Resource uptake and the evolution of moderately efficient enzymes

Florian Labourel[*,1] and Etienne Rajon[1]

[1]Univ Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR5558, F-69622 Villeurbanne, France

November 8, 2020

Abstract

Enzymes speed up reactions that would otherwise be too slow to sustain the metabolism of self-replicators. Yet, most enzymes seem only moderately efficient, exhibiting kinetic parameters orders of magnitude lower than their expected physically achievable maxima. Here, we question how these parameters evolve using a mechanistic model where enzyme efficiency is a key component of individual competition for resources. We show that kinetic parameters are under strong directional selection only up to a point, above which enzymes appear to evolve under near-neutrality. A majority of kinetic parameters compiled elsewhere do spread onto this plateau. Nonetheless, using a population genetics model that includes genetic drift and mutational biases, we show that this is a very unlikely outcome of evolution on a common landscape, as even very moderate biases towards lower efficiency should prevent the occurrence of such a diversity. Instead, differences between species, and within a species between metabolic pathways and the reactions to perform, should be involved. Our results point to drift playing an important role, along with the kinetics of nutrient transporters, the tolerance to high concentrations of intermediate metabolites, and the reversibility of reactions. Enzyme concentration also shapes selection on kinetic parameters, suggesting that the joint evolution of concentration and efficiency, facilitated by the plateau, should matter. Interestingly, the position of an enzyme along the metabolic pathway is not key for its evolution, contrasting with the prediction of models assuming that fitness depends on a precise level of flux control, rather than on competitive abilities.

1 Introduction

Living organisms need to uptake nutrients and metabolize them, relying on enzymes to catalyze chemical reactions along metabolic pathways. Accordingly, and despite being intrinsically reversible (Haldane, 1930; Klipp and Heinrich, 1994), *in vivo* enzyme-catalyzed reactions are commonly thought of as an irreversible two-step process (Bar-Even *et al.*, 2011, 2015; Michaelis and Menten, 1913; Johnson and Goody, 2011):

$$\mathrm{E} + \mathrm{S} \underset{\mathrm{k}_r}{\overset{\mathrm{k}_f}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{\mathrm{k}_{\mathrm{cat}}} \mathrm{E} + \mathrm{P}, \tag{1}$$

*Corresponding author: `florian.labourel@univ-lyon1.fr`

where k_f and k_r are the rates of association and dissociation between enzyme and substrate, and k_{cat} is the turnover number, that is the rate of formation of the product P from ES complexes. The first part of this chemical equation describes the encounters between the enzyme E and the substrate S; the enzyme will be efficient if ES complexes form often and do not dissociate before the substrate has been turned into a product, which is reflected by the constant $K_{\text{M}} = \frac{k_{\text{r}} + k_{\text{cat}}}{k_{\text{f}}}$. The efficiency v of an enzyme – the rate at which it makes a product P from S – depends on these two constants through equation (2):

$$v = k_{cat}.[E_{\text{tot}}].\frac{[S]}{K_{\text{M}} + [S]}, \tag{2}$$

under the assumption that the concentration $[S]$ is approximately constant and that of $[ES]$ is at equilibrium (Michaelis and Menten, 1913; Briggs and Haldane, 1925).

At first glance, Natural selection is presumed to optimize enzymatic efficiency by pushing k_{cat} upwards and K_{M} downwards to universal physical limits. Enzyme efficiencies are for instance limited by the diffusion properties of their aqueous environment, which sets an upper bound of approximately $10^8 - 10^{10} M^{-1} s^{-1}$ for the ratio k_{cat}/K_M (Alberty and Hammes, 1958; Zhou and Zhong, 1982).

Nearly optimal enzymes indeed seem to exist, as exemplified by triosephosphate isomerase (TIM) whose ratio is close to this theoretical ceiling (Knowles and Albery, 1977). But they are uncommon: most enzymes appear to be only moderately efficient and far off these physical limits – including enzymes immediately flanking TIM in the glycolysis metabolic pathway (Davidi *et al.*, 2018). Indeed, Bar-Even *et al.* (2011) have analysed a dataset of several hundreds of enzymes and found a wide diversity among enzyme parameters, thus sketching a puzzling pattern that has far more in common with a zoo than it looks like variations around an archetypal form.

To date, three sources of explanations have been invoked to explain both this variance in parameter estimates and the fact that many look suboptimal: (1) *in vitro* estimates do not reflect the *in vivo* functioning of enzymes, (2) physical constraints vary between enzymes such that their achievable optima differ, and (3) Evolution fails to produce optimum properties.

Surely, differences between enzyme behavior *in vivo* and *in vitro* are expected, first because diffusion in a test tube is hardly comparable to diffusion in the cytoplasm (Ellis, 2001; Rivas *et al.*, 2004; Zhou *et al.*, 2008; Rivas and Minton, 2018). As the cytoplasm gets packed, the cell approaches a state where molecules are less mobile, hindering substrate-enzyme encounters (Muramatsu and Minton, 1988; Zimmerman and Minton, 1993; Blanco *et al.*, 2018). In this regard, K_{M} values are thus underestimated *in vitro*, and enzymes should perform less efficiently *in vivo*. Simultaneously, macromolecular crowding can sometimes improve catalytic activity *in vivo*, making specificity constants k_{cat}/K_M higher than their *in vitro* estimates (Ralston, 1990; Ellis, 2001; Jiang and Guo, 2007; Pozdnyakova and Wittung-Stafshede, 2010). These effects are obviously important for our understanding of enzyme evolution but do not alone provide a reasonable explanation to the reported patterns insofar as some of them should be partially correlated between enzymes (Tabaka *et al.*, 2014), and their reported estimates are rather weak Davidi *et al.* (2016).

Besides the diffusion limit already mentioned, constraints on enzyme evolution also include an intrinsic trade-off that originates from the dependency of both k_{cat} and K_M on intermediate energy profiles (Heinrich *et al.*, 1991). Nonetheless, this trade-off is scarcely observable among evolved enzymes – Bar-Even *et al.* (2011) report a coefficient of determination around 0.09 for the correlation between $\log_{10}(k_{\text{cat}})$ and $\log_{10}(K_M)$ – suggesting that

it can be overcome. Other constraints may exist and be specific of a given reaction (Klipp and Heinrich, 1994) – *e.g.* reaction reversibility – potentially explaining a part of the variance in enzyme efficiencies. Estimating constraints on all individual enzymes appears like a daunting task, which could be guided, in part, by the identification of deviations from evolutionary predictions.

Enzyme evolution has often been considered theoretically through the lens of flux control (Burns *et al.*, 1985; Fell, 1992; Kacser *et al.*, 1995). Indeed, Kacser and Burns (1973) along with Heinrich and Rapoport (1974) brought to light that the control of the flux in a metabolic pathway is shared between all enzymes, in what is known as the summation theorem. Thence, because the sum of control coefficients must equal 1 within a pathway, if all enzymes have similar kinetic parameters, none of them exerts a strong influence. But if one enzyme departs from this trend and becomes inefficient, it exerts a strong control at the expense of others. This leads to diminishing-returns epistasis because, as an enzyme becomes more efficient, subsequent mutations have smaller effects (Kacser and Burns, 1973; Dykhuizen *et al.*, 1987; Tokuriki *et al.*, 2012), a finding that have since received empirical confirmation (Fell, 1992; Lunzer *et al.*, 2005).

Hartl *et al.* (1985) have considered such a fitness landscape under a population genetics framework to conclude that enzymes may be evolving on nearly neutral landscapes (Ohta, 1992). More recently, based on similar evolutionary models and a fitness function coming from flux balance analysis, Heckmann *et al.* (2018) have furthered the idea, compellingly showing how diminishing returns epistasis could generate the wide expected variety of k_{cat}s at the level of an organism. In these models as in most, an enzyme's efficiency is captured by its activity, a composite of k_{cat}, K_M and enzyme concentration (Hartl *et al.*, 1985; Kaltenbach and Tokuriki, 2014), such that the distinct evolutionary dynamics of these parameters is ignored.

Perhaps more importantly, Heinrich *et al.* (1991) and Schuster *et al.* (2008) have pointed out that these frameworks assume a constant value for either or both concentrations of the first substrate and of the final product (Orth *et al.*, 2010), whereas Evolution should instead maximize the amount of final products generated. Klipp and Heinrich (1994) found that the aforementioned concentrations indeed have a major influence on optimal rate constants under certain assumptions. Likewise, nutrient uptake is most often not considered explicitly in existing models of enzyme evolution, while it is obviously critical in the competition for resources.

Nutrient uptake occurs when metabolites move inwards across cell membranes; it may rely on membrane permeability only (passive diffusion) or involve channels and carrier proteins, be they transporters or cotransporters (Stein, 1986a). Here we build a model that explicitly includes passive (PD hereafter) or facilitated diffusion (FD) followed by an unbranched metabolic pathway to study how resource availability coupled to transport modulates the evolution of enzymes along the pathway. In ecological scenarios where individuals compete for resources, Natural Selection should favour genotypes that maximize the net intake of molecules and their transformation, which are linked under both PD and FD. Based on this premise, we show that the evolution of the first enzyme's kinetic parameters k_f and k_{cat} takes place on cliff-like fitness landscapes where a fitness plateau covers a wide part of the relevant parameter space. Kinetic parameters have co-dependent but distinct evolutionary dynamics – and thus distinct sensitivities to certain parameters of the model – such that the shape of the plateau can be modulated by changing parameters of the model within realistic ranges.

We further show that most empirical estimates for kinetic parameters compiled by Bar-Even *et al.* (2011) are spread on the fitness plateau of a generic fitness landscape. But we also demonstrate, using a simple population genetics model, that this is a highly unexpected evolutionary outcome when we introduce biases – even very

mild – for the mutations of kinetic parameters. We thus explore the effect of several parameters that may make the fitness landscape of each enzyme unique, thus explaining, at least in part, their exceptional diversity.

2 Results

2.1 Passive diffusion is generally inadequate to sustain cell metabolism

In the version of our model in which intake relies on passive diffusion (PD), the net uptake of a nutrient is a direct outcome of its concentration gradient, and therefore of how efficiently the first enzyme catalyses its transformation inside the cell. Assuming that fitness is proportional to the flux of product of this reaction, we find that the fitness landscape has a cliff-like shape with fitness increasing steeply as parameters k_{cat} and k_f increase (see Supplementary materials, section Text S1). The precise shape will not be commented in detail here, for it is very similar to landscapes obtained under facilitated diffusion (FD, treated in the rest of this manuscript).

Importantly, our results indicate that PD can only sustain a small part of the metabolism of most living cells given cell permeabilities reported in the literature (Wood *et al.*, 1968; Milo *et al.*, 2010), suggesting that this process may not be a determining factor in the evolution of enzymes along metabolic pathways. Indeed, even extremely efficient enzymes, harbouring values of k_{cat} and $k_{\text{cat}}/K_{\text{M}}$ close to their physical limits, yield low inward fluxes that approach $10^{-2}\ mM.s^{-1}$ when considering a spherical cell with a diameter $D = 1\mu m$. To get a sense of how low these fluxes are, we calculated the maximum cell size that they can theoretically sustain. Considering that basal metabolic demands are approximately proportional to the cell volume and using estimates by Lynch and Marinov (2015) for this relationship, we predicted the maximum size enabled by sugar passive diffusion (see 4Materials and Methods). Setting a (conservatively high) medium concentration in glucose $[G] = 1$M yields a theoretical volume ceiling $V_{est} = 0.84\mu m^3$.

Nearly all eukaryotes, and most prokaryotes are *de facto* larger than this threshold (Heim *et al.*, 2017), which might help explain the apparent ubiquity of FD. While this demonstration hinges on numbers for sugar uptake, which may arguably be the task requiring the highest flux, PD may be limiting for many other metabolites (Boer *et al.*, 2010), depending on their permeability and availability in the environment: even for very high amino-acids concentrations that may only be met in multicellular organisms (Schmidt *et al.*, 2016) and assuming the highest observed permeability for such metabolites (Chakrabarti, 1994), these levels are orders of magnitude lower than with FD (see Supplementary material - section Text S1 for PD results).

Our calculations therefore suggest that primordial cells, preceding the evolution of FD, must have been extremely small. Current-sized cells, however, must predominantly rely on FD when concentration gradients can be maintained, and on active transport otherwise.

2.2 General shape of the fitness landscape under facilitated diffusion

For most metabolites, FD relies on the specific binding of the substrate to transmembrane carrier proteins (transporters hereafter), followed by its translocation to the other side of the membrane (Danielli, 1954; Kotyk, 1967; Stein, 1986b). Our model builds on ter Kuile and Cook (1994)'s approach to model FD, considering the simplifying assumption of symmetric transport. Within this framework, FD operates on the concentration gradient and

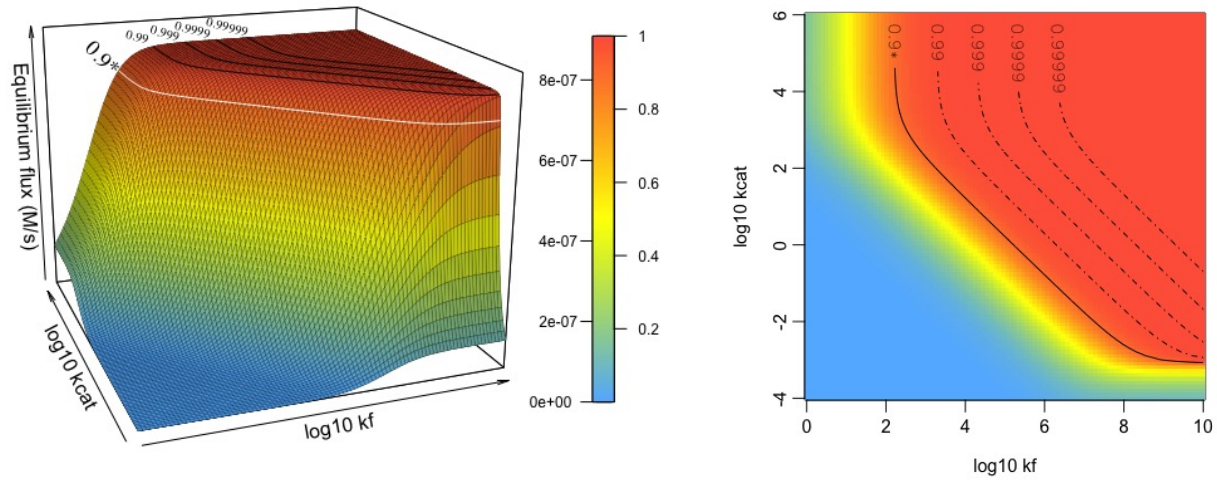

Figure 1: The flux of product following substrate uptake by transporters and conversion by a dedicated enzyme depends on the kinetic parameters k_f and k_{cat} of the enzyme. This landscape is generic in the sense that we avoid focusing on sugars – arguably extreme cases in terms of flux intensity – for a general description, considering a moderate flux at saturation $V_{Tm} = 1\mu M.s^{-1}$ close to those measured for amino acids and nucleosides in E.*coli* (Zampieri *et al.*, 2019). We also set the transport saturation ratio $[S_{\text{out}}]/K_{\text{T}}$ to 10 such that the FD process approaches saturation, and high transporter affinity $K_{\text{T}} = 50\mu M$, also in line with estimates for nucleosides (Griffith and Jarvis, 1996; Xie *et al.*, 2004) (see Material and Methods). Other parameter values include $k_r = 10^3 s^{-1}$ and $[E_{tot}] = 1mM$. The color gradient shows the absolute and normalized (such that the maximum flux equals 1) values of equilibrium flux.

is susceptible to saturation, represented by constant K_T – similar to K_M in the Michaelis-Menten equation – and an interaction constant α (see Methods for details). We assessed how this saturation phenomenon influences the selection pressure acting on forward enzyme kinetic parameters (k_f and k_{cat}) under various scenarios.

In order to depict a fitness landscape representative of an average enzyme, we first consider a situation where transporters saturate at an intermediate external concentration of the substrate with an intermediate affinity (FIG. 1). In this situation, the inward equilibrium flux (which, as argued in the introduction, can be considered representative of fitness) forms a plateau when the upstream enzyme in the metabolic pathway has high k_{cat} and k_f. This low equilibrium flux elasticity coincides with the saturation theory (Wright, 1934; Kacser and Burns, 1973), especially with its version incorporating facilitated diffusion (ter Kuile and Cook, 1994). The flux plateau is delineated by parallel isoclines (solid and interrupted lines in FIG. 1) oriented in the bottom-right direction of the landscape for intermediate values of k_{cat} and k_f, such that decreasing k_f by one order of magnitude can be compensated by a similar increase in k_{cat}. While this mutual dependency holds even for high k_f values as long as k_{cat} is not critically low (*i.e.* when $k_{\text{cat}} > 10^{-3}$), it stops when $k_{\text{cat}} \geq 10^3$, where increasing k_{cat} no longer improves fitness. Besides, the influence of k_{cat} and k_f is not strictly equivalent, since the increase in flux is more gradual in response to k_f.

Furthermore, and contrary to the textbook picture whereby most biological reactions are not limited by diffusion at all (Bar-Even *et al.*, 2011; Sweetlove and Fernie, 2018), increasing an enzyme's association rate k_f – be it through its diffusivity or its binding rate – may still enhance the equilibrium flux when diffusion is substantially faster than catalysis.

2.3 Properties of facilitated diffusion modulate the landscape

To explore the effect of FD kinetics on the evolution of enzymes in the metabolic pathway, we studied the influence of K_T – the affinity of the transporter for the substrate – and V_{Tm} – the maximum transport rate – still assuming that the substrate is close to saturation ($[S_{\text{out}}]/K_T = 10$). We considered ranges of empirical estimates for sugars (high flux with low to moderate affinity in FIG. 2, (Stein, 1986b; Maier *et al.*, 2002)), nucleosides (Griffith and Jarvis, 1996) and amino acids (Stein, 1986b; Zampieri *et al.*, 2019) (weak to moderate flux with moderate to high affinity in FIG. 2).

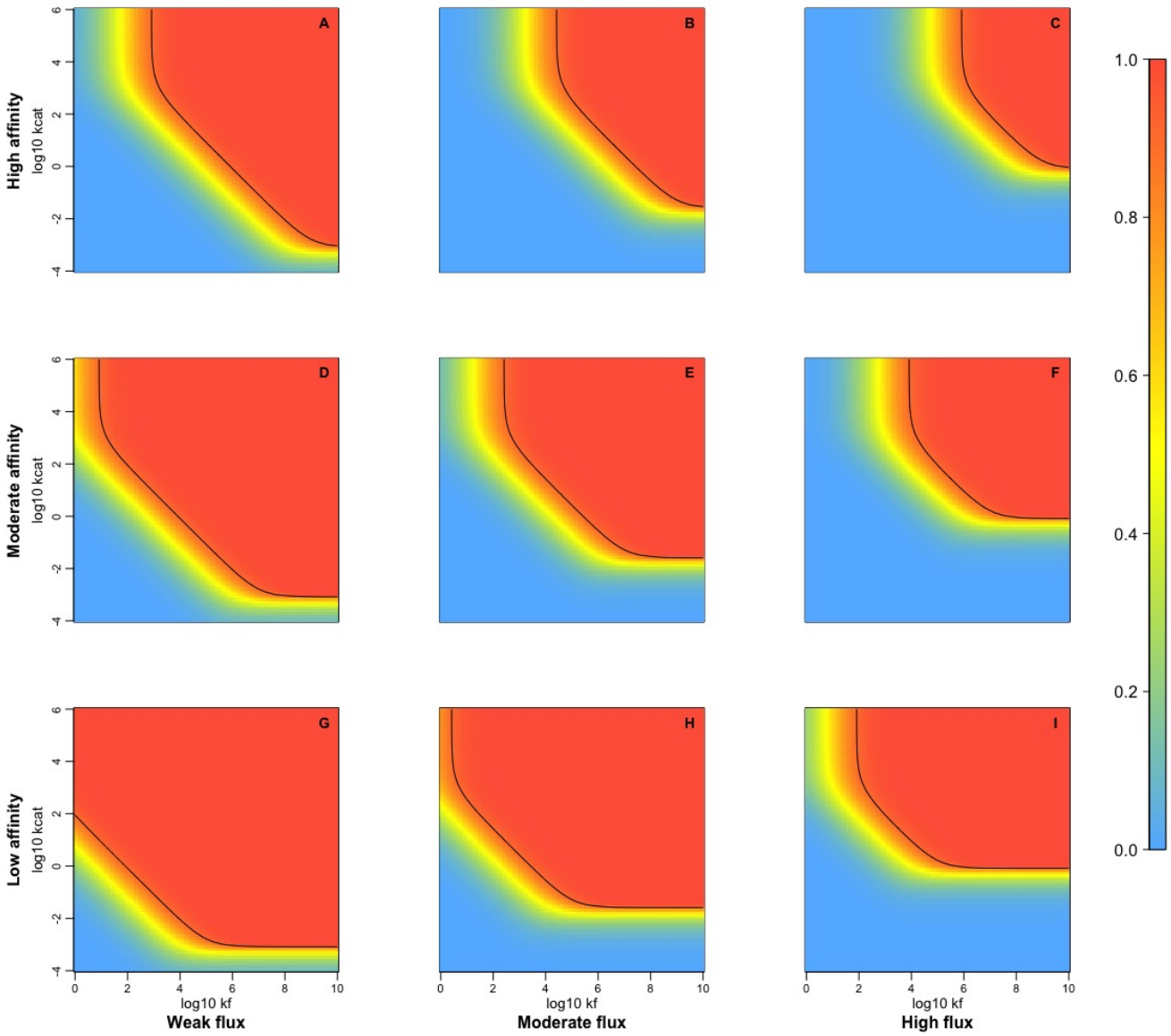

Figure 2: Both the affinity and rate of a transporter have an impact on the (normalized) flux landscape for upstream enzymes, the black isocline (corresponding to 0.9) delineates the fitness plateau. Each plot represents the landscape obtained with a pair of values for transporters affinity K_T and saturation V_{Tm}. Moving one step to the right means that V_{Tm} increases by 1.5 orders of magnitude – from 10^{-6} (low flux) to $10^{-3} M.s^{-1}$ (high) – and one step up means that K_T decreases by 2 orders of magnitude, starting at $10^{-1} M$ (low affinity). Increasing K_T extends the plateau only towards the left part of the landscape, allowing enzymes with lower k_f on the plateau, whereas decreasing V_{Tm} extends the plateau in both directions. Other parameter values: $k_r = 1000/s$, $[E_{tot}] = 1mM$ and $[S_{env}] = 10 \times K_T$.

We find that increasing the transport flux V_{Tm} exerts a positive selection pressure on kinetic parameters for the upstream enzyme (*i.e.* for increasing k_{cat} and k_f). The plateau is shifted accordingly, towards the top-right corner of the landscape, at a distance that corresponds to the magnitude of the change in V_{Tm}. Increasing the affinity of the transporter (*i.e.* decreasing K_T), however, selects for higher k_f (the isoclines are displaced to the right and the fold change is similar to that of K_T) but has no visible influence on k_{cat}, a result that holds regardless of the flux at saturation V_{Tm}. This specific effect on the affinity of the upstream enzyme is likely due to a competition between the transporter – which can transport the substrate in both directions – and the enzyme, which harvests the substrate at a rate that depends on the dissociation constant $K_D = k_r/k_f$. It should be noted that nutrients under lower demands – *e.g.* amino acids – are generally less concentrated in the environment, often coinciding with a higher affinity of their transporter. Therefore, the possible combinations of flux and affinity likely occupy a restricted space of possibilities where flux and affinity are negatively linked, which as can be seen in FIG. 2-A,E,I results in landscapes that mainly differ by the minimum value of k_{cat} on the plateau.

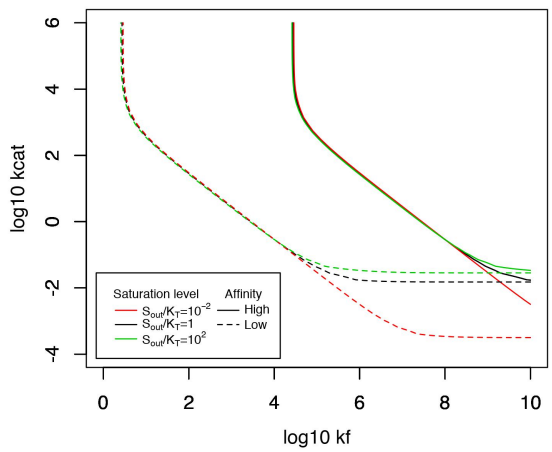

Figure 3: The shape of the fitness plateau – represented for each condition through the 0.9 isocline (see FIG. 1 for details) – is little dependent on the saturation of the transporter. Only enzymes downstream low affinity transporters show slightly relaxed selection on k_{cat} when the external concentration of the substrate is lower than the affinity constant of the transporter, K_T. The example shown is for a transporter with moderate flux ($V_{Tm} = 10^{-4.5} M.s^{-1}$); the effect is identical for other V_{Tm} (see Figure S2 of SM).

So far we have considered transporters saturated by high external substrate concentrations. Relaxing this assumption has little impact on the fitness landscape, except that very low values of k_{cat} (*i.e.* lower than 10^{-2}) can only sustain high fluxes at saturation (see FIG. 3). Diminishing returns shall thus be the rule for the first enzyme whatever the extracellular environment, and this phenomenon typically occurs within the known behavioral range for both catalysts and transporters.

2.4 Downstream enzymes also evolve on cliff-like fitness landscapes

In this section, we intend to draw fitness landscapes for enzymes catalysing reactions downstream in the metabolic pathway. Because the flux of the first product in a pathway increases with the substrate gradient across the cell membrane, the upstream enzyme of a given metabolic pathway is selected for efficiency within the limits mentioned in the previous two sections. In contrast, this selective pressure does not apply directly downstream; at

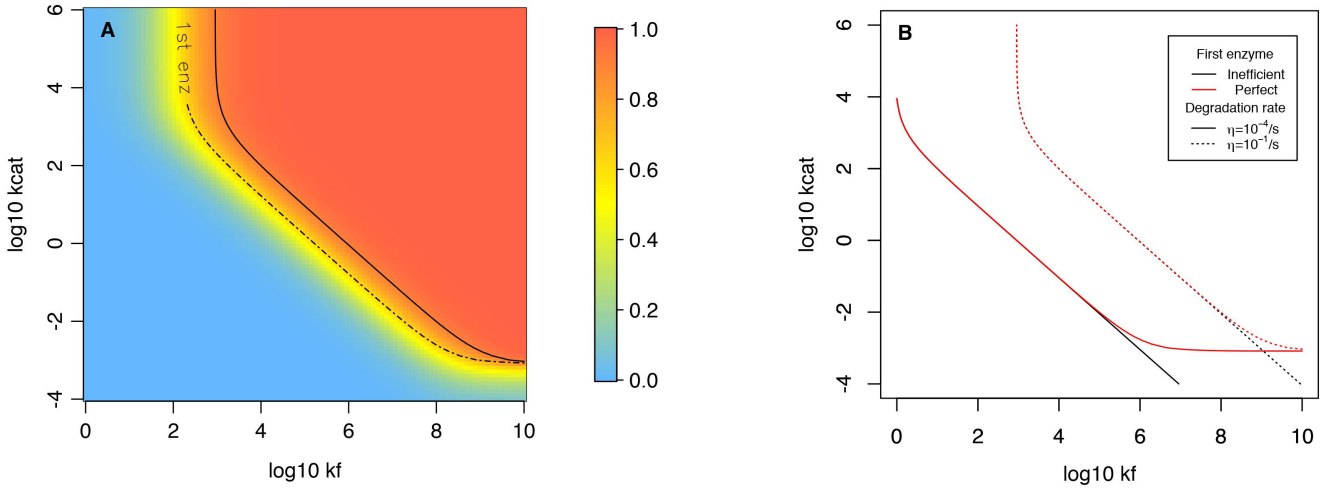

Figure 4: Downstream enzymes exhibit similar fitness landscapes as upstream, with a dependency to degradation parameter η_d. These enzymes are involved in a moderate flux pathway whose pace is driven by the same uptake kinetics as in FIG. 1 where $V_{Tm} = 1\mu M/s$, $K_T = 50\mu M$ and $[S_{\text{out}}] = 10K_T$. Settings for the second enzyme also correspond to a model case with $k_r = 1000s^{-1}$ and $[E_{tot}] = 1mM$. (A) corresponds to a high degradation rate $\eta_d = 10^{-1}s^{-1}$ that limits metabolite intracellular concentration to $[S]_{in} = 10\mu M$ (for the selected V_{Tm} value), with the upstream enzyme being perfect and concentrated ($k_f = 10^{10}M^{-1}.s^{-1}$, $k_{cat} = 10^6s^{-1}$, $k_r = 10^3s^{-1}$, $[E_{tot}] = 1mM$) and results in a fitness landscape very similar to that of the first enzyme when it faces similar conditions (findings are relatively similar for sugar like transporters, as reported in SM - Figure S4), while plateau isoclines drawn in (B) also show results for a second degradation rate limiting this concentration to $[S]_{in} = 10mM$, in the context of an upstream concentrated enzyme being either perfect (see above) or inefficient ($k_f = 10^2M^{-1}.s^{-1}$, $k_{cat} = 10^{-2}s^{-1}$, $k_r = 10^3s^{-1}$). Decreasing η_d makes the cell tolerant to higher concentrations of intermediate metabolites (the product of the first reaction), hence relieving the selection pressure on the second enzyme and moving the plateau towards the bottomleft corner.

steady-state, even inefficient enzymes can in principle process newly formed substrate molecules at an elevated rate, assuming that the concentration of the substrate is allowed to reach any equilibrium value. This is an obviously unreasonable assumption, since a part of this standing substrate should be lost by outward diffusion or degradation, or may interfere with other cellular processes due to the surging involvement in non-specific or promiscuous interactions (Khersonsky and Tawfik, 2010; Schäuble *et al.*, 2013). We included a degradation term in the model, proportional to the constant η_d (assuming that the external environment is infinite, the degradation term can as well represent an efflux). Modelling the cost of accumulation this way is likely also consistent with non-specific interactions that should follow Michaelis-Menten kinetics albeit with much lower affinities and rates than the focal reaction, which means their rates should follow a linear relationship up to very high cellular concentrations.

We first consider a "perfect", highly concentrated upstream enzyme ($k_f = 10^{10}M^{-1}s^{-1}$, $k_{cat} = 10^6s^{-1}$, $k_r = 10^3s^{-1}$, $[E_{tot}] = 10^{-3}M$) and study the second enzyme in the pathway, showing that it evolves on a fitness landscape that has a similar shape than described above, still hitting a plateau (FIG. 4, with the same parameterization as FIG. 1). The degradation rate creates a ceiling for the concentration of the product of the first reaction, such that reducing η_d allows for higher concentrations (see SM - Figure S3) and makes the flux tolerant to second enzymes with lower k_fs, whereas selection on k_{cat} is barely impacted by this parameter. The

plateau is therefore extended to the left when high product concentrations are enabled at low η_d (see FIG. 4-B). The shape of the plateau is little impacted by changes in the efficiency of the first enzyme, especially when it stands on the plateau. These conclusions are independent of characteristics of the transporter initiating the pathway, although degradation rates limiting metabolite concentrations to similar levels pushes the plateau to the right slightly more when compared to the landscape of the first enzyme (see SM - Figure S4 for the case of low affinity - high flux transporters). We can notice that the second enzyme's kinetic parameters evolve in response to the net rate v of the upstream reaction and to the presence of any process in competition for its product (like efflux or degradation), such that these results are expected to hold for any enzyme downstream.

2.5 Most enzymes lie on the fitness plateau

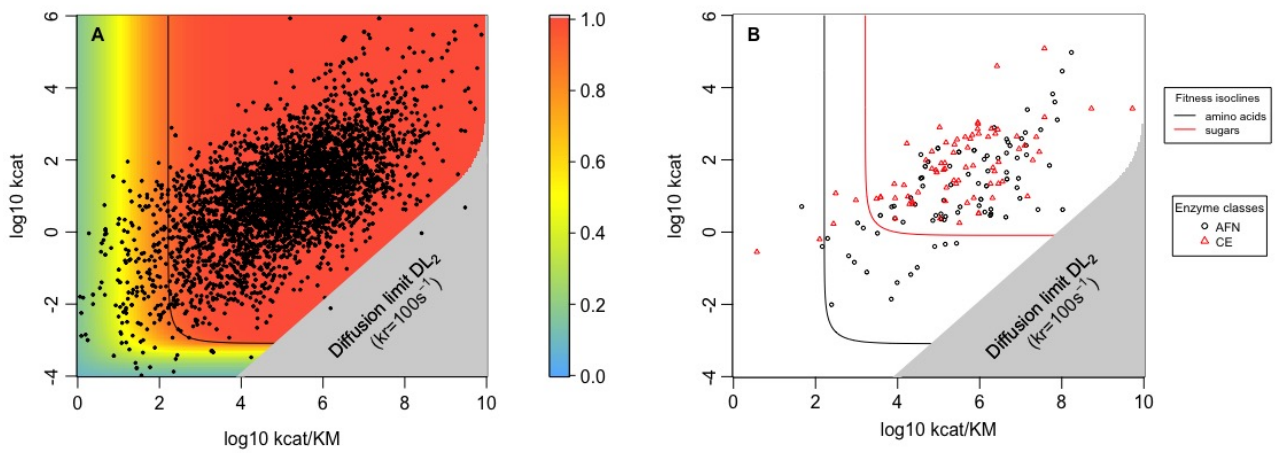

Figure 5: *in vitro* experimental estimates for enzyme kinetic parameters spread on the plateau of a generic theoretical fitness landscape. The parameter space considered corresponds to the usual empirical estimates for enzyme kinetic parameters k_{cat} and $kcat/K_M$. The space is narrowed down here due to the diffusion limit, outlined by the grey area when k_r is low ($10^2 s^{-1}$), that impedes high values for k_{cat}/K_M under low values for the ratio between k_{cat} and k_r. In (A), the full dataset from (Bar-Even *et al.*, 2011) is plotted in the parameter space and contrasted with the fitness landscape of the most upstream enzyme involved in amino acid or nucleoside metabolism (same parameters for transport than in Fig. 1, $[E_{tot}] = 10^{-3}M$ and $k_r = 10^2 s^{-1}$). Most enzymes fall on the plateau of the fitness landscape. A few outliers overcome the diffusion limit, suggesting that k_r might be lower than $10^2 s^{-1}$ for these enzymes. In (B), the dataset is divided into classes identified by Bar-Even *et al.* (2011), with the acronym denoting enzymes involved in the central metabolism including (AFN): amino acids, fatty acids, and nucleotides - and (CE): carbohydrates and energy. The isoclines shown correspond to parameter values typical of sugar transporters ($K_T = 5mM$,$V_{Tm} = 1mM.s^{-1}$) (Maier *et al.*, 2002) or amino acids transporters (same as in (A)). With a few exceptions, enzymes seem to spread on the corresponding fitness plateau.

We then superimposed empirical estimates of kinetic parameters over our theoretical fitness landscapes, after substituting parameter k_f for its usual empirical counterpart, k_{cat}/K_M. Because $k_{cat}/K_M \approx k_f k_{cat}/k_r$, this approximation only holds when $k_{\text{cat}} \gg k_r$. No less significantly, representing the fitness landscape in this parameter space sets an area in the bottomright part of the landscapes where k_f exceeds the diffusion limit (grey area on FIG. 5). For purposes of inclusiveness, we used $k_r = 10^2 s^{-1}$ by default – noting that this limit would be displaced upwards for larger k_r – and otherwise used the same set of parameters as in FIG. 1 to obtain FIG. 5.

We see that a vast majority of the enzymes in the Bar-Even *et al.* (2011) dataset belong to the fitness plateau. The implicit suggestion that this could be a result of Evolution on a common fitness landscape is surprising, as the dataset includes enzymes from many Eukaryote and Prokaryote species that contribute to various pathways. We instead anticipate that differences in evolutionary context should explain a part of the observed variance. We also note a strong linear positive relationship between k_{cat} and k_{cat}/K_M ($R^2 = 0.4$). This illustrates how focusing on k_{cat}/K_M can be misleading if one's interest lies in mechanistic processes and their implications. Indeed, as obvious as it may sound, k_{cat} and k_{cat}/K_M are not independent from each other, especially when $k_{\text{cat}} < k_r$ (which is the case for a majority of the data points in FIG. 5). Concluding on whether a positive relationship between k_f and k_{cat} does exist would require to estimate k_r for each enzyme.

Because we have previously shown that different transporters may yield plateaus with different shapes (*e.g.* for sugars and amino acids), our model predicts that enzymes involved in the corresponding pathways should have their own specific distributions. Representing findings by (Bar-Even *et al.*, 2011) suggests that there may indeed be an explainable difference between enzymes contributing to carbohydrate processing (in red in FIG. 5-B) and to that of amino acids (a part of the black dots in FIG. 5-B). This apparent discrepancy in fitness landscapes across pathways starting with different transporters can only explain a small part of the wide distribution of kinetic parameters reported by Bar-Even *et al.* (2011), thus requiring a careful study of other evolutionary, ecological and chemical parameters.

2.6 Evolutionary dynamics of enzyme kinetic parameters

To approach how populations evolve on the fitness landscapes drawn through our mathematical model, we built a simple population genetics model that includes, besides selection, mutation and drift. In this individual based model, absolute fitness is directly proportional to the flux of product at equilibrium – which itself equals the inward flux of nutrients. We consider the flux of the first enzyme (see Methods and previous Results sections for details), as presumably the evolutionary dynamics should be similar for downstream enzymes with similar landscapes. Two different levels of metabolic demands were considered, corresponding to parameter values used to draw panels (A) and (I) in FIG. 2 in order to control that differences in the delineation of the plateau (in the parameter space) do not affect the evolutionary equilibrium. For the enzyme of a given individual, only k_{cat} and k_f were susceptible to evolve through mutations. Mutational effects on $\log_{10} k_{cat}$ and $\log_{10} k_f$ were drawn from independent normal distributions with mean $b \leq 0$, with the absolute value of b setting the intensity of mutational towards less efficient parameter values. The standard deviation of the distribution of mutational effects was set to 0.3 such that most mutations explore the neighbouring parameter space, with rare mutations allowing changes by more than one order of magnitude (one $\log_{10}$ unit) at a time, accordingly to values reported in (Carlin *et al.*, 2016).

Without any mutational bias ($b = 0$), simulated enzymes spread all over the fitness plateau, as expected (FIG. 6A for low flux, Figure S7A of SM otherwise). The onset of the plateau depends on the strength of drift and hence derive from the effective population size N_e, following the classical expectation that selection becomes inefficient when $N_e \times s < 1$ (Kimura, 1968). Introducing a mutational bias that makes enzyme kinetics less efficient on average has a strong effect on both k_{cat} and k_f, preventing simulated enzymes from improving far above the drift barrier (FIG. 6B for low flux, Figure S7B of SM otherwise). Even weak biases ($b = 0.1$) lead to enzymes evolving in the vicinity of the isocline where $N_e \times s \approx 1$. Increasing the strength of this bias to 0.2 only

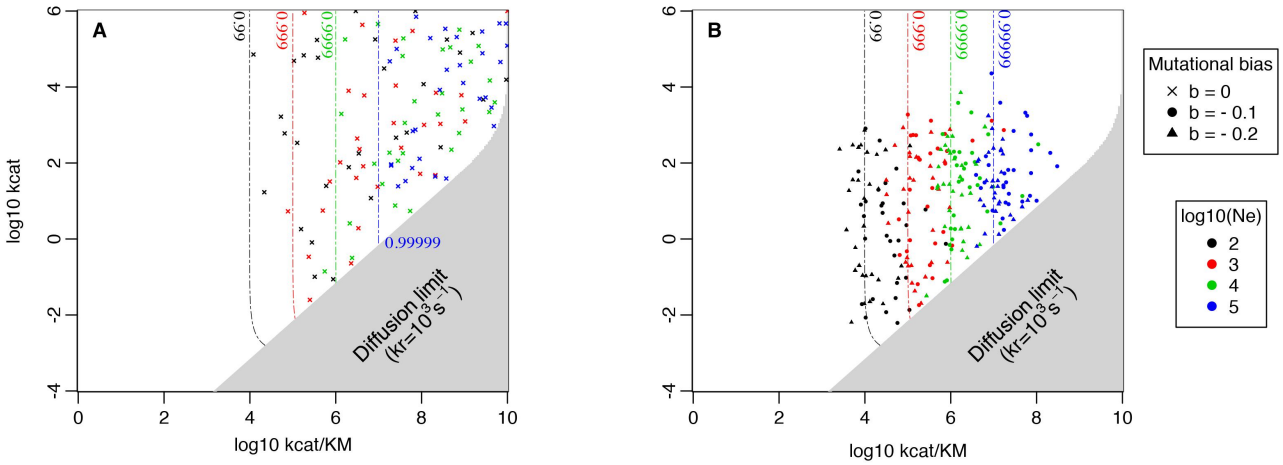

Figure 6: Simulations of a population genetics model indicate that enzymes should spread on a plateau corresponding to their effective population size N_e only when mutations are unbiased; mutations biases towards lower efficiencies make enzymes stick to isoclines, such that their evolution is predictable. The case considered here is that of a transporter with a low flux at saturation and high affinity ($V_{Tm} = 1\mu M s^{-1}$ and $K_T = 10\mu M$) – under various scenarios: four effective population sizes from 10^2 to 10^5 (different colors) and three cases of mutational biases(($b = 0$ corresponds to the absence of mutational bias). We ran 30 independent simulations for each scenario, each represented a dot in the "empirical" parameter space (k_{cat}, k_{cat}/K_M). Only k_{cat} and k_f were susceptible to evolve, while k_r was set to $10^3 s^{-1}$ such that the grey part of the parameter space is inaccessible to enzymes due to the diffusion limit. At evolutionary equilibrium, enzyme efficiencies reach a plateau delineated by an isocline indicative of effective selection. In (A), no mutational bias results in enzymes spreading onto the plateau, some reaching very high k_{cat} and/or k_{cat}/K_M values, while in (B) enzyme efficiencies stick to their predicted isoclines – under the Nearly Neutral Theory of Evolution (Ohta, 1992) – owing to the over-representation of mutations reducing enzyme efficiencies.

slightly decreases the population variance around this expectation.

In light of this effect of the effective population size on enzyme evolution, the seeming adequation between the plateau and the data in FIG. 5 appears puzzling, as the lowest N_e considered here do not encompass those traditionally found for unicellular organisms that mostly exceed 10^5 (Bobay and Ochman, 2018). In the same vein, Eukaryotes and Prokaryotes datasets display similar ranges despite large differences in effective population sizes (Bar-Even *et al.*, 2011). It is thus possible that the role of genetic drift is not well captured by a model that considers small parts of a large system in isolation, or more generally that the link between fitness and flux is not as straightforward as we have assumed. Notwithstanding this issue, the prediction of enzymes evolving a predictable set of kinetics parameters when mutational biases are considered contrasts with the broad variability reported (Davidi *et al.*, 2018), thereupon requiring further investigation.

2.7 Explaining the variance within a pathway

So far, our model predicts that enzymes intervening in a given pathway should evolve on a common fitness landscape, provided that the accumulation of intermediate metabolites is highly detrimental. This should produce similar enzymes along a pathway, which is directly contradicted by empirical estimates, suggesting that each enzyme instead evolves on its own fitness landscape. In this section we show that (1) physical constraints, among which in first place is the reversibility of reactions and (2) the joint evolution of kinetic parameters and other

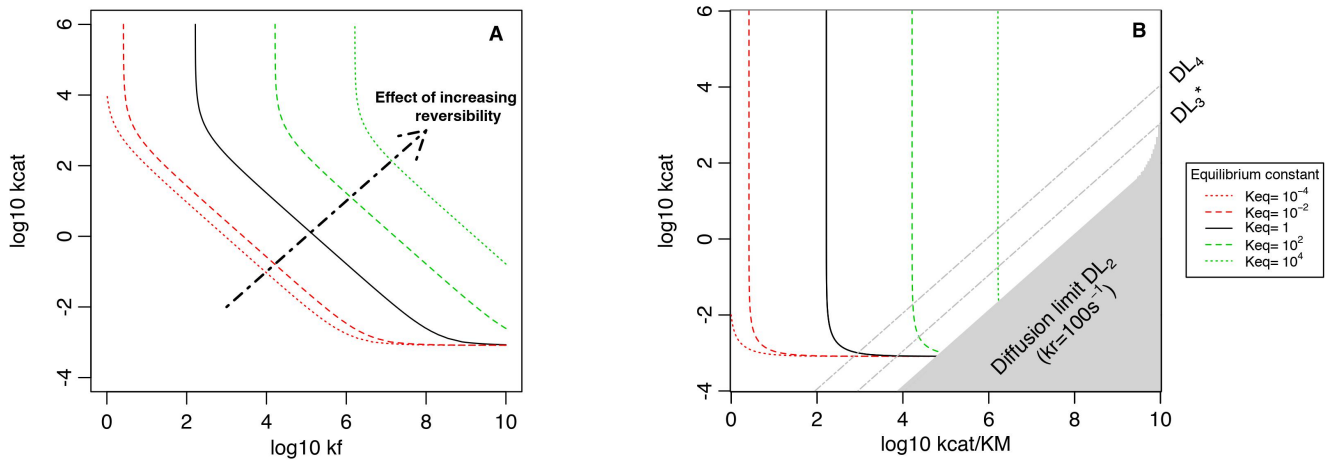

Figure 7: Backwards reaction rates of an enzyme directly upstream have a strong impact on the fitness landscape. Both plots show results of the influence of the reversibility of the first reaction on the fitness landscape of the next enzyme. Parameters are identical as in FIG. 1, corresponding to the model case for amino acids, with $k_r = 10^3 s^{-1}$. K_{eq} equals $[S]_{eq}/[P]_{eq} = k_r k_{inh}/k_{\text{cat}} k_f$ (Klipp and Heinrich, 1994) and quantifies the degree of reversibility, a low K_{eq} featuring low reversibility and vice versa. The first enzyme is a nearly perfect forward enzyme with $k_f = 10^8 M^{-1}$ and $k_{cat} = 10^4 s^{-1}$. Reversibility was equally spread between the two backwards parameters (*e.g.* $K_{eq} = 10^2$ yields $k_r = 10^1 k_{cat}$ and $k_{inh} = 10^1 k_f$), and a low degradation rate was considered ($\eta_d = 10^{-4} s^{-1}$). In (A), results are plotted in the theoretical parameter space (k_f, k_{cat}), showing that any increase in reversibility increases the pressure on enzyme kinetics by the same magnitude – except when reactions are highly non-reversible (in red) . In (B), the same results are shown in the experimenter parameter space of the second enzyme, showing that there is an increased pressure on k_{cat}/K_M under higher reversibility. While the plateau is not moved upwards, indicative of a selection on k_{cat} independent on reversibility when k_{cat}/K_M is fixed, positive selection for this parameter may still arise due to the diffusion limit that precludes the access to the lowest k_{cat} at high k_{cat}/K_M. The diffusion limit should play an important role for enzymes with a high dissociation rate k_r, as illustrated by the delineation of the diffusion limit (grey area or grey dashed lines) corresponding to several k_r values (eg. DL_3 stands for $k_r = 10^3 s^{-1}$; the star indicates that it is the case represented in (A)).

evolving parameters, can explain large differences among enzyme kinetic parameters by influencing actual *in vivo* efficiencies.

Reversibility is an intrinsic feature of chemical reactions that cannot be directly overcome by Evolution (Haldane, 1930; Cornish-Bowden, 1979). A highly reversible reaction corresponds to a large intrinsic equilibrium constant $K_{eq} = [S]_{eq}/[P]_{eq}$ (Klipp and Heinrich, 1994), and results in higher backward than forward rates in the following chemical equation:

$$\text{E} + \text{S} \underset{\text{k}_r}{\overset{\text{k}_f}{\rightleftharpoons}} \text{ES} \underset{\text{k}_{\text{inh}}}{\overset{\text{k}_{\text{cat}}}{\rightleftharpoons}} \text{E} + \text{P}_1, \tag{3}$$

where k_{inh} represents the rate at which enzyme and product combine back. Such a (reversible) reaction could in principle influence the selective pressure acting on the following enzyme in the pathway, for both enzymes compete to process the same metabolite P_1. We thus quantified how reversibility affects the evolution of an enzyme downstream (FIG. 7 ; using the same model case parameters used throughout the paper (see FIG. 1 for details) and considering a nearly perfect first enzyme for consistency: $k_{cat} = 10^4 s^{-1}$,$k_f = 10^8 M^{-1} s^{-1}$).

The equilibrium constant K_{eq} has a similar (non-linear) impact on the fitness landscape of the second enzyme

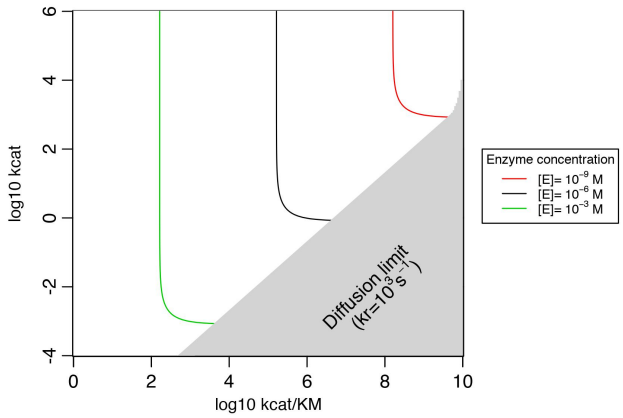

Figure 8: High enzyme concentration $[E_{tot}]$ alleviate selection on both k_{cat}/K_M and k_{cat}. The parameters for transport are the same than in FIG. 1, with $k_r = 10^3 s^{-1}$.

than the degradation rate, with a highly reversible upstream enzyme exerting a selection pressure downstream towards an increase of kinetic parameters (FIG. 7-A). Indeed, increasing K_{eq} moves the fitness plateau toward the upper-right corner in the (k_f,k_{cat}) parameter space, hence selecting for more efficient downstream enzymes. The effect appears linear, except for very low values of K_{eq} where it fades – such that a 100-fold change in K_{eq} has little impact – because the main issue when reactions are highly non-reversible becomes, again, metabolite accumulation. Therefore, the reversibility of the upstream reaction appears like a critical parameter for the evolution of an enzyme, able to generate large changes in evolutionarily expected kinetic parameters. When using the "empirical" parameter space (FIG. 7-B), we observe that increasing reversibility only selects for higher k_{cat}/K_M. Interestingly, the unbinding rate k_r of an enzyme, which is correlated to its reversibility, may eventually favour a positive codependency between the forward rates k_{cat} and k_f as the latter is no longer sufficient to ensure a high k_{cat}/K_M (see FIG. 7-B and Figure S9 of SM for this influence in the theoretical parameter space).

Hitherto, we have considered enzymes to be highly concentrated, an assumption that we relax here. Predictably, increasing the concentration of the first or second enzyme in a pathway releases the selection on their kinetic parameters (Noor *et al.*, 2016), producing larger fitness plateaus as an enzyme concentration increases (see FIG. 8 for the first enzyme and Figure S10 of SM for the second). Because the concentration is subject to heritable changes (Schaefke *et al.*, 2013), we anticipate that the joint evolutionary dynamics of the concentration and kinetic parameters should yield a negative correlation between them, as reported by Davidi *et al.* (2016, 2018) (who focused on $k_{\text{cat}}[E]_{tot}$), due to the compensatory effect of concentration.

3 Discussion

Most enzymes have been considered to be only moderately efficient, if not sloppy (Bar-Even *et al.*, 2011, 2015). This claim was put into perspective by Newton *et al.* (2018) who argued that the link between fitness and enzyme efficiencies is complex and may be partly enzyme-dependent. Through this work, we have developed a model where enzyme efficiencies are mechanistically linked to fitness through the impact of nutrient gradients on the production of metabolites. Our results emphasize that enzymes might in fact be quite efficient and, possibly, as efficient as Evolution allows.

We have modelled two processes of nutrient uptake that rely heavily on concentration gradients: passive (PD)

and facilitated diffusion (FD). Our results apply to the numerous metabolic pathways that start with them, and also possibly to the widely used secondary active transport (SAT) in which FD of ion-substrate complexes follows the active extraction of a cotransported ion. Indeed, SAT can be described by similar equations as FD, with the difference that energy is required for extraction (Stein, 1986c).

Our results indicate that an enzyme's forward kinetic parameters – k_{cat} and k_f – evolve on cliff-like fitness landscapes with a plateau – where the flux of energy generated increases on a log scale – surrounded by a steep slope. This conclusion seems to agree qualitatively with the assumptions of previous models (Hartl *et al.*, 1985; Kaltenbach and Tokuriki, 2014), which we confirm here from a nutrient driven perspective. Importantly, our framework allows to consider the joint evolution of two kinetic parameters, k_f and k_{cat} instead of a composite "efficiency" whose relevance is questionable (Koshland, 2002; Eisenthal *et al.*, 2007), and to pinpoint model parameters that have a differential impact on each. This, along with our broad comparison of model predictions with a massive dataset, initiates an attempt to predict evolutionary trends at the level of individual enzymes. We show that this should be possible in principle, under the assumption that mutations of efficient enzymes are slightly biased towards making them less efficient, which limits evolutionary predictions to a narrow range in the parameter space.

Overall, the evolution of an enzyme should result from an interplay between (1) ecological and biochemical factors – metabolic demands, richness of the environment, reversibility of reactions – that draw the landscape, and (2) evolutionary factors – effective population size, distribution of mutational effects – governing how organisms move through evolutionary times on these landscapes. We have effectively found that the shape of the fitness landscape depends on features of the transporter, namely the maximum flux they can sustain and their affinity for the substrate, each having different effects on k_{cat} and k_f. These features are themselves evolving quantities, and the joint evolution of transporters and intracellular enzymes remains to be studied. Yet, using parameters that correspond to empirical estimates for sugars and amino acids, we have found that enzymes intervening in subsequent metabolic pathways should be different, with those in the "sugars" pathway being selected for faster kinetics. The focus in this study was not on interactions within a pathway and high-order epistasis, but our framework suggests that all enzymes but the first in a pathway should evolve on fitness landscapes that do not depend on their position, as long as all upstream enzymes belong to the fitness plateau. If one enzyme is noticeably inefficient, however, selection will push downstream enzymes towards a wider plateau, as a larger region of the parameter space can accommodate such a lower flux.

We also notice that different constraints may be met along a pathway and make each enzyme evolve on its own fitness landscape, possibly explaining a part of the large variance reported at this scale. Physical constraints may for instance act differentially on different enzymes, as exemplified by the intrinsic reversibility of a reaction that fuels the selective pressure towards higher efficiency in downstream enzymes. This may contribute to explain the high efficiency of a few specific enzymes, and thereby why there exists some "equilibrium enzymes" pushing reactions close to equilibrium (Williamson *et al.*, 1967; Davidi *et al.*, 2018). Aside from physical constraints, actual enzyme kinetic parameters are entangled with many more biochemical parameters, making the prediction of their evolution complex.

More specifically, we have examined the effect of one such parameter (but not its evolution), the concentration of the enzyme, showing that it indeed exerts an important role and is likely involved in the variance in k_f and k_{cat}, for low efficiency can be buffered by high levels of expression and vice versa. How enzyme levels of expression

and efficiencies jointly evolve, including the role of the cellular content on these efficiencies, should be the next step forward in the understanding of enzyme evolution. Because a cell is packed with many organites and biomolecules, making diffusion less efficient and k_fs lower *in vivo* than their *in vitro* counterparts, opposing forces govern the evolution of enzyme concentrations. Increasing concentrations may thus come at a great cost by slightly but steadily decreasing the actual efficiency of many catalysts at once, in turn favouring the evolution of highly efficient enzymes unless physical constraints prevent such an improvement.

Our work provides a mechanistic model for enzyme evolution, based on the assumption that enzymatic efficiency helps ingesting nutrients and win the competition for resources. Unfortunately, mechanistic does not mean free of a definition of fitness, as we have assumed that the latter is proportional to metabolic flux, hence considering each flux in isolation. Fitness instead results from a wide range of metabolic pathways that combine together and should all be competitive to certain degrees. How high-order epistasis builds up, and genetic drift acts in this context, is far from obvious; but this should not impact much how enzymes evolve in old, overall efficient pathways, as any impediment in efficiency should have an independent effect on fitness in this context, as captured by our model. Yet, this could make selection weaker in some or all pathways, as exemplified by Heckmann *et al.* (2018), and therefore help explain why they appear to be evolutionary suboptimal. But more importantly, such a model would be crucial to understand how metabolic pathways arose and improved, likely from a highly inefficient state during early steps in the evolution of life on Earth.

4 Materials and Methods

4.1 Quantifying the maximum size for cells using passive diffusion

If a cell is to be viable, it has, at least, to uptake enough glucose to compensate for basal metabolism – metabolism that allows to maintain the same cell size for non-actively growing cells (Lynch and Marinov, 2015) – leading to the following equation: $\phi_{PD} = C_M$, with ϕ_{PD} the uptake through passive diffusion and C_M the basal metabolism demand. To calculate the maximum size a cell can reach using only passive diffusion, we relied on the formula $C_M = 0.39V^{0.88}(10^9 ATP/hr)$ estimated in (Lynch and Marinov, 2015). We also assumed the cell to be of spherical shape, both concentrations – inside and outside the cell – to be constant with the cellular concentration staying so low that it can be overlooked, meaning that the uptake resulting from passive diffusion can merely be written as $\phi_{PD} = P.[S_{\text{out}}].\frac{SA_{\text{sphere}}}{V_{\text{sphere}}}$, where SA_{sphere} and V_{sphere} are the surface area and the volume of a sphere, and P represents the cell permeability and was measured to $10^{-6}\mu m^{-1}$ (Wood *et al.*, 1968) for glucose. Finally, we considered that each glucose yields 30 ATP molecules (Rich, 2003).

4.2 Flux sustained by the first enzyme

When assessing the flux of product made by the first enzyme in a pathway, both (PD) and (FD) result in similar sets of equations; we focus on FD here (see Text S5 - Mathematical appendix in Supplementary material for a comparison with PD). FD typically relies on the specific binding of substrate molecules – located outside the cell – by transmembrane carrier proteins, followed by their translocation into the cytoplasm (Danielli, 1954; Wilbrandt and Rosenberg, 1961; Kotyk, 1967). This specific process obeys Michaelis Menten-like kinetics when transport is assumed to be symmetric (Kotyk, 1967), which can be modelled through Briggs-Haldane equations

(Briggs and Haldane, 1925; Haldane, 1930; Stein, 1986b):

$$\frac{d[S_{\text{in}}]}{dt} = V_{Tm}.\frac{[S_{\text{out}}] - [S_{\text{in}}]}{K_T + ([S_{\text{out}}] + [S_{\text{in}}]) + \alpha.\frac{[S_{\text{out}}][S_{\text{in}}]}{K_T}} \tag{4}$$

with:

$$\begin{cases} V_{Tm}\text{: the maximum rate of a given carrier protein;} \\ K_T\text{: a constant inversely proportional to the transporter efficiency;} \\ \alpha : \text{the Kotyk interactive constant capturing the disequilibrium between bound and free transporters.} \end{cases}$$

By construction, α cannot exceed 1 Kotyk (1967) and is close to this upper limit for sugars (*e.g.* $\alpha = 0.91$ for glucose (Teusink *et al.*, 1998), so we set $\alpha = 1$ by default in this study, maximizing the effect of interaction).

A model including both FD and irreversible substrate conversion by an enzyme therefore corresponds to the following chemical equation:

$$\mathrm{S}_{out} \underset{\alpha}{\overset{V_{Tm}, K_T}{\rightleftharpoons}} \mathrm{S}_{in} + \mathrm{E} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_{\text{cat}}} \mathrm{E} + \mathrm{P} \tag{5}$$

Using analytical tools (see ter Kuile and Cook (1994), rederived in Supplementary material - Text S5 Mathematical appendix), the flux can be determined through the following set of equations:

$$[ES]^* = \frac{k_f[S_{in}]^*}{k_r + k_{cat} + k_f[S_{\text{in}}]^*}.[E_{tot}] \tag{6}$$

$$v = \frac{d[P]}{dt} = k_{cat}[ES]^* \tag{7}$$

where:

$$[S_{\text{in}}]^* = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{8}$$

with:

$$\begin{cases} a = k_f k_{cat}[E_{tot}](1 + \frac{[S_{in}]}{K_T}) + k_f V_{Tm} \\ b = k_f k_{cat}[E_{tot}]([S_{in}] + K_T) + (k_{cat} + k_r - k_f[S_{in}])V_{Tm} \\ c = -V_{Tm}[S_{in}](k_r + k_{cat}) \end{cases} \tag{9}$$

4.3 Multiple enzymes model

In order to study the evolution of downstream enzymes, we considered an unbranched metabolic pathway in which the product formed by the first reaction serves as the substrate for a second reaction. Theoretically, as there is nothing prohibiting increase in product concentrations – for it is not considered reversible at this point – any second enzyme should be able to sustain any metabolic demand. We

penalized large increases in cellular concentrations through a degradation process of the product of the first reaction, occurring at rate η_d ($\times$ this concentration). Under these new assumptions, the chemical reactions occurring after uptake (Michaelis Menten part of Eq.5) are described by the following equations:

$$\mathrm{S}_{in} + \mathrm{E}_1 \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} \mathrm{E}_1 S \xrightarrow{k_{\mathrm{cat1}}} \mathrm{E}_1 + \mathrm{P}_1 \tag{10}$$

$$\begin{array}{ccccc} \mathrm{P}_1 + \mathrm{E}_2 & \underset{k_{r2}}{\overset{k_{f2}}{\rightleftharpoons}} & \mathrm{E}_2\mathrm{P}_1 & \xrightarrow{k_{\mathrm{cat2}}} & \mathrm{E}_2 + \mathrm{P}_2 \\ \downarrow \eta_d & & & & \\ \mathrm{P}_{1out} & & & & \end{array} \tag{11}$$

Such a system may reach a steady-state at which the cellular concentrations of the substrate S_{in} and of the first product P_1 are constant. At this point, the net instantaneous uptake of substrate (through the membrane) equals the instantaneous production of P_1 which, in turn, equals the addition of the instaneous amount of degraded P_1 and the instantaneous production of P_2, according to the principle of mass conservation. It yields the following system of equations:

$$V_{Tm}.\frac{([S_{out}] - [S_{in}])}{K_T + ([S_{\mathrm{out}}] + [S_{\mathrm{in}}]) + \alpha.\frac{[S_{\mathrm{out}}][S_{\mathrm{in}}]}{K_T}} = V_{m1}.\frac{[S_{in}]}{K_{M1} + [S_{in}]} \tag{12}$$

$$V_{m1}.\frac{[S_{in}]}{K_{M1} + [S_{in}]} = V_{m2}.\frac{[P_1]}{K_{M2} + [P_1]} + \eta_d.[P_1] \tag{13}$$

where appear the traditional Michaelis-Menten kinetic parameters for the (i^{eth}) reaction:

$$\begin{cases} V_{m_i} = k_{cat_i}[E_{tot_i}] \\ K_{M_i} = \dfrac{k_{r_i} + k_{cat_i}}{k_{f_i}} \end{cases}$$

In an independent section, we also introduced reversibility to further the analysis, through the modification of equation 10, which becomes:

$$\mathrm{S}_{in} + \mathrm{E}_1 \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} \mathrm{E}_1 S \underset{k_{inh1}}{\overset{k_{\mathrm{cat1}}}{\rightleftharpoons}} \mathrm{E}_1 + \mathrm{P}_1 \tag{14}$$

Such a phenomenon is described by the more general form of Haldane equation (Haldane, 1930; Cornish-Bowden, 1979), which changes the contribution of the first reaction ($V_{m1}.\frac{[S_{in}]}{K_{M1}+[S_{in}]}$) in both equations (12) and (13) to:

$$V_{m1^+}.\frac{[S_{in}]}{K_{M1^+} + [S_{in}] + K_I[P_1]} - V_{m1^-}.\quad \frac{[P_1]}{K_{M1^-} + [P_1] + [S_{in}]/K_I}$$

with V_{m1^+} and K_{M1^+} respectively correponding to the previous V_{m1} and K_{M1}, while:

$$\begin{cases} V_{m1^-} = k_{r1}[E_{tot1}] \\ K_I = k_{inh1}/k_{f1} \\ K_{M1^-} = K_{M1^+}/K_I \end{cases}$$

To solve these systems, which are not easily tractable analytically – for they rely on cubic equations – we implemented the Newton method (Atkinson, 1989) aiming to find $[S_{in}]^*$ and $[P_1]^*$. We ran the algorithm starting from very low values of concentration (both set to $10^{-20}M$) to determine numerically the equilibrium without facing convergence problems. The final flux can then be determined through the "production" part of equation (13), *i.e.* $V_{m2} \cdot \frac{[P_1]}{K_{M2}+[P_1]}$.

4.4 Validation of the method and computing of the fitness landscapes

To validate the approach, we compared equilibrium results obtained with Raphson-Newton algorithm to those obtained when simulating the process with Euler explicit schemes for a set of (3x3) kinetic values – k_{cat} and k_f – encompassing three orders of magnitude (see SM - Section Text 5 for further details).

In each section, we then determined the flux achieved for enzyme kinetic parameters k_{cat} and k_f varying by 10 orders of magnitude, setting k_r to $10^3 s^{-1}$ – within the range found for several enzymes (Klipp and Heinrich, 1994; Knowles and Albery, 1977) – unless stated otherwise. Except in the section dedicated to the influence of enzyme concentration, we set the enzyme concentration such that $[E_{tot}] = 1mM$, lying nearby the highest observed values (Albe *et al.*, 1990; Noor *et al.*, 2016). Other parameters are detailed on a case-by-case basis as they may change depending on the goal of each section. To compare with the data and visualize results in the experimenter's parameter space, we also determined the flux and plotted simulation results using k_{cat} and $\frac{k_{\text{cat}}}{K_M}$, also making them vary by 10 orders of magnitude. Fitness was considered to be linearly related to the flux in a pathway. To compute fitness landscapes, we thus divided the parameter space in 100 equidistant values on a log-scale (250 for representations with k_{cat}/K_M to draw a cleaner demarcation for the diffusion limit).

4.5 Population genetics model

We built an evolutionary model based on a Wright-Fisher process including selection, to determine how enzymes should evolve on fitness landscapes. We considered the two fitness landscapes corresponding to opposite cases – weak flux, high affinity (A on Fig.2) and high flux, low affinity (I on Fig.2) – for saturated facilitated diffusion ($S_{out} = 10K_T$), with the following constant parameters: $k_r = 10^3 s^{-1}$ and $[E_{tot}] = 10mM$. Mutations followed reproduction and were drawn at random from Gaussian distributions, at rates $\mu = 10^{-1}/N_e$ making the quantity of mutations per unit time similar between

simulations. Along with a similar dispersion ($\sigma = 0.3$), three cases of average mutational biaises were considered for the log-values of forward parameters (k_{cat} and k_f), which affected them independently: no mutational bias ($m = 0$), a low mutational bias ($m = -0.1$ so that approximately 1 mutation out of 3 is positive) and a high mutational bias ($m = -0.2$ so that approximately 1 mutation out of 4 is positive and 1% of them exceed 0.5). To determine the effect of N_e, we tested 4 different values of the parameters, ranging from $N_e = 10^2$ to 10^5. Kinetic parameters were initiated to the inefficient values of $k_{cat} = 10^{-3}s^{-1}$ and $k_f = 10^2M^{-1}s^{-1}$ and k_f was limited to values under that allowed by the diffusion limit – $10^{10}M^{-1}s^{-1}$ (k_{cat} was also limited to 10^6s^{-1} when $b = 0$ to avoid physical outliers). To analyse simulation outcomes, we picked out the kinetic and fitness values of the most represented genotype when multiple variants were segregating. 30 simulations were ran for each set of parameters. Finally, we verified that evolutionary steady-states were reached and considered it was the case when at least the average fitnesses (over all simulations) of the last three time-steps were not significantly different one from another (see Figures S5 and S6 of SM).

4.6 Data sets

All the enzyme data used in this work to compare fitness landscapes and measured values were recovered from (Bar-Even *et al.*, 2011), and so was the classification of reactions with regards to metabolic groups.

5 Supplementary Material

Supplementary material is available online.

References

Albe, K. R., Butler, M. H., and Wright, B. E. 1990. Cellular concentrations of enzymes and their substrates. *J Theor Biol*, 143(2): 163–95.

Alberty, R. A. and Hammes, G. G. 1958. Application of the theory of diffusion-controlled reactions to enzyme kinetics. *The Journal of Physical Chemistry*, 62(2): 154–159.

Atkinson, K. 1989. *An Introduction to Numerical Analysis*. John Wiley and Sons, New York, 2nd edition.

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., and Milo, R. 2011. The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*, 50(21): 4402–4410.

Bar-Even, A., Milo, R., Noor, E., and Tawfik, D. S. 2015. The moderately efficient enzyme: Futile encounters and enzyme floppiness. *Biochemistry*, 54(32): 4969–4977.

Blanco, P. M., Garcés, J. L., Madurga, S., and Mas, F. 2018. Macromolecular diffusion in crowded media beyond the hard-sphere model. *Soft Matter*, 14: 3105–3114.

Bobay, L.-M. and Ochman, H. 2018. Factors driving effective population size and pan-genome evolution in bacteria. *BMC Evolutionary Biology*, 18(1): 153.

Boer, V. M., Crutchfield, C. A., Bradley, P. H., Botstein, D., and Rabinowitz, J. D. 2010. Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. *Molecular biology of the cell*, 21(1): 198–211.

Briggs, G. E. and Haldane, J. B. 1925. A note on the kinetics of enzyme action. *The Biochemical journal*, 19(2): 338–339.

Burns, J., Cornish-Bowden, A., Groen, A., Heinrich, R., Kacser, H., Porteous, J., Rapoport, S., Rapoport, T., Stucki, J., Tager, J., Wanders, R., and Westerhoff, H. 1985. Control analysis of metabolic systems. *Trends in Biochemical Sciences*, 10(1): 16.

Carlin, D. A., Caster, R. W., Wang, X., Betzenderfer, S. A., Chen, C. X., Duong, V. M., Ryklansky, C. V., Alpekin, A., Beaumont, N., Kapoor, H., Kim, N., Mohabbot, H., Pang, B., Teel, R., Whithaus, L., Tagkopoulos, I., and Siegel, J. B. 2016. Kinetic characterization of 100 glycoside hydrolase mutants enables the discovery of structural features correlated with kinetic constants. *PLOS ONE*, 11(1): e0147596–.

Chakrabarti, A. C. 1994. Permeability of membranes to amino acids and modified amino acids: Mechanisms involved in translocation. *Amino Acids*, 6(3): 213–229.

Cornish-Bowden, A. 1979. *Chapter 2 - Introduction to enzyme kinetics*, pages 16–38. Butterworth-Heinemann.

Danielli, J. 1954. Morphological and molecular aspects of active transport. In *Symposia of the Society for Experimental Biology*, volume 8, pages 502–516. COMPANY BIOLOGISTS LTD UNIV CAMBRIDGE DEPT ZOOLOGY, DOWNING ST, CAMBRIDGE ....

Davidi, D., Noor, E., Liebermeister, W., Bar-Even, A., Flamholz, A., Tummler, K., Barenholz, U., Goldenfeld, M., Shlomi, T., and Milo, R. 2016. Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro k_{cat} measurements. *Proceedings of the National Academy of Sciences*, 113(12): 3401.

Davidi, D., Longo, L. M., Jabłońska, J., Milo, R., and Tawfik, D. S. 2018. A bird's-eye view of enzyme evolution: Chemical, physicochemical, and physiological considerations. *Chemical Reviews*, 118(18): 8786–8797.

Dykhuizen, D. E., Dean, A. M., and Hartl, D. L. 1987. Metabolic flux and fitness. *Genetics*, 115(1): 25–31.

Eisenthal, R., Danson, M. J., and Hough, D. W. 2007. Catalytic efficiency and k_{cat}/k_M: a useful comparator? *Trends in Biotechnology*, 25(6): 247–249.

Ellis, R. J. 2001. Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences*, 26(10): 597–604.

Fell, D. A. 1992. Metabolic control analysis: a survey of its theoretical and experimental development. *The Biochemical journal*, 286 (Pt 2)(Pt 2): 313–330.

Griffith, D. A. and Jarvis, S. M. 1996. Nucleoside and nucleobase transport systems of mammalian cells. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 1286(3): 153–181.

Haldane, J. B. S. 1930. *Enzymes*. Longmans, Green, London, New York.

Hartl, D. L., Dykhuizen, D. E., and Dean, A. M. 1985. Limits of adaptation: the evolution of selective neutrality. *Genetics*, 111(3): 655–674.

Heckmann, D., Zielinski, D. C., and Palsson, B. O. 2018. Modeling genome-wide enzyme evolution predicts strong epistasis underlying catalytic turnover rates. *Nature Communications*, 9(1): 5270.

Heim, N. A., Payne, J. L., Finnegan, S., Knope, M. L., Kowalewski, M., Lyons, S. K., McShea, D. W., Novack-Gottshall, P. M., Smith, F. A., and Wang, S. C. 2017. Hierarchical complexity and the size limits of life. *Proceedings of the Royal Society B: Biological Sciences*, 284(1857): 20171039.

Heinrich, R. and Rapoport, T. A. 1974. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *Eur J Biochem*, 42(1): 89–95.

Heinrich, R., Schuster, S., and Holzhütter, H. G. 1991. Mathematical analysis of enzymic reaction systems using optimization principles. *Eur J Biochem*, 201(1): 1–21.

Jiang, M. and Guo, Z. 2007. Effects of macromolecular crowding on the intrinsic catalytic efficiency and structure of enterobactin-specific isochorismate synthase. *Journal of the American Chemical Society*, 129(4): 730–731.

Johnson, K. A. and Goody, R. S. 2011. The original michaelis constant: Translation of the 1913 michaelis–menten paper. *Biochemistry*, 50(39): 8264–8269.

Kacser, H. and Burns, J. A. 1973. The control of flux. *Symp Soc Exp Biol*, 27: 65–104.

Kacser, H., Burns, J. A., Kacser, H., and Fell, D. A. 1995. The control of flux. *Biochemical Society Transactions*, 23(2): 341–366.

Kaltenbach, M. and Tokuriki, N. 2014. Dynamics and constraints of enzyme evolution. *J Exp Zool B Mol Dev Evol*, 322(7): 468–487.

Khersonsky, O. and Tawfik, D. S. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem*, 79: 471–505.

Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature*, 217(5129): 624–626.

Klipp, E. and Heinrich, R. 1994. Evolutionary optimization of enzyme kinetic parameters; effect of constraints. *Journal of Theoretical Biology*, 171(3): 309–323.

Knowles, J. R. and Albery, W. J. 1977. Perfection in enzyme catalysis: the energetics of triosephosphate isomerase. *Accounts of Chemical Research*, 10(4): 105–111.

Koshland, D. E. 2002. The application and usefulness of the ratio kcat/km. *Bioorganic Chemistry*, 30(3): 211–213.

Kotyk, A. 1967. Mobility of the free and of the loaded monosaccharide carrier in saccharomyces cerevisiae. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 135(1): 112–119.

Lunzer, M., Miller, S. P., Felsheim, R., and Dean, A. M. 2005. The biochemical architecture of an ancient adaptive landscape. *Science*, 310(5747): 499–501.

Lynch, M. and Marinov, G. K. 2015. The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences of the United States of America*, 112(51): 15690–15695.

Maier, A., Völker, B., Boles, E., and Fuhrmann, G. F. 2002. Characterisation of glucose transport in Saccharomyces cerevisiae with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS Yeast Research*, 2(4): 539–550.

Michaelis, L. and Menten, M. 1913. Die kinetik der invertinwirkung. biochemische zeitschrift. *Biochemische Zeitschrift*, 49: 333–369.

Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. 2010. Bionumbers–the database of key numbers in molecular and cell biology. *Nucleic acids research*, 38(Database issue): D750–D753.

Muramatsu, N. and Minton, A. P. 1988. Tracer diffusion of globular proteins in concentrated protein solutions. *Proceedings of the National Academy of Sciences*, 85(9): 2984.

Newton, M. S., Arcus, V. L., Gerth, M. L., and Patrick, W. M. 2018. Enzyme evolution: innovation is easy, optimization is complicated. *Current Opinion in Structural Biology*, 48: 110–116.

Noor, E., Flamholz, A., Bar-Even, A., Davidi, D., Milo, R., and Liebermeister, W. 2016. The protein cost of metabolic fluxes: Prediction from enzymatic rate laws and cost minimization. *PLOS Computational Biology*, 12(11): e1005167.

Ohta, T. 1992. The nearly neutral theory of molecular evolution. *Annual Review of Ecology and Systematics*, 23(1): 263–286.

Orth, J. D., Thiele, I., and Palsson, B. Ø. 2010. What is flux balance analysis? *Nature biotechnology*, 28(3): 245–248.

Pozdnyakova, I. and Wittung-Stafshede, P. 2010. Non-linear effects of macromolecular crowding on enzymatic activity of multi-copper oxidase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1804(4): 740–744.

Ralston, G. B. 1990. Effects of "crowding" in protein solutions. *Journal of Chemical Education*, 67(10): 857.

Rich, P. R. 2003. The molecular machinery of keilin's respiratory chain. *Biochemical Society Transactions*, 31(6): 1095–1105.

Rivas, G. and Minton, A. P. 2018. Toward an understanding of biochemical equilibria within living cells. *Biophysical reviews*, 10(2): 241–253.

Rivas, G., Ferrone, F., and Herzfeld, J. 2004. Life in a crowded world. *EMBO reports*, 5(1): 23–27.

Schaefke, B., Emerson, J. J., Wang, T.-Y., Lu, M.-Y. J., Hsieh, L.-C., and Li, W.-H. 2013. Inheritance of gene expression level and selective constraints on trans- and cis-regulatory changes in yeast. *Molecular Biology and Evolution*, 30(9): 2121–2133.

Schäuble, S., Stavrum, A. K., Puntervoll, P., Schuster, S., and Heiland, I. 2013. Effect of substrate competition in kinetic models of metabolic networks. *FEBS Letters*, 587(17): 2818–2824.

Schmidt, J. A., Rinaldi, S., Scalbert, A., Ferrari, P., Achaintre, D., Gunter, M. J., Appleby, P. N., Key, T. J., and Travis, R. C. 2016. Plasma concentrations and intakes of amino acids in male meat-eaters, fish-eaters, vegetarians and vegans: a cross-sectional analysis in the epic-oxford cohort. *European Journal of Clinical Nutrition*, 70(3): 306–312.

Schuster, S., Pfeiffer, T., and Fell, D. A. 2008. Is maximization of molar yield in metabolic networks favoured by evolution? *J Theor Biol*, 252(3): 497–504.

Stein, W. D. 1986a. *CHAPTER 1 - Physical Basis of Movement across Cell Membranes*, chapter 1, pages 1–68. Academic Press.

Stein, W. D. 1986b. *CHAPTER 4 - Facilitated Diffusion: The Simple Carrier*, chapter 4, pages 231–361. Academic Press.

Stein, W. D. 1986c. *CHAPTER 5 - The Cotransport Systems: Two Substrates That Are Carried on a Single Transporter*, chapter 5, pages 363–474. Academic Press.

Sweetlove, L. J. and Fernie, A. R. 2018. The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nature Communications*, 9(1): 2136.

Tabaka, M., Kalwarczyk, T., Szymanski, J., Hou, S., and Holyst, R. 2014. The effect of macromolecular crowding on mobility of biomolecules, association kinetics, and gene expression in living cells. *Frontiers in Physics*, 2: 54.

ter Kuile, B. H. and Cook, M. 1994. The kinetics of facilitated diffusion followed by enzymatic conversion of the substrate. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1193(2): 235–239.

Teusink, B., Diderich, J. A., Westerhoff, H. V., van Dam, K., and Walsh, M. C. 1998. Intracellular glucose concentration in derepressed yeast cells consuming glucose is high enough to reduce the glucose transport rate by 50%. *Journal of bacteriology*, 180(3): 556–562.

Tokuriki, N., Jackson, C. J., Afriat-Jurnou, L., Wyganowski, K. T., Tang, R., and Tawfik, D. S. 2012. Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nature Communications*, 3(1): 1257.

Wilbrandt, W. and Rosenberg, T. 1961. The concept of carrier transport and its corollaries in pharmacology. *Pharmacological Reviews*, 13(2): 109.

Williamson, D. H., Lund, P., and Krebs, H. A. 1967. The redox state of free nicotinamide-adenine dinucleotide in the cytoplasm and mitochondria of rat liver. *Biochem J*, 103(2): 514–527.

Wood, R. E., Wirth, F. P. J., and Morgan, H. E. 1968. Glucose permeability of lipid bilayer membranes. *Biochim Biophys Acta*, 163(2): 171–178.

Wright, S. 1934. Physiological and evolutionary theories of dominance. 68(714): 24–53.

Xie, H., Patching, S. G., Gallagher, M. P., Litherland, G. J., Brough, A. R., Venter, H., Yao, S. Y. M., Ng, A. M. L., Young, J. D., Herbert, R. B., Henderson, P. J. F., and Baldwin, S. A. 2004. Purification and properties of the escherichia coli nucleoside transporter nupg, a paradigm for a major facilitator transporter sub-family. *Mol Membr Biol*, 21(5): 323–336.

Zampieri, M., Hörl, M., Hotz, F., Müller, N. F., and Sauer, U. 2019. Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in escherichia coli. *Nature Communications*, 10(1): 3354.

Zhou, G.-Q. and Zhong, W.-Z. 1982. Diffusion-controlled reactions of enzymes. *European Journal of Biochemistry*, 128(2-3): 383–387.

Zhou, H.-X., Rivas, G., and Minton, A. P. 2008. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics*, 37(1): 375–397.

Zimmerman, S. B. and Minton, A. P. 1993. Macromolecular crowding: Biochemical, biophysical, and physiological consequences. *Annual Review of Biophysics and Biomolecular Structure*, 22(1): 27–65.