The limits of enzyme specificity and the evolution of metabolism

The limits of enzyme specificity and the evolution of metabolism

Alessio Peracchi

Department of Chemistry, Life Sciences and Environmental Sustainability,

University of Parma, 43124 Parma, Italy.

Phone: (39) 0521905137.

Fax: (39) 0521905151.

Email: alessio.peracchi@unipr.it

Keywords: substrate specificity; substrate promiscuity; chemical-physical limits; metabolite repair enzymes; underground metabolism

This is a post-print version of the article eventually published on *Trends Biochem Sci* 2018 **43** (12), pp 984–996 (doi: 10.1016/j.tibs.2018.09.015)

Abstract

The substrate specificity of enzymes is bound to be imperfect, because of unavoidable physical-chemical limits. In extant metabolic enzymes, furthermore, such limits are seldom approached, suggesting that, on average, the degree of specificity of these enzymes is much lower than it could be attained. Along biological evolution, the reactivity of a single enzyme with available alternative substrates may be preserved to a significant or even substantial level for different reasons – for example, when the alternative reaction contributes to fitness, or when its undesirable products are nevertheless dispatched by metabolite repair enzymes. In turn, the widespread occurrence of promiscuous reactions is a consistent source of metabolic ‘messiness’, from which both liabilities and opportunities ensue in the evolution of metabolic systems.

Many metabolic enzymes are not strictly substrate-specific

Today it is well appreciated that a substantial fraction of metabolic enzymes can catalyze reactions of different types and/or with different substrates [1, 2]. The former behavior is termed catalytic promiscuity [3], the latter is usually called substrate promiscuity [4] (as these terms may be equivocal, Text Box 1 explains the definition of 'promiscuity' used here and compares it with a stricter definition accepted by evolutionary biochemists). Although this paper is essentially concerned with substrate promiscuity, it must be noted that the two behaviors are interrelated, often co-occur (e.g., [5-7]), and have analogous impacts on metabolism, so that most of the points concerning substrate promiscuity are similarly applicable to catalytic promiscuity.

Why is it that many metabolic enzymes can transform different substrates? Is it simply because an absolute substrate specificity cannot be attained, due to the inherent imperfection of enzymes [8, 9]?. Or is it mostly the result of selective pressures (or lack thereof)? Finally, what are the consequences and implications of the recurrence of substrate promiscuity for the global evolution of metabolism? Related to these issues, this review will begin by showing that substrate specificity is indeed inherently limited, for reasons rooted in physical chemistry, but also that, in many cases, metabolic enzymes are less selective than they could be. Subsequently, the review will examine how different evolutionary factors (both positive and negative selection, as well as neutral drift) may help shape the degree to which enzymes discriminate between potential substrates. Finally, it will be suggested that the universal tendency of enzymes to show substrate promiscuity is an important source of metabolome complexity and helps fuel an 'underground' network of reactions which may represent a premise to the further evolution and diversification of metabolism.

Substrate specificity, discrimination and binding energy

Contrary to catalytic efficiency, which can be gauged in reference to an absolute scale of 'catalytic perfection' [10, 11], specificity is a relative concept, as it requires a comparison between given alternative substrates. In fact, specificity is formally defined as the ability of an enzyme to discriminate between reaction with two potential substrates, in the presence of both compounds [12, 13]. In a biological context, specificity entails reacting with a single substrate in preference to a multitude of other metabolites in the cell. Intuitively, this can be a very difficult exercise. For example, an enzyme intended to be specific for aspartate should discriminate against (among others) asparagine, glutamate,

homoserine, homocysteine, phosphoserine, alanine, L-malate, oxaloacetate and succinate, all of which are common metabolites with obvious structural similarities to the proper (cognate) substrate.

Discrimination (and hence specificity) does not depend simply on the relative affinity of the substrates for the enzyme, indeed it is acknowledged that the substrates should be compared based on the ratio of k_{cat}/K_M values for their reactions [14] – a ratio also called the discrimination factor [8]. A pertinent question therefore is whether there are intrinsic limits to this ratio. One potential way to address this issue is through application of transition state theory. According to this theory, the logarithm of k_{cat}/k_M is proportional to the free-energy difference ($\Delta G^{\neq}$) between the free enzyme and substrate and the transition state complex [12] (Text Box 2). Thus, whenever transition state theory is applicable, differences in k_{cat}/K_M between different substrates reflect their different binding energies in *the transition state*.

Additionally, the amount of binding energy provided by simple groups is finite. Hence, when comparing substrates with similar structures, the difference in binding energies ($\Delta\Delta G^{\neq}$) must also be finite, relatively small and, in some cases, calculable. To illustrate this point, I will next analyze three exemplary cases, in which an enzyme is asked to discriminate against alternative substrates that are either slightly smaller or slightly larger than the cognate substrate (Figure 1).

Theoretical and empirical limits of substrate specificity

That substrate specificity is inherently limited was first glimpsed by Pauling sixty years ago [15], in relation to the process of aminoacyl-tRNA synthesis. He focused in particular on the case of isoleucyl-tRNA synthetase having to distinguish between isoleucine and valine. The two amino acids differ only by one methyl group, so the difference in binding energy between them could not be greater than the energy provided by the terminal CH_3 group of L-Ile. This has been estimated to be worth at most 3 kcal/mole ([12] and references therein), which sets an upper limit of about 160-fold in the discrimination between valine and isoleucine by simple binding (at 25°C). Such a relatively low discrimination seems biologically unacceptable, in fact extant isoleucyl-tRNA synthetases possess a distinct proofreading function, that selectively deacylates any mis-aminoacylated $tRNA^{Ile}$, achieving a better accuracy in an energy-expensive manner [16].

Evidently, the limit above does not apply solely to aminoacyl-tRNA synthetases, but is general for enzymes that must distinguish between two substrates differing by just one

methyl group. Couples of this kind are not infrequent among common metabolites. A survey of the Brenda database [17] shows that only rarely the observed discrimination index reaches the theoretical limit (Figure 1a).

One can also try to calculate theoretical limits of discrimination in cases when the alternative substrate is missing a hydroxyl group compared to the cognate substrate. The binding energy provided by a hydroxyl depends essentially on the group's ability to form hydrogen bonds, and the thermodynamic stability of hydrogen bonds within macromolecules has been addressed in a number of mutagenesis studies. The maximum energy associated to hydrogen bonding in an uncharged donor-acceptor pair is about 2 kcal/mol, whereas a hydrogen bond to a charged partner can be worth up to 5 kcal/mol [18]. Since interactions with more than one charged partner would generate electrostatic interferences, and since hydroxyl groups can form up to three hydrogen bonds, one rough estimate is that an extra hydroxyl group could be worth up to 9 kcal/mol of binding energy, or about a $4x10^6$-fold factor in selectivity [19]. Figure 2B shows a set of data from the literature pertaining to this case. Although the sample may be somewhat biased (because catalytic parameters for substrates showing a very low reactivity are less likely to be determined) it nevertheless suggests that discrimination is often much lower than the theoretical maximum (Figure 1b and Supplementary Table 2).

Repulsive interactions towards undesirable substrates are arguably a most efficient means to implement specificity [8]. In particular, it could be assumed that discrimination against a substrate that is larger than the cognate substrate may be achieved easily by restricting the active site and exploiting steric repulsion [12]. However, enzymes are quite flexible (and there are experimental evidences that flexibility may be correlated with substrate promiscuity - e.g. [20, 21] [22]) while active sites must have an opening towards the solvent; it has been suggested that such features may allow in specific cases the reaction with bulkier alternative substrates [9].

In this instance actual limits to specificity are nearly impossible to estimate *a priori*. Fig. 1c examines a practical case, considering how enzymes that act on a linear substrate discriminate against a competitor metabolite containing one additional methylene group. In this case, too, many enzymes show a relatively modest selectivity. Note that alternative activities may occur even with substrates that are much larger than the canonical substrate. For example, it was shown that several mammalian transaminases, whose standard substrates are simple amino acids, can transaminate with low efficiency the tripeptide glutathione [23].

Reactivity with alternative substrates may be advantageous for fitness

It is assumed that in the earliest phases of metabolism enzymes were generalists, reacting on many substrates. In the course of evolution, most of them tended to become specialist - more and more selective (and arguably efficient) towards a given substrate [1, 24, 25]. Nevertheless, as shown in Fig. 1, it seems that the maximum possible discrimination between available substrates is seldom reached by extant metabolic enzymes. This observation is consistent with the idea that the actual degree of substrate specificity in these catalysts largely depends on evolutionary factors and dynamics.

In the context of metabolism, the reactivity of an enzyme with alternative substrates could conceivably produce three types of evolutionary responses. Sometimes, reactivity towards different substrates is simply advantageous for fitness and hence retained or selected for during evolution. Besides digestive enzymes or detoxifying catalysts such as cytochromes P450 [26], there are textbook instances of substrate promiscuous enzymes operating in the middle of primary metabolism, such as ketol-acid reductoisomerase [27], and branched-chain amino acid transaminase [28], each of which is involved in the biosynthesis of both valine and isoleucine, catalyzing distinct (but analogous) reactions in the two pathways. Similarly transketolase [29] and adenylosuccinate lyase [30] catalyze reactions with different substrates in the same pathway. Other enzymes catalyzing distinct reactions in primary metabolism have been described, often in specific organisms (e.g., [27] [31-35]). And there have even been reports of bifunctional pathways, such as one in which the biosynthesis of both L-lysine and L-ornithine is accomplished using the same set of enzymes [36].

In order to provide an advantage, a substrate-promiscuous enzyme needs not to react with comparable efficiency against two alternative substrates, as even a slow side reaction may sometimes be beneficial, e.g., when its product is required at low levels by the cell. For example, in microorganisms, certain enzymes from primary metabolism can also participate (acting on different substrates) in the biosynthesis of antibiotics [37, 38]. In mammals, a similar case may perhaps be the synthesis of D-aspartate. This D-amino acid occurs at small but significant levels in the brain and acts as an endogenous co-agonist for synaptic N-Methyl D-Aspartate receptors [39]. Contrary to earlier claims on the occurrence of a specific aspartate racemase, it is now believed that D-aspartate may be mainly produced by serine racemase, even though this enzyme reacts with L-aspartate ~50-fold less efficiently than with L-serine [40].

The immediate advantage of using of a single catalyst for different metabolic reactions is parsimony [41]; however, there are also obvious drawbacks. One potential problem is related to flux efficiency - both because it may be difficult to combine in the same enzyme substrate versatility with high rate of the catalyzed reactions ([8, 42] and references therein) and because the alternative substrates may interfere with each other (e.g., [30, 43]). Another drawback is that two activities carried out by a single enzyme cannot be regulated separately by modulating expression levels or subcellular localization (intuitively, this is much less of a problem if the two activities are required in the same pathway). A network analysis study on *Escherichia coli* [1] has provided support to the notion that promiscuous enzymes are less abundant in pathways where a high metabolic flux is needed, or where flux must be regulated more tightly.

Detrimental activities with alternative substrates may be redressed by repair enzymes

An opposite scenario arises when the reaction of a metabolic enzyme with an alternative substrate is seriously detrimental for fitness. In fact, every such reaction may generate a product that is useless for the cell (a metabolic dead end, implying a waste of resources) or even toxic [44], which is expected to elicit a strong evolutionary pressure to improve the enzyme specificity. However, as discussed, there are limits to such an improvement. Preventing access of the enzyme to the alternative substrate by compartmentalization may also be often impossible or insufficient.

Instead, evolution can lead to the development of *ad hoc* enzymes that destroy or recycle the unwanted products of these side reactions. In fact, there is a growing list of 'metabolite repair' enzymes whose sole purpose appears to be the correction of ‘errors’ committed by enzymes of intermediary metabolism [45, 46], most often attributable to imperfect substrate specificity. These resemble the proofreading activities that improve the accuracy of aminoacyl-tRNA synthesis [45]; sometimes, the repair task is even allocated to a separate domain of the promiscuous metabolic enzyme, not unlike what happens in many aminoacyl-tRNA synthetases [47].

One exemplary case is the reaction of malate dehydrogenase with α-ketoglutarate instead of the standard substrate oxaloacetate. The relative efficiency of the promiscuous reaction is very low, the discrimination index of malate dehydrogenase is estimatedly around 10^6 [48], i.e. higher than for any of the enzymes in Fig 1c. Despite this, the alternative reaction is not physiologically insignificant, as both α-ketoglutarate and malate

dehydrogenase are abundant in cells while L-2 hydroxyglutarate is a metabolic dead end, which favors its buildup over time. Normally, such a buildup is prevented by a dedicated repair enzyme, a FAD-dependent dehydrogenase that irreversibly re-oxidizes L-2-hydroxyglutarate to α-ketoglutarate. The fact that, in humans, a deficiency of the repair enzyme causes L-2-hydroxyglutaric aciduria (a severe neurological disorder) [49], underscores the potential dangers of even slow side activities of substrate-promiscuous enzymes.

Similarly, the slow transamination of glutathione carried out by many transaminases (mentioned earlier) conduces to an apparently useless product, deaminated glutathione. This compound, however, can be hydrolyzed and recycled by a 'repair' amidase, identified in mammals, yeast and several glutathione-producing bacteria [23]. It is notable that the compound processed by this amidase appears to originate from the side activities of an entire class of enzymes. This attests to the efficiency of metabolite repair as an evolutionary solution to the inherent imperfection of metabolic catalysts. An analogous and even more striking example is that of a 'repair' phosphatase that degrades the inhibitory compounds generated by the substrate-promiscuous activities of two glycolytic enzymes – glyceraldehydes 3-phosphate dehydrogenase and pyruvate kinase [46].

It must be noted that metabolite repair enzymes, by themselves, do not prevent the formation of useless or toxic side-reaction products, but simply curb their accumulation. Furthermore, the very presence of a repair system that remedies the effects of an undesired side reaction, may be expected, quite paradoxically, to lower the evolutionary pressure for metabolic enzymes to maximize their specificity.

When does reactivity with alternative substrates depend on neutral drift?

In a final scenario, the alternative reaction of an enzyme could have no significant (positive or negative) effects on the system fitness, so that it would be invisible to natural selection and essentially subject to neutral drift. This is often assumed as the 'default' case [11], but positive proofs are scarce. We have just seen that some reactions which, based on the discrimination factor, would appear very negligible, do become liabilities because the promiscuous enzyme and the alternative substrate abound in the cell. As a minimum, an estimate of the biological significance of the alternative reaction (and hence of the selective pressure to which it is subject) should take into account the effective rate of its occurrence *in vivo*.

Also, one may think that secondary activities that generate mainstream metabolites should be essentially irrelevant for fitness and therefore depend on neutral drift. Perhaps the best case in point comes from a study by Khanal et al. [50], who analyzed the ability of nine microbial γ-glutamylphosphate reductases (ProA) to use the alternative substrate N-acetyl-glutamyl phosphate (which is usually processed by a distinct reductase, ArgC). The discrimination indices for the nine enzymes were always very high (>40,000) while the absolute efficiency of the secondary activity varied up to 50-fold between species. Furthermore, the enzyme from *E. coli* could not compensate for the loss of ArgC. All of these features were consistent with the side activity of ProA being biologically irrelevant [50].

In contrast to the ProA case, however, enzymes whose secondary activities are metabolically redundant and can, under some circumstances, surrogate the function of more specialized catalysts are observed often (e.g. [51, 52]) and it cannot be excluded that the non-strict substrate specificity of these enzymes may be retained because beneficial. Speaking more generally,there are suggestions that side activities generating mainstream metabolites, may provide advantages or disadvantages under particular conditions and therefore be under active selection [44].

For example, let's consider an aminotransferase whose primary substrate is not alanine. It might be assumed that a side activity on alanine may not be strongly favored or unfavored by evolution because the product, pyruvate, is a standard metabolic intermediate. But the alternative reaction can contribute to metabolic redundancy, which may increase fitness under particular circumstances while ensuring the resilience of the metabolic system. In *E. coli* for example, there are three major alanine aminotransferases plus up to seven other enzymes with substantial activity towards alanine. Such a remarkable redundancy is proposedly important to ensure a supply of D-alanine for peptidoglycan synthesis [53]. Conversely, there may be significant counterselection towards side activities that, despite generating standard metabolites, represent an objective waste of energy or resources. An example could be the transamination of glutamine, which tends to be reserved to processes that need to be metabolically irreversible [54].

Another case worth considering is that of kinases and ligases; many of these enzymes show a strong specificity for ATP, while others can use with comparable efficiencies different nucleotides triphosphates (NTPs), (e.g., [55, 56]). These different behaviors could be the random results of distinct evolutionary histories, but again positive

selection for different levels of specificity cannot be ruled out *a priori*, ATP is usually the most abundant NTP, so the activity of enzymes strictly dependent on ATP will respond more promptly to the energy status of the cell, which may be important in the perspective of regulating metabolic fluxes. On the other hand, kinases or ligases that are less selective may allow an organism to scavenge alternative NTPs depending upon availability, thus responding to the need of preserving the efficiency of an important reaction under variable metabolic conditions [55, 56].

In sum, while it is reasonable to assume that often the activities of an enzyme on alternative (metabolically available) substrates may depend on neutral drift, the point is hard to prove because various kinds of selective pressures can be at play, different in different organismal lineages [57, 58]. For many enzymes, the boundaries between an evolutionarily selected degree of substrate promiscuity and an evolutionarily irrelevant degree of sloppiness may be difficult to establish (See Outstanding Questions Box).

Substrate promiscuity contributes to underground metabolism

As seen in the sections above, enzymes have an unavoidable tendency to react with alternative, available substrates. Even when (perhaps not too often) this tendency is repressed near the minimum level allowed by chemistry, and even in the presence of 'repair' systems, these enzymes will generate alternative products, many of which will be nonstandard metabolites, Substrate promiscuity must be hence considered a major contribution to the complexity of the metabolome - together with catalytic promiscuity and other factors such as the nonenzymic transformation of metabolites and the introduction of xenobiotics [44, 45]. The contribution of substrate promiscuity is in part predictable [59] and understanding it seems needed for a thorough mapping the metabolome, which is one of the current challenges of biomolecular sciences [60].

Sometimes, the compounds generated by side activities of metabolic enzymes contribute to a subterranean ensemble of reactions, proceeding in general at extremely low fluxes, that are collectively termed 'underground metabolism' [61, 62]. It is believed that this messy network of reactions, while somehow unavoidable, may represent a premise to the further evolution and diversification of metabolism [54, 63].

That messiness may be key to evolve new beneficial functions is well exemplified by many pathways in secondary (specialized) metabolism. Contrary to primary metabolism, specialized metabolism contributes to fitness by producing molecules that interfere with the biology of other organisms (e.g., environmental competitors). Many

pathways in specialized metabolism tend to include multiple enzymes with non-strict substrate specificity and to produce an entire suite of chemically similar, but distinct, compounds [64-66] . The most direct advantage, in this case, is that the ability to generate a multiplicity of secondary metabolites offers more flexibility in responding to different, and evolving, organisms in the environment.

In primary metabolism, the 'messy' underground reactions could also, in some cases, facilitate an evolutionary adaptation to new environments or provide bypass to genetic lesions [61, 62]. For example, Kim et al. found that the deletion of gene PdxB, required in *E.coli* for the biosynthesis of pyridoxal phosphate, could be bypassed by at least three 'underground' routes (one of which was elucidated in detail and found to involve the promiscuous activity of homoserine kinase on an alternative substrate) [67]. A fourth route was reported by another group, using a promiscuity prediction software [68]. Earlier, Patrick and co-workers had found that 41 out of 104 *E. coli* knockout strains, unable to grow on minimal medium, could be rescued by overexpression of at least one gene different from the deleted one. In several cases, the multicopy suppressor was not homologous to the deleted gene and its effect was attributed to a promiscuous activity of the encoded enzyme or to a metabolic pathway bypass [69].

Underground reactions may also become - at some point and in some organisms - incorporated into mainstream primary metabolism. A possible example is the promiscuous reaction of phosphoglycerate dehydrogenase with α-ketoglutarate, yielding D- 2-hydroxyglutarate. In humans, D-2-hydroxyglutarate is an oncometabolite [70] and requires a repair enzyme to be removed [45]. In bacteria such as *E. coli*, however, the promiscuous activity of phosphoglycerate dehydrogenase is enhanced and tightly coupled to the repair activity, to drive thermodynamically L-serine biosynthesis [71].

Concluding remarks

The results and arguments reviewed here show that, while perfect substrate specificity is essentially unattainable, metabolic enzymes are often much less selective than they could be. Furthermore, the activities with alternative substrates are subject to distinct selective pressures. They can be fostered by natural selection until they reach levels that are most useful for fitness, or repressed to levels at which they are no longer harmful. When deleterious side reactions cannot be controlled efficiently enough, metabolite repair enzymes may evolve, which limit the buildup of unwanted products. In any case, enzymes that react with alternative substrates create complexity in the systems

they contribute to. While this complicates our understanding of cellular behavior, it seems important for future advantageous metabolic adaptations of the host organisms.

References

1. Nam, H.*, et al.* (2012) Network context and selection in the evolution to enzyme specificity. *Science* 337, 1101-1104
2. Martínez-Núñez, M.A.*, et al.* (2017) Tracing the repertoire of promiscuous enzymes along the metabolic pathways in archaeal organisms. *Life (Basel).* 7, 30
3. O'Brien, P.J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* 6, R91-R105.
4. Hult, K., and Berglund, P. (2007) Enzyme promiscuity: mechanism and applications. *Trends Biotechnol.* 25, 231-238
5. Taylor Ringia, E.A.*, et al.* (2004) Evolution of enzymatic activity in the enolase superfamily: functional studies of the promiscuous o-succinylbenzoate synthase from *Amycolatopsis*. *Biochemistry* 43, 224-229
6. Magnani, G.*, et al.* (2013) Completing the folate biosynthesis pathway in *Plasmodium falciparum*: *p*-aminobenzoate is produced by a highly divergent promiscuous aminodeoxychorismate lyase. *Biochem. J.* 455, 149-155
7. Tarcz, S.*, et al.* (2014) Substrate and catalytic promiscuity of secondary metabolite enzymes:O-prenylation of hydroxyxanthones with different prenyl donors by a bisindolyl benzoquinone C- and N-prenyltransferase. *RSC Adv.* 4, 17986
8. Tawfik, D.S. (2014) Accuracy-rate tradeoffs: how do enzymes meet demands of selectivity and catalytic efficiency? *Curr. Opin. Chem. Biol.* 21C, 73-80
9. Copley, S.D. (2015) An evolutionary biochemist's perspective on promiscuity. *Trends Biochem. Sci.* 40, 72-78
10. Bar-Even, A.*, et al.* (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* 50, 4402-4410
11. Newton, M.*, et al.* (2018) Enzyme evolution: innovation is easy, optimization is complicated. *Curr. Opin. Struct. Biol.* 48, 110-116
12. Fersht, A.R. (1999) *Structure and mechanism in protein science - A guide to enzyme catalysis and protein folding*. Freeman & Co.
13. Herschlag, D. (1988) The role of induced fit and conformational changes of enzymes in specificity and catalysis. *Bioorg. Chem.* 16, 62-96
14. Cornish-Bowden, A., and Cardenas, M.L. (2010) Specificity of non-Michaelis-Menten enzymes: necessary information for analyzing metabolic pathways. *J. Phys. Chem. B* 114, 16209-16213

15. Pauling, L. (1958) The probability of errors in the process of synthesis of protein molecules. In *Festschrift fur Prof Dr Arthur Stoll*, 597-602, Birkhauser Verlag
16. Perona, J.J., and Gruic-Sovulj, I. (2014) Synthetic and editing mechanisms of aminoacyl-tRNA synthetases. *Top. Curr. Chem.* 344, 1-41
17. Schomburg, I.*, et al.* (2017) The BRENDA enzyme information system-From a database to an expert system. *J. Biotechnol.* 261, 194-206
18. Fersht, A.R., and Serrano, L. (1993) Principles of protein stability derived from protein engineering experiments. *Curr. Opin. Struct. Biol.* 3, 75-83
19. Wolfenden, R., and Kati, W.M. (1991) Testing the limits of protein-ligand binding discrimination with transition-state analogue inhibitors. *Acc. Chem. Res.* 24, 209–215
20. Kristiansson, H., and Timson, D.J. (2011) Increased promiscuity of human galactokinase following alteration of a single amino acid residue distant from the active site. *Chembiochem* 12, 2081-2087
21. Noda-García, L.*, et al.* (2013) Evolution of substrate specificity in a recipient's enzyme following horizontal gene transfer. *Mol. Biol. Evol.* 30, 2024-2034
22. Zou, T.*, et al.* (2015) Evolution of conformational dynamics determines the conversion of a promiscuous generalist into a specialist enzyme. *Mol. Biol. Evol.* 32, 132-143
23. Peracchi, A.*, et al.* (2017) Nit1 is a metabolite repair enzyme that hydrolyzes deaminated glutathione. *Proc. Natl. Acad. Sci. U.S.A.* 114, E3233-E3242
24. Mannervik, B., and Runarsdottir, A. (2010) The quest for molecular quasi-species in ligand-activity space and its application to directed enzyme evolution. *FEBS Lett.* 584, 2565-2571
25. Khersonsky, O., and Tawfik, D.S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* 79, 471-505
26. Behrendorff, J.*, et al.* (2015) Directed evolution of cytochrome P450 enzymes for biocatalysis: exploiting the catalytic versatility of enzymes with relaxed substrate specificity. *Biochem. J.* 467, 1-15
27. Verdel-Aranda, K.*, et al.* (2015) Molecular annotation of ketol-acid reductoisomerases from *Streptomyces* reveals a novel amino acid biosynthesis interlock mediated by enzyme promiscuity. *Microb. Biotechnol.* 8, 239-252
28. Yu, X.*, et al.* (2014) The specificity and kinetic mechanism of branched-chain amino acid aminotransferase from *Escherichia coli* studied with a new improved coupled assay procedure and the enzyme's potential for biocatalysis. *FEBS J.* 281, 391-400

29. Kochetov, G.A., and Solovjeva, O.N. (2014) Structure and functioning mechanism of transketolase. *Biochim. Biophys. Acta* 1844, 1608-1618
30. Ray, S.P.*, et al.* (2013) Inherent properties of adenylosuccinate lyase could explain S-Ado/SAICAr ratio due to homozygous R426H and R303C mutations. *Biochim. Biophys. Acta* 1834, 1545-1553
31. Poon, W.W.*, et al.* (1999) Yeast and rat Coq3 and *Escherichia coli* UbiG polypeptides catalyze both O-methyltransferase steps in coenzyme Q biosynthesis. *J. Biol. Chem.* 274, 21665-21672
32. Barona-Gómez, F., and Hodgson, D.A. (2003) Occurrence of a putative ancient-like isomerase involved in histidine and tryptophan biosynthesis. *EMBO Rep.* 4, 296-300
33. Adams, N.E.*, et al.* (2014) Promiscuous and adaptable enzymes fill "holes" in the tetrahydrofolate pathway in *Chlamydia* species. *MBio* 5, e01378-01314
34. Shimizu, T.*, et al.* (2017) Structure and function of an ancestral-type β-decarboxylating dehydrogenase from *Thermococcus kodakarensis*. *Biochem. J.* 474, 105-122
35. Ferla, M.P.*, et al.* (2017) Primordial-like enzymes from bacteria with reduced genomes. *Mol Microbiol* 105, 508-524
36. Yoshida, A.*, et al.* (2016) Lysine biosynthesis of *Thermococcus kodakarensis* with the capacity to function as an ornithine biosynthetic system. *J. Biol. Chem.* 291, 21630-21643
37. Mahlert, C.*, et al.* (2007) Stereospecific enzymatic transformation of α-ketoglutarate to (2S,3R)-3-methyl glutamate during acidic lipopeptide biosynthesis. *J. Am. Chem. Soc.* 129, 12011-12018
38. Ishikawa, F.*, et al.* (2016) *In vitro* investigation of crosstalk between fatty acid and polyketide synthases in the andrimid biosynthetic assembly line. *Chembiochem* 17, 2137-2142
39. Errico, F.*, et al.* (2015) D-Aspartate: An endogenous NMDA receptor agonist enriched in the developing brain with potential involvement in schizophrenia. *J. Pharm. Biomed. Anal.* 116, 7-17
40. Uda, K.*, et al.* (2016) Distribution and evolution of the serine/aspartate racemase family in invertebrates. *Amino Acids* 48, 387-402
41. Price, D.R., and Wilson, A.C. (2014) A substrate ambiguous enzyme facilitates genome reduction in an intracellular symbiont. *BMC Biol.* 12, 110
42. Gatti-Lafranconi, P., and Hollfelder, F. (2013) Flexibility and reactivity in promiscuous enzymes. *Chembiochem* 14, 285-292

43. Sun, R., and Wang, L. (2014) Thymidine kinase 2 enzyme kinetics elucidate the mechanism of thymidine-induced mitochondrial DNA depletion. *Biochemistry* 53, 6142-6150
44. Piedrafita, G.*, et al.* (2015) The Impact of non-enzymatic reactions and enzyme promiscuity on cellular metabolism during (oxidative) stress conditions. *Biomolecules* 5, :2101-2122
45. Linster, C.L.*, et al.* (2013) Metabolite damage and its repair or pre-emption. *Nat. Chem. Biol.* 9, 72-80
46. Collard, F.*, et al.* (2016) A conserved phosphatase destroys toxic glycolytic side products in mammals and yeast. *Nat. Chem. Biol.* 12, 601-607
47. Manandhar, M., and Cronan, J.E. (2013) Proofreading of noncognate acyl adenylates by an acyl-coenzyme a ligase. *Chem. Biol.* 20, 1441-1446
48. Van Schaftingen, E.*, et al.* (2013) Metabolite proofreading, a neglected aspect of intermediary metabolism. *J. Inherit. Metab. Dis.* 36, 427-434
49. Rzem, R.*, et al.* (2004) A gene encoding a putative FAD-dependent L-2-hydroxyglutarate dehydrogenase is mutated in L-2-hydroxyglutaric aciduria. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16849-16854
50. Khanal, A.*, et al.* (2015) Differential effects of a mutation on the normal and promiscuous activities of orthologs: implications for natural and directed evolution. *Mol. Biol. Evol.* 32, 100-108
51. Vorobieva, A.A.*, et al.* (2014) *Escherichia coli* D-malate dehydrogenase, a generalist enzyme active in the leucine biosynthesis pathway. *J. Biol. Chem.* 289, 29086-29096
52. Jerlstrom Hultqvist, J.*, et al.* (2018) A bacteriophage enzyme induces bacterial metabolic perturbation that confers a novel promiscuous function. *Nat. Ecol. Evol.* 2, 1321-1330
53. Lal, P.B.*, et al.* (2014) The redundant aminotransferases in lysine and arginine synthesis and the extent of aminotransferase redundancy in *Escherichia coli*. *Mol. Microbiol.* 94, 843-856
54. Danchin, A., and Sekowska, A. (2013) The logic of metabolism and its fuzzy consequences. *Environ. Microbiol.* 16, 19-28
55. Gallois-Montbrun, S.*, et al.* (2004) Broad specificity of human phosphoglycerate kinase for antiviral nucleoside analogs. *Biochem. Pharmacol.* 68, 1749-1756
56. Salaemae, W.*, et al.* (2015) Nucleotide triphosphate promiscuity in *Mycobacterium tuberculosis* dethiobiotin synthetase. *Tuberculosis* 95, 259-266

57. Verduzco-Castro, E.A.*, et al.* (2016) Co-occurrence of analogous enzymes determines evolution of a novel (βα)8-isomerase sub-family after non-conserved mutations in flexible loop. *Biochem. J.* 473, 1141-1152
58. Juárez-Vázquez, A.L.*, et al.* (2017) Evolution of substrate specificity in a retained enzyme driven by gene loss. *Elife* 6, e22679
59. Jeffryes, J.*, et al.* (2015) MINEs: open access databases of computationally predicted enzyme promiscuity products for untargeted metabolomics. *J. Cheminform.* 28, 44
60. da Silva, R.R.*, et al.* (2015) Illuminating the dark matter in metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* 112, 12549-12550
61. D'Ari, R., and Casadesús, J. (1998) Underground metabolism. *Bioessays* 20, 181-186
62. Notebaart, R.A.*, et al.* (2018) Underground metabolism: network-level perspective and biotechnological potential. *Curr. Opin. Biotechnol.* 49, 108-114
63. Noda-Garcia, L.*, et al.* (2018) Metabolite-enzyme coevolution: from single enzymes to metabolic pathways and networks. *Annu. Rev. Biochem.* 87, 187-216
64. Fischbach, M.A., and Clardy, J. (2007) One pathway, many products. *Nat. Chem. Biol.* 3, 353-355
65. Weng, J.K., and Noel, J.P. (2012) The remarkable pliability and promiscuity of specialized metabolism *Cold Spring Harb. Symp. Quant. Biol.* 77, 309–320
66. Guo, J.*, et al.* (2014) Specificity and promiscuity at the branch point in gentamicin biosynthesis. *Chem. Biol.* 22, 608-618
67. Kim, J.*, et al.* (2010) Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5'-phosphate synthesis. *Mol. Syst. Biol.* 6, 436
68. Oberhardt, M.A.*, et al.* (2016) Systems-wide prediction of enzyme promiscuity reveals a new underground alternative route for pyridoxal 5'-phosphate production in *E. coli*. *PLoS Comput. Biol.* 12, e1004705
69. Patrick, W.M.*, et al.* (2007) Multicopy suppression underpins metabolic evolvability. *Mol. Biol. Evol.* 24, 2716-2722
70. Fan, J.*, et al.* (2015) Human phosphoglycerate dehydrogenase produces the oncometabolite D-2-hydroxyglutarate. *ACS Chem. Biol.* 10, 510-516
71. Zhang, W.*, et al.* (2017) Coupling between D-3-phosphoglycerate dehydrogenase and D-2-hydroxyglutarate dehydrogenase drives bacterial L-serine synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7574-E7582

72. Grant, G.A. (2018) Elucidation of a self-sustaining cycle in *Escherichia coli* L-serine biosynthesis that results in the conservation of the coenzyme, NAD^+. *Biochemistry* 57, 1798-1806
73. Holman, K.M.*, et al.* (2017) Insights into substrate promiscuity of human seryl-tRNA synthetase. *RNA* 23, 1685-1699
74. Copley, S.D. (2017) Shining a light on enzyme promiscuity. *Curr. Opin. Struct. Biol.* 47, 167-175
75. Flores, C.L., and Gancedo, C. (2011) Unraveling moonlighting functions with yeasts. *IUBMB Life* 63, 457-462
76. Lienhard, G.E. (1973) Enzymatic catalysis and transition-state theory. *Science* 180, 149-154

Figure legends

Figure 1: A survey of experimentally determined abilities of enzymes to discriminate between substrates having similar structures. The substrates considered are metabolically available and, within each pair of substrates, the structural differences occur away from the reaction center. Information about the individual enzymes and substrates, as well as the references to the original studies is provided in supplementary Tables 1-3. (a) Cases in which the alternative substrate lacks one CH_3 group with respect to the reference substrate. The dashed line represents a discrimination factor of 160, which is the approximate limit for selectivity at 25°C (the limit will vary with temperature). (b) Cases in which the alternative substrate lacks one OH with respect to the reference substrate. The dashed line represents a discrimination factor of 4×10^6. (c) Cases in which the reference substrate is a linear molecule and the alternative substrate is longer by one methylene (-CH_2-) group.

Figure 2 – Detrimental activities of two glycolytic enzymes on alternative substrates are corrected by a single repair enzyme. Panels (a) and (b) show the side activities of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) on erithrose 4-phosphate and of pyruvate kinase (PK) on lactate, to yield 4-phosphoerythronate and 2-phospho-L-lactate, respectively. As detailed by Collard et al. [46] these nonstandard metabolites are inhibitory for other enzymes and severely interfere with in carbohydrate metabolism. (c) A repair enzyme, known as phosphoglycolate phosphatase (PGP) in mammals, is able to dephosphorylate both compounds, reconverting them to harmless products. (d) The mammalian PGP is also very active with phosphoglycolate, but it shows little or no detectable activity towards standard glycolytic intermediates such as 2-phosphoglycerate.

Figure 3 – The canonical reaction of 3-phosphoglycerate dehydrogenase (PGDH) interconverts the glycolytic intermediate D-3-phosphoglycerate into phosphohydroxypyruvate (PHP), which is the first step of the phosphorylated pathway for L-serine synthesis (upper part of the panel). The human enzyme also catalyzes, as a side reaction, the reduction of α-ketoglutarate (α-KG) to D-2-hydroxyglutarate (D-2-HG; middle part of the panel) [70]; this compound is considered an oncometabolite and its accumulation in the cell is countered by a dedicated repair enzyme, D-2-hydroxyglutarate dehydrogenase (D2HGDH; bottom part of the panel) [45]. While in humans the canonical

and side reaction of PGDH are not connected, they are tightly coupled in the *E. coli* enzyme. This enzyme does not release NADH at the end of the canonical reaction, but requires that the coenzyme be re-oxidized *in situ* by α-KG [72]. The cycle that ensue (including the activity of D2GDH to recycle D-2-HG) helps drive thermodynamically the production of PHP and the overall pathway [71].

Text Box 1 – **Usage of the term 'promiscuity' in this review and in the literature**

This paper examines the capacity of enzymes to transform, with variable efficiency, different, physiologically available substrates. As a shorthand to indicate such a property, I use herein the expression "substrate promiscuity". This is close to the inclusive (non-rigorous) usage of the term promiscuity adopted in a large part of the metabolic literature (e.g., [1, 7, 33, 44, 52, 53, 56, 65, 66, 73]). Note however that evolutionary biochemists assign to 'promiscuity' a much more restricted definition. According to this definition, promiscuous enzymes are solely those possessing side activities that are physiologically irrelevant (e.g., because the alternative substrates are never available in the cell, or because the secondary activity is too weak to have an impact on organismal fitness) [25, 74]. Conversely, enzymes catalyzing multiple reactions that are biologically relevant are termed 'multifunctional' [35, 74]. This distinction assumes the absence of evolutionary pressures on 'truly promiscuous' activities – an assumption that is useful in many analyses. However, several examples in this review will suggest that (i),when dealing with enzymes that act on different metabolites, drawing clear boundaries between 'relevant' and 'irrelevant' alternative reactions may be difficult or nearly impossible, and (ii) overall, true promiscuity (as defined by evolutionary biochemists) might be less common than frequently assumed. Note also that the term 'multifunctionality' (or moonlighting [75]) suggests a usefulness of the alternative activities of an enzyme; however sometimes side activities may be biologically relevant *in the sense that they are detrimental* and therefore subject to negative selective pressures – again, as suggested by some examples in this review.

Text Box 2 – Specificity and binding energy

In transition state theory, the transition state is treated as if it were a stable entity in thermodynamic quasi-equilibrium with the ground state (i.e., the reactant(s)). In a simple situation in which an enzyme reacts with a single substrate, k_{cat}/K_M describes precisely such an equilibrium (panel a below) [76]. Accordingly, the logarithm of k_{cat}/k_M will be proportional to the free-energy difference between the free enzyme and substrate and the transition state complex ($\Delta G^{\neq}$; see panel b in the scheme below) [12]. Thus, whenever transition state theory is applicable, differences in k_{cat}/K_M between alternative substrates reflect their differential binding interactions in *the transition state.* Specificity can arise from

differential interactions already formed in the initial enzyme-substrate complex (in which case, they will contribute mostly to K_M; panel c) or forming only in the transition state complex (hence contributing to k_{cat}; panel d) but the total amount of binding energy associated to a specific group will be the same irrespective of where, along the reaction coordinates, the group establishes interactions with the enzyme.