the first assumption regards the evolvability of promiscuous functions, namely, once a promiscuous function becomes available, it can be easily improved through one or just few mutations (Section 8.03.8.1). A key related issue regards the notion that mutations leading to improvements in promiscuous functions need not induce parallel decreases in the native function.⁸⁶ For the reasons explained below, we dubbed this feature as 'the evolvability of promiscuous functions'. Negative trade-offs between the evolving function and existing function are a dominant factor in evolution.¹⁵⁶ It is therefore of crucial importance that in many cases, promiscuous protein functions evolve with negative trade-offs that are weaker than generally assumed. Indeed, the weak trade-off hypothesis directly inflicts on our understanding of how new genes carrying new functions emerge.

Evolvability, or evolutionary adaptability, is the capacity of biological systems, be they organisms, cells, or proteins, to evolve. Evolvability comprises of two elements:^{31,157} the first one is the induction of novel phenotypic traits by a relatively low number of mutations (this feature is often dubbed 'plasticity'). As mentioned in Section 8.03.8.1, this property of promiscuous enzyme functions has been established by numerous directed evolution experiments. Not only single proteins, but whole metabolic pathways were found to be plastic and evolvable. Moreover, it seems that the more plastic is a metabolic pathway, the more evolvable are the proteins comprising it.¹⁵⁸ However, plasticity is in conflict with the fact that most mutations are deleterious.^{159–161} Since organisms must constantly endure a significant number of mutations while maintaining their fitness, and the structure and function of their proteins, they have evolved a certain level of resistance to the effects of mutations ('robustness').³¹ It appears that proteins exhibit both traits, namely plasticity and robustness, and the two need not be mutually exclusive.^{31,162} The promiscuous, accidental functions of the protein are highly plastic. They can be reshaped through few mutations that significantly increase or decrease them. However, these mutations need not have a large effect on the protein's native activity. Indeed, the results of many directed evolution experiments indicate that, in clear contrast to the dramatic shifts observed with the promiscuous substrates, the native activities, that take place in the very same active site, often show comparatively small changes. The robustness of the native function is observed despite the fact that the only selection criterion applied in these experiments was an increase in one of the promiscuous activities of the target enzyme.

We have initially described this trend in three different enzymes subjected to a selection for an increase in six different promiscuous activities.⁸⁶ The same trend was identified in other laboratory experiments aimed at increasing promiscuous enzymatic and binding activities of various proteins (see Supplementary Table 8 in Aharoni *et al.*⁸⁶). Averaging eighteen cases in which data was provided for the effect of the selected mutations on both the evolving promiscuous activity and the original, native function (for both binding and enzymatic functions) indicated that, 1–4 mutations increased the promiscuous activity that was under selection by $>$ 1000-fold, on average, whereas the original activity of these proteins decreased, by 3.2-fold. More recent examples

are listed in **Table 3**. They show a similar trend: 1–11 mutations increased the promiscuous activity under selection by 10–10⁶-fold, whereas the original activity of these proteins decreased, by 0.8–42-fold. In the majority of cases, the ratio of increase in the selected promiscuous function to decrease in the original one is >10. In fewer cases, the change in the evolving promiscuous function, and in the original function, is comparable. Besides, other variants from the same selection show weaker trade-offs (e.g., **Table 3**, entry 9). There are also cases in which large negative trade-offs were observed, and these are addressed in detail in Section 8.03.8.4.

Although out of the scope of this review, it is notable that similar trends can be clearly seen in various receptors, where the acquisition of specificity for a new effector exploits the promiscuity of existing receptors.^{163,164} As also demonstrated with bacterial transcription factors, new effector specificities can then be acquired by natural evolution, or laboratory rounds of mutagenesis and selection, often with weak negative trade-offs with respect to the original effector.¹⁶⁵

There is little doubt that ultimately, the acquisition of a highly proficient new enzyme comes at the expense of the old function. However, the *relative rates* by which a new function is gained, and the old one is lost, matters. The model depicted in **Figure 6** assumes that trade-offs can be determined and quantified, in particular with enzymatic activities. It suggests that in those cases where the negative trade-off is weak (red line), the divergence of a new function can proceed through a ‘generalist’ intermediate that exhibits broad specificity. Gene duplication may then follow this process, rather than initiate it, and lead to divergence of a new ‘specialist’. The results of several directed evolution experiments also convincingly demonstrate that, the concave route – ‘strong negative trade-off’, is also applicable (**Figure 6**, blue line), in particular when a *dual* selection pressure applies, namely, when a parallel selection for an increase of a promiscuous activity, and decrease in the native activity, is applied.^{143,166–168} Thus, in the face of selection for specialization, proteins can evolve to yield ‘specialists’, sometimes with a surprisingly abrupt shift in selectivity.¹⁴³

Although the convex route denoted in **Figure 6** might be common, in other cases, a single amino acid exchange can completely switch the specificity of an enzyme (see also Section 8.03.8.4). For example, the His89Phe mutation

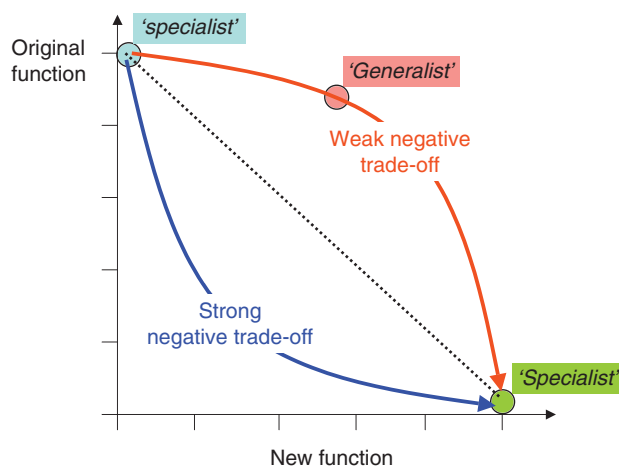


Figure 6 Possible routes to the divergence of a new function. Under selection, a weak, promiscuous activity of a protein with an existing function (blue circle) gradually evolves. By the end of this process, that typically requires many generations of mutation and selection, the ‘new’ function has traded off with the original one (green circle). However, the dynamics of this process may vary. The gain–loss of the new versus old function, and the conversion of one ‘specialist’ protein into another, may trade-off linearly (dashed line), or follow either concave, or convex, routes. Results of numerous directed evolution experiments indicate that, the convex route (‘weak negative trade-offs’) is the most probable one – large increases in the promiscuous function under selection (‘new function’) are accompanied by significantly smaller decreases in the original function (**Table 3**). By virtue of gaining a ‘new’ function without losing the original one (and often gaining other new functions not selected for), the intermediates of these routes are ‘generalists’, and their evolution can therefore proceed *prior* to gene duplication. In contrast, the concave route implies that gene duplication is a necessary prerequisite, because acquisition of even low levels of the ‘new’ function is accompanied by large losses of the original one. This route is also observed in the laboratory, in particular under a dual selection, for gain of a new function and loss of the old one.

in the active site of tyrosine ammonia lyase (TAL) switched its substrate selectivity from tyrosine to phenylalanine, thereby converting it into an enzyme whose kinetic parameters and selectivity are comparable to native PAL (phenylalanine ammonia lyase).¹⁶⁹ Indeed, in a living cell, the toll of a ‘generalist’ on fitness might be too high, and the driving force for specialization might be stronger than observed under *in vitro* selection.¹⁶⁶

The different effects of mutations on the native versus the promiscuous functions are particularly striking in view of the fact that many of these mutations are found within the active site, typically at the wall and perimeter. Structural and thermodynamic insights into the effects of these ‘generalist’ mutations are needed before any definite statements could be made. Yet it seems probable that the plasticity of these residues lies in the fact that they are not part of the protein’s scaffold, or of the catalytic machinery of the enzyme. The mutated residues are typically located on surface loops that exhibit high conformational flexibility and comprise the substrate binding part of the active site.^{86,102,103,170,151} As discussed in Section 8.03.6.3, there exist fundamental differences between the mode of binding of the native substrate – that is typically mediated by several independent, enthalpy-driven interactions such as hydrogen bonds – versus the promiscuous substrates where hydrophobic, and other entropy-driven interactions, play an important role.^{42,171} It is therefore probable that the same mutation could affect the native and the promiscuous substrates in a very different way. Although, an in-depth understanding of the effect of mutations awaits a sufficient number of cases (or even one case at this stage) in which structures become available for both the wild-type protein and its evolved mutants, in complex with analogues of both the native and promiscuous substrates.

Altogether, the above observations support the hypothesis of evolutionary progenitors and intermediates being of broad specificity, or high promiscuity,²⁵ and that, a frequent (but not exclusive) evolutionary route leads from a ‘specialist’ to a ‘generalist’, and, in turn, to a new ‘specialist’ (Figures 5 and 6). The reconstruction of evolutionary ancestors of both enzymes, and receptors, also supports the idea of ‘generalist’ progenitors.^{172,173} The implications of negative trade-offs, and of ‘generalists’ and ‘specialists’, are further discussed in Sections 8.03.8.4–8.03.8.6.

8.03.8.4 Exceptions to Weak Negative Trade-offs

Although the above strongly argues in favor of ‘weak negative trade-offs’, this generalization has notable exceptions.

8.03.8.4.1 Size and charge considerations

The magnitude of trade-offs is obviously dependent on the structural differences such as size and charge between the original and promiscuous substrates.^{96,144} For example, most reported studies involve promiscuous substrates that are larger than the native one, and cases in which both the native and the promiscuous substrates are hydrophobic. In these cases it is easy to see how a mutation that makes the active site larger, and thus increases the activity toward the promiscuous substrate, would not have a drastic effect on the native substrate. However, what about promiscuous substrates that are smaller than the native one? In this case, mutations that reshape the active site to minimize its volume and increase contacts with the smaller substrate might significantly reduce the activity with the larger native substrate. Other cases in which the native and promiscuous activities might trade-off involve differences in charge. Imagine a charged native substrate, and a neutral (let alone hydrophobic) promiscuous substrate. Mutations that favor the charged form are probable to restrict binding of the hydrophobic one, and vice versa. It is probable therefore that as the number of enzymes studied and mutations increase, more refined trade-off rules will be established.

8.03.8.4.2 Stability trade-offs

An important facet of the trade-off concerns the effect of mutations on stability. Most mutations destabilize, and mutations that affect function often exhibit even higher destabilizing effects. Destabilization may result in reduced enzyme levels, due to misfolding and aggregation, proteolytic digestion, or clearance. Thus, some of the mutations that show little effect on specific activity as measured with purified proteins *in vitro* (k_{cat} or K_M), may lower the enzyme concentration ($[E]_0$), and hence decrease the levels of enzymatic activity *in vivo*. This phenomenon was first highlighted by Wang *et al.*,²³ who studied various mutations found in clinical isolates of TEM-1 β -lactamase that evolved to degrade third-generation cephalosporin antibiotics such as cefotaxime.

The mutations that alter TEM-1's active site are all destabilizing. The key cefotaxime resistance mutation that appeared in the clinic (Gly238Ser) increases $k_{\text{cat}}/K_{\text{M}}$ by 86-fold (Wang *et al.*²³, **Table 3**, entry 11), and only reduces $k_{\text{cat}}/K_{\text{M}}$ for ampicillin (a native substrate of TEM-1) by sixfold. However, this mutation has a significant detrimental effect on bacterial growth under ampicillin due to destabilization of the enzyme by ~ 2 kcal mol⁻¹, and a significant reduction in the levels of soluble, active TEM-1. At later stages, the loss of stability was compensated by Met182Thr, which increases TEM-1 stability by 2.7 kcal mol⁻¹.²³

Following the TEM-1 analysis, Wang *et al.*²³ suggested that, in general, mutations that endow a new or improved function, trade-off with protein stability. Thus, for the evolutionary process to continue (in nature or in the laboratory) this loss of stability must be compensated. A more recent study was based on computational predictions of the stability changes for >500 mutations that arose from the directed evolution of 22 different enzymes.¹⁷⁴ The stability effects of function-altering mutations ($\Delta\Delta G$ values) were compared to stability changes arising from any random point mutation in the same enzymes. It was found that, as was the case with TEM-1, mutations that modulate enzymatic functions are mostly destabilizing (average $\Delta\Delta G = +1.1$ kcal mol⁻¹). Although the effects of function-altering mutations are actually not more destabilizing than the 'average' mutation in these enzymes (+1.3 kcal mol⁻¹), these mutations weaken stability. The analysis also indicated that many mutations that appear in directed evolution variants with no obvious role in the new function exert stabilizing effects that may compensate for the destabilizing effects of the crucial function-altering mutations. Thus, despite a lack of a specific trade-off between function and stability, the evolution of new enzymatic activities, both in nature and in the laboratory, is dependent on the compensatory, stabilizing effect of apparently 'silent' mutations in regions of the protein that are irrelevant to its function. Indeed, other works showed that limited protein stability constrains the acquisition of new function,¹⁷⁵ and highlighted the role of stability compensators such as Met182Thr of TEM-1 in expediting the evolution of new functions.¹⁷⁶

Thus, reductions in activity in terms of k_{cat} and K_{M} (**Tables 3 and 4**) may not reflect the full impact of these mutations. Loss of stability is also part of the trade-off that may accompany a gain in the evolving promiscuous activity.

8.03.8.4.3 Targeted versus random mutagenesis

It is also the case that mutations incorporated through rational design show larger trade-offs relative to mutations obtained by selection from random repertoires (**Table 4**, in comparison with **Table 3**). This is not surprising given that rational design usually aims at the replacement of key active-site residues. Exchanges in such key residues yield drastic changes, including dramatic enhancements of a promiscuous function at the expense of the native one. However, mutations isolated in directed evolution rarely occur in key active-site residues, and they typically exhibit more subtle effects. It is probable that key active-site mutations that yield more drastic changes occur at later stages, typically at the last stages of divergence when relatively small improvements in the evolving function trade-off with larger drops in the old one (**Figure 6**, convex route). Nevertheless, some of the engineered mutations (e.g., **Table 4**, entries 2, 6, 7, 10, 11) clearly reflect changes that may lead, or may have actually led, to the divergence of a new function through the strong negative trade-off route (**Figure 6**, concave route).

8.03.8.5 On- and Off-Pathway Evolutionary Intermediates

A notable case of large negative trade-offs was described for the directed evolution of HisA to yield TrpF activity (**Table 3**, entry 15; **Figure 7**). The gene was randomly mutated, and the Asp127Val mutant of *Thermotoga maritima* HisA isolated from the selection exhibited measurable TrpF activity ($k_{\text{cat}}/K_{\text{M}} = 120$ mol⁻¹ s⁻¹) that was sufficient to complement the *E. coli* TrpF knockout strain used for the selection. The newly evolving TrpF activity led to a dramatic drop in the original HisA activity ($\sim 10^4$ -fold). Perhaps this effect is not so surprising given that the starting point had no measurable TrpF activity, and that the mutation occurred in a key active-site residue: Asp127 is the putative acid catalyst in the Amadori rearrangement catalyzed by HisA.¹⁸⁵

Interestingly, few years later after this directed evolution experiment, a bi-functional enzyme dubbed *PriA* was discovered that performs both reactions with high efficiency, and within the very same active site^{186, 187}. In

Table 4 Examples for trade-offs in enzymatic functions following targeted mutations

Enzyme	Approach	Native activity (k_{cat}/K_M of wild type, $\text{mol}^{-1}\text{I s}^{-1}$)	Promiscuous/target activity (k_{cat}/K_M of wild type, $\text{mol}^{-1}\text{I s}^{-1}$)	Incorporated mutations	Effect on native activity ($k_{cat}/K_M^{\text{variant}} / k_{cat}/K_M^{\text{wt}}$)	Effect on the evolved promiscuous activity ($k_{cat}/K_M^{\text{variant}} / k_{cat}/K_M^{\text{wt}}$)	Comments	Reference(s)
1 <i>HaeIII</i> methyltransferase	Semi-rational (saturation mutagenesis of 8 positions in the target recognition domain, based on crystal structure)	Methylation of GGCC sites ($26\,000\ \text{mol}^{-1}\text{I s}^{-1}$)	Methylation of AGCC sites ($1180\ \text{mol}^{-1}\text{I s}^{-1}$)	Arg225Ala, Asn260Leu, Leu261Met, Asn262Trp	Ninefold higher	670-fold higher	With the exception of Arg225, the mutated positions do not seem to make direct DNA contact.	76
2 tHisA and tHisF from <i>Termatoga maritima</i>	Semi-rational (saturation mutagenesis of Asp127 in HisA and Asp130 in HisF, based on previous results)	HisA: ProFAR isomerase (Amadori re-arrangement). HisF: imidazole glycerol phosphate synthase.	Isomerization of phosphoribosyl-anthranilate (TrpF activity) No detectable promiscuous activity	HisA: Asp127Val, Thr164His HisF: Asp130Val/ Thr/ Pro	Not tested	The wild-type activity is below detection limits	A Val residue at position 127 in HisA, and Val, Thr, or Pro residues at position 130 in HisF, are not compatible with HisA or HisF native activities, which require a negatively charged residue at that position	177
3 Aspartate aminotransferase from <i>E. coli</i> (eAATase)	Semi-rational: combination of seven of the most frequent mutations when selecting for aromatic specificity (Rothman <i>et al.</i> , ¹⁷⁹), plus Ala293Asp	Transamination of dicarboxylic substrates ($9.1\ \text{mol}^{-1}\text{I s}^{-1}$)	Transamination of phenylalanine (eTATase activity, $0.012\ \text{mol}^{-1}\text{I s}^{-1}$)	Ala12Thr, Pro13Thr, Asn34Asp, Thr109Ser, Gly261Ala, Ser285Gly, Ala293Asp, Asn297Ser	40-fold lower	280-fold higher	The selected mutations not only alter the direct contact with the ligand, but also promote flexibility of large regions of the active site.	178

Author's personal copy

4	Leucine aminopeptidase from <i>Streptomyces streptatus</i>	Semi-rational (saturation mutagenesis at position 221, based on predicted structure)	Hydrolysis of L-Leu peptides (or L-Leu-pNA derivative, $1.8 \times 10^5 \text{ mol}^{-1} \text{ s}^{-1}$)	Hydrolysis of L-Phe peptides (or L-Phe-pNA derivative, $5 \times 10^4 \text{ mol}^{-1} \text{ s}^{-1}$)	Phe221Ala	3.7-fold lower	10.2-fold higher	Residue at position 221 was predicted to interact with the side chain of the substrate.	180
5	L-Ala-D/L-Glu epimerase (AEE)	Rational design based on structural alignment with α -succinylbenzoate synthase (OSBS, 30% sequence identity)	1,1-Proton transfer ($7.7 \times 10^4 \text{ mol}^{-1} \text{ s}^{-1}$)	β -Elimination (OSBS, $<5 \times 10^{-3} \text{ mol}^{-1} \text{ s}^{-1}$)	Asp297Gly	7800-fold lower	>2400-fold higher	Large effects of single mutations	85, 144, 181
					Asp297Gly Ile19Phe	>790 000-fold lower	>17 000-fold higher		
					Asp297Gly Ile19Phe Arg24Trp Tyr265Ala	> 10^6 -fold lower	>404 000-fold higher		
6	Alanine racemase from <i>G. stearothermophilus</i>	Rational design (based on structure and proposed mechanism)	Isomerase	Aldolase ($5 \times 10^{-5} \text{ mol}^{-1} \text{ s}^{-1}$)		3000-fold lower	2.3×10^5 -fold higher	A single point mutation can change both substrate specificity and reaction profile simultaneously	182
7	Lipase B from <i>C. antarctica</i>	Rational design (molecular dynamic simulations and automated docking), based on a previously described mutant (Branneby <i>et al.</i> ⁷⁸).	Hydrolysis of triglyceride esters	Michael-type addition of thiols to α,β -unsaturated carbonyl compounds	Ser105Ala	Not tested, but probable to be > 10^3 -fold lower	6–1600-fold higher (k_{cat}) relative to wild-type, for the addition of thiols to methyl acrylate.		79

(Continued)

Table 4 (Continued)

Enzyme	Approach	Native activity (k_{cat}/K_M of wild type, $\text{mol}^{-1} \text{ l s}^{-1}$)	Promiscuous/target activity (k_{cat}/K_M of wild type, $\text{mol}^{-1} \text{ l s}^{-1}$)	Incorporated mutations	Effect on native activity ($k_{cat}/K_M^{variant}/k_{cat}/K_M^{wt}$)	Effect on the evolved promiscuous activity ($k_{cat}/K_M^{variant}/k_{cat}/K_M^{wt}$)	Comments	Reference(s)	
8	3-Keto-L-gulonate 6-phosphate decarboxylase (KGPDC) from <i>E. coli</i> .	Rational design by sequence alignment with D-arabino-hex-3-ulose 6-phosphate synthase (HPS, 30% identity)	Decarboxylation of 3-keto-L-gulonate 6-phosphate ($7.7 \times 10^4 \text{ mol}^{-1} \text{ l s}^{-1}$)	Aldol condensation of D-ribulose 5-phosphate with formaldehyde ($0.082 \text{ mol}^{-1} \text{ l s}^{-1}$)	Glu112Asp Arg139Val Thr169Ala	30-fold lower	260-fold higher		133
9	Human GST A2-2	Rational design (based on homology with GST A3-3)	Glutathione peroxidase (2.9×10^4 with cumene peroxide)	Steroid double bond isomerase (1×10^3 with Δ^5 -androstene-3,17-dione)	Ser10Phe Ile12Gly Phe111Leu Met208Ala Ser216Ala	Twofold lower	3500-fold higher	Restricted to substrate-binding residues	183
10	Rat liver 3 α -hydroxy-steroid dehydro-genase (HSD) – AKR1C9	Rational design (based on homology with steroid 5 β -reductase)	Position and stereo-specific interconversion of steroid ketones and alcohols (1.6×10^5)	Steroid 5 β -reductase (reduction of C–C double bonds), no detectable promiscuous activity	His117Glu	>600-fold lower	The wild type activity is below detection limits	Single-point mutation of catalytic residue	184
11	Tyrosine ammonia lyase (TAL) from <i>Rba. sphaeroides</i>	Rational design by homology with phenylalanine ammonia lyase (PAL)	Deamination of tyrosine to 4-coumaric acid and ammonia (1.1×10^5)	Deamination of phenylalanine to cinnamic acid and ammonia (403)	His89Phe	>19 000-fold lower	220-fold higher	Single-point mutation of catalytic residue that completely switches specificity	169
12	N-acetyl -neuraminate lyase (NAL)	Rational design (by homology with DHDPS)	Cleavage of N-acetyl-neuraminate to produce pyruvate and N-acetyl mannosamine (3.1×10^3)	Aldol condensation of pyruvate and L-aspartate – β -semialdehyde (DHDPS activity)	Leu142Arg Tyr190Asp Glu192Ala	62-fold lower	Sixfold higher	Despite very modest improvement, a single mutation Leu142Arg is enough for <i>in vivo</i> complementation	134

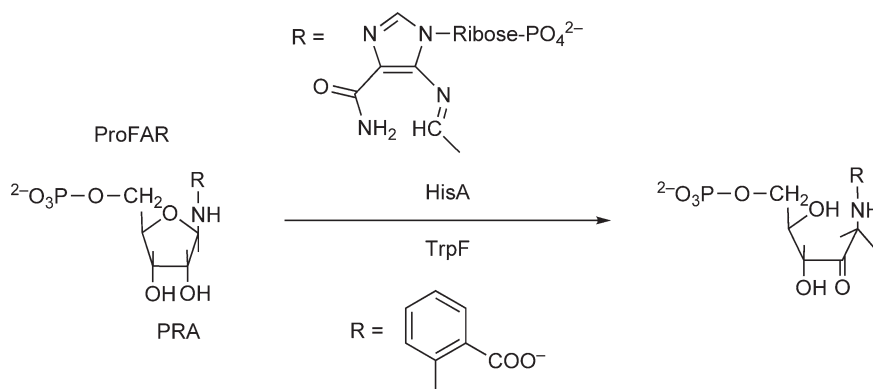


Figure 7 The reactions catalyzed by HisA (isomerization of *N'*-[(5'-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide) and TrpF (isomerization of phosphoribosyl-anthranilate).

contrast to the laboratory selection for TrpF activity, in nature, a generalist enzyme evolved under selection to maintain both the HisA and TrpF functions. Thus, the 'generalist' intermediate originally proposed by Jurgens *et al.*¹⁵² exists in certain bacteria. However, it seems that the Asp127Val mutant isolated in the laboratory evolution experiment does not reflect the sequence and structural features of such a generalist intermediate.

In a recent directed evolution experiment, a HisA Asp127Val mutant was further evolved for TrpF activity¹⁸⁸. However, it led to a complete change in the reaction mechanism, and the evolved variants completely lost the original HisA activity. Thus, although this mutant provides a clear example of how TrpF activity could emerge in HisA, it seems to comprise an 'off-pathway' intermediate – namely, an intermediate that provides a temporary advantage but cannot lead to the eventual divergence of a proficient bifunctional TrpF-HisA enzyme. Off-pathway intermediates might be observed in nature (for example, see the E3 esterase example discussed in Section 8.03.8.6.3). They are as likely, if not more likely, to appear in laboratory evolution experiments, and even more so in rationally designed enzyme variants. It could may well be that some of the 'generalist' intermediates observed in laboratory evolution experiments are also 'off-pathway'. Indeed, the ultimate proof for 'on-pathway' evolutionary intermediates lies in the ability to complete the divergence process and generate a new 'specialist' enzyme with native-like kinetic parameters. Completing the process, however, involves many rounds of mutation and selection, and numerous mutations. This is very rarely pursued, and when it is, the aim is the final enzyme product, and not the pathway leading to it^{168,189,190}. Nonetheless, these applicative engineering projects have the potential to provide interesting insights into the pathway, the role and order of individual mutations, and the nature of the intermediates.

8.03.8.6 Promiscuity and the Mechanisms for the Divergence of New Gene Functions

8.03.8.6.1 Ohno's model

The mechanisms governing the divergence of new gene/protein functions comprise a central part of evolutionary theory. Early models that have become known as Ohno's model,¹⁹¹ were later expanded by Kimura and Ohta,¹⁹² and are currently the textbook paradigm. This model assumes that duplication is a frequent event, which is largely neutral – that is, duplication provides no fitness advantage, or disadvantage, and is therefore not under selection. The redundant duplicated copy is therefore free to accumulate mutations, including deleterious ones. If and when the need arises, some of these mutations that endow a new function come under positive, adaptive selection, thus leading to the divergence of the new gene and function (neo-functionalization).

The fact that duplication and relief from selection is a prerequisite stems from the negative trade-off assumption, namely, maintaining selection for the existing function is expected to purge mutations with adaptive potential. However, as noted in Section 8.03.8.3, in many cases, a promiscuous function can further evolve with little effect on the original function. Indeed, the underlining assumptions of Ohno's model differ fundamentally from what is described in the previous sections. Contrary to the notion that promiscuous

activities can provide an immediate advantage (Section 8.03.8.2), in Ohno's model, the original gene plays a completely passive role in the emergence of a new function.

Indeed, the part of Ohno's model surmises that gene duplication and the subsequent mutational drift occur under no selection is being intensively questioned. First, in contrast to Ohno's model, most duplicated genes found in existing genomes appear to drift under functional selection that purges deleterious mutations.^{193–195} Second, expression of redundant copies carries substantial energetic costs,^{34,44} and there exists a strong selection pressure to inactivate their expression.^{43,196} Third, as discussed in Section 8.03.8.6.4, many cases have been recorded in which gene duplication is not a neutral event, but is rather positively selected under demands for higher protein doses.^{197,198} Finally, about a third of random mutations in a given protein are deleterious,^{159–161} whereas beneficial mutations that can promote new functions are very rare. Thus, when drifting in the absence of any selection, loss of all functions (nonfunctionalization) primary due to mutations that diminish gene expression, or the ability to form a stable fold,^{161,199} is orders of magnitude more probable than neo-functionalization. In the sections that follow we describe alternative mechanisms of divergence that are based on promiscuity and its evolutionary features.

8.03.8.6.2 Gene sharing

The first evidence indicating lack of trade-off, and thus the emergence of a new protein function prior to, or even without, gene duplication, led to the hypothesis of 'gene sharing', by which, a gene with a given function is recruited for a very different function without significant changes in the coding region.²⁰⁰ Gene sharing is a feasible event, as indicated by the classical example of crystallins whereby metabolic enzymes (e.g., argininosuccinate lyase) were recruited later in evolution as structural proteins in eye lenses, with no sequence changes and while retaining their enzymatic activity.^{69,200} This is certainly not a singular example. Several other examples for the recruitment of a protein, for example, an enzyme, for completely different tasks, and under different regulation regimes are currently known.⁷ Secondary roles of aminoacyl-tRNA synthetases are mentioned in Section 8.03.2. Another notable example is thymidine phosphorylase that also acts as an endothelial growth factor. Indeed, inhibitors of this enzyme also inhibit cell growth.^{67,68}

Piatigorsky's findings opened the door for a reconsideration of Ohno's model, but at the same time, it is clear that in most cases, a new function must eventually trade-off with the existing one (Figure 6). Hence, the 'gene sharing' model was extended to include cases in which divergence is dependent on duplication.

These alternative models are detailed in the sections that follow. They all include duplication. However, they fundamentally differ from Ohno's model in that duplication is not a neutral event, but it rather provides an immediate fitness advantage. Consequently, both gene copies (the original and the newly duplicated gene) are maintained under selection throughout the process.

8.03.8.6.3 Divergence prior to duplication

This model, which comprises an extension of 'gene sharing', gathers growing levels of support.^{90,156,193} The 'divergence prior to duplication' model, or IAD (innovation–amplification–divergence)¹⁹⁷ model, assumes that the very first step toward divergence is the selection of a mutant enzyme (an allele) with higher secondary, promiscuous activity. Using this model, the mutant appeared in the population well before the need for a new function had appeared, primarily because the mutation had little, or no, effect on the existing function of this enzyme. Thus, a mutation that initially accumulated as neutral becomes adaptive if and when a change is required. The 'generalist' nature of these apparently neutral mutants is such that duplication may become necessary only at the later stages of further divergence and the emergence of a new specialist. If mutations with adaptive potential accumulate prior to duplication, as neutral (namely with no, or little effect on fitness under a current state), adaptive events become more frequent.

The 'divergence prior to duplication' model is strongly supported by the notion of neutral networks. Theoretical, computational, and experimental works indicate that a drift under selection to maintain the existing function and structure (a 'neutral drift') can increase the potential for adaptation (see Means and Bender⁹³, Smith,¹⁴² Nei,²⁰¹ and Wroe *et al.*²⁰² and references therein). The existence of this latent pleiotropy – a range of promiscuous functions that were neither selected for, nor against (and conformational isomers that can mediate these promiscuous activities) – facilitates evolution by providing ample starting points for new functions while retaining the primary activity.

An intriguing example for the feasibility of 'latent adaptations' came from the recent adaptation of the sheep blowfly *Lucilia cuprina* to organophosphate (OP) pesticides.^{16,18} Early in the twentieth century, OP resistance in blowflies evolved through two separate pathways. Initially, the introduction of OPs resulted in the rapid enrichment of the Trp251Leu mutation in carboxylesterase E3 that endowed this enzyme with weak OP hydrolase activity, and thus led to pesticide resistance. Sequencing of blowfly specimens preserved from the period before the introduction of OPs revealed that the Trp251Leu mutation was already present in the population at high frequency. Indeed, this mutation allowed retention of the native esterase activity while improving the promiscuous OP-hydrolase activity. Interestingly, after several years, a second mutation arose (Gly137Asp) that rapidly superseded Trp251Leu in OP-treated populations. In contrast to the Trp251Leu mutation, Gly137Asp leads to a very significant reduction in the esterase activity, and is associated with developmental defects that were later relieved by a suppressor mutation elsewhere in the blowfly genome. Future experiments may reveal whether these two mutations can be combined to yield even higher levels of resistance, or perhaps the neutral Trp251Leu mutation comprises an 'off-pathway' mutation that cannot lead to an enzyme variant with much higher rates of OP hydrolysis (see Section 8.03.8.5).

Recent laboratory experiments followed the notion of a 'neutral drift' by placing an enzyme under mutation and selection to maintain its native function. The data provide empirical evidence in support of the hypothesis that neutrality enables the formation of latent changes, or 'latent adaptation'. It was found^{93,203} that latent evolutionary potentials are indeed very frequent within a neutral set of related enzyme mutants, and that these potentials are most often seen as changes in specificity for one or more promiscuous substrates.

The initial manifestation of such 'latent adaptation' might be in providing an immediate selective advantage by expanding the range of activities of existing enzymes. As demonstrated by the E3 esterase case, once a latent promiscuous function has become advantageous due to a change in the environment, neutral mutants in which this function is higher can be rapidly selected. Duplication may follow, and enable further divergence toward a completely new function. The acquisition of neutral mutations therefore shortens the adaptation and enables facile transitions between one function to another (Figure 8).

8.03.8.6.4 Duplication is positively selected

Contrary to the above-described model (Section 8.03.8.6.3), in certain cases, duplication may precede divergence. However, contrary to Ohno's model which assumes that duplication is a neutral event, duplication can be under positive selection. Specifically, when divergence capitalizes on a minor or promiscuous activity in an existing protein, immediate selective advantage can be provided by increasing protein doses.⁵⁰ Duplication is known to occur under such selection pressures, and under this scenario, both duplicates will be maintained under selection to maintain both the primary function and the minor one.¹⁹⁷

8.03.8.6.5 Subfunctionalization

By virtue of increasing enzyme doses, duplication can also have a key role in enabling a wider variety of function altering mutations to accumulate. Despite the generally weak trade-offs, at the end of the day, mutations that endow new enzymatic functions have a measurable effect on the existing enzymatic function, and on the enzyme's stability. In particular, mutations that alter enzymatic functions tend to be destabilizing,^{23,174} and can thus reduce the levels of soluble active enzyme (Section 8.03.8.6.4). Indeed, in the neutral drift experiment described in Section 8.03.8.6.3 and Figure 8, most neutral variants tested *in vitro* exhibited specificity changes (namely, higher ratios of the promiscuous relative to the native activity), and not necessarily improvements in the absolute levels of promiscuous activity. This indicates that the mutations that improved these activities in terms of k_{cat} or K_M values for the promiscuous function, may have also reduced expression ($[E]_0$) and hence the total activity (v_0) remained largely unchanged. Even if the effects on enzyme stability, expression levels, and catalytic efficiency in relation with the native activity, are minor relative to the potential innovation benefit, the acquisition of beneficial mutations can only continue as long as the existing function is reduced to an extent that does not compromise organismal fitness. Following that not all genes contribute to organismal fitness at all times, long periods of relaxation in the selection pressure is a conceivable scenario. During such periods, mutations exhibiting weak trade-offs can accumulate. An alternative solution to this problem is gene duplication, provided that, contrary to Ohno's model, the two genes remain under selection. By virtue of the two genes carrying the same level of function, duplication can offer a margin that allows a wider variety of potentially beneficial mutations to accumulate.

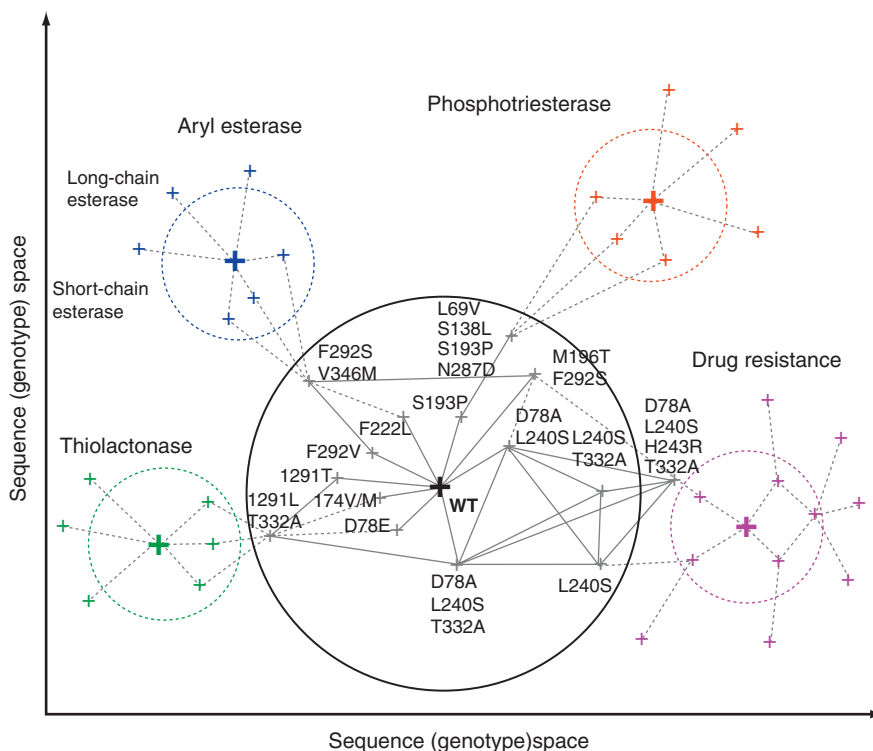


Figure 8 Schematic representation of a putative neutral network of an enzyme. The scheme is based on the results of a ‘neutral drift’ of serum paraoxonase (PON1) performed in the laboratory, and PON1 mutants identified in various directed evolution experiments. The mutations of neutral variants with significantly altered phenotypes are noted. The large circle denotes the hypothetical boundaries of the neutral network of PON1’s native phenotype (lipo-lactonase). The length of the edges corresponds to the degree of neutrality – variants that are more distant from the center (WT) vary in their phenotype to a larger degree. Connecting edges illustrate common mutations between two or more nodes. The other neutral regions relate to new phenotypes that can stem from PON1 (e.g., esterase, phosphotriesterase, and a ‘drug-resistant PON1’). The dashed nodes and edges are hypothetical, whereas the continuous ones relate to the apparently neutral variants that were characterized in the experiment. Adapted from G. Amitai; R. Devi-Gupta; D. S. Tawfik, *HFSP J.* **2007**, *1*, 67–78.

Indeed, as previously observed, once duplication has occurred, both copies are probable to be maintained under purifying selection. This may occur in the subfunctionalization (or DDC; duplication–degeneration–complementation) model.^{204,205} Following duplication, one, or even both copies may acquire loss-of-function mutations, such that both genes are now required to maintain the level of function provided by the single ancestral gene. Although this model was devised for complex gene functions and regulatory elements, it can be readily extended to simple enzymatic functions. Indeed, a study of yeast genome duplications indicated that the relaxation in selection pressure afforded by redundant gene copies led to the increased accumulation of activity-reducing mutations (ARMs).¹⁹⁵ The appearance of these mutations enforces both copies to be maintained viable as in the DDC model. However, whereas the DDC model describes loss-of-function mutations that can be very difficult to regain (such as mutations that knockout an entire protein domain), the ARMs are dose-dependent mutations. These mutations decrease enzyme dose and activity, and are easily compensated by the enzymatic activity of the second gene copy. As discussed here, most mutations that increase an existing, latent promiscuous activity, and may therefore possess adaptive potential, belong to the category of ARMs in the sense that they reduce in one way or another, but do not abolish, the enzyme’s primary activity. Duplication, followed by the fixation of both copies under purifying selection, can provide the key to rapid divergence owing to the fact that duplication allows, and is sustained, by such ARMs.

Acknowledgments

We acknowledge research funds granted by the Israel Science Foundation and the NIH Contract No. W81XWH-07-2-0020 for the study.

Abbreviations

4-OT	4-oxalocrotonate tautomerase
AEE	L-Ala-D/L-Glu epimerase
Ap4A	two adenosines linked through four phosphates
ARM	activity-reducing mutation
CaaD	<i>trans</i> -3-chloroacrylic acid dehalogenase
CALB	<i>Candida antarctica</i> lipase B
DAOCS	deacetoxycephalosporin C synthase
DDC	duplication–degeneration–complementation model
DFPase	diisopropyl fluorophosphatase
DHDPS	dihydrodi-picolinate synthase
DHO	dihydroorotase
eAATase	aspartate aminotransferase
GADPH	glyceraldehyde-3-phosphate dehydrogenase
GMSF	guanidino-modifying enzyme superfamily
GST	glutathione S-transferase
GUS	β -glucuronidase
HAD	superfamily haloacid dehalogenase-like hydrolases
HisA	ProFAR isomerase
HisF	imidazole glycerol phosphate synthase
HPS	D- <i>arabino</i> -hex-3-ulose 6-phosphate synthase
HSD	3 α -hydroxy-steroid dehydrogenase
KDGA	D-2-keto-3-deoxy-gluconate aldolase
KGPDS	3' keto L-gluconate 6-phosphate decarboxylase
MLE	muconate lactonizing enzyme
MSAD	malonate semialdehyde decarboxylase
NAL	<i>N</i> -acetyl-neuraminic lyase
NPP	nucleotide pyrophosphatase
OMPDC	orotidine 5' monophosphate decarboxylase suprafamily
OP	organophosphate
OSBS	<i>o</i> -succinylbenzoate synthase
PaADI	arginine deiminase
PaAgDI	agmatine deiminase
PaDDAH	<i>N</i> ^w , <i>N</i> ^w -dimethyl-arginine dimethyl-aminohydrolase
PAL	phenylalanine ammonia lyase
PchB	isochorismate pyruvate-lyase
PHP	phosphotriesterase homology protein
PLL	PTE-like lactonase
PON1	serum paraoxonase
PTE	phosphotriesterase
TAL	tyrosine ammonia lyase
TrpF	phosphoribosyl-anthranilate isomerase
WT	wild type

References

1. K. Hult; P. Berglund, *Trends Biotechnol.* **2007**, *25*, 231–238.
2. Y. Pocker; J. T. Stone, *J. Am. Chem. Soc.* **1965**, *87*, 5497–5498.
3. T. W. Reid; D. Fahrney, *J. Am. Chem. Soc.* **1967**, *89*, 3941–3943.
4. Y. Nakagawa; M. L. Bender, *J. Am. Chem. Soc.* **1969**, *91*, 1566–1567.
5. R. C. Jackson; R. E. Handschumacher, *Biochemistry* **1970**, *9*, 3585–3590.
6. P. J. O'Brien; D. Herschlag, *Chem. Biol.* **1999**, *6*, R91–R105.
7. S. D. Copley, *Curr. Opin. Chem. Biol.* **2003**, *7*, 265–272.
8. U. T. Bornscheuer; R. J. Kazlauskas, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 6032–6040.
9. R. J. Kazlauskas, *Curr. Opin. Chem. Biol.* **2005**, *9*, 195–201.
10. C. A. Bonner; T. Disz; K. Hwang; J. Song; V. Vonstein; R. Overbeek; R. A. Jensen, *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 13–53, table of contents.
11. M. E. Glasner; J. A. Gerlt; P. C. Babbitt, *Curr. Opin. Chem. Biol.* **2006**, *10*, 492–497.
12. P. J. O'Brien, *Chem. Rev.* **2006**, *106*, 720–752.
13. D. R. Palmer; J. B. Garrett; V. Sharma; R. Meganathan; P. C. Babbitt; J. A. Gerlt, *Biochemistry* **1999**, *38*, 4252–4258.
14. O. Khersonsky; C. Roodveldt; D. S. Tawfik, *Curr. Opin. Chem. Biol.* **2006**, *10*, 498–508.
15. M. D. Toscano; K. J. Woycechowsky; D. Hilvert, *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 3212–3236.
16. C. J. Hartley; R. D. Newcomb; R. J. Russell; C. G. Yong; J. R. Stevens; D. K. Yeates; J. La Salle; J. G. Oakeshott, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8757–8762.
17. D. B. Janssen; I. J. Dinkla; G. J. Poelarends; P. Terpstra, *Environ. Microbiol.* **2005**, *7*, 1868–1882.
18. R. D. Newcomb; P. M. Campbell; D. L. Ollis; E. Cheah; R. J. Russell; J. G. Oakeshott, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7464–7468.
19. F. M. Raushel; H. M. Holden, *Adv. Enzymol. Relat. Areas Mol. Biol.* **2000**, *74*, 51–93.
20. L. P. Wackett, *J. Biol. Chem.* **2004**, *279*, 41259–41262.
21. M. Barlow; B. G. Hall, *J. Mol. Evol.* **2002**, *55*, 314–321.
22. B. G. Hall, *Nat. Rev. Microbiol.* **2004**, *2*, 430–435.
23. X. Wang; G. Minasov; B. K. Shoichet, *J. Mol. Biol.* **2002**, *320*, 85–95.
24. D. M. Weinreich; N. F. Delaney; M. A. Depristo; D. L. Hartl, *Science* **2006**, *312*, 111–114.
25. R. A. Jensen, *Annu. Rev. Microbiol.* **1974**, *30*, 409–425.
26. A. Fersht, *Structure and Mechanism in Protein Science*; W. H. Freeman and Company: New York, 1999.
27. O. Kotik-Kogan; N. Moor; D. Tworowski; M. Safo, *Structure* **2005**, *13*, 1799–1807.
28. H. Yang; G. Zheng; X. Peng; B. Qiang; J. Yuan, *FEBS Lett.* **2003**, *552*, 95–98.
29. L. Reshetnikova; N. Moor; O. Lavrik; D. G. Vassilyev, *J. Mol. Biol.* **1999**, *287*, 555–568.
30. J. Piatigorsky, *Gene Sharing and Evolution: The Diversity of Protein Functions*; Harvard University Press: Cambridge, MA; London, 2007.
31. A. Wagner, *FEBS Lett.* **2005**, *579*, 1772–1778.
32. A. Brevet; P. Plateau; B. Cirakoglu; J. P. Pailliez; S. Blanquet, *J. Biol. Chem.* **1982**, *257*, 14613–14615.
33. Y. N. Lee; H. Nechushtan; N. Figov; E. Razin, *Immunity* **2004**, *20*, 145–151.
34. A. Wagner, *Mol. Biol. Evol.* **2005**, *22*, 1365–1374.
35. G. A. Michaud; M. Salcius; F. Zhou; R. Bangham; J. Bonin; H. Guo; M. Snyder; P. F. Predki; B. I. Schweitzer, *Nat. Biotechnol.* **2003**, *21*, 1509–1512.
36. J. M. Varga; G. Kalchschmid; G. F. Klein; P. Fritsch, *Mol. Immunol.* **1991**, *28*, 655–659.
37. J. M. Varga; G. Kalchschmid; G. F. Klein; P. Fritsch, *Mol. Immunol.* **1991**, *28*, 641–654.
38. J. Inman, The Antibody Combining Region: Speculation on the Hypothesis of General Multispecificity. In *Theoretical Immunology*; G. I. Bell, A. S. Perelson; G. H. Pimbley, Jr., Eds.; Marcel Dekker, Inc.: New York, 1978.
39. D. Lancet; E. Sadovsky; E. Seidemann, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3715–3719.
40. A. S. Perelson; G. F. Oster, *J. Theor. Biol.* **1979**, *81*, 645–670.
41. A. D. Griffiths; D. S. Tawfik, *Curr. Opin. Biotechnol.* **2000**, *11*, 338–353.
42. L. C. James; D. S. Tawfik, *Protein Sci.* **2003**, *12*, 2183–2193.
43. E. Dekel; U. Alon, *Nature* **2005**, *436*, 588–592.
44. D. M. Stoebel; A. M. Dean; D. E. Dykhuizen, *Genetics* **2008**, *178*, 1653–1660.
45. E. Kuznetsova; M. Proudfoot; C. F. Gonzalez; G. Brown; M. V. Omelchenko; I. Borozan; L. Carmel; Y. I. Wolf; H. Mori; A. V. Savchenko; C. H. Arrowsmith; E. V. Koonin; A. M. Edwards; A. F. Yakunin, *J. Biol. Chem.* **2006**, *281*, 36149–36161.
46. L. W. Tremblay; D. Dunaway-Mariano; K. N. Allen, *Biochemistry* **2006**, *45*, 1183–1193.
47. K. Yang; W. W. Metcalf, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7919–7924.
48. R. D'Ari; J. Casadesus, *Bioessays* **1998**, *20*, 181–186.
49. A. Macchiarulo; I. Nobeli; J. M. Thornton, *Nat. Biotechnol.* **2004**, *22*, 1039–1045.
50. W. M. Patrick; E. M. Quandt; D. B. Swartzlander; I. Matsumura, *Mol. Biol. Evol.* **2007**, *24*, 2716–2722.
51. A. Kurakin, *J. Mol. Recog.* **2007**, *20*, 205–214.
52. R. J. Greenspan, *Nat. Rev. Genet.* **2001**, *2*, 383–387.
53. B. Van Swinderen; R. J. Greenspan, *Genetics* **2005**, *169*, 2151–2163.
54. R. Bernhardt, *J. Biotechnol.* **2006**, *124*, 128–145.
55. M. Ekroos; T. Sjogren, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13682–13687.
56. K. E. Griswold; N. S. Aiyappan; B. L. Iverson; G. Georgiou, *J. Mol. Biol.* **2006**, *364*, 400–410.
57. L. Hou; M. T. Honaker; L. M. Shireman; L. M. Balogh; A. G. Roberts; K. C. Ng; A. Nath; W. M. Atkins, *J. Biol. Chem.* **2007**, *282*, 23264–23274.
58. A. Nath; W. M. Atkins, *Biochemistry* **2008**, *47*, 157–166.

59. S. Lodeiro; Q. Xiong; W. K. Wilson; M. D. Kolesnikova; C. S. Onak; S. P. Matsuda, *J. Am. Chem. Soc.* **2007**, *129*, 11213–11222.
60. C. L. Steele; J. Crock; J. Bohlmann; R. Croteau, *J. Biol. Chem.* **1998**, *273*, 2078–2089.
61. S. Pilkington; H. Dalton, Soluble Methane Monooxygenase from *Methylococcus Capsulatus* Bath. In *Methods in Enzymology*; M. Lindstrom, Ed.; Academic Press: San Diego, 1990; Vol. 188, pp 181–190.
62. M. Konarzycka-Bessler; U. T. Bornscheuer, *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 1418–1420.
63. M. T. Reetz, *Curr. Opin. Chem. Biol.* **2002**, *6*, 145–150.
64. G. Hasnaoui-Dijoux; M. Majeric Elenkov; J. H. Lutje Spelberg; B. Hauer; D. B. Janssen, *ChemBioChem* **2008**, *9*, 1048–1051.
65. J. Domenech; J. Ferrer, *Biochim. Biophys. Acta* **2006**, *1760*, 1667–1674.
66. C. Jeffrey, *Trends Biochem. Sci.* **1999**, *24*, 8–11.
67. M. Haraguchi; K. Miyadera; K. Uemura; T. Sumizawa; T. Furukawa; K. Yamada; S. Akiyama; Y. Yamada, *Nature* **1994**, *368*, 198.
68. S. Liekens; A. Bronckaers; M. J. Perez-Perez; J. Balzarini, *Biochem. Pharmacol.* **2007**, *74*, 1555–1567.
69. J. Piatigorsky; W. E. O'Brien; B. L. Norman; K. Kalumuck; G. J. Wistow; T. Borrás; J. M. Nickerson; E. F. Wawrousek, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3479–3483.
70. A. Allali-Hassani; P. W. Pan; L. Dombrowski; R. Najmanovich; W. Tempel; A. Dong; P. Loppnau; F. Martin; J. Thornton; A. M. Edwards; A. Bochkarev; A. N. Plotnikov; M. Vedadi; C. H. Arrowsmith, *PLoS Biol.* **2007**, *5*, e97.
71. B. G. Miller; R. T. Raines, *Biochemistry* **2004**, *43*, 6387–6392.
72. B. G. Miller; R. T. Raines, *Biochemistry* **2005**, *44*, 10776–10783.
73. S. C. Rothman; J. F. Kirsch, *J. Mol. Biol.* **2003**, *327*, 593–608.
74. I. Matsumura; A. D. Ellington, *J. Mol. Biol.* **2001**, *305*, 331–339.
75. L. A. Rowe; M. L. Geddie; O. B. Alexander; I. Matsumura, *J. Mol. Biol.* **2003**, *332*, 851–860.
76. H. M. Cohen; D. S. Tawfik; A. D. Griffiths, *Protein Eng. Des. Sel.* **2004**, *17*, 3–11.
77. E. Henke; U. T. Bornscheuer, *Anal. Chem.* **2003**, *75*, 255–260.
78. C. Branneby; P. Carlqvist; A. Magnusson; K. Hult; T. Brinck; P. Berglund, *J. Am. Chem. Soc.* **2003**, *125*, 874–875.
79. P. Carlqvist; M. Svedendahl; C. Branneby; K. Hult; T. Brinck; P. Berglund, *ChemBioChem* **2005**, *6*, 331–336.
80. O. Torre; I. Alfonso; V. Gotor, *Chem. Commun. (Camb.)* **2004**, *15*, 1724–1725.
81. A. R. Bassindale; K. F. Brandstadt; T. H. Lane; P. G. Taylor, *J. Inorg. Biochem.* **2003**, *96*, 401–406.
82. P. J. O'Brien; D. Herschlag, *Biochemistry* **2001**, *40*, 5691–5699.
83. I. Catrina; P. J. O'Brien; J. Purcell; I. Nikolic-Hughes; J. G. Zalatan; A. C. Hengge; D. Herschlag, *J. Am. Chem. Soc.* **2007**, *129*, 5760–5765.
84. P. J. O'Brien; D. Herschlag, *J. Am. Chem. Soc.* **1998**, *120*, 12369–12370.
85. D. M. Schmidt; E. C. Mundorff; M. Dojka; E. Bermudez; J. E. Ness; S. Govindarajan; P. C. Babbitt; J. Minshall; J. A. Gerlt, *Biochemistry* **2003**, *42*, 8387–8393.
86. A. Aharoni; L. Gaidukov; O. Khersonsky; Q. G. S. Mc; C. Roodveldt; D. S. Tawfik, *Nat. Genet.* **2005**, *37*, 73–76.
87. C. Roodveldt; D. S. Tawfik, *Biochemistry* **2005**, *44*, 12728–12736.
88. L. Afriat; C. Roodveldt; G. Manco; D. S. Tawfik, *Biochemistry* **2006**, *45*, 13677–13686.
89. D. Liu; B. W. Lepore; G. A. Petsko; P. W. Thomas; E. M. Stone; W. Fast; D. Ringe, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11882–11887.
90. G. Amitai; R. Devi-Gupta; D. S. Tawfik, *HFSP J.* **2007**, *1*, 67–78.
91. D. I. Draganov; J. F. Teiber; A. Speelman; Y. Osawa; R. Sunahara; B. N. La Du, *J. Lipid Res.* **2005**, *46*, 1239–1247.
92. O. Khersonsky; D. S. Tawfik, *Biochemistry* **2005**, *44*, 6371–6382.
93. G. E. Means; M. L. Bender, *Biochemistry* **1975**, *14*, 4989–4994.
94. M. A. Sogorb; V. Carrera; M. Benabent; E. Vilanova, *Chem. Res. Toxicol.* **2002**, *15*, 520–526.
95. F. Hollfelder; A. J. Kirby; D. S. Tawfik, *Nature* **1996**, *383*, 60–62.
96. S. M. Gould; D. S. Tawfik, *Biochemistry* **2005**, *44*, 5444–5452.
97. A. Fernandez-Gacio; A. Codina; J. Fastrez; O. Riant; P. Soumillion, *ChemBioChem* **2006**, *7*, 1013–1016.
98. E. Morett; G. Saab-Rincon; L. Olvera; M. Olvera; H. Flores; R. Grande, *J. Mol. Biol.* **2008**, *376*, 839–853.
99. W. M. Patrick; I. Matsumura, *J. Mol. Biol.* **2008**, *377*, 323–336.
100. F. M. Menger; M. Ladika, *J. Am. Chem. Soc.* **1978**, *100*, 3145–3146.
101. J. C. Hermann; R. Marti-Arbona; A. A. Fedorov; E. Fedorov; S. C. Almo; B. K. Shoichet; F. M. Raushel, *Nature* **2007**, *448*, 775–779.
102. N. U. Gamage; S. Tsvetanov; R. G. Duggleby; M. E. McManus; J. L. Martin, *J. Biol. Chem.* **2005**, *280*, 41482–41486.
103. L. C. James; D. S. Tawfik, *Trends Biochem. Sci.* **2003**, *28*, 361–368.
104. S. Meier; S. Ozbek, *Bioessays* **2007**, *29*, 1095–1104.
105. R. Bone; J. L. Silen; D. A. Agard, *Nature* **1989**, *339*, 191–195.
106. R. Bone; D. Frank; C. A. Kettner; D. A. Agard, *Biochemistry* **1989**, *28*, 7600–7609.
107. J. J. Perona; A. M. Martin, *J. Mol. Biol.* **1997**, *273*, 207–225.
108. Y. Yasutake; M. Yao; N. Sakai; T. Kirita; I. Tanaka, *J. Mol. Biol.* **2004**, *344*, 325–333.
109. J. M. Turner; J. Graziano; G. Spraggon; P. G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6483–6488.
110. X. Lu; L. Li; R. Wu; X. Feng; Z. Li; H. Yang; C. Wang; H. Guo; A. Galkin; O. Herzberg; P. S. Mariano; B. M. Martin; D. Dunaway-Mariano, *Biochemistry* **2006**, *45*, 1162–1172.
111. H. J. Lambie; N. I. Heyer; S. D. Bull; D. W. Hough; M. J. Danson, *J. Biol. Chem.* **2003**, *278*, 34066–34072.
112. A. Theodossis; H. Walden; E. J. Westwick; H. Connaris; H. J. Lambie; D. W. Hough; M. J. Danson; G. L. Taylor, *J. Biol. Chem.* **2004**, *279*, 43886–43892.
113. S. C. Wang; W. H. Johnson; C. P. Whitman, Jr., *J. Am. Chem. Soc.* **2003**, *125*, 14282–14283.
114. S. C. Wang; M. D. Person; W. H. Johnson; C. P. Whitman, Jr., *Biochemistry* **2003**, *42*, 8762–8773.
115. G. J. Poelarends; V. P. Veetil; C. P. Whitman, *Cell Mol. Life Sci.* **2008**, *65*, 3606–3618.
116. G. J. Poelarends; H. Serrano; W. H. Johnson; D. W. Hoffman, Jr.; C. P. Whitman, *J. Am. Chem. Soc.* **2004**, *126*, 15658–15659.
117. G. J. Poelarends; H. Serrano; W. H. Johnson; C. P. Whitman, Jr., *Biochemistry* **2005**, *44*, 9375–9381.
118. O. Khersonsky; D. S. Tawfik, *J. Biol. Chem.* **2006**, *281*, 7649–7656.

119. M. M. Blum; F. Lohr; A. Richardt; H. Ruterjans; J. C. Chen, *J. Am. Chem. Soc.* **2006**, *128*, 12750–12757.
120. G. Amitai; L. Gaidukov; R. Adani; S. Yishay; G. Yacov; M. Kushnir; S. Teitlboim; M. Lindenbaum; P. Bel; O. Khersonsky; D. S. Tawfik; H. Meshulam, *FEBS J.* **2006**, *273*, 1906–1919.
121. D. T. Yeung; D. E. Lenz; D. M. Cerasoli, *FEBS J.* **2005**, *272*, 2225–2230.
122. M. M. Blum; C. M. Timperley; G. R. Williams; H. Thiermann; F. Worek, *Biochemistry* **2008**, *47*, 5216–5224.
123. E. T. Kaiser; D. S. Lawrence, *Science* **1984**, *226*, 505–511.
124. M. Bakker; F. Van Rantwijk; R. A. Sheldon, *Can. J. Chem./Rev. Can. Chim.* **2002**, *80*, 622–625.
125. G. F. Da Silva; L. J. Ming, *J. Am. Chem. Soc.* **2005**, *127*, 16380–16381.
126. Z. P. Wu; D. Hilvert, *J. Am. Chem. Soc.* **1990**, *112*, 5647–5648.
127. S. Boschi-Muller; S. Muller; A. Van Dorselaer; A. Bock; G. Branlant, *FEBS Lett.* **1998**, *439*, 241–245.
128. H. J. Yu; J. Q. Liu; A. Bock; J. Li; G. M. Luo; J. C. Shen, *J. Biol. Chem.* **2005**, *280*, 11930–11935.
129. R. Kourist; S. Bartsch; L. Fransson; K. Hult; U. T. Bornscheuer, *ChemBioChem* **2008**, *9*, 67–69.
130. M. Forconi; D. Herschlag, *J. Am. Chem. Soc.* **2005**, *127*, 6160–6161.
131. D. E. Kunzler; S. Sasso; M. Gamper; D. Hilvert; P. Kast, *J. Biol. Chem.* **2005**, *280*, 32827–32834.
132. E. A. Taylor Ringia; J. B. Garrett; J. B. Thoden; H. M. Holden; I. Rayment; J. A. Gerlt, *Biochemistry* **2004**, *43*, 224–229.
133. W. S. Yew; K. Akana; E. L. Wise; I. Rayment; J. A. Gerlt, *Biochemistry* **2005**, *44*, 1807–1815.
134. A. C. Joerges; S. Mayer; A. R. Fersht, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5694–5699.
135. R. De Mot; A. De Schrijver; G. Schoofs; A. H. A. Parret, *FEMS Microbiol. Lett.* **2003**, *224*, 197–203.
136. M. Kataoka; K. Honda; S. Shimizu, *Eur. J. Biochem.* **2000**, *267*, 3–10.
137. C. Li; M. Hassler; T. D. H. Bugg, *ChemBioChem* **2008**, *9*, 71–76.
138. B. Seelig; J. W. Szostak, *Nature* **2007**, *448*, 828–831.
139. H. S. Park; S. H. Nam; J. K. Lee; C. N. Yoon; B. Mannervik; S. J. Benkovic; H. S. Kim, *Science* **2006**, *311*, 535–538.
140. L. Jiang; E. A. Althoff; F. R. Clemente; L. Doyle; D. Rothlisberger; A. Zanghellini; J. L. Gallaher; J. L. Betker; F. Tanaka; C. F. Barbas; D. Hilvert; K. N. Houk; B. L. Stoddard; D. Baker, *Science* **2008**, *319*, 1387–1391.
141. D. Rothlisberger; O. Khersonsky; A. M. Wollacott; L. Jiang; J. DeChancie; J. Betker; J. L. Gallaher; E. A. Althoff; A. Zanghellini; O. Dym; S. Albeck; K. N. Houk; D. S. Tawfik; D. Baker, *Nature* **2008**, *453*, 190–195.
142. J. M. Smith, *Nature* **1970**, *225*, 563–564.
143. N. Varadarajan; J. Gam; M. J. Olsen; G. Georgiou; B. L. Iverson, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6855–6860.
144. J. E. Vick; D. M. Schmidt; J. A. Gerlt, *Biochemistry* **2005**, *44*, 11722–11729.
145. D. Hoffmeister; J. Yang; L. Liu; J. S. Thorson, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13184–13189.
146. E. Timar; G. Groma; A. Kiss; P. Venetianer, *Nucleic Acids Res.* **2004**, *32*, 3898–3903.
147. A. Aharoni; L. Gaidukov; S. Yagur; I. Silman; D. S. Tawfik, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 482–487.
148. C. L. Wei; Y. B. Yang; C. H. Deng; W. C. Liu; J. S. Hsu; Y. C. Lin; S. H. Liaw; Y. C. Tsai, *Appl. Environ. Microbiol.* **2005**, *71*, 8873–8880.
149. J. Delmas; F. Robin; F. Carvalho; C. Mongaret; R. Bonnet, *Antimicrob. Agents Chemother.* **2006**, *50*, 731–738.
150. J. C. Samuelson; R. D. Morgan; J. S. Benner; T. E. Claus; S. L. Packard; S. Y. Xu, *Nucleic Acids Res.* **2006**, *34*, 796–805.
151. M. Larion; L. B. Moore; S. M. Thompson; B. G. Miller, *Biochemistry* **2007**, *46*, 13564–13572.
152. C. Jurgens; A. Strom; D. Wegener; S. Hettwer; M. Wilmanns; R. Sterner, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9925–9930.
153. L. Gaidukov; D. S. Tawfik, *Biochemistry* **2005**, *44*, 11843–11854.
154. L. C. James; D. S. Tawfik, *Pro. Sci.* **2001**, *10*, 2600–2607.
155. B. G. Hall, *FEMS Microbiol. Lett.* **1999**, *174*, 1–8.
156. F. A. Kondrashov, *Nat. Genet.* **2005**, *37*, 9–10.
157. M. Kirschner; J. Gerhart, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8420–8427.
158. D. Umeno; A. V. Tobias; F. H. Arnold, *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 51–78.
159. S. Bershtein; M. Segal; R. Bekerman; N. Tokuriki; D. S. Tawfik, *Nature* **2006**, *444*, 929–932.
160. M. Camps; A. Herman; E. Loh; L. A. Loeb, *Crit. Rev. Biochem. Mol. Biol.* **2007**, *42*, 313–326.
161. N. Tokuriki; F. Stricher; J. Schymkowitz; L. Serrano; D. S. Tawfik, *J. Mol. Biol.* **2007**, *369*, 1318–1332.
162. A. Wagner, *Proc. Biol. Sci.* **2008**, *275*, 91–100.
163. C. Adami, *Science* **2006**, *312*, 61–63.
164. J. T. Bridgham; S. M. Carroll; J. W. Thornton, *Science* **2006**, *312*, 97–101.
165. T. C. Galvao; V. de Lorenzo, *Curr. Opin. Biotechnol.* **2006**, *17*, 34–42.
166. T. L. O’Loughlin; D. N. Greene; I. Matsumura, *Mol. Biol. Evol.* **2006**, *23*, 764–772.
167. N. Ran; K. M. Draths; J. W. Frost, *J. Am. Chem. Soc.* **2004**, *126*, 6856–6865.
168. N. Varadarajan; S. Rodriguez; B. Y. Hwang; G. Georgiou; B. L. Iverson, *Nat. Chem. Biol.* **2008**, *4*, 290–294.
169. K. T. Watts; B. N. Mijts; P. C. Lee; A. J. Manning; C. Schmidt-Dannert, *Chem. Biol.* **2006**, *13*, 1317–1326.
170. I. Bertini; V. Calderone; M. Cosenza; M. Fragai; Y. M. Lee; C. Luchinat; S. Mangani; B. Terni; P. Turano, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5334–5339.
171. L. C. James; D. S. Tawfik, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12730–12735.
172. J. W. Thornton; E. Need; D. Crews, *Science* **2003**, *301*, 1714–1717.
173. M. A. Wouters; K. Liu; P. Riek; A. Husain, *Mol. Cell.* **2003**, *12*, 343–354.
174. N. Tokuriki; F. Stritcher; L. Serrano; D. S. Tawfik, *PLoS Comput. Biol.* **2008**, *4*, e1000002.
175. J. D. Bloom; J. J. Silberg; C. O. Wilke; D. A. Drummond; C. Adami; F. H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 606–611.
176. S. Bershtein; K. Goldin; D. S. Tawfik, *J. Mol. Biol.* **2008**, *379*, 1029–1044.
177. S. Leopoldseeder; J. Claren; C. Jurgens; R. Sterner, *J. Mol. Biol.* **2004**, *337*, 871–879.
178. M. A. Chow; K. E. McElroy; K. D. Corbett; J. M. Berger; J. F. Kirsch, *Biochemistry* **2004**, *43*, 12780–12787.
179. S. Rothman; M. Voorhies; J. Kirsch, *Protein Science* **2004**, *13* (3), 763.
180. J. Arima; Y. Uesugi; M. Iwabuchi; T. Hatanaka, *Appl. Environ. Microbiol.* **2005**, *71*, 7229–7235.
181. J. E. Vick; J. A. Gerlt, *Biochemistry* **2007**, *46*, 14589–14597.

182. F. P. Seebeck; D. Hilvert, *J. Am. Chem. Soc.* **2003**, *125*, 10158–10159.
183. P. L. Pettersson; A. S. Johansson; B. Mannervik, *J. Biol. Chem.* **2002**, *277*, 30019–30022.
184. J. M. Jez; T. M. Penning, *Biochemistry* **1998**, *37*, 9695–9703.
185. M. Henn-Sax; R. Thoma; S. Schmidt; M. Hennig; K. Kirschner; R. Sterner, *Biochemistry* **2002**, *41*, 12032–12042.
186. F. Barona-Gomez; D. A. Hodgson, *EMBO Rep.* **2003**, *4*, 296–300.
187. H. Wright; L. Noda-Garcia; A. Ochoa-Leyva; D. A. Hodgson; V. Fulop; F. Barona-Gomez, *Biochem. Biophys. Res. Commun.* **2008**, *365*, 16–21.
188. J. Claren; C. Malisi; B. Höcker; R. Sterner, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3704–3709.
189. L. A. Castle; D. L. Siehl; R. Gorton; P. A. Patten; Y. H. Chen; S. Bertain; H. J. Cho; N. Duck; J. Wong; D. L. Liu; M. W. Lassner, *Science* **2004**, *304*, 1151–1154.
190. I. Sarkar; I. Hauber; J. Hauber; F. Buchholz, *Science* **2007**, *316*, 1912–1915.
191. S. Ohno, *Evolution by Gene Duplication*; Allen & Unwin, Springer-Verlag: London, New York, 1970.
192. M. Kimura; T. Ota, *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2848–2852.
193. A. L. Hughes, *Trends Genet.* **2002**, *18*, 433–434.
194. M. Lynch; V. Katju, *Trends Genet.* **2004**, *20*, 544–549.
195. D. R. Scannell; K. H. Wolfe, *Genome Res.* **2008**, *18*, 137–147.
196. V. S. Cooper; R. E. Lenski, *Nature* **2000**, *407*, 736–739.
197. U. Bergthorsson; D. I. Andersson; J. R. Roth, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17004–17009.
198. S. Y. McLoughlin; D. L. Ollis, *Chem. Biol.* **2004**, *11*, 735–737.
199. P. Yue; Z. Li; J. Moulton, *J. Mol. Biol.* **2005**, *353*, 459–473.
200. J. Piatigorsky, *Gene Sharing and Evolution*; Harvard University Press: Boston, 2007.
201. M. Nei, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12235–12242.
202. R. Wroe; H. S. Chan; E. Bornberg-Bauer, *HFSP J.* **2007**, *1*, 79–87.
203. J. D. Bloom; Z. Lu; D. Chen; A. Raval; O. S. Venturelli; F. H. Arnold, *BMC Biol.* **2007**, *5*, 29.
204. A. Force; M. Lynch; F. B. Pickett; A. Amores; Y. L. Yan; J. Postlethwait, *Genetics* **1999**, *151*, 1531–1545.
205. M. Lynch; A. Force, *Genetics* **2000**, *154*, 459–473.

Biographical Sketches



Olga Khersonsky is a Ph.D. student in the group of Professor Dan S. Tawfik at the Department of Biological Chemistry in the Weizmann Institute of Science. She received her B.Sc. in chemistry from the Hebrew University of Jerusalem. Olga Khersonsky is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities.



Dan S. Tawfik is at the Department of Biological Chemistry in the Weizmann Institute of Science. He received his B.Sc. and M.Sc. degrees in chemistry and biochemistry from the Hebrew University of Jerusalem, and his Ph.D. from the Weizmann Institute of Science. He was a postdoctoral fellow and group leader at Cambridge University, and the MRC Centre for Protein Engineering, under the directorship of Professor Sir Alan Fersht. His research interests include protein evolution, mechanistic enzymology, and enzyme engineering.